Exploiting genetic diversity and gene synthesis to identify superior nitrogenase NifH protein variants to engineer N$_2$-fixation in plants

Xi Jiang$^{1,2}$, Lucía Payá-Tormo$^{1,2}$, Diana Coroian$^1$, Inés García-Rubio$^3$, Rocío Castellanos-Rueda$^{1,4}$, Álvaro Eseverri$^{1,2}$, Gema López-Torrejón$^{1,2}$, Stefan Burén$^5$ & Luis Manuel Rubio$^{1,2}$

Engineering nitrogen fixation in eukaryotes requires high expression of functional nitrogenase structural proteins, a goal that has not yet been achieved. Here we build a knowledge-based library containing 32 nitrogenase nifH sequences from prokaryotes of diverse ecological niches and metabolic features and combine with rapid screening in tobacco to identify superior NifH variants for plant mitochondria expression. Three NifH variants outperform in tobacco mitochondria and are further tested in yeast. Hydrogenobacter thermophilus (Aquificae) NifH is isolated in large quantities from yeast mitochondria and fulfills NifH protein requirements for efficient N$_2$ fixation, including electron transfer for substrate reduction, P-cluster maturation, and FeMo-co biosynthesis. H. thermophilus NifH expressed in tobacco leaves shows lower nitrogenase activity than that from yeast. However, transfer of [Fe$_4$S$_4$] clusters from NifU to NifH in vitro increases 10-fold the activity of the tobacco-isolated NifH, revealing that plant mitochondria [Fe-S] cluster availability constitutes a bottleneck to engineer plant nitrogenases.

1 Centro de Biotecnología y Genómica de Plantas, Universidad Politécnica de Madrid, Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria, Pozuelo de Alarcón, 28223 Madrid, Spain. 2 Departamento de Biotecnología-Biología Vegetal, Escuela Técnica Superior de Ingeniería Agronómica, Alimentaria y de Biosistemas, Universidad Politécnica de Madrid, 28040 Madrid, Spain. 3 Centro Universitario de la Defensa, Ctra. de Huesca s/n, 50090 Zaragoza, Spain. 4 Department of Biosystems Science and Engineering, ETH Zürich, 4058 Basel, Switzerland. 5 email: stefan.buren@upm.es; lm.rubio@upm.es
Nitrogen (N) fertilizers used to increase crop productivity in intensive agriculture practices pollute groundwater and release greenhouse gases. On the other hand, subsistence agriculture practices including poor N fertilization produce low and inconsistent yields causing malnutrition and poverty. There is large interest in engineering cereal crop varieties capable of acquiring their own N. One approach to this outcome relies on functional expression of a nitrogenase enzyme by the cereal plant. Nitrogenases are prokaryotic, O₂-sensitive, two-component metalloproteins that convert inert N₂ into biologically useful NH₃. The most efficient and widespread variant, the molybdogenum nitrogenase, is composed of an Fe protein (nifH-encoded) and a MoFe protein (encoded by nifD and nifK). The Fe protein (NifH) donates electrons to the MoFe protein (NifDK) that in turn reduces N₂. Nascent NifH and NifDK polypeptides need to acquire proper quaternary structure to receive metal clusters, one [Fe₄S₄] cluster per NifH homodimer and two pairs of P-cluster and FeMo-co per NifDK heterotetramer, for functionality. We have recently reviewed the mechanisms and genetic requirements to assemble these cofactors and to mature NifH and NifDK into active Mo nitrogenase. The large number of nitrogen fixation (nif) genes involved, and the sensitivity of most of the protein products towards O₂, makes nitrogenase engineering a daunting task with issues that need to be solved stepwise.

To date, functional NifH, NifU, and NifB have been purified from mitochondria of aerobically cultured Saccharomyces cerevisiae cells, while active NifU and NifH were isolated from chloroplasts of Nicotiana benthamiana at the end of the dark period. Also, the reported low stability of the NifD protein has now been improved in two recent studies that identified key residues in the NifD sequence as susceptible to cleavage upon mitochondrial import. Notwithstanding these achievements, detailed analysis of yeast mitochondria-targeted Azotobacter vinelandii NifH has been hampered by low protein solubility resulting in suboptimal yields. Accumulation of mostly insoluble NifH was also reported when Klebsiella oxytoca NifH was targeted to the tobacco mitochondria. The difficulty of expressing high levels of soluble and functional NifH in yeast and tobacco poses a major problem for eukaryotic nitrogenase engineering as it is the most abundant Nif protein during N₂ fixation. The problem is exacerbated because, in addition to serving NifDK with electrons for substrate reduction, NifH is required to mature P-clusters onto NifDK and for the final steps of FeMo-co biosynthesis in complex with NifEN. For these reasons it is essential to identify a NifH variant that is highly soluble and stable when expressed at very high levels in a plant cell, and that can perform all three NifH-dependent activities. One approach to achieve this outcome would be protein engineering of well-studied NifH from model diazotrophs (e.g., A. vinelandii or K. oxytoca) aimed to introduce sequences that improve stability in the mitochondria. Protein engineering has been extensively employed to obtain glyphosate resistance, another important trait for crops. Alternatively, mining of phylogenetically diverse nifH sources can be undertaken in order to find natural NifH proteins with superior properties, a strategy that was successful for NifB and for increasing carotenoid levels in "Golden Rice".

Here, 32 distinct nifH genes were screened for expression level and solubility in mitochondria of N. benthamiana. The nifM, nifU, and nifS genes were co-expressed because their protein products are involved in NifH folding and in the biosynthesis and delivery of its [Fe₄S₄] cluster. The Hydrogenobacter thermophilus NifH was identified as vastly superior to the A. vinelandii NifH in terms of expression levels, solubility, and functionality both in tobacco and yeast mitochondria. Mitochondria-targeted H. thermophilus NifH satisfied all functional and spectroscopic requirements of a nitrogenase Fe protein when purified from yeast. The screening also pinpointed the plant mitochondria [Fe-S] cluster assembly as a bottleneck for further engineering.

**Results**

Library design and strategy for expression of mitochondria-targeted NifH in *N. benthamiana*. A library of 32 nifH sequences from phylogenetically diverse prokaryotes was designed considering one or several of the following criteria: (i) nifH genes found in confirmed diazotrophs, in plant-associated bacteria; (ii) nifH genes from phototrophs or plant-associated bacteria; (iii) nifH genes from aerobic organisms; (iv) growth temperature of the nifH host; (v) nifH genes from archaeal representatives (Supplementary Data 1). Organized by phylog, the selection included genes from 1 Aquificae, 4 Firmicutes, 1 Actinobacteria, 15 Proteobacteria, 6 Cyanobacteria, 1 Chlorobi, 1 Chloroflexi, and 3 Euryarchaeota (Fig. 1a).

The workflow of this study is described in Fig. 1b. The gene sequences encoding the 32 NifH variants were cloned into plant vectors for Agrobacterium tumefaciens infiltration-mediated NifH expression in *N. benthamiana* leaves (Supplementary Table 1, see Methods section for details). The nifH sequences were codon-optimized for *S. cerevisiae* because codon-usage is similar to tobacco and the workflow included downstream expression of tobacco-selected NifH variants in yeast for biochemical characterization.

The genes were under control of the strong and constitutive E35S promoter. Amino-terminal COX4-TS extensions were added to NifH proteins. COX4 is the 29 amino acid transit peptide of the *S. cerevisiae* mitochondria protein cytochrome c oxidase subunit IV (MLSRQSLRKKPATRTLCSSRYLLQK), whereas TS denotes the 28 amino acid Twin-Strep-Tag peptide (WSPHQFEKGGSGGGGGGSAWHPQFEK). COX4 targeted NifH proteins to the mitochondria matrix and TS was used to enable variant-independent immunoblot detection of NifH and to facilitate its purification. Importantly, the TS-tag has been shown to not significantly affect NifH functionality. COX4-TS-NifH variants are hereafter denoted as NbNifH, where Nb stands for the host *N. benthamiana*, and denotes variants collectively, and other superscripts indicate the species from which NifH sequence was obtained. Vectors with NbNifH constructs additionally contained a transcriptional unit for expression of the green fluorescent protein (GFP) that was used as indicator of successful leaf infiltration (Supplementary Table 1).

An auxiliary vector was constructed to co-express A. vinelandii nifM, nifU, and nifS and target their protein products to mitochondria via N-terminal SU9 extensions. Similar to COX4, the mitochondrial prescience of subunit 9 of the Neurospora crassa F₀-ATPase (SU9) has been shown to deliver Nif proteins to *N. benthamiana* mitochondria. NifU and NifS assemble [Fe-S] clusters destined for Nif proteins in A. vinelandii. While not essential for expression of functional NifH in *S. cerevisiae* mitochondria, they were required to generate high amounts of active NifB in yeast. As we aimed to identify NifH variants accumulating at higher levels than NifH, NifU and NifS were included in this study. In *A. vinelandii* and other well-studied diazotrophs NifM is involved in NifH folding or dimerization prior [Fe₄S₄] cluster acquisition. Despite NbNifM not being present in organisms of some selected nifH variants (Supplementary Data 1), this gene was always included in infiltration experiments for consistency.

**Identification of NifH proteins suitable for expression in *N. benthamiana***

*N. benthamiana* leaves were co-infiltrated with a 1:1:1 mixture of three distinct *A. tumefaciens* cultures for expression of, respectively, one NbNifH variant plus GFP, the auxiliary proteins NbNifMa, NbNifUa, and NbNifSa, and the
RNA silencing suppressor p19 to enhance the nif transgene expression (Fig. 1b). Protein extracts were prepared from the N. benthamiana leaves three days after infiltration and analyzed for accumulation of soluble NbNifHXX using antibodies recognizing the TS-tag. Only two NifH variants were consistently detected among experiments (Fig. 1c, Supplementary Fig. 1a), namely those originating from Methanococcales and Hydrogenobacter thermophilus. A third NifH variant from Methanothermobacter marburgensis (NbNifHMM) was detected at low levels at one occasion. In contrast, analysis of total extracts prepared from the infiltrated tobacco leaves showed that, although accumulation levels of the NbNifHXX proteins varied significantly, 25 of the 32 variants could be detected (Supplementary Fig. 1b). Only NbNifH expression of variants from Bradyrhizobium janicum, Rhizobium leguminosarum bv. trifolii, Herbaspirillum seropedicae, Gloeothecae sp. KO68DGA, Rhodopseudomonas palustris, Methanothermobacter thermautotrophicus, and Frankia sp. (strain FaC1) could not be demonstrated. Sequence alignments and 3D-modeling of NbNifHMM, NbNifHIII, and NbNifHII are shown in Supplementary Fig. 2. The 3D-models did not reveal any specific feature that would explain their superior accumulation as soluble protein in tobacco mitochondria, but all three proteins originate from thermophilic organisms (Supplementary Data 1) which could possibly explain their stability and solubility.

Activity of NifH variants isolated from mitochondria of aerobically cultured S. cerevisiae. N. benthamiana screening-identified variants and NbNifHIV were expressed in S. cerevisiae and purified by Strep-tag affinity chromatography (STAC) to evaluate functionality when targeted to mitochondria. For this, genes encoding COX4-TS-NifH constructs were transferred to expression vectors together with su9-nifMAv, su9-nifHAv, and su9-nifSAv under the control of galactose-inducible GAL1 or GAL10 promoters (Supplementary Table 2, Supplementary Fig. 3a). These COX4-TS-NifH variants expressed in aerobic S. cerevisiae cultures are hereafter denoted ScNifHMM, ScNifHIII, and ScNifHII, and ScNifHAv (ScNifHXX collectively).

While ScNifHMM, ScNifHIII, and ScNifHII were purified to near homogeneity (Fig. 2a), SDS-PAGE analysis of ScNifHAv showed additional slower migrating co-eluting proteins. Mass spectrometry confirmed that these were contaminants (Fig. 2a). ScNifHAv solubility was low and much protein was lost to the pellet fraction when preparing the soluble cell-free extract (CFE) explaining its poor purification yield (about 11 mg per kg of S. cerevisiae cells) (Supplementary Fig. 3b-e, Supplementary Table 3). The yield of ScNifHMM was also relatively low, in line with the inferior result in the N. benthamiana screening. In contrast, the yields of ScNifHIII and ScNifHII were ca. 20 times higher. Iron (Fe) quantification of purified samples was variable but indicated that ScNifHIII was isolated largely as holo-protein containing one [Fe4S4] cluster per dimer (Supplementary Table 3). Consistently, immunoblot analysis showed that ScNifHIII, ScNifMAv, ScNifUAv, and ScNifSAv had been efficiently targeted to the mitochondria (Supplementary Fig. 4).

Activities of purified ScNifHXX variants were determined in vitro using the acetylene reduction assay (ARA) and compared...
to that of NifH purified from *A. vinelandii* (denoted NifH<sub>Av</sub>). In all cases NifDK purified from *A. vinelandii* (denoted NifDK<sub>Av</sub>) was used as MoFe protein component. ScNifH<sup>Av</sup> activity was 85% of NifH<sup>Av</sup> (Fig. 2b), supporting previous observations that STAC is suitable for purification of metal-cluster containing Nif proteins expressed in yeast<sup>11,28</sup>. Reconstitution of ScNifH<sup>Mm</sup> and ScNifH<sup>Mi</sup> showed very low activities (Fig. 2b). The assay did not determine whether lower activities were due to NifH variant defects, or to incompatibility with ScNifU<sup>Sc</sup> in vivo (resulting in apo-NifH protein with low [Fe<sub>4</sub>S<sub>4</sub>] cluster occupancy) or NifDK<sup>Av</sup> in vitro (resulting in poor electron donation). Reconstitution of ScNifH<sup>Mm</sup> clusters in vitro by either mixing with Fe<sub>4</sub>S<sub>4</sub>-complex (Fig. 2b), and ScNifH<sup>Mi</sup> by direct reconstitution (Supplementary Fig. 5), indicating that this NifH variant is not compatible with NifDK<sup>Av</sup>. In contrast, ScNifH<sup>Mi</sup> was activated to some extent by NifU<sup>Av</sup>, and further by direct reconstitution, indicating that the *A. vinelandii* NiU<sup>Sc</sup> machinery is not optimal for NifH<sup>Mi</sup> (Supplementary Fig. 5). However, activities were very low compared to the as-isolated ScNifH<sup>Av</sup> protein (Fig. 2b). This could be explained by NifH<sup>Av</sup> harboring more of the conserved amino acid residues known to be important for the interaction with NifDK<sup>Av</sup> (Supplementary Fig. 2).

Importantly, soluble accumulation of ScNifH<sup>Av</sup> in mitochondria was 20-fold higher than ScNifH<sup>Av</sup> (Supplementary Table 3), which translates into at least 10-fold higher in vivo activity and fulfills NifH quantity requirements for nitrogenase engineering. Thus, ScNifH<sup>Av</sup> was further characterized.

### ScNifH<sup>Av</sup> exhibits NifH-characteristic spectroscopic signals and is functional in vivo.  
Purified ScNifH<sup>Av</sup> protein presented ultraviolet–visible (UV–vis) absorption spectra typical of O<sub>2</sub>-sensitive [Fe-S] cluster-containing proteins (Fig. 3a). Amino-terminal sequencing revealed that amino acid residues EQKP remained after COX<sub>4</sub> processing (Fig. 3b), where conversion of glutamine (Q) to glutamic acid (E) could be due to deamination performed by the mitochondrial matrix N-terminal amidase NAT<sub>Av</sub><sup>129</sup>. Electron paramagnetic resonance (EPR) confirmed that ScNifH<sup>Av</sup> protein contained an [Fe<sub>4</sub>S<sub>4</sub>] cluster with similar signal intensity and g-values as NifH<sup>Av</sup> (Fig. 3c), suggestive of successful maturation into functional Fe protein.

The NifH variant chosen to engineer N<sub>2</sub>-fixing plants must perform P-cluster maturation and FeMo-co biosynthesis in addition to serve as electron donor for substrate reduction. We therefore tested whether *H. thermophilus* NifH could revert the Ni<sup>+</sup><sup>+</sup>-phenotype of *A. vinelandii* DJ77 (ΔnifH strain)<sup>30</sup>. For this, *ts-nifH<sup>Av</sup>* was introduced by transformation into DJ77 and the resulting strain UW481 was tested for diazotrophic growth and in vivo acetylene reduction activity. UW481 showed diazotrophic growth both in solid and liquid media (Supplementary Fig. 6a, b), and immunoblot analysis demonstrated sustained AvNifH<sup>Av</sup> expression and acetylene reducing activity indicative of active nitrogenase (Supplementary Fig. 6c, d). These data strongly indicate that NifH<sup>Av</sup> can replace the functions of native *A. vinelandii* NifH to some extent, which requires productive interactions with at least apo-NifDK<sup>Av</sup>, NifDK<sup>Av</sup>, and NifEN<sup>Av</sup> proteins.

### ScNifH<sup>Av</sup> is active in substrate reduction, P-cluster formation and FeMo-co synthesis.  
Each individual NifH-dependent activity was then analyzed in vitro using pure ScNifH<sup>Av</sup> preparations (Fig. 3d). P-cluster maturation was determined by supplementing CFE of *A. vinelandii* DJ77 (ΔnifH) with ScNifH<sup>Av</sup>. The DJ77 extract is devoid of FeMo-co and contains inactive apo-NifDK<sup>Av</sup> with immature P-clusters. The P-cluster maturation assay using DJ77 CFE relies on positive outcomes of three distinct activities performed in two sequential reactions (Fig. 3d). In the first reaction (Step I + II) pure NifH and FeMo-co are added to DJ77 CFE resulting in NifH-dependent reductive coupling of the two [Fe<sub>4</sub>S<sub>4</sub>] P-cluster precursors to form mature P-clusters (Step I), followed by FeMo-co insertion into P-cluster containing apo-NifDK<sup>Av</sup> to generate active NifDK<sup>Av</sup> (Step II) (Fig. 3d). Tetra-thiomolybdate is then added to prevent further FeMo-co insertion, separating the maturation (Step I + II) and activity (Step III) reactions. Activation of DJ77 apo-NifDK<sup>Av</sup> by ScNifH<sup>Av</sup> demonstrated its P-cluster maturation activity (Fig. 3e).

In vitro FeMo-co synthesis (Fig. 3d, Step II) was determined by combining purified preparations of ScNifH<sup>Av</sup>, apo-NifDK<sup>Av</sup> containing P-clusters but devoid of FeMo-co<sup>31</sup>, apo-NifEN<sup>Av</sup> containing permanent [Fe<sub>4</sub>S<sub>4</sub>] clusters but lacking FeMo-co precursor<sup>32</sup>, Mo, homocitrate, and either the FeMo-co precursor
NiF-B co-bound to the carrier protein NiFXA3 or NiFB protein supplemented with Fe and S34. As for the P-cluster maturation assay, tetramethylidylobate was added before the ARA (Fig. 3d, Step III). Figure 3f shows that ScNiHFlH supported FeMo-co synthesis in vitro. Importantly, ScNiHFlH and ScNiBMM (Methanothrix thermotoeutrophics NiFB isolated from S. cerevisiae)14 acted together in the NiFB-dependent in vitro FeMo-co synthesis assay in which NiFB-co was concomitantly synthesized by ScNiBMM rather than added in purified form. This result proved compatibility of two essential proteins for N2 fixation, ScNiHFlH and ScNiBMM, when produced in yeast mitochondria. It also showed interspecies compatibility with NiFDK4 and NiFEA4, altogether constituting the conserved biochemical core of nitrogenase.

ScNiHFlH activity in substrate reduction was demonstrated by the ARA and by reduction of N2 to NH3. ARA titration was carried out with a fixed quantity of NiFDK4 and increasing amounts of ScNiHFlH. Maximum NiFDK4 activity was achieved at molar ScNiHFlH to NiFDK4 ratios larger than 40 (Fig. 3g), similar to reactions with the natural counterpart NiHFlH35. This result suggests that the maximum activity that can be achieved combining ScNiHFlH with NiFDK4 is 1000 units (i.e., half of the activity with NiHFlH). In addition, ScNiHFlH supported N2 reduction into NH3 by NiFDK4. Importantly, the ratio of NH3 to ethylene produced by NiFDK4 was similar independently of using NiHFlH or ScNiHFlH (Fig. 3h).

As-isolated NbNiHFlH was inactive but could be activated by [Fe3S4] cluster reconstitution. NbNiHFlH was purified from A. tumefaciens-infiltrated leaves of N. benthamiana. Plants were grown under long-day conditions (16 h light/8 h dark) and leaves were processed at the end of the dark period. Genes encoding

Fig. 3 Characterization of ScNiHFlH. a UV–vis absorption spectra of as-isolated and air-exposed ScNiHFlH. b Processing site (black arrow) of COX4 mitochondria targeting signal (blue) as determined by N-terminal sequencing of ScNiHFlH. The conversion of Q to E (marked by *) could be due to deamination processes. c EPR signal of as-isolated ScNiHFlH (190 μM) compared to NiHFlH (71.2 μM). Experimental data (black lines) and simulations (blue lines) of each protein are shown together with g values. d Schematic representation of NiHFlH-dependent activities tested in e–h. Figure adapted from Burén et al.9. Copyright 2020 ACS under CC BY 4.0 http://creativecommons.org/licenses/by/4.0. e In vitro P-cluster maturation of apo-NiFDK4 present in CFE of A. vinelandii DJ77 (ΔnifH) after addition of ScNiHFlH or NiHFlH as indicated (+). Tetramethylidylobate was added (following step I + II) to inhibit further apo-NiFDK4 activation during the ARA (step III). Data represent mean values (n = 4 technical replicates, ±). Tetrathiomolybdate was added (following step I) to inhibit further apo-NiFDK4 activation during the ARA (step III). A 20:1 molar ratio of ScNiHFlH to NiFDK4 was used in the ARA (step III). Data represent mean values (n = 2 technical replicates (NiFB-co + NiHFlH), n = 3 technical replicates (NiFB-co + ScNiHFlH), n = 4 technical replicates (NiFB-co + ScNiHFlH), n = 5 technical replicates (ScNiBMM + ScNiHFlH), n = 3 technical replicates (ScNiBMM), n = 4 technical replicates (ScNiHFlH)). Blue and red dots correspond to independent experiments. f Titration of NiFDK4 activity with ScNiHFlH. Positive control reactions performed with NiHFlH and NiFDK4 at 4:1 molar ratio gave 1692 ± 4 units (nmol ethylene formed per min and mg of Fe protein). g Fe protein concentrations were processed at the end of the dark period. Genes encoding...
NbNifH<sup>Hi</sup>, NbNifM<sup>Av</sup>, NbNifU<sup>Av</sup>, and NbNifS<sup>Av</sup> (together with p19 and GFP) were piled up in a single plant-expression vector for co-expression (Methods section and Supplementary Table 1). Purified NbNifH<sup>Hi</sup> did not exhibit brown color of [Fe-S] clusters and was inactive in the ARA when combined with NifDK<sup>Av</sup> (Fig. 4a). Therefore, we reconstituted NbNifH<sup>Hi</sup> [Fe<sub>4</sub>S<sub>4</sub>] cluster in vitro either by mixing with Fe, L-cysteine, DTT, and EcNifS<sup>Av</sup> (direct reconstitution), or by incubating with [Fe<sub>4</sub>S<sub>4</sub>] cluster-loaded EcNifU<sup>Av</sup> (NifU-mediated reconstitution). Both methods activated the NbNifH<sup>Hi</sup> as determined by the ARA (Fig. 4a), demonstrating that the protein was correctly folded but lacked its [Fe<sub>4</sub>S<sub>4</sub>] cluster. This result suggested that insertion and/or stability of NbNifH<sup>Hi</sup> [Fe<sub>4</sub>S<sub>4</sub>] cluster was poor in mitochondria of leaves.
Fe fertilization of the soil increases soluble NifU in mitochondria of *N. benthamiana*. One explanation for the low [Fe₄S₄] cluster content of NifH in tobacco mitochondria could be insufficient Fe availability in the soil. We observed that accumulation of NhNifUAv, but not of NhNifAv, increased when the water used to irrigate the *A. tumefaciens*-infiltrated plants was supplemented with Fe (Fig. 4b). Sulfur was not supplemented in soil as the infiltration solution contained MgSO₄. Although Fe fertilization tripled the yield of STAC-isolated NhNifUAv, the average Fe content of 2 Fe atoms per protein was not affected (Supplementary Fig. 7a, b, Supplementary Table 4). This could be due to the loss of transient NifU [Fe-S] clusters during purification, and it is not a surprising outcome as isolation of EcNifUAv containing only the permanent [Fe₂S₂] clusters has been previously observed. Immunoblots detected two differently migrating NhNifUAv species in purifications from tobacco leaves (Fig. 4c). Amino-terminal sequencing showed that both species were cleaved either one or seven amino acids into the TS-tag (Fig. 4d). As both NhNifUAv species showed the same N-termini processing, we concluded that the faster migrating polypeptide was truncated at the C-terminus.

Extended dark period combined with Fe fertilization produced active NhNifHtHt in mitochondria of *N. benthamiana* leaves. The soil of *N. benthamiana* plants expressing NhNifHtHt was fertilized with Fe to increase Fe availability. In addition, the dark period preceding leaf harvest was extended from 8 h to 16 h to hypothesizing that longer darkness would lower intracellular O₂ and stabilize NhNifHtHt [Fe₄S₄] cluster. Dark period extension did not increase NhNifHtHt accumulation (Fig. 4e) but allowed for isolation of active protein as shown below. About 6 mg of NhNifHtHt was consistently isolated per kg of N. benthamiana leaves (Fig. 4f, Supplementary Fig. 7c, Supplementary Table 5). Amino-terminal sequencing showed that mitochondria-targeted NhNifHtHt accumulated as two species (similar to NhNifUAv), one in which two amino acid residues from the TS-tag were removed with the COX4 signal and another that was processed five amino acid residues further into the TS-tag (Fig. 4g).

Functionality of NhNifHtHt isolated from leaves of Fe fertilized tobacco plants following 16 h of darkness was determined using ARA. NhNifHtHt preparations consistently showed activities but these were low compared to those of [Fe₄S₄] cluster-reconstituted NhNifHtHt (Fig. 4h). This result suggested that NhNifHtHt accumulated as two species in tobacco mitochondria, whereas inactive protein likely lacking [Fe₄S₄] cluster was more abundant than functional and [Fe₄S₄] cluster-containing NhNifHtHt. Consistently, the Fe content of purified NhNifHtHt preparations were below detection limit (Supplementary Table 5). Altogether the results indicate that while soluble NhNifHtHt accumulates in good quantity in mitochondria of *N. benthamiana* leaves, engineering of additional protein components or biosynthetic pathways will be required to improve [Fe₄S₄] cluster acquisition or stability.

Discussion

The first study reporting production of active NifH in yeast proved that mitochondria is a suitable organelle for hosting O₂-sensitive Nif proteins under aerobic growth conditions. Despite being a valid proof-of-concept, further developments with *A. vinelandii* NifH were limited by low yields as only a small portion was soluble in the mitochondrial matrix. Similar solubility issues were later reported for *K. oxytoca* NifH targeted to *N. benthamiana* mitochondria and are confirmed in this study using immunoblot screening and STAC. Identifying the best possible NifH protein for eukaryotic (plant) expression was therefore of uttermost importance. NifH is the most abundant Nif protein required for N₂ fixation in *A. vinelandii*. Besides being the Fe protein component of Mo nitrogenase, NifH is essential to the assembly of both NifDK cofactors, namely the P-cluster and the FeMo-co.

NifH proteins for nitrogenase engineering in plants should: (i) be stable and soluble at high levels in the mitochondrial matrix, and (ii) be compatible with the NifDK component from a well-studied model-diazotroph if their own NifDK components are not available in purified form. Compatibility is important when evaluating function of candidate NifH variants. In our case it meant that any selected NifH XS should be compatible with Nif-M Av (if NifH XS is not NifM-independent), NifU Av for maturaton and [Fe₄S₄] cluster synthesis/insertion, and NifDK Av for nitrogenase activity measurements. We note that this requirement introduces a selection bias and that the screening could have overlooked NifH variants that were superior to that of *H. thermophilus* if combined with different NifDK.

The NifH variants tested in this study were selected from a curated dataset of hundreds of NifH sequences by favoring aerobic or plant-associated origins, to overcome the inherent O₂-sensitivity of NifH, and functionality at moderate temperatures. We also hypothesized that NifH variants from archaea could function better in a eukaryotic environment as this domain of life is believed to be more closely related to the Eukarya than, and because our previous work expressing archaean NifB variants in yeast had shown them to be superior to those of bacterial origin.

We expected that most NifH variants would be partly soluble in tobacco mitochondria when expressed together with the accessory proteins NifU Av, NifS Av, and NifM Av. However, only NifH from *M. infernus* and *H. thermophilus* were consistently detected in soluble tobacco extracts, in addition to *M. marinum*. This was occasionally detected at lower levels. Two of these NifH proteins originated from archaea and the third from a bacterium. One possibility could be that the NifM Av protein was not expressed at sufficient levels in the tobacco mitochondria and that only these three NifH variants did not require NifM for maturation. However, low levels of NifM expression appear to be enough for NifH maturation in *K. oxytoca*. A more plausible explanation can be found in the thermophilic nature of *M. infernus*, *H. thermophilus*, and *M. marinum*. It has recently been hypothesized that these organisms evolved the capacity to tolerate high temperatures.
been reported that the temperature inside respiring mitochondria of cultured human cells is around 50 °C, even when the external medium is maintained at 38 °C. Whether the same drastic effect on temperature holds true for mitochondria of a leaf cell is not known to us, but it could explain in part the outcome of our NifH screening. None of the two highest expressed NifH proteins originated from proven diazotrophs. We are not aware of any study investigating diazotrophy in the archaon M. infernus. However, NifB cured the Nif- phenotype of an A. vinelandii nifB mutant strain and, as NifB has no other known function than biosynthesis of nitrogenase active-site cofactors, it is likely that M. infernus is in fact a diazotroph. On the other hand, N2- fixation has been tested but not observed in H. thermophilus TK-68. Interestingly only six NifH variants in our library originated from organisms having genes with high similarity to A. vinelandii nifM. Perhaps other prolyl isomerases could substitute for NifM in these organisms. Whether NifM (and the NifUS machinery) is required for maturation of the three selected NifH proteins (especially NifH(H)) in mitochondria will be investigated in future work.

Mitochondria-expressed ScNifH(H) was the only variant that supported relevant nitrogenase activity when combined with NifDK(Av). Its activity corresponded to roughly half of that using ScNifH(Av) even if the ScNifH(H) to NifDK(Av) molar ratio was increased well above 40 normally used for ARA. Emerich and Burris showed that NifH proteins can function with NifDK from other bacteria. An optimal growth temperature of 72 °C has been reported for H. thermophilus TK-68, which could explain lower ScNifH(H) activity in substrate reduction assays. However, our prediction from this study and previous work on NifH is that suboptimal working temperature of Nif proteins from thermophiles is a price worth paying when engineering nitrogenase in eukaryotes, as solubility and stability of these variants is so much improved.

One observation of this study was that the specific activity of as-isolated NbNifH(H) protein was lower than ScNifH(H). We think this was caused by poor [Fe₃S₄] cluster availability – and hence inefficient incorporation – or poor NifH [Fe₃S₄] cluster stability within the leaf cell mitochondria. In this context, it is not known how Fe fertilization increased accumulation of soluble NbNifU(Av). More available Fe could increase mitochondria [Fe-S] clusters biosynthesis and [Fe₃S₄] cluster occupancy in NbNifU(Av) which, in turn, would provide stability to the protein. A compatibility issue between NbNifH(H) and NbNifU(Av) and NbNifS(Av) is unlikely since EcnifU(Av) could effectively activate NbNifH(H) in vitro. NbNifH(H) misfolding in mitochondria is also unlikely as it was efficiently activated by reconstitution of its [Fe₃S₄] cluster. It is however likely that protection by respiratory O₂-consumption in leaf is lower than in yeast. NbNifH(H) exposure to O₂ during leaf processing is also a possibility making this a purely technical problem. While leaves were kept in liquid nitrogen and lysis and purification were performed inside an anaerobic glove box, it is difficult to completely rule out that some O₂ trapped within the leaf was released during tissue disruption.

In conclusion, this study shows that genetic diversity can be exploited to identify, from a very large pool of sequences, the most adequate Nif protein components to engineer a eukaryotic nitrogenase. Modular cloning techniques, gene synthesis with codon optimization, and other synthetic biology tools permit building multi-protein pathways with components of very diverse origin. In this case the NifH protein from H. thermophilus was identified as soluble in mitochondria of both S. cerevisiae and N. benthamiana accumulating at much higher levels than the A. vinelandii homologue. This example is relevant not only because the identified variant performed all three NifH-essential reactions, namely P-cluster maturation, FeMo-co biosynthesis, and NiFD-K₄V reduction, but also because NifH(H) formed functional interspecies interactions with NifB, NifE, and NifDK proteins, altogether representing the four proteins constituting the core of diazotrophy.

**Methods**

**Design, assembly, and cloning of the nifH library.** A curated dataset of diazotrophs was used to collect nifH candidates and design the library. Genes encoding nifH variants were codon optimized for expression in S. cerevisiae and the sequence encoding pE35Scos:twinstrep was codon optimized for expression in tobacco (Supplementary Data 1). All genetic parts were optimized using the GeneOptimizer tool (ThermoFisher) and synthesized by ThermoFisher via the Engineering Nitrogen Symbiosis for Africa (ENSA) project. The nifH genes were synthesized as pMA cloning vectors with BamHI and BstXI restriction sites flanking each gene. The pE35Scos:twinstrep sequence was flanked by HindIII and BglII restriction sites.

pGFPuGusPlus (plasmid #64401, Addgene) and the pMA vector containing pE35Scos:twinstrep were digested with HindIII and BglII and used to generate the parental vector pN2SB41, containing a pE35Scos:twinstrep-gus-nOS transcriptional unit in which gus was flanked by BamHI and BstElII restriction sites. The parental vector pN2SB41 and all pMA vectors containing nifH variants were digested with BamHI and BstElII and used to generate vectors pN2XJ81 and pN2XJ112 (Supplementary Table 1).

pN2XJ165 was used to generate vector pN2XJ165 containing transcription units for mitochondria-targeted accessory Nif proteins (A. vinelandii NifU, NiF, and NifM). The su9-nifAv(AAvAAAGGTCACTTATGGCGCTCAGCTGCTCG, AAAAGGTCTCTAGATCAGCTCAGCTCAGCTGCTCG) and su9-nifAv(AAACTATGATGGCCTCAGCTCAGCTCAGCTCAGCTGCTCG, AAAAGGCTCTACACAGCTCAGCTCAGCTCAGCTGCTCG) were amplified by PCR from the yeast vector pN2GLT418. Amplification reactions added flanking BamHI and BstElII (for su9-nifAv) or SpeI and SacI (for su9-nifAm) sites. The DNA fragment containing the su9-nifAv(AAv sequence was created by overlapping PCR using primers introducing sequences homologous to those flanking the XhoI site of pGFPuGusPlus (ATTATGGAGAAGATCTACCTGCTATGTTCC, TACAAAAATCTATCTCTTGCTTGCAGTGCTCAGCTGCTCG, AAAAGGTCTCTAGATCAGCTCAGCTCAGCTCAGCTGCTCG) and was digested with NheI and XhoI restriction sites (Supplementary Table 3).

**Growth of S. cerevisiae, mitochondria isolations, and ScNifH purifications.** S. cerevisiae for galactose-induced expression of ScNifH(H), ScNifU(H), ScNifH(H), and ScNifU(Av) together with Su9-NifAv, Su9-NifAm, and Su9-NifAm (XJ1-XJ4, Supplementary Table 3) were cultured in 4-l fermenters under aerobic conditions (0.625 i of air per minute and 1 i of culture, 230 rpm stirring) and used for mitochondrial isolations or NifH purifications as previously described. Preparation of CFE and STAC purifications were performed at O₃ levels below 1 ppm in anaerobic chambers (Coy systems or Mbraun). Typically, cells were resuspended in lysis buffer (100 mM Tris-HCl (pH 8.6), 200 mM NaCl, 10% glycerol, 2 mM sodium dithionite (DTH), 1 mM PMSF, 1 mM 2-mercaptoethanol (DTE), 10% sucrose) and centrifuged at 5000 g for 1 h at 4 °C (Avanti J-26 SP). The supernatant was filtered using O₃-containing cups with a pore size of 0.22 μm, washing cell-free extract (CFE) of soluble proteins that was loaded at 2.5 ml/min into a 5 ml Strep-Tactin XP column (IBA LifeSciences) attached to an AKTA FPLC (GE Healthcare). The column was washed using 75 ml washing buffer (100 mM Tris-HCl pH 8.0, 200 mM NaCl, 10% glycerol, 2 mM DTE). Strep-Tactin XP column-bound proteins were eluted with 15 ml washing buffer supplemented with 50 mM biotin (IBA LifeSciences). The elution fraction was concentrated, and biotin removed, by passing the protein through PD-10 desalting columns (GE Healthcare). Desalted eluate was further concentrated using centrifugal filters (Amicon, Millipore) with 30 kDa cutoff. Finally, the concentrated protein was snap-frozen in cryovials (Nalgene) and stored in liquid N₂.

**Soil Fe fertilization, preparation of anaerobic N. benthamiana leaf cell-free extracts, and purification of NbNifH(H) and NbNifU(Av).** N. benthamiana plants were kept in the dark under long day conditions and kept in 16 light:8 dark cycles. Plants were irrigated (2 l per week) with tap water supplemented with 1 g/l Sequestrene G100.
Conditions were carefully monitored to avoid over-modulation or saturation effects. mW and a magnetic cryostat refrigerated with helium. For measurements, a microwave power of 2.5 μP, pH 7.8, as buffer. Mixtures were prepared in volumes of 750 μl of the blank (30–30) were added in duplicates to 200 μl ATP-regenerating mixture as described above. Reduction of N2 to NH3 was determined in reaction mixtures prepared as for the ARA but containing 100 mM 3-(N-morpholino)propanesulfonic acid (MOPS), pH 7.8, as buffer. Mixtures were prepared in volumes of 750 μl from which 100 μl was removed at 9 min to serve as background (t0) for N2 measurements. After exchanging via atmosphere for N2, mixtures were incubated at 30 °C for 30 min, and reactions were stopped by addition of 100 μl of 5 mM EDTA. Twenty-five μl of the blank (t0) and the reaction (t) were added in duplicates to 200 μl o-phthalaldehyde reagent solution (ThermoFisher Scientific) in 96-well microplate for fluorescence-based assays (Nunc). Fluorescence (Ex 390 nm, Em 472 nm) was measured using a Varianous LUX plate reader (ThermoFisher Scientific). NH3 production was determined from the increase in fluorescence (t−t0) against standards prepared with NH4Cl and recorded in the same plate.

In vitro NifH activity. NifH activity was described as measured by Shah et al. with slight modifications.48 Reactions were prepared inside anaerobic chambers. Purified NifH proteins were analyzed by ARA after addition of NidDK46 and ATP-regenerating mixture (1.23 mM ATP, 18 mM phosphocreatine, 2.2 mM MgCl2, 3 mM DTH and 46 μg/ml of creatine phosphokinase, 22 mM Tris-HCl pH 7.5) in a final volume of 9 μl and incubated at 37 °C for 30 min. Then, reactions were supplemented with 1 mM L-cysteine, 1 mM DTT, 400 μM (NH4)2Fe(SO4)2, and 255 mM NifS purified from E. coli (EcNifS)10, and incubated at 37 °C overnight. Finally, the proteins were diluted 1000-fold in 22 mM Tris-HCl (pH 7.5) buffer, and then concentrated using centrifugal filters (Amicon, Millipore) with 30 kDa cutoff to remove excess reagents.

For "direct reconstitution" activity assays, the activity of [Fe-S] cluster reconstituted NifH protein was determined using ARA. For "NiF-mediated reconstitution", as-isolated NifH protein was mixed with [Fe-S] cluster reconstituted EcNifS, and then immediately used for ARA.

Statistics and reproducibility. Distinct samples were used for in vitro activity measurements and sample sizes are indicated by n, where each distinct sample was measured at least two times. Mean of measured activities are shown. The data presented in the figure graphs are listed in Supplementary Data 2.

Data availability
The authors declare that the data supporting the findings of this study are available within the article, its supplementary information and data, and upon request.

Received: 15 June 2020; Accepted: 25 November 2020; Published online: 04 January 2021

COMMUNICATIONS BIOLOGY | (2021) 4:4 | https://doi.org/10.1038/s42003-020-01536-6 | www.nature.com/commsbio

Protein methods, antibodies, UV-vis absorption spectrum, and electron paramagnetic resonance. Protein concentrations were measured as performed by a protein assay (PIERCE) in combination with iodoacetamide to eliminate the interference of DTH. Colorimetric Fe determination was measured as reported36, and the N-terminal amine acid sequences were determined by Edman degradation (Proteome Factory AG).

Antibodies used in this study and their dilutions for immunoblotting were as follows: polyclonal antibodies detecting NifU (used at 12,000 in 5% BSA), NifS (used at 1,200 in 5% BSA), NifH (used at 12,000 in 5% BSA), and NifDK (used at 1,200 in 5% BSA) were raised against purified preparations of A. vinelandii proteins (generated in house). A Streptag II ("Strep-MAB", 2-1502-001, IBA Lifesciences®, 1:50,000 in TBS-T), GFP (sc-9996, Santa Cruz Biotechnology, 1:500 in TBS-T), α-Tubulin (3H3087, sc-69971, Santa Cruz Biotechnology). Anti-Tubulin antibody was used to measure tubulin level in the cells.

Strept-Avidin conjugated to horseradish peroxidase (Strep-HRP, sc-2240, Santa Cruz Biotechnology). For tubulin level measurement in the cells, Anti-Tubulin antibody was used to measure tubulin level in the cells.

In vitro Fe-Moco synthesis and apo-NidDK reconstitution. NidB-dependent Fe-Moco synthesis assays were performed inside anaerobic chambers as described by Curati et al., with slight modifications5. One hundred μl reactions contained 3.0 μM NidB, 20.4 μM Fe(I), 1.5 μM apo-NidEN44, 0.6 μM apo-NidDK45, 17.5 μM NidH, 175 μM R-homocitrate, 1 mg/ml BSA, and ATP-regenerating mixture (1.23 mM ATP, 18 mM phosphocreatine disodium salt, 2.2 mM MgCl2, 3 mM DTH, 46 μg/ml creatine phosphokinase, final concentrations in 22 mM Tris-HCl (pH 7.5)) buffer at 30 °C for 60 min. NidB-dependent Fe-Moco synthesis assays were performed as described above. NidB-dependent assay replacing GST-NidX-NidB-co by 10.0 μM NiB monomer, 125 μM FeSO4, 125 μM Na2S2O3, and 125 μM SAM.

Following in vitro synthesis of Fe-Moco-co, 17.5 μM (NH4)2MoS4 was added to prevent further Fe-Moco incorporation into apo-NidDK46, and incubated for 10 min at 25 °C. Activation of apo-NidDK was measured by addition of 50 μl ATP-regenerating mixture and ScNidH46 (2.0 μM final concentration) in 9 ml vials containing Ar and 500 μl acetylene. The ARA were performed at 30 °C for 20 min. Positive control reactions for ARA contained NidDK46 and NidE46. Ethylene formed was measured in 50 μl gas phase samples using a Porapak N 80/100 column in a gas chromatograph (Shimadzu).

In vitro [Fe-S] cluster reconstitution and NifH activity. In vitro [Fe-S] cluster reconstitutions of NifH and NifU were performed with E. coli (EcNifH)10,10 were performed in anaerobic chambers as described by Zheng and Dean39 with slight modifications. NifH or NifU (20 μM) was added to 22 mM Tris-HCl (pH 7.5) buffer supplemented with 8 mM 1,4-diithiothreitol (DTT) in a final volume of 100 μl and incubated at 37 °C for 30 min. Then, reactions were supplemented with 1 mM L-cysteine, 1 mM DTT, 400 μM (NH4)2Fe(SO4)2, and 225 mM NifS purified from E. coli (EcNifS)10, and incubated at 37 °C overnight. Finally, the proteins were diluted 1000-fold in 22 mM Tris-HCl (pH 7.5) buffer, and then concentrated using centrifugal filters (Amicon, Millipore) with 30 kDa cutoff to remove excess reagents.

For "direct reconstitution" activity assays, the activity of [Fe-S] cluster reconstituted NifH protein was determined using ARA. For "NiF-mediated reconstitution", as-isolated NifH protein was mixed with [Fe-S] cluster reconstituted EcNifH, and then immediately used for ARA.

References
1. Erisman, J. W. et al. Nitrogen: Too much of a vital resource. (Science Brief. WWF Netherlands, Zeist, The Netherlands, 2015).
2. Borlaug, N. E. Feeding a world of 10 billion people: the miracle ahead. 228 (2002).
3. Sibbata, K. T. & Qaim, M. Rural food security, subsistence agriculture, and seasonality. PloS One 12, e186406 (2017).
4. Bloch, S. E., Ryu, M. H., Ozaydin, B. & Broglie, R. Harnessing atmospheric nitrogen for cereal crop production. Curr. Opin. Biotechnol. 62, 181–188 (2020).
5. Burén, S. & Rubio, L. M. State of the art in eukaryotic nitrogenase engineering. FEMS Microbiol. Lett. 365, fnx274–fnx274 (2018).
6. Bulen, W. A. & LeComte, J. R. The nitrogenase system from Azotobacter: two-enzyme requirement for N2 reduction, ATP-dependent H2 evolution, and ATP hydrolysis. Proc. Natl Acad. Sci. USA 56, 979–986 (1966).
7. Hoffman, B. M., Lukoyanov, D., Yang, Z. Y., Dean, D. R. & Seefeldt, L. C. Mechanism of nitrogen fixation by nitrogenase: the next stage. *Chem. Rev.* **114**, 4041–4062 (2014).

8. Seefeldt, L. C. et al. Reduction of substrates by nitrogenases. *Chem. Rev.* **120**, 5082–5120 (2020).

9. Burén, S., Jimenez-Vicente, E., Echavarri-Erasun, C. & Rubio, L. M. Biosynthesis of nitrogenase cofactors. *Chem. Rev.* **120**, 4921–4968 (2020).

10. Lopez-Torrejon, G. et al. Expression of a functional oxygen-labile nitrogenase component in the mitochondrial matrix of aerobically grown yeast. *Nat. Commun.* **7**, 11426 (2016).

11. Burén, S. et al. Biosynthesis of the nitrogenase active-site cofactor precursor NiFe-b in *Saccharomyces cerevisiae*. *Proc. Natl Acad. Sci. USA* **116**, 29078–29086 (2019).

12. Euseverri, A. et al. Use of synthetic biology tools to optimize the production of active nitrogenase Fe protein in chloroplasts of tobacco leaf cells. *Plant Biotechnol. J.* **18**, 1882–1896 (2020).

13. Burén, S. et al. Formation of nitrogenase NiFeDK tetramers in the mitochondria of *Saccharomyces cerevisiae*. *ACS Synth. Biol.* **6**, 1043–1055 (2017).

14. Allen, R. S. et al. Plant expression of NiFD protein variants resistant to mitochondrial degradation. *Proc. Natl Acad. Sci. USA* **117**, 23165–23173 (2020).

15. Xiang, N. et al. Using synthetic biology to overcome barriers to stable expression of nitrogenase in eukaryotic organelles. *Proc. Natl Acad. Sci. USA* **117**, 16537–16545 (2020).

16. Okada, S. et al. A Synthetic Biology Workflow Reveals Variation in Processing and Solubility of Nitrogenase Proteins Targeted to Plant Mitochondria, and Differing Tolerance of Targeting Sequences in a Bacterial Nitrogenase Assay. *Front Plant Sci* **11**, 552160 (2020).

17. Poza-Carrion, C., Jimenez-Vicente, E., Navarro-Rodriguez, M., Echavarri-Erasun, C. & Rubio, L. M. Kinetics of NiF gene expression in a nitrogen-fixing bacterium. *J. Bacteriol.* **196**, 595–603 (2014).

18. Vogtle, F. N. et al. Global analysis of the mitochondrial N-proteome identifies a processing peptidase critical for protein stability. *Cell* **139**, 428–439 (2009).

19. Pollegioni, L., Schonbrunn, E. & Siehl, D. Molecular basis of glyphosate resistance-different approaches through protein engineering. *FEBS J.* **278**, 2753–2766 (2011).

20. Paine, J. A. et al. Improving the nutritional value of golden rice through increased pro-vitamin A content. *Nat. Biotechnol.* **23**, 482–487 (2005).

21. Nakamura, Y., Gojobori, T. & Ikemura, T. Codon usage tabulated from protein: a diamagnetic cluster containing eight iron atoms. *Int. J. Biochem. Cell Biol.* **34**, 8, 1421 (2012).

22. Schmidt, T. G. et al. Development of the Twin-Strep-tag(R) and its application for purification of recombinant proteins from cell culture supernatants. *Protein Expr. Purif.* **92**, 54–61 (2013).

23. Westermann, B. & Neupert, W. Mitochondria-targeted green fluorescent proteins: convenient tools for the study of organelle biogenesis in *Saccharomyces cerevisiae*. *Yeast* **16**, 1421–1427 (2000).

24. Burén, S., Jiang, X., Lopez-Torrejon, G., Echavarri-Erasun, C. & Rubio, L. M. Purification and in vitro activity of mitochondria targeted nitrogenase cofactor mature NiF. *Front. Plant Sci.* **8**, 1567 (2017).

25. Kennedy, C. & Dean, D. The niFU, niF and niF gene products are required for activity of all three nitrogenases of *Azotobacter vinelandii*; a homodimer that contains two identical [Fe–Fe]–5 clusters. *Biochemistry* **33**, 13455–13463 (1994).

26. Hill, H. D. & Straka, J. G. Protein determination using bicinchoninic acid in the presence of sulphydryl reagents. *Anal. Biochem.* **170**, 203–208 (1988).

27. Fish, W. W. Rapid colorimetric micromethod for the quantitation of complexed iron in biological samples. *Methods Enzymol.* **158**, 357–364 (1988).

28. Stoll, S. & Schweiger, A. EasySpin, a comprehensive software package for spectral simulation and analysis in EPR. *J. Magn. Reson.* **178**, 42–55 (2006).

29. Shah, V. K. & Brill, W. J. Nitrogenase. IV. Simple method of purification of homogeneity of nitrogenase components from *Azotobacter vinelandii*. *Biochim. Biophys. Acta* **305**, 445–454 (1973).

30. Zheng, L. & Dean, D. R. Catalytic formation of a nitrogenase iron-sulfur cluster. *J. Biol. Chem.* **269**, 18723–18726 (1994).

Acknowledgements
This paper is dedicated to the memory of Prof. Tomás Ruiz Argüeso. We thank Marcel Veldhuizen for yeast fermentations and Carlos Echavarri-Erasun for helpful discussions. Funding for this research was provided by Bill & Melinda Gates Foundation Grant OPP1143172 (L.M.R.). X.J. is supported by a doctoral fellowship from Universidad de Alcalá-Spain and L.P.T. is recipient of the PPU16/00284 from Ministerio de Ciencia, Innovación y Universidades. I.G.R. acknowledges financial contributions from MINECO (CTQ2015-64486-R), Gobierno de Aragón (E35_17R) and Fondo Social Europeo “Construyendo Europa desde Aragón”.

Author contributions
X.J., L.P.T., D.C., I.G.R., R.C.R., A.E., G.L.T., and S.B. performed the experiments; X.J., L.P.T., I.G.R., S.B., and L.M.R. designed experiments and analyzed data; X.J., S.B., and L.M.R. wrote the paper.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s42003-020-01536-6.
