



(51) International Patent Classification:

C05F 11/08 (2006.01) C12N 1/20 (2006.01)
C07K 14/195 (2006.01)

(21) International Application Number:

PCT/US2016/042170

(22) International Filing Date:

13 July 2016 (13.07.2016)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/192,009 13 July 2015 (13.07.2015) US
62/213,567 2 September 2015 (02.09.2015) US

(71) Applicant: **PIVOT BIO, INC.** [US/US]; 2929 7th Street,
Suite 120, Berkeley, CA 94710 (US).

(72) Inventors: **TEMME, Karsten**; 6015 La Salle Avenue,
Oakland, CA 94611 (US). **TAMSIR, Alvin**; 55 Page
Street, Apartment 329, San Francisco, CA 94102 (US).
BLOCH, Sarah; 121 Embarcadero West, Apartment 2203,
Oakland, CA 94607 (US). **CLARK, Rosemary**; 418 Al-
bemarle Street, El Cerrito, CA 94530 (US). **TUNG, Emily**;
1358 Manzanita Drive, Millbrae, CA 94030 (US).

(74) Agents: **STRONCEK, Jacqueline** et al.; Wilson Sonsini
Goodrich & Rosati, 650 Page Mill Road, Palo Alto, CA
94304-1050 (US).

(81) Designated States (unless otherwise indicated, for every
kind of national protection available): AE, AG, AL, AM,
AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY,
BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM,
DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT,
HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR,
KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG,
MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM,
PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC,
SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN,
TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every
kind of regional protection available): ARIPO (BW, GH,
GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ,
TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU,
TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE,
DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU,
LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK,
SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ,
GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— with international search report (Art. 21(3))

[Continued on next page]

(54) Title: METHODS AND COMPOSITIONS FOR IMPROVING PLANT TRAITS

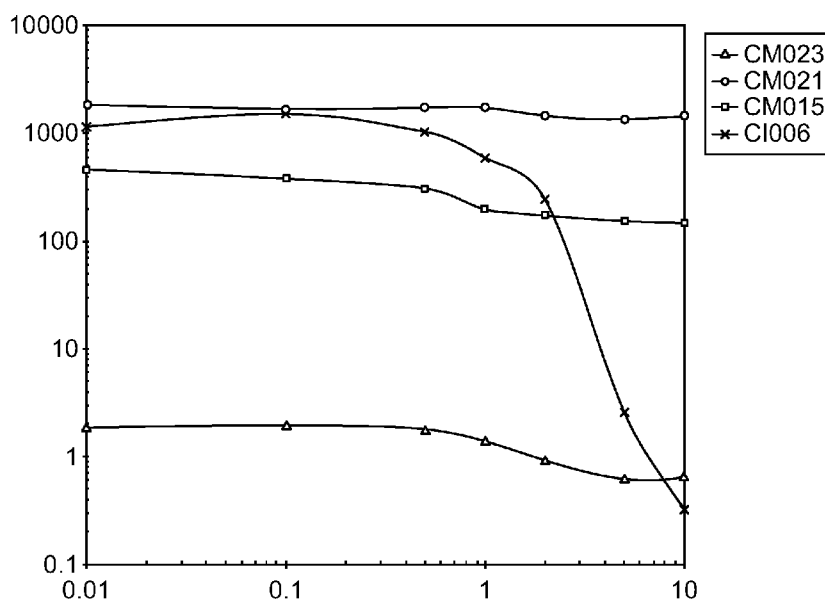


FIG. 4B

(57) Abstract: Disclosed herein are methods of increasing nitrogen fixation in a non-leguminous plant. The methods can comprise exposing the plant to a plurality of bacteria. Each member of the plurality comprises one or more genetic variations introduced into one or more genes or non-coding polynucleotides of the bacteria's nitrogen fixation or assimilation genetic regulatory network, such that the bacteria are capable of fixing atmospheric nitrogen in the presence of exogenous nitrogen. The bacteria are not intergeneric microorganisms. Additionally, the bacteria, *in planta*, produce 1% or more of the fixed nitrogen in the plant.



-
- *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

METHODS AND COMPOSITIONS FOR IMPROVING PLANT TRAITS

CROSS-REFERENCE

[0001] This application claims priority to U.S. Provisional Patent Application No. 62/192,009, filed July 13, 2015, and U.S. Provisional Patent Application No. 62/213,567, filed September 2, 2015, each of which is entirely incorporated herein by reference.

STATEMENT AS TO FEDERALLY SPONSORED RESEARCH

[0002] This invention was made with the support of the United States government under SBIR grant 1520545 awarded by the National Science Foundation. The government has certain rights in the disclosed subject matter.

BACKGROUND OF THE INVENTION

[0003] Plants are linked to the microbiome via a shared metabolome. A multidimensional relationship between a particular crop trait and the underlying metabolome is characterized by a landscape with numerous local maxima. Optimizing from an inferior local maximum to another representing a better trait by altering the influence of the microbiome on the metabolome may be desirable for a variety of reasons, such as for crop optimization. Economically-, environmentally-, and socially-sustainable approaches to agriculture and food production are required to meet the needs of a growing global population. By 2050 the United Nations' Food and Agriculture Organization projects that total food production must increase by 70% to meet the needs of the growing population, a challenge that is exacerbated by numerous factors, including diminishing freshwater resources, increasing competition for arable land, rising energy prices, increasing input costs, and the likely need for crops to adapt to the pressures of a drier, hotter, and more extreme global climate.

[0004] One area of interest is in the improvement of nitrogen fixation. Nitrogen gas (N_2) is a major component of the atmosphere of Earth. In addition, elemental nitrogen (N) is an important component of many chemical compounds which make up living organisms. However, many organisms cannot use N_2 directly to synthesize the chemicals used in physiological processes, such as growth and reproduction. In order to utilize the N_2 , the N_2 must be combined with hydrogen. The combining of hydrogen with N_2 is referred to as nitrogen fixation. Nitrogen fixation, whether accomplished chemically or biologically, requires an investment of large amounts of energy. In biological systems, an enzyme known as nitrogenase catalyzes the reaction which results in nitrogen fixation. An important goal of nitrogen fixation research is the

extension of this phenotype to non-leguminous plants, particularly to important agronomic grasses such as wheat, rice, and maize. Despite enormous progress in understanding the development of the nitrogen-fixing symbiosis between rhizobia and legumes, the path to use that knowledge to induce nitrogen-fixing nodules on non-leguminous crops is still not clear. Meanwhile, the challenge of providing sufficient supplemental sources of nitrogen, such as in fertilizer, will continue to increase with the growing need for increased food production.

SUMMARY OF THE INVENTION

[0005] In view of the foregoing, there is a need to improve the traits of plants imparted by an associated microbiome. The present disclosure addresses this need, and provides additional advantages as well. In some cases, both the species composing the microbiome and their underlying genetics are targets for modulating microbial influence on the metabolome.

[0006] In one aspect, the present disclosure provides a method of increasing nitrogen fixation in a non-leguminous plant, the method comprising exposing the plant to a plurality of bacteria, each member of the plurality comprising one or more genetic variations introduced into one or more genes or non-coding polynucleotides of the bacteria's nitrogen fixation or assimilation genetic regulatory network, such that the bacteria are capable of fixing atmospheric nitrogen in the presence of exogenous nitrogen; wherein the bacteria are not intergeneric microorganisms; and wherein the bacteria, *in planta*, produce 1% or more of the fixed nitrogen in the plant.

[0007] In some embodiments, the bacteria, *in planta*, produce 5% or more of the fixed nitrogen in the plant. In some embodiments, the bacteria, *in planta*, produce 10% or more of the fixed nitrogen in the plant.

[0008] In some embodiments, the one or more genetic variations comprise an introduced control sequence operably linked to said one or more genes of the nitrogen fixation or assimilation genetic regulatory network. In further embodiments, the control sequence is a promoter. In further embodiments, the promoter is an inducible promoter. In some embodiments, the bacteria do not comprise a constitutive promoter operably linked to a gene of the nitrogen fixation or assimilation genetic regulatory network. In some embodiments, the bacteria do not comprise a constitutive promoter operably linked to a gene in the *nif* gene cluster.

[0009] In some embodiments, the bacteria, *in planta*, excrete the nitrogen-containing products of nitrogen fixation. In some embodiments, the plurality of bacteria exposed to the plant do not stimulate an increase in the uptake of exogenous non-atmospheric nitrogen.

[0010] In some embodiments, the plant is grown in soil from a field which has been administered about 50 lbs of nitrogen-containing fertilizer per acre, and wherein the nitrogen-containing fertilizer comprises at least 5% nitrogen by weight. In further embodiments, the nitrogen-containing fertilizer comprises ammonium or an ammonium containing molecule. In some embodiments, the exogenous nitrogen is selected from fertilizer comprising one or more of glutamine, ammonia, ammonium, urea, nitrate, nitrite, ammonium-containing molecules, nitrate-containing molecules, and nitrite-containing molecules.

[0011] In some embodiments, the plurality of bacteria comprise at least two different species of bacteria. In some embodiments, the plurality of bacteria comprise at least two different strains of the same species of bacteria. In some embodiments, the plurality of bacteria are of the genus *Enterobacter*. In some embodiments, the plurality of bacteria are endophytic, epiphytic, or rhizospheric. In some embodiments, the plurality of bacteria colonize the plant such that the bacteria are present in the plant at least 10^5 cfu per gram of fresh weight of the plant.

[0012] In some embodiments, the one or more genes or non-coding polynucleotides of the bacteria's nitrogen fixation or assimilation genetic regulatory network are selected from the group consisting of: *nifA*, *nifL*, *ntrB*, *ntrC*, polynucleotide encoding glutamine synthetase, *glnA*, *glnB*, *glnK*, *drat*, *amtB*, polynucleotide encoding glutaminase, *glnD*, *glnE*, *nifJ*, *nifH*, *nifD*, *nifK*, *nifY*, *nifE*, *nifN*, *nifU*, *nifS*, *nifV*, *nifW*, *nifZ*, *nifM*, *nifF*, *nifB*, and *nifQ*. In some embodiments, the one or more genetic variations is a mutation that results in one or more of: increased expression or activity of NifA or glutaminase; decreased expression or activity of NifL, NtrB, glutamine synthetase, GlnB, GlnK, DraT, AmtB; decreased adenylyl-removing activity of GlnE; or decreased uridylyl-removing activity of GlnD. In some embodiments, the one or more genetic variations is (A) a knock-out mutation; (B) alters or abolishes a regulatory sequence of a target gene; or (C) comprises the insertion of a heterologous regulatory sequence.

[0013] In some embodiments, the plant is an agricultural crop plant. In further embodiments, the agricultural crop plant is selected from sorghum, canola, tomato, strawberry, barley, rice, maize, and wheat. In further embodiments, the plant is a genetically modified organism. In further embodiments, the plant is not a genetically modified organism. In some embodiments, the plant has been genetically engineered or bred for efficient nitrogen use.

[0014] In one aspect, the present disclosure provides a bacterial population comprising bacteria comprising one or more genetic variations introduced into one or more genes or non-coding polynucleotides of the bacteria's nitrogen fixation or assimilation genetic regulatory network, such that the bacteria are capable of fixing atmospheric nitrogen in the presence of exogenous

nitrogen; wherein the bacteria are not intergeneric microorganisms; and wherein the bacteria, *in planta*, produce 1% or more of the fixed nitrogen in a plant grown in the presence of the population of bacteria.

[0015] In some embodiments, the bacteria, *in planta*, produce 5% or more of the fixed nitrogen in the plant. In some embodiments, the bacteria, *in planta*, produce 10% or more of the fixed nitrogen in the plant.

[0016] In some embodiments, the one or more genetic variations comprise an introduced control sequence operably linked to said one or more genes of the nitrogen fixation or assimilation genetic regulatory network. In further embodiments, the control sequence is a promoter. In further embodiments, the promoter is an inducible promoter. In some embodiments, the bacteria do not comprise a constitutive promoter operably linked to a gene of the nitrogen fixation or assimilation genetic regulatory network. In some embodiments, the bacteria do not comprise a constitutive promoter operably linked to a gene in the *nif* gene cluster.

[0017] In some embodiments, the bacteria, *in planta*, excrete the nitrogen-containing products of nitrogen fixation. In some embodiments, the plurality of bacteria exposed to the plant do not stimulate an increase in the uptake of exogenous non-atmospheric nitrogen. In some embodiments, the exogenous nitrogen is selected from fertilizer comprising one or more of glutamine, ammonia, ammonium, urea, nitrate, nitrite, ammonium-containing molecules, nitrate-containing molecules, and nitrite-containing molecules.

[0018] In some embodiments, the bacterial population comprises at least two different species of bacteria. In some embodiments, the bacterial population comprises at least two different strains of the same species of bacteria. In some embodiments, the plurality of bacteria are of the genus *Enterobacter*. In some embodiments, the plurality of bacteria are endophytic, epiphytic, or rhizospheric. In some embodiments, the plurality of bacteria colonize the plant such that the bacteria are present in the plant at least 10^5 cfu per gram of fresh weight of the plant.

[0019] In some embodiments, the one or more genes or non-coding polynucleotides of the bacteria's nitrogen fixation or assimilation genetic regulatory network are selected from the group consisting of: *nifA*, *nifL*, *ntrB*, *ntrC*, polynucleotide encoding glutamine synthetase, *glnA*, *glnB*, *glnK*, *drat*, *amtB*, polynucleotide encoding glutaminase, *glnD*, *glnE*, *nifJ*, *nifH*, *nifD*, *nifK*, *nifY*, *nifE*, *nifN*, *nifU*, *nifS*, *nifV*, *nifW*, *nifZ*, *nifM*, *nifF*, *nifB*, and *nifQ*. In some embodiments, the one or more genetic variations is a mutation that results in one or more of: increased expression or activity of NifA or glutaminase; decreased expression or activity of NifL, NtrB,

glutamine synthetase, GlnB, GlnK, DraT, AmtB; decreased adenylyl-removing activity of GlnE; or decreased uridylyl-removing activity of GlnD. In some embodiments, the one or more genetic variations is (A) a knock-out mutation; (B) alters or abolishes a regulatory sequence of a target gene; or (C) comprises the insertion of a heterologous regulatory sequence.

[0020] In some embodiments, the plant is an agricultural crop plant. In further embodiments, the agricultural crop plant is selected from sorghum, canola, tomato, strawberry, barley, rice, maize, and wheat. In further embodiments, the plant is a genetically modified organism. In further embodiments, the plant is not a genetically modified organism. In some embodiments, the plant has been genetically engineered or bred for efficient nitrogen use.

[0021] In one aspect, the present disclosure provides a composition comprising a bacterial population of the present disclosure. In some embodiments, the composition comprises the bacterial population coated on a surface of a seed. In some embodiments, the composition is formulated as a liquid or powder.

[0022] In one aspect, the present disclosure provides an isolated bacterium deposited as ATCC Accession Deposit No. PTA-122293 or PTA-122294.

[0023] In one aspect, the present disclosure provides a non-intergenic bacterium comprising one or more genetic variations introduced into one or more genes or non-coding polynucleotides of the bacteria's nitrogen fixation or assimilation genetic regulatory network, such that the bacterium is capable of fixing atmospheric nitrogen in the presence of exogenous nitrogen.

[0024] In some embodiments, the one or more genetic variations comprise an introduced control sequence operably linked to said one or more genes of the nitrogen fixation or assimilation genetic regulatory network. In further embodiments, the control sequence is a promoter. In further embodiments, the promoter is an inducible promoter. In some embodiments, the bacteria do not comprise a constitutive promoter operably linked to a gene of the nitrogen fixation or assimilation genetic regulatory network. In some embodiments, the bacteria do not comprise a constitutive promoter operably linked to a gene in the *nif* gene cluster.

[0025] In some embodiments, the one or more genes or non-coding polynucleotides of the bacteria's nitrogen fixation or assimilation genetic regulatory network are selected from the group consisting of: *nifA*, *nifL*, *ntrB*, *ntrC*, polynucleotide encoding glutamine synthetase, *glnA*, *glnB*, *glnK*, *drat*, *amtB*, polynucleotide encoding glutaminase, *glnD*, *glnE*, *nifJ*, *nifH*, *nifD*, *nifK*, *nifY*, *nifE*, *nifN*, *nifU*, *nifS*, *nifV*, *nifW*, *nifZ*, *nifM*, *nifF*, *nifB*, and *nifQ*. In some embodiments, the one or more genetic variations is a mutation that results in one or more of: increased

expression or activity of NifA or glutaminase; decreased expression or activity of NifL, NtrB, glutamine synthetase, GlnB, GlnK, DraT, AmtB; decreased adenylyl-removing activity of GlnE; or decreased uridylyl-removing activity of GlnD. In some embodiments, the one or more genetic variations is (A) a knock-out mutation; (B) alters or abolishes a regulatory sequence of a target gene; or (C) comprises the insertion of a heterologous regulatory sequence.

