Construction of genetically-engineered *Escherichia coli* for sustainable ammonia production

Yuki Tatemichi

2022

Contents

Abbreviations

BNF : **biological nitrogen fixation**

cAMP : **cyclic adenosine monophosphate**

cGMP : **cyclic guanosine monophosphate**

CRP : **cyclic AMP receptor protein**

E. coli : ***Escherichia coli***

GC-MS : **gas chromatography-mass spectrometry**

HPLC : **high-performance liquid chromatography**

IPTG : **isopropyl β-D-1-thiogalactopyranoside**

LC-MS/MS : **liquid chromatography-tandem mass spectrometry**

MRM : **multiple reaction monitoring**

Nac : **nitrogen assimilation control**

NAD : **nicotinamide adenine dinucleotide**

PBS : **phosphate-buffered saline**

PCR : **polymerase chain reaction**

PLS : **partial least squares**

PTS : **phosphotransferase system**

RT-qPCR : **reverse transcription-quantitative polymerase chain reaction**

VIP : **variable importance in the projection**

Introduction

The Haber-Bosch process is the most common process of ammonia production. However, it consumes large amounts of oil and natural gas for the steam reforming process, progresses nitrogen fixation, and emits carbon dioxide. Depending on the consumption of fossil fuels, the current ammonia production process raises concerns on sustainability and carbon neutrality. In recent years, ammonia production by microorganisms has attracted attention as an alternative process. For example, microbial catabolism of amino acids can efficiently produce ammonia from unused biomass. In addition, the development of a biological nitrogen fixation process via heterologous expression of genes derived from nitrogen-fixing bacteria has been reported [1]. While these biological ammonia production processes are expected to reduce the use of fossil fuels, their production efficiency is still insufficient. In producing ammonia from unused biomass via the catabolism of amino acids, a wide range of issues should be addressed, including the selection of biomass feedstock, pretreatment methods, optimum conditions for efficient amino acid catabolism, and extraction and purification of ammonia from the reaction solution. In the biological nitrogen fixation process, several issues are observed, such as the oxygen-mediated irreversible inactivation of nitrogenase, which proceeds nitrogen fixation, the low efficiency of nitrogen fixation, and low yield due to ammonia assimilation.

Importance of ammonia and concerns regarding fossil fuels consumption

Ammonia is a colorless gas at normal temperature and pressure and is mainly used as a raw material in fertilizers and chemical products. As the global population increases, the importance of ammonia as a fertilizer for agricultural products does not change due to food security. Ammonia is also a promising carrier of hydrogen, which is a next-generation energy source and is an alternative to fossil fuels. Particularly, ammonia is attracting attention as a “carbon-free” substance, since it does not emit CO_2 when burned; moreover, mixing ammonia with coal-fired power generation (co-firing) can reduce CO_2 emissions [2-4].

Ammonia is synthesized very efficiently via the Haber-Bosch process; however, the energy and environmental impacts of this process are often a concern. This process requires a high pressure (20–40 MPa) and temperature (400–600 °C), causing a heavy consumption of energy. Substantial energy consumption in Haber-Bosch process, coupled with the high consumption of ammonia due to economic development and population growth, accounts for 1% of the global energy consumption [5]. Efficient catalysts and

improved ammonia synthesis processes have been developed to overcome the high energy consumption of the Haber-Bosch process; ammonia production at low temperatures and pressures are now partially possible [6-10]. However, issues are still present, such as the low efficiency of ammonia production, difficulty of large-scale production, and difficulty in acquiring and high cost of rare metals needed as new catalysts.

Ammonia production from food by-products

Biomass, an organic resource derived from plants and animals, is currently attracting attention as a renewable energy source. Food by-products are rich in protein and fiber, and some are used in health foods due to their properties. However, most biomass are feed or waste products [11]. The efficient use of protein-rich food by-products as raw materials for ammonia production would enable the effective utilization of unused resources and reduce their environmental burden of the Haber-Bosch process. Okara is the residue produced during soymilk and tofu production; 1.1 kg of wet okara is produced from 1 kg of soybeans [12]. Despite having high protein concentrations (27–38%), most of the okara produced is either used as animal feed or discarded [13]. These proteins found in the water extract of okara are insoluble and difficult to handle, but they can be efficiently extracted via hydrolysis of the cell wall using cellulase [14]. Specifically, the biotechnological processes involved in decomposing amino acids into ammonia do not require hydrogen generation and nitrogen triple bond dissociation, although both are essential in the Haber-Bosch process. Therefore, the biotechnological production of ammonia from food by-products is expected to be an alternative to the Haber-Bosch process in terms of energy consumption.

Biomass utilization has environmental advantages; however, several issues in industrial applications should be addressed. Several studies have constructed genetically modified *Escherichia coli* or *Bacillus subtilis* with activated amino acid metabolism, specifically the expression of aminotransferase, to produce ammonia and biofuel from amino acids in biomass; however, the production efficiency is still insufficient [15, 16]. The low yield could be due to the impurities in the biomass. However, since inhibitors associated with impurities vary greatly depending on the types of biomass and microorganisms, it is necessary to identify inhibitors in each production process, which requires considerable effort. Therefore, efficient biotechnological production of ammonia from food by-products requires identification of impurities and establishment of inhibitor-resistant strains. An efficient process is yet to be demonstrated.

Identification of key components using metabolic profiling and partial least squares (PLS) regression analysis

In recent years, metabolomics has been developed as a tool for understanding life phenomena and are used to study microbes [17], plants [18], and mammals [19]. In metabolomics, primary and secondary metabolites are comprehensively quantified using liquid chromatography (LC)-mass spectrometry (MS), capillary electrophoresis (CE)-MS, and gas chromatography (GC)-MS analysis with derivatization. Metabolic profiling, which was developed based on metabolomics technology and expanded more compounds other than metabolites, can identify important components at the same low cost and high throughput as metabolomics [20] and facilitate enhanced comprehension of the relationship between product quality and the chemical components in the product [21, 22]. Using metabolic profiling, it is possible to predict the production volume from raw materials and search for components that have significant impact on production.

In recent years, PLS regression analysis is a statistical analysis method often used to screen for key components from metabolic profiling data in recent years. PLS regression combines the features of principal component analysis (PCA) and multiple regression. PCA is a multivariate analysis method used to reduce the dimensionality of data by synthesizing a small number of uncorrelated variables from a large number of correlated variables, called principal components, which best represent the overall variability. PLS enables the analysis or prediction of a set of dependent variables from a set of independent or predictor variables. This prediction is accomplished by extracting from the predictors a set of orthogonal factors, called latent variables, that have the highest predictive probability. Variable importance in projection (VIP) values obtained using PLS regression analysis can be used to predict important factors in different fields, including epidemiology [23, 24], food science [22, 25], and biotechnology [26, 27].The rapid and comprehensive acquisition of component data using metabolic profiling and the selection of variables using PLS regression analysis and VIP can provide more efficient and faster screening of important factors than conventional methods.

Biological nitrogen fixation by diazotrophic bacteria

Biological nitrogen fixation is usually performed at moderate temperatures by the activity of molybdenum (Mo) nitrogenase in diazotrophic bacteria, which are widely distributed in nature [28, 29]. The Mo nitrogenase enzyme is composed of two interacting metalloproteins (MoFe and Fe proteins) and three cofactors (FeMo-co, P-cluster, and [4Fe-4S] cluster) that catalyze the reduction of N_2 to NH_3 using energy from intra- and intermolecular electron transfer and ATP hydrolysis [30]. In addition, some diazotrophic

microorganisms harbor alternative vanadium and/or iron-only nitrogenases [31].

Klebsiella oxytoca, the most studied diazotrophic bacterium, fixes nitrogen under anaerobic conditions using Mo nitrogenase. Twenty nitrogen fixation (nif)-related genes consisting of multiple transcription units, *nifJHDKTYENXUSVWZMFLABQ*, are clustered in a 23-kb region on the chromosome of *K. oxytoca* (Fig. 1) [32]. This *nif* gene cluster of *K. oxytoca* can be cloned and has been frequently used as a gene source for heterologous expression of nitrogenase. In contrast to other anaerobic nitrogen-fixing diazotrophic bacteria, *Azotobacter vinelandii* can fix nitrogen under aerobic conditions. The *nif* genes of *A. vinelandii* exist as major and minor clusters in the chromosomes, namely *nifHDKTYENXUSVWZMF* and *nifLABOQ*, respectively. Other important *nif*-related genes were also included in each cluster (Fig. 1).

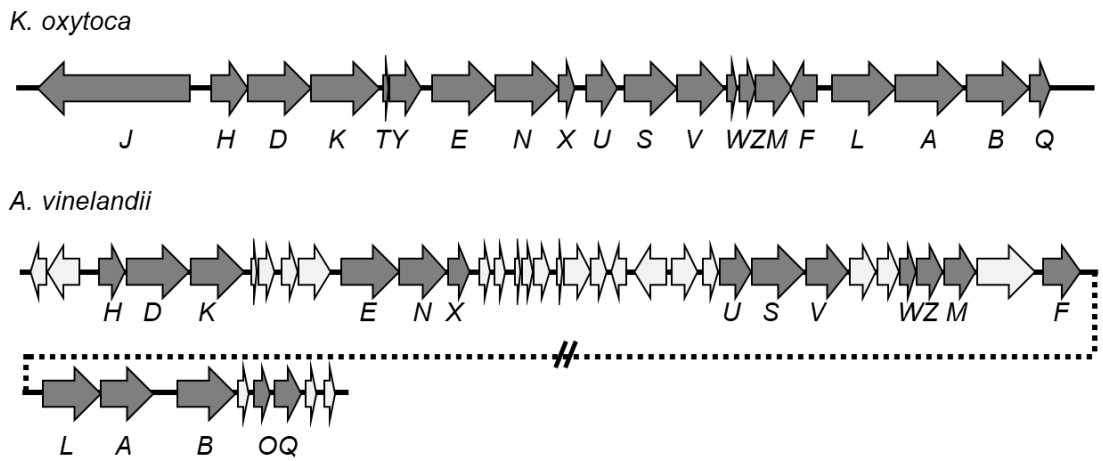

Fig. 1 The *nif* cluster of *K. oxytoca* and *A. vinelandii*. In *K. oxytoca*, the *nif* genes (grey) are clustered together (top). In *A. vinelandii*, the *nif* genes are separated into major (middle) and minor clusters (bottom), and other *nif*-related or genes with unknown function (white) are present between the *nif* genes.

The *nifHDK* gene encodes a structural member of the nitrogenase enzyme complex and is essential for nitrogen fixation (Fig. 2). The biosynthesis of cofactors and maturation of NifDK require complex biosynthetic pathways, and the products of *nif* genes are involved in the biosynthesis of Mo nitrogenase [33]. NifEN and NifB are involved in the FeMo-co synthesis from iron, sulfur, molybdenum, and homocitrate [34]. The activator and inactivator proteins, NifA and NifL, respectively, form a pair and regulate the expression of *nif* genes in response to environmental signals, such as the rate of nitrogen assimilation, presence of oxygen, redox state of the cell, and ATP levels [35]. Typically, under nitrogen starvation conditions, NifA deviates from NifL and promotes

transcription of nitrogenase genes such as *nifH*. Flavodoxin (NifF), which is required to supply electrons to nitrogenase, is not essential for the nitrogen-fixing activity of *A. vinelandii* but is known to be highly transcribed during nitrogen fixation [36]. NifU and NifS, homologs of IscU and IscS proteins, respectively, are commonly involved in the formation of [Fe-S] clusters in various organisms and in the synthesis of [Fe-S] clusters, which are cofactors of nitrogenase [37].

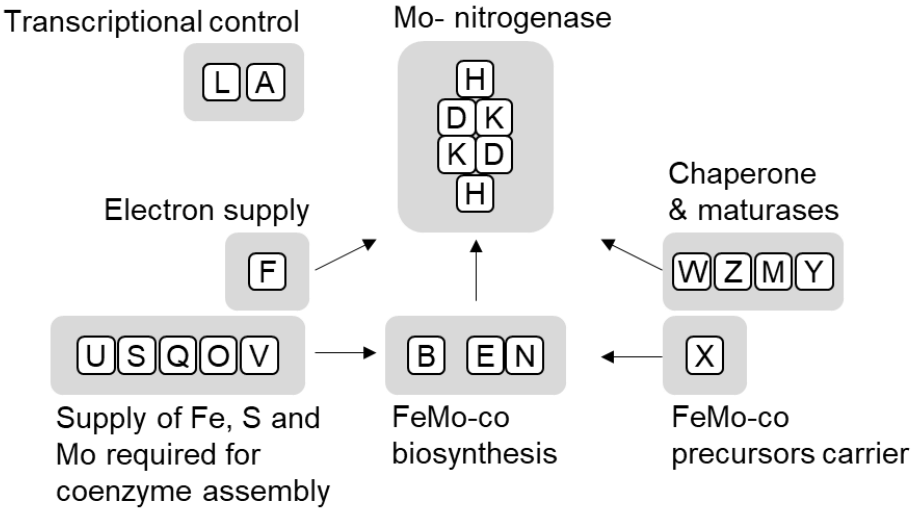

Fig. 2 The process of Mo-nitrogenase biosynthesis in *A. vinelandii*. The functions of the *nif* genes are classified into *nifHDK*, which constitutes Mo-nitrogenase, and *nifLAFUSQOVBENWZMYX,* which are involved in the maturation of Mo-nitrogenase and coenzyme synthesis. NifQ, NifO, and NifV contribute to the biosynthesis of the FeMo cofactor and its insertion into the MoFe protein. In addition, NifZ plays an important role in the synthesis of metalloclusters, and NifM and NifW are required for the proper folding of the Fe protein in nitrogenase [32]. Both NifY and NifX have chaperone-like functions, with NifY and NifX binding to Mo nitrogenase and the NifEN complex, respectively [38].

Heterologous expression of *nif*-related genes

To further understand the biological nitrogen fixation reactions by nitrogenases and explore their industrial applications, heterologous expression of nitrogenase and related proteins has been studied using *K. oxytoca* or *Paenibacillus* sp. [1, 39-43]. Although these studies have revealed the chemical characteristics of nitrogenase and the factors essential for its maturation, nitrogenase activity and productivity still need to be improved for industrial applications.

Unlike *K. oxytoca*, *A. vinelandii* can fix nitrogen even in the presence of oxygen

owing to its unique tolerance to oxygen, which has not been fully elucidated [44]. By understanding the mechanism and introducing genes that improve oxygen tolerance into microorganisms with heterologous expression of *nif*-related genes, it may be possible to establish microorganisms that can fix nitrogen in the presence of oxygen. However, unlike *K. oxytoca* that has a continuous *nif*-related gene cluster, the *nif* and related genes of *A. vinelandii* are scattered in large and small clusters on the genomic DNA, making them difficult to clone and express.

To establish a sustainable ammonia production process, two approaches were explored. The first approach was to improve the efficiency of ammonia production from food by-products. Inhibitors of ammonia production by *E. coli* were identified in food by-products, and the ammonia production efficiency of metabolically engineered *E. coli* was improved. Next, we established recombinant *E. coli* expressing *nif* genes derived from *A. vinelandii*. Nitrogenase activity was detected in the recombinant *E. coli* and was further enhanced by selecting and additionally expressing useful genes from *A. vinelandii*.

References

1. Dixon RA, Postgate JR. Genetic transfer of nitrogen fixation from *Klebsiella pneumoniae* to *Escherichia coli*. Nature 1972; 237:102-103.
2. Lan R, Irvine JTS, Tao S. Ammonia and related chemicals as potential indirect hydrogen storage materials. Int J Hydr Energy 2012; 37:1482-1494.
3. Miura D, Tezuka T. A comparative study of ammonia energy systems as a future energy carrier, with particular reference to vehicle use in Japan. Energy 2014; 68:428-436.
4. Elishav O, Mosevitzky Lis B, Miller EM, Arent DJ, Valera-Medina A, Grinberg Dana A, Shter GE, Grader GS. Progress and prospective of nitrogen-based alternative fuels. Chem Rev 2020; 120:5352-5436.
5. Schrock RR. Reduction of dinitrogen. Proc Natl Acad Sci U S A 2006; 103:17087.
6. Yandulov DV, Schrock RR. Catalytic reduction of dinitrogen to ammonia at a single molybdenum center. Science 2003; 301:76-78.
7. Arashiba K, Miyake Y, Nishibayashi Y. A molybdenum complex bearing PNP-type pincer ligands leads to the catalytic reduction of dinitrogen into ammonia. Nat Chem 2011; 3:120-125.
8. Anderson JS, Rittle J, Peters JC. Catalytic conversion of nitrogen to ammonia by an iron model complex. Nature 2013; 501:84-87.

9. Fajardo J, Jr., Peters JC. Catalytic nitrogen-to-ammonia conversion by osmium and ruthenium complexes. J Am Chem Soc 2017; 139:16105-16108.
10. Ashida Y, Arashiba K, Nakajima K, Nishibayashi Y. Molybdenum-catalysed ammonia production with samarium diiodide and alcohols or water. Nature 2019; 568:536-540.
11. Kibler KM, Reinhart D, Hawkins C, Motlagh AM, Wright J. Food waste and the food-energy-water nexus: a review of food waste management alternatives. Waste Manag 2018; 74:52-62.
12. Khare SK, Jha K, Gandhi AP. Citric acid production from okara (soy-residue) by solid-state fermentation. Bioresour Technol 1995; 54:323-325.
13. Surel O, Couplet B. Influence of the dehydration process on active compounds of okara during its fractionation. J Sci Food Agric 2005; 85:1343-1349.
14. de Figueiredo VRG, Yamashita F, Vanzela ALL, Ida EI, Kurozawa LE. Action of multi-enzyme complex on protein extraction to obtain a protein concentrate from okara. J Food Sci Technol 2018; 55:1508-1517.
15. Choi KY, Wernick DG, Tat CA, Liao JC. Consolidated conversion of protein waste into biofuels and ammonia using *Bacillus subtilis*. Metab Eng 2014; 23:53-61.
16. Huo YX, Cho KM, Rivera JG, Monte E, Shen CR, Yan Y, Liao JC. Conversion of proteins into biofuels by engineering nitrogen flux. Nat Biotechnol 2011; 29:346-351.
17. Allen J, Davey HM, Broadhurst D, Heald JK, Rowland JJ, Oliver SG, Kell DB. High-throughput classification of yeast mutants for functional genomics using metabolic footprinting. Nat Biotechnol 2003; 21:692-696.
18. Fiehn O, Kopka J, Dörmann P, Altmann T, Trethewey RN, Willmitzer L. Metabolite profiling for plant functional genomics. Nat Biotechnol 2000; 18:1157-1161.
19. Atherton HJ, Bailey NJ, Zhang W, Taylor J, Major H, Shockcor J, Clarke K, Griffin JL. A combined 1H-NMR spectroscopy- and mass spectrometry-based metabolomic study of the PPAR-alpha null mutant mouse defines profound systemic changes in metabolism linked to the metabolic syndrome. Physiol Genomics 2006; 27:178-186.
20. Dunn WB, Broadhurst D, Begley P, Zelena E, Francis-McIntyre S, Anderson N, Brown M, Knowles JD, Halsall A, Haselden JN *et al*. Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry. Nat Protoc 2011; 6:1060-1083.

