Global investigation of an engineered nitrogen-fixing Escherichia coli strain reveals regulatory coupling between host and heterologous nitrogen-fixation genes

Zhimin Yang1,2, Yunlei Han2, Yao Ma2, Qinghua Chen2, Yuhua Zhan2, Wei Lu2, Li Cai2, Mingsheng Hou2, Sanfeng Chen3, Yongliang Yan2 & Min Lin2

Transfer of nitrogen fixation (nif) genes from diazotrophs to amenable heterologous hosts is of increasing interest to genetically engineer nitrogen fixation. However, how the non-diazotrophic host maximizes opportunities to fine-tune the acquired capacity for nitrogen fixation has not been fully explored. In this study, a global investigation of an engineered nitrogen-fixing Escherichia coli strain EN-01 harboring a heterologous nif island from Pseudomonas stutzeri was performed via transcriptomics and proteomics analyses. A total of 1156 genes and 206 discriminative proteins were found to be significantly altered when cells were incubated under nitrogen-fixation conditions. Pathways for regulation, metabolic flux and oxygen protection to nitrogenase were particularly discussed. An NtrC-dependent regulatory coupling between E. coli nitrogen regulation system and nif genes was established. Additionally, pentose phosphate pathway was proposed to serve as the primary route for glucose catabolism and energy supply to nitrogenase. Meanwhile, HPLC analysis indicated that organic acids produced by EN-01 might have negative effects on nitrogenase activity. This study provides a global view of the complex network underlying the acquired nif genes in the recombinant E. coli and also provides clues for the optimization and redesign of robust nitrogen-fixing organisms to improve nitrogenase efficiency by overcoming regulatory or metabolic obstacles.

In nature, a variety of genes and islands can be rapidly and frequently horizontally transferred among bacteria, resulting in the acquisition of certain properties such as nitrogen fixation, antimicrobial resistance and pathogenesis, which help bacteria to succeed in altered habitats or new niches1–3. However, newly acquired genes or islands become a burden for bacteria if they are not properly integrated with host regulatory systems. A greater understanding of the physiological alterations that occur in an engineered cell following the insertion of large fragments of foreign DNA, especially from distant species, is needed to pave the way towards the goal of biological engineering.

Biological nitrogen fixation is catalyzed, in most cases, by the molybdenum nitrogenase encoded by a highly conserved nifHDK gene cluster. Previous studies have shown that the nif genes encoding active nitrogenase can be transferred to non-nitrogen-fixing prokaryotes to impart the ability to reduce atmospheric nitrogen gas into ammonia as a nitrogen source4–12. From the perspective of synthetic biology, one key goal of studying biological
nitrogen fixation is to facilitate the introduction of this ability into organisms of great importance for human beings, for instance, engineering autonomous nitrogen-fixing cereal crops. Eventually, a successfully engineered N₂-fixing non-legume crop may significantly cut down the use of chemical fertilizers for a cleaner environment and higher yields.

In recent years, the synthesis of nitrogen-fixing systems has become increasingly common due to advances in synthetic biology. Normally, a functional entity or pathway is first detected in prokaryotes before being transferred to eukaryotes and plants. Because of its well-studied genetic background, *Escherichia coli* is frequently used as the preferred first-step research model. Following the pioneering work on nitrogen-fixation engineering in 1970s, several groups have reported successful gene transfer of *nif* genes to *E. coli* in the past five years. However, these recombinant *E. coli* strains showed much lower nitrogenase activity compared with the original host, and the horizontally acquired ability was insufficient to enable diazotrophic growth on nitrogen-free medium, implying the presence of (i) regulatory coupling between the host and heterologous nitrogen-fixation systems, as well as (ii) a regulatory/or metabolic barrier that results in reduced nitrogenase activity in the engineered cells. To date, how the non-diazotrophic host maximizes opportunities to fine-tune the acquired capacity for nitrogen fixation has not yet been fully explored.

*Pseudomonas stutzeri* A1501 is a root-associated bacterium that exhibits an unusual feature, for a *Pseudomonas* strain, the ability to fix nitrogen. Similar to previous reports, the nitrogenase activity of recombinant *E. coli* strain EN-01 was much lower than that of A1501. In the present study, to better understand the global regulatory effect of the host on NFI expression, we monitored the global transcriptional and proteomic profiles of recombinant *E. coli* grown anaerobically under nitrogen-fixation and nitrogen-repression conditions. Furthermore, the metabolic flux shift of *E. coli* EN-01 under nitrogen-fixation conditions was also determined by HPLC. To the best of our knowledge, this is the first report of a global investigation into the regulatory cascade of *nif* genes in an engineered nitrogen-fixing organism. These data are particularly useful for providing a more comprehensive understanding of how the *E. coli* host intervenes in the transcriptional regulation of the “foreign” NFI to support a functional nitrogenase complex. Our results will also help guide efforts to more successfully remodel and optimize similar systems in other species.

**Results**

**Overview of the *E. coli* EN-01 transcriptome and proteome under nitrogen-fixation and nitrogen-repression conditions.** The expression of 1156 genes was significantly altered (≥2-fold, P ≤ 0.05) under nitrogen-fixation conditions compared with nitrogen-repression conditions, including 789 up-regulated and 367 down-regulated genes. The up-regulated genes mainly belonged to three major functional categories: nitrogen metabolism, transport or membrane protein and unknown function, while the down-regulated genes were mainly involved in energy synthesis, transport, protein synthesis and regulation. These altered genes were further classified according to the COG functional classification system. As shown in Fig. 1, 259 genes, of which 102 were down-regulated and 157 were up-regulated, were involved in bacterial preservation and processing of genetic information (DNA replication, duplication, repair and transcription, expression, etc.). A total of 361 genes (31% of the total altered genes) were involved in transport and metabolic pathways. An additional 90 genes were involved in energy synthesis and transformation processes, with 67 genes induced under nitrogen-fixation conditions (Fig. 1A). Moreover, the expression of 151 genes encoding proteins of unknown function was also significantly altered.

Transcriptome analysis has revealed that 255 genes were significantly up-regulated and 294 genes were severely down-regulated in *P. stutzeri* A1501 under nitrogen-fixation conditions. Subsequent comparison of the *E. coli* EN-01 and *P. stutzeri* A1501 transcriptomes revealed that at least 24 genes in the *E. coli* EN-01 transcriptome showed a similar expression pattern to that in *P. stutzeri* A1501 (Supplementary Table S1), including the glnK-amtB operon, the two-component regulatory system *ntrBC* and the serine protein kinase-coding gene *prkA*, implying that the overlapping genes or systems might be essential for NFI expression in *E. coli*.