[0026] In some embodiments, the bacterium is from the genus *Enterobacter*. In some embodiments, the bacterium is endophytic, epiphytic, or rhizospheric.

[0027] In one aspect, the present disclosure provides a method of producing one or more bacteria. In one embodiment, the method comprises (a) isolating bacteria from tissue or soil of a first plant; (b) introducing a genetic variation (e.g. one or more genetic variations) into one or more of the bacteria to produce one or more variant bacteria; (c) exposing a plurality of plants to the variant bacteria; (d) isolating bacteria from tissue or soil of one of the plurality of plants, wherein the plant from which the bacteria is isolated has an improved trait relative to other plants in the plurality of plants; and (e) repeating steps (b) to (d) with bacteria isolated in step (d). The improved trait may be enhanced nitrogen fixation in the plant from which bacteria are isolated, and/or in plants exposed to the bacteria. The genetic variation can be variation in a gene selected from the group consisting of: nifA, nifL, ntrB, ntrC, glnA, glnB, glnK, draT, amtB, glnD, glnE, nifJ, nifH, nifD, nifK, nifY, nifE, nifN, nifU, nifS, nifV, nifW, nifZ, nifM, nifF, nifB, and nifQ. The genetic variation can be a variation in a gene encoding a protein with functionality selected from the group consisting of: glutamine synthetase, glutaminase, glutamine synthetase adenylyltransferase, transcriptional activator, anti-transcriptional activator, pyruvate flavodoxin oxidoreductase, flavodoxin, or NAD⁺-dinitrogen-reductase ADP-D-ribosyltransferase. In some embodiments, the genetic variation is a mutation that results in one or more of: increased expression or activity of NifA or glutaminase; decreased expression or activity of NifL, NtrB, glutamine synthetase, GlnB, GlnK, DraT, AmtB; decreased adenylyl-removing activity of GlnE; or decreased uridylyl-removing activity of GlnD. The genetic variation can be a knock-out mutation, result in elimination or abolishment of activity of a protein domain, alter or abolish a regulatory sequence of a target gene, and/or comprise insertion of a heterologous regulatory sequence. In some embodiments, the genetic variation comprises insertion of a regulatory sequence found within a genome of a bacterial species or genus corresponding to the bacteria into which the genetic variation is introduced. The regulatory sequence may optionally be selected based on expression level of a gene in a bacterial culture or within plant tissue. Genetic variation can be a random mutation at a random location, a random mutation at a target site, or a

predetermined genetic variation specifically introduced to a target site. The genetic variation can comprise insertion, deletion, or replacement of one or more nucleotides, or any combination of these. The genetic variation can be produced by chemical mutagenesis. In some embodiments, the method further comprises exposing the plants to biotic or abiotic stressors. In some embodiments, bacteria isolated after repeating steps (b) to (d) one or more times produce 1% or more (e.g. at least 2%, 5%, 10%, or more) of nitrogen in a second plant of the same type as the first plant, or in a plant exposed to the bacteria. Such production may still be achieved when the second plant is grown in the presence of fertilizer supplemented with glutamine, ammonia, or other chemical source of nitrogen. In some embodiments, bacteria isolated after repeating steps (b) to (d) one or more times exhibit at least a 2-fold increase (e.g. at least 5-fold increase) in nitrogen fixation as compared to bacteria isolated from the first plant. The first plant, or plants in the plurality of plants, can be an agricultural crop plant, such as a plant selected from barley, rice, maize, wheat, sorghum, sweet corn, sugar cane, onions, tomatoes, strawberries, or asparagus. The first plant, or plants in the plurality of plants, can be a model plant, such as a plant selected from *Setaria*, *Brachypodium*, or *Arabidopsis*. In some embodiments, step (a) further comprises performing genetic analysis of isolated bacteria. In some embodiments, step (b) further comprises applying a selection pressure to enrich for bacteria comprising the genetic variation, and optionally isolating bacteria that survive the selection pressure. The selection pressure can comprise cleaving genomes lacking the genetic variation introduced to a target site, wherein cleavage occurs within 100 nucleotides of the target site. The cleavage can be directed by a site-specific nuclease, such as a nuclease selected from the group consisting of a Zinc Finger nuclease, a CRISPR nuclease, a TALE nuclease, or a meganuclease. In some cases, a CRISPR nuclease may be preferred. Bacteria isolated after repeating steps (b) to (d) one or more times are endophytic, epiphytic, or rhizospheric. The bacteria may be isolated from plant tissue (e.g. seeds). The bacteria may comprise a plurality of different bacterial taxa. In some embodiments, isolating bacteria in step (a) comprises isolating bacteria from a seed of the first plant.

[0028] In one aspect, the present disclosure provides a method of increasing nitrogen fixation in a plant. In one embodiment, the method comprises exposing the plant to bacteria comprising one or more genetic variations introduced into one or more genes regulating nitrogen fixation, wherein the bacteria produce 1% or more (e.g. at least 2%, 5%, 10%, or more) of nitrogen in the plant. The bacteria may produce the nitrogen in the presence of fertilizer supplemented with glutamine, ammonia, or other chemical source of supplemental nitrogen. In some embodiments, genetic variation is a variation in a gene selected from the group consisting of: *nifA*, *nifL*, *ntrB*,

ntrC, glutamine synthetase, glnA, glnB, glnK, draT, amtB, glutaminase, glnD, glnE, nifJ, nifH, nifD, nifK, nifY, nifE, nifN, nifU, nifS, nifV, nifW, nifZ, nifM, nifF, nifB, and nifQ. The genetic variation can be a mutation that results in one or more of: increased expression or activity of nifA or glutaminase; decreased expression or activity of nifL, ntrB, glutamine synthetase, glnB, glnK, draT, amtB; decreased adenylyl-removing activity of GlnE; or decreased uridylyl-removing activity of GlnD. In some embodiments, the genetic variation (a) is a knock-out mutation; (b) alters or abolishes a regulatory sequence of a target gene; or (c) comprises insertion of a heterologous regulatory sequence. The bacteria can be endophytic, epiphytic, or rhizospheric. In some cases, the bacteria are of the genus *Enterobacter* or *Rahnella*. The bacteria can comprise a plurality of different bacterial taxa. In some embodiments, the plant is an agricultural crop plant, such as a plant selected from sorghum, canola, tomato, strawberry, barley, rice, maize, and wheat. The plant can be a non-leguminous plant. The plant can be a genetically modified organism (a GMO; e.g. a plant having a genome altered to carry a heterologous gene), a non-genetically modified organism (non-GMO), or have been genetically engineered or bred for efficient nitrogen use.

[0029] In one aspect, the present disclosure provides a bacterial population. In one embodiment, the bacterial population comprises bacteria comprising one or more genetic variations introduced into one or more genes regulating nitrogen fixation, wherein the bacteria produce 1% or more (e.g. at least 2%, 5%, 10%, or more) of nitrogen in a plant grown in the presence of the population of bacteria. The bacteria may produce the nitrogen in the presence of fertilizer supplemented with glutamine, ammonia, or other chemical source of supplemental nitrogen. In some embodiments, the genetic variation is a variation in a gene selected from the group consisting of: nifA, nifL, ntrB, ntrC, glutamine synthetase, glnA, glnB, glnK, draT, amtB, glutaminase, glnD, glnE, nifJ, nifH, nifD, nifK, nifY, nifE, nifN, nifU, nifS, nifV, nifW, nifZ, nifM, nifF, nifB, and nifQ. The genetic variation can be a mutation that results in one or more of: increased expression of nifA or glutaminase; decreased expression of nifL, ntrB, glutamine synthetase, glnB, glnK, draT, amtB; decreased adenylyl-removing activity of GlnE; or decreased uridylyl-removing activity of GlnD. In some embodiments, the genetic variation (a) is a knock-out mutation; (b) alters or abolishes a regulatory sequence of a target gene; or (c) comprises insertion of a heterologous regulatory sequence. The bacteria can be endophytic, epiphytic, or rhizospheric. In some cases, the bacteria are of the genus *Enterobacter* or *Rahnella*. The bacteria can comprise a plurality of different bacterial taxa.

[0030] In one aspect, the present disclosure provides a composition comprising a bacterial population, such as a bacterial population as described herein. The composition can comprise

the bacterial population coated on a surface of a seed. In some embodiments, the composition is formulated as a liquid or a powder.

[0031] In one aspect, the present disclosure provides a bacterium having ATCC deposit number PTA-122293. In one aspect, the present disclosure provides a bacterium having ATCC deposit number PTA-122294.

INCORPORATION BY REFERENCE

[0032] All publications, patents, and patent applications mentioned in this specification are herein incorporated by reference to the same extent as if each individual publication, patent, or patent application was specifically and individually indicated to be incorporated by reference.

BRIEF DESCRIPTION OF THE DRAWINGS

[0033] The novel features of the invention are set forth with particularity in the appended claims. A better understanding of the features and advantages of the present invention will be obtained by reference to the following detailed description that sets forth illustrative embodiments, in which the principles of the invention are utilized, and the accompanying drawings of which:

[0034] Figure 1A-B depicts enrichment and isolation of nitrogen fixing bacteria. (A) Nfb agar plate was used to isolate single colonies of nitrogen fixing bacteria. (B) Semi-solid Nfb agar casted in Balch tube. The arrow points to pellicle of enriched nitrogen fixing bacteria.

[0035] Figure 2 depicts a representative *nifH* PCR screen. Positive bands were observed at ~350bp for two colonies in this screen. Lower bands represent primer-dimers.

[0036] Figure 3 depicts an example of a PCR screen of colonies from CRISPR-Cas-selected mutagenesis. CI006 colonies were screened with primers specific for the *nifL* locus. The wild type PCR product is expected at ~2.2kb, whereas the mutant is expected at ~1.1kb. Seven of ten colonies screened unambiguously show the desired deletion.

[0037] Figures 4A-D depict in vitro phenotypes of various strains. The Acetylene Reduction Assay (ARA) activities of mutants of strain CI010 (**Figure 4A**) and mutants of strain CI006 (**Figure 4B**) grown in nitrogen fixation media supplemented with 0 to 10mM glutamine. ARA activities of additional strains are shown in **Figure 4C**, and the ammonium excretion profile across time of two strains is shown in **Figure 4D**.

[0038] Figure 5 depicts in culture expression profile of 9 different genes in strains CI006 involved in diazotrophic nitrogen fixation. Numbers represent counts of each transcript. Various conditions (0, 1, 10 mM Glutamine and 0%, 10%, 20% atmospheric air in N₂) are indicated.

[0039] **Figure 6** depicts CI006 colonization of corn roots. Corn seedlings were inoculated with CI006 harboring an RFP expression plasmid. After two weeks of growth and plasmid maintenance through watering with the appropriate antibiotic, roots were harvested and imaged through fluorescence microscopy. Colonization of the root intercellular space is observed.

[0040] **Figure 7** depicts nitrogen derived from microbe level in WT (CI050) and optimized (CM002) strain.

[0041] **Figure 8** shows an experimental setup for a Micro-Tom fruiting mass assay.

[0042] **Figure 9** shows a screen of 10 strains for increase in Micro-Tom plant fruit mass. Results for six replicates are presented. For column 3, $p = 0.07$. For column 7, $p = 0.05$.

[0043] **Figures 10A-C** depicts additional results for ARA activities of candidate microbes and counterpart candidate mutants grown in nitrogen fixation media supplemented with 0 to 10mM glutamine.

[0044] **Figure 11** depicts a double mutant that exhibits higher ammonia excretion than the single mutant from which it was derived.

[0045] **Figure 12** depicts NDFA obtained from 15N Gas Uptake experiment (extrapolated back using days exposed) to measure NDFA in Corn plants in fertilized condition.

[0046] **Figure 13** depicts NDFA value obtained from 15N Gas Uptake experiment (extrapolated back using days exposed) to measure NDFA in Setaria plants in fertilized condition.

[0047] **Figure 14A** depicts rate of incorporation of 15N gas. Plants inoculated with evolved strain showed increase in 15N gas incorporation compared to uninoculated plants.

[0048] **Figure 14B** depicts 4 weeks after planting, up to 7% of the nitrogen in plants inoculated with an evolved strain is derived from microbially fixed nitrogen.

[0049] **Figure 14C** depicts leaf area (and other biomass measurement, data not shown) is increased in plants inoculated with an evolved strain when compared to uninoculated or wild type inoculated plants.

[0050] **Figure 15A** depicts evolved strains that show significantly higher nifH production in the root tissue, as measured by in planta transcriptomic study.

[0051] **Figure 15B** depicts that rate of fixed nitrogen found in plant tissue is correlated with the rate in which that particular plant is colonized by HoME optimized strain.

[0052] **Figure 16A** depicts a soil texture map of various field soils tested for colonization. Soils in which a few microbes were originally source from are indicated as stars.

[0053] **Figure 16B** depicts the colonization rate of Strain 1 and Strain 5 that are tested across four different soil types (circles). Both strains showed relatively robust colonization profile across diverse soil types.

[0054] **Figure 16C** depicts colonization of Strain 1 as tested in a field trial over the span of a growing season. Strain 1 persists in the corn tissue up to week 12 after planting and starts to show decline in colonization after that time.

DETAILED DESCRIPTION OF THE INVENTION

[0055] The terms “polynucleotide”, “nucleotide”, “nucleotide sequence”, “nucleic acid” and “oligonucleotide” are used interchangeably. They refer to a polymeric form of nucleotides of any length, either deoxyribonucleotides or ribonucleotides, or analogs thereof. Polynucleotides may have any three dimensional structure, and may perform any function, known or unknown. The following are non-limiting examples of polynucleotides: coding or non-coding regions of a gene or gene fragment, loci (locus) defined from linkage analysis, exons, introns, messenger RNA (mRNA), transfer RNA (tRNA), ribosomal RNA (rRNA), short interfering RNA (siRNA), short-hairpin RNA (shRNA), micro-RNA (miRNA), ribozymes, cDNA, recombinant polynucleotides, branched polynucleotides, plasmids, vectors, isolated DNA of any sequence, isolated RNA of any sequence, nucleic acid probes, and primers. A polynucleotide may comprise one or more modified nucleotides, such as methylated nucleotides and nucleotide analogs. If present, modifications to the nucleotide structure may be imparted before or after assembly of the polymer. The sequence of nucleotides may be interrupted by non-nucleotide components. A polynucleotide may be further modified after polymerization, such as by conjugation with a labeling component.

[0056] “Hybridization” refers to a reaction in which one or more polynucleotides react to form a complex that is stabilized via hydrogen bonding between the bases of the nucleotide residues. The hydrogen bonding may occur by Watson Crick base pairing, Hoogsteen binding, or in any other sequence specific manner according to base complementarity. The complex may comprise two strands forming a duplex structure, three or more strands forming a multi stranded complex, a single self-hybridizing strand, or any combination of these. A hybridization reaction may constitute a step in a more extensive process, such as the initiation of PCR, or the enzymatic cleavage of a polynucleotide by an endonuclease. A second sequence that is complementary to a first sequence is referred to as the “complement” of the first sequence. The term “hybridizable” as applied to a polynucleotide refers to the ability of the polynucleotide to form a complex that is stabilized via hydrogen bonding between the bases of the nucleotide residues in a hybridization reaction.

[0057] “Complementarity” refers to the ability of a nucleic acid to form hydrogen bond(s) with another nucleic acid sequence by either traditional Watson-Crick or other non-traditional types. A percent complementarity indicates the percentage of residues in a nucleic acid molecule which can form hydrogen bonds (e.g., Watson-Crick base pairing) with a second nucleic acid sequence (e.g., 5, 6, 7, 8, 9, 10 out of 10 being 50%, 60%, 70%, 80%, 90%, and 100% complementary, respectively). “Perfectly complementary” means that all the contiguous residues of a nucleic acid sequence will hydrogen bond with the same number of contiguous residues in a second nucleic acid sequence. “Substantially complementary” as used herein refers to a degree of complementarity that is at least 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, or 100% over a region of 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 35, 40, 45, 50, or more nucleotides, or refers to two nucleic acids that hybridize under stringent conditions. Sequence identity, such as for the purpose of assessing percent complementarity, may be measured by any suitable alignment algorithm, including but not limited to the Needleman-Wunsch algorithm (see e.g. the EMBOSS Needle aligner available at www.ebi.ac.uk/Tools/psa/emboss_needle/nucleotide.html, optionally with default settings), the BLAST algorithm (see e.g. the BLAST alignment tool available at blast.ncbi.nlm.nih.gov/Blast.cgi, optionally with default settings), or the Smith-Waterman algorithm (see e.g. the EMBOSS Water aligner available at www.ebi.ac.uk/Tools/psa/emboss_water/nucleotide.html, optionally with default settings). Optimal alignment may be assessed using any suitable parameters of a chosen algorithm, including default parameters.