21. Karpe AV, Beale DJ, Godhani NB, Morrison PD, Harding IH, Palombo EA. Untargeted metabolic profiling of winery-derived biomass waste degradation by *Penicillium chrysogenum*. J Agric Food Chem 2015; 63:10696-10704.
22. Shiga K, Yamamoto S, Nakajima A, Kodama Y, Imamura M, Sato T, Uchida R, Obata A, Bamba T, Fukusaki E. Metabolic profiling approach to explore compounds related to the umami intensity of soy sauce. J Agric Food Chem 2014; 62:7317-7322.
23. Jonasson A, Eriksson C, Jenkinson HF, Källestål C, Johansson I, Strömberg N. Innate immunity glycoprotein gp-340 variants may modulate human susceptibility to dental caries. BMC Infect Dis 2007; 7:57.
24. Yan Q, Liew Z, Uppal K, Cui X, Ling C, Heck JE, von Ehrenstein OS, Wu J, Walker DI, Jones DP *et al*. Maternal serum metabolome and traffic-related air pollution exposure in pregnancy. Environ Int 2019; 130:104872.
25. Harsha PSCS, Wahab RA, Cuparencu C, Dragsted LO, Brennan L. A metabolomics approach to the identification of urinary biomarkers of pea intake. Nutrients 2018; 10.
26. Chen Z, Zheng Z, Wang FL, Niu YP, Miao JL, Li H. Intracellular metabolic changes of *Rhodococcus* sp. LH during the biodegradation of diesel oil. Mar Biotechnol (NY) 2018; 20:803-812.
27. Zalai D, Koczka K, Párta L, Wechselberger P, Klein T, Herwig C. Combining mechanistic and data-driven approaches to gain process knowledge on the control of the metabolic shift to lactate uptake in a fed-batch CHO process. Biotechnol Prog 2015; 31:1657-1668.
28. Raymond J, Siefert JL, Staples CR, Blankenship RE. The natural history of nitrogen fixation. Mol Biol Evol 2004; 21:541-554.
29. Bulen WA, LeComte JR. The nitrogenase system from *Azotobacter*: two-enzyme requirement for N_2 reduction, ATP-dependent H_2 evolution, and ATP hydrolysis. Proc Natl Acad Sci U S A 1966; 56:979-986.
30. Hoffman BM, Lukoyanov D, Yang ZY, Dean DR, Seefeldt LC. Mechanism of nitrogen fixation by nitrogenase: the next stage. Chem Rev 2014; 114:4041-4062.
31. Bothe H, Schmitz O, Yates MG, Newton WE. Nitrogen fixation and hydrogen metabolism in cyanobacteria. Microbiol Mol Biol Rev 2010; 74:529-551.
32. Rubio LM, Ludden PW. Biosynthesis of the iron–molybdenum cofactor of nitrogenase. Annu Rev Microbiol 2008; 62:93-111.
33. Rubio LM, Ludden PW. Maturation of nitrogenase: a biochemical puzzle. J Bacteriol 2005; 187:405-414.

34. Hu Y, Yoshizawa JM, Fay AW, Lee CC, Wiig JA, Ribbe MW. Catalytic activities of NifEN: implications for nitrogenase evolution and mechanism. Proc Natl Acad Sci U S A 2009; 106:16962-16966.
35. Martinez-Argudo I, Little R, Shearer N, Johnson P, Dixon R. The NifL-NifA system: a multidomain transcriptional regulatory complex that integrates environmental signals. J Bacteriol 2004; 186:601-610.
36. Mus F, Alleman AB, Pence N, Seefeldt LC, Peters JW. Exploring the alternatives of biological nitrogen fixation. Metallomics 2018; 10:523-538.
37. Zheng L, Cash VL, Flint DH, Dean DR. Assembly of iron–sulfur clusters. Identification of an *iscSUA-hscBA-fdx* gene cluster from *Azotobacter vinelandii*. J Biol Chem 1998; 273:13264-13272.
38. Fani R, Gallo R, Liò P. Molecular evolution of nitrogen fixation: the evolutionary history of the *nifD, nifK*, *nifE*, and *nifN* genes. J Mol Evol 2000; 51:1-11.
39. Temme K, Zhao D, Voigt CA. Refactoring the nitrogen fixation gene cluster from *Klebsiella oxytoca*. Proc Natl Acad Sci U S A 2012; 109:7085-7090.
40. Wang L, Zhang L, Liu Z, Zhao D, Liu X, Zhang B, Xie J, Hong Y, Li P, Chen S *et al*. A minimal nitrogen fixation gene cluster from *Paenibacillus* sp. WLY78 enables expression of active nitrogenase in *Escherichia coli*. PLOS Genet 2013; 9:e1003865.
41. Ivleva NB, Groat J, Staub JM, Stephens M. Expression of active subunit of nitrogenase via integration into plant organelle genome. PLOS ONE 2016; 11:e0160951.
42. Yang J, Xie X, Xiang N, Tian ZX, Dixon R, Wang YP. Polyprotein strategy for stoichiometric assembly of nitrogen fixation components for synthetic biology. Proc Natl Acad Sci U S A 2018; 115:E8509-E8517.
43. Liu X, Wang M, Song Y, Li Y, Liu P, Shi H, Li Y, Hao T, Zhang H, Jiang W *et al*. Combined assembly and targeted integration of multigene for nitrogenase biosynthetic pathway in *Klebsiella pneumoniae*. ACS Synth Biol 2019; 8:1766-1775.
44. Gallon JR. Reconciling the incompatible: N_2 fixation and O_2. New Phytol 2006; 122:571-609.

Chapter I

Efficient Ammonia Production from Food By-products by Engineered *Escherichia coli*

In this chapter, I demonstrated ammonia production from food by-products using metabolic profiling and metabolic engineering to prevent the action of inhibitors. *E. coli* was selected as the host strain for ammonia production due to its potential for efficient ammonia production [1] and ease of genetic modification. Since our experiments of ammonia production from various food by-products suggested the presence of ammonia production inhibitors, I attempted to utilize metabolic profiling to clarify the relationship between the components of food by-products and ammonia production. This approach revealed that glucose in food by-products negatively affected ammonia production. This finding prompted us to engineer *E. coli* to prevent the action of ammonia production inhibitors by disrupting *ptsG* and *glnA*, thereby achieving efficient ammonia production from food by-products.

Materials and methods

Strains, plasmids, and media

E. coli and plasmids used in this study are described in Table 1. Food by-products of soy sauce cake, *mirin* cake, and tomato peel were provided by Kikkoman (Noda, Japan), and okara was purchased from Nippon beans (Isesaki, Japan). The enzymes described below were purchased from Amano-enzyme (Nagoya, Japan). The enzyme mixture was prepared by mixing 20 mg/mL of peptidase 'ProteAX', peptidase 'Peptidase R', protease 'Protin SD-AY10', protease 'Protease M Amano SD', protease 'Protin SD-NY10', protease 'Thermoase PC10F', protease 'Protease A Amano SD', hemicellulase 'Hemicellulase Amano 90', cellulase 'Cellulase A Amano 3', cellulase 'Cellulase T Amano 4', mannanase 'Mannanase BGM Amano 10' and pectinase 'Pectinase G Amano' in 50 mM MES (pH 5.5). To provide an equal amount of the food by-products by dry weight to the pretreatment reaction, a total of 70 g were selected from the following: 3 g soy sauce cake, 3 g *mirin* cake, 9.2 g tomato peel, or 10 g okara (each wet weight) were suspended in MES (pH 5.5) and reacted with 30 mL enzyme mixture at 55 °C for 72 h. The mixture was incubated at 80 °C for 30 min and filtered using ADVANTEC2, AVDANTEC131 (Toyo Roshi Kaisha, Tokyo, Japan), and Millex-HV Syringe Filter Unit, 0.45 μm, PVDF, 33 mm (Merck Millipore, MA, USA). LB broth, YPD broth, M9-yeast

extract, M9-tryptone, M9-peptone, and M9-casamino acids were prepared per previous studies [1, 2]. Ampicillin (100 μg/ mL) (Meiji Seika Pharma, Tokyo, Japan) and kanamycin (25 μg/ mL) (Nacalai Tesque, Kyoto, Japan) were added as appropriate.

Table 1 *E.coli* strains and plasmids used in this study

Strains/plasmids	Description	Source
E. coli strains		
DH10B	F⁻ *mcrA* Δ(*mrr-hsdRMS-mcrBC*) φ80d *lacZ*ΔM15 Δ*lacX74 recA1 endA1 araD139* Δ(*ara leu*)7697 *galU galK λ⁻ rpsL nupG*	Thermo fisher
DH10B(pKD46)	DH10B harboring pKD46	[1]
Δ*glnA*	DH10B, Δ*glnA*::FRT-km^R-FRT	[1]
Δ*ptsG*	DH10B, Δ*ptsG*::FRT-km^R-FRT	This study
Δ*ptsG-kan*	DH10B, Δ*ptsG*::FRT	This study
Δ*ptsG*Δ*glnA*	DH10B, Δ*ptsG*::FRT Δ*glnA*::FRT-km^R-FRT	This study
Plasmids		
pKD46	Red recombinase expression plasmid	[3]
pKD13	Km^R template plasmid	[3]
pCP20	Temperature-sensitive replication and thermal induction of FLP synthesis	[36]

Construction of *E. coli* strains

Gene deletion mutants were constructed using the previously reported method [3]. A *ptsG*-Km-resistant fragment was amplified from pKD13 [3] through polymerase chain reaction (PCR) with primers ptsGF and ptsGR containing homologous sequences upstream and downstream of the *ptsG* coding region and flippase recognition target (FRT) sequence. *E. coli* DH10B (pKD46) was electroporated with the *ptsG*-Km resistant fragment and plated on LB agar containing ampicillin and kanamycin at 30 °C. After incubation at 37 °C to eliminate the temperature-sensitive plasmid pKD46, the resulting ampicillin-sensitive strain (*ptsG*: FRT-*kan*-FRT) was designated Δ*ptsG*. Δ*ptsG*-kan strain was constructed from Δ*ptsG* using flippase (FLP) helper plasmid pCP20 [3]. A *glnA*-Km resistant fragment was amplified from pKD13 using PCR with primers glnAF and glnAR containing homologous sequences upstream and downstream of the *glnA* coding region and FRT sequence (Table 2). Strains, Δ*glnA* and Δ*ptsG*Δ*glnA*, were constructed from DH10B and Δ*ptsG* respectively, using the *glnA*-Km resistant fragment. To confirm the desired genome insertion of the resistance fragments, the insertion region was amplified by PCR and sequenced using insertion-checking primers.

Table 2 Primers used in this study

Primer	Sequence (5′-3′)	Note
glnAF	GCCAGAGACAGGCGAAAAGTTTCCACGGCA-ACTAAAACACGTGTAGGCTGGAGCTGCTTC	Construction for *glnA*-Km cassette
glnAR	GTTACCACGACGACCATGACCAATCCAGGAG-AGTTAAAGTCTGTCAAACATGAGAATTAA	Construction for *glnA*-Km cassette
ptsGF	AACGTAAAAAAAGCACCCATACTCAGGAGC-ACTCTCAATTCTGTCAAACATGAGAATTAA	Construction for *ptsG*-Km cassette
ptsGR	CAGCCATCTGGCTGCCTTAGTCTCCCCAACG-TCTTACGGAGTGTAGGCTGGAGCTGCTTC	Construction for *ptsG*-Km cassette
glnAseqF	AGCTGACAAACTTCACGTTG	Check the genome insertion of *glnA*-Km cassette
glnAseqR	GCAACATTCACATCGTGGTG	Check the genome insertion of *glnA*-Km cassette
ptsGseqF	GTCAAACAAATTGGCACTG	Check the genome insertion of *ptsG*-Km cassette
ptsGseqR	CAATAGCAGCCAGTCCCTTC	Check the genome insertion of *ptsG*-Km cassette

Evaluation of ammonia production and glucose uptake

Ammonia production experiments were conducted as previously described [1]. *E. coli* strains were initially grown overnight in M9-yeast extract, then the cells were washed with sterile water and inoculated into 2 mL of pretreated food by-products or semisynthesized medium to a final OD_{600} of 0.5. After incubation at 37 °C for 26.5 h with shaking, the concentrations of ammonia in the supernatants were measured by F-kit ammonia (J.K. International, Tokyo, Japan). To evaluate the glucose uptake, each of the strains cultured overnight in M9-yeast extract was added to 50 mM glucose with OD_{600} of 0.5 and incubated for 2 h. After centrifugation at 5,000 g, 4 °C for five min, the supernatant was collected and diluted with water five times. The amount of residual glucose was measured using the glucose CII-test-Wako (Wako, Osaka, Japan) to evaluate glucose uptake.

Quantification of ammonia, amino acids, organic acids, and sugars in medium and culture supernatant

The medium and culture supernatants were filtered using a 0.45 μm DISMIC filter (Advantec Toyo, Tokyo, Japan). The amino acid analysis was conducted on an L-8900 amino acid analyzer (Hitachi, Tokyo, Japan). Organic acids were quantified with an ST-3 (Showa Denko, Tokyo, Japan) post-column reaction system [4, 5] using prominence high-performance liquid chromatography (HPLC) (Shimadzu, Kyoto, Japan) equipped with an RSpak KC-811 column (30 mm × 8.0 mm, 6 μm; Showa Denko, Tokyo, Japan).

Analysis of sugars was conducted using the phenylhydrazine post-column method [6] by employing a chromstar 5510 (Hitachi, Tokyo, Japan) equipped with an Asahipak NH2P-50 4E column (4.6 mm × 250 mm) (Showa Denko, Tokyo, Japan).

Statistics

The statistical screening was conducted by metabolic profiling [7] with modifications as described briefly below. Partial least squares (PLS) regression analysis was carried out with SIMCA-P software (version 12.0, Umetrics, Umeå, Sweden). The data of ammonia production and its medium components were analyzed using Pareto scaling and PLS regression analysis model. The variable importance in the projection (VIP) [8] and regression coefficient were calculated to select candidate inhibitors of ammonia production. The VIP_{Ak} value of the k_{th} explanatory variable in K variables of PLS with the A_{th} component was calculated using the following equation:

$$\mathrm{VIP}=\sqrt{\sum_{a=1}^{A}(w_{ak}{}^{2}\times(SSY_{a\text{-}1}\text{-}SSY_{a}))\times\frac{K}{(SSY_{0}\text{-}SSY_{A})}}$$

A is the total of the latent variable, w_{ak} is the PLS weight of the k_{th} variable in the a_{th} latent variable, and SSY indicates the variance of the predicted residuals by each PLS component. The VIP value indicates the contribution of each explanatory variable to the model [8]. The "VIP scores >1" rule is generally used as the criterion for important variable selection [9]. The variable, *Y*, the variable, *X*, the coefficients, *B*, and error matrix, *F**, are represented by the internal relationship $Y = XB + F^*$ [10]. For interpreting the influence of the variables *X* on *Y,* the coefficients are calculated as coeffCS in SIMCA-P [8]. Dunnett's and Tukey's tests were performed using the JMP software version 14 (SAS, North Carolina, U.S.) for multiple comparisons.

Results

Screening for inhibitors of ammonia production from pretreated food by-products

The amount of ammonia produced from food by-products should correlate with the available amino acids, a major nitrogen source in food by-products, because the amino acid concentration in the medium affects *E. coli* ammonia production. [1, 11]. However, ammonia production could be inhibited by other components present in food by-products. To examine whether inhibitory components of ammonia production are present in food by-products, ammonia was produced by an *E. coli* wild-type strain (DH10B) from four

pretreated food by-products and six media containing rich amino acids. Ammonia production was not correlated with the concentration of total amino acids in the media (R^2 = 0.05, Fig. 1), suggesting that the pretreated food by-products contain potent ammonia production inhibitors.

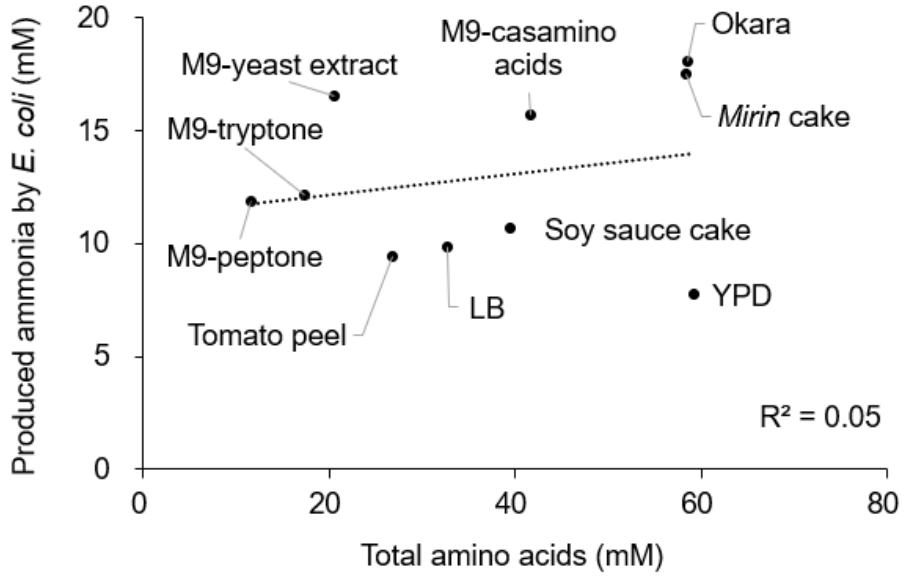

Figure 1 Relationship between ammonia production and total amino acids. Ammonia was produced by *E. coli* DH10B from four food by-products and six amino acid-containing media. *E. coli* strains were initially grown overnight in M9-yeast extract, then the cells were washed with sterile water and inoculated into 2 mL of pretreated food by-products or semisynthesized medium to a final OD_{600} of 0.5. After incubation at 37 °C for 26.5 h with shaking, the concentrations of ammonia in the supernatants were measured. The data of ammonia concentration are represented as the mean values (n = 3), and total amino acid concentration in the medium was determined at n = 1.

Metabolic profiling by PLS regression analysis was then performed to identify the inhibitors. VIP values obtained by PLS regression analysis can be used to predict important factors in a wide range of fields, including epidemiological studies [12, 13], food science [7, 14], and biotechnology [15, 16]. To clarify the relationship between metabolite data and ammonia production, PLS regression and VIP analysis were performed using analytical data for amino acids, sugars, and organic acids as explanatory variables and ammonia production as response variables. Fig. 2 shows the predicted value and the measured value of ammonia production. The correlation of the measured amount of ammonia production with the amount predicted by PLS regression (R^2 = 0.69) was

better than that with the amount of total amino acids (R^2 = 0.05).