Proteomic analysis of *E. coli* EN-01 was also performed under identical conditions. A total of 110 protein spots increased by more than 2-fold under nitrogen-fixation conditions; notably, 55 proteins were only expressed under nitrogen-fixation conditions. In addition, 96 proteins were significantly down-regulated under nitrogen-fixation conditions, and the expression of 24 protein spots was abolished on the 2D-PAGE gel (Supplementary Fig. S1). A total of 138 proteins were further identified by MALDI-TOF mass spectrometry, of which 63 were up-regulated and 75 were down-regulated (Supplementary Table S2). In addition, 132 identified proteins showed similar expression patterns to the transcriptional analysis. These proteins included the glutamine synthetase (GS) GlnA, nitrogen regulatory PI I-pl membrane protein GlnK, hydperoxidase II Kat E, oligopeptidase ABC transporter OppA, glucose-6-phosphate 1-dehydrogenase Zwf and the dihydrolipoamide dehydrogenase LpdA.

Taken together, the above data revealed dramatic changes to the global transcriptional regulatory networks and cellular metabolic pathways in the recombinant nitrogen-fixing *E. coli* EN-01 under nitrogen-fixation conditions. Subsequent analyses focused on the nitrogen regulatory system, energy production and conversion systems, and the oxygen protection-related pathway, which have been reported to be crucial for nitrogen fixation.
Expression patterns of *E. coli* nitrogen metabolic pathways and regulatory system under nitrogen-fixation conditions.

Expression of *P. stutzeri* NFI genes in recombinant *E. coli* strain EN-01. EN-01 exhibits nitrogenase activity under glucose-fermentation and nitrogen-fixation conditions, indicating the heterologous production of a functional nitrogenase complex [12]. To further test the heterologous expression of nif genes, we determined the induction ratio of 43 genes within the island using quantitative real-time RT-PCR (Table 1). Previously, we found that the A1501 NFI genes are organized into 11 NifA-σ54-dependent operons and are up-regulated under nitrogen-fixation conditions [24]. Similarly, except for a 7-gene operon (PST1302-1306), all NifA-σ54-dependent operons were strongly induced in *E. coli*, suggesting that the nif-specific transcriptional activator NifA is controlled by the *E. coli* nitrogen regulatory system. Consistent with the RT-PCR results, proteomic analysis also revealed that the NifS, NifD and PST1336 proteins encoded by the NFI genes were up-regulated under nitrogen-fixation conditions (Fig. 2). Why the PST1302-1306 operon (including the nifQ and nifB genes) was not regulated in EN-01 is still unclear. The nifQ mutant of *Klebsiella pneumoniae* is defective in nitrogen fixation due to an elevated requirement for molybdenum [29]. In addition, nifB has been long recognized as crucial.

**Figure 1.** Overview of expression profiling analysis in recombinant nitrogen-fixing *E. coli* EN-01. (A) Functional categories of nitrogen fixation-induced genes (P ≤ 0.05 and fold change ≥ 2) in *E. coli* EN-01. (B) Functional categories of core subset of down-regulated genes (P ≤ 0.05 and fold change ≥ 2) under nitrogen-fixation conditions. The percentage of genes in each section is depicted.
for nitrogen fixation because NifB participates in the biosynthesis of the FeMo-co factor\(^\text{30}\). We therefore postulate that the relatively low nitrogenase activity of EN-01 might be partly attributed to low expression of these factors within the operon.

**Regulatory coupling between the E. coli general nitrogen regulatory system and the heterologous P. stutzeri nif island.** Regulation of the nitrogen-fixation process in free-living diazotrophs facilitates the stringent control that is necessary to maximize the physiological benefits from diazotrophy\(^\text{28}\). In *P. stutzeri* A1501, the nitrogen regulatory cascade comprises the AmtB–GlnK–NtrBC global nitrogen regulation proteins, which sense the nitrogen signal and subsequently control the expression of the *nif*-specific regulatory proteins NifLA\(^\text{31}\). However, the nitrogen regulatory network in *E. coli* appears to be an expanded version compared to that in *Pseudomonas*.
with one additional PII protein and cascade regulation achieved by a Nac regulatory protein.\textsuperscript{32,33} Acquisition of a nif-specific regulatory system in the recombinant \textit{E. coli} raises the question of nitrogen-fixation efficiency by these two regulatory systems of different evolutionary origins.

Changes in the expression of \textit{E. coli} nitrogen regulatory system genes are summarized in Supplementary Table S3. Of these, the \textit{glnLG} genes, which are homologs of \textit{Pseudomonas ntrBC}, were up-regulated by more than 10-fold, the sigma factor \textit{rpoN} was up-regulated by 3.64-fold, the ammonium transporter \textit{amtB} was up-regulated by 55.28-fold, and the two PII protein genes \textit{glnK} and \textit{glnB} were both up-regulated under nitrogen-fixation conditions. Consistent with the transcriptional change, proteomic data indicated that GlnK and GlnA expression was enhanced (Supplementary Table S2). Based on the observations in the present study and those from previous experiments from a wide range of model systems, we propose a regulatory cascade that controls heterologous nif transcription in \textit{E. coli} EN-01 (Fig. 3): The UTase/URase enzyme GlnD (\textit{glnD} gene was up-regulated 5.27-fold) is activated by low internal glutamine concentrations and further modifies the PII protein GlnB (\textit{glnB} was up-regulated 2.12-fold) and GlnK (\textit{glnK} was up-regulated 86.88-fold) to PII-UMP. Subsequently, PII-UMP phosphorylates NtrC. As a global transcriptional regulator, NtrC activates the expression of a series of nitrogen

Figure 2. Enhanced expression of nitrogen fixation related proteins in recombinant \textit{E. coli} EN-01 under nitrogen-fixation conditions. Magnified regions of 2-D gel images were sliced from Supplementary Fig. S1. Protein spots of interest are indicated with circles with the corresponding protein names pointed out on the left. Proteins were identified by MALDITOF-MS analysis.
metabolic genes, including the GS-encoding gene glnA (up-regulated 7.61-fold), the nitrogen transcriptional regulation (Ntr) genes, the ammonium transporter gene amtB (up-regulated 55.28-fold), the transcriptional regulator nac (up-regulated 124.98-fold) and the nif-specific regulator genes nifLA. In fact, we previously demonstrated cross-complementation of the P. stutzeri/or E. coli ntrC (glnG) mutant by the E. coli/or the P. stutzeri ntrC gene. We have also shown that the purified E. coli NtrC protein can bind specifically to the nifLA operon promoter region. In a sense, this working model supports a similar expression pattern for the NFI genes in E. coli with that observed in P. stutzeri A1501, in that the nif genes were subject to regulation by the NtrC protein via the nif-specific regulator NifA, suggesting a regulatory coupling of these two different evolutionary systems through a direct activating interaction.