[0058] In general, “stringent conditions” for hybridization refer to conditions under which a nucleic acid having complementarity to a target sequence predominantly hybridizes with a target sequence, and substantially does not hybridize to non-target sequences. Stringent conditions are generally sequence-dependent, and vary depending on a number of factors. In general, the longer the sequence, the higher the temperature at which the sequence specifically hybridizes to its target sequence. Non-limiting examples of stringent conditions are described in detail in Tijssen (1993), *Laboratory Techniques In Biochemistry And Molecular Biology-Hybridization With Nucleic Acid Probes Part I, Second Chapter “Overview of principles of hybridization and the strategy of nucleic acid probe assay”*, Elsevier, N.Y.

[0059] As used herein, “expression” refers to the process by which a polynucleotide is transcribed from a DNA template (such as into and mRNA or other RNA transcript) and/or the process by which a transcribed mRNA is subsequently translated into peptides, polypeptides, or proteins. Transcripts and encoded polypeptides may be collectively referred to as “gene

product.” If the polynucleotide is derived from genomic DNA, expression may include splicing of the mRNA in a eukaryotic cell.

[0060] The terms “polypeptide”, “peptide” and “protein” are used interchangeably herein to refer to polymers of amino acids of any length. The polymer may be linear or branched, it may comprise modified amino acids, and it may be interrupted by non amino acids. The terms also encompass an amino acid polymer that has been modified; for example, disulfide bond formation, glycosylation, lipidation, acetylation, phosphorylation, or any other manipulation, such as conjugation with a labeling component. As used herein the term “amino acid” includes natural and/or unnatural or synthetic amino acids, including glycine and both the D or L optical isomers, and amino acid analogs and peptidomimetics.

[0061] As used herein, the term “about” is used synonymously with the term “approximately.” Illustratively, the use of the term “about” with regard to an amount indicates that values slightly outside the cited values, e.g., plus or minus 0.1% to 10%.

[0062] The term “biologically pure culture” or “substantially pure culture” refers to a culture of a bacterial species described herein containing no other bacterial species in quantities sufficient to interfere with the replication of the culture or be detected by normal bacteriological techniques.

[0063] “Plant productivity” refers generally to any aspect of growth or development of a plant that is a reason for which the plant is grown. For food crops, such as grains or vegetables, “plant productivity” can refer to the yield of grain or fruit harvested from a particular crop. As used herein, improved plant productivity refers broadly to improvements in yield of grain, fruit, flowers, or other plant parts harvested for various purposes, improvements in growth of plant parts, including stems, leaves and roots, promotion of plant growth, maintenance of high chlorophyll content in leaves, increasing fruit or seed numbers, increasing fruit or seed unit weight, reducing NO₂ emission due to reduced nitrogen fertilizer usage and similar improvements of the growth and development of plants.

[0064] Microbes in and around food crops can influence the traits of those crops. Plant traits that may be influenced by microbes include: yield (*e.g.*, grain production, biomass generation, fruit development, flower set); nutrition (*e.g.*, nitrogen, phosphorus, potassium, iron, micronutrient acquisition); abiotic stress management (*e.g.*, drought tolerance, salt tolerance, heat tolerance); and biotic stress management (*e.g.*, pest, weeds, insects, fungi, and bacteria). Strategies for altering crop traits include: increasing key metabolite concentrations; changing temporal dynamics of microbe influence on key metabolites; linking microbial metabolite

production/degradation to new environmental cues; reducing negative metabolites; and improving the balance of metabolites or underlying proteins.

[0065] As used herein, a “control sequence” refers to an operator, promoter, silencer, or terminator.

[0066] As used herein, “*in planta*” refers to in the plant, and wherein the plant further comprises leaves, roots, stems, seed, ovules, pollen, flowers, fruit, etc.

[0067] In some embodiments, native or endogenous control sequences of genes of the present disclosure are replaced with one or more intragenomic control sequences.

[0068] As used herein, “introduced” refers to the introduction by means of modern biotechnology, and not a naturally occurring introduction.

[0069] In some embodiments, the bacteria of the present disclosure have been modified such that they are not naturally occurring bacteria.

[0070] In some embodiments, the bacteria of the present disclosure are present in the plant in an amount of at least 10^3 cfu, 10^4 cfu, 10^5 cfu, 10^6 cfu, 10^7 cfu, 10^8 cfu, 10^9 cfu, 10^{10} cfu, 10^{11} cfu, or 10^{12} cfu per gram of fresh or dry weight of the plant. In some embodiments, the bacteria of the present disclosure are present in the plant in an amount of at least about 10^3 cfu, about 10^4 cfu, about 10^5 cfu, about 10^6 cfu, about 10^7 cfu, about 10^8 cfu, about 10^9 cfu, about 10^{10} cfu, about 10^{11} cfu, or about 10^{12} cfu per gram of fresh or dry weight of the plant. In some embodiments, the bacteria of the present disclosure are present in the plant in an amount of at least 10^3 to 10^9 , 10^3 to 10^7 , 10^3 to 10^5 , 10^5 to 10^9 , 10^5 to 10^7 , 10^6 to 10^{10} , 10^6 to 10^7 cfu per gram of fresh or dry weight of the plant.

[0071] Fertilizers and exogenous nitrogen of the present disclosure may comprise the following nitrogen-containing molecules: ammonium, nitrate, nitrite, ammonia, glutamine, etc. Nitrogen sources of the present disclosure may include anhydrous ammonia, ammonia sulfate, urea, diammonium phosphate, urea-form, monoammonium phosphate, ammonium nitrate, nitrogen solutions, calcium nitrate, potassium nitrate, sodium nitrate, etc.

[0072] As used herein, “exogenous nitrogen” refers to non-atmospheric nitrogen readily available in the soil, field, or growth medium that is present under non-nitrogen limiting conditions, including ammonia, ammonium, nitrate, nitrite, urea, uric acid, ammonium acids, etc.

[0073] As used herein, “non-nitrogen limiting conditions” refers to non-atmospheric nitrogen available in the soil, field, media at concentrations greater than about 4 mM nitrogen, as

disclosed by Kant *et al.* (2010. J. Exp. Biol. 62(4):1499-1509), which is incorporated herein by reference.

[0074] As used herein, an “intergeneric microorganism” is a microorganism that is formed by the deliberate combination of genetic material originally isolated from organisms of different taxonomic genera. An “intergeneric mutant” can be used interchangeably with “intergeneric microorganism”. An exemplary “intergeneric microorganism” includes a microorganism containing a mobile genetic element which was first identified in a microorganism in a genus different from the recipient microorganism. Further explanation can be found, *inter alia*, in 40 C.F.R. § 725.3.

[0075] As used herein, an “intrageneric microorganism” is a microorganism that is formed by the deliberate combination of genetic material originally isolated from organisms of the same taxonomic genera. An “intrageneric mutant” can be used interchangeably with “intrageneric microorganism”.

[0076] As used herein, “introduced genetic material” means genetic material that is added to, and remains as a component of, the genome of the recipient.

[0077] In some embodiments, the nitrogen fixation and assimilation genetic regulatory network comprises polynucleotides encoding genes and non-coding sequences that direct, modulate, and/or regulate microbial nitrogen fixation and/or assimilation and can comprise polynucleotide sequences of the *nif* cluster (e.g., *nifA*, *nifB*, *nifC*,.....*nifZ*), polynucleotides encoding nitrogen regulatory protein C, polynucleotides encoding nitrogen regulatory protein B, polynucleotide sequences of the *gln* cluster (e.g. *glnA* and *glnD*), *draT*, and ammonia transporters/permeases.

[0078] In some embodiments, fertilizer of the present disclosure comprises at least 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% nitrogen by weight.

[0079] In some embodiments, fertilizer of the present disclosure comprises at least about 5%, about 6%, about 7%, about 8%, about 9%, about 10%, about 11%, about 12%, about 13%, about 14%, about 15%, about 16%, about 17%, about 18%, about 19%, about 20%, about 21%, about 22%, about 23%, about 24%, about 25%, about 26%, about 27%, about 28%, about 29%, about

30%, about 31%, about 32%, about 33%, about 34%, about 35%, about 36%, about 37%, about 38%, about 39%, about 40%, about 41%, about 42%, about 43%, about 44%, about 45%, about 46%, about 47%, about 48%, about 49%, about 50%, about 51%, about 52%, about 53%, about 54%, about 55%, about 56%, about 57%, about 58%, about 59%, about 60%, about 61%, about 62%, about 63%, about 64%, about 65%, about 66%, about 67%, about 68%, about 69%, about 70%, about 71%, about 72%, about 73%, about 74%, about 75%, about 76%, about 77%, about 78%, about 79%, about 80%, about 81%, about 82%, about 83%, about 84%, about 85%, about 86%, about 87%, about 88%, about 89%, about 90%, about 91%, about 92%, about 93%, about 94%, about 95%, about 96%, about 97%, about 98%, or about 99% nitrogen by weight.

[0080] In some embodiments, fertilizer of the present disclosure comprises about 5% to 50%, about 5% to 75%, about 10% to 50%, about 10% to 75%, about 15% to 50%, about 15% to 75%, about 20% to 50%, about 20% to 75%, about 25% to 50%, about 25% to 75%, about 30% to 50%, about 30% to 75%, about 35% to 50%, about 35% to 75%, about 40% to 50%, about 40% to 75%, about 45% to 50%, about 45% to 75%, or about 50% to 75% nitrogen by weight.

[0081] In some embodiments, the increase of nitrogen fixation and/or the production of 1% or more of the nitrogen in the plant are measured relative to control plants which have not been exposed to the bacteria of the present disclosure. All increases or decreases in bacteria are measured relative to control bacteria. All increases or decreases in plants are measured relative to control plants.

[0082] As used herein, a “constitutive promoter” is a promoter which is active under most conditions and/or during most development stages. There are several advantages to using constitutive promoters in expression vectors used in biotechnology, such as: high level of production of proteins used to select transgenic cells or organisms; high level of expression of reporter proteins or scorable markers, allowing easy detection and quantification; high level of production of a transcription factor that is part of a regulatory transcription system; production of compounds that requires ubiquitous activity in the organism; and production of compounds that are required during all stages of development. Non-limiting exemplary constitutive promoters include, CaMV 35S promoter, opine promoters, ubiquitin promoter, alcohol dehydrogenase promoter, etc.

[0083] As used herein, a “non-constitutive promoter” is a promoter which is active under certain conditions, in certain types of cells, and/or during certain development stages. For example, tissue specific, tissue preferred, cell type specific, cell type preferred, inducible promoters, and promoters under development control are non-constitutive promoters. Examples of promoters

under developmental control include promoters that preferentially initiate transcription in certain tissues.

[0084] As used herein, “inducible” or “repressible” promoter is a promoter which is under chemical or environmental factors control. Examples of environmental conditions that may affect transcription by inducible promoters include anaerobic conditions, certain chemicals, the presence of light, acidic or basic conditions, etc.

[0085] As used herein, a “tissue specific” promoter is a promoter that initiates transcription only in certain tissues. Unlike constitutive expression of genes, tissue-specific expression is the result of several interacting levels of gene regulation. As such, in the art sometimes it is preferable to use promoters from homologous or closely related species to achieve efficient and reliable expression of transgenes in particular tissues. This is one of the main reasons for the large amount of tissue-specific promoters isolated from particular tissues found in both scientific and patent literature.

[0086] As used herein, the term “operably linked” refers to the association of nucleic acid sequences on a single nucleic acid fragment so that the function of one is regulated by the other. For example, a promoter is operably linked with a coding sequence when it is capable of regulating the expression of that coding sequence (i.e., that the coding sequence is under the transcriptional control of the promoter). Coding sequences can be operably linked to regulatory sequences in a sense or antisense orientation. In another example, the complementary RNA regions of the disclosure can be operably linked, either directly or indirectly, 5' to the target mRNA, or 3' to the target mRNA, or within the target mRNA, or a first complementary region is 5' and its complement is 3' to the target mRNA

[0087] One trait that may be targeted for regulation by the methods described herein is nitrogen fixation. Nitrogen fertilizer is the largest operational expense on a farm and the biggest driver of higher yields in row crops like corn and wheat. Described herein are microbial products that can deliver renewable forms of nitrogen in non-leguminous crops. While some endophytes have the genetics necessary for fixing nitrogen in pure culture, the fundamental technical challenge is that wild-type endophytes of cereals and grasses stop fixing nitrogen in fertilized fields. The application of chemical fertilizers and residual nitrogen levels in field soils signal the microbe to shut down the biochemical pathway for nitrogen fixation.

[0088] Changes to the transcriptional and post-translational levels of nitrogen fixation regulatory network are required to develop a microbe capable of fixing and transferring nitrogen to corn in the presence of fertilizer. To that end, described herein is Host-Microbe Evolution (HoME)

technology to precisely evolve regulatory networks and elicit novel phenotypes. Also described herein are unique, proprietary libraries of nitrogen-fixing endophytes isolated from corn, paired with extensive omics data surrounding the interaction of microbes and host plant under different environmental conditions like nitrogen stress and excess. This enables precision evolution of the genetic regulatory network of endophytes to produce microbes that actively fix nitrogen even in the presence of fertilizer in the field. Also described herein are evaluations of the technical potential of evolving microbes that colonize corn root tissues and produce nitrogen for fertilized plants and evaluations of the compatibility of endophytes with standard formulation practices and diverse soils to determine feasibility of integrating the microbes into modern nitrogen management strategies.

[0089] In order to utilize elemental nitrogen (N) for chemical synthesis, life forms combine nitrogen gas (N₂) available in the atmosphere with hydrogen in a process known as nitrogen fixation. Because of the energy-intensive nature of biological nitrogen fixation, diazotrophs (bacteria and archaea that fix atmospheric nitrogen gas) have evolved sophisticated and tight regulation of the *nif* gene cluster in response to environmental oxygen and available nitrogen. *Nif* genes encode enzymes involved in nitrogen fixation (such as the nitrogenase complex) and proteins that regulate nitrogen fixation. Shamseldin (2013. Global J. Biotechnol. Biochem. 8(4):84-94) discloses detailed descriptions of *nif* genes and their products, and is incorporated herein by reference. Described herein are methods of producing a plant with an improved trait comprising isolating bacteria from a first plant, introducing a genetic variation into a *nif* gene of the isolated bacteria, exposing a second plant to the variant bacteria, isolating bacteria from the second plant having an improved trait relative to the first plant, and repeating the steps with bacteria isolated from the second plant.

[0090] In Proteobacteria, regulation of nitrogen fixation centers around the σ_{54} -dependent enhancer-binding protein NifA, the positive transcriptional regulator of the *nif* cluster. Intracellular levels of active NifA are controlled by two key factors: transcription of the *nifLA* operon, and inhibition of NifA activity by protein-protein interaction with NifL. Both of these processes are responsive to intracellular glutamine levels *via* the PII protein signaling cascade. This cascade is mediated by GlnD, which directly senses glutamine and catalyzes the uridylylation or deuridylylation of two PII regulatory proteins – GlnB and GlnK – in response to the absence or presence, respectively, of bound glutamine. Under conditions of nitrogen excess, unmodified GlnB signals the deactivation of the *nifLA* promoter. However, under conditions of nitrogen limitation, GlnB is post-translationally modified, which inhibits its activity and leads to transcription of the *nifLA* operon. In this way, *nifLA* transcription is tightly controlled in

response to environmental nitrogen via the PII protein signaling cascade. On the post-translational level of NifA regulation, GlnK inhibits the NifL/NifA interaction in a manner dependent on the overall level of free GlnK within the cell.

[0091] NifA is transcribed from the *nifLA* operon, whose promoter is activated by phosphorylated NtrC, another σ_{54} -dependent regulator. The phosphorylation state of NtrC is mediated by the histidine kinase NtrB, which interacts with deuridylylated GlnB, but not uridylylated GlnB. Under conditions of nitrogen excess, a high intracellular level of glutamine leads to deuridylylation of GlnB, which then interacts with NtrB to deactivate its phosphorylation activity and activate its phosphatase activity, resulting in dephosphorylation of NtrC and the deactivation of the *nifLA* promoter. However, under conditions of nitrogen limitation, a low level of intracellular glutamine results in uridylylation of GlnB, which inhibits its interaction with NtrB and allows the phosphorylation of NtrC and transcription of the *nifLA* operon. In this way, *nifLA* expression is tightly controlled in response to environmental nitrogen via the PII protein signaling cascade. *nifA*, *ntrB*, *ntrC*, and *glnB*, are all genes that can be mutated in the methods described herein.