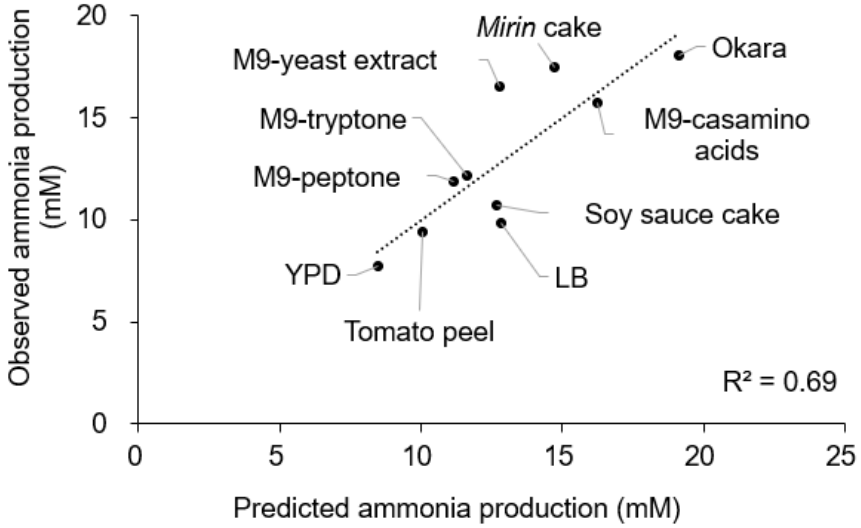

Figure 2 Relationship between actual and predicted ammonia production using the PLS model. The concentration of ammonia produced by *E. coli* DH10B from four food by-products and six amino acid-containing media was plotted against the concentration of ammonia predicted by PLS regression. The data of ammonia concentration are expressed as the mean values (n = 3), and the predicted ammonia production was calculated by PLS regression analysis using the data (n=1) of the concentration of amino acids, sugars, and organic acids in the medium.

The VIP and regression coefficient were used to detect the inhibitor. VIP is an indicator of the effect of explanatory variables on model response variables and is used to select important metabolites in metabolomics [17]. The regression coefficient determines whether an explanatory variable has a positive or negative effect on a response variable. Glucose showed the highest VIP with a negative regression coefficient, indicating that glucose is the most probable ammonia production inhibitor (Table 3). In bioethanol production, glucose concentrations up to 2.78 M are produced from food by-product digests [18]. In contrast, pretreated food by-products, which are mainly composed of amino acids, have a lower glucose concentration than biomass digests (74 mM to 123 mM) (Table 3 and 4).

Table 3 Concentration of glucose in various medium

Medium	Type	Glucose (mM)	Fructose (mM)	Total amino acids (mM)
Soy sauce cake	Food by-product	88.6	3.1	39.4
Mirin cake	Food by-product	123.3	2.3	58.3
Tomato peel	Food by-product	95.7	4.0	26.8
Okara	Food by-product	74.4	12.9	58.5
M9-yeast extract	Semisynthesized medium	n.d.	n.d.	20.5
M9-tryptone	Semisynthesized medium	n.d.	n.d.	17.4
M9-peptone	Semisynthesized medium	n.d.	n.d.	11.6
M9-casamino acids	Semisynthesized medium	n.d.	n.d.	41.7
LB	Semisynthesized medium	n.d.	n.d.	32.6
YPD	Semisynthesized medium	108.6	n.d.	59.2

n.d.: not detected

Table 4 Concentration of amino acids in various medium (expressed as mM)

	Soy sauce cake	*Mirin* cake	Tomato peel	Okara	M9-yeast extract	M9-tryptone	M9-peptone	M9-casamino acids	LB	YPD
Aspartic acid	1.91	2.71	1.85	2.85	1.17	0.42	0.29	3.74	1.19	2.56
Threonine	2.50	3.46	1.80	3.61	1.05	1.02	0.39	2.45	1.62	2.86
Serine	2.86	4.40	1.98	4.62	1.40	0.87	0.46	3.37	1.80	3.16
Glutamic acid	2.94	4.40	2.70	6.08	3.64	0.95	0.75	10.38	3.46	7.31
Glycine	2.10	3.54	1.40	3.34	1.23	0.36	1.28	1.76	1.18	5.09
Alanine	3.73	5.55	2.43	4.95	3.43	1.01	1.30	2.67	3.36	8.84
Cysteine	0.41	0.98	0.39	0.80	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
Valine	3.11	5.18	2.08	4.64	1.69	1.59	0.83	3.65	2.85	4.73
Methionine	0.40	1.20	0.35	0.71	0.40	0.33	0.13	0.76	0.89	0.17
Isoleucine	2.18	3.26	1.34	3.21	1.20	0.87	0.43	1.88	1.73	2.89
Leucine	3.16	5.31	2.17	4.97	2.01	3.29	1.11	3.15	4.96	5.97
Tyrosine	5.48	2.92	1.33	2.30	0.39	0.25	0.26	0.22	0.51	1.17
Phenylalanine	2.04	2.80	1.11	2.59	0.96	1.49	0.54	1.34	2.24	2.86
Lysine	1.68	1.84	1.61	3.14	0.06	0.03	n.d.	0.02	0.06	0.22
Histidine	0.48	0.72	0.39	0.96	0.09	n.d.	0.03	n.d.	0.13	0.16
Arginine	1.15	2.44	1.08	2.69	0.67	2.54	0.95	3.21	3.34	3.32
Proline	1.46	2.67	0.93	2.85	0.19	0.31	0.11	0.47	0.39	0.56
Asparagine	1.30	2.52	1.21	2.83	0.20	0.24	0.13	0.97	0.49	0.34
Glutamine	0.48	2.42	0.61	1.41	0.70	1.78	2.64	1.68	2.45	7.01
Total	39.4	58.3	26.8	58.5	20.5	17.4	11.6	41.7	32.6	59.2

n.d.: not detected

To confirm whether glucose is a true inhibitor of ammonia production from food by-products, the effect of glucose addition on ammonia production was examined in an M9-yeast extract medium. Ammonia production was inhibited in a glucose concentration-dependent manner and decreased to 31% of that before glucose was added (Fig. 3). Furthermore, the quantification of the remaining amino acids in the supernatant indicated that the consumption of eight amino acids decreased in a glucose concentration-

dependent manner. These results suggest that glucose inhibits ammonia production from pretreated food by-products by inhibiting amino acid cellular metabolism.

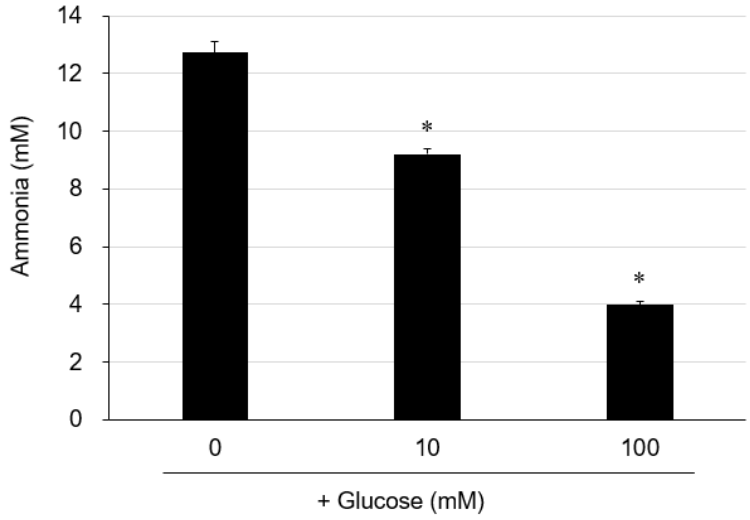

Figure 3 Inhibitory effect of glucose on ammonia production by *E. coli* DH10B. *E. coli* DH10B was incubated in M9-yeast extract with or without glucose (10 or 100 mM) for 26.5 h to produce ammonia. Values are expressed as the mean ± SD (n=3). * Statistical significance was determined by Dunnett's test ($P < 0.001$).

Gene deletion of *E. coli* to improve ammonia production

Based on the above experimental results, I attempted to develop *E. coli* for more efficient ammonia production from pretreated food by-products. To avoid the negative effects of glucose, I suppressed the glucose uptake function in *E. coli* by impairing the phosphoenolpyruvate sugar phosphotransferase system (PTS) that transports major sugars, such as glucose. PTS plays an important role in favoring glucose uptake over other carbon sources [19]. PTS consists of a phosphohistidine carrier protein, an enzyme I component, and enzymes EIIA, EIIB, and EIIC, which are single polypeptide chains encoded by *ptsG* [20]. It was reported that the inactivation of *ptsG* reduces glucose uptake efficiency [21]. Therefore, I disrupted *ptsG* to impair the PTS in *E. coli.* An additional gene, *glnA* encoding glutamine synthetase, was also disrupted to promote the conversion of amino acids to ammonia because a previous study [1] showed that ammonia production during incubation in M9-yeast extract is greatly increased by disrupting *glnA* in *E. coli*.

To investigate whether suppression of glucose uptake and synthesis of glutamine by glutamine synthetase improves ammonia production from glucose-containing media, I performed ammonia production by Δ*ptsG*, Δ*glnA*, and Δ*ptsG*Δ*glnA* strains. In an M9-yeast extract medium containing 50 mM glucose, Δ*ptsG* and Δ*ptsG*Δ*glnA* strains showed

lower glucose uptake rates than the wild-type and Δ*glnA* strains (Fig. 4A). Furthermore, these strains showed higher ammonia productivity than the wild-type strain (Fig. 4B). In particular, Δ*ptsG*Δ*glnA* strain produced about 3.2 fold more ammonia than the wild-type strain. These results suggest that the disruptions of *ptsG* and *glnA* can improve ammonia production even in the medium containing high concentrations of glucose, such as food by-products.

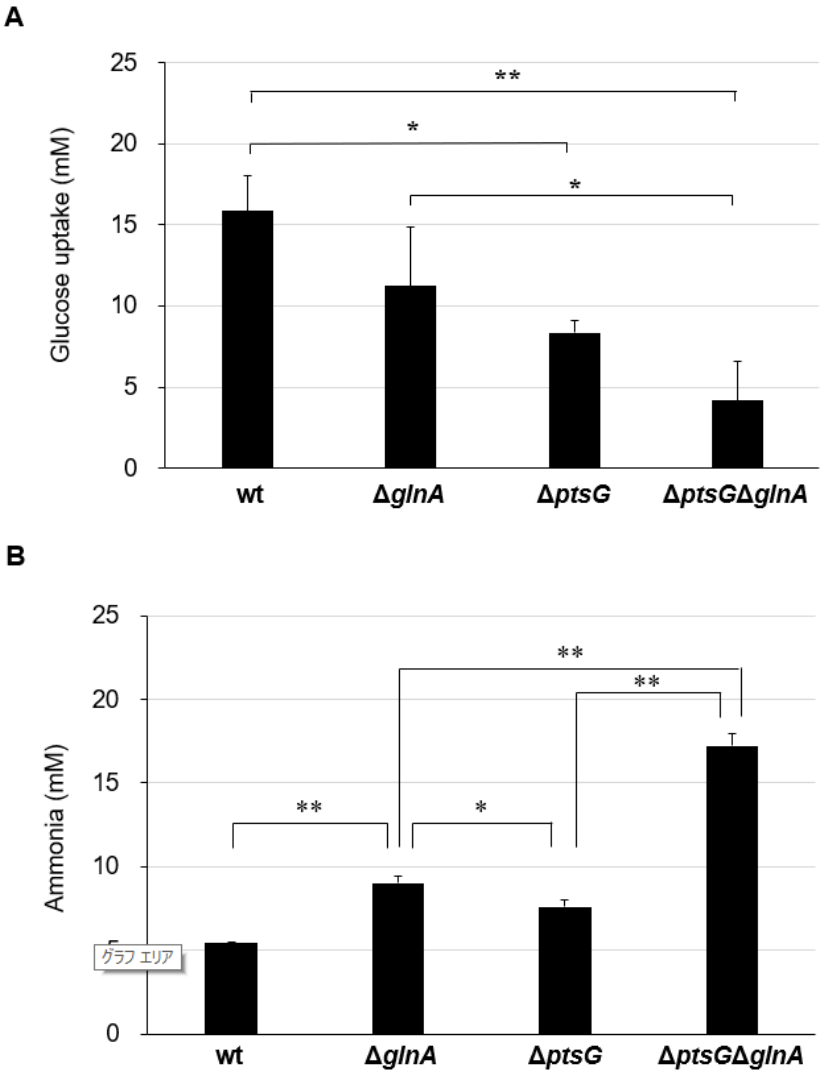

Figure 4 Effects of *ptsG* and *glnA* disruption on glucose uptake (A) and ammonia production (B) in the M9-yeast extract medium with 50 mM glucose. Glucose uptake and ammonia production was measured 2 h and 26.5 h after the start of incubation at 37 ºC in the M9-yeast extract medium, respectively. Values are expressed as the mean ± SD (n=3). Statistical significance was determined by Tukey's test (*P < 0.05, **P<0.01).

Ammonia production from pretreated food by-products by engineered *E. coli*

To evaluate the ammonia productivity of the engineered *E. coli* strains (Δ*ptsG*, Δ*glnA*, and Δ*ptsG*Δ*glnA*), these strains were utilized to produce ammonia from okara, a model of food by-products. Okara is the residue produced in soy milk and tofu production; 1.1 kg of wet okara is produced from 1 kg of soybeans and water [22]. Due to high protein concentrations (27% to 38%), most of the produced okara is used as animal feed or discarded [23]. The proteins included in okara are water-insoluble and cumbersome but can be efficiently extracted by hydrolyzing the cell wall with cellulase [24]. Accordingly, the extracted proteins were treated with protease to produce amino acids, which can be a nitrogen source for ammonia production. I obtained 58.5 mM free amino acids by solubilizing okara with cellulase and protease (Table 3).

The pretreated okara was then incubated with the engineered *E. coli* for ammonia production. The Δ*ptsG* and Δ*glnA* strains slightly increased ammonia production compared with the wild-type strain, suggesting that the strategies for disruption of glucose transporters and the promotion of amino acid catabolism are also effective for ammonia production from the pretreated okara (Fig. 5).

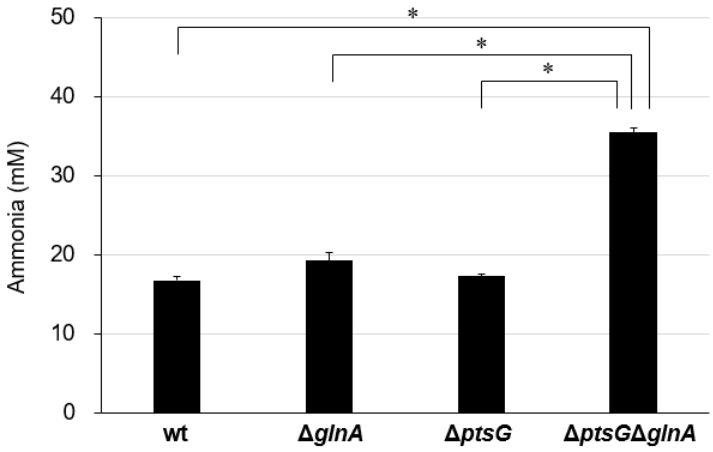

Figure 5 Ammonia production from pretreated okara by engineered *E. coli* strains. Ammonia was produced by incubating cells in the pretreated okara solution at 37 ºC for 26.5 h. Values are expressed as the mean ± SD (n=3). Statistical significance was determined by Tukey's test (*P < 0.0001).

Furthermore, the Δ*ptsG*Δ*glnA* strain, which had the highest ammonia production efficiency in the M9-yeast extract containing 50 mM glucose, also exhibited the highest ammonia production efficiency in the pretreated okara. This represents a 2.1-fold improvement over the wild type, with a conversion efficiency from amino acids of about

47%. These results indicate that the combination of reducing glucose uptake and promoting amino acid catabolism is an effective strategy of efficient ammonia production from pretreated food by-products.

Discussion

The effective use of food by-products is expected to become increasingly important, considering the extensive efforts required for the realization of a sustainable society [25]. In this study, I attempted to screen ammonia production inhibitors included in pretreated food by-products using metabolic profiling. The biomarker screen using VIP and coefficients calculated from PLS regression analysis has been widely used [7, 12-16] , and it was also effective in our screening method for inhibitors from food by-products. Glucose has been known to exert important effects on amino acid metabolism, but this is the first study to highlight the negative effects of glucose on ammonia production from food by-products. Based on this finding, the recombinant *E. coli* with high ammonia productivity was successfully constructed.

Our experiments showed that glucose in food by-products suppressed ammonia production by inhibiting the metabolism of several amino acids (Fig. 2). Ammonia metabolism is tightly regulated in *E. coli*, with the major operon *glnALG* encoding the glutamine synthetases, nitrogen regulatory protein B (NtrB), and nitrogen regulatory protein C (NtrC) [26]. Transcription of the *glnALG* operon is regulated by tandem promoters *glnA*p1 and *glnA*p2. *glnA*p1 is a σ70 dependent promoter that is activated by the cyclic AMP receptor protein (CRP) under high glucose conditions and repressed by NtrC-phosphate under nitrogen-deficient conditions. *glnA*p2 is a σ54 dependent promoter that is strongly activated by NtrC-phosphate and repressed by CRP [27]. Glutamine synthetase is regulated by the *glnA*p1 promoter and activated through CRP when glucose is added to M9-yeast extract. This may promote ammonia assimilation and lead to a decrease in free ammonia.

Additionally, the metabolism of amino acids, especially Gly, was inhibited by the M9-yeast extract containing 50 mM glucose. This was possibly due to the inhibition of purine nucleotide synthesis repressor gene (*purR*) translation by small regulatory RNA (*sgrS*) in response to glucose stress [28, 29]. Since *purR* inhibits the expression of glycine cleavage system proteins encoded by *gcvTHP* that induces Gly degradation [30], there is a possibility that ammonia production by Gly degradation was inhibited under glucose-added conditions.

To suppress ammonia assimilation and promote amino acid metabolism, two strategies were developed to reduce glucose uptake and inhibit glutamine synthetase. To reduce glucose uptake, I disrupted *ptsG*, which encodes a major component of the glucose transporter. However, glucose is also taken up into *E. coli* cells though other transporters such as the galactose transporter (encoded by *mglBAC*), galactose permease (encoded by *galP*), and maltose transporter (encoded by *malEFG*) [31, 32]. In fact, the ammonia production efficiency from *mirin* cake containing excess glucose (123.3 mM) by the *ΔptsGΔglnA* strain was 23%, which was lower than that from okara (47%). By disrupting these transporters simultaneously, glucose uptake efficiency could be further reduced, and ammonia production efficiency could be further improved.

A single disruption of *glnA* improved ammonia production from glucose-added M9-yeast extract, while dual disruption of *ptsG* and *glnA* markedly improved ammonia production efficiency (Fig. 4B). In Δ*ptsG*, inhibition of glucose uptake might lead to a decrease in CRP and activation of *glnA*p2. This activates NtrC and nitrogen assimilation control (Nac) proteins, which control approximately 2% of the genes involved in nitrogen catabolism [33], resulting in the increased production of free ammonia. However, the produced ammonia can be assimilated by glutamine synthetase into amino acids in the cells. Additional disruption of *glnA,* which encodes glutamine synthetase, would increase the free ammonia concentration by inhibiting the assimilation of ammonia to glutamine in the Δ*ptsG* strain, resulting in a synergistic improvement in ammonia productivity. Disrupting genes involved in ammonia assimilation, such as *gdhA* and *gltBD*, is expected to further improve ammonia productivity, but there are concerns about growth retardation and lethality issues. In fact, growth delays have been observed in *ΔglnA* and *ΔglnAΔptsG*, but knockdown of the ammonia assimilation genes by CRISPRi [34, 35] or other methods could solve these problems.