The PII protein is recognized as a critical signal transduction protein in bacterial nitrogen metabolism. E. coli encodes two PII counterparts, GlbK and GlnK, which show 68% and 78% identity, respectively, with the sole A1501 PII protein GlnK and share 67% identity with each other (Fig. 4). Under nitrogen-fixation conditions, E. coli EN-01 glnK mRNA expression was increased by 86.88-fold, while the glnB gene was up-regulated by 2.12-fold. GlnK protein was only detected under nitrogen-fixation conditions (Fig. 2). In A1501, an interaction between GlnK and the C-terminal domain of NifL was observed using the yeast two-hybrid system, suggesting Glk-dependent control of NifA activity by NifL. Based on the overall transcription of the nitrogen regulation system in E. coli EN-01, we conjectured that, as reported before, GlbK and GlnK may constitute a two-tiered regulatory system and that their functions may depend on the timing of expression and the levels of accumulation, while glnK may play key roles in nitrogen regulation in E. coli to adjust to perturbations that occur under nitrogen-fixation conditions.
Rearrangement of the pathways for nitrogen assimilation and alternative nitrogen sources. Glutamate and glutamine provide approximately 75% and 25% of cellular nitrogen, respectively. Nitrogen assimilation must therefore result in the synthesis of these two nitrogen donors. The GDH pathway, encoded by the glutamate dehydrogenase gene gdhA, is generally associated with a nitrogen-rich, aerobic environment, whereas the GS-GOGAT pathway, encoded by the GS gene glnA and the glutamate synthase (GOGAT) genes gldB, is often associated with a nitrogen-limited, microaerobic environment. Transcriptomic analyses showed that both operons of the GS-GOGAT system were up-regulated more than 7-fold under nitrogen-fixation conditions, however, the gdhA gene was down-regulated by 2.83-fold (Supplementary Table S3). In accordance with the mRNA results, we also observed that GlnA was up-regulated 7.61-fold and GdhA, encoding GDH, was down-regulated 9.1-fold in the proteomic analysis (Supplementary Table S2). Clearly, the recombinant E. coli EN-01 chooses the high-affinity GS-GOGAT pathway rather than the GDH pathway when environmental nitrogen is depleted.

Zimmer et al. previously concluded that E. coli use a range of pathways to scavenge for nitrogen-containing compounds as a first line of defense against nitrogen starvation. In this work, we observed a rearrangement of transport capacity by EN-01 under nitrogen-fixation conditions, with alterations of 299 transport or membrane protein genes. Notably, the ammonium transporter gene amtB, which forms a transcriptional operon with the PII protein-coding gene gltK, showed a 55-fold enhancement (Fig. 3). Glutamate uptake systems, including the gltJK operon, which encodes a glutamate/aspartate uptake system; the glnK operon, which is responsible for a high-affinity glutamine transport system; the gadC gene, which encodes a glutamic acid/γ-aminobutyrate antiporter; and the ghp operon, which encodes a proton/glutamate-aspartate symporter, were up-regulated under nitrogen-fixation conditions. Alternative nitrogen-transport systems, for instance, lysine, arginine, ornithine, histidine, serine, polyamine, peptide, dipeptide and oligopeptide transporter genes, were all significantly induced. Additionally, the astABCDE operon, which is responsible for the degradation of arginine into \( \text{N}_2 \)-succinyl-L-ornithine, CO₂ and ammonia, was also up-regulated (Supplementary Table S3).

Nac is a gltK-controlled Ntr protein widely existed in enterobacteria. Nac is well known to be involved in regulating the expression of multiple pathways to modulate the levels of important cellular metabolic intermediates and thus integrates nitrogen assimilation with other aspects of metabolism. As expected, nac gene expression was drastically up-regulated (over 100-fold) under nitrogen-fixation conditions, gabDTP (γ-aminobutyrate transport and degradation), dppABD (dipeptide transport), and fklB (peptidyl-prolyl cis-trans isomerase) gene expression was also increased (Supplementary Table S3). However, codBA (cytosine metabolism), nupC (nucleoside transport) and ydkSTUVW (putative putrescine transport and degradation) gene expression did not appear to be significantly altered. These genes were recognized to be controlled by the nac gene in E. coli. The unclear inconsistent gene expression obtained in this study may provide important clues into the regulatory disturbances of the NFI island by the nac-conferred nitrogen stress response in recombinant E. coli EN-01.

Metabolic flux shift of E. coli EN-01 under nitrogen-fixation conditions. Biological nitrogen fixation is an energy-dependent process that requires ATP, produced by the catalytic reaction of carbon sources, to break the \( \text{N} = \text{N} \) bond. Heterologous NFI expression may place a considerable metabolic burden on the host E. coli. In the present work, EN-01 was observed to shift its metabolic flux to overcome the energy barrier for nitrogen fixation. As the central carbon pathways constitute the backbone of cell metabolism by providing energy, metabolic building blocks, and reducing power for biomass synthesis, we investigated the responses of central carbon metabolism in EN-01. As shown in Fig. 5, a stronger flux toward the pentose phosphate pathway (PPP) was found under nitrogen-fixation conditions. Thus, the PPP was postulated to serve as the primary route for glucose catabolism, since the key genes encoding the rate-limiting enzymes glucose-6-phosphate 1-dehydrogenase (zwf), transketolase (tktAB), transaldolase (talA) and glucose-6-phosphate isomerase (pgi) were highly induced. The enhanced expression of glucose-6-phosphate 1-dehydrogenase was also detected in proteome analyses (Supplementary Table S2). Conversely, the glyoxylate shunt and Entner-Doudoroff pathway were found to be inhibited in EN-01 under nitrogen-fixation conditions. Furthermore, the electron transfer systems commonly involved in energy production were up-regulated, such as the appABC (encoding cytochrome bd II oxidase), hyaABCDCE (encoding the hydrogenase), fhaA (encoding the transcriptional activator of formate hydrogenlyase), fdhF and hycABCDEFGH (encoding formate hydrogenlyase complex), ndh (encoding the NADH dehydrogenase), napH (encoding the ferredoxin-type protein), and adhE (encoding the bifunctional acetalddehyde-CoA/alcohol dehydrogenase) genes. Meanwhile, both the transcription and translation of the formate acetyltransferase gene pfIB were enhanced (Supplementary Table S2). Additionally, three alcohol dehydrogenase genes that are reportedly involved in ethanol fermentation were up-regulated.

E. coli is an anaerobic fermentative bacterium that produces a variety of organic acids by utilizing glucose under nitrogen-fixation conditions; however, these acids are harmful to nitrogenase activity. In contrast to E. coli, P. stutzeri A1501 is an alkali-producing strain that supplies an optimal environment for nitrogen fixation at pH 7.0–9.0 (Supplementary Fig. S2). The organic acid content in E. coli wild-type and EN-01, as well as P. stutzeri A1501, were determined by HPLC after growth under nitrogen-fixation conditions. E. coli wild-type and EN-01 produced a large number of organic acids under nitrogen-fixation conditions (Supplementary Table S4), including pyruvic acid, succinic acid, lactic acid, formic acid and acetic acid, whereas A1501 does not generate these organic acids. The presence of these organic acids could certainly have negative effects on nitrogenase activity in the recombinant EN-01 strain. Altering the intracellular acid environment may be a feasible method to increase the nitrogen-fixing ability of E. coli EN-01.