[0092] The activity of NifA is also regulated post-translationally in response to environmental nitrogen, most typically through NifL-mediated inhibition of NifA activity. In general, the interaction of NifL and NifA is influenced by the PII protein signaling cascade via GlnK, although the nature of the interactions between GlnK and NifL/NifA varies significantly between diazotrophs. In *Klebsiella pneumoniae*, both forms of GlnK inhibit the NifL/NifA interaction, and the interaction between GlnK and NifL/NifA is determined by the overall level of free GlnK within the cell. Under nitrogen-excess conditions, deuridylylated GlnK interacts with the ammonium transporter AmtB, which serves to both block ammonium uptake by AmtB and sequester GlnK to the membrane, allowing inhibition of NifA by NifL. On the other hand, in *Azotobacter vinelandii*, interaction with deuridylylated GlnK is required for the NifL/NifA interaction and NifA inhibition, while uridylylation of GlnK inhibits its interaction with NifL. In diazotrophs lacking the *nifL* gene, there is evidence that NifA activity is inhibited directly by interaction with the deuridylylated forms of both GlnK and GlnB under nitrogen-excess conditions. Regardless of the mechanism, post-translational inhibition of NifA is an important regulator of the *nif* cluster in most known diazotrophs. Additionally, *nifL*, *amtB*, and *glnK*, are genes that can be mutated in the methods described herein.

[0093] In addition to regulating the transcription of the *nif* gene cluster, many diazotrophs have evolved a mechanism for the direct post-translational modification and inhibition of the nitrogenase enzyme itself, known as nitrogenase shutoff. This is mediated by ADP-ribosylation

of the Fe protein (NifH) under nitrogen-excess conditions, which disrupts its interaction with the MoFe protein complex (NifDK) and abolishes nitrogenase activity. DraT catalyzes the ADP-ribosylation of the Fe protein and shutoff of nitrogenase, while DraG catalyzes the removal of ADP-ribose and reactivation of nitrogenase. As with *nifLA* transcription and NifA inhibition, nitrogenase shutoff is also regulated *via* the PII protein signaling cascade. Under nitrogen-excess conditions, deuridylylated GlnB interacts with and activates DraT, while deuridylylated GlnK interacts with both DraG and AmtB to form a complex, sequestering DraG to the membrane. Under nitrogen-limiting conditions, the uridylylated forms of GlnB and GlnK do not interact with DraT and DraG, respectively, leading to the inactivation of DraT and the diffusion of DraG to the Fe protein, where it removes the ADP-ribose and activates nitrogenase. The methods described herein also contemplate introducing genetic variation into the *nifH*, *nifD*, *nifK*, and *draT* genes.

[0094] Although some endophytes have the ability to fix nitrogen *in vitro*, often the genetics are silenced in the field by high levels of exogenous chemical fertilizers. One can decouple the sensing of exogenous nitrogen from expression of the nitrogenase enzyme to facilitate field-based nitrogen fixation. Improving the integral of nitrogenase activity across time further serves to augment the production of nitrogen for utilization by the crop. Specific targets for genetic variation to facilitate field-based nitrogen fixation using the methods described herein include one or more genes selected from the group consisting of *nifA*, *nifL*, *ntrB*, *ntrC*, *glnA*, *glnB*, *glnK*, *draT*, *amtB*, *glnD*, *glnE*, *nifJ*, *nifH*, *nifD*, *nifK*, *nifY*, *nifE*, *nifN*, *nifU*, *nifS*, *nifV*, *nifW*, *nifZ*, *nifM*, *nifF*, *nifB*, and *nifQ*.

[0095] An additional target for genetic variation to facilitate field-based nitrogen fixation using the methods described herein is the NifA protein. The NifA protein is typically the activator for expression of nitrogen fixation genes. Increasing the production of NifA (either constitutively or during high ammonia condition) circumvents the native ammonia-sensing pathway. In addition, reducing the production of NifL proteins, a known inhibitor of NifA, also leads to an increased level of freely active NifA. In addition, increasing the transcription level of the *nifAL* operon (either constitutively or during high ammonia condition) also leads to an overall higher level of NifA proteins. Elevated level of *nifAL* expression is achieved by altering the promoter itself or by reducing the expression of NtrB (part of *ntrB* and *ntrC* signaling cascade that originally would result in the shutoff of *nifAL* operon during high nitrogen condition). High level of NifA achieved by these or any other methods described herein increases the nitrogen fixation activity of the endophytes.

[0096] Another target for genetic variation to facilitate field-based nitrogen fixation using the methods described herein is the GlnD/GlnB/GlnK PII signaling cascade. The intracellular glutamine level is sensed through the GlnD/GlnB/GlnK PII signaling cascade. Active site mutations in GlnD that abolish the uridylyl-removing activity of GlnD disrupt the nitrogen-sensing cascade. In addition, reduction of the GlnB concentration short circuits the glutamine-sensing cascade. These mutations “trick” the cells into perceiving a nitrogen-limited state, thereby increasing the nitrogen fixation level activity.

[0097] The amtB protein is also a target for genetic variation to facilitate field-based nitrogen fixation using the methods described herein. Ammonia uptake from the environment can be reduced by decreasing the expression level of amtB protein. Without intracellular ammonia, the endophyte is not able to sense the high level of ammonia, preventing the down-regulation of nitrogen fixation genes. Any ammonia that manages to get into the intracellular compartment is converted into glutamine. Intracellular glutamine level is the major currency of nitrogen sensing. Decreasing the intracellular glutamine level prevents the cells from sensing high ammonium levels in the environment. This can be done by increasing the expression level of glutaminase, an enzyme that converts glutamine into glutamate. In addition, intracellular glutamine can also be reduced by decreasing glutamine synthase (an enzyme that converts ammonia into glutamine). In diazotrophs, fixed ammonia is quickly assimilated into glutamine and glutamate to be used for cellular processes. Disruptions to ammonia assimilation may enable diversion of fixed nitrogen to be exported from the cell as ammonia. The fixed ammonia is predominantly assimilated into glutamine by glutamine synthetase (GS), encoded by *glnA*, and subsequently into glutamate by glutamine oxoglutarate aminotransferase (GOGAT). In some examples, *glnS* encodes a glutamine synthetase. GS is regulated post-translationally by GS adenylyl transferase (GlnE), a bi-functional enzyme encoded by *glnE* that catalyzes both the adenylylation and de-adenylylation of GS through activity of its adenylyl-transferase (AT) and adenylyl-removing (AR) domains, respectively. Under nitrogen limiting conditions, *glnA* is expressed, and GlnE's AR domain de-adenylylates GS, allowing it to be active. Under conditions of nitrogen excess, *glnA* expression is turned off, and GlnE's AT domain is activated allosterically by glutamine, causing the adenylylation and deactivation of GS.

[0098] Furthermore, the *draT* gene may also be a target for genetic variation to facilitate field-based nitrogen fixation using the methods described herein. Once nitrogen fixing enzymes are produced by the cell, nitrogenase shut-off represents another level in which cell downregulates fixation activity in high nitrogen condition. This shut-off could be removed by decreasing the expression level of DraT.

[0099] Methods for imparting new microbial phenotypes can be performed at the transcriptional, translational, and post-translational levels. The transcriptional level includes changes at the promoter (such as changing sigma factor affinity or binding sites for transcription factors, including deletion of all or a portion of the promoter) or changing transcription terminators and attenuators. The translational level includes changes at the ribosome binding sites and changing mRNA degradation signals. The post-translational level includes mutating an enzyme's active site and changing protein-protein interactions. These changes can be achieved in a multitude of ways. Reduction of expression level (or complete abolishment) can be achieved by swapping the native ribosome binding site (RBS) or promoter with another with lower strength/efficiency. ATG start sites can be swapped to a GTG, TTG, or CTG start codon, which results in reduction in translational activity of the coding region. Complete abolishment of expression can be done by knocking out (deleting) the coding region of a gene. Frameshifting the open reading frame (ORF) likely will result in a premature stop codon along the ORF, thereby creating a non-functional truncated product. Insertion of in-frame stop codons will also similarly create a non-functional truncated product. Addition of a degradation tag at the N or C terminal can also be done to reduce the effective concentration of a particular gene.

[00100] Conversely, expression level of the genes described herein can be achieved by using a stronger promoter. To ensure high promoter activity during high nitrogen level condition (or any other condition), a transcription profile of the whole genome in a high nitrogen level condition could be obtained, and active promoters with a desired transcription level can be chosen from that dataset to replace the weak promoter. Weak start codons can be swapped out with an ATG start codon for better translation initiation efficiency. Weak ribosomal binding sites (RBS) can also be swapped out with a different RBS with higher translation initiation efficiency. In addition, site specific mutagenesis can also be performed to alter the activity of an enzyme.

[00101] Increasing the level of nitrogen fixation that occurs in a plant can lead to a reduction in the amount of chemical fertilizer needed for crop production and reduce greenhouse gas emissions (*e.g.*, nitrous oxide).

Serial Passage

[00102] Production of bacteria to improve plant traits (*e.g.*, nitrogen fixation) can be achieved through serial passage. This can be done by selecting plants which have a particular improved trait which is influenced by the microbial flora, in addition to identifying bacteria and/or compositions that are capable of imparting one or more improved traits to one or more plants. One method of producing a bacteria to improve a plant trait includes the steps of: (a)

isolating bacteria from tissue or soil of a first plant; (b) introducing a genetic variation into one or more of the bacteria to produce one or more variant bacteria; (c) exposing a plurality of plants to the variant bacteria; (d) isolating bacteria from tissue or soil of one of the plurality of plants, wherein the plant from which the bacteria is isolated has an improved trait relative to other plants in the plurality of plants; and (e) repeating steps (b) to (d) with bacteria isolated from the plant with an improved trait (step (d)). Steps (b) to (d) can be repeated any number of times (*e.g.*, once, twice, three times, four times, five times, ten times, or more) until the improved trait in a plant reaches a desired level. Further, the plurality of plants can be more than two plants, such as 10 to 20 plants, or 20 or more, 50 or more, 100 or more, 300 or more, 500 or more, or 1000 or more plants.

[00103] In addition to obtaining a plant with an improved trait, a bacterial population comprising bacteria comprising one or more genetic variations introduced into one or more genes (*e.g.*, genes regulating nitrogen fixation) is obtained. By repeating the steps described above, a population of bacteria can be obtained that include the most appropriate members of the population that correlate with a plant trait of interest. The bacteria in this population can be identified and their beneficial properties determined, such as by genetic and/or phenotypic analysis. Genetic analysis may occur of isolated bacteria in step (a). Phenotypic and/or genotypic information may be obtained using techniques including: high through-put screening of chemical components of plant origin, sequencing techniques including high throughput sequencing of genetic material, differential display techniques (including DDRT-PCR, and DD-PCR), nucleic acid microarray techniques, RNA-seq (Whole Transcriptome Shotgun Sequencing), and qRT-PCR (quantitative real time PCR). Information gained can be used to obtain community profiling information on the identity and activity of bacteria present, such as phylogenetic analysis or microarray-based screening of nucleic acids coding for components of rRNA operons or other taxonomically informative loci. Examples of taxonomically informative loci include 16S rRNA gene, 23S rRNA gene, 5S rRNA gene, 5.8S rRNA gene, 12S rRNA gene, 18S rRNA gene, 28S rRNA gene, *gyrB* gene, *rpoB* gene, *fusA* gene, *recA* gene, *coxI* gene, *nifD* gene. Example processes of taxonomic profiling to determine taxa present in a population are described in US20140155283. Bacterial identification may comprise characterizing activity of one or more genes or one or more signaling pathways, such as genes associated with the nitrogen fixation pathway. Synergistic interactions (where two components, by virtue of their combination, increase a desired effect by more than an additive amount) between different bacterial species may also be present in the bacterial populations.

[00104] The genetic variation may be a gene selected from the group consisting of: *nifA*, *nifL*, *ntrB*, *ntrC*, *glnA*, *glnB*, *glnK*, *draT*, *amtB*, *glnD*, *glnE*, *nifJ*, *nifH*, *nifD*, *nifK*, *nifY*, *nifE*, *nifN*, *nifU*, *nifS*, *nifV*, *nifW*, *nifZ*, *nifM*, *nifF*, *nifB*, and *nifQ*. The genetic variation may be a variation in a gene encoding a protein with functionality selected from the group consisting of: glutamine synthetase, glutaminase, glutamine synthetase adenylyltransferase, transcriptional activator, anti-transcriptional activator, pyruvate flavodoxin oxidoreductase, flavodoxin, or NAD⁺-dinitrogen-reductase aDP-D-ribosyltransferase. The genetic variation may be a mutation that results in one or more of: increased expression or activity of NifA or glutaminase; decreased expression or activity of NifL, NtrB, glutamine synthetase, GlnB, GlnK, DraT, AmtB; decreased adenylyl-removing activity of GlnE; or decreased uridylyl-removing activity of GlnD. Introducing a genetic variation may comprise insertion and/or deletion of one or more nucleotides at a target site, such as 1, 2, 3, 4, 5, 10, 25, 50, 100, 250, 500, or more nucleotides. The genetic variation introduced into one or more bacteria of the methods disclosed herein may be a knock-out mutation (e.g. deletion of a promoter, insertion or deletion to produce a premature stop codon, deletion of an entire gene), or it may be elimination or abolishment of activity of a protein domain (e.g. point mutation affecting an active site, or deletion of a portion of a gene encoding the relevant portion of the protein product), or it may alter or abolish a regulatory sequence of a target gene. One or more regulatory sequences may also be inserted, including heterologous regulatory sequences and regulatory sequences found within a genome of a bacterial species or genus corresponding to the bacteria into which the genetic variation is introduced. Moreover, regulatory sequences may be selected based on the expression level of a gene in a bacterial culture or within a plant tissue. The genetic variation may be a pre-determined genetic variation that is specifically introduced to a target site. The genetic variation may be a random mutation within the target site. The genetic variation may be an insertion or deletion of one or more nucleotides. In some cases, a plurality of different genetic variations (e.g. 2, 3, 4, 5, 10, or more) are introduced into one or more of the isolated bacteria before exposing the bacteria to plants for assessing trait improvement. The plurality of genetic variations can be any of the above types, the same or different types, and in any combination. In some cases, a plurality of different genetic variations are introduced serially, introducing a first genetic variation after a first isolation step, a second genetic variation after a second isolation step, and so forth so as to accumulate a plurality of genetic variations in bacteria imparting progressively improved traits on the associated plants.

[00105] In general, the term “genetic variation” refers to any change introduced into a polynucleotide sequence relative to a reference polynucleotide, such as a reference genome or

portion thereof, or reference gene or portion thereof. A genetic variation may be referred to as a “mutation,” and a sequence or organism comprising a genetic variation may be referred to as a “genetic variant” or “mutant”. Genetic variations can have any number of effects, such as the increase or decrease of some biological activity, including gene expression, metabolism, and cell signaling. Genetic variations can be specifically introduced to a target site, or introduced randomly. A variety of molecular tools and methods are available for introducing genetic variation. For example, genetic variation can be introduced via polymerase chain reaction mutagenesis, oligonucleotide-directed mutagenesis, saturation mutagenesis, fragment shuffling mutagenesis, homologous recombination, CRISPR/Cas9 systems, chemical mutagenesis, and combinations thereof. Chemical methods of introducing genetic variation include exposure of DNA to a chemical mutagen, e.g., ethyl methanesulfonate (EMS), methyl methanesulfonate (MMS), N-nitrosourea (EN U), N-methyl-N-nitro-N'-nitrosoguanidine, 4-nitroquinoline N-oxide, diethylsulfate, benzopyrene, cyclophosphamide, bleomycin, triethylmelamine, acrylamide monomer, nitrogen mustard, vincristine, diepoxyalkanes (for example, diepoxybutane), ICR-170, formaldehyde, procarbazine hydrochloride, ethylene oxide, dimethylnitrosamine, 7,12 dimethylbenz(a)anthracene, chlorambucil, hexamethylphosphoramide, bisulfan, and the like. Radiation mutation-inducing agents include ultraviolet radiation, γ -irradiation, X-rays, and fast neutron bombardment. Genetic variation can also be introduced into a nucleic acid using, e.g., trimethylpsoralen with ultraviolet light. Random or targeted insertion of a mobile DNA element, e.g., a transposable element, is another suitable method for generating genetic variation. Genetic variations can be introduced into a nucleic acid during amplification in a cell-free in vitro system, e.g., using a polymerase chain reaction (PCR) technique such as error-prone PCR. Genetic variations can be introduced into a nucleic acid in vitro using DNA shuffling techniques (e.g., exon shuffling, domain swapping, and the like). Genetic variations can also be introduced into a nucleic acid as a result of a deficiency in a DNA repair enzyme in a cell, e.g., the presence in a cell of a mutant gene encoding a mutant DNA repair enzyme is expected to generate a high frequency of mutations (i.e., about 1 mutation/100 genes-1 mutation/10,000 genes) in the genome of the cell. Examples of genes encoding DNA repair enzymes include but are not limited to Mut H, Mut S, Mut L, and Mut U, and the homologs thereof in other species (e.g., MSH 1 6, PMS 1 2, MLH 1, GTBP, ERCC-1, and the like). Example descriptions of various methods for introducing genetic variations are provided in e.g., Stemple (2004) *Nature* 5:1-7; Chiang et al. (1993) *PCR Methods Appl* 2(3): 210-217; Stemmer (1994) *Proc. Natl. Acad. Sci. USA* 91:10747-10751; and U.S. Pat. Nos. 6,033,861, and 6,773,900.