The Δ*ptsG*Δ*glnA* strain produced 35.4 mM ammonia from 58.5 mM amino acids in the pretreated okara (Fig. 5). At the end of incubation, the ammonia production rate was 1.34 mM/h and the conversion efficiency from amino acids to ammonia was approximately 47%, and further improvement of the efficiency is required for industrial use. This result was lower than the conversion efficiency of approximately 73% in the M9-yeast extract containing 50 mM glucose (Fig. 4B), suggesting that other ammonia production inhibitors were present in the pretreated okara to suppress amino acid metabolism or promote ammonia assimilation. To improve ammonia productivity, monitoring of amino acid metabolism by amino acid flux analysis, large-scale analysis by LC-MS and GC-MS, identification of novel inhibitors, and subsequent genetic modification are needed.

In this study, I identified glucose as the cause of inhibitory effects on ammonia production from food by-products using an efficient screening strategy. I improved ammonia production by *E. coli* through the reduction of glucose uptake and the promotion of amino acid catabolism by gene disruptions. These findings provide important insights into ammonia production in *E. coli* as well as an industrially available strain (e.g., *B. subtilis and S. cerevisiae*) and would contribute to the study of ammonia production from various biomasses or other unused resources, potentially expanding the application of ammonia bioproduction.

Summary

Using metabolic profiling, glucose was identified as a potential inhibitor of ammonia production from impure food by-products. I constructed the recombinant *Escherichia coli*, in which glucose uptake was reduced by *ptsG* gene disruption and amino acid catabolism was promoted by *glnA* gene disruption. Ammonia production efficiency from okara, a food by-product, was improved in this strain; 35.4 mM ammonia was produced (47% yield). This study might provide a strategy for efficient ammonia production from food by-products.

References

1. Mikami Y, Yoneda H, Tatsukami Y, Aoki W, Ueda M. Ammonia production from amino acid-based biomass-like sources by engineered *Escherichia coli*. AMB Express 2017; 7:83.
2. Matsui K, Hirayama T, Kuroda K, Shirahige K, Ashikari T, Ueda M. Screening for candidate genes involved in tolerance to organic solvents in yeast. Appl Microbiol Biotechnol 2006; 71:75-79.
3. Datsenko KA, Wanner BL. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc Natl Acad Sci U S A 2000; 97:6640-6645.
4. Wada A, Bonoshita M, Tanaka Y, Hibi K. A study of a reaction system for organic acid analysis using a pH indicator as post-column reagent. J Chromatogr A 1984; 291:111-118.
5. Sano A, Satoh T, Oguma T, Nakatoh A, Satoh J-i, Ohgawara T. Determination of levulinic acid in soy sauce by liquid chromatography with mass spectrometric detection. Food Chem 2007; 105:1242-1247.

6. Suzuki H, Kato E, Matsuzaki A, Ishikawa M, Harada Y, Tanikawa K, Nakagawa H. Analysis of saccharides possessing post-translational protein modifications by phenylhydrazine labeling using high-performance liquid chromatography. Anal Sci 2009; 25:1039-1042.
7. Shiga K, Yamamoto S, Nakajima A, Kodama Y, Imamura M, Sato T, Uchida R, Obata A, Bamba T, Fukusaki E. Metabolic profiling approach to explore compounds related to the umami intensity of soy sauce. J Agric Food Chem 2014; 62:7317-7322.
8. Eriksson L, Johansson E, Kettaneh-Wold N, Trygg J, Wikström C, Wold S Multivariate and Megavariate Data Analysis Basic Principles and Applications (Part I) Umetrics Sweden; 2006.
9. Afanador NL, Tran TN, Buydens LM. Use of the bootstrap and permutation methods for a more robust variable importance in the projection metric for partial least squares regression. Anal Chim Acta 2013; 768:49-56.
10. Palermo G, Piraino P, Zucht HD. Performance of PLS regression coefficients in selecting variables for each response of a multivariate PLS for omics-type data. Adv Appl Bioinform Chem 2009; 2:57-70.
11. Jeremy MB, John LT, Lubert S Biochemistry. In. Edited by Freeman WH, vol. 5th edition. New York; 2002.
12. Jonasson A, Eriksson C, Jenkinson HF, Källestål C, Johansson I, Strömberg N. Innate immunity glycoprotein gp-340 variants may modulate human susceptibility to dental caries. BMC Infect Dis 2007; 7:57.
13. Yan Q, Liew Z, Uppal K, Cui X, Ling C, Heck JE, von Ehrenstein OS, Wu J, Walker DI, Jones DP *et al*. Maternal serum metabolome and traffic-related air pollution exposure in pregnancy. Environ Int 2019; 130:104872.
14. S C Sri Harsha P, Abdul Wahab R, Cuparencu C, Dragsted LO, Brennan L. A metabolomics approach to the identification of urinary biomarkers of pea intake. Nutrients 2018; 10:1911.
15. Chen Z, Zheng Z, Wang F-L, Niu Y-P, Miao J-L, Li H. Intracellular metabolic changes of *Rhodococcus* sp. LH during the biodegradation of diesel oil. Mar Biotechnol 2018; 20:803-812.
16. Zalai D, Koczka K, Párta L, Wechselberger P, Klein T, Herwig C. Combining mechanistic and data-driven approaches to gain process knowledge on the control of the metabolic shift to lactate uptake in a fed-batch CHO process. Biotechnol Prog 2015; 31:1657-1668.
17. Fonville J, Richards S, Barton R, Boulange C, Ebbels T, Nicholson J, Holmes E,

Dumas M-E. The evolution of partial least squares models and related chemometric approaches in metabonomics and metabolic phenotyping. J Chemom 2010; 24:636-649.

18. Mohd Azhar SH, Abdulla R, Jambo SA, Marbawi H, Gansau JA, Mohd Faik AA, Rodrigues KF. Yeasts in sustainable bioethanol production: A review. Biochem Biophys Rep 2017; 10:52-61.
19. Postma PW, Lengeler JW, Jacobson GR. Phosphoenolpyruvate:carbohydrate phosphotransferase systems of bacteria. Microbiol Rev 1993; 57:543-594.
20. Escalante A, Salinas Cervantes A, Gosset G, Bolívar F. Current knowledge of the *Escherichia coli* phosphoenolpyruvate-carbohydrate phosphotransferase system: peculiarities of regulation and impact on growth and product formation. Appl Microbiol Biotechnol 2012; 94:1483-1494.
21. Steinsiek S, Bettenbrock K. Glucose transport in *Escherichia coli* mutant strains with defects in sugar transport systems. J Bacteriol 2012; 194:5897-5908.
22. Khare S, Jha K, Gandhi A. Citric acid production from okara (soy-residue) by solid-state fermentation. Bioresour Technol 1995; 54:323-325.
23. Surel O, Couplet B. Influence of the dehydration process on active compounds of okara during its fractionation. J Sci Food Agric 2005; 85:1343-1349.
24. de Figueiredo VRG, Yamashita F, Vanzela ALL, Ida EI, Kurozawa LE. Action of multi-enzyme complex on protein extraction to obtain a protein concentrate from okara. J Food Sci Technol 2018; 55:1508-1517.
25. Avtar R, Aggarwal R, Kharrazi A, Kumar P, Kurniawan TA. Utilizing geospatial information to implement SDGs and monitor their progress. Environ Monit Assess 2019; 192:35.
26. Reitzer LJ, Magasanik B. Transcription of *glnA* in *E. coli* is stimulated by activator bound to sites far from the promoter. Cell 1986; 45:785-792.
27. Tian ZX, Li QS, Buck M, Kolb A, Wang YP. The CRP–cAMP complex and downregulation of the *glnA*p2 promoter provides a novel regulatory linkage between carbon metabolism and nitrogen assimilation in *Escherichia coli*. Mol Microbiol 2001; 41:911-924.
28. Vanderpool CK, Gottesman S. Involvement of a novel transcriptional activator and small RNA in post-transcriptional regulation of the glucose phosphoenolpyruvate phosphotransferase system. Mol Microbiol 2004; 54:1076-1089.
29. Bobrovskyy M, Vanderpool CK. Diverse mechanisms of post-transcriptional repression by the small RNA regulator of glucose-phosphate stress. Mol

Microbiol 2016; 99:254-273.
30. Cho B-K, Federowicz SA, Embree M, Park Y-S, Kim D, Palsson BØ. The *PurR* regulon in *Escherichia coli* K-12 MG1655. Nucleic Acids Res 2011; 39:6456-6464.
31. Death A, Ferenci T. The importance of the binding-protein-dependent Mgl system to the transport of glucose in *Escherichia coli* growing on low sugar concentrations. Res Microbiol 1993; 144:529-537.
32. Ferenci T. Adaptation to life at micromolar nutrient levels: the regulation of *Escherichia coli* glucose transport by endoinduction and cAMP. FEMS Microbiol Rev 1996; 18:301-317.
33. Zimmer DP, Soupene E, Lee HL, Wendisch VF, Khodursky AB, Peter BJ, Bender RA, Kustu S. Nitrogen regulatory protein C-controlled genes of *Escherichia coli*: scavenging as a defense against nitrogen limitation. Proc Natl Acad Sci U S A 2000; 97:14674-14679.
34. Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, Lim WA. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell 2013; 152:1173-1183.
35. Hawkins JS, Wong S, Peters JM, Almeida R, Qi LS. Targeted Transcriptional Repression in Bacteria Using CRISPR Interference (CRISPRi). Methods Mol Biol 2015; 1311:349-362.
36. Cherepanov PP, Wackernagel W. Gene disruption in *Escherichia coli*: Tc^R and Km^R cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant. Gene 1995; 158:9-14.

Chapter II

Construction of recombinant *Escherichia coli* producing nitrogenase-related proteins from *Azotobacter vinelandii*

In this chapter, I constructed an expression plasmid containing the minimum number of *A. vinelandii* genes required for nitrogen fixation for heterologous expression in *E. coli*. *E. coli* is a gram-negative bacterium commonly studied in nitrogen fixation and metabolic engineering [1, 2]. Moreover, *E. coli* can be further modified to produce useful substances such as ammonia with relative ease, because an extensive genetic toolbox is available in *E. coli*, unlike *A. vinelandii*. The assembly of the scattered nif-related genes into a synthetic cluster was achieved by a multi-step seamless cloning method, and the expression of each gene in *E. coli* was quantified at the transcription and protein levels. The nitrogenase-producing *E. coli* constructed in this study successfully showed nitrogenase activity. The constructed *E. coli* would be useful as a platform microorganism for studying the expression of additional *A. vinelandii* genes and identifying potential factors that may improve nitrogenase activity.

Materials and methods

Strains and media

The *A. vinelandii* and *E. coli* strains and plasmids used in this study are described in Table 1. *A. vinelandii* was aerobically grown at 30°C in modified Burk's (MB) medium [3] consisting of 20 g/L sucrose, 0.2 g/L NaCl, 0.05 g/L $CaSO_4 \cdot 2H_2O$, 0.2 g/L $MgSO_4 \cdot 7H_2O$, 2.9 mg/L $Na_2MoO_4 \cdot 2H_2O$, 27 mg/L $FeSO_4 \cdot 7H_2O$, 0.66 g/L K_2HPO_4, and 0.16 g/L KH_2PO_4. *E. coli* strains were aerobically grown at 37°C in LB broth (BD Difco, NJ, USA). For measurements of transcript levels, protein levels, and nitrogenase activity in *E. coli* strain, cells were grown in the previously reported medium [4] modified as follows (10.4 g/L Na_2HPO_4, 3.4 g/L KH_2PO_4, 26 mg/L $CaCl_2 \cdot 2H_2O$, 30 mg/L $MgSO_4$, 0.3 mg/L $MnSO_4$, 41 mg/L $FeSO_4 \cdot 7H_2O$, 10 mg/L para-aminobenzoic acid, 5 μg/L biotin, 337 mg/L vitamin B1, 10 mM glutamate, 4 g/L glucose, and 7.6 mg/L Na_2MoO_4). Ampicillin (100 μg/mL), kanamycin (25 μg/mL), and chloramphenicol (10 μg/mL) (Fujifilm Wako Pure Chemical, Osaka, Japan) were added to the medium as appropriate.

Plasmid construction

To construct plasmids for expressing nif-related genes in *E. coli*, genomic DNA

of *A. vinelandii* was extracted using ISOGEN (NIPPON GENE, Tokyo, Japan). Seventeen nif-related genes were classified into eight clusters (*nifHDK, nifB, iscA-nifUSV, nifQ, nifENX, nifY, nifWZM,* and *nifF)* and amplified by PCR from a genomic DNA template using KOD-One polymerase (Toyobo, Osaka, Japan) and primers that contained 15-50 bp regions homologous to neighboring vectors sequences or nif-related genes. In the primer annealing to the 5' region of *iscA-nifUSV*, the *lac* promoter sequence was inserted between the homologous region and the coding region. The PCR amplicons of *nifHDK*, *nifB*, *iscA-nifUSV*, *and nifQ* were assembled into a single DNA fragment in this order by overlap extension (OE)-PCR [5] and fused to PCR-amplified pTrcHis2-TOPO (Thermo Fisher Scientific, MA, USA) using the In-Fusion HD Cloning Kit (Takara Bio, Shiga, Japan). The resulting plasmid was named pTrc-nif001 (Fig. 1). Likewise, the PCR amplicons of *nifENX*, *nifY*, *nifWZM*, and *nifF* were assembled in this order and fused to PCR-amplified pMW219 (NIPPON GENE, Tokyo, Japan). The *nifWZM* cluster was amplified using a forward primer with a *lac* promoter. The resulting plasmid was named pMW-nif002 (Fig. 1). The sequences of the assembled PCR amplicons in the constructed plasmids were confirmed by Sanger sequencing using sequencing primers.

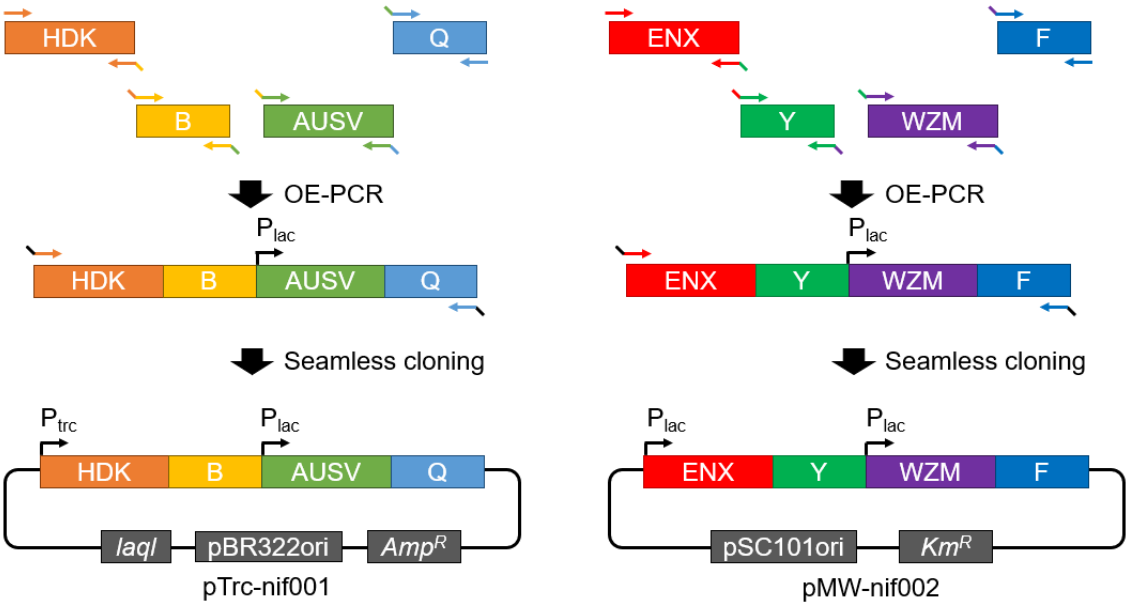

Figure 1. The assembly of 17 nif-related genes into expression plasmids for *E. coli* was performed by a combination of OE-PCR and seamless cloning. Expression of nif-related genes in *E. coli* was performed using two plasmids with different copy numbers. Artificial gene clusters of 4 genes (*nif, H, D, K,* and *B*) controlled by the trc promoter and 5 genes (*iscA, nifU, S, V,* and *Q*) controlled by the lac promoter were assembled in medium copy

plasmid pTrcHis2-TOPO/lacZ. Artificial gene clusters of 4 genes (*nifE, N, X,* and *Y*) controlled by the *lac* promoter and 4 genes (*nifW, Z, M,* and *F*) controlled by the *lac* promoter were assembled in low copy plasmid pMW219.

Measurement of transcription and protein levels from nif-related genes

E. coli JM109 transformants harboring pTrc-nif001and pMW-nif002 were cultured in LB at 37°C with overnight shaking at 200 rpm. *E. coli* or pre-cultured *A. vinelandii* cells were inoculated into 10 mL nitrogen-deficient medium with or without 0.1 mM IPTG or MB medium, respectively, at a final OD_{600} of 0.5. Incubation was performed at 30°C for 6 h to measure the transcription levels or 24 h for the measurement of protein levels. The cells were collected by centrifugation at 5,000 × *g*, washed twice with saline solution, and stored at -80°C.

To measure the transcription levels of nif-related genes, cells were lysed in ISOGEN (NIPPON GENE) to extract RNA, which was then used as template for cDNA synthesis by reverse transcription using PrimeScript RT reagent Kit with gDNA Eraser (Takara Bio). RT-qPCR was performed using TB Green Premix Ex Taq II (Takara Bio) on an MX3005 real-time PCR system (Agilent Technologies, CA, USA) with cDNA and specific primers listed in Table S1. Fold change values of transcription levels were calculated using a template standard curve and normalized to the transcription level of *ropA* that is often regarded as a housekeeping gene in prokaryotes [6, 7].

For the measurement of protein levels from nif-related genes, the cells were sonicated in 200 µL lysis buffer (2% SDS and 7 M urea) with a bioruptor (Sonic bio, Kanagawa, Japan), and the supernatant after centrifugation at 15,000 × *g* at 4°C for 1 min was transferred to a microtube. The lysate was mixed with 800 µL methanol, 200 µL chloroform, and 600 µL water and centrifuged at 15,000 × *g* at 4°C for 1 min. After removing the aqueous phase, 800 µL methanol was added and the mixture was centrifuged at 15,000 × *g* for 2 min at 4°C. The pellet was dried using SpeedVac vacuum concentrators (Thermo Fisher Scientific) and re-dissolved in 120 µL of 7 M urea/50 mM Tris-HCl (pH 8.0), and the total protein concentration was measured using a BCA Protein Assay Kit (Thermo Fisher Scientific). The protein extract (100 µL) was treated with 5 mM dithiothreitol (final concentration) at 37°C for 30 min, followed by treatment with 15 mM 2-iodoacetamide at 37°C for 30 min under light shielding. Trypsin/Lys-C Mix (Promega, Wisconsin, USA) was then added at a protein : protease ratio of 25 : 1 (w/w). After incubation at 37°C for 3 h, the reaction solution was diluted with 600 µL of 50 mM Tris-HCl (pH 8) and incubated at 37°C overnight. Digestion was stopped by adding trifluoroacetic acid (1% v/v) and desalting with Monospin C18 (GL Sciences, Tokyo,

Japan) to obtain peptide samples for LC-MS/MS analysis.