Oxygen protection strategy of E. coli EN-01 to cope with \( \text{O}_2 \) toxicity to nitrogenase. Nitrogenase is extremely oxygen sensitive. In fact, purified nitrogenase, regardless of its source, undergoes extremely rapid and irreversible inactivation by \( \text{O}_2 \). E. coli EN-01 can grow under both aerobic and anaerobic conditions but only executes nitrogen-fixing activity anaerobically. Given the importance of oxygen in nitrogen fixation,
we explored the potential mechanism for oxygen protection in *E. coli* EN-01 under anaerobic nitrogen-fixing conditions. In the present study, the scavenging capacity of *E. coli* EN-01 for superoxide radicals, hydrogen peroxide, and hydroxyl radicals appeared to increase under nitrogen-fixation conditions. The genes *sodC* and *katE*, which encode superoxide dismutase and the hydroperoxidase HPII, were up-regulated 2.94-fold and 15.47-fold, respectively, under nitrogen-fixation conditions (Supplementary Table S3). Furthermore, proteomic analysis revealed these two gene products were increased as well (Supplementary Table S2). Other genes such as the hydrogenase 1-coding genes *hyaABCDEF*, cytochrome bd-II synthetic-related genes, which have higher oxygen affinity than ATP terminal oxidase, and all cytochrome c synthetic genes, were also up-regulated. The *osmC* gene, whose product is involved in the detoxification of organic hydroperoxides, was also up-regulated 5.81-fold under nitrogen-fixation conditions (Supplementary Table S3). In addition, the *yhbO* gene was up-regulated 7.5-fold, and its product was only detected under nitrogen-fixation conditions (Supplementary Table S2). YhbO has been shown to function in the response to various stresses. Mutating the *yhbO* gene in *E. coli* leads to increased sensitivity to heat, oxidative, hyperosmotic, pH and UV stresses.

Although no significant changes were detected at the mRNA level, proteomic analysis revealed that Dps was induced 4.06-fold under nitrogen-fixation conditions (Supplementary Table S2). The Dps protein, a nonspecific DNA-binding protein, is reported to reduce the production of oxidative radicals through ferroxidase activity and therefore to protect the cell from oxidative stress, UV irradiation, iron and copper toxicity, among others. Taken together, the genes or proteins mentioned above, despite their unclear functions, are thought to be involved in the oxygen stress response under nitrogen-fixation conditions and to simultaneously protect the nitrogenase complex from *O*₂ inhibition by providing a tolerable oxygen environment in *E. coli* EN-01.

Figure 5. Bioreaction metabolic flux shifts of EN-01 central carbon metabolism under nitrogen-fixation conditions. Changes in the central carbon metabolism network were constructed from microarray data. Arrows indicate the physiological directions of reactions. Red arrows, enhanced expression (P ≤ 0.05 and fold change ≥ 2); blue arrows, reduced expression (P ≤ 0.05 and fold-change ≥ 2). The accumulation of organic acids highlighted in gray was detected in the medium by HPLC. Abbreviations: G6P, glucose 6-phosphate; F6P, fructose 6-phosphate; F1,6P2, fructose 1,6-phosphate; G3P, glyceraldehyde 3-phosphate; 1,3-BPG, 1,3 diphosphoglycerate; 3PG, 3-phosphoglycerate; PEP, phosphoenolpyruvic acid; PYR, pyruvate; AcCoA, acetyl coenzyme A; OAA, oxaloacetate; ICT, isocitrate; AKG, α-ketoglutarate; SUC, succinate; MAL, malate; GLX, glyoxylic acid; 6PG, 6-phosphogluconate; E4P, erythrose 4-phosphate; Ru5P, ribulose-5-phosphate; R5P, ribose-5-phosphate; X5P, xylulose-5-phosphate; 2-keto-3-deoxy-6-phosphogluconic acid; Acetyl-P, acetyl phosphate.
Discussion

To date, the ability to fix nitrogen is exclusively found among bacteria, including green sulfur bacteria, firmibacteria, actinomycetes, cyanobacteria and all subdivisions of proteobacteria. Genomic analyses have revealed that nitrogen-fixation genes are clustered in most diazotrophs. The nif genes were further presumed to have been lost in most bacterial and archaeal lineages during evolution and that horizontal gene transfer played a pivotal role in the recent acquisition of nif genes in some species. In agreement with this hypothesis, several groups have found that transferring nif genes from non-nitrogen fixers successfully yields functional recombinant nitrogen-fixing strains, albeit with far lower nitrogenase activity compared with the donor strains. Therefore, for recombinant nitrogen-fixing strains harboring a heterologous gene system, a variety of metabolic penalties and regulatory barriers may exist between the two systems. Understanding how to overcome and escape the metabolic penalties and regulatory barriers is essential to fine-tune acquired nitrogen-fixing capacity of the non-diazotrophic host.

In this work, transcriptomic and proteomic analysis separately identified 1156 genes and 138 proteins whose expression was significantly altered when E. coli EN-01 was incubated under nitrogen-fixation conditions. Our results highlight the metabolic penalties and regulatory barriers for nitrogen signal transduction and metabolism brought by excess fixed nitrogen in the medium. These genes and proteins, particularly the 24 genes that overlapped between the two transcriptomic analyses (the previous transcriptomic analysis of the EN-01 strain with the E. coli analysis presented in this work), are potential targets for subsequent modification in E. coli EN-01 to optimize nitrogen-fixing capacity.

The nitrogen regulatory cascade of P. stutzeri A1501 comprises the AmtB–GlnB–NtrBC general nitrogen regulation proteins and the nif-specific regulatory protein NifLA. E. coli has a similar nitrogen regulatory system to P. stutzeri but harbors an additional PII protein and a cascade regulation by the Nac regulatory protein. In the present work, a similar expression network of NFI genes in E. coli to that operating in P. stutzeri A1501 was predicted based on the expression patterns of rpoN, glnD, glnA, glnK, ntrBC, nifLA and other nitrogen regulation-related genes (Fig. 3). Given that the purified E. coli NtrC protein can bind specifically to the A1501 nifLA operon promoter region, we postulated the regulatory coupling of these two evolutionary divergent systems through a direct activating interaction. Additionally, the nac gene was also predicted to be involved in nif gene regulation in E. coli although the mechanism requires further investigation. Interestingly, an expression difference was observed in the two PII counterparts GlnB and GlnK in E. coli. In principle, the functions of GlnK and GlnB overlap in Enterobacteriaceae. However, both proteins exhibited distinct physiological roles in the regulation of nitrogen assimilation over a wide range of environmental conditions, despite having distinct expression patterns. The glnB promoter contains a putative -10 sequence for RpoS activation, while the transcription of glnK is induced by NtrC and RpoN. We therefore presumed that glnK may play a more important role than GlnB in nitrogen regulation in E. coli under nitrogen-fixation conditions. Notably, the PST1302-1306 cluster within the NFI showed unregulated expression in E. coli EN-01 but was highly up-regulated in P. stutzeri A1501 under nitrogen-fixation conditions. The weak transcription of these nitrogenase component genes may predict imperfections and the necessity for further rebuilding of nif gene expression regulation in E. coli EN-01.