[00106] As a cyclic amplification technique, polymerase chain reaction (PCR) mutagenesis uses mutagenic primers to introduce desired mutations. PCR is performed by cycles of denaturation, annealing, and extension. After amplification by PCR, selection of mutated DNA and removal of parental plasmid DNA can be accomplished by: 1) replacement of dCTP by hydroxymethylated-dCTP during PCR, followed by digestion with restriction enzymes to remove non-hydroxymethylated parent DNA only; 2) simultaneous mutagenesis of both an antibiotic resistance gene and the studied gene changing the plasmid to a different antibiotic resistance, the new antibiotic resistance facilitating the selection of the desired mutation thereafter; 3) after introducing a desired mutation, digestion of the parent methylated template DNA by restriction enzyme DpnI which cleaves only methylated DNA, by which the mutagenized unmethylated chains are recovered; or 4) circularization of the mutated PCR products in an additional ligation reaction to increase the transformation efficiency of mutated DNA. Further description of exemplary methods can be found in e.g. US7132265, US6713285, US6673610, US6391548, US5789166, US5780270, US5354670, US5071743, and US20100267147.

[00107] Oligonucleotide-directed mutagenesis, also called site-directed mutagenesis, typically utilizes a synthetic DNA primer. This synthetic primer contains the desired mutation and is complementary to the template DNA around the mutation site so that it can hybridize with the DNA in the gene of interest. The mutation may be a single base change (a point mutation), multiple base changes, deletion, or insertion, or a combination of these. The single-strand primer is then extended using a DNA polymerase, which copies the rest of the gene. The gene thus copied contains the mutated site, and may then be introduced into a host cell as a vector and cloned. Finally, mutants can be selected by DNA sequencing to check that they contain the desired mutation.

[00108] Genetic variations can be introduced using error-prone PCR. In this technique the gene of interest is amplified using a DNA polymerase under conditions that are deficient in the fidelity of replication of sequence. The result is that the amplification products contain at least one error in the sequence. When a gene is amplified and the resulting product(s) of the reaction contain one or more alterations in sequence when compared to the template molecule, the resulting products are mutagenized as compared to the template. Another means of introducing random mutations is exposing cells to a chemical mutagen, such as nitrosoguanidine or ethyl methanesulfonate (Nestmann, *Mutat Res* 1975 June; 28(3):323-30), and the vector containing the gene is then isolated from the host.

[00109] Saturation mutagenesis is another form of random mutagenesis, in which one tries to generate all or nearly all possible mutations at a specific site, or narrow region of a gene. In a general sense, saturation mutagenesis is comprised of mutagenizing a complete set of mutagenic cassettes (wherein each cassette is, for example, 1-500 bases in length) in defined polynucleotide sequence to be mutagenized (wherein the sequence to be mutagenized is, for example, from 15 to 100, 000 bases in length). Thusly, a group of mutations (e.g. ranging from 1 to 100 mutations) is introduced into each cassette to be mutagenized. A grouping of mutations to be introduced into one cassette can be different or the same from a second grouping of mutations to be introduced into a second cassette during the application of one round of saturation mutagenesis. Such groupings are exemplified by deletions, additions, groupings of particular codons, and groupings of particular nucleotide cassettes.

[00110] Fragment shuffling mutagenesis, also called DNA shuffling, is a way to rapidly propagate beneficial mutations. In an example of a shuffling process, DNase is used to fragment a set of parent genes into pieces of e.g. about 50-100 bp in length. This is then followed by a polymerase chain reaction (PCR) without primers--DNA fragments with sufficient overlapping homologous sequence will anneal to each other and are then be extended by DNA polymerase. Several rounds of this PCR extension are allowed to occur, after some of the DNA molecules reach the size of the parental genes. These genes can then be amplified with another PCR, this time with the addition of primers that are designed to complement the ends of the strands. The primers may have additional sequences added to their 5' ends, such as sequences for restriction enzyme recognition sites needed for ligation into a cloning vector. Further examples of shuffling techniques are provided in US20050266541.

[00111] Homologous recombination mutagenesis involves recombination between an exogenous DNA fragment and the targeted polynucleotide sequence. After a double-strand break occurs, sections of DNA around the 5' ends of the break are cut away in a process called resection. In the strand invasion step that follows, an overhanging 3' end of the broken DNA molecule then "invades" a similar or identical DNA molecule that is not broken. The method can be used to delete a gene, remove exons, add a gene, and introduce point mutations. Homologous recombination mutagenesis can be permanent or conditional. Typically, a recombination template is also provided. A recombination template may be a component of another vector, contained in a separate vector, or provided as a separate polynucleotide. In some embodiments, a recombination template is designed to serve as a template in homologous recombination, such as within or near a target sequence nicked or cleaved by a site-specific nuclease. A template polynucleotide may be of any suitable length, such as about or more than

about 10, 15, 20, 25, 50, 75, 100, 150, 200, 500, 1000, or more nucleotides in length. In some embodiments, the template polynucleotide is complementary to a portion of a polynucleotide comprising the target sequence. When optimally aligned, a template polynucleotide might overlap with one or more nucleotides of a target sequences (e.g. about or more than about 1, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100 or more nucleotides). In some embodiments, when a template sequence and a polynucleotide comprising a target sequence are optimally aligned, the nearest nucleotide of the template polynucleotide is within about 1, 5, 10, 15, 20, 25, 50, 75, 100, 200, 300, 400, 500, 1000, 5000, 10000, or more nucleotides from the target sequence. Non-limiting examples of site-directed nucleases useful in methods of homologous recombination include zinc finger nucleases, CRISPR nucleases, TALE nucleases, and meganuclease. For a further description of the use of such nucleases, see e.g. US8795965 and US20140301990.

[00112] CRISPR/Cas9 (Clustered regularly interspaced short palindromic repeats)/CRISPR-associated (Cas) systems provide bacteria and archaea with adaptive immunity against viruses and plasmids by using CRISPR RNAs (crRNAs) to guide the silencing of invading nucleic acids. The Cas9 protein (or functional equivalent and/or variant thereof, i.e., Cas9-like protein) naturally contains DNA endonuclease activity that depends on association of the protein with two naturally occurring or synthetic RNA molecules called crRNA and tracrRNA (also called guide RNAs). In some cases, the two molecules are covalently linked to form a single molecule (also called a single guide RNA ("sgRNA")). Thus, the Cas9 or Cas9-like protein associates with a DNA-targeting RNA (which term encompasses both the two-molecule guide RNA configuration and the single-molecule guide RNA configuration), which activates the Cas9 or Cas9-like protein and guides the protein to a target nucleic acid sequence. If the Cas9 or Cas9-like protein retains its natural enzymatic function, it will cleave target DNA to create a double-strand break, which can lead to genome alteration (i.e., editing: deletion, insertion (when a donor polynucleotide is present), replacement, etc.), thereby altering gene expression. Some variants of Cas9 (which variants are encompassed by the term Cas9-like) have been altered such that they have a decreased DNA cleaving activity (in some cases, they cleave a single strand instead of both strands of the target DNA, while in other cases, they have severely reduced to no DNA cleavage activity). Further exemplary descriptions of CRISPR systems for introducing genetic variation can be found in, e.g. US8795965.

[00113] Mutagens that create primarily point mutations and short deletions, insertions, transversions, and/or transitions, including chemical mutagens or radiation, may be used to create genetic variations. Mutagens include, but are not limited to, ethyl methanesulfonate,

methylnmethane sulfonate, N-ethyl-N-nitrosourea, triethylmelamine, N-methyl-N-nitrosourea, procarbazine, chlorambucil, cyclophosphamide, diethyl sulfate, acrylamide monomer, melphalan, nitrogen mustard, vincristine, dimethylnitrosamine, N-methyl-N'-nitro-Nitrosoguanidine, nitrosoguanidine, 2-aminopurine, 7,12 dimethyl-benz(a)anthracene, ethylene oxide, hexamethylphosphoramide, bisulfan, diepoxyalkanes (diepoxyoctane, diepoxybutane, and the like), 2-methoxy-6-chloro-9[3-(ethyl-2-chloro-ethyl)aminopropylamino]acridine dihydrochloride and formaldehyde.

[00114] Introducing genetic variation may be an incomplete process, such that some bacteria in a treated population of bacteria carry a desired mutation while others do not. In some cases, it is desirable to apply a selection pressure so as to enrich for bacteria carrying a desired genetic variation. Traditionally, selection for successful genetic variants involved selection for or against some functionality imparted or abolished by the genetic variation, such as in the case of inserting antibiotic resistance gene or abolishing a metabolic activity capable of converting a non-lethal compound into a lethal metabolite. It is also possible to apply a selection pressure based on a polynucleotide sequence itself, such that only a desired genetic variation need be introduced (e.g. without also requiring a selectable marker). In this case, the selection pressure can comprise cleaving genomes lacking the genetic variation introduced to a target site, such that selection is effectively directed against the reference sequence into which the genetic variation is sought to be introduced. Typically, cleavage occurs within 100 nucleotides of the target site (e.g. within 75, 50, 25, 10, or fewer nucleotides from the target site, including cleavage at or within the target site). Cleaving may be directed by a site-specific nuclease selected from the group consisting of a Zinc Finger nuclease, a CRISPR nuclease, a TALE nuclease (TALEN), or a meganuclease. Such a process is similar to processes for enhancing homologous recombination at a target site, except that no template for homologous recombination is provided. As a result, bacteria lacking the desired genetic variation are more likely to undergo cleavage that, left unrepaired, results in cell death. Bacteria surviving selection may then be isolated for use in exposing to plants for assessing conferral of an improved trait.

[00115] A CRISPR nuclease may be used as the site-specific nuclease to direct cleavage to a target site. An improved selection of mutated microbes can be obtained by using Cas9 to kill non-mutated cells. Plants are then inoculated with the mutated microbes to re-confirm symbiosis and create evolutionary pressure to select for efficient symbionts. Microbes can then be re-isolated from plant tissues. CRISPR nuclease systems employed for selection against non-variants can employ similar elements to those described above with respect to introducing

genetic variation, except that no template for homologous recombination is provided. Cleavage directed to the target site thus enhances death of affected cells.

[00116] Other options for specifically inducing cleavage at a target site are available, such as zinc finger nucleases, TALE nuclease (TALEN) systems, and meganuclease. Zinc-finger nucleases (ZFNs) are artificial DNA endonucleases generated by fusing a zinc finger DNA binding domain to a DNA cleavage domain. ZFNs can be engineered to target desired DNA sequences and this enables zinc-finger nucleases to cleave unique target sequences. When introduced into a cell, ZFNs can be used to edit target DNA in the cell (e.g., the cell's genome) by inducing double strand breaks. Transcription activator-like effector nucleases (TALENs) are artificial DNA endonucleases generated by fusing a TAL (Transcription activator-like) effector DNA binding domain to a DNA cleavage domain. TALENS can be quickly engineered to bind practically any desired DNA sequence and when introduced into a cell, TALENS can be used to edit target DNA in the cell (e.g., the cell's genome) by inducing double strand breaks.

Meganucleases (homing endonuclease) are endodeoxyribonucleases characterized by a large recognition site (double-stranded DNA sequences of 12 to 40 base pairs). Meganucleases can be used to replace, eliminate or modify sequences in a highly targeted way. By modifying their recognition sequence through protein engineering, the targeted sequence can be changed.

Meganucleases can be used to modify all genome types, whether bacterial, plant or animal and are commonly grouped into four families: the LAGLIDADG family, the GIY-YIG family, the His-Cyst box family and the HNH family. Exemplary homing endonucleases include I-SceI, I-CeuI, PI-PspI, PI-Sce, I-SceIV, I-CsmI, I-PanI, I-SceII, I-PpoI, I-SceIII, I-CreI, I-TevI, I-TevII and I-TevIII.

[00117] Methods of the present disclosure may be employed to introduce or improve one or more of a variety of desirable traits. Examples of traits that may introduced or improved include: root biomass, root length, height, shoot length, leaf number, water use efficiency, overall biomass, yield, fruit size, grain size, photosynthesis rate, tolerance to drought, heat tolerance, salt tolerance, resistance to nematode stress, resistance to a fungal pathogen, resistance to a bacterial pathogen, resistance to a viral pathogen, level of a metabolite, and proteome expression. The desirable traits, including height, overall biomass, root and/or shoot biomass, seed germination, seedling survival, photosynthetic efficiency, transpiration rate, seed/fruit number or mass, plant grain or fruit yield, leaf chlorophyll content, photosynthetic rate, root length, or any combination thereof, can be used to measure growth, and compared with the growth rate of reference agricultural plants (e.g., plants without the improved traits) grown under identical conditions. A preferred trait to be introduced or improved is nitrogen fixation, as described herein. In some

cases, a plant resulting from the methods described herein exhibits a difference in the trait that is at least about 5% greater, for example at least about 5%, at least about 8%, at least about 10%, at least about 15%, at least about 20%, at least about 25%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 75%, at least about 80%, at least about 80%, at least about 90%, or at least 100%, at least about 200%, at least about 300%, at least about 400% or greater than a reference agricultural plant grown under the same conditions in the soil.

[00118] The trait to be improved may be assessed under conditions including the application of one or more biotic or abiotic stressors. Examples of stressors include abiotic stresses (such as heat stress, salt stress, drought stress, cold stress, and low nutrient stress) and biotic stresses (such as nematode stress, insect herbivory stress, fungal pathogen stress, bacterial pathogen stress, and viral pathogen stress).

[00119] The trait improved by methods and compositions of the present disclosure may be nitrogen fixation, including in a plant not previously capable of nitrogen fixation. In some cases, bacteria isolated according to a method described herein produce 1% or more (e.g. 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 15%, 20%, or more) of a plant's nitrogen, which may represent an increase in nitrogen fixation capability of at least 2-fold (e.g. 3-fold, 4-fold, 5-fold, 6-fold, 7-fold, 8-fold, 9-fold, 10-fold, 20-fold, 50-fold, 100-fold, 1000-fold, or more) as compared to bacteria isolated from the first plant before introducing any genetic variation. In some cases, the bacteria produce 5% or more of a plant's nitrogen. The desired level of nitrogen fixation may be achieved after repeating the steps of introducing genetic variation, exposure to a plurality of plants, and isolating bacteria from plants with an improved trait one or more times (e.g. 1, 2, 3, 4, 5, 10, 15, 25, or more times). In some cases, enhanced levels of nitrogen fixation are achieved in the presence of fertilizer supplemented with glutamine, ammonia, or other chemical source of nitrogen. Methods for assessing degree of nitrogen fixation are known, examples of which are described herein.

Nitrogen Fixation

[00120] Described herein are methods of increasing nitrogen fixation in a plant, comprising exposing the plant to bacteria comprising one or more genetic variations introduced into one or more genes regulating nitrogen fixation, wherein the bacteria produce 1% or more of nitrogen in the plant (e.g. 2%, 5%, 10%, or more), which may represent a nitrogen-fixation capability of at least 2-fold as compared to the plant in the absence of the bacteria. The bacteria may produce the nitrogen in the presence of fertilizer supplemented with glutamine or ammonia. Genetic

variations can be any genetic variation described herein, including examples provided above, in any number and any combination. The genetic variation may be introduced into a gene selected from the group consisting of *nifA*, *nifL*, *ntrB*, *ntrC*, glutamine synthetase, *glnA*, *glnB*, *glnK*, *draT*, *amtB*, glutaminase, *glnD*, *glnE*, *nifJ*, *nifH*, *nifD*, *nifK*, *nifY*, *nifE*, *nifN*, *nifU*, *nifS*, *nifV*, *nifW*, *nifZ*, *nifM*, *nifF*, *nifB*, and *nifQ*. The genetic variation may be a mutation that results in one or more of: increased expression or activity of *nifA* or glutaminase; decreased expression or activity of *nifL*, *ntrB*, glutamine synthetase, *glnB*, *glnK*, *draT*, *amtB*; decreased adenylyl-removing activity of *GlnE*; or decreased uridylyl-removing activity of *GlnD*. The genetic variation introduced into one or more bacteria of the methods disclosed herein may be a knock-out mutation or it may abolish a regulatory sequence of a target gene, or it may comprise insertion of a heterologous regulatory sequence, for example, insertion of a regulatory sequence found within the genome of the same bacterial species or genus. The regulatory sequence can be chosen based on the expression level of a gene in a bacterial culture or within plant tissue. The genetic variation may be produced by chemical mutagenesis. The plants grown in step (c) may be exposed to biotic or abiotic stressors.

[00121] The amount of nitrogen fixation that occurs in the plants described herein may be measured in several ways, for example by an acetylene-reduction (AR) assay. An acetylene-reduction assay can be performed *in vitro* or *in vivo*. Evidence that a particular bacterium is providing fixed nitrogen to a plant can include: 1) total plant N significantly increases upon inoculation, preferably with a concomitant increase in N concentration in the plant; 2) nitrogen deficiency symptoms are relieved under N-limiting conditions upon inoculation (which should include an increase in dry matter); 3) N_2 fixation is documented through the use of an ^{15}N approach (which can be isotope dilution experiments, $^{15}N_2$ reduction assays, or ^{15}N natural abundance assays); 4) fixed N is incorporated into a plant protein or metabolite; and 5) all of these effects are not be seen in uninoculated plants or in plants inoculated with a mutant of the inoculum strain.

[00122] The wild-type nitrogen fixation regulatory cascade can be represented as a digital logic circuit where the inputs O_2 and NH_4^+ pass through a NOR gate, the output of which enters an AND gate in addition to ATP. In some embodiments, the methods disclosed herein disrupt the influence of NH_4^+ on this circuit, at multiple points in the regulatory cascade, so that microbes can produce nitrogen even in fertilized fields. However, the methods disclosed herein also envision altering the impact of ATP or O_2 on the circuitry, or replacing the circuitry with other regulatory cascades in the cell, or altering genetic circuits other than nitrogen fixation. Gene clusters can be re-engineered to generate functional products under the control of a heterologous

regulatory system. By eliminating native regulatory elements outside of, and within, coding sequences of gene clusters, and replacing them with alternative regulatory systems, the functional products of complex genetic operons and other gene clusters can be controlled and/or moved to heterologous cells, including cells of different species other than the species from which the native genes were derived. Once re-engineered, the synthetic gene clusters can be controlled by genetic circuits or other inducible regulatory systems, thereby controlling the products' expression as desired. The expression cassettes can be designed to act as logic gates, pulse generators, oscillators, switches, or memory devices. The controlling expression cassette can be linked to a promoter such that the expression cassette functions as an environmental sensor, such as an oxygen, temperature, touch, osmotic stress, membrane stress, or redox sensor.

[00123] As an example, the *nifL*, *nifA*, *nifT*, and *nifX* genes can be eliminated from the *nif* gene cluster. Synthetic genes can be designed by codon randomizing the DNA encoding each amino acid sequence. Codon selection is performed, specifying that codon usage be as divergent as possible from the codon usage in the native gene. Proposed sequences are scanned for any undesired features, such as restriction enzyme recognition sites, transposon recognition sites, repetitive sequences, sigma 54 and sigma 70 promoters, cryptic ribosome binding sites, and rho independent terminators. Synthetic ribosome binding sites are chosen to match the strength of each corresponding native ribosome binding site, such as by constructing a fluorescent reporter plasmid in which the 150 bp surrounding a gene's start codon (from -60 to +90) is fused to a fluorescent gene. This chimera can be expressed under control of the Ptac promoter, and fluorescence measured via flow cytometry. To generate synthetic ribosome binding sites, a library of reporter plasmids using 150 bp (-60 to +90) of a synthetic expression cassette is generated. Briefly, a synthetic expression cassette can consist of a random DNA spacer, a degenerate sequence encoding an RBS library, and the coding sequence for each synthetic gene. Multiple clones are screened to identify the synthetic ribosome binding site that best matched the native ribosome binding site. Synthetic operons that consist of the same genes as the native operons are thus constructed and tested for functional complementation. A further exemplary description of synthetic operons is provided in US20140329326.

Bacterial Species

[00124] Microbes useful in the methods and compositions disclosed herein can be obtained by extracting microbes from surfaces or tissues of native plants; grinding seeds to isolate microbes; planting seeds in diverse soil samples and recovering microbes from tissues; or inoculating plants with exogenous microbes and determining which microbes appear in plant tissues. Non-

limiting examples of plant tissues include a seed, seedling, leaf, cutting, plant, bulb or tuber. In some cases, bacteria are isolated from a seed. The parameters for processing samples may be varied to isolate different types of associative microbes, such as rhizospheric, epiphytes, or endophytes. Bacteria may also be sourced from a repository, such as environmental strain collections, instead of initially isolating from a first plant. The microbes can be genotyped and phenotyped, via sequencing the genomes of isolated microbes; profiling the composition of communities *in planta*; characterizing the transcriptomic functionality of communities or isolated microbes; or screening microbial features using selective or phenotypic media (e.g., nitrogen fixation or phosphate solubilization phenotypes). Selected candidate strains or populations can be obtained via sequence data; phenotype data; plant data (e.g., genome, phenotype, and/or yield data); soil data (e.g., pH, N/P/K content, and/or bulk soil biotic communities); or any combination of these.

[00125] The bacteria and methods of producing bacteria described herein may apply to bacteria able to self-propagate efficiently on the leaf surface, root surface, or inside plant tissues without inducing a damaging plant defense reaction, or bacteria that are resistant to plant defense responses. The bacteria described herein may be isolated by culturing a plant tissue extract or leaf surface wash in a medium with no added nitrogen. However, the bacteria may be unculturable, that is, not known to be culturable or difficult to culture using standard methods known in the art. The bacteria described herein may be an endophyte or an epiphyte or a bacterium inhabiting the plant rhizosphere (rhizospheric bacteria). The bacteria obtained after repeating the steps of introducing genetic variation, exposure to a plurality of plants, and isolating bacteria from plants with an improved trait one or more times (e.g. 1, 2, 3, 4, 5, 10, 15, 25, or more times) may be endophytic, epiphytic, or rhizospheric. Endophytes are organisms that enter the interior of plants without causing disease symptoms or eliciting the formation of symbiotic structures, and are of agronomic interest because they can enhance plant growth and improve the nutrition of plants (e.g., through nitrogen fixation). The bacteria can be a seed-borne endophyte. Seed-borne endophytes include bacteria associated with or derived from the seed of a grass or plant, such as a seed-borne bacterial endophyte found in mature, dry, undamaged (e.g., no cracks, visible fungal infection, or prematurely germinated) seeds. The seed-borne bacterial endophyte can be associated with or derived from the surface of the seed; alternatively, or in addition, it can be associated with or derived from the interior seed compartment (e.g., of a surface-sterilized seed). In some cases, a seed-borne bacterial endophyte is capable of replicating within the plant tissue, for example, the interior of the seed. Also, in some cases, the seed-borne bacterial endophyte is capable of surviving desiccation.

[00126] The bacterial isolated according to methods of the disclosure can comprise a plurality of different bacterial taxa in combination. By way of example, the bacteria may include Proteobacteria (such as *Pseudomonas*, *Enterobacter*, *Stenotrophomonas*, *Burkholderia*, *Rhizobium*, *Herbaspirillum*, *Pantoea*, *Serratia*, *Rahnella*, *Azospirillum*, *Azorhizobium*, *Azotobacter*, *Duganella*, *Delftia*, *Bradyrhizobium*, *Sinorhizobium* and *Halomonas*), Firmicutes (such as *Bacillus*, *Paenibacillus*, *Lactobacillus*, *Mycoplasma*, and *Acetabacterium*), and Actinobacteria (such as *Streptomyces*, *Rhodococcus*, *Microbacterium*, and *Curtobacterium*). Bacteria that can be produced by the methods disclosed herein include *Azotobacter* sp., *Bradyrhizobium* sp., *Klebsiella* sp., and *Sinorhizobium* sp. The bacteria may be selected from the group consisting of: *Azotobacter vinelandii*, *Bradyrhizobium japonicum*, *Klebsiella pneumoniae*, and *Sinorhizobium meliloti*. The bacteria may be of the genus *Enterobacter* and *Rahnella*.

[00127] The bacteria may be obtained from any general terrestrial environment, including its soils, plants, fungi, animals (including invertebrates) and other biota, including the sediments, water and biota of lakes and rivers; from the marine environment, its biota and sediments (for example, sea water, marine muds, marine plants, marine invertebrates (for example, sponges), marine vertebrates (for example, fish)); the terrestrial and marine geosphere (regolith and rock, for example, crushed subterranean rocks, sand and clays); the cryosphere and its meltwater; the atmosphere (for example, filtered aerial dusts, cloud and rain droplets); urban, industrial and other man-made environments (for example, accumulated organic and mineral matter on concrete, roadside gutters, roof surfaces, and road surfaces).

[00128] The plants from which the bacteria are obtained may be a plant having one or more desirable traits, for example a plant which naturally grows in a particular environment or under certain conditions of interest. By way of example, a certain plant may naturally grow in sandy soil or sand of high salinity, or under extreme temperatures, or with little water, or it may be resistant to certain pests or disease present in the environment, and it may be desirable for a commercial crop to be grown in such conditions, particularly if they are, for example, the only conditions available in a particular geographic location. By way of further example, the bacteria may be collected from commercial crops grown in such environments, or more specifically from individual crop plants best displaying a trait of interest amongst a crop grown in any specific environment: for example the fastest-growing plants amongst a crop grown in saline-limiting soils, or the least damaged plants in crops exposed to severe insect damage or disease epidemic, or plants having desired quantities of certain metabolites and other compounds, including fibre content, oil content, and the like, or plants displaying desirable colors, taste or smell. The

bacteria may be collected from a plant of interest or any material occurring in the environment of interest, including fungi and other animal and plant biota, soil, water, sediments, and other elements of the environment as referred to previously.

[00129] The bacteria may be isolated from plant tissue. This isolation can occur from any appropriate tissue in the plant, including for example root, stem and leaves, and plant reproductive tissues. By way of example, conventional methods for isolation from plants typically include the sterile excision of the plant material of interest (e.g. root or stem lengths, leaves), surface sterilization with an appropriate solution (e.g. 2% sodium hypochlorite), after which the plant material is placed on nutrient medium for microbial growth. Alternatively, the surface-sterilized plant material can be crushed in a sterile liquid (usually water) and the liquid suspension, including small pieces of the crushed plant material spread over the surface of a suitable solid agar medium, or media, which may or may not be selective (e.g. contain only phytic acid as a source of phosphorus). This approach is especially useful for bacteria which form isolated colonies and can be picked off individually to separate plates of nutrient medium, and further purified to a single species by well-known methods. Alternatively, the plant root or foliage samples may not be surface sterilized but only washed gently thus including surface-dwelling epiphytic microorganisms in the isolation process, or the epiphytic microbes can be isolated separately, by imprinting and lifting off pieces of plant roots, stem or leaves onto the surface of an agar medium and then isolating individual colonies as above. This approach is especially useful for bacteria, for example. Alternatively, the roots may be processed without washing off small quantities of soil attached to the roots, thus including microbes that colonize the plant rhizosphere. Otherwise, soil adhering to the roots can be removed, diluted and spread out onto agar of suitable selective and non-selective media to isolate individual colonies of rhizospheric bacteria.

[00130] Biologically pure cultures of *Rahnella aquatilis* and *Enterobacter sacchari* were deposited on July 14, 2015 with the American Type Culture Collection (ATCC; an International Depositary Authority), Manassas, VA, USA, and assigned ATTC Patent Deposit Designation numbers PTA-122293 and PTA-122294, respectively. These deposits were made under the provisions of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purpose of Patent Procedure and the Regulations (Budapest Treaty).

Compositions

[00131] Compositions comprising bacteria or bacterial populations produced according to methods described herein and/or having characteristics as described herein may also be used to

improve plant traits. The compositions comprising bacterial populations may be coated on a surface of a seed, and may be in liquid form. The compositions include seed coatings for commercially important agricultural crops, for example, sorghum, canola, tomato, strawberry, barley, rice, maize, and wheat. The compositions may also be sprayed on the plant aerial parts, or applied to the roots by inserting into furrows in which the plant seeds are planted, watering to the soil, or dipping the roots in a suspension of the composition. The compositions may be dehydrated in a suitable manner that maintains cell viability and the ability to artificially inoculate and colonize host plants. The bacterial species may be present in the compositions at a concentration of between 10^8 to 10^{10} CFU/ml. The compositions may be supplemented with trace metal ions, such as molybdenum ions, iron ions, manganese ions, or combinations of these ions. The concentration of ions in the compositions described herein may be between about 0.1 mM and about 50 mM. The compositions may also be formulated with a carrier, such as beta-glucan, carboxymethyl cellulose (CMC), bacterial extracellular polymeric substance (EPS), sugar, animal milk, or other suitable carriers. Alternatively, peat or planting materials can be used as a carrier, or biopolymers in which the composition is entrapped in the biopolymer can be used as a carrier. The compositions comprising the bacterial populations described herein can improve plant traits, such as promoting plant growth, maintaining high chlorophyll content in leaves, increasing fruit or seed numbers, and increasing fruit or seed unit weight.

[00132] The compositions comprising the bacterial populations described herein may be coated onto the surface of a seed. As such, compositions comprising a seed coated with one or more bacteria described herein are also contemplated. The seed coating can be formed by mixing the bacterial population with a porous, chemically inert granular carrier. Alternatively, the compositions may be inserted directly into the furrows into which the seed is planted or sprayed onto the plant leaves or applied by dipping the roots into a suspension of the composition. An effective amount of the composition can be used to populate the sub-soil region adjacent to the roots of the plant with viable bacterial growth, or populate the leaves of the plant with viable bacterial growth. In general, an effective amount is an amount sufficient to result in plants with improved traits (e.g. a desired level of nitrogen fixation).

[00133] Bacterial compositions described herein can be formulated using an agriculturally acceptable carrier. The formulation useful for these embodiments may include at least one member selected from the group consisting of a tackifier, a microbial stabilizer, a fungicide, an antibacterial agent, an herbicide, a nematicide, an insecticide, a plant growth regulator, a fertilizer, a rodenticide, a dessicant, and a nutrient. For example, any of the compositions described herein can include an agriculturally acceptable carrier (e.g., one or more of a fertilizer

such as a non-naturally occurring fertilizer, an adhesion agent such as a non-naturally occurring adhesion agent, and a pesticide such as a non-naturally occurring pesticide). A non-naturally occurring adhesion agent can be, for example, a polymer, copolymer, or synthetic wax. For example, any of the coated seeds, seedlings, or plants described herein can contain such an agriculturally acceptable carrier in the seed coating. In any of the compositions or methods described herein, an agriculturally acceptable carrier can be or can include a non-naturally occurring compound (e.g., a non-naturally occurring fertilizer, a non-naturally occurring adhesion agent such as a polymer, copolymer, or synthetic wax, or a non-naturally occurring pesticide). Non-limiting examples of agriculturally acceptable carriers are described below. Additional examples of agriculturally acceptable carriers are known in the art.

[00134] In some cases, bacteria are mixed with an agriculturally acceptable carrier. The carrier can be a solid carrier or liquid carrier, and in various forms including microspheres, powders, emulsions and the like. The carrier may be any one or more of a number of carriers that confer a variety of properties, such as increased stability, wettability, or dispersability. Wetting agents such as natural or synthetic surfactants, which can be nonionic or ionic surfactants, or a combination thereof can be included in the composition. Water-in-oil emulsions can also be used to formulate a composition that includes the isolated bacteria (see, for example, U.S. Patent No. 7,485,451). Suitable formulations that may be prepared include wettable powders, granules, gels, agar strips or pellets, thickeners, and the like, microencapsulated particles, and the like, liquids such as aqueous flowables, aqueous suspensions, water-in-oil emulsions, etc. The formulation may include grain or legume products, for example, ground grain or beans, broth or flour derived from grain or beans, starch, sugar, or oil.

[00135] In some embodiments, the agricultural carrier may be soil or a plant growth medium. Other agricultural carriers that may be used include water, fertilizers, plant-based oils, humectants, or combinations thereof. Alternatively, the agricultural carrier may be a solid, such as diatomaceous earth, loam, silica, alginate, clay, bentonite, vermiculite, seed cases, other plant and animal products, or combinations, including granules, pellets, or suspensions. Mixtures of any of the aforementioned ingredients are also contemplated as carriers, such as but not limited to, pesta (flour and kaolin clay), agar or flour-based pellets in loam, sand, or clay, etc. Formulations may include food sources for the bacteria, such as barley, rice, or other biological materials such as seed, plant parts, sugar cane bagasse, hulls or stalks from grain processing, ground plant material or wood from building site refuse, sawdust or small fibers from recycling of paper, fabric, or wood.