LC-MS/MS analysis was performed on an LCMS8050 (Shimadzu, Kyoto, Japan) connected to a UHPLC Nexera (Shimadzu). Peptides were separated on an InertSustain AQ-C18 column (150 mm × 2.1 mm, 1.0 μm; GL Sciences) by gradient elution of mobile phases A (0.1% formic acid/water, v/v) and B (0.1% formic acid/acetonitrile, v/v). Gradient elution was programmed as follows: isotropic elution with 5% B for 3 min, linear gradient elution with 5% to 95% B for 12 min, wash with 95% B for 2 min, and finally equilibration with starting conditions for 5 min. The column temperature was set at 40°C throughout the analysis. The separated peptides were quantified using multiple reaction monitoring (MRM) mode by triple quadrupole mass spectrometry equipped with an electrospray ionization source. The MS conditions were as follows: electrospray voltage, 4.0 kV; capillary temperature, 300°C; and sheath gas, N_2, 10 L/min. Candidate peptides derived from trypsin digestion of nif-related proteins were identified using Skyline software [8] and narrowed down to peptides with a length of 5-20 amino acids. Among the candidate peptides, the peptides showing the highest detection sensitivity in LC-MS/MS analysis were synthesized by Bio-Synthesis (TX, USA), and the elution time and transitions were determined (Table 2).

Evaluation of nitrogenase activity

In vivo nitrogenase activity was measured using the acetylene reduction method as described previously [9]. In advance, a headspace vial was filled with 4.5 mL assay medium supplemented with antibiotics and inducers, sealed with a butyl rubber septum and crimp cap, and replaced with argon gas several times. The cells of *A. vinelandii* pre-cultured in MB medium at 30°C and *E. coli* pre-cultured in LB at 37°C were centrifuged at 3,000 rpm at room temperature. After discarding the supernatant, the cell pellet was suspended in the fresh assay medium to an OD_{600} of 0.5, and 0.5 mL of the cell suspension was added to the above-mentioned vial using a syringe. Acetylene gas was prepared in a sagittal gas generator with calcium carbide (Sigma-Aldrich, MO, USA) and water in a vacuum. Acetylene gas (200 mL) and ethane gas (1 mL, internal standard) were mixed in a sample bag and 1 mL of the mixture was injected into the vials containing the cells using a gas-tight syringe. After incubation at 30°C for 16 h, the reaction was stopped by injecting 300 μL of 4 M NaOH into each vial. The produced ethylene (m/z: 26) and the internal standard ethane (m/z: 30) were detected using GCMS QP2020Ultra (Shimadzu) connected to a TurboMatrix HS110 headspace autosampler (PerkinElmer, MA, USA) with the injector temperature, column oven temperature, and detector temperature set at 40°C, 40°C, and 250°C, respectively, and a GS GasPro column (0.32 mm, 30 m) (Agilent,

CA, USA); the ethylene concentration was determined by the external standard method.

Results

Construction of artificial gene cluster of nif-related genes from *A. vinelandii*

To reconstitute nitrogenase from *A. vinelandii* in *E. coli*, nif-related genes scattered in the *A. vinelandii* genome were assembled into a gene cluster by OE-PCR and inserted into an expression plasmid by seamless cloning. In the case of reconstitution of nif-related genes from *K. oxytoca* in *E. coli*, 18 or 20 genes (*nifHDKYENJBQFUSVWZMT* with or without *nifTX*) were selected [1, 4]. In this study, 17 genes (*nifHDKBUSVQENXYWZMF* and *iscA*) were selected from *A. vinelandii* based on transcriptional analysis [10] and phenotypic analysis of nif-related gene deletion mutants [11]. The expression of nif-related genes in *E. coli* was controlled by an IPTG-inducible promoter (Fig. 1). For expression in *E. coli*, nine nif-related genes (*nifHDK*, which constitute nitrogenase; *nifBUSVQ* and *iscA*, which are involved in the maturation of nitrogenase) that are highly transcribed under the nitrogen-fixation culture conditions in *A. vinelandii* were incorporated into a medium-copy plasmid to construct pTrc-nif001, and eight genes (*nifENXYWZMF*, which are involved in the maturation of nitrogenase) with relatively low transcript levels compared to *nifHDK* were integrated into a low-copy plasmid to construct pMW-ni002 (Fig. 1).

Confirmation of gene expression at transcription and protein levels

Expression of the nif-related genes at the transcription level was analyzed using RT-qPCR. IPTG induction resulted in a 2.9- to 8.7-fold increase in the transcription levels of 17 nif-related genes in recombinant *E. coli* (Fig. 2). In addition, expression of the 17 nif-related genes was confirmed at the protein level by quantification of their translated products in total protein extracts of recombinant *E. coli* after induction, employing *A. vinelandii* extract as a control. The peptides obtained by trypsin digestion of the respective protein extracts were quantified using LC-MS/MS. The levels of nitrogenase MoFe protein α-subunit (NifD), β-subunit (NifK), and the Fe protein (NifH) that supplies electrons to MoFe proteins were higher in recombinant *E. coli* than those in *A. vinelandii*. The levels of NifS, NifV, NifB, NifQ, NifU, NifE and NifN in recombinant *E. coli* were also significantly higher than those in *A. vinelandii* (Fig. 3).

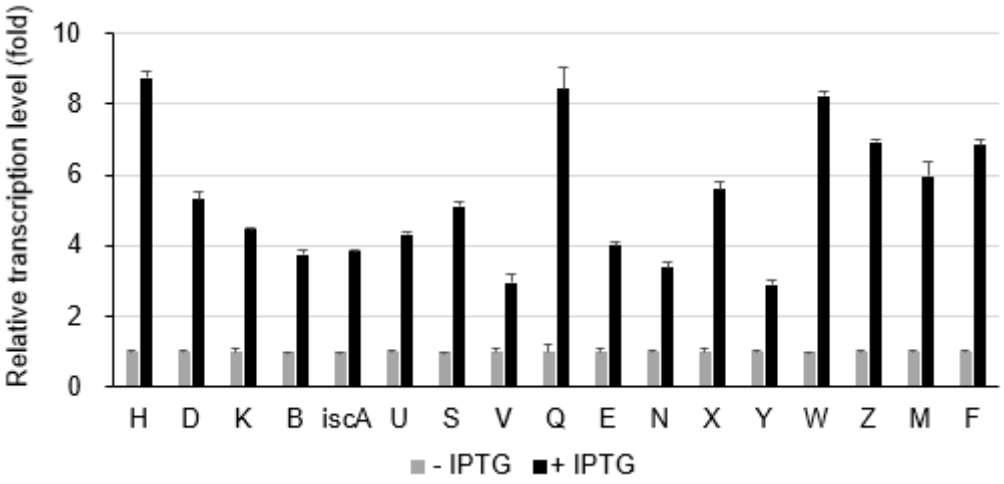

Figure 2. Evaluation of the transcription levels of nif-related genes in the *E. coli* transformants. Cells harboring pTrc-nif001 and pMW-nif002 were collected 6 h after IPTG induction, and the transcript levels of nif-related genes were evaluated by RT-qPCR. Transcript levels of each gene were normalized to *rpoA*, and are shown as fold change relative to uninduced control. Data are represented as means ± SD of three independent experiments.

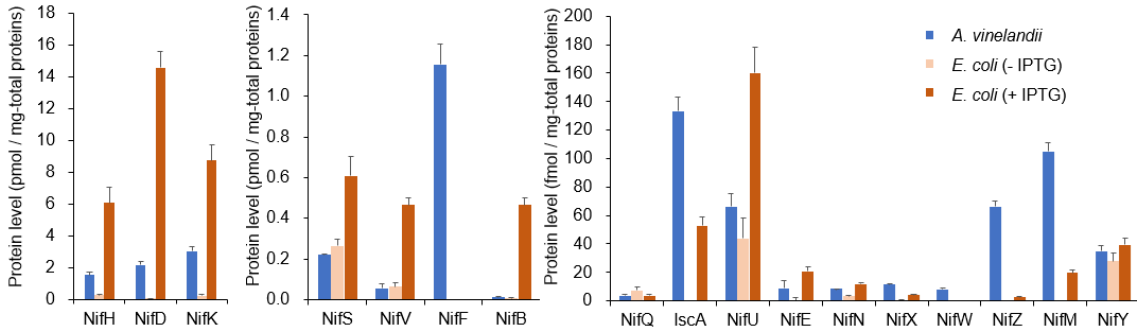

Figure 3. Evaluation of protein level of the introduced nif-related genes in the *E. coli* transformant. Cells harboring pTrc-nif001 and pMW-nif002 were collected 24 h after the IPTG induction, and subjected to total protein extraction. The protein extracts were digested with trypsin to obtain peptides with nif-related protein-specific sequences and quantified by LC-MS/MS. The expression level of each protein was normalized to total protein content. Data are represented as means ± SD of three independent experiments.

In contrast, the protein levels of NifF, iscA, NifX, NifW, NifZ, NifM and NifY in recombinant *E. coli* were lower than those in *A. vinelandii.* Interestingly, more than 10-fold variability was observed in protein levels (NifH, NifD, and NifK vs. NifQ, NifU, and IscA), even though the corresponding genes were integrated into the same plasmid in recombinant *E. coli*. These results showed that the major nif-related genes exhibited an expression profile similar to that in *A. vinelandii* when expressed in *E. coli* by the plasmids integrating the scattered nif-related genes from *A. vinelandii*.

***In vivo* nitrogenase activity under anaerobic and microaerobic conditions**

In vivo nitrogenase activity of the constructed recombinant strains was evaluated by monitoring the reduction of acetylene to ethylene under anaerobic conditions, in which the gas phase was completely replaced by argon gas, and under microaerobic conditions, in which 20% (v/v) of the gas phase was replaced by oxygen gas after argon gas replacement (Fig. 4). The recombinant *E. coli* produced 4.96 nmol-ethylene / mL-culture volume under anaerobic conditions, which was comparable to the production by *A. vinelandii* (5.69 nmol/mL). In contrast, ethylene production by the recombinant *E. coli* decreased to 4.13 nmol/mL under microaerobic conditions, which was much lower than by *A. vinelandii* under similar conditions (25.8 nmol/mL). Nevertheless, these results showed that the recombinant *E. coli* expressing the assembled nif-related genes from *A. vinelandii* had sufficient nitrogenase activity under anaerobic conditions.

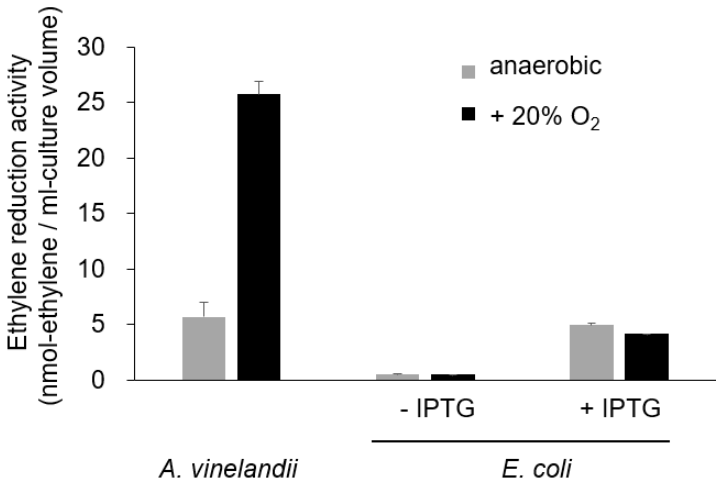

Figure 4. Nitrogenase activity of the wild- type *A. vinelandii* and the *E. coli* transformant harboring pTrc-nif001 and pMW-nif002 was determined under anaerobic and microaerophilic conditions by the acetylene reduction method. The nitrogenase activity

of each strains was normalized to the culture volume. Data are represented as means ± SD of three independent experiments.

Discussion

In this study, I attempted to introduce nitrogenase activity in *E. coli* by assembling 17 nif-related genes scattered in the *A. vinelandii* genome into expression plasmids. The introduction of the expression plasmids constructed by OE-PCR and seamless cloning drove nif-related gene transcription and translation in *E. coli*. In previous studies on nitrogenase reconstruction in heterologous microorganisms, nif-related genes were partially derived from *K. oxytoca* or *Paenibacillus* sp. and contained extra or missing genes from *A. vinelandii* [12-16]. These recombinant microorganisms are not always appropriate for elucidating the function of *Azotobacter* genes because they express nitrogenase-related genes derived from other species and more genes than necessary. This study is the first to reconstitute nitrogenase with a minimal gene cassette from a single species, and our nitrogenase-producing *E. coli* may be used to further understand the nitrogen fixation reaction and oxygen tolerance mechanism.

The reconstituted recombinant *E. coli* had sufficient nitrogenase activity under anaerobic conditions. However, the protein levels of some of the nif-related genes were insufficient compared to those in *A. vinelandii*. In particular, the protein levels of NifW and NifZ (proteins involved in the maturation of P clusters) and NifF (flavodoxin protein supplying electron donor of MoFe protein) were very low (Fig. 3). *nifZ* was reported to be functionally compensated by *nifM* [17], and it was possible that the nitrogenase activity in recombinant *E. coli* was not significantly affected. By contrast, the nitrogenase activity of *nifF* and *nifW* gene-disrupted strains of *A. vinelandii* have been reported to be significantly reduced relative to the wild type [17-19]. However, considering the low nitrogenase activity of *A. vinelandii* (Fig. 4), the effect of low expression levels of *nifF* and *nifW* on the nitrogenase activity of recombinant *E. coli* might be limited under anaerobic conditions.

Under microaerobic conditions, recombinant *E. coli* showed much lower nitrogenase activity than *A. vinelandii*, which may be due to low expression levels of *nifF* and *nifW*, as well as insufficient electron supply to nitrogenase and lack of a system protecting nitrogenase from oxygen. The gene cluster consisting of *rnf1*, *rnf2*, and *fix* encodes an electron transport system that may supply reducing equivalents to nitrogenase in *A. vinelandii* [20]. Single or multiple disruptions of *rnf1*, *rnf2*, or *fix* significantly affect cell growth dependent on nitrogen fixation under aerobic conditions [21]. In recombinant

E. coli, the electron supply pathway is limited to the intrinsic *fix*, and thus the scarcity of this pathway may lead to weak nitrogenase activity. In addition, because nitrogenase is irreversibly inactivated by oxygen, *A. vinelandii* is thought to have physical and chemical systems to protect nitrogenase from oxygen. In *A. vinelandii*, it has been suggested that nitrogenase stabilization by cell aggregation [22], complex stabilization by binding of FeSII to nitrogenase [23-25], and oxygen removal by high respiratory activity ensure high nitrogenase activity even under microaerobic conditions [26-28]. By increasing the expression of *nifF* and *nifW* using high copy plasmid, strong promoter (e.g. T7 promoter) or codon-optimized gene, adding the electron supply pathways and oxygen protection systems to the constructed recombinant *E. coli*, nitrogenase activity may be improved even under microaerobic conditions.

In conclusion, I have constructed a recombinant *E. coli* strain that harbors artificial gene clusters consisting of nif-related genes from *A. vinelandii* and reconstitutes a functional nitrogenase. This strain showed nitrogenase activity comparable to that of *A. vinelandii* under anaerobic conditions. Therefore, our study provides a useful platform to explore additional factors that influence nitrogen fixation. In the future, it may be possible to efficiently fix nitrogen even in the presence of oxygen by mimicking the oxygen tolerance mechanism of *A. vinelandii* in recombinant *E. coli*.

Summary

Azotobacter vinelandii, which produces an oxygen-sensitive nitrogenase, can fix nitrogen even under aerobic conditions; therefore, the heterologous expression of nif-related genes from *A. vinelandii* is a promising strategy for developing a biological nitrogen fixation method. I assembled 17 nif-related genes, which are scattered throughout the genome of *A. vinelandii*, into synthetic gene clusters by overlap-extension-PCR and seamless cloning and expressed them in *Escherichia coli*. The transcription and translation of the 17 nif-related genes were evaluated by RT-qPCR and LC-MS/MS, respectively. The constructed *E. coli* showed nitrogenase activity under anaerobic and microaerobic conditions. This strain would be a useful model for examining the effect of other genes from *A. vinelandii* on nitrogen fixation by expressing them in addition to the minimal set of nif-related genes.