A core challenge for diazotrophs is the need to generate sufficient energy to drive nitrogen fixation. In the present work a strong shift in the metabolic flux of the EN-01 strain was observed under nitrogen-fixation conditions. The key enzyme of the PPP was up-regulated, but the glyoxylate shunt and Entner-Doudoroff pathway were inhibited. Thus, the PPP pathway was concluded to serve as the primary route for glucose catabolism to support nitrogen fixation. Moreover, anaerobic respiration-related electron transfer systems and substances were induced. These data highlight the energy acquisition pathway of EN-01 to adapt to nitrogen fixation. However, as an anaerobic fermentative bacterium, EN-01 produces a large number of organic acids in the media, which may negatively affect nitrogenase activity, since the donor strain P. stutzeri A1501 is an alkali strain. Changing the intracellular acid environment may thus be a feasible method to increase the nitrogen-fixing ability of E. coli EN-01.

We identified a set of genes that might be involved in protecting nitrogenase from oxygen via scavenging of superoxide radicals, hydrogen peroxide, and hydroxyl radicals. Two proteins, YhbO and Dps, also appeared to be involved in the oxygen protection of nitrogenase by protecting the cell from oxidative stress. These genes or proteins may protect the nitrogenase complex from O$_2$ inhibition by establishing a tolerable oxygen environment in E. coli EN-01. Therefore, we speculate that improving the expression of antioxidant capacity-related pathway genes may satisfy the need for oxygen protection under microaerobic conditions and perhaps drive the metabolic flux shift from anaerobic respiration to aerobic respiration to produce more energy and reduce the energy requirements of the nitrogen fixation.

The nitrogenase activity of E. coli EN-01 was much lower than that in P. stutzeri A1501. According to the data in this work, this difference may be due to the regulatory divergence of nitrogen assimilation, the inactivation of the gene operon PST1302-1306, and the production of organic acids. Another factor may also be involved. In our recent work, a species-specific regulatory mRNA nfiS identified in P. stutzeri A1501, could directly base paired with the mRNA of the nitrogenase component nifK to enhance translation efficiency and transcript half-life, thereby regulating nitrogenase biosynthesis. Blast results showed that no sequence homologous to A1501 nfiS was present in the E. coli genome. Introduction of nfiS regulation into E. coli EN-01 may be a useful way to refine NFI gene regulation.

From an agri-biotechnology perspective, the successful engineering of an N$_2$-fixing organism would significantly reduce the need for nitrogen fertilizers. Biosynthetic techniques have been used by several groups to improve or optimize nitrogenase activity in metabolically engineered strains. Wang X. et al. replaced the nif regulatory elements in the recombinant E. coli strain with a T$_7$ RNA polymerase–LacI expression system to overcome regulatory barriers, and the newly constructed T$_7$-dependent nif system bypassed the original complex.
regulatory circuits with minor physiological limitations. Subsequent work from the same group pinpointed the electron-transfer components from plant organelles that can be used to support nitrogenase activity and reductase from various target genes required to engineer nitrogen fixation in plants. Another group observed that five separate gene clusters, as well as a combined cluster from *Paenibacillus sp.* WLY78 and *Klebsiella oxytoca*, increase the nitrogenase activity of the recombinant *E. coli* 78-7, which harbors a minimal *nif* gene cluster from *Paenibacillus sp.* WLY78.

In summary, the results presented in this work demonstrate the overall expression pattern of *E. coli* EN-01 genes under nitrogen fixation and revealed the regulatory coupling of host genes and the heterologous NFI that supports functional nitrogenase activity in the engineered *E. coli* strain. We also uncovered the metabolic penalties and regulatory barriers that lead to the reduced nitrogenase activity in *E. coli* EN-01 and provided clues for biotechnological applications to generate new hypotheses concerning gene regulation and flux, which may lead to an improved ability to fix nitrogen in *E. coli*. This work may shed light on the metabolic penalties and coordinated patterns between host and horizontally transferred genes and provide the knowledge necessary to pave the way for engineered nitrogen fixation.

### Methods

#### Bacterial strains and growth conditions.

The strains and plasmids used in this study are described in Table S5. The recombinant strain *E. coli* EN-01 bearing an NFI was grown at 37°C in LB medium. Chloramphenicol, tetracycline, and hygromycin were added to media at a concentration of 20 μg/mL, 10 μg/mL, and 100 μg/mL, respectively, as required. *P. stutzeri* A1501 was cultured at 30°C in LB medium or minimal lactate-containing medium (medium K) as described previously.

#### Nitrogenase activity assays.

To measure nitrogenase activity in *E. coli* EN-01, cells were cultured overnight to an OD600 of 2.0 at 37°C in LB and 0.2 mL cultures were directly added to 6 mL NFD medium containing 20 mM ammonium (nitrogen-repression conditions) or without ammonium (nitrogen-fixation conditions) in 10 mL serum bottles. These cultures were sealed and incubated at 30°C for 20 h with constant shaking (200 rpm), after which 10% acetylene without oxygen was injected to determine ethylene production by gas chromatography, as described by Cannon et al. The nitrogenase activity of *P. stutzeri* was determined in N-free minimal medium at an OD600 of 0.1 at 30°C under an argon atmosphere containing 0.5% oxygen and 10% acetylene, according to the protocol described by Desnoues et al. Nitrogenase-specific activity is expressed as nmol ethylene per mg protein per h. Each experiment was repeated at least three times.

#### Total RNA preparation.

*E. coli* EN-01 cells were cultured overnight to an OD600 of 2.0 at 37°C in LB, and 0.2 mL cultures were directly added to 6 mL NFD medium containing 20 mM ammonium (nitrogen-repression conditions) or without ammonium (nitrogen-fixation conditions) in 10 mL serum bottles, respectively. These cultures were sealed and incubated at 30°C for 20 h with constant shaking (200 rpm). The cell densities of EN-01 cultures were sealed and incubated at 30 °C for 20 h with constant shaking (200 rpm). The cell densities of EN-01 cultures were sealed and incubated at 30 °C for 20 h with constant shaking (200 rpm). The cell densities of EN-01 cultures were sealed and incubated at 30 °C for 20 h with constant shaking (200 rpm). The cell densities of EN-01 cultures were sealed and incubated at 30 °C for 20 h with constant shaking (200 rpm).