[00136] For example, a fertilizer can be used to help promote the growth or provide nutrients to a seed, seedling, or plant. Non-limiting examples of fertilizers include nitrogen, phosphorous, potassium, calcium, sulfur, magnesium, boron, chloride, manganese, iron, zinc, copper, molybdenum, and selenium (or a salt thereof). Additional examples of fertilizers include one or more amino acids, salts, carbohydrates, vitamins, glucose, NaCl, yeast extract, $\text{NH}_4\text{H}_2\text{PO}_4$, $(\text{NH}_4)_2\text{SO}_4$, glycerol, valine, L-leucine, lactic acid, propionic acid, succinic acid, malic acid, citric acid, KH tartrate, xylose, lyxose, and lecithin. In one embodiment, the formulation can include a tackifier or adherent (referred to as an adhesive agent) to help bind other active agents to a substance (e.g., a surface of a seed). Such agents are useful for combining bacteria with carriers that can contain other compounds (e.g., control agents that are not biologic), to yield a coating composition. Such compositions help create coatings around the plant or seed to maintain contact between the microbe and other agents with the plant or plant part. In one embodiment, adhesives are selected from the group consisting of: alginate, gums, starches, lecithins, formononetin, polyvinyl alcohol, alkali formononetinate, hesperetin, polyvinyl acetate, cephalins, Gum Arabic, Xanthan Gum, Mineral Oil, Polyethylene Glycol (PEG), Polyvinyl pyrrolidone (PVP), Arabino-galactan, Methyl Cellulose, PEG 400, Chitosan, Polyacrylamide, Polyacrylate, Polyacrylonitrile, Glycerol, Triethylene glycol, Vinyl Acetate, Gellan Gum, Polystyrene, Polyvinyl, Carboxymethyl cellulose, Gum Ghatti, and polyoxyethylene-polyoxybutylene block copolymers.

[00137] In some embodiments, the adhesives can be, e.g. a wax such as carnauba wax, beeswax, Chinese wax, shellac wax, spermaceti wax, candelilla wax, castor wax, ouricury wax, and rice bran wax, a polysaccharide (e.g., starch, dextrans, maltodextrins, alginate, and chitosans), a fat, oil, a protein (e.g., gelatin and zeins), gum arabic, and shellacs. Adhesive agents can be non-naturally occurring compounds, e.g., polymers, copolymers, and waxes. For example, non-limiting examples of polymers that can be used as an adhesive agent include: polyvinyl acetates, polyvinyl acetate copolymers, ethylene vinyl acetate (EVA) copolymers, polyvinyl alcohols, polyvinyl alcohol copolymers, celluloses (e.g., ethylcelluloses, methylcelluloses, hydroxymethylcelluloses, hydroxypropylcelluloses, and carboxymethylcelluloses), polyvinylpyrrolidones, vinyl chloride, vinylidene chloride copolymers, calcium lignosulfonates, acrylic copolymers, polyvinylacrylates, polyethylene oxide, acylamide polymers and copolymers, polyhydroxyethyl acrylate, methylacrylamide monomers, and polychloroprene.

[00138] In some examples, one or more of the adhesion agents, anti-fungal agents, growth regulation agents, and pesticides (e.g., insecticide) are non-naturally occurring compounds (e.g.,

in any combination). Additional examples of agriculturally acceptable carriers include dispersants (e.g., polyvinylpyrrolidone/vinyl acetate PVPIVA S-630), surfactants, binders, and filler agents.

[00139] The formulation can also contain a surfactant. Non-limiting examples of surfactants include nitrogen-surfactant blends such as Prefer 28 (Cenex), Surf-N(US), Inhance (Brandt), P-28 (Wilfarm) and Patrol (Helena); esterified seed oils include Sun-It II (AmCy), MSO (UAP), Scoil (Agasco), Hasten (Wilfarm) and Mes-100 (Drexel); and organo-silicone surfactants include Silwet L77 (UAP), Silikin (Terra), Dyne-Amic (Helena), Kinetic (Helena), Sylgard 309 (Wilbur-Ellis) and Century (Precision). In one embodiment, the surfactant is present at a concentration of between 0.01% v/v to 10% v/v. In another embodiment, the surfactant is present at a concentration of between 0.1% v/v to 1% v/v.

[00140] In certain cases, the formulation includes a microbial stabilizer. Such an agent can include a desiccant, which can include any compound or mixture of compounds that can be classified as a desiccant regardless of whether the compound or compounds are used in such concentrations that they in fact have a desiccating effect on a liquid inoculant. Such desiccants are ideally compatible with the bacterial population used, and should promote the ability of the microbial population to survive application on the seeds and to survive desiccation. Examples of suitable desiccants include one or more of trehalose, sucrose, glycerol, and Methylene glycol. Other suitable desiccants include, but are not limited to, non reducing sugars and sugar alcohols (e.g., mannitol or sorbitol). The amount of desiccant introduced into the formulation can range from about 5% to about 50% by weight/volume, for example, between about 10% to about 40%, between about 15% to about 35%, or between about 20% to about 30%. In some cases, it is advantageous for the formulation to contain agents such as a fungicide, an antibacterial agent, an herbicide, a nematocide, an insecticide, a plant growth regulator, a rodenticide, or a nutrient. Non-limiting examples of growth regulators include brassinosteroids, cytokinines (e.g., kinetin and zeatin), auxins (e.g., indolylacetic acid and indolylacetyl aspartate), flavonoids and isoflavanoids (e.g., formononetin and diosmetin), phytoalexins (e.g., glyceolline), and phytoalexin-inducing oligosaccharides (e.g., pectin, chitin, chitosan, polygalacturonic acid, and oligogalacturonic acid), and gibberellins. Such agents are ideally compatible with the agricultural seed or seedling onto which the formulation is applied (e.g., it should not be deleterious to the growth or health of the plant). Furthermore, the agent is ideally one which does not cause safety concerns for human, animal or industrial use (e.g., no safety issues, or the compound is sufficiently labile that the commodity plant product derived from the plant contains negligible amounts of the compound).

[00141] In the liquid form, for example, solutions or suspensions, bacterial populations can be mixed or suspended in water or in aqueous solutions. Suitable liquid diluents or carriers include water, aqueous solutions, petroleum distillates, or other liquid carriers.

[00142] Solid compositions can be prepared by dispersing the bacterial populations in and on an appropriately divided solid carrier, such as peat, wheat, bran, vermiculite, clay, talc, bentonite, diatomaceous earth, fuller's earth, pasteurized soil, and the like. When such formulations are used as wettable powders, biologically compatible dispersing agents such as non-ionic, anionic, amphoteric, or cationic dispersing and emulsifying agents can be used.

[00143] The solid carriers used upon formulation include, for example, mineral carriers such as kaolin clay, pyrophyllite, bentonite, montmorillonite, diatomaceous earth, acid white soil, vermiculite, and perlite, and inorganic salts such as ammonium sulfate, ammonium phosphate, ammonium nitrate, urea, ammonium chloride, and calcium carbonate. Also, organic fine powders such as wheat flour, wheat bran, and rice bran may be used. The liquid carriers include vegetable oils such as soybean oil and cottonseed oil, glycerol, ethylene glycol, polyethylene glycol, propylene glycol, polypropylene glycol, etc.

Plant Species

[00144] The methods and bacteria described herein are suitable for any of a variety of plants, such as plants in the genera *Hordeum*, *Oryza*, *Zea*, and *Triticeae*. Other non-limiting examples of suitable plants include mosses, lichens, and algae. In some cases, the plants have economic, social and/or environmental value, such as food crops, fiber crops, oil crops, plants in the forestry or pulp and paper industries, feedstock for biofuel production and/or ornamental plants. Non-limiting examples of crop plants include maize, rice, wheat, barley, sorghum, millet, oats, rye triticale, buckwheat, sweet corn, sugar cane, onions, tomatoes, strawberries, and asparagus.

[00145] Plants that may be obtained or improved using the methods and composition disclosed herein also include pineapple, banana, coconut, lily, and grass; and dicotyledonous plants, such as, for example, peas, alfalfa, tomatillo, melon, chickpea, chicory, clover, kale, lentil, soybean, tobacco, potato, sweet potato, radish, cabbage, rape, apple trees, grape, cotton, sunflower, thale cress, canola, citrus (including orange, mandarin, kumquat, lemon, lime, grapefruit, tangerine, tangelo, citron, and pomelo), pepper, bean, and lettuce.

[00146] In some cases, the plant to be improved is not readily amenable to experimental conditions. For example, a crop plant may take too long to grow enough to practically assess an improved trait serially over multiple iterations. Accordingly, a first plant from which bacteria

are initially isolated, and/or the plurality of plants to which genetically manipulated bacteria are applied may be a model plant, such as a plant more amenable to evaluation under desired conditions. Non-limiting examples of model plants include *Setaria*, *Brachypodium*, and *Arabidopsis*. Ability of bacteria isolated according to a method of the disclosure using a model plant may then be applied to a plant of another type (e.g. a crop plant) to confirm conferral of the improved trait.

[00147] Traits that may be improved by the methods disclosed herein include any observable characteristic of the plant, including, for example, growth rate, height, weight, color, taste, smell, changes in the production of one or more compounds by the plant (including for example, metabolites, proteins, drugs, carbohydrates, oils, and any other compounds). Selecting plants based on genotypic information is also envisaged (for example, including the pattern of plant gene expression in response to the bacteria, or identifying the presence of genetic markers, such as those associated with increased nitrogen fixation). Plants may also be selected based on the absence, suppression or inhibition of a certain feature or trait (such as an undesirable feature or trait) as opposed to the presence of a certain feature or trait (such as a desirable feature or trait).

EXAMPLES

[00148] The examples provided herein describe methods of bacterial isolation, bacterial and plant analysis, and plant trait improvement. The examples are for illustrative purposes only and are not to be construed as limiting in any way.

Example 1: Isolation of Microbes from Plant Tissue

[00149] Topsoil was obtained from various agricultural areas in central California. Twenty soils with diverse texture characteristics were collected, including heavy clay, peaty clay loam, silty clay, and sandy loam. Seeds of various field corn, sweet corn, heritage corn and tomato were planted into each soil, as shown in **Table 1**.

Crop Type	Field Corn	Sweet Corn	Heritage Corn	Tomato
Varieties	Mo17	Ferry-Morse 'Golden Cross Bantam T-51'	Victory Seeds 'Moseby Prolific'	Ferry-Morse Roma VF
	B73	Ferry-Morse 'Silver Queen Hybrid'	Victory Seeds 'Reid's Yellow Dent'	Stover Roma
	DKC 66-40	Ferry-Morse 'Sugar Dots'	Victory Seeds 'Hickory King'	Totally Tomatoes 'Micro Tom Hybrid'
	DKC 67-07			Heinz 1015

	DKC 70-01			Heinz 2401
				Heinz 3402
				Heinz 5508
				Heinz 5608
				Heinz 8504

Table 1: Crop Type and Varieties planted into soil with diverse characteristics

[00150] Plants were uprooted after 2-4 weeks of growth and excess soil on root surfaces was removed with deionized water. Following soil removal, plants were surface sterilized with bleach and rinsed vigorously in sterile water. A cleaned, 1 cm section of root was excised from the plant and placed in a phosphate buffered saline solution containing 3 mm steel beads. A slurry was generated by vigorous shaking of the solution with a Qiagen TissueLyser II.

[00151] The root and saline slurry was diluted and inoculated onto various types of growth media to isolate rhizospheric, endophytic, epiphytic, and other plant-associated microbes. R2A and Nfb agar media were used to obtain single colonies, and semisolid Nfb media slants were used to obtain populations of nitrogen fixing bacteria. After 2-4 weeks incubation in semi-solid Nfb media slants, microbial populations were collected and streaked to obtain single colonies on R2A agar, as shown in **Figure 1A-B**. Single colonies were resuspended in a mixture of R2A and glycerol, subjected to PCR analysis, and frozen at -80°C for later analysis. Approximately 1,000 single colonies were obtained and designated “isolated microbes.”

[00152] Isolates were then subjected to a colony PCR screen to detect the presence of the *nifH* gene in order to identify diazotrophs. The previously-described primer set Ueda 19F/388R, which has been shown to detect over 90% of diazotrophs in screens, was used to probe the presence of the *nif* cluster in each isolate (Ueda et al. 1995; J. Bacteriol. 177: 1414–1417). Single colonies of purified isolates were picked, resuspended in PBS, and used as a template for colony PCR, as shown in **Figure 2**. Colonies of isolates that gave positive PCR bands were re-streaked, and the colony PCR and re-streaking process was repeated twice to prevent false positive identification of diazotrophs. Purified isolates were then designated “candidate microbes.”

Example 2: Characterization of Isolated Microbes

Sequencing, Analysis and Phylogenetic Characterization

[00153] Sequencing of 16S rDNA with the 515f-806r primer set was used to generate preliminary phylogenetic identities for isolated and candidate microbes (see e.g. Vernon et al.;

BMC Microbiol. 2002 Dec 23;2:39.). The microbes comprise diverse genera including: Enterobacter, Burkholderia, Klebsiella, Bradyrhizobium, Rahnella, Xanthomonas, Raoultella, Pantoea, Pseudomonas, Brevundimonas, Agrobacterium, and Paenibacillus, as shown in **Table 2**.

Genus	Isolates	Genus	Isolates
<i>Achromobacter</i>	7	<i>Paenibacillus</i>	1
<i>Agrobacterium</i>	117	<i>Paenisporsarcina</i>	3
<i>Agromyces</i>	1	<i>Pantoea</i>	14
<i>Alicyclobacillus</i>	1	<i>Pedobacter</i>	16
<i>Asticcacaulis</i>	6	<i>Pimelobacter</i>	2
<i>Bacillus</i>	131	<i>Pseudomonas</i>	212
<i>Bradyrhizobium</i>	2	<i>Rhizobium</i>	4
<i>Brevibacillus</i>	2	<i>Rhodoferax</i>	1
<i>Burkholderia</i>	2	<i>Sphingobacterium</i>	13
<i>Caulobacter</i>	17	<i>Sphingobium</i>	23
<i>Chryseobacterium</i>	42	<i>Sphingomonas</i>	3
<i>Comamonas</i>	1	<i>Sphingopyxis</i>	1
<i>Dyadobacter</i>	2	<i>Stenotrophomonas</i>	59
<i>Flavobacterium</i>	46	<i>Streptococcus</i>	3
<i>Halomonas</i>	3	<i>Variovorax</i>	37
<i>Leptothrix</i>	3	<i>Xylanimicrobium</i>	1
<i>Lysobacter</i>	2	unidentified	75
<i>Neisseria</i>	13		

Table 2: Diversity of microbes isolated from tomato plants as determined by deep 16S rDNA sequencing.

[00154] Subsequently, the genomes of 39 candidate microbes were sequenced using Illumina Miseq platform. Genomic DNA from pure cultures was extracted using the QIAmp DNA mini kit (QIAGEN), and total DNA libraries for sequencing were prepared through a third party vendor (SeqMatic, Hayward). Genome assembly was then carried out via the A5 pipeline (Tritt et al. 2012; PLoS One 7(9):e42304). Genes were identified and annotated, and those related to regulation and expression of nitrogen fixation were noted as targets for mutagenesis.

Transcriptomic Profiling of Candidate Microbes

[00155] Transcriptomic profiling of strain CI010 was performed to identify promoters that are active in the presence of environmental nitrogen. Strain CI010 was cultured in a defined, nitrogen-free media supplemented with 10 mM glutamine. Total RNA was extracted from these cultures (QIAGEN RNeasy kit) and subjected to RNAseq sequencing via Illumina HiSeq (SeqMatic, Fremont CA). Sequencing reads were mapped to CI010 genome data using Geneious, and highly expressed genes under control of proximal transcriptional promoters were

identified. Tables 3A-C lists genes and their relative expression level as measured through RNASeq sequencing of total RNA. Sequences of the proximal promoters were recorded for use in mutagenesis of *nif* pathways, nitrogen utilization related pathways, or other genes with a desired expression level.

Assessment of Genetic Tractability

[00156] Candidate microbes were characterized based on transformability and genetic tractability. First, optimal carbon source utilization was determined by growth on a small panel of relevant media as well as a growth curve in both nitrogen-free and rich media. Second, the natural antibiotic resistance of each strain was determined through spot-plating and growth in liquid culture containing a panel of antibiotics used as selective markers for mutagenesis. Third, each strain was tested for its transformability through electroporation of a collection of plasmids. The plasmid collection comprises the combinatorial expansion of seven origins of replication, *i.e.*, p15a, pSC101, CloDF, colA, RK2, pBBR1, and pRO1600 and four antibiotic resistance markers, *i.e.*, CmR, KmR, SpecR, and TetR. This systematic evaluation of origin and resistance marker compatibility was used to identify vectors for plasmid-based mutagenesis in candidate microbes.