References

1. Dixon RA, Postgate JR. Genetic transfer of nitrogen fixation from *Klebsiella pneumoniae* to *Escherichia coli*. Nature 1972; 237:102-103.
2. Pontrelli S, Chiu TY, Lan EI, Chen FY, Chang P, Liao JC. *Escherichia coli* as a host for metabolic engineering. Metab Eng 2018; 50:16-46.
3. Peña C, Campos N, Galindo E. Changes in alginate molecular mass distributions, broth viscosity and morphology of *Azotobacter vinelandii* cultured in shake flasks. Appl Microbiol Biotechnol 1997; 48:510-515.
4. Wang X, Yang JG, Chen L, Wang JL, Cheng Q, Dixon R, Wang YP. Using synthetic biology to distinguish and overcome regulatory and functional barriers related to nitrogen fixation. PLOS ONE 2013; 8:e68677.
5. Shao Z, Zhao H, Zhao H. DNA assembler, an *in vivo* genetic method for rapid construction of biochemical pathways. Nucleic Acids Res 2009; 37:e16.
6. Mulder J, Wels M, Kuipers OP, Kleerebezem M, Bron PA. Unleashing Natural Competence in Lactococcus lactis by Induction of the Competence Regulator ComX. Appl Environ Microbiol 2017; 83.
7. Valat C, Hirchaud E, Drapeau A, Touzain F, de Boisseson C, Haenni M, Blanchard Y, Madec JY. Overall changes in the transcriptome of *Escherichia coli* O26:H11 induced by a subinhibitory concentration of ciprofloxacin. J Appl Microbiol 2020; 129:1577-1588.
8. MacLean B, Tomazela DM, Shulman N, Chambers M, Finney GL, Frewen B, Kern R, Tabb DL, Liebler DC, MacCoss MJ. Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. Bioinformatics 2010; 26:966-968.
9. Temme K, Zhao D, Voigt CA. Refactoring the nitrogen fixation gene cluster from *Klebsiella oxytoca*. Proc Natl Acad Sci USA 2012; 109:7085-7090.
10. Hamilton TL, Ludwig M, Dixon R, Boyd ES, Dos Santos PC, Setubal JC, Bryant DA, Dean DR, Peters JW. Transcriptional profiling of nitrogen fixation in *Azotobacter vinelandii*. J Bacteriol 2011; 193:4477-4486.
11. Burén S, Jiménez-Vicente E, Echavarri-Erasun C, Rubio LM. Biosynthesis of nitrogenase cofactors. Chem Rev 2020; 120:4921-4968.
12. Burén S, Jiang X, López-Torrejón G, Echavarri-Erasun C, Rubio LM. Purification and *in vitro* activity of mitochondria targeted nitrogenase cofactor maturase *nifB*. Front Plant Sci 2017; 8:1567.
13. López-Torrejón G, Jiménez-Vicente E, Buesa JM, Hernandez JA, Verma HK,

Rubio LM. Expression of a functional oxygen-labile nitrogenase component in the mitochondrial matrix of aerobically grown yeast. Nat Commun 2016; 7:11426.
14. Ryu MH, Zhang J, Toth T, Khokhani D, Geddes BA, Mus F, Garcia-Costas A, Peters JW, Poole PS, Ane JM *et al*. Control of nitrogen fixation in bacteria that associate with cereals. Nat Microbiol 2020; 5:314-330.
15. Yang J, Xie X, Yang M, Dixon R, Wang YP. Modular electron-transport chains from eukaryotic organelles function to support nitrogenase activity. Proc Natl Acad Sci USA 2017; 114:E2460-2465.
16. Yang J, Xie X, Wang X, Dixon R, Wang YP. Reconstruction and minimal gene requirements for the alternative iron-only nitrogenase in *Escherichia coli*. Proc Natl Acad Sci USA 2014; 111:E3718-3725.
17. Jacobson MR, Cash VL, Weiss MC, Laird NF, Newton WE, Dean DR. Biochemical and genetic analysis of the *nifUSVWZM* cluster from *Azotobacter vinelandii*. Mol Gen Genet 1989; 219:49-57.
18. Robinson AC, Dean DR, Burgess BK. Iron-molybdenum cofactor biosynthesis in *Azotobacter vinelandii* requires the iron protein of nitrogenase. J Biol Chem 1987; 262:14327-14332.
19. Bennett LT, Jacobson MR, Dean DR. Isolation, sequencing, and mutagenesis of the *nifF* gene encoding flavodoxin from *Azotobacter vinelandii*. J Biol Chem 1988; 263:1364-1369.
20. Mus F, Alleman AB, Pence N, Seefeldt LC, Peters JW. Exploring the alternatives of biological nitrogen fixation. Metallomics 2018; 10:523-538.
21. Ledbetter RN, Garcia Costas AM, Lubner CE, Mulder DW, Tokmina-Lukaszewska M, Artz JH, Patterson A, Magnuson TS, Jay ZJ, Duan HD *et al*. The electron bifurcating FixABCX protein complex from *Azotobacter vinelandii*: generation of low-potential reducing equivalents for nitrogenase catalysis. Biochemistry 2017; 56:4177-4190.
22. Sabra W, Zeng AP, Lünsdorf H, Deckwer WD. Effect of oxygen on formation and structure of *Azotobacter vinelandii* alginate and its role in protecting nitrogenase. Appl Environ Microbiol 2000; 66:4037-4044.
23. Robson RL. Characterization of an oxygen-stable nitrogenase complex isolated from *Azotobacter chroococcum*. Biochem J 1979; 181:569-575.
24. Maier RJ, Moshiri F. Role of the *Azotobacter vinelandii* nitrogenase-protective shethna protein in preventing oxygen-mediated cell death. J Bacteriol 2000; 182:3854-3857.
25. Schlesier J, Rohde M, Gerhardt S, Einsle O. A conformational switch triggers

nitrogenase protection from oxygen damage by shethna protein II (FeSII). J Am Chem Soc 2016; 138:239-247.

26. Kelly MJ, Poole RK, Yates MG, Kennedy C. Cloning and mutagenesis of genes encoding the cytochrome bd terminal oxidase complex in *Azotobacter vinelandii*: mutants deficient in the cytochrome d complex are unable to fix nitrogen in air. J Bacteriol 1990; 172:6010-6019.

27. Leung D, van der Oost J, Kelly M, Saraste M, Hill S, Poole RK. Mutagenesis of a gene encoding a cytochrome o-like terminal oxidase of *Azotobacter vinelandii*: a cytochrome o mutant is aero-tolerant during nitrogen fixation. FEMS Microbiol Lett 1994; 119:351-357.

28. Gallon JR. Reconciling the incompatible: N_2 fixation and O_2. New Phytol 2006; 122:571-609.

Chapter III

Identification of genes from *Azotobacter vinelandii* for improving nitrogenase activity in *Escherichia coli* expressing *nif* genes

In this chapter, I identified genes from *A. vinelandii* that enhance nitrogenase activity in *E. coli* harboring 17 nitrogenase genes. In our previous study, RNA-seq screening suggested 210 genes showing a similar transcription pattern to *nif* genes, such as *nifH*, *nifD*, and *nifK* (Takimoto, R. in revision). I assumed that the transcription pattern of the genes contributing to the maintenance or enhancement of nitrogenase activity would be the same as that of *nif* genes. Therefore, in this study, I focused on the 210 genes and constructed a plasmid library expressing these genes, except for rRNA coding and non-coding RNA genes.

Materials and methods

Strains and media

A. vinelandii and *E. coli* strains used in this study are described in Table 1. *A. vinelandii* was aerobically grown at 30°C in modified Burk (MB) medium [1], slightly modified as described previously [2]. *E. coli* strains were aerobically grown at 37°C in Luria Bertani (LB) broth (BD Difco, NJ, USA). To measure the nitrogenase activity and metabolites in the *E. coli* strain, cells were grown in a nitrogen-deficient medium composed of 10.4 g/L Na_2HPO_4, 3.4 g/L KH_2PO_4, 26 mg/L $CaCl_2 \cdot 2H_2O$, 30 mg/L $MgSO_4$, 0.3 mg/L $MnSO_4$, 41 mg/L $FeSO_4 \cdot 7H_2O$, 10 mg/L para-aminobenzoic acid, 5 μg/L biotin, 337 mg/L vitamin B1, 10 mM glutamate, 4 g/L glucose, and 7.6 mg/L Na_2MoO_4 [3]. Ampicillin (100 μg/mL), kanamycin (25 μg/mL), chloramphenicol (10 μg/mL), and isopropyl β-D-1-thiogalactopyranoside (IPTG; 0.1 mM) (Fujifilm Wako Pure Chemical, Osaka, Japan) were added to the medium as needed.

Construction of a plasmid library

Recombinant *E. coli* expressing 17 genes of *A. vinelandii* and nitrogenase activity was used as a host to evaluate the 210 genes presumed to be important for nitrogen fixation in *A. vinelandii* in our previous study (Takimoto, R. in revision). Of the 210 genes, 17 genes (*nifHDKBUSVQENXYWZMF* and *iscA*) that have already been expressed in *E. coli* [2] and 25 genes of rRNA and non-coding RNA were excluded. Then,

34 genes that were classified into a different cluster from the 210 genes using hierarchical clustering analysis (HCA) but might be transcribed continuously with the target gene under the same promoter were included, resulting in a total of 202 target genes in the plasmid library (Table 2).

Table 1. Strains and plasmids used in this study

Plasmids	Relative nitrogenase activity (%)	Gene	Function
pSTV-KC033	911	*zwf*	Glucose-6-phosphate dehydrogenase
		pgi	Glucose-6-phosphate isomerase
pSTV-KC085	664	*aceE*	Pyruvate dehydrogenase, E1 component
		aceF	Pyruvate dehydrogenase complex, E2 component
pSTV-KC014	590	*Avin_04140*	Glyoxalase/Bleomycin resistance/Dioxygenase superfamily protein
		gluP	Glucose/galactose transporter protein
pSTV-KC022	560	*fruB*	Fructose-specific multiphosphoryl transfer protein
		fruK	1-phosphofructokinase
		fruA	Fructose phosphotransferase system IIBC component

Table 2. List of genes included in plasmid library

Plasmid name	Gene name or locus tag	Locus tag	Function
pSTV-KC001	*Avin_00570*	Avin_00570	Hypothetical protein
pSTV-KC002	*Avin_01240*	Avin_01240	OprE-like outer membrane porin
pSTV-KC002	*Avin_01250*	Avin_01250	Hypothetical protein
pSTV-KC003	*Avin_01360*	Avin_01360	Conserved hypothetical protein
pSTV-KC003	*Avin_01370*	Avin_01370	Conserved hypothetical protein
pSTV-KC004	*nifT*	Avin_01410	Nitrogen fixation protein

pSTV-KC004	*nifY*	Avin_01420	Nitrogenase iron molybdenum cofactor biosynthesis protein
pSTV-KC004	*Avin_01430*	Avin_01430	Conserved hypothetical protein
pSTV-KC004	*lrv*	Avin_01440	Nitrogen fixing leucine rich variant repeat 4Fe-4S cluster protein
pSTV-KC005	*Avin_01490*	Avin_01490	Conserved hypothetical protein
pSTV-KC005	*Avin_01500*	Avin_01500	Conserved hypothetical protein
pSTV-KC006	*Avin_01510*	Avin_01510	Nitrogen fixation (4Fe-4S) ferredoxin-like protein
pSTV-KC006	*fesII*	Avin_01520	Nitrogen fixation (2Fe-2S) ferredoxin (Shethna protein), FeSII
pSTV-KC006	*Avin_01530*	Avin_01530	Conserved hypothetical protein
pSTV-KC006	*Avin_01540*	Avin_01540	Conserved hypothetical protein
pSTV-KC006	*Avin_01550*	Avin_01550	Conserved hypothetical protein
pSTV-KC006	*Avin_01560*	Avin_01560	Conserved hypothetical protein
pSTV-KC007	*Avin_01570*	Avin_01570	Conserved hypothetical protein
pSTV-KC007	*Avin_01580*	Avin_01580	ABC transporter, ATP-binding protein
pSTV-KC008	*cysE1nif*	Avin_01650	Nitrogen fixation serine O-acetyltransferase CysE1
pSTV-KC008	*Avin_01660*	Avin_01660	Conserved hypothetical protein
pSTV-KC009	*clpX*	Avin_01700	Nitrogen fixation protein orf9, ClpX
pSTV-KC010	*nafU*	Avin_01720	Hypothetical protein
pSTV-KC011	*Avin_02460*	Avin_02460	Conserved hypothetical protein
pSTV-KC011	*aglA*	Avin_02470	Alpha-glucosidase
pSTV-KC012	*vnfF*	Avin_02650	Vanadium nitrogenase ferredoxin, vnfF
pSTV-KC012	*vnfH*	Avin_02660	Vanadium nitrogenase iron protein
pSTV-KC013	*glk*	Avin_04130	Glucokinase
pSTV-KC014	*Avin_04140*	Avin_04140	Glyoxalase/Bleomycin resistance/Dioxygenase superfamily protein
pSTV-KC014	*gluP*	Avin_04150	Glucose/galactose transporter protein
pSTV-KC015	*relE*	Avin_04230	Transcription repressor protein
pSTV-KC015	*Avin_04240*	Avin_04240	Transcriptional regulator with Fis-type helix-turn-helix motif

pSTV-KC016	*cooJ*	Avin_04450	Nickel transporter for carbon monoxide+E51 dehydrogenase, HupE/UreJ family
pSTV-KC016	*cooT*	Avin_04460	Carbon monoxide dehydrogenase assessory protein
pSTV-KC016	*cooC*	Avin_04470	Carbon monoxide dehydrogenae assesory protein
pSTV-KC016	*Avin_04480*	Avin_04480	FAD-dependent pyridine nucleotide-disulphide
pSTV-KC016	*cooS*	Avin_04490	Carbon-monoxide dehydrogenase, catalytic subunit
pSTV-KC016	*cooF*	Avin_04500	Iron-sulfur cluster-binding protein CooF
pSTV-KC017	*Avin_05520*	Avin_05520	Hypothetical protein
pSTV-KC017	*metK*	Avin_05530	S-adenosylmethionine synthetase
pSTV-KC017	*Avin_05540*	Avin_05540	Bacterial regulatory protein, ArsR family
pSTV-KC018	*Avin_09350*	Avin_09350	TonB-dependent siderophore receptor
pSTV-KC019	*Avin_09770*	Avin_09770	Transposase IS3/IS911
pSTV-KC019	*Avin_09780*	Avin_09780	Transposase IS4 family
pSTV-KC019	*Avin_09790*	Avin_09790	Transposase
pSTV-KC020	*Avin_10380*	Avin_10380	Conserved hypothetical protein
pSTV-KC021	*Avin_11660*	Avin_11660	Hypothetical protein
pSTV-KC021	*lptF*	Avin_11670	Permease YjgP/YjgQ family
pSTV-KC021	*lptG*	Avin_11680	Permease YjgP/YjgQ family
pSTV-KC022	*fruB*	Avin_12190	Fructose-specific multiphosphoryl transfer protein
pSTV-KC022	*fruK*	Avin_12200	1-Phosphofructokinase
pSTV-KC022	*fruA*	Avin_12210	Fructose phosphotransferase system IIBC component
pSTV-KC023	*Avin_12220*	Avin_12220	Short-chain dehydrogenase/reductase SDR
pSTV-KC024	*Avin_12340*	Avin_12340	2-Oxoglutarate and Fe(II)-dependent dioxygenase superfamily
pSTV-KC025	*Avin_13130*	Avin_13130	Conserved hypothetical protein
pSTV-KC025	*Avin_13140*	Avin_13140	LppC lipoprotein
pSTV-KC026	*Avin_13200*	Avin_13200	UDP-N-acetylmuramyl-tripeptide synthetase

pSTV-KC026	*murF*	Avin_13210	UDP-N-acetylmuramoylalanyl-D-glutamyl-2, 6-diaminopimelate-D-alanyl-D-alanyl ligase
pSTV-KC027	*Avin_13880*	Avin_13880	Transcriptional regulatory protein, IclR family
pSTV-KC028	*Avin_14380*	Avin_14380	Hypothetical protein
pSTV-KC029	*Avin_15520*	Avin_15520	Heavy metal sensor histidine protein kinase, two-component
pSTV-KC029	*Avin_15530*	Avin_15530	Heavy metal response regulator, two-component
pSTV-KC030	*hexR-2*	Avin_15700	Transcriptional regulatory protein, RpiR family
pSTV-KC031	*mtnA*	Avin_15800	Initiation factor 2B
pSTV-KC032	*Avin_16160*	Avin_16160	Type II secretion system protein
pSTV-KC033	*zwf*	Avin_16620	Glucose-6-phosphate dehydrogenase
pSTV-KC033	*pgi*	Avin_16630	Glucose-6-phosphate isomerase
pSTV-KC034	*Avin_17160*	Avin_17160	Hypothetical conserved
pSTV-KC035	*Avin_17620*	Avin_17620	Enoyl-CoA hydratase/isomerase family protein
pSTV-KC036	*zwf*	Avin_17630	Glucose-6-phosphate dehydrogenase
pSTV-KC036	*pgl*	Avin_17640	6-Phosphogluconolactonase
pSTV-KC036	*pgi*	Avin_17650	Glucose-6-phosphate isomerase
pSTV-KC036	*Avin_17660*	Avin_17660	Fucose permease
pSTV-KC036	*Avin_17670*	Avin_17670	PfkB-family carbohydrate kinase
pSTV-KC036	*Avin_17680*	Avin_17680	Conserved hypothetical protein
pSTV-KC037	*Avin_18070*	Avin_18070	ATPase, AAA superfamily
pSTV-KC038	*Avin_19860*	Avin_19860	Hypothetical protein
pSTV-KC038	*Avin_19870*	Avin_19870	Hypothetical transmembrane protein
pSTV-KC038	*cydB*	Avin_19880	Cytochrome bd ubiquinol oxidase, subunit II
pSTV-KC038	*cydA*	Avin_19890	Cytochrome bd ubiquinol oxidase, subunit I
pSTV-KC039	*Avin_22040*	Avin_22040	NAD-dependent aldehyde dehydrogenase
pSTV-KC040	*pykA*	Avin_22190	Pyruvate kinase
pSTV-KC040	*Avin_22200*	Avin_22200	Hypothetical protein
pSTV-KC040	*eno*	Avin_22210	Enolase

pSTV-KC041	*Avin_22570*	Avin_22570	Hypothetical protein
pSTV-KC042	*Avin_24970*	Avin_24970	Sulphate transporter-SulP-type
pSTV-KC043	*Avin_24980*	Avin_24980	RHH-domain repressor protein
pSTV-KC044	*Avin_25120*	Avin_25120	Conserved hypothetical protein
pSTV-KC045	*Avin_25690*	Avin_25690	Tn4652, cointegrate resolution protein T
pSTV-KC046	*Avin_25710*	Avin_25710	Cointegrate resolution protein S
pSTV-KC047	*Avin_25970*	Avin_25970	Major facilitator superfamily protein
pSTV-KC048	*Avin_26080*	Avin_26080	Hypothetical protein
pSTV-KC048	*Avin_26090*	Avin_26090	ABC transporter, ATP-binding protein
pSTV-KC048	*Avin_26100*	Avin_26100	ABC-transporter ATP-binding protein
pSTV-KC048	*Avin_26110*	Avin_26110	Bacterial ABC-transporter, permease protein
pSTV-KC048	*Avin_26120*	Avin_26120	Bacterial ABC-transporter, permease prtein
pSTV-KC048	*Avin_26140*	Avin_26140	ABC trasporter substrate binding protein
pSTV-KC049	*Avin_26330*	Avin_26330	Hypothetical protein
pSTV-KC050	*Avin_26980*	Avin_26980	Glucan 1,4-alpha-glucosidase
pSTV-KC050	*Avin_26990*	Avin_26990	Hypothetical protein
pSTV-KC051	*Avin_27190*	Avin_27190	DNA topoisomerase III
pSTV-KC052	*Avin_27200*	Avin_27200	LamB type porin
pSTV-KC053	*Avin_27210*	Avin_27210	Glyceraldehyde-3-phosphate dehydrogenase
pSTV-KC053	*eno*	Avin_27220	Enolase
pSTV-KC053	*Avin_27230*	Avin_27230	Hypothetical protein
pSTV-KC053	*pykA*	Avin_27240	Pyruvate kinase
pSTV-KC053	*eda*	Avin_27250	KDPG aldolase
pSTV-KC054	*zwf*	Avin_27260	Glucose-6-phosphate dehydrogenase
pSTV-KC055	*hexR*	Avin_27270	Transcriptional regulatory protein, RpiR family
pSTV-KC056	*edd*	Avin_27280	6-Phosphogluconate dehydratase
pSTV-KC057	*Avin_27300*	Avin_27300	Transcriptional regulator, LacI family
pSTV-KC058	*pgm*	Avin_27440	2,3-Bisphosphoglycerate-independent phosphoglycerate mutase
pSTV-KC059	*Avin_27460*	Avin_27460	Conserved hypothetical protein
pSTV-KC060	*Avin_27500*	Avin_27500	Conserved hypothetical protein

pSTV-KC061	*mnmA*	Avin_28360	TRNA (5-methylaminomethyl-2-thiouridylate)-methyltransferase
pSTV-KC061	*Avin_28370*	Avin_28370	Conserved hypothetical protein
pSTV-KC062	*Avin_28680*	Avin_28680	Hypothetical protein
pSTV-KC063	*Avin_29020*	Avin_29020	Phosphohistidine phosphatase SixA
pSTV-KC063	*gpsA*	Avin_29030	Glycerol-3-phosphate dehydrogenase (NAD(P)-dependent)
pSTV-KC064	*Avin_29220*	Avin_29220	N-acetylmuramoyl-L-alanine amidase
pSTV-KC065	*Avin_29610*	Avin_29610	Conserved hypothetical protein
pSTV-KC066	*lapL*	Avin_30750	Multi-component phenol hydoxylase, beta subunit; LapL
pSTV-KC066	*lapK*	Avin_30760	Multi-component phenol hydoxylase, assembly subunit; LapK
pSTV-KC067	*lapR*	Avin_30810	Sigma54-dependent activator protein
pSTV-KC068	*Avin_32320*	Avin_32320	Transposase of insertion sequence ISRm10-1, orfA protein
pSTV-KC069	*Avin_32400*	Avin_32400	Conserved hypothetical protein
pSTV-KC069	*Avin_32410*	Avin_32410	Conserved hypothetical protein
pSTV-KC070	*pcpS*	Avin_33280	4-Phosphopantetheinyl transferase protein
pSTV-KC071	*Avin_33570*	Avin_33570	Hypothetical protein
pSTV-KC072	*Avin_35600*	Avin_35600	Transposase
pSTV-KC073	*Avin_36020*	Avin_36020	Transcriptional regulator Cro/CI-like protein
pSTV-KC074	*Avin_36050*	Avin_36050	C-5 cytosine-specific DNA methylase
pSTV-KC075	*Avin_36060*	Avin_36060	Hypothetical protein
pSTV-KC076	*Avin_36070*	Avin_36070	Hypothetical protein
pSTV-KC076	*Avin_36080*	Avin_36080	ATP binding/ATPase (HATPase_c) domain-containing protein
pSTV-KC076	*Avin_36090*	Avin_36090	Conserved hypothetical protein
pSTV-KC077	*Avin_36910*	Avin_36910	Conserved hypothetical protein
pSTV-KC078	*Avin_37940*	Avin_37940	Hypothetical protein
pSTV-KC079	*Avin_37950*	Avin_37950	Zinc-containing alcohol dehydrogenase
pSTV-KC079	*Avin_37960*	Avin_37960	O-methyltransferase
pSTV-KC079	*Avin_37970*	Avin_37970	Dienelactone hydrolase-like protein
pSTV-KC079	*Avin_37980*	Avin_37980	Hypothetical protein

pSTV-KC080	*Avin_38660*	Avin_38660	Bacterial chemotaxis sensory transducer protein
pSTV-KC080	*aerP*	Avin_38670	Aerotaxis sesnor; AerP
pSTV-KC081	*Avin_39240*	Avin_39240	Conserved hypothetical protein
pSTV-KC081	*Avin_39250*	Avin_39250	Conserved hypothetical protein
pSTV-KC081	*Avin_39260*	Avin_39260	Conserved hypothetical protein
pSTV-KC081	*yeaZ*	Avin_39270	Peptidase M22, glycoprotease
pSTV-KC081	*adk*	Avin_39280	Adenylate kinase
pSTV-KC081	*Avin_39290*	Avin_39290	Hypothetical protein
pSTV-KC081	*ppc*	Avin_39300	Phosphoenolpyruvate carboxylase
pSTV-KC082	*Avin_43180*	Avin_43180	Hypothetical protein
pSTV-KC082	*Avin_43190*	Avin_43190	Hypothetical protein
pSTV-KC083	*Avin_43220*	Avin_43220	Hypothetical protein
pSTV-KC084	*eno*	Avin_43350	Enolase
pSTV-KC084	*Avin_43360*	Avin_43360	Hypothetical protein
pSTV-KC084	*pykA*	Avin_43370	Pyruvate kinase
pSTV-KC085	*aceE*	Avin_44900	Pyruvate dehydrogenase, E1 component
pSTV-KC085	*aceF*	Avin_44910	Pyruvate dehydrogenase complex, E2 component
pSTV-KC086	*Avin_44930*	Avin_44930	Usp-like protein
pSTV-KC087	*relE*	Avin_44980	Cytotoxin, RelE protein
pSTV-KC088	*Avin_45750*	Avin_45750	Hypothetical protein
pSTV-KC088	*eno*	Avin_45760	Enolase
pSTV-KC088	*Avin_45780*	Avin_45780	Aldehyde dehydrogenase
pSTV-KC089	*Avin_45970*	Avin_45970	Hypothetical protein
pSTV-KC090	*speD*	Avin_46080	S-adenosylmethionine decarboxylase
pSTV-KC090	*Avin_46090*	Avin_46090	OsmC family protein
pSTV-KC091	*Avin_48680*	Avin_48680	McbC-like_oxidoreductase
pSTV-KC091	*Avin_48690*	Avin_48690	Rhodanese-like protein with Ankyrin repeat
pSTV-KC091	*Avin_48700*	Avin_48700	Metallocluster binding protein of the NafY/NifY/NifX/VnfX family
pSTV-KC091	*Avin_48710*	Avin_48710	Conserved hypothetical protein
pSTV-KC091	*Avin_48720*	Avin_48720	Glutathione S-transferase
pSTV-KC092	*tctA*	Avin_49480	Tripartite tricarboxylate transporter large transmembrane protein

pSTV-KC093	*Avin_49950*	Avin_49950	Aldehyde dehydrogenase
pSTV-KC093	*eno*	Avin_49970	Enolase
pSTV-KC093	*Avin_49980*	Avin_49980	Hypothetical protein
pSTV-KC093	*pykA*	Avin_49990	Pyruvate kinase
pSTV-KC094	*Avin_50410*	Avin_50410	Hypothetical membrane protein
pSTV-KC095	*Avin_50870*	Avin_50870	Rhs element Vgr protein
pSTV-KC095	*Avin_50880*	Avin_50880	Conserved hypothetical protein
pSTV-KC096	*Avin_50890*	Avin_50890	Conserved hypothetical protein
pSTV-KC097	*Avin_50900*	Avin_50900	Nitrogen fixation-related protein
pSTV-KC097	*nafY*	Avin_50910	Nitrogen fixation-related protein, gamma subunit
pSTV-KC097	*rnfH*	Avin_50920	Electron transport complex, RnfABCDGE type, H subunit
pSTV-KC097	*rnfE*	Avin_50930	Electron transport complex, RnfABCDGE type, E subunit
pSTV-KC097	*rnfG*	Avin_50940	Electron transport complex, RnfABCDGE type, G subunit
pSTV-KC097	*rnfD*	Avin_50950	Electron transport complex, RnfABCDGE type, D subunit
pSTV-KC097	*rnfC*	Avin_50960	Electron transport complex, RnfABCDGE type, C subunit
pSTV-KC097	*rnfB*	Avin_50970	Electron transport complex, RnfABCDGE type, B subunit
pSTV-KC097	*rnfA*	Avin_50980	Electron transport complex, RnfABCDGEFH, A subunit
pSTV-KC098	*fdx*	Avin_51020	Ferredoxin protein
pSTV-KC098	*nifO*	Avin_51030	Nitrogenase-associated protein
pSTV-KC099	*pykA*	Avin_51360	Pyruvate kinase
pSTV-KC100	*Avin_51550*	Avin_51550	Multidrug efflux pump membrane fusion protein
pSTV-KC100	*Avin_51560*	Avin_51560	Multidrug efflux pump RND-family transporter protein
pSTV-KC100	*Avin_51570*	Avin_51570	Multidrug efflux pump RND-family outer membrane protein
pSTV-KC101	*Avin_51600*	Avin_51600	3-Oxoacid CoA-transferase

pSTV-KC101	*Avin_51610*	Avin_51610	3-Oxoacid CoA-transferase
pSTV-KC102	*Avin_51780*	Avin_51780	Alpha-glucosidase
pSTV-KC103	*scrY*	Avin_51790	Sucrose porin
pSTV-KC104	*scrT*	Avin_51800	Proton/sucrose symporter
pSTV-KC104	*scrB*	Avin_51810	Sucrose or/and sucrose-6-phosphate hydrolase

The open reading frame of the target gene fused with a 15 bp sequence homologous to the pSTV29 cloning site at both ends less than 500 bp in length was artificially synthesized using the codon optimization for *E. coli* by Integrated DNA Technologies (IA, USA). The other genes were amplified using polymerase chain reaction (PCR) template with KOD-One polymerase (TOYOBO, Osaka, Japan) and primers with 15 bp homologous sequences, with the *A. vinelandii* genomic DNA extracted by ISOGEN (NIPPON GENE, Tokyo, Japan) as a template. Multiple genes, which are transcribed as a single mRNA under the same promoter, were designed to be inserted together into a single plasmid. The artificially synthesized genes and PCR products were fused with the linearized pSTV29 prepared using PCR with the In-Fusion HD Cloning Kit (Takara Bio, Shiga, Japan). The resulting plasmids were sequenced and named as pSTV-KC001 to pSTV-KC104 (Table 1 and 2).

Acetylene reduction activity assay

The C_2H_2 reduction activity assay was performed as described previously [2]. Vials with 21 mL headspace were filled with 4.5 mL of nitrogen-deficient media supplemented with antibiotics and IPTG, sealed with a butyl rubber septum and crimp cap, and refilled with argon gas several times to achieve anaerobic conditions. For the activity assay under microaerobic conditions, 3.5 mL of gas was syringe-drained from the headspace and the same volume of oxygen gas was added. The cell pellet of recombinant *E. coli* after aerobic pre-culture in LB broth at 37°C overnight was suspended in the nitrogen-deficient media to reach an OD_{600} of 5, and then 0.5 mL of the cell suspension was added to the headspace vial using a syringe (final OD_{600} of 0.5). Acetylene gas was prepared in a sagittal gas generator using calcium carbide (Sigma-Aldrich, MO, USA) and water in a vacuum. As an internal standard, 1 mL of ethane gas (GL Sciences, Tokyo, Japan) was mixed with 200 mL of acetylene gas in a sample bag, and 1 mL of the gas mixture was injected into the vial containing the cells using a gas-tight syringe. After incubation at 30°C for 16 h, the reaction was stopped by injecting 300 µL of 4 M NaOH into each vial. The produced ethylene (m/z: 26) and the internal standard ethane (m/z: 30)

were detected using the GCMS QP2020Ultra (Shimadzu, Kyoto, Japan) connected to the TurboMatrix HS110 headspace autosampler (PerkinElmer, MA, USA) with the injector temperature, column oven temperature, and detector temperature set at 40°C, 40°C, and 250°C, respectively, and a GS GasPro column (0.32 mm, 30 m) (Agilent, CA, USA). Ethylene concentration was determined using the external standard method.

Measurement of intracellular metabolites

E. coli cells cultured for 6 h under microaerophilic conditions were collected via centrifugation at 800 × *g* for 5 min at 4°C. The pellets were washed multiple times with phosphate buffered saline (PBS), frozen in liquid nitrogen, and stored at −80°C until analysis. Quantification of nicotinamide adenine dinucleotide (NAD^+) and NADH in the cells was performed using the NAD/NADH Assay Kit-WST (Dojindo Laboratories, Kumamoto, Japan). For the measurement of intracellular metabolites, unit cells (OD_{600} = 8) were suspended in 1.5 mL of 0.2 µg ribitol/70% methanol solution and frozen at −80°C overnight. The suspension was repeatedly thawed 25°C for 1 h, frozen at −80°C for 1 h, and then vigorously stirred for 1 min. After centrifugation at 10,000 × *g* for 5 min at 4°C, the supernatant was filtered through Amicon Ultra 3k filter (Merck Group, Darmstadt, Germany). Then, the filtrate was vacuum-dried and resolved in 200 µL of 50% (v/v) acetonitrile. The sample (5 µL) was analyzed on an LC-MS8050 (Shimadzu) equipped with a BEH amide column (2.1 x 150 mm, 17 µm; Waters, MA, USA). The relative levels of metabolites normalized by ribitol were subjected to principal component analysis (PCA) using Jmp16 (SAS Institute, NC, USA).

Results

Screening of genes from *A. vinelandii* that enhance nitrogenase activity

To screen for *A. vinelandii* genes that enhance nitrogenase activity in the *E. coli* strain expressing 17 *nif* genes constructed in our previous study [2], I first prepared a plasmid library using pSTV29 as a vector. For an efficient screening, I hypothesized that genes showing the same transcriptional pattern as *nifHDK* from a transcriptome analysis of *A. vinelandii* (Takimoto, R. in revision) would be important for nitrogen fixation, and some may enhance the nitrogenase activity of *E. coli* expressing 17 *nif* genes. Based on this, a total of 202 genes were selected, including 21 genes involved in nitrogen fixation, 35 genes involved in sugar transport and metabolism, 3 genes involved in respiration, 6 genes related to transcription, 137 genes with other functions, and other genes with

unknown functions (Table 2). Several genes that may be transcribed and translated consecutively, such as operons, were inserted into a single expression plasmid. As a result, I constructed a library consisting of 104 plasmids, each of which contained single or multiple genes under the control of the lac promoter in pSTV29 and whose expression was regulated by IPTG. The plasmid library or empty plasmid (pSTV29) was introduced into *E. coli* JM109 harboring plasmids (pTrc-nif001 and pMW-nif002) for the expression of 17 *nif* genes. The nitrogenase activities of the transformants were measured under anaerobic and microaerobic conditions.

Using the *E. coli* strain harboring pTrc-nif001, pMW-nif002, and pSTV29 as a reference, a scatter plot was generated based on the relative nitrogenase activity of the strains harboring the plasmid library under anaerobic and microaerobic conditions (Fig. 1).

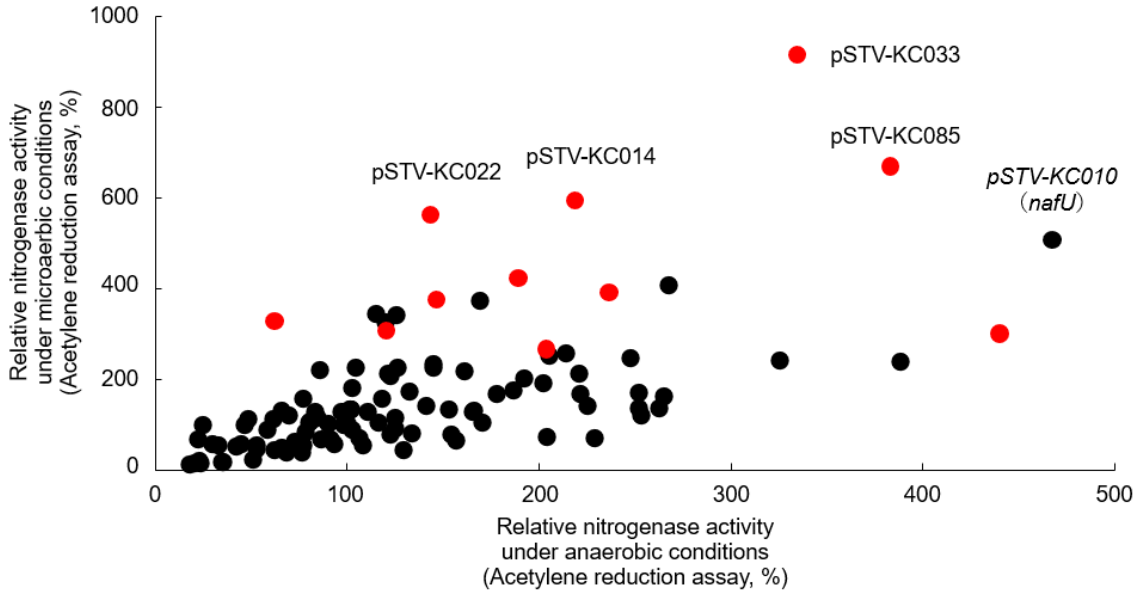

Figure 1. Nitrogenase activity was measured under anaerobic and microaerobic conditions and the relative activity of pTrc-nif001, pMW-nif002, and pSTV29 expressing strains are plotted. Values are expressed as the mean (n=3).

The highest relative nitrogenase activity under anaerobic conditions and the fifth highest under microaerobic conditions was observed in the strain harboring the pSTV-KC010 plasmid containing *nafU*. Although it is unclear how *nafU* is involved in nitrogen fixation, its involvement in the maintenance of nitrogenase activity of *A. vinelandii* under

aerobic conditions has recently been proposed (Takimoto, R. in revision). In contrast, nitrogenase activity was higher under microaerobic conditions in the strains harboring pSTV-KC033, pSTV-KC085, pSTV-KC014, and pSTV-KC022 (Table 3).

Table 3 Function of genes contained in plasmids with high nitrogenase activity under microaerophilic conditions

Plasmids	Relative nitrogenase activity (%)	Gene	Function
pSTV-KC033	911	*zwf*	Glucose-6-phosphate dehydrogenase
		pgi	Glucose-6-phosphate isomerase
pSTV-KC085	664	*aceE*	Pyruvate dehydrogenase, E1 component
		aceF	Pyruvate dehydrogenase complex, E2 component
pSTV-KC014	590	*Avin_04140*	Glyoxalase/Bleomycin resistance/Dioxygenase superfamily protein
		gluP	Glucose/galactose transporter protein
pSTV-KC022	560	*fruB*	Fructose-specific multiphosphoryl transfer protein
		fruK	1-phosphofructokinase
		fruA	Fructose phosphotransferase system IIBC component

The pSTV-KC033 contained *zwf* and *pgi*, which encode for glucose-6-phosphate dehydrogenase and glucose-6-phosphate isomerase, respectively, while pSTV-KC033 contained *aceEF*, which encodes for pyruvate dehydrogenase E1 and E2 components. The pSTV-KC014 contained *gluP*, which encodes a glucose/galactose transporter protein, and *Avin_0410*, which has an unknown function but contains domains of the glyoxalase/bleomycin resistance/dioxygenase superfamily. The pSTV-KC022 contained *fruB* (fructose-specific multiphosphoryl transfer protein), *fruK* (1-phosphofructokinase), and *fruA* (fructose phosphotransferase system IIBC component). These results suggest that additional expression of the genes involved in sugar metabolism was effective in

enhancing nitrogenase activity under microaerobic conditions in *E. coli* expressing 17 *nif* genes.

NAD^+/NADH levels in *E. coli* showing an improved nitrogenase activity under microaerophilic conditions

To understand the mechanisms underlying the expression of genes related to glycolysis that improved nitrogenase activity, intracellular metabolites of the transformants were analyzed. Among the various intracellular metabolites, I first focused on NAD. NAD^+ is an essential cofactor in many major oxidative reactions in living organisms and plays a central role in glycolysis [4]. In contrast, NADH is used for ATP production and various reduction reactions. During nitrogen fixation of *A. vinelandii*, electrons from NADH are transferred to nitrogenase via ferredoxin, and nitrogen is finally reduced to ammonia [5]. Thus, I hypothesized that intracellular NAD^+ and NADH concentrations, which have a strong influence on nitrogen fixation reactions and glycolytic systems, were altered when the genes related to glucose metabolism were expressed in the strain expressing 17 *nif* genes. Total NAD (NAD^+ and NADH) and NADH concentrations in the lysates of the strains harboring pSTV-KC033, pSTV-KC085, pSTV-KC014, pSTV-KC022, or control pSTV29 grown under microaerophilic conditions in nitrogen-deficient medium were measured using enzymatic methods, and NAD^+ concentration was determined by subtracting the NADH concentration from NAD concentration. In addition, NAD^+/NADH is used as an indicator of the oxidation/reduction state in cells. The NAD^+/NADH levels of the strains showing improved nitrogenase activity were significantly lower than those of the control strain (Fig. 2). These results indicated that the expression of genes related to glucose metabolism promoted the shift to a reducing state in the cells. Reducing power is a driving force for the nitrogen fixation reaction, suggesting that the transition of the cells to a reducing state is a part of the mechanism underlying the enhanced nitrogenase activity.

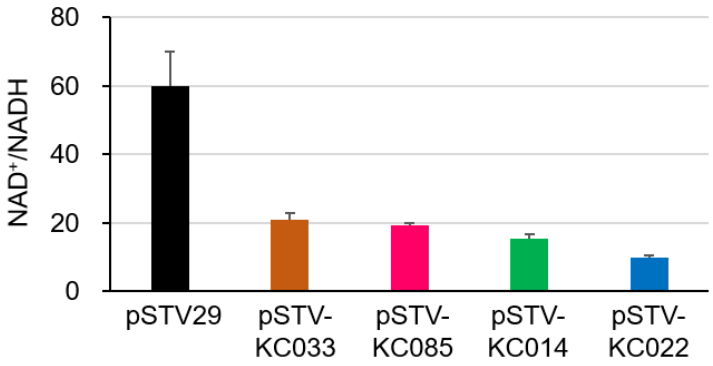

Figure 2. Intracellular (a) NAD^+, (b) NADH, and (c) NAD^+ / NADH of recombinant *E. coli* that showed high nitrogenase activity under microaerophilic conditions was measured. Values are expressed as the mean ± SD (n=3).

Analysis of a wide range of intracellular metabolites

To further understand the mechanism underlying the enhanced nitrogenase activity in the cells, in which cells are transferred to a reducing state by transformation of genes related to glucose metabolism, I analyzed a wide range of hydrophilic metabolites using LC-MS/MS. The amounts of 64 components, including sugar phosphates, amine acids, and bases, were measured relative to the control strain. First, PCA was performed using the total metabolite data to characterize the control strain and the strain showing an improved nitrogenase activity (Fig. 3). The first principal component (PC1) and the second principal component (PC2) accounted for 34.1% and 18.0%, respectively, indicating that the two principal components explained 52.1% of the total variance. The plots of PC1 and PC2 showed that the cellular metabolite changes in each *E. coli* strain could be classified into three: pSTV29 (group 1), pSTV-KC033 (group 2); and pSTV-KC085, pSTV-KC014, pSTV-KC022 (group 3). Group 1 was separated from groups 2 and 3 by PC1, showing the contribution of cAMP and cGMP. Specifically, the relative values of cAMP and cGMP in groups 2 and 3 were lower than those in group 1 (Fig. 4). On the other hand, group 2 was separated from groups 1 and 3 by PC2, showing the contribution of glucose 6-phosphate and other factors. The relative value of glucose 6-phosphate in group 2 was higher than that in the other strains. These results indicate that the strains showing improved nitrogenase activity through the enhanced reducing power commonly had relatively low levels of cAMP and cGMP, and the strain showing the highest nitrogenase activity (pSTV-033) had the highest level of glucose 6-phosphate.

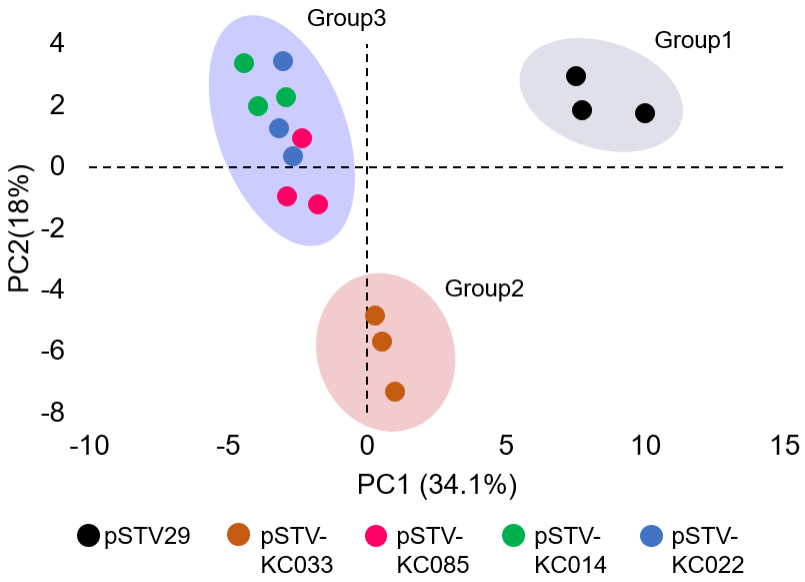

Figure 3. PCA analysis using the intracellular component data of recombinant *E. coli* showing high nitrogenase activity under microaerophilic conditions. The contribution rate of PC1 and PC2 was 34.1% and 18.0%, respectively.

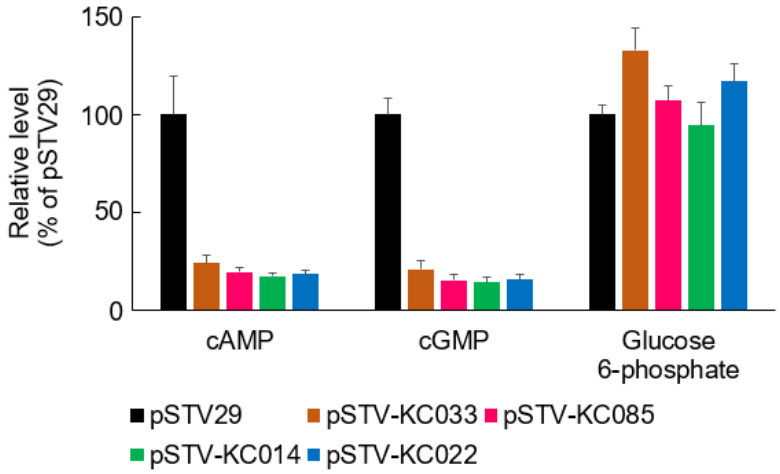

Figure 4. Relative levels of cAMP, cGMP, and glucose 6-phosphate in recombinant *E. coli.* The expression strains pTrc-nif001, pMW-nif002, and pSTV29 are shown. Values are expressed as the mean ± SD (n=3).

Discussion

Nitrogen fixing-related genes in *A. vinelandii* have been used for heterologous expression of nitrogenase in other microorganisms [3, 6]. However, heterologous

expression have been limited to the genes included in large and small *nif* clusters (e.g., *nif* and *naf*) or genes encoding nitrogenase isotypes (e.g., *anf* and *vnf*). Some genes (e.g., *cydBD* and *alg*AC), which are required for efficient nitrogen fixation under aerobic conditions in *A. vinelandii* [7] [8], are located far from the large/small clusters of *nif* [9]. To our knowledge, this is the first study to express a wide range of genes from *A. vinelandii*, including *nif*, *naf*, *anf*, and *vnf*, to enhance the activity of recombinant nitrogenase. I observed that the genes increasing the nitrogenase activity in recombinant *E. coli* are involved in glucose metabolism, suggesting the importance of an improved glucose metabolism in nitrogenase heterologous expression strains.

In our study, I focused on a total of 202 genes based on the transcription pattern of *nifHDK* and constructed a plasmid library for the expression of these genes. Among the genes expressed in *E. coli* expressing 17 *nif* genes, genes related to glucose metabolism, such as *zwf* and *pgi*, had the highest positive effect on nitrogenase activity under microaerobic conditions. Our findings may contribute to an effective strategy for improving the activity of recombinant nitrogenases.

In the measurement of intracellular metabolites, the NAD^{+}/NADH ratio of the control strain was 60.0, which is higher than those reported previously (0.1-8) [10, 11]. Although this study used different *E. coli* strains and culture conditions from the previous study, the most significant difference was that the bacteria were cultured in nitrogen-deficient medium in this study. Under nitrogen-deficient conditions, intracellular amino acids and glutamine are drastically reduced, while α-ketoglutarate (αKG) is increased [12, 13]. αKG is involved in the transcriptional regulation of nitrogen metabolism genes and inhibits PtsI, which plays an important role in glucose uptake [14]. Therefore, the significantly higher NAD^{+}/NADH ratio compared to previous reports could be attributed to the reduction in intracellular glucose concentration caused by the suppression of glucose uptake and glucose metabolism.

The high NAD^{+}/NADH ratio in the control strain was reduced by 10.0–20.8 by expression of glucose metabolism-related genes (Fig. 2), suggesting that glucose metabolism may be partially improved. Interestingly, cyclic nucleotide levels, such as cAMP and cGMP, were reduced in the transformants expressing sugar metabolism-related genes (Fig. 4). These cyclic nucleotides act as secondary messengers and are involved in the transcription of glucose metabolism-related genes in collaboration with cAMP receptor protein (CRP) and others [15]. These cyclic nucleotides decreased inversely to the increase in intracellular glucose content. Therefore, pSTV-KC033, pSTV-KC085, pSTV-KC014, and pSTV-KC022 strains were thought to have increased intracellular glucose and its metabolites, which ultimately led to increased NADH and high

nitrogenase activity (Fig. 5). Nitrogenase activity has been reported to decrease when ammonia or glutamine is added to the medium to alleviate nitrogen starvation [16]. In our study, the activity was also significantly decreased (data not shown). Therefore, the control of sugar metabolism is an important issue for improving the nitrogenase activity of *E. coli* expressing 17 *nif* genes in nitrogen-deficient media. A similar control could be effective for nitrogen fixation by nitrogenase produced in yeast and plant cells.

In conclusion, I have improved nitrogenase production in *E. coli* by enhancing intracellular reducing power through the expression of sugar metabolism-related genes from *A. vinelandii*. This study provides a new perspective for nitrogen fixation research using recombinant *E. coli*. Moreover, this study may contributed in nitrogen fixation research using other organisms.

Summary

In this chapter, I screened for genes that enhance nitrogenase activity from the 202 genes previously reported using transcriptome analysis of *A. vinelandii*, and selected several genes related to sugar metabolism. NADH levels were increased in the activity-enhanced strains, indicating that heterologous expression of nitrogenase is important for the efficient supply of reducing power. These results may be an important factor in the heterologous expression of nitrogenase for an efficient nitrogen fixation.

References

1. Peña C, Campos N, Galindo E. Changes in alginate molecular mass distributions, broth viscosity and morphology of *Azotobacter vinelandii* cultured in shake flasks. Appl Microbiol Biotechnol 1997; 48:510-515.
2. Tatemichi Y, Nakahara T, Ueda M, Kuroda K. Construction of recombinant *Escherichia coli* producing nitrogenase-related proteins from *Azotobacter vinelandii*. Biosci Biotechnol Biochem 2021; 85:2209-2216.
3. Yang J, Xie X, Wang X, Dixon R, Wang YP. Reconstruction and minimal gene requirements for the alternative iron-only nitrogenase in *Escherichia coli*. Proc Natl Acad Sci USA 2014; 111:E3718-3725.
4. Katoh A, Hashimoto T. Molecular biology of pyridine nucleotide and nicotine biosynthesis. Front Biosci 2004; 9:1577-1586.
5. Ledbetter RN, Garcia Costas AM, Lubner CE, Mulder DW, Tokmina-

Lukaszewska M, Artz JH, Patterson A, Magnuson TS, Jay ZJ, Duan HD *et al*. The electron bifurcating FixABCX protein complex from *Azotobacter vinelandii*: generation of low-potential reducing equivalents for nitrogenase catalysis. Biochemistry 2017; 56:4177-4190.

6. Ryu MH, Zhang J, Toth T, Khokhani D, Geddes BA, Mus F, Garcia-Costas A, Peters JW, Poole PS, Ane JM *et al*. Control of nitrogen fixation in bacteria that associate with cereals. Nat Microbiol 2020; 5:314-330.
7. Kelly MJ, Poole RK, Yates MG, Kennedy C. Cloning and mutagenesis of genes encoding the cytochrome bd terminal oxidase complex in *Azotobacter vinelandii*: mutants deficient in the cytochrome d complex are unable to fix nitrogen in air. J Bacteriol 1990; 172:6010-6019.
8. Galindo E, Peña C, Núñez C, Segura D, Espín G. Molecular and bioengineering strategies to improve alginate and polydydroxyalkanoate production by *Azotobacter vinelandii*. Microb Cell Fact 2007; 6:7.
9. Mus F, Alleman AB, Pence N, Seefeldt LC, Peters JW. Exploring the alternatives of biological nitrogen fixation. Metallomics 2018; 10:523-538.
10. Zhou YJ, Yang W, Wang L, Zhu Z, Zhang S, Zhao ZK. Engineering NAD^+ availability for *Escherichia coli* whole-cell biocatalysis: a case study for dihydroxyacetone production. Microb Cell Fact 2013; 12:103-103.
11. Han Q, Eiteman MA. Coupling xylitol dehydrogenase with NADH oxidase improves l-xylulose production in *Escherichia coli* culture. Enzyme Microb Technol 2017; 106:106-113.
12. Ninfa AJ, Jiang P. PII signal transduction proteins: sensors of alpha-ketoglutarate that regulate nitrogen metabolism. Curr Opin Microbiol 2005; 8:168-173.
13. Chubukov V, Desmarais JJ, Wang G, Chan LJG, Baidoo EE, Petzold CJ, Keasling JD, Mukhopadhyay A. Engineering glucose metabolism of *Escherichia coli* under nitrogen starvation. NPJ Syst Biol Appl 2017; 3:16035.
14. Doucette CD, Schwab DJ, Wingreen NS, Rabinowitz JD. α-Ketoglutarate coordinates carbon and nitrogen utilization via enzyme I inhibition. Nat Chem Biol 2011; 7:894-901.
15. Botsford JL, Harman JG. Cyclic AMP in prokaryotes. Microbiol Rev 1992; 56:100-122.
16. Wang X, Yang JG, Chen L, Wang JL, Cheng Q, Dixon R, Wang YP. Using synthetic biology to distinguish and overcome regulatory and functional barriers related to nitrogen fixation. PLOS ONE 2013; 8:e68677.

Conclusion

The aim of this study was to modify *E.coli* for sustainable ammonia production. There were two main strategies for ammonia production using *E.coli*. The first strategy is catabolism of amino acids in food by-products by metabolic reactions in E. coli, and the second is biological nitrogen fixation by nitrogenase. However, both methods has its own productivity issues. In the former strategy, I improved productivity by identifying inhibitors of ammonia production in biomass and constructing the *E. coli* strain with a gene disruption. In the latter strategy, I constructed the nitrogen-fixing *E. coli* strain by expressing the *nif* genes from *A. vinelandii* and identified the genes whose expressions improve nitrogenase activity under microaerobic conditions.

In Chapter I, I searched for inhibitors of ammonia production present in pre-processed food by-products. Regression analysis using the amount of hydrophilic components in the food by-products and their ammonia production showed that glucose in the food by-products inhibited ammonia production by *E. coli*. Based on this finding, the double disruption of the glucose transporter (PtsG) and glutaminase (GlnA) genes improved the efficiency of ammonia production from food byproduct (okara) pretreatment.

In Chapter II, an artificial gene cluster consisting of *A. vinelandii* nitrogenase-related genes (*nifHDKBUSVQENXYWZMF* and *iscA*) was constructed and introduced into *E. coli* for expression. In the analysis of the transcription and protein levels of *nif* genes using RT-qPCR and LC-MS, the transcription and expression levels of each gene were compared between the constructed *E. coli* and wild-type *A. vinelandii*. The constructed *E. coli* produced 4.96 nmol/mL of ethylene under anaerobic conditions and 4.13 nmol/mL under microaerobic conditions.

In Chapter III, to improve the nitrogenase activity of *E. coli* expressing nitrogenase, I attempted to identify genes contributing to the activity from *A. vinelandii*. Based on the results of RNA-seq analysis, 213 genes were selected as a candidate gene. A plasmid library was constructed by cloning of these genes, and additionally introduced into nitrogenase-expressing *E. coli* to evaluate the effect on nitrogenase activity. I have identified several genes improving nitrogenase activity by their overexpression, such as genes related to glucose metabolism. In the constructed *E. coli* with enhanced nitrogenase activity by the gene overexpression, the amount of NADH was greatly increased, suggesting that it would contribute to the enhanced activity by increasing the intracellular reducing power.

In this study, I successfully improved the efficiency of ammonia production from food byproducts by *E. coli*, reconstructed nitrogenase in *E. coli* by heterologously expressing *nif*-related genes from *A.vinelandii*, and identified the genes that improve the nitrogenase activity in the nitrogenase-expressing *E. coli*. Further development of this technology is expected to lead to the establishment of a sustainable ammonia production process.

Acknowledgements

This doctoral thesis was submitted by the author to Kyoto University for the Doctoral Degree of Agriculture. The studies presented here were carried out under the direction of professor Mitsuyoshi Ueda and associate professor Kouichi Kuroda at the Laboratory of Biomacromolecular Chemistry. Division of Applied Life Sciences, Graduate School of Agriculture, Kyoto University, during 2019-2022.

I would also like to express my gratitude to professor Tatsuo Kurihara, assistant professor Wataru Aoki, Dr. Keiko Gomi, Dr. Takeharu Nakahara, Dr. Tomohiro Kawaguchi, and Dr. Miho Imamura (Touhara) for their critical and insightful comments. Without their encouragement, I would not have been able to accomplish this research. I would also like to express my gratitude to my secretary, Fukuko Suzuki, and other colleagues in the laboratory, as they have supported me countless times in my research life.

Finally, I would like to express my greatest appreciation to my wife and children, my parents and siblings for their warm support and encouragement.

Yuki TATEMICHI

Laboratory of Biomacromolecular Chemistry
Division of Applied Life Sciences
Graduate School of Agriculture
Kyoto University

Publications

Chapter I

Tatemichi Y, Kuroda K, Nakahara T and Ueda M.

Efficient ammonia production from food by-products by engineered *Escherichia coli*. *AMB Express* 2020;10:150.

Chapter II

Tatemichi Y, Nakahara T, Ueda M and Kuroda K. Construction of recombinant *Escherichia coli* producing nitrogenase-related proteins from *Azotobacter vinelandii*. *Biosci Biotechnol Biochem* 2021;85:2209-16.

Chapter III

Tatemichi Y, Nakahara T, Ueda M and Kuroda K.

Identification of genes from *Azotobacter vinelandii* for improving nitrogenase activity in *Escherichia coli* expressing nif genes.

In preparation

Other publication

Takimoto R, Tatemichi Y, Aoki W, Kosaka Y, Minakuchi H, Ueda M, Kuroda K

Critical role of an oxygen-responsive gene for aerobic nitrogenase activity in *Azotobacter vinelandii* and its application to *Escherichia coli*

In revision