#### Microarray analysis.

For microarrays, standard methods were used for cDNA synthesis, fragmentation, and terminal biotin labeling based on Affymetrix protocols. Labeled cDNA was hybridized to the Affymetrix *E. coli* Genome 2.0 array. Hybridized arrays were stained with streptavidin-phycocerythrin using an Affymetrix Fluidic Station. After staining, arrays were scanned with an Affymetrix GeneChip Scanner 3000 based on the total signal intensity. The resulting microarray data were analyzed using Affymetrix software (MAS 5.0). Consensus “detection p-values”, “change p-values”, and “mean expression ratios” were calculated. All signal intensities with mean expression ratios above 2 were considered significant changes if the p-value was below 0.05. Complete microarray data have been deposited in the Gene Expression Omnibus (GEO) database under accession number GSE37780.

#### Quantitative real-time PCR.

To investigate the expression of NFI genes in *E. coli* EN-01, real-time quantitative PCR (qPCR) was performed using an ABI 7500 Real-Time PCR System. Primers pairs used for qPCR were designed using DNAman and are listed in Table S6. RNA was extracted from cells treated under the same conditions described for the microarray analyses. Reverse transcription was performed using a cDNA synthesis kit (Promega, USA). SYBR Green premix was used to detect PCR amplification. The 16S RNA gene was used to normalize the results. Additionally, qPCR was also used to verify microarray results.

#### Proteome analysis by two-dimensional PAGE.

As previously described, *E. coli* EN-01 cells were cultured under the anaerobic nitrogen-fixation or nitrogen-repression conditions. The cells were harvested via high-speed centrifugation (12,000 × g, 5 min and 4 °C). Cell pellets were washed three times with PBS buffer and suspended in lysis buffer (8 M urea, 2 M thiourea, 0.5% w/v CHAPS, 2% v/v carrier ampholyte, 1% w/v DTT and 1 mM PMSF) for sonication on ice. The samples were centrifuged at 14,000 × g for 30 min. Subsequently, cell pellets were rapidly frozen in liquid nitrogen and stored at −80°C. Total RNA was isolated using TRIzol Reagent (Invitrogen, Carlsbad, CA). To avoid possible DNA contamination, an additional DNase I digestion was performed at 37°C for 30 min, and the samples were chilled on ice. Total RNA was purified using RNeasy columns, resuspended in RNase-free water and quantitated on a NanoVue plus.
(UMAX Technologies Inc., Dallas, TX) and the gel images were analyzed with PDQuest V7.3.0 (Bio-Rad Laboratories) software according to the manufacturer’s protocol. The protein spots with intensity levels greater than 2.0 or less than 0.5 were picked and subjected to mass spectrometric (MS) analysis. For MALDI-TOF MS analysis, the protein spots with significant MS signals were excised from the gels for tryptic digestion as described[24]. The peptides were analyzed on a 4700 Proteomics Analyzer (Applied Biosystems, USA). Proteins were identified using GPS Explorer software V3.5 (Applied Biosystems) and the function was quoted from the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov) database.

Measurement of intercellular organic acid content. E. coli DH10B and its recombinant strain EN-01 were cultured under anaerobic nitrogen-fixation conditions, and cells were collected at 0 h and 45 h. A1501 was grown under microaerobic nitrogen-fixation conditions, and the bacteria were centrifuged at 0 h and 10 h. The cell suspension was centrifuged at 12000 g, 4°C for 10 min, and the supernatant was used to analyze the organic acid content. Organic acid concentrations were determined using an HPLC (1200 series, Agilent Technologies) equipped with a Bio-Rad HPX-87H column with 5 mM H2SO4 as the mobile phase (0.6 ml/min, 50°C).

References
1. Ochman, H., Lawrence, J. G. & Groisman, E. A. Lateral gene transfer and the nature of bacterial innovation. Nature 405, 299–304, https://doi.org/10.1038/35012500 (2000).
2. Piremee, R., Cronje, L., Beuzidt, O. & Reva, O. N. Pre_GL: a global map of ontological links between horizontally transferred genomic islands in bacterial and archaeal genomes. Database 58, https://doi.org/10.1093/database/bav058 (2015).
3. Melnyk, R. A. & Coates, J. D. The Perchlorate Reduction Genomic Island: Mechanisms and Pathways of Evolution by Horizontal Gene Transfer. BMC Genomics 16, 862, https://doi.org/10.1186/s12864-015-1211-5 (2015).
4. Dixon, R. & Postgate, J. Transfer of nitrogen-fixation genes by conjugation in Klebsiella pneumoniae. Nature 234, 47–48, https://doi.org/10.1038/234047a0 (1971).
5. Dixon, R. & Postgate, J. Genetic transfer of nitrogen fixation from Klebsiella pneumoniae to Escherichia coli. Nature 237, 102–103, https://doi.org/10.1038/237102a0 (1972).
6. Streicher, S., Gurney, E. & Valentine, R. C. Transduction of the nitrogen fixation genes in Klebsiella pneumoniae. Proc Natl Acad Sci USA 68, 1174–1177 (1971).
7. Postgate, J. & Kent, H. M. Qualitative Evidence for Expression of Nitrogenase in nifQ on the Nitrogen Fixation Genes of Pseudomonas putida. Journal of General Microbiology 133, 2563–2566, https://doi.org/10.1099/00221287-133-3-2563 (1987).
8. Temme, K., Zhao, D. & Voigt, C. A. Refactoring the nitrogen fixation gene cluster from Klebsiella oxytoca. Proc Natl Acad Sci USA 109, 7085–7090, https://doi.org/10.1073/pnas.1107881109 (2012).
9. Wang, X. et al. Using Synthetic Biology to Distinguish and Overcome Regulatory and Functional Barriers Related to Nitrogen Fixation. Plos One 8, e68677, https://doi.org/10.1371/journal.pone.0068677 (2013).
10. Setten, L. et al. Engineering Pseudomonas protegens Pf 5 for Nitrogen Fixation and its Application to Improve Plant Growth under Nitrogen-Deficient Conditions. Plos One 8, e36666, https://doi.org/10.1371/journal.pone.0036666 (2013).
11. Wang, L. et al. A Minimal Nitrogen Fixation Gene Cluster from Paenibacillus sp. WLY78 Enables Expression of Active Nitrogenase in Escherichia coli. Plos Genetics 9, e1003865, https://doi.org/10.1371/journal.pgen.1003865 (2013).
12. Han, Y. L. et al. Interspecies Transfer and Regulation of Pseudomonas stutzeri A1501 Nitrogen Fixation Island in Escherichia coli. J. Microbiol. Biotechnol 25, 1339–1348, https://doi.org/10.4014/jm.1502.02027 (2015).
13. Cheng, Q., Day, A., Dowson-Day, M., Shen, G. & Dixon, R. The Klebsiella pneumoniae nitrogenase Fe protein gene (nfiH) functionally substitute for the chlH gene in Chlamydomonas reinhardtii. Biochemical and Biophysical Research Communication 329, 966–975, https://doi.org/10.1016/j.bbrc.2005.02.064 (2005).
14. Cheng, Q. Perspectives in Biological Nitrogen Fixation Research. Journal of Integrative Plant Biology 50, 784–796, https://doi.org/10.1111/j.1744-7909.2008.00710.x (2008).
15. Geddes, B. A. et al. Use of plant colonizing bacteria as chassis for transfer of N2-fixation to cereals. Current opinion in biotechnology 32, 216–222, https://doi.org/10.1016/j.copbio.2015.01.004 (2015).
16. Vicente, E. J. & Dean, D. R. Keeping the nitrogen-fixation dream alive. Proc Natl Acad Sci USA 114, 3009–3011, https://doi.org/10.1073/pnas.1701560114 (2017).
17. Mus, F. et al. Symbiotic nitrogen fixation and the challenges to its extension to nonlegumes. Applied and environmental microbiology 82, 3698–3710, https://doi.org/10.1128/AEM.01055-16 (2016).
18. Yang, J., Xie, X., Wang, X., Dixon, R. & Wang, Y. Reconstruction and minimal gene requirements for the alternative iron-only nitrogenase in Escherichia coli. Proc Natl Acad Sci USA 111, E3718–E3725, https://doi.org/10.1073/pnas.1411185114 (2014).
19. Li, X., Liu, Q., Liu, X., Shi, H. & Chen, S. Using synthetic biology to increase nitrogenase activity. Microbiol Cell Factors 15, 43, 1–7, https://doi.org/10.1186/s12934-016-0412-6 (2016).
20. Lin, M. & You, C. Denitrification and nitrogen fixation by Alcaligenes faecalis. Acta Agriculturae Sinica 1, 1–7 (1987).
21. Rediers, H. et al. Development and application of a dapB-based in vivo expression technology system to study colonization of rice by the endophytic nitrogen-fixing bacterium Pseudomonas stutzeri A15. Appl Environ Microbiol 69, 64–6874, https://doi.org/10.1128/AEM.69.1.64-6874 (2003).
22. Xie, Z. H. et al. Interaction between NifL and NifA in the nitrogen-fixing Pseudomonas stutzeri A1501. Microbiology 152, 3535–3542, https://doi.org/10.1099/micro.0.29171-0 (2006).
23. He, S. et al. Involvement of GlnK, a P4 protein, in control of nitrogen fixation and ammonia assimilation in Pseudomonas stutzeri A1501. Arch Microbiol 190, 1–10, https://doi.org/10.1007/s00203-008-0354-x (2008).
24. Yan, Y. et al. Global transcriptional analysis of nitrogen fixation and ammonium repress in root-associated Pseudomonas stutzeri A1501. BMC Genomics 11, 11, https://doi.org/10.1186/1471-2164-11-11 (2011).
25. Yan, Y. et al. Nitrogen fixation island and rhizosphere competence traits in the genome of root-associated Pseudomonas stutzeri A1501. Proc Natl Acad Sci USA 105, 7564–7569, https://doi.org/10.1073/pnas.0801093105 (2008).
26. Zhang, T. et al. Involvement of the ammonium transporter AmtB in nitrogenase regulation and ammonium excretion in Pseudomonas stutzeri A1501. Research in Microbiology 163, 332–339, https://doi.org/10.1016/j.resmic.2012.05.002 (2012).
27. Zhan, Y. et al. The novel regulatory ncRNA, Nf2S, optimizes nitrogen fixation via base pairing with the nitrogenase nifH mRNA in Pseudomonas stutzeri A1501. Proc Natl Acad Sci USA 113, E4348–E4356, https://doi.org/10.1073/pnas.1604514113 (2016).
28. Dixon, R. & Kahn, D. Genetic regulation of biological nitrogen fixation. Nat. Rev. Microbiol 2, 621–631, https://doi.org/10.1038/nrmicro954 (2004).
29. Imperiali, J., Ugalde, R. A., Shah, V. K. & Brill, W. J. Role of the nifQ Gene Product in the Incorporation of Molybdenum into Nitrogenase in Klebsiella pneumonia. Journal of Bacteriology 158, 187–194 (1984).
30. Roberts, G. P., MacNeil, T., MacNeil, D. & Brill, W. J. Regulation and Characterization of Protein Products Coded by the nif (Nitrogen Fixation) Genes of Klebsiella pneumoniae. Journal of Bacteriology 136, 267–279 (1978).
31. Lin, M., Yan, Y., Lu, W., Zhan, Y. & Zhang, Y. Regulatory Coupling of the Nitrogen and Carbon Metabolism in Nitrogen-Fixing *Pseudomonas stutzeri* A1501. *Biological Nitrogen Fixation* (ed. De Brujin Frans,) I, 109–119 (Wiley-Blackwell, Hoboken, NJ, 2015).

32. Van Heeswijk, W. C. et al. An alternative P*β* protein in the regulation of glutamine synthetase in *Escherichia coli*. *Molecular Microbiology* 21, 133–146, https://doi.org/10.1111/j.1365-2958.1996.tb01349.x (1996).

33. Muse, W. B. & Bender, R. A. The nuc (Nucleotide Assimilation Control) Gene from *Escherichia coli*. *Journal of Bacteriology* 180, 1166–1173 (1998).

34. Ke, Q., Chen, Q., Zhan, Y., Lu, W. & Yan, Y. Functional Complement Between General Nitrogen Regulator GlnG in *Escherichia coli* and General Nitrogen Regulator NtrC in *Pseudomonas*. *Biotechnology Bulletin* 32, 190–197, https://doi.org/10.13560/j.ckbiotech.bull.1985.2016.04.025 (2016).

35. Arcondeguy, T., Jack, R. & Merrick, M. P*β*, Signal Transduction Proteins, Pivotal Players in Microbial Nitrogen Control. *Microbiology and Molecular Biology Reviews* 65, 80–105, https://doi.org/10.1128/MMBR.65.1.80-105.2001 (2001).

36. Akinison, M. R., Blauskamp, T. A. & Ninfa, A. J. Context-Dependent Functions of the PII and GlnK Signal Transduction Proteins in *Escherichia coli*. *Journal of Bacteriology* 184, 5364–5375, https://doi.org/10.1128/JB.184.15.5364-5375.2002 (2002).

37. Reitzer, L. & Schneider, B. L. Metabolic Context and Possible Physiological Themes of δ*δ*1-Dependent Genes in *Escherichia coli*. *Microbiology and Molecular Biology Reviews* 65, 422–444 (2001).

38. Reitzer, L. Nitrogen Assimilation and Global Regulation in *Escherichia coli*. *Annual Review of Microbiology* 57, 155–176, https://doi.org/10.1146/annurev.micro.57.030506.190820 (2003).

39. Zimmer, D. P. et al. Nitrogen regulatory protein C-controlled genes of *Escherichia coli*: Scavenging as a defense against nitrogen limitation. *Proc Natl Acad Sci USA* 97, 14674–14679, https://doi.org/10.1073/pnas.97.26.14674 (2000).

40. Yang, J., Yan, D., Dixon, R. & Wang, Y. Deciphering the Principles of Bacterial Nitrogen Dietary Preferences: A Strategy for Nutrient Containment. *mBio* 7, e00792–16, https://doi.org/10.1128/mBio.00792-16 (2016).

41. Muse, W. B. & Bender, R. A. The nuc (nitrogen assimilation control) gene from *Escherichia coli*. *Journal of Bacteriology* 180, 1166–1173 (1998).

42. Blauskamp, T. A. & Ninfa, A. J. Physiological role of the GlnK signal transduction protein of *Escherichia coli*: survival of nitrogen starvation. *Molecular Microbiology* 46, 203–214, https://doi.org/10.1111/j.1365-2958.2002.03153.x (2002).

43. Goss, T. J., Perez-Matos, A. & Bender, R. A. Roles of glutamate synthase, gbdB, and gldB in nitrogen metabolism of *Escherichia coli* and *Klebsiella aerogenes*. *Journal of Bacteriology* 183, 6607–6619, https://doi.org/10.1128/JB.183.22.6607-6619.2001 (2001).

44. Poggio, S., Domeinzain, C., Osorio, A. & Camarena, L. The nitrogen assimilation control (Nac) protein represses asetC and asetA transcription in *Escherichia coli*. *FEBS Letters* 206, 151–156, https://doi.org/10.10111/j.1574-9689.2002.tb11001.x (2002).

45. Canfield, D. E., Glazer, A. N. & Falkowski, P. G. The Evolution and Future of Earth’s Nitrogen Cycle. *Science* 330(192–196), 330: 192–196 (2010).

46. Borisova, V. B., Gennis, R. B., Hemp, J. & Verkhovsky, M. I. The cytochrome bd respiratory oxygen reductases. *Biochimica et Biophysica Acta* 1807, 1398–1413, https://doi.org/10.1016/j.bbabo.2011.06.016 (2011).

47. Wissenbawch, U., Kröger, A. & Uniden, G. The specific functions of menaquinone and dimenaquinone in anaerobic respiration with fumarate, dimethylsulfoxide, trimethylamic N-oxide and nitrate by *Escherichia coli*. *FEMS Microbiol. Lett.* 192–196 (2010).

48. Gallon, J. R. Reconciling the incompatible: N2 fixation and O2. *Science* 289, 1775–1782, https://doi.org/10.1126/science.10022187-147.7-1775 (2001).

49. Abdallah, J., Caldas, T., Khirli, E., Kern, R. & Richarme, G. YhbO Protects Cells against Multiple Stresses. *American Society for Microbiology* 189, 9140–9144, https://doi.org/10.1128/JB.01208-07 (2007).

50. Buda, F. B., Ensing, B. E., Gribnau, M. C. M. & Baerends, E. J. O2 evolution in the Fenton reaction. *Chemistry* 9, 3436–3444, https://doi.org/10.1002/chem.200204444 (2003).

51. Nair, S. & Finkel, S. E. Dps Protects Cells against Multiple Stresses during Stationary Phase. *Journal of Bacteriology* 186, 4192–4198 (2004).

52. Claude, E. Associative and endophytic nitrogen-fixing bacteria and cyanobacterial associations (ed. Claude, E. & Newton, W. E.) 1–16, Vol. 5. *Dordrecht: Springer* (2007).

53. Raymond, J., Siefer, J. L., Staples, C. R. & Blankenship, R. E. The Natural History of Nitrogen Fixation. *Molecular Biology & Evolution* 21, 541–554, https://doi.org/10.1093/molbev/msi87 (2004).

54. Venieraki, A. et al. The nitrogen-fixation island insertion site is conserved in diatrophic *Pseudomonas stutzeri* and *Pseudomonas* sp. isolated from distal and close geographical regions. *Plos One* 9, e105837, https://doi.org/10.1371/journal.pone.0105837 (2014).

55. Xiong, J., Fischer, W. M., Inoue, K., Nakahara, M. & Bauer, C. E. Molecular evidence for the early evolution of photosynthesis. *Science* 289, 1724–1730, https://doi.org/10.1126/science.289.5485.1724 (2000).

56. Yang, J., Xie, X., Yang, M., Dixon, R. & Wang, Y. Modular electron-transport chains from eukaryotic organelles function to support nitrogenase activity. *Proceedings of the National Academy of Sciences of the United States of America* 114, 2460–2465, https://doi.org/10.1073/pnas.1300581114 (2017).

57. Brennan, C. F., Dixon, R., Postgate, J. R. & Primrose, S. B. Chromosomal integration of *Klebsiella* niitrogen fixation genes in *Escherichia coli*. *Journal of General Microbiology* 80, 220–239, https://doi.org/10.1099/0221287-80-1-227 (1974).

58. Desmons, N. et al. Nitrogen fixation genetics and regulation in a *Pseudomonas stutzeri* strain associated with rice. *Microbiology* 149, 2251–2262, https://doi.org/10.1099/mic.0.26270-0 (2003).

59. Zhou, Z. F. et al. Genome-wide transcriptome and proteome analysis of *Escherichia coli* expressing IrE, a global regulator of *Deinococcus radiodurans*. *Molecular Systems* 7, 1613–1620, https://doi.org/10.1039/c0mb00336e (2011).

**Acknowledgements**

This work was supported by National Natural Science Foundation of China (31230004, 31470205, 31470174 and 31770067), the National Basic Research Program of China (2015CB755700), and Ministry of Agriculture (Transgenic Program, No. 2016ZX08009003-002). This study was also conducted with the support of the Agricultural Science and Technology Innovation Program (2014–2017), Fundamental Research Funds for Central Non-profit Scientific Institution (039201702) and the Guangdong Innovative and Entrepreneurial Research Team Program (No. 2013S033). We would like to thank Dr. Claudine Elmerich and Dr. Ray Dixon for many helpful discussions.
Author Contributions
Y.Y. and M.L. conceived and designed the experiments. Z.Y., Y.H., Y.M., Q.C. and Y.Z. performed the experiments. Z.Y., Y.Y. and M.L. contributed to the writing of the manuscript. Z.Y., W.L., L.C., M.H., Y.Y. and S.C. contributed to the collection and analysis of data. All authors (Z.Y., Y.H., Y.M., Q.C., Y.Z., W.L., L.C., M.H., S.C., Y.Y. & M.L.) discussed the results, commented and approved the version of final manuscript.

Additional Information
Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-018-29204-0.

Competing Interests: The authors declare no competing interests.

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2018