Example 3: Mutagenesis of Candidate Microbes

Lambda-Red Mediated Knockouts

[00157] Several mutants of candidate microbes were generated using the plasmid pKD46 or a derivative containing a kanamycin resistance marker (Datsenko et al. 2000; PNAS 97(12): 6640-6645). Knockout cassettes were designed with 250bp homology flanking the target gene and generated via overlap extension PCR. Candidate microbes were transformed with pKD46, cultured in the presence of arabinose to induce Lambda-Red machinery expression, prepped for electroporation, and transformed with the knockout cassettes to produce candidate mutant strains. Four candidate microbes and one laboratory strain, *Klebsiella oxytoca* M5A1, were used to generate thirteen candidate mutants of the nitrogen fixation regulatory genes *nifL*, *glnB*, and *amtB*, as shown in **Table 4**.

Strain	nifL	glnB	amtB
M5A1	X	X	X
CI006	X	X	X
CI010	X	X	X
CI019	X	X	
CI028	X	X	

Table 4: List of single knockout mutants created through Lambda-red mutagenesis

Oligo-Directed Mutagenesis with Cas9 Selection

[00158] Oligo-directed mutagenesis was used to target genomic changes to the *rpoB* gene in *E. coli* DH10B, and mutants were selected with a CRISPR-Cas system. A mutagenic oligo (ss1283:

“G*T*T*G*ATCAGACCGATGTTTCGGACCTTCcaagGTTTCGATCGGACATACGCGACCGTAGTGGGTCGGGTGTACGTCTCGAACTTCAAAGCC”, where * denotes phosphorothioate bond) was designed to confer rifampicin resistance through a 4-bp mutation to the *rpoB* gene. Cells containing a plasmid encoding Cas9 were induced for Cas9 expression, prepped for electroporation, and then electroporated with both the mutagenic oligo and a plasmid encoding constitutive expression of a guide RNA (gRNA) that targets Cas9 cleavage of the WT *rpoB* sequence. Electroporated cells were recovered in nonselective media overnight to allow sufficient segregation of the resulting mutant chromosomes. After plating on selection for the gRNA-encoding plasmid, two out of ten colonies screened were shown to contain the desired mutation, while the rest were shown to be escape mutants generated through protospacer mutation in the gRNA plasmid or Cas9 plasmid loss.

Lambda-Red Mutagenesis with Cas9 Selection

[00159] Mutants of candidate microbes CI006 and CI010 were generated via lambda-red mutagenesis with selection by CRISPR-Cas. Knockout cassettes contained an endogenous promoter identified through transcriptional profiling (as described in **Example 2** and depicted in **Table 3**) and ~250bp homology regions flanking the deletion target. CI006 and CI010 were transformed with plasmids encoding the Lambda-red recombination system (exo, beta, gam genes) under control of an arabinose inducible promoter and Cas9 under control of an IPTG

inducible promoter. The Red recombination and Cas9 systems were induced in resulting transformants, and strains were prepared for electroporation. Knockout cassettes and a plasmid-encoded selection gRNA were subsequently transformed into the competent cells. After plating on antibiotics selective for both the Cas9 plasmid and the gRNA plasmid, 7 of the 10 colonies screened showed the intended knockout mutation, as shown in **Figure 3**.

Example 4: *In Vitro* Phenotyping of Candidate Molecules

[00160] The impact of exogenous nitrogen on nitrogenase biosynthesis and activity in various mutants was assessed. The Acetylene Reduction Assay (ARA) (Temme et. al. 2012; 109(18): 7085–7090) was used to measure nitrogenase activity in pure culture conditions. Strains were grown in air-tight test tubes, and reduction of acetylene to ethylene was quantified with an Agilent 6890 gas chromatograph. ARA activities of candidate microbes and counterpart candidate mutants grown in nitrogen fixation media supplemented with 0 to 10mM glutamine are shown in **Figures 4A-B** and **Figures 10A-C**.

[00161] Under anaerobic culture conditions, a range of glutamine and ammonia concentrations was tested to quantify impact on nitrogen fixation activity. In wild-type cells, activity quickly diminished as glutamine concentrations increased. However, in a series of initial knock-out mutations, a class of mutation was validated enabling expression of nitrogen fixation genes under concentrations of glutamine that would otherwise shut off activity in wild type. This profile was generated in four different species of diazotrophs, as seen in **Figure 4C**. In addition, by rewiring the regulatory network using genetic parts that have been identified, the nitrogen fixation activity level was tuned predictably. This is seen in **Figure 4B**, which illustrates strains CM023, CM021, CM015, and CI006. Strain CM023 is an evolved strain low; strain CM021 is an evolved strain high; strain CM015 is an evolved strain mid; strain CI006 is a wild-type (strain 2). Ammonia excreted into culture supernatants was tested using an enzymatic-based assay (MEGAZYME). The assay measures the amount of NADPH consumed in the absorbance of 340 nm. The assay was conducted on bacterial cultures grown in nitrogen-free, anaerobic environment with a starting density of 1E9 CFU/ml. Across a panel of six evolved strains, one strain excreted up to 100 μ M of ammonia over a course of a 48 hour period, as seen in **Figure 4D**. Further, a double mutant exhibited higher ammonia excretion than the single mutant from which it was derived, as seen in **Figure 11**. This demonstrates a microbial capacity to produce ammonia in excess of its physiological needs.

Transcription Profiling of Pure Cultures

[00162] Transcriptional activity of CI006 was measured using the Nanostring Elements platform. Cells were grown in nitrogen-free media and 10E8 cells were collected after 4 hours incubation. Total RNA was extracted using the Qiagen RNeasy kit. Purified RNA was submitted to Core Diagnostics in Palo Alto, CA, for probe hybridization and Digital Analyzer analysis, as shown in **Figure 5**.

Example 5: *In Planta* Phenotyping of Candidate Microbes

Colonization of Plants by Candidate Microbes

[00163] Colonization of desired host plants by a candidate microbe was quantified through short-term plant growth experiments. Corn plants were inoculated with strains expressing RFP either from a plasmid or from a Tn5-integrated RFP expression cassette. Plants were grown in both sterilized sand and nonsterile peat medium, and inoculation was performed by pipetting 1 mL of cell culture directly over the emerging plant coleoptile three days post-germination. Plasmids were maintained by watering plants with a solution containing the appropriate antibiotic. After three weeks, plant roots were collected, rinsed three times in sterile water to remove visible soil, and split into two samples. One root sample was analyzed via fluorescence microscopy to identify localization patterns of candidate microbes. Microscopy was performed on 10mm lengths of the finest intact plant roots, as shown in **Figure 6**.

[00164] A second quantitative method for assessing colonization was developed. A quantitative PCR assay was performed on whole DNA preparations from the roots of plants inoculated with the endophytes. Seeds of corn (Dekalb DKC-66-40) were germinated in previously autoclaved sand in a 2.5 inch by 2.5 inch by 10 inch pot. One day after planting, 1ml of endophyte overnight culture (SOB media) was drenched right at the spot of where the seed was located. 1mL of this overnight culture is roughly equivalent to about 10^9 cfu, varying within 3-fold of each other, depending on which strain is being used. Each seedling was fertilized 3x weekly with 50mL modified Hoagland's solution supplemented with either 2.5mM or 0.25mM ammonium nitrate. At four weeks after planting, root samples were collected for DNA extraction. Soil debris were washed away using pressurized water spray. These tissue samples were then homogenized using QIAGEN Tissuelyzer and the DNA was then extracted using QIAmp DNA Mini Kit (QIAGEN) according to the recommended protocol. qPCR assay was performed using Stratagene Mx3005P RT-PCR on these DNA extracts using primers that were designed (using NCBI's Primer BLAST) to be specific to a loci in each of the endophyte's genome. The presence of the genome copies of the endophytes was quantified. To further

confirm the identity of the endophytes, the PCR amplification products were sequenced and are confirmed to have the correct sequence. The summary of the colonization profile of strain CI006 and CI008 from candidate microbes are presented in **Table 5**. Colonization rate as high as $10^7 \times \text{cfu} / \text{g fw}$ of root was demonstrated in strain CI008.

Strain	Colonization Rate (CFU / g fw)
CI006	1.45×10^5
CI008	1.24×10^7

Table 5: Colonization of corn as measured by qPCR

In Planta RNA Profiling

[00165] Biosynthesis of nif pathway components *in planta* was estimated by measuring the transcription of nif genes. Total RNA was obtained from root plant tissue of CI006 inoculated plants (planting methods as described previously). RNA extraction was performed using RNEasy Mini Kit according to the recommended protocol (QIAGEN). Total RNA from these plant tissues was then assayed using Nanostring Elements kits (NanoString Technologies, Inc.) using probes that were specific to the nif genes in the genome of strain CI006. The data of nif gene expression *in planta* is summarized in **Table 6**. Expression of nifH genes was detected in plants inoculated by CM013 strains whereas nifH expression was not detectable in CI006 inoculated plants. Strain CM013 is a derivative of strain CI006 in which the nifL gene has been knocked out.

[00166] Highly expressed genes of CM011, ranked by transcripts per kilobase million (TPM), were measured in planta under fertilized condition. The promoters controlling expression of some of these highly expressed genes were used as templates for homologous recombination into targeted nitrogen fixation and assimilation loci. RNA samples from greenhouse grown CM011 inoculated plant were extracted, rRNA removed using Ribo-Zero kit, sequenced using Illumina's Truseq platform and mapped back to the genome of CM011. Highly expressed genes from CM011 are listed in **Table 7**.

Strains	Relative Transcript Expression
CI006	9.4
CM013	103.25

Table 6: Expression of *nifH* in *planta*

Gene Name	Gene Location	Direction	Raw Read Count	TPM (Transcripts Per Kilobase Million)
rpsH CDS	18196 - 18588	reverse	4841.5	27206.4
rplQ CDS	11650 - 12039	reverse	4333	24536.2
rpsJ CDS	25013 - 25324	reverse	3423	24229
rplV CDS	21946 - 22278	reverse	3367.5	22333
rpsN CDS	18622 - 18927	reverse	2792	20150.1
rplN CDS	19820 - 20191	reverse	3317	19691.8
rplF CDS	17649 - 18182	reverse	4504.5	18628.9
rpsD CDS	13095 - 13715	reverse	5091.5	18106.6
rpmF CDS	8326 - 8493	forward	1363.5	17923.8
rplW CDS	23429 - 23731	reverse	2252	16413.8
rpsM CDS	14153 - 14509	reverse	2269	14036.2
rplR CDS	17286 - 17639	reverse	2243.5	13996.1
rplC CDS	24350 - 24979	reverse	3985	13969.2
rplK CDS	25526 - 25954	reverse	2648.5	13634.1
rplP CDS	20807 - 21217	reverse	2423	13019.5
rplX CDS	19495 - 19809	reverse	1824	12787.8
rpsQ CDS	20362 - 20616	reverse	1460.5	12648.7
bhsA 3 CDS	79720 - 79977	reverse	1464	12531.5
rpmC CDS	20616 - 20807	reverse	998.5	11485
rpoA CDS	12080 - 13069	reverse	4855	10830.2
rplD CDS	23728 - 24333	reverse	2916.5	10628.5
bhsA 1 CDS	78883 - 79140	reverse	1068	9141.9
rpsS CDS	22293 - 22571	reverse	1138.5	9011.8
rpmA CDS	2210 - 2467	forward	1028.5	8803.7
rpmD CDS	16585 - 16764	reverse	694.5	8520.8
rplB CDS	22586 - 23410	reverse	3132	8384
rpsC CDS	21230 - 21928	reverse	2574.5	8133.9
rplE CDS	18941 - 19480	reverse	1972.5	8066.9
rplO CDS	16147 - 16581	reverse	1551	7874.2
preprotein translocase subunit SecY CDS	14808 - 16139	reverse	4657	7721.2
rpsE CDS	16771 - 17271	reverse	1671.5	7368

Gene Name	Gene Location	Direction	Raw Read Count	TPM (Transcripts Per Kilobase Million)
rpsK CDS	13746 - 14135	reverse	1223.5	6928.2
tufA CDS	27318 - 28229	reverse	2850	6901.3
rpmI CDS	38574 - 38771	forward	615	6859.5
rplU CDS	1880 - 2191	forward	935.5	6621.7
rplT CDS	38814 - 39170	forward	1045	6464.4
bhsA 2 CDS	79293 - 79550	reverse	754	6454.1
rpmB CDS	8391 - 8627	reverse	682	6355.1
rplJ CDS	23983 - 24480	reverse	1408	6243.9
fusA 2 CDS	481 - 2595	reverse	5832	6089.6
rpsA CDS	25062 - 26771	reverse	4613	5957.6
rpmJ CDS	14658 - 14774	reverse	314	5926.9
rpsR CDS	52990 - 53217	forward	603	5840.7
rpsG CDS	2692 - 3162	reverse	1243	5828.2
rpsI CDS	11354 - 11746	reverse	980.5	5509.8
cspC 1 CDS	8091 - 8300	reverse	509	5352.8
rpsF CDS	52270 - 52662	forward	916	5147.4
rpsT CDS	55208 - 55471	reverse	602	5035.9
infC CDS	38128 - 38478	forward	755	4750.3
cspG CDS	30148 - 30360	forward	446	4624.2

Table 7**¹⁵N Assay**

[00167] The primary method for demonstrating fixation uses the nitrogen isotope ¹⁵N, which is found in the atmosphere at a set rate relative to ¹⁴N. By supplementing either fertilizer or atmosphere with enriched levels of ¹⁵N, one can observe fixation either directly, in heightened amounts of ¹⁵N fixed from an atmosphere supplemented with ¹⁵N₂ gas (Yoshida 1980), or inversely, through dilution of enriched fertilizer by atmospheric N₂ gas in plant tissues (Iniguez 2004). The dilution method allows for the observation of cumulative fixed nitrogen over the course of plant growth, while the ¹⁵N₂ gas method is restricted to measuring the fixation that occurs over the short interval that a plant can be grown in a contained atmosphere (rate measurement). Therefore, the gas method is superior in specificity (as any elevated ¹⁵N₂ levels in the plant above the atmospheric rate can be attributed unambiguously to fixation) but cannot show cumulative activity.

[00168] Both types of assay has been performed to measure fixation activity of improved strains relative to wild-type and uninoculated corn plants, and elevated fixation rates were observed *in*

planta for several of the improved strains (**Figure 12, Figure 14A, and Figure 14B**). These assays are instrumental in demonstrating that the activity of the strains observed *in vitro* translates to *in vivo* results. Furthermore, these assays allow measurement of the impact of fertilizer on strain activity, suggesting suitable functionality in an agricultural setting. Similar results were observed when setaria plants were inoculated with wild-type and improved strains (**Figure 13**). *In planta* fixation activity shown in **Figures 14A-14C** is further backed up by transcriptomic data. Evolved strains exhibit increased *nifH* transcript level relative to wild-type counterparts. Furthermore, the microbe derived nitrogen level *in planta* is also correlated with the colonization level on a plant by plant basis. These results (**Figure 12, Figure 13, Figures 14A-14C, Figure 15A, and Figure 15B**) support the hypothesis that the microbe, through the improved regulation of the *nif* gene cluster, is the likely reason for the increase in atmospheric derived nitrogen seen in the plant tissue. In addition to measuring fixation directly, the impact of inoculating plants with the improved strains in a nitrogen-stressed plant biomass assay was measured. While plant biomass may be related to many possible microbe interactions with the plant, one would expect that the addition of fixed nitrogen would impact the plant phenotype when nitrogen is limited. Inoculated plants were grown in the complete absence of nitrogen, and significant increases in leaf area, shoot fresh and dry weight, and root fresh and dry weight in inoculated plants relative to untreated controls was observed (**Figure 14C**). Although these differences cannot be attributed to nitrogen fixation exclusively, they support the conclusion that the improved strains are actively providing nitrogen to the plant. Corn and setaria plants were grown and inoculated as described above. Fertilizer comprising 1.2% ^{15}N was regularly supplied to plants via watering. Nitrogen fixation by microbes was quantified by measuring the ^{15}N level in the plant tissue. Fourth leaf tissue was collected and dried at 4 weeks after planting. Dried leaf samples were homogenized using beads (QIAGEN Tissuelyzer) and aliquoted out into tin capsules for IRMS (MBL Stable Isotope Laboratory at The Ecosystