Endowing plants with the capacity for autogenic nitrogen fixation

Quan-hong Yao (✉ yaoquanhong88@163.com )
Shanghai Academy of Agricultural Sciences

Ri-he Peng
Shanghai Academy of Agricultural Sciences

Bo Wang
Shanghai Academy of Agricultural Sciences

Yong-Sheng Tian
Shanghai Key Laboratory of Agricultural Genetics and Breeding

Yan-man Zhu
Shanghai Key Laboratory of Agricultural Genetics and Breeding

Jian-jie Gao
Shanghai Key Laboratory of Agricultural Genetics and Breeding

Zhen-jun Li
Shanghai Key Laboratory of Agricultural Genetics and Breeding

Xiao-Yan Fu
Shanghai Key Laboratory of Agricultural Genetics and Breeding

Jing Xu
Shanghai Academy of Agricultural Sciences

Hong-juan Han
Shanghai Academy of Agricultural Sciences

Li-juan Wang
Shanghai Academy of Agricultural Sciences

Yu Wang
Shanghai Academy of Agricultural Sciences

Fu-jian Zhang
Shanghai Academy of Agricultural Sciences

Wen-hui Zhang
Shanghai Academy of Agricultural Sciences

Yong-dong Deng
Shanghai Key Laboratory of Agricultural Genetics and Breeding

Biological Sciences - Article
Abstract

Biologically available nitrogen is a common limitation to crop productivity in modern agriculture. The endowment of higher plants with the ability to produce their own nitrogenous fertilizers has been attempted for nearly half a century\textsuperscript{1–4}. Here we report that a minimal nitrogen fixation system from \textit{Paenibacillus polymyxa}\textsuperscript{5–8} can be used to create an autogenic nitrogen-fixing plant through synthetic biology. We found that the genetically modified \textit{Arabidopsis} containing the cassette of all nine \textit{nif} genes (\textit{nif}BHDKENXhesAnifV) showed some activity of nitrogenase and caused higher biomasses and chlorophyll contents than wild-type plants grown in low-nitrogen or nitrogen-free medium. Then we found that the engineered \textit{Arabidopsis} displayed resistance to KCN and NaN\textsubscript{3}, two substrates of nitrogenase\textsuperscript{9}. Furthermore, overexpression of electron transfer component\textsuperscript{10} in the engineered \textit{nif} gene-carrying plants resulted in higher nitrogen fixation efficiency. Isotopic labeling analysis using liquid chromatography-tandem mass spectrometry showed that the fixed nitrogen can flow to amino acids and chlorophyll\textsuperscript{11,12}. This study represents a milestone toward realizing the goal of endowing plants with the capacity for self-fertilization.

Main Text

Nitrogen is one of the most important components of living cells. Sixty percent of total fixed nitrogen from all natural and industrial sources on earth comes from biological nitrogen fixation (BNF), the conversion of atmospheric N\textsubscript{2} to NH\textsubscript{3}. However, BNF occurs only in limited species of bacteria and archaea\textsuperscript{13–15}. Plants cannot directly use atmospheric N\textsubscript{2}. The limited bioavailability of nitrogen and the dependence of crop growth on this element have resulted in extremely high ammonia factors worldwide\textsuperscript{16,17}. Scientists have long desired to cultivate nitrogen-self- sufficient/semi-self-sufficient plants. Currently, two major approaches are considered for the engineering of diazotrophic plants. The first method for developing nitrogen-sufficient plants focuses on the engineering of symbiotic interactions between nonlegume plants and nitrogen-fixing bacteria, which would improve the nodule formation that currently occurs in soybeans and other legumes\textsuperscript{18–20}. This approach is extremely difficult to execute due to the complexity of plant-microbe interactions\textsuperscript{21–29}. The second approach for developing nitrogen-sufficient plants involves the direct introduction of bacterial nitrogen fixation genes into a plant to enable it to produce nitrogen-fixing enzymes and fix atmospheric nitrogen.

To date, the goal of introducing nitrogen fixation genes directly into a plant remains unrealized\textsuperscript{30,31}. Several challenges must be overcome to obtain a nitrogen-fixing plant. The smallest number of \textit{nif} genes necessary to produce a functional nitrogenase in plants and the optimal expression levels of various genes necessary to produce an efficient nitrogenase remain unknown\textsuperscript{32–37}. The toxicity of oxygen to nitrogenase is another challenge of achieving high nitrogen fixation activity in plants\textsuperscript{2,38–41}. Moreover, it is an arduous project to construct a plant multi-gene expression system for the reason that the vector became increasingly cumbersome and unstable when too much gene is inserted\textsuperscript{42,43}. In this study, we
selected the compact BNF system from *Paenibacillus* sp. to design a strategy enabling the constitutive expression of components required for the biosynthesis and activity of nitrogenase in the cytoplasm. We also designed to introduce the ferredoxin, which can act as an electron donor for the Fe protein of nitrogenase and reductant for oxygen protection\(^{44-46}\), into the transgenic nif plants to improve nitrogen fixation efficiency.

**Gene synthesis, vector construction, and plant transformation**

To develop a first-generation nitrogen-fixing plant, the simplest nitrogen fixation system in the gram-positive bacterium *Paenibacillus* sp. (GenBank accession number CP017967.3) was chosen for plant transformation. Its minimal compact nif gene cluster contained nine genes organized in the order \(\text{nifBHDKENXhesAnifV}\) (Figure 1A). The structures of the Fe protein and MoFe protein in the nitrogenase of *Paenibacillus* are similar to those of *Azotobacter vinelandii*\(^{47,48}\) (Figure 1B), but the processing of nitrogenase metalloclusters and the synthesis of Mo cofactor (FeMo-co) is relatively simple\(^{49}\) (Figure 1C). To test whether the minimal nif gene cluster can be used in plant nitrogen fixation, we reconstructed the expression system of the nitrogenase with the constitutive CaMV 35S \(\Omega\) promoter after modifying and synthesizing all genes using codons at high frequency in plants (Figure 1D, Supplementary materials and methods). Then the entire nitrogenase synthetic pathway was introduced into *Arabidopsis*. More than 100 hygromycin B-resistance plants were further screened at a low level of nitrogen. Three lines showing a normal vegetative phenotype were selected from homozygous T\(_2\) plants containing a stable T-DNA insertion for further analysis. Molecular analysis by PCR amplification confirmed that all nif genes were present in these transgenic plants (Figure S1A) and all lines exhibited single insertions (Figure 1E). Furthermore we revealed that transgenic lines 37 and 54 exhibited a T-DNA insertion on chromosome 5, while line 58 exhibited a T-DNA insertion on chromosome 3 by resequencing method (Figure 1F, S1B).

**Expression of nitrogenase genes in plants**

To determine the transcription efficiency of every nif gene in individual transgenic plants, quantitative real-time PCR and Northern blotting were conducted. The results of Northern blotting analysis indicated that the RNAs were stable and intact (Figure 2A), but the expression level of each exogenous gene was different (Figure 2B, S1C). The levels of the nif genes were lower in transgenic line 58. For revealing the mechanism of different transcription efficiency of transgenic lines, we assessed the methylation of the CG, CHG and CHH sites in the CaMV 35S \(\Omega\) promoter following the reaction of genomic DNA with bisulphate and the amplification and sequencing of the promoter\(^{50,51}\). The average DNA methylation level in transgenic line 58 reached 93.1%, which is higher than those in line 37 (45.8%) and line 54 (53.6%). The CpG islands in the CaMV 35S promoter showed a higher level of methylation than those of CHG and CHH (Figure 2C).
To determine whether all Nif proteins required for nitrogenase biosynthesis and function could be expressed in plants given the successful transcription of all nif genes in transgenic plants, Western blotting and immunogold labeling were conducted. Western blotting analysis with rabbit polyclonal antibodies against each protein showed that each encoded Nif polypeptide in transgenic lines was approximately the predicted size though the abundance of every polypeptide varied (Figure 2D). Immunogold labeling of structural proteins of nitrogenase and their cofactors in TEM sections using their specific antibody showed associations of Au particles in the cytoplasm, cell membrane, endoplasmic reticulum, and other organelle membranes (Figure 2E, S2). These finding proved that all Nif proteins were expressed and distributed across the cells of nif transgenic Arabidopsis.

Detection of activity of nitrogenase in transgenic plants

To test this hypothesis that the complete Paenibacillus Nif system is capable of N₂ fixation in transgenic plants, we sowed Arabidopsis seeds in low- or no-nitrogen medium and compared the growth potential of the plants. Under low-nitrogen conditions (50 mg/L KNO₃), the most evident difference between transgenic and wild-type plants was that the former exhibit greener and larger leaves (Figures 3A, S3). After growth for 15 days, the fresh weight of the transgenic plants was 1.12-1.28 times of wild-type plants (Figure 3D); moreover, the total chlorophyll content of the former was 1.32-1.56 times of the latter (Figure 3F). After growth for 60 days, the dry weight of the transgenic line 54 was increased to 1.37 times of wild-type plants (Figure S3). An identical phenotype change was detected in the six-leaf stage. When the plants were transferred to medium with low-nitrogen conditions (50 mg/L KNO₃) and grown for 30 days, the total chlorophyll content of the transgenic plants was 2–4 times that of the wild-type plants (Figures 3C, 3G). Surprisingly, when Arabidopsis was grown on nitrogen-free medium, the leaves of the transgenic seedlings presented watery lesions, and their biomass declined significantly (Figure 3B). The symptoms may be related to ammonium toxicity in cells. The fresh weight of the wild-type seedlings in turn was 16%–24% higher than that of transgenic plants grown in nitrogen-free conditions. With the increase in growing time, the symptoms worsened (Figure S3B). The fresh weight of transgenic line 54 was only 72% that of wild-type seedlings when seedlings were grown in nitrogen-free conditions for 60 days. However, the dry wet ratios of transgenic plants were obviously higher than that of wild-type plants (Figure S3C). In addition, the content of the proteins in solution and ammonium concentrations in the nif gene-transgenic plants were higher than those of wild-type plants regardless of whether they were grown under low-nitrogen conditions or nitrogen-free conditions (Figure 3E, S4).

To verify that the growth difference between transgenic and wild-type plants come from nitrogenase, we changed the oxygen concentration in the environment where plants grow based on its sensitive to oxygen. We found that the growth difference between transgenic and wild type seedlings became more obvious with the decrease in time of illumination from 16 h to 8 h under low-nitrogen conditions (Figure S5). Another experiment showed that the fresh weight and total chlorophyll content of healthy seedlings
of transgenic lines increased more obviously than that of wild-type with the decrease in oxygen concentration from 60% to 5% (Figure S6).

To determine whether the nine Nif proteins could be assembled in transgenic plants and function as in *Paenibacillus*, we adopted two independent methods for assessing nitrogenase activity: (1) the acetylene reduction method, wherein acetylene is reduced to ethylene, which can be readily quantified by gas chromatography, and (2) the $^{15}$N isotope enrichment method, which directly measures the incorporation of this tracer into organic nitrogen. When grown in nitrogen-deficient and high-Mo-concentration medium, the crude proteins of all transgenic plants exhibited acetylene reduction. Transgenic line 37 and 54, which expresses the nif genes at significantly higher levels, showed approximately 2–3-fold-higher acetylene reduction activity than line 58. The high nitrogenase activity of line 37 and 54 was confirmed by the plant's level of assimilation of $^{15}$N$_2$ (Figures 3H, 3I). In addition, the nitrogenase activity in underground parts was higher than that in its aboveground parts (Figure S7). The poor expression of NifB and NifK in line 58 may interfere with the assembly of the multiprotein complex of nitrogenase (Figure 2D).

**The expressed nitrogenase in transgenic plants shows wide substrate specificity**

In addition to reducing N$_2$ and protons, nitrogenase can reduce azide and a wide array of carbon-containing compounds such as alkynes and carbon–nitrogen substrates. In these substrates, KCN and NaN$_3$ can decrease plant respiration rates by inhibiting the electron transfer pathway through cytochrome. To verify the activity of nitrogenase to catalyze the decomposition of the two respiratory inhibitors, we tested the tolerance of transgenic plants to KCN and NaN$_3$. We found that the death rates of wild-type *Arabidopsis* seedlings reached approximately 63.5% and 41.6%, while the death rates of nif gene-transgenic *Arabidopsis* seedlings were only 24.2%–28.5% and 3.5%–14.7% when treated with 20 mg/L KCN and 1.35 mg/L NaN$_3$, respectively (Figure 4AB). In addition, the respiratory inhibitors KCN and NaN$_3$ exerted more significant effects on the growth of wild-type plants than on transgenic plants. The fresh weights of the wild-type seedlings were reduced by 12% at low nitrogen concentrations (30 mg/L KNO$_3$) and 16.7% at higher nitrogen concentrations (200 mg/L KNO$_3$) when 10 mg/L KCN was added to the medium. By comparison, the fresh weight of the transgenic plants remained unchanged or even increased when KCN was added to the medium. As the reduction of HCN was observed with the production of ammonia (NH$_3$) for Mo-nitrogenases, low KCN concentrations can stimulate the growth of nif gene-transgenic plants at low nitrogen concentrations. The fresh weight of wild-type seedlings was 4.5%–11.5% lower than that of engineered nif gene-carrying plants when 0.65 mg/L NaN$_3$ was added to the medium containing a higher nitrogen concentration (200 mg/L KNO$_3$), but the reduction was increased to 24.4%–29.7% at a low nitrogen concentration (30 mg/L KNO$_3$) (Figure 4C, 4D).
NifXVhaesA assists in nitrogenase assembly and ferredoxin is a useful electron donor for nitrogenase in plants

NifH, NifD, NifK, NifE, NifN, and NifB are conserved in diazotrophs. In theory, the expression of these six proteins could allow the breeding of a nitrogen-fixing plant. To determine whether plants can complete the assembly of nitrogenase with the six proteins, we constructed a constitutive expression vector with four nifBHDK genes and six nifBHDKEN (Figure 5A). No difference in plant growth in nutrient medium was found among the wild-type, nifBHDK, nifBHDKEN, and nifBHDKENXVhesA transgenic plants. However, when the growth environment exhibited low-nitrogen conditions, wild-type and transgenic seedlings containing incomplete Paenibacillus nitrogen fixation gene clusters showed reduced growth and leaf chlorosis compared with plants expressing all genes in the cluster (Figure 5B). The fresh weight of the plants containing only nifBHDKEN and nifBHDK was 24%–33% lower, and their total chlorophyll content was 30%–37% lower than those of plants containing the complete nitrogen fixation gene cluster nifBHDKENXVhesA when grown with 50 mg/L KNO₃ for 15 days (Figures 5D, 5E). In addition, transgenic seedlings containing nifBHDKEN showed no resistance to KCN and NaN₃, as observed in the wild-type (Figure 5C). When grown on low-nitrogen medium containing 10 mg/L KCN or 0.65 mg/L NaN₃, more than 20% of nifBHDKEN transgenic plants died, and their fresh weights were only 52% and 64% of that of nifBHDKENXVhesA plants. Similar results were obtained in studies of the nifBHKD and nifXVhaesA transgenic lines (Figure 5F). No nitrogenase activity was detected in the nifBHKD and nifBHDKEN transgenic lines by either the acetylene reduction method or the ¹⁵N₂ enrichment method. NifXVhaesA might participate in the assembly of nitrogenase and act as an activating factor for its activity in plants.

Electron transfer component (ETC) is important for the supply of reducing power for catalysis of nitrogen fixation. However, the nifF gene encoding flavodoxin, a direct electron donor to nitrogenase, is deleted in the minimal Paenibacillus nif gene cluster genome. In addition, flavodoxin does not exist alone in plants. To provide efficient electron donors to Paenibacillus nitrogenase expressed in the plant cytoplasm, we introduced exogenous ferredoxin to the engineered nif gene-carrying plants based on the fact that ferredoxins were found to be competent electron donors to nitrogenase in Rhodospirillum rubrum. At first, we transferred genes of ETC into nif gene-carrying plants by hybridization. Transgenic plants containing ferredoxin and the ferredoxin oxidoreductase of the naphthalene dioxygenase from Pseudomonas putida were used as the male parents in crosses with engineered nif gene-carrying plants. The hybrids grew better than the engineered nif gene-carrying plants under low-nitrogen conditions (Figure 5G). In addition, the engineered nif gene-carrying plants showed greater resistance to KCN and NaN₃ upon the introduction of ETC (Figure 5H). The level of ¹⁵N₂ assimilation in the hybrids was higher than that in engineered nif gene-carrying plants (Figure 5I). Then we introduced genes of ETC into nif gene-carrying plants by Agrobacterium-mediated transformation. The activity of nitrogenase also increased in plants with high expression of genes of ETC. In addition, we found no significant difference
in effects on the active level of nitrogenase between ferredoxin and flavodoxin (Figure 5J). We speculated that ferredoxin can be act electron donor as flavodoxin to nitrogenase in vivo.

Mass spectrometry reveals that $^{15}\text{N}_2$ can be transformed to the source of nitrogen for amino acids and chlorophyll biosynthesis in plants

Inorganic nitrogen produced by nitrogen fixation can be involved in the synthesis of amino acids and other important nitrogen-containing compounds, such as amino acids, chlorophyll, and many products of secondary metabolism in nature. To test whether some intracellular pool of nitrogen is indeed synthesized from N$_2$ fixation, we investigated the metabolism of organisms using tracer techniques with stable isotopes. We found that the nine tested amino acids and chlorophyll of nif transgenic plants have higher level of $^{15}\text{N}$ than that of wild type plants after cultured in nitrogen free condition supplemented with isotopically labeled $^{15}\text{N}_2$ (20% air). Moreover, the hybrids of line 37 and ferredoxin plants (404-37+FD) have a higher level of $^{15}\text{N}$ labeled amino acids and chlorophyll than nif transgenic line 37 (Figure 6 and Supplementary Table S6). For example, glutamate, the most important amino acid for nitrogen metabolism, showed a $^{15}\text{N}$ isotopic abundance of 4.9% in the hybrids 404-37+FD, while at natural abundance levels (about 0.37%) in wild-type plants, asparagine, the principal nitrogen transport amino acid in higher plants, also showed a higher $^{15}\text{N}$ isotopic abundance in the hybrids 404-37+FD, alanine, which played a key role in storing carbon and nitrogen, showed a $^{15}\text{N}$ isotopic abundance more than 5 times higher in the hybrids 404-37+FD than in wild-type plants. Other amino acids produced by intermediates in the EMP pathway, such as isoleucine and valine formed from pyruvate, serine and glycine from 3-phosphoglyceric acid, and phenylalanine from phosphoenolpyruvate (PEP), have more or less been labeled by $^{15}\text{N}$ (Figures 6 and S8; Supplementary Table S6). Higher levels of $^{15}\text{N}$ isotopic abundance in chlorophyll a and chlorophyll b were also detected in hybrids 404-37+FD than in control plants (Figures 6 and S9; Supplementary Table S6).

Discussion

Despite the results of pioneering work, the ability to reliably engineer micro-BNF systems in plants remains limited. Here, we constructed a plant multigene expression vector with *Paenibacillus* nitrogen fixation genes to constitutively and simultaneously express nitrogenase catalytic proteins and their assembly cofactors in cells. When we grew plants in a low-nitrogen environment, the complete-nif gene-transgenic plants overcame the general nitrogen deficiency and achieved a greater biomass and higher chlorophyll content. We have detected the activity of nitrogenase via protein extraction from the transgenic lines in vitro through the acetylene reduction method and N$_2$ assimilation in vivo by the $^{15}\text{N}$ enrichment method. Furthermore, we found that the transgenic lines show resistance to the respiratory inhibitors KCN and NaN$_3$, two substrates of nitrogenase. Through further research using tracer techniques, we also found that nitrogen can flow to the anabolism of amino acids and chlorophyll in
engineered plants. Based on our analysis, we are confident that the complete set of *Paenibacillus* nitrogen fixation genes enables plants to synthesize their own nitrogen fertilizer.

The transgenic lines described in this work may serve as a unique starting point for the introduction of the nif gene cluster from bacteria to enable nitrogen fixation activity in plants. Although numerous challenges and many barriers remain before plants can efficiently fix atmospheric nitrogen, for example protecting nitrogenase from toxicity of oxygen and fine-tuning the expression of target genes, the work here let us see a glimmer of light to produce nitrogen autotrophic plants in future.

**References**

1. Beatty, P.H. & Good, A. Future prospects for cereals that fix nitrogen. *Science* **333**, 416-417 (2011).
2. Curatti, L. & Rubio, L.M. Challenges to develop nitrogen fixing cereals by direct nif-gene transfer. *Plant Sci.* **225**, 130-137 (2014).
3. Vicentea, E.J. & Deana, D.R. Keeping the nitrogen-fixation dream alive. *Natl. Acad. Sci. USA* **114**, 3009-3011 (2017).
4. Good, A. Toward nitrogen-fixing plants. *Science* **359**, 869-870 (2018).
5. Wang, L. et al. A minimal nitrogen fixation gene cluster from *Paenibacillus* sp. WLY78 enables expression of active nitrogenase in *Escherichia coli*. *PLOS Genet.* **9**, e1003865 (2013).
6. Xie, J.B. et al. Comparative genomic analysis of N$_2$-fixing and non-N$_2$-fixing *Paenibacillus* spp.: organization, evolution and expression of the nitrogen fixation genes. *PLOS Genet.* **10**, e1004231 (2014).
7. Shi, H. et al. Genome-wide transcriptome profiling of nitrogen fixation in *Paenibacillus* WLY78. *BMC Microbiol.* **16**, 25 (2016).
8. Liu, H. et al. Complete genome sequence of *Paenibacillus polymyxa* YC0136, a plant growth-promoting rhizobacterium isolated from tobacco rhizosphere. *Genome Announc.* **5**, e01635-16 (2017).
9. Seefeldt, L.C., Yang, Z.Y., Duval, S. & Dean, D.R. Nitrogenase reduction of carbon-containing compounds. *BBA* **1827**, 1102-1111 (2013).
10. Peng, R. et al. Metabolic engineering of Arabidopsis for remediation of different polycyclic aromatic hydrocarbons using a hybrid bacterial dioxygenase complex. *Eng.*** **26**, 100-110 (2014).
11. Le, A. et al. Glucose-independent glutamine metabolism via TCA cycling for proliferation and survival in B cells. *Cell Metab.* **15**, 110–121 (2012).
12. Gleizer, S. et al. Conversion of *Escherichia coli* to generate all biomass carbon from CO$_2$. *Cell* **179**, 1255–1263 (2019).
13. Dixon, R. & Kahn, D. Genetic regulation of biological nitrogen fixation. *Rev. Microbiol.* **2**, 621-631 (2004).
14. Dos Santos, P.C. et al. Distribution of nitrogen fixation and nitrogenase-like sequences amongst microbial genomes. *BMC Genom.* **13**, 162 (2012).
15. Leigh, J.A. Nitrogen fixation in methanogens: the archaeal perspective. *Rev. Microbiol.* **2**, 125-131 (2000).
16. Galloway, J.N. et al. Transformation of the nitrogen cycle: Recent trends, questions, and potential solutions. *Science* **320**, 889-892 (2008).
17. Tilman, D., Balzer, C., Hill, J. & Befort, B.L. Global food demand and the sustainable intensification of agriculture. *Natl. Acad. Sci. USA* **108**, 20260-20264 (2011).
18. Oldroyd, G.E.D., Harrison, M.J. & Paszkowski, U. Reprogramming plant cells for endosymbiosis. *Science* **324**, 753-754 (2009).
19. Marchetti, M. et al. Experimental evolution of a plant pathogen into a legume symbiont. *Biol.* **8**, e1000280, 1-10 (2010).
20. Guan, S.H. et al. Experimental evolution of nodule intracellular infection in legume symbionts. *ISME J.* **7**, 1367-1377 (2013).
21. Godfroy, O., Debelle, F., Timmers, T. & Rosenberg, C. A rice calcium- and calmodulin-dependent protein kinase restores noduleation to a legume mutant. *Plant. Microbe. In.* **19**, 495-501 (2006).
22. Chen, C., Gao, M., Liu, J. & Zhu, H. Fungal symbiosis in rice requires an ortholog of a legume common symbiosis gene encoding a Ca$^{2+}$/calmodulin-dependent protein kinase. *Plant Physiol.* **145**, 1619-1628 (2007).
23. Jones, K.M., Kobayashi, H., Davies, B.W. & Walker, G.C. How rhizobial symbionts invade plants: the Sinorhizobium-Medicago *Nat. Rev. Microbiol.* **5**, 619-633 (2007).
24. Deakin, W.J. & Broughton, W.J. Symbiotic use of pathogenic strategies: rhizobial protein secretion systems. *Rev. Microbiol.* **7**, 312-320 (2009).
25. Heath, K.D. & Tiffin, P. Stabilizing mechanisms in a legume rhizobium mutualism. *Evolution* **63**, 652-662 (2009).
26. Oldroyd, G.E.D, Murray, J.D., Poole, P.S. & Downie, J.A. The rules of engagement in the legume-rhizobial symbiosis. *Rev.Genetic.* **45**,119-144 (2010).
27. van Velzena, R. et al. Comparative genomics of the nonlegume Parasponia reveals insights into evolution of nitrogen-fixing rhizobium symbioses. *Natl. Acad. Sci. USA* **115**, 4700-4709 (2018).
28. Masson-Boivin, C. & Sachs, J.L. Symbiotic nitrogen fixation by rhizobia- the roots of a success story. *Opin. Plant Biol.* **44**, 7-15 (2018).
29. Dong W. et al. An SHR–SCR module species legume cortical cell fate to enable nodulation. *Nature* **589**, 586-590 (2021).
30. Ivleva, N.B., Groat, J., Staub, J.M. & Stephens, M. Expression of active subunit of nitrogenase via integration into plant organelle genome. *PLOS one* **11**, e0160951 (2016).
31. Allen, R.S. et al. Expression of 16 nitrogenase proteins within the plant mitochondrial matrix. *Plant. Sci.* **8**, 00287 (2017).
32. Hoover, T.R., Imperial, J., Ludden, P.W. & Shah, V.K. Homocitrate is a component of the iron-molybdenum cofactor of nitrogenase. *Biochemistry* **28**, 2768-2771 (1989).

33. Raymond, J., Siefert, J.L., Staples, C.R. & Blankenship, R.E. The natural history of nitrogen fixation. *Biol. Evol.* **21**, 541-554 (2004).

34. Hu, Y. et al. Assembly of nitrogenase MoFe protein. *Biochemistry* **47**, 3973-3981 (2008).

35. Rubio, L.M. & Ludden, P.W. Biosynthesis of the iron-molybdenum cofactor of nitrogenase. *Rev. Microbiol.* **62**, 93-111 (2008).

36. Yang, J. et al. Reconstruction and minimal gene requirements for the alternative iron-only nitrogenase in *Escherichia coli*. *Natl. Acad. Sci. USA* **111**, 3718-3725 (2014).

37. Yang, J. et al. Polyprotein strategy for stoichiometric assembly of nitrogen fixation components for synthetic biology. *Proc. Natl. Acad. Sci. USA* **115**, 8509-8517 (2018).

38. Thorneley, R.N. & Lowe, D.J. Nitrogenase of *Klebsiella pneumoniae*. Kinetics of the dissociation of oxidized iron protein from molybdenum-iron protein: Identification of the rate-limiting step for substrate reduction. *J.* **215**, 393-403 (1983).

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

40. Burén, S. et al. Purification and in vitro activity of mitochondria targeted nitrogenase cofactor maturase NifB. *Front. Sci.* **8**, 1567 (2017).

41. Burén, S. et al. Biosynthesis of the nitrogenase active-site cofactor precursor NifB-co in *Saccharomyces cerevisiae*. *Natl. Acad. Sci. USA* **116**, 25078-25086 (2019).

42. Naqvi, S. et al. Transgenic multivitamin corn through biofortification of endosperm with three vitamins representing three distinct metabolic pathways. *Natl. Acad. Sci. USA* **106**, 7762-7767 (2009).

43. Zorrilla-López, U. et al. Engineering metabolic pathways in plants by multigene transformation. *J. Dev. Biol.* **57**, 565-576 (2013).

44. Edgren, T. & Nordlund, S. Two pathways of electron transport to nitrogenase in *Rhodospirillum rubrum*: the major pathway is dependent on the fix gene products. *FEMS Microbiol. Lett.* **260**, 30-35 (2006).

45. Hanke, G. & Mulo, P. Plant type ferredoxins and ferredoxin-dependent metabolism. *Plant, Cell Environ.* **36**, 1071-1084 (2013).

46. Yang, J. et al. Modular electron-transport chains from eukaryotic organelles function to support nitrogenase activity. *Natl. Acad. Sci. USA* **114**, 2460-2465 (2017).

47. Sippel, D. & Einsle, O. The structure of vanadium nitrogenase reveals an unusual bridging ligand. *Chem. Biol.* **13**, 956-960 (2017).

48. Rohde, M. et al. Crystal structure of VnfH, the iron protein component of vanadium nitrogenase. *Biol. Inorg. Chem.* **23**, 1049-1056 (2018).

49. Hernandez, J.A. et al. NifX and NifEN exchange NifB cofactor and the VK-cluster, a newly isolated intermediate of the iron-molybdenum cofactor biosynthetic pathway. *Microbiol.* **63**, 177-192 (2007).
50. Morel, J. B., Mourrain, P., Béclin, C., & Vaucheret, H. DNA methylation and chromatin structure affect transcriptional and post-transcriptional transgene silencing in Arabidopsis. *Biol. 10*, 1591-1594 (2000).

51. Lou, S.K. et al. Whole-genome bisulfite sequencing of multiple individuals reveals complementary roles of promoter and gene body methylation in transcriptional regulation. *Genome Biol. 15*, 408 (2014).

52. Sun, D. et al. Kinase SnRK1.1 Regulates nitrate channel SLAH3 engaged in nitrate-dependent alleviation of ammonium toxicity. *Plant Physiol.* doi: 10.1093/plphys/kiab057 (2021).

53. Willison, J.C. & Vignais, P.M. The use of metronidazole to isolate Nif- mutants of *Rhodopseudomonas capsulata*, and the identification of a mutant with altered regulatory properties of nitrogenase. *Gen. Microbial. 128*, 3001-3010 (1982).

54. Montoya, J.P., Voss, M., Kahler, P. & Capone, D.G. A Simple, high-precision, high-sensitivity tracer assay for N2 *Appl. Environ. Microbiol. 62*, 986-993 (1996).

55. Lambers, H. (1980). The physiological significance of cyanide-resistant respiration in higher plants. *Plant Cell Environ.* 3, 293-302.

56. Arnold, W. et al. Nucleotide sequence of a 24,206-base-pair DNA fragment carrying the entire nitrogen fixation gene cluster of *Klebsiella pneumoniae*. *Mol. Biol. 203*, 715-738 (1988).

57. Jacobson, M.R. et al. Physical and genetic map of the major nif gene cluster from *Azotobacter vinelandii*. *Bacteriol. 171*, 1017-1027 (1989).

58. Kaiser, J.T. et al. Structure of precursor bound NifEN: a nitrogenase FeMo cofactor maturase/insertase. *Science 331*, 91-94 (2011).

59. Kennedy, C. & Dean. D. The nifU, nifS and nifV gene products are required for activity of all three nitrogenases of *Azotobacter vinelandii*. *Gen. Genet. 231*, 494-498 (1992).

60. Andi, B., West, A.H. & Cook, P.F. Stabilization and characterization of histidine-tagged homocitrate synthase from *Saccharomyces cerevisiae*. *Biochem. Biophys. 421*, 243-254 (2004).

61. Nouwen, N. et al. The role of rhizobial (NifV) and plant (FEN1) homocitrate synthases in *Aeschynomene* photosynthetic *Bradyrhizobium Sci. Rep. 7*, 448 (2017).

62. Liu, X., Zhao, X., Li, X. & Chen, S. Suppression of hesA mutation on nitrogenase activity in *Paenibacillus polymyxa* WLY78 with the addition of high levels of molybdate or cystine. *Peer J. 7*, e6294 (2019).

63. Deitung, J. & Thorneley, R.N.F. Electron transfer to nitrogenase. *J. 239*, 69-75 (1986).

64. Karlusich, J.J.P., Lodeyro, A.F. & Carrillo, N. The long goodbye: the rise and fall of flavodoxin during plant evolution. *Exp. Bot. 65*, 5161-5178 (2014).

65. Tegeder, M. & Masclaux-Daubresse, C. Source and sink mechanisms of nitrogen transport and use. *New Phytol. 217*, 35–53 (2018).

66. Gleizer, S. et al. Conversion of Escherichia coli to generate all biomass carbon from CO2. *Cell 179*, 1255–1263 (2019).
Methods

Construction of plasmids for the reconstitution of nitrogenase in plants

The genomic DNA of *Paenibacillus polymyxa* YC0136 (CP017967) was used to search the nif cluster. All nine genes were chemically synthesized via the polymerase chain reaction (PCR)-based accurate synthesis (PAS) method. To obtain the most favorable expression levels of the various genes necessary to produce nitrogenase, the nine chemically synthesized genes were placed under the control of the CaMV 35S Ω promoter and the NOS terminator. Every gene was seamlessly connected with the promoter and terminator to form gene expression cassettes. Four nif-gene expression plasmids were constructed from the pYP674 plasmid as follows: pYB654, containing a four-gene expression cassette (35S:nifB:nos: 35S:nifH:nos: 35S:nifD:nos: 35S:nifK:nos), pYB1538, containing a six-gene expression cassette (35S:nifB:nos: 35S:nifH:nos: 35S:nifD:nos: 35S:nifK:nos: 35S:nifE:nos: 35S:nifN:nos), pYB2082, containing a nine-gene expression cassette (35S:nifB:nos: 35S:nifH:nos: 35S:nifD:nos: 35S:nifK:nos: 35S:nifE:nos: 35S:nifN:nos: 35S:nifX:nos: 35S:nifV:nos: 35S:hesA:nos), and pYB1876, containing a three-gene expression cassette (35S:nifX:nos: 35S:nifV:nos: 35S:hesA:nos). ETC-gene expression plasmid YB60 contained the expression cassette of ferredoxin and ferredoxin oxidoreductase from *Pseudomonas pudita* (35S:ferredoxin:nos: 35S:ferredoxin oxidoreductase:nos) and plasmid YB9771 contained the expression cassette of flavodoxin and flavodoxin oxidoreductase from *Klebsiella pneumoniae* (35S:flavodoxin:nos: 35S:flavodoxin oxidoreductase:nos).

Transformation of plants

The plasmids were transformed into *Agrobacterium* strain GV3101 (pMP90) using electroporation and then transformed into *Arabidopsis thaliana* using the floral dip method mediated by *Agrobacterium*. Transgenic *Arabidopsis* was selected on solid medium (1/2 MS containing 50 mg/L hygromycin B). Homozygous genotypes of the transgenic plants were obtained through self-fertilization and confirmed by segregation analysis. Homozygous transgenic plants were used to detect the expression and activity of nitrogenase. Different crosses were obtained using homozygous nif transgenic *Arabidopsis* as the
female parent and ETC transgenic *Arabidopsis* containing ferredoxin and ferredoxin oxidoreductase of naphthalene dioxygenase from *Pseudomonas putida*\(^1\) as the male.

**Southern blot analysis**

Genomic DNA was extracted from wild-type *Arabidopsis* and transgenic lines using the Qiagen DNEasy Plant Maxi Kit and quantified on a Nanodrop 2000 spectrophotometer (Thermo). Genomic DNA was digested with *Eco*RI or *Hind*III and precipitated with ethanol. The digestes were loaded on 0.8% agarose gels with DIG-labeled MarkerII (Roche). Gels were run overnight and blotted for 4 h using the alkaline transfer method with a HyBond N\(^+\) membrane (Amersham). DNA was UV cross linked to the membrane. The probe sequences of the *snifD* gene were amplified via PCR from pYB2802 using oligonucleotide primers 5'‘- CACTGGCTAAGGGTATCTC-3’ and 5'‘-CATCTGACGGAATGGAAT-3’. The probes were labeled using the Roche DIG Probe Synthesis Kit, and the DIG procedure was performed according to the manufacturer’s recommendations.

**Identification of T-DNA insertion sites in transgenic plants by genome resequencing\(^74\)**

Three genomic DNA samples from different transgenic lines were mixed equally with approximately 2 g of DNA per sample. The DNA was randomly sheared into 200–300 bp fragments by using a Covaris E210 ultrasonicator (Covaris, Inc., Woburn, MA, USA). The ends of the fragments were converted to blunt ends using the T4 DNA polymerase and DNA polymerase I Klenow fragments, after which the fragments were ligated with adaptors and amplified using DNA polymerase to generate amplicon libraries according to the instructions of the Nextera DNA Library (Illumina Corporation, USA). Short read sequence data were generated using the Illumina HiSeq 2500 platform (Hangzhou Guhe Information Technology Co., Ltd.). Each read was 150 bp in length, and at least 12 GB of data were obtained. The Q30 value of each sample was up to 80 (the general perception of a high error rate [> 0.1%] is less than 20%).

The T-DNA sequence of the vector was compared with the sequence obtained by using Bowtie2 software (http://Bowtie-bio.sourceforge.net/bowtie2/index.shtml). The obtained sequence was further assembled and screened to determine the sequence containing the vector sequences. A possible T-DNA insertion site was identified by BLAST searches at the *Arabidopsis* genome website. The genomic DNA of the transgenic lines was verified by PCR. Primers were designed based on genomic DNA sequences near T-DNA insertion sites and vector sequences.

**Real-time fluorescent quantitative PCR assay and northern blotting**
Total RNA was extracted from 100 mg samples of the transgenic plants using the Universal Plant Total RNA Extraction Kit (spin-column) from Shenggong (Shanghai, China) following the protocol provided by the manufacturer. The possibility of contamination of genomic DNA was eliminated by digestion with RNase-free DNase I (Takara Bio). cDNA was synthesized in a reaction volume of 20 µl by using a cDNA Synthesis Kit (Takara, China).

Real-time PCR was carried out in an iCycler IQTM real-time detection system (BioRad), and the Premix Ex Taq™ protocol (Takara) was used throughout in 25 µl reactions with minor changes. The reactions contained 12.5 µl of SYBR® Premix Ex Taq™ II (2×), 1 µl of the forward and reverse primers at 10 µM, 8.5 µl of sterilized deionized water and 2 µl of cDNA. The reaction conditions were as follows: 94 °C predenaturation for 5 min and 40 cycles of 94 °C for 1 min, 60 °C for 30 s and 60 °C for 15 s. The real-time assays were analyzed using iCycler iQ version 3.1 system software (BioRad). To generate standard quantitation curves, Ct values were plotted proportionally to the logarithm of the input copy numbers. The coefficients of variation (CV) were calculated by dividing the standard error by the mean Ct value for two replicates of each sample from three separate experiments to evaluate reproducibility. Water and healthy plants were used as checks. An *A. thaliana* actin gene (*AtAc2*, accession number NM112764) synthesized with two primers (AtAc2Z1: 5'-GCACCCTGTTCTTCTTACCGAG-3'; AtAc2F1: 5'-AGTAAGGTCACGTCCAGCAG-3') was used as an internal standard gene. The experiments were repeated three times.

RNA (15 µg) was separated on 1% denaturing agarose–formaldehyde gels. Equal loading was confirmed by staining the gels with ethidium bromide. After the RNA was transferred to nylon membranes, it was probed with digoxigenin (DIG)-labeled cDNA probes obtained by PCR (PCR DIG Probe Synthesis Kit, Roche, Mannheim, Germany). To amplify the respective probes, sequence-specific primers were used. Colorimetric detection was performed using nitro blue tetrazolium (NBT) and 5-bromo-4-chloro-3-indolyl-phosphate (BCIP) as substrates for alkaline phosphatase. Quantitative analysis of the northern blot results was performed by using Gel Analyser 2010.

**Western blot analysis**

To demonstrate the production of Nif proteins, the total protein fraction generated from the *Arabidopsis* seedlings was extracted with SDS-Tris buffer (0.1 M Tris-HCl, 5% SDS, 2% β-mercaptoethanol and pH 6.8) and adjusted for consistency of the protein concentration. For western blot analysis, the protein samples were separated on a 3% gel and a 10%–15% separation gel using Tris/glycine/SDS buffer. The protein was transferred to 0.2 µm nitrocellulose (NC) membranes using a Trans-Blot Turbo Transfer System (Bio-Rad). Three percent skim milk was used as the blocking agent, and TBST (10 mM Tris, 150 mM NaCl, 0.05% Tween-20, pH 7.6) was used as the washing buffer. Specific antibodies (Abclonal Biotechnology Co, China) raised in rabbit against nine different nif proteins were used for analysis at a 1:1,000 dilution. The NC membranes were fixed, closed, treated with the anti-nif protein antibody and then incubated
overnight at 4 °C. Finally, the NC membranes were reacted with a horseradish peroxidase-conjugated goat anti-rabbit IgG antibody and colored via the ECL method.

**Immunogold labeling** of *Nif* proteins

To determine the location of the expressed proteins, leaf tissue (2×2 mm) was cut, fixed overnight in 0.5% (v/v) glutaraldehyde and 4% (v/v) paraformaldehyde in a vacuum bottle and then washed three times in 0.1 M phosphate buffer (423 mM NaH$_2$PO$_4$, 577 mM Na$_2$HPO$_4$, pH 7.2). The samples were fixed with 1% (w/v) osmic acid, serially dehydrated in 30%, 50%, 70%, 90% and 100% ethanol and then embedded in LR white resin (London resin company, England). Ultrathin resin sections of approximately 70–80 nm were cut and transferred to grids. TEM sections were labeled with the protein-A conjugate Au reagent (Aurion, US) and antibodies specific to each *nif* protein. The grids were blocked with PBS supplemented with 5% BSA and 0.1% fish skin gelatin and then transferred to drops of specific primary antibodies. After washing six times with PBS, the grids were transferred to the appropriate Au-conjugated reagent for 2 h. Finally, the grids were washed with PBS, post-fixed in 2% glutaraldehyde and stained with 2% (w/v) uranyl acetate and lead citrate. Leaf sections were observed at 100 KV using a Hitachi HM7100 TEM.

**Methylation analysis of the CaMV 35S promoter**

The sulfonation of genomic DNA was carried out according to the instructions of the EZ DNA Methylation–Gold Kit. The CaMV 35S promoter was amplified with the primers 35S-MZ: AGATAGTGAAACAGGAGGTGG and 35S-MF: CCTCTCCAAATRAAATRAACTTCCT using a methylated PCR kit (Tiangen, Beijing). The reaction system consisted of 1 µl of primers (10 µM), 1.6 µl of dNTPs (2.5 mM), 1 µl of MSP DNA polymerase (2.5 U/µl), 2 µl of MSP PCR buffer, 1 µl of DNA and 13.4 µl of H$_2$O. The amplified products were cloned, sequenced and analyzed.

**Nitrogenase activity analysis *in vivo***

Seeds of *A. thaliana* were surface sterilized with bleaching powder (5%, w/v) for 20 min, washed three times with sterile water and grown on solid MS medium with 0–200 mg/L KNO$_3$, 15 µM Na$_2$MoO$_4$, 50 mM Fe(III)C$_6$H$_5$O$_7$ and no NH$_4$NO$_3$. The plants were grown in a petri dish or flask under the following conditions: 22 °C, an 8 or 16 h photoperiod and light intensity of 550 µmol/m$^2$/s.

To assay nitrogenase activity, transgenic plants were transferred to nitrogen-deficient or low-nitrogen medium at the germination and seedling stages. To test the inhibition of oxygen on nitrogenase, we cultured plants in a sealed bag (Mitsubishi Gas Chemical Company, Japan), in which the mixed gas with
different concentrations of oxygen were filled. After 2 weeks, the transgenic Arabidopsis plants were used for biomass, chlorophyll, proteins in solution and free amino acid analyses.

Chlorophyll content was determined spectrophotometrically at 649 and 665 nm, according to the Lichtenthaler and Welburn method\textsuperscript{77}. The protein content was determined from the OD\textsubscript{280}\textsuperscript{78}; bovine serum albumin was used as the standard. Ammonium concentrations in Arabidopsis seedlings were determined by fluorescence spectroscopy at a neutral pH based on \textit{o}-phthalaldialdehyde (OPA)\textsuperscript{79}. Arabidopsis seedlings (50 mg) were ground and homogenized with 1 ml of cold 10 mM formic acid solution. The homogenate was centrifuged at 13,000×g (4 °C) for 15 min, and the supernatant was transferred to a new eppendorf tube and stored on ice. OPA (16 mg) was dissolved in ethanol, and 40 ml of 0.1 M phosphate buffer containing 10 mM \textit{β}-mercaptoethanol (pH 6.8). A 10 µl volume of the extract was mixed with 400 µl of OPA solution, reacted at 80 °C for 15 min and immediately cooled on ice. Fluorescence was measured at 470 nm with excitation at 410 nm using an F-2700 fluorescence spectrophotometer (Hitachi, Japan). Ultrapure (NH\textsubscript{4})\textsubscript{2}SO\textsubscript{4} was used as the standard.

Cyanide (CN\textsuperscript{−}) is an inhibitor of electron flow and is toxic to plant cells. To assay nitrogenase activities \textit{in vivo}, transgenic Arabidopsis plants were transferred to low-nitrogen medium containing 10 mg/L KCN, and the KCN resistance of seedlings was investigated. Similar to cyanide, sodium azide is a highly toxic cellular respiratory inhibitor that can hinder seed germination and growth. Azide can also act as a nitrogenase substrate; it is catalytically reduced to produce N\textsubscript{2}H\textsubscript{2} and NH\textsubscript{3}. Thus, transgenic Arabidopsis plants were transferred to low-nitrogen medium containing 0.65 mg/L NaN\textsubscript{3} to investigate its effects on seed germination and plant growth.

\textbf{\textsuperscript{15}N\textsubscript{2} incorporation assay}

Seeds of \textit{A. thaliana} were surface sterilized and then sown in 50 ml flasks containing 20 ml of solid nitrogen-deficient MS medium with 9 µM Na\textsubscript{2}MoO\textsubscript{4} and 50 mM Fe(III)C\textsubscript{6}H\textsubscript{5}O\textsubscript{7}. The flasks were sealed with a rubber stopper. Fifty percent of the air in the flasks was replaced with \textsuperscript{15}N\textsubscript{2} (99%+, Shanghai Engineering Research Centre for Stable Isotope). After 30 days of incubation at 25 °C, the cultivated samples were dried at 80 °C and gradually ground into fine powder.

Five milligrams of sample and 50 mg of copper oxide particles were put into a glass reformer tube, and then the air in the tube was extracted with a vacuum system. The sample reformer tube was sealed when its vacuum degree reached 0.01Pa. The sealed tube was heated in a muffle furnace at 530 °C for 4 h. After the reaction, the tube was cooled for mass spectrometry analysis. The gas stream from the reformer tube was injected under the conditions of high vacuum and low seepage rate into the ion source of a Finnigan MAT-271 mass spectrometer for isotope ratio measurement. Each test contained 3 samples and was repeated 3 times. The isotope ratios of low (\textsuperscript{15}N=10.0%) and high abundance\textsuperscript{15}N labeled semicarbazide (\textsuperscript{15}N>99.0%, Shanghai Research Institute of Chemical Industry) provided a check on the
internal precision of the mass spectrometer. The isotopic composition of $^{15}\text{N}$ was determined by the isotope ratio of $^{15}\text{N}/^{14}\text{N}$. Due to the low-level $^{15}\text{N}$ content in the sample, isotope ratios were expressed as $\delta^{15}\text{N}\%$ values:

$$\delta^{15}\text{N}\% = \left( \frac{^{15}\text{N}/^{14}\text{N}}{^{15}\text{N}/^{14}\text{N}}_{\text{atmosphere}} - 1 \right) \times 1000,$$

which represents the per mille difference between the isotope ratios in a sample and in atmospheric $\text{N}_2$, where $(^{15}\text{N}/^{14}\text{N})_{\text{atmosphere}} = 0.366 \text{ atom}\%$.

**In vitro acetylene reduction assay**

For nitrogenase activity assays, plants grown in soil (9:3:1 mixture of peat moss/vermiculite/perlite) were transferred to an anaerobic chamber, irrigated with PNS medium containing 9 µM $\text{Na}_2\text{MoO}_4$ and 50 mM Fe(III)$_6\text{C}_6\text{H}_5\text{O}_7$ and incubated for 3 days under hypoxic conditions (10% $\text{O}_2$, 0.3% $\text{CO}_2$, balance $\text{N}_2$). Protein was extracted with plant extraction buffer containing Tris/HCl (0.1 M, pH 8.0), sodium dithionite (2 mM) and dithiothreitol (0.5 mM) and then centrifuged at 12,000 rpm for 15 min. To maintain anaerobic conditions during protein extraction, the samples and buffers were maintained inside either a bag or sealed centrifuge tubes filled with argon (Ar). All samples and buffers were washed out with Ar to remove $\text{O}_2$ from the solutions and then stored at 4 °C throughout the assay.

**In vitro** nitrogenase activities were tested by using an ATP-regeneration system with dithionite as the artificial electron donor$^{52}$. Exactly 0.2 ml of crude protein and 0.8 ml of the enzyme reaction solution containing ATP, MgCl$_2$, creatine phosphate (Sigma), creatine phosphokinase (Sigma, 324 u/mg) and 40 mM MOPS-KOH (pH 7.4) were placed in a serum bottle (10 ml), which was sealed with a rubber plug and deoxygenated several times with high-purity Ar. $\text{C}_2\text{H}_2$ (10% of the headspace volume) was injected into the test tubes. After incubating the cultures at 30 °C with shaking at 250 rpm for 1 h, the reaction was stopped with 30% TCA. Thereafter, 1 ml of the culture headspace was withdrawn through the rubber stopper with a gas-tight syringe and manually injected into an Agilent 7890B gas chromatograph to quantify ethylene production. All treatments were performed with three replicates, and all experiments were repeated at least three times.

**Sample preparation for GC-MS analysis of isotope-labeled amino acids and chlorophyll**

Four-week-old *Arabidopsis* seedlings were transferred to 50 ml flasks containing nitrogen-deficient MS medium with 9 µM $\text{Na}_2\text{MoO}_4$ and 50 mM Fe(III)$_6\text{C}_6\text{H}_5\text{O}_7$ and grown for 30 days. Then 20% air in the flasks was replaced with isotopic-labeled $\text{N}_2$ (99%+, Shanghai Engineering Research Centre for Stable Isotope). Seedlings were grown under such condition for one week. For isotopic analysis of amino acids, samples
must be pretreated with silylation reactions. After freeze-drying treatment, the sample from 3 different seedlings was homogenized with distilled deionized water at a ratio of 1/25 (W/V) and sonicated for 30 min in an ice bath. The homogenized samples were then centrifuged for 10 min at 12000 r/min. The supernatant was transferred to a new glass tube and treated with vacuum drying. The dried sample was efficiently derivatized by N-methyl-N-(trimethylsilyl) trifluoroacetamide (MSTFA)/pyridine (500 µL) at 60 °C for 1h, filtered through a 0.22 micron hole membrane filter, and injected (10 µL) into a gas chromatography-mass spectrometer (GC-MS) (Agilent 7890B/5977A, Agilent, Santa Clara, CA) equipped with an HP-5MS capillary column (30 m ×0.25 mm ×0.25 µm). For isotopic analysis of chlorophyll, total chlorophyll was extracted with 96% ethanol according to the Lichtenthaler and Welburn method. Helium (99.999%) was used as the carrier gas at a constant flow rate of 1.0 mL/min. The GC oven temperature programmed from 50 °C (3 min) to 160 °C at 40 °C/min, ramped from 160 °C to 250 °C at 10 °C/min, continued to 300 °C at 20 °C/min, and then held at 300 °C for 16 min. The obtained derivative of 15N-labeled amino acid was detected by selected ion monitoring mode between 50 and 500 m/z and chlorophyll detected by ion monitoring mode between 850 and 950 m/z.

Isotopic analysis of different amino acids and chlorophyll

Total isotope abundance of compounds was calculated by the isotope peak cluster algorithm. The characteristic fragment ions, which contain N and have a base peak or a higher peak intensity, were chosen for isotopic abundance calculation. There are different numbers of N atoms labeled by isotopes for different compounds. The 15N isotopic labeling number of amino acid (except for Gln, Asn, Trp, and basic amino acids) is only one, but the number is four for chlorophyll. For example, C5H1414NSi+; the alanine derivatives of MSTFA have a higher peak intensity of the characteristic ion M/Z 116 and has natural isotopic distribution 116 and 117. In sample, the peak intensity of isotopic distribution included the peak intensity of natural fragment and 15N labeled fragment. In order to obtain the isotopic abundance of alanine in the sample, the ratio of natural isotopic distribution of the alanine peak clusters with M/Z 116 and 117 must be calculated firstly according to the electron impact (EI) spectra of GC-MS.

If the peak intensity of natural alanine represented as $A_0^{116}$: $A_0^{117}$=a:b (a+b=1), the peak intensity of alanine in sample was:

$$A_{mix}^{116} = X_0 \cdot a$$  \hspace{1cm} (1)

$$A_{mix}^{117} = X_0 \cdot b + X_1 \cdot a$$  \hspace{1cm} (2)

where $A_{mix}^{116}$ and $A_{mix}^{117}$ are the detected peak intensity of alanine in sample at M/Z 116 and 117, respectively.

The 15N isotopic abundance of alanine can be obtained from formula (3).
\[^{15}\text{N} \text{ atom } \% = \frac{X_1}{(X_0 + X_1)} \] \hspace{1cm} (3)

The peak cluster of chlorophyll a includes 893.5, 894.5, 895.5, 896.5, and 897.5. If the ratio of the natural isotopic distribution of the standard of chlorophyll a is a:b:c:d:e, the \[^{15}\text{N} \text{ isotope } \% \text{ abundance of chlorophyll a in the sample can be obtained from formula (4).} \]

\[^{15}\text{N} \text{ atom } \% = \frac{(1 \cdot X_1 + 2 \cdot X_2 + 3 \cdot X_3 + 4 \cdot X_4)}{4(X_0 + X_1 + X_2 + X_3 + X_4)} \] \hspace{1cm} (4)

\[A_{\text{mix}}^{894.5} = X_0 \cdot a \] \hspace{1cm} (5)

\[A_{\text{mix}}^{895.5} = X_0 \cdot b + X_1 \cdot a \] \hspace{1cm} (6)

\[A_{\text{mix}}^{896.5} = X_0 \cdot c + X_1 \cdot b + X_2 \cdot a \] \hspace{1cm} (7)

\[A_{\text{mix}}^{897.5} = X_0 \cdot d + X_1 \cdot c + X_2 \cdot b + X_3 \cdot a \] \hspace{1cm} (8)

\[A_{\text{mix}}^{898.5} = X_0 \cdot e + X_1 \cdot d + X_2 \cdot c + X_3 \cdot b + X_4 \cdot a \] \hspace{1cm} (9)

where \(A_{\text{mix}} \) is the peak intensity of the measured mass spectrum for the sample.

The \[^{15}\text{N} \text{ isotope } \% \text{ abundance of chlorophyll b can be determined in the same way.} \]

1. Xiong, AS. et al. PCR-Based Accurate Synthesis of Long DNA Sequences. *protoc.* 1, 791-797 (2006).
2. Peng, R., Xiong, A. & Yao, Q. A direct and efficient PAGE-mediated overlap extension PCR (POEP) method for gene multiple-site mutagenesis. *Appl. Microbiol.* 73, 234-240 (2006).
3. Clough, S.J., Bent, A.F. Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. *Plant J.* 16, 735-743 (1998).
4. Polko, JK. et al. Illumina sequencing technology as a method of identifying T-DNA insertion loci in activation-tagged Arabidopsis thaliana plants. *Mol Plant* 5, 948–950 (2012).
5. Vernooy-Gerritsen, M., Leunissen, J.L., Veldink, G.A. & Vliegenthart, J.F.: Intracellular localization of lipoxygenases-1 and -2 in germinating soybean seeds by indirect labeling with protein A colloidal gold complexes. *Plant Physiol.* 76, 1070-1078 (1984).
6. Gruntman, E. et al. Analyzer of plant methylation states through bisulfite sequencing. *BMC Bioinformatics* 9, 371 (2008).
7. Lichtenthaler, H.K. & Welburn, A.R. Determination of total carotenoids and chlorophylls a and b of extracts in different solvents. *Soc. Trans.* 11, 591-592 (1983).
8. Pace, C.N. et al. How to measure and predict the molar absorption coefficient of a protein. *Protein Sci.* 4, 2411-2423 (1995).
9. Husted, S. et al. Photorespiratory \(\text{NH}_4^+\) production in leaves of wild-type and glutamine synthetase. *Plant Physiol.* 130, 989-998 (2002).
80. Le, W., Hou, J., Du, X.N. & Zhang, W.B. Determination of isotope abundance of $^{15}$N labeled amino Acids by gas chromatography-mass spectrometry. *Chinese J. Anal. Chem.* **43**, 1465-1471 (2015).

81. Antonovsky, N. et al. Sugar Synthesis from CO$_2$ in Escherichia coli. *Cell* **166**, 115–125 (2016).

**Declarations**

**SUPPLEMENTAL INFORMATION**

The Supplemental information includes extended experimental procedures, results, discussion `and figures and can be found with this article online.

**ACKNOWLEDGMENTS**

The research was supported by the Key Project Fund of the Shanghai Municipal Committee of Agriculture (chuangzi 2018 1-5), Shanghai Academic Technology Research Leader (19XD1432300), Shanghai Natural Science Foundation (18ZR1413100) and National Natural Science Foundation (31672439). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

The authors wish to acknowledge Dr. Wen Lei, Jie Hou and Yan-Qiu Chen from the Shanghai Research Institute of Chemical Industry Co., Ltd. for their help in the qualitative and quantitative analysis of different isotopologues, and Dr. Zhi Liu from Viva Biotech Ltd. of China for help in performing the Immune electron microscopy sample preparation and observation in this study.

**Figures**
Figure 1

The formation of nitrogenase by the minimal nif gene cluster in Paenibacillus sp. and strategy for the construction of nitrogen-fixing plants. (A) Modular arrangement of genes in Paenibacillus sp. nif gene cluster. Letters within the box represent the corresponding nif genes. (B) Maturation of functional nitrogenase. NifB generates NifB-co with the help of HesA. Then, NifB-co binds NifX with high affinity and transfers to NifEN, where it is converted into the VK-cluster; the new FeMo-co precursor shows electronic paramagnetic resonance (EPR). The VK-cluster on NifEN ultimately transforms FeMo-co into apo-nitrogenase to generate mature nitrogenase. (C) Structural comparison of the Fe and MoFe proteins of Paenibacillus nitrogenase with Azotobacter vinelandii nitrogenase (PDB code 6q93 and 5N6Y). Both compounds of the A. vinelandii nitrogenase proteins are shown in blue. The Fe protein and MoFe protein α subunit of Paenibacillus nitrogenase is shown in red, and the MoFe protein β subunit is shown in yellow. (D) Structures of the T-DNA region of pYB2082 containing nif genes of Paenibacillus sp used in the transformations into Arabidopsis. LB, left border; RB, right border; p, promoters. (E) Representative Southern blots of independent transgenic T2-plants. E, EcoRI; H, HindIII. (F) Sequence information of insertion sites obtained from the analysis of genome resequencing data.
Figure 2

Overexpression of nif genes in transgenic Arabidopsis. (A) Northern blot analyses (30 µg of total RNA per lane) of the expression of nine transgenes in wild-type (CK) and transgenic (404-37, 405-54, 405-58) plants. rRNA stained with ethidium bromide was used as the loading control. “+”, special fragment from each gene. (B) qRT-PCR analysis of all nine genes of the nif cluster in Arabidopsis plants. The Atactin2 gene was used as the reference gene for normalization. (C) Dot map of the methylation of the Ca MV35S
Comparison of the growth of engineered nif gene-carrying Arabidopsis with wild-type under low nitrogen conditions. (A, B) Arabidopsis seedlings were grown in MS medium containing 50 mg/L KNO3 (A) or no nitrogen (B) for 15 days. (C) Six-leaf-stage Arabidopsis plants were transplanted to MS medium with low-nitrogen conditions (50 mg/L KNO3) for 30 days. (D, E) Comparison of the fresh weight (D) and proteins in solution (E) of wild-type and nif gene-transgenic plants. Seedlings were grown under 50 mg/L KNO3 or nitrogen-free conditions for 15 days. (F, G) Comparison of the total chlorophyll content of CK and nif gene-transgenic seedlings grown in MS medium containing 50 mg/L KNO3 (F) and six-leaf-stage plants grown in MS medium containing 50 mg/L KNO3 (G). (H, I) Assay of nitrogenase activities in engineered nif gene-carrying plants using the method of acetylene reduction (H) or 15N2 incorporation (I). Error bars indicate the standard deviation observed from three independent experiments.
Figure 4

Tolerance of engineered nif gene-carrying plants to toxic cellular respiratory inhibitors. (A) Arabidopsis seedlings were grown in medium containing 0.65 mg/L NaN3, 1.35 mg/L NaN3, 10 mg/L KCN or 20 mg/L NaN3 for 15 days. Here, 50 mg/L KNO3 was used as the nitrogen source in the medium. (B) Mortality of seedlings grown in different concentrations of toxic cellular respiratory inhibitors (0.65 mg/L NaN3, 1.35 mg/L NaN3, 10 mg/L KCN or 20 mg/L NaN3). (C) Arabidopsis seedlings were grown in low-nitrogen (30 mg/L KNO3) or high-nitrogen (200 mg/L KNO3) medium under the stress of 0.65 mg/L NaN3 or 10 mg/L KCN for 30 days. (D) Comparison of the fresh weights of wild-type and engineered nif
gene seedlings when grown in medium containing different concentrations of nitrogen under the stress of 0.65 mg/L NaN3 (top) or 10 mg/L KCN (bottom).

Figure 5

Engineered plants and their nitrogen-fixing abilities. (A) Scheme showing the genetic organization of the engineered plants. (B, C) Different engineered and wild-type Arabidopsis seedlings were grown in MS medium, low-nitrogen MS medium containing 30 mg/L KNO3, low-nitrogen MS medium containing 0.65
mg/L NaN3 or low-nitrogen MS medium containing 10 mg/L KCN for 30 days. (D, E) Comparison of the fresh weight and total chlorophyll content of wild-type and different engineered plants grown in low-nitrogen medium for 30 days. (F) Comparison of the fresh weight of wild-type and different engineered plants grown in low-nitrogen medium containing NaN3 and KNO3 for 30 days. (G) The addition of an electron transfer component (ETC) of ferredoxin and ferredoxin reductase improves the growth of engineered nif gene-carrying plants in low-nitrogen medium and their resistance to toxic cellular respiratory inhibitors (NaN3 and KNO3). ETC was added to engineered nif gene-carrying plants by crossing ferredoxin and ferredoxin reductase-containing plants as the male parent and nif gene-transgenic plants as the female parent. (H) The insertion of ETC of ferredoxin and ferredoxin reductase into engineered nif gene-carrying plants increased their fresh weight. Plants were grown in low-nitrogen medium containing NaN3 and KNO3 at 26 °C for 30 days. (I) The insertion of ETC of ferredoxin and ferredoxin reductase into the engineered nif gene-carrying plants improved their nitrogen-fixing abilities. (J) No significant difference in effects on the active level of nitrogenase between the ETC of ferredoxin and flavodoxin. FD, ferredoxin and ferredoxin reductase; FJ, flavodoxin and flavodoxin oxidoreductase. Error bars indicate the standard deviation observed from three independent experiments.
Isotopic labeling experiments using 15N2 show that N2 are the nitrogen donors of amino acids and chlorophyll. (A) The wide distribution of 15N to various amino acids and chlorophyll indicated that N2 can be fixed and involved in nitrogen metabolism of plants. The amino acids and chlorophyll with higher 15N isotope abundance in nif gene-transgenic plants are indicated in red. (B) Isotope abundance for 15N-
labeled amino acids and chlorophyll. Values are based on GC-MS analysis of stable 15N-labeled amino acids and chlorophyll. Data are presented as the mean of two independent experiments.

**Supplementary Files**

This is a list of supplementary files associated with this preprint. Click to download.

- [suplementaryinformationnature.docx](#)
- [rawdatafile1.xlsx](#)
- [rawdatafile2.xlsx](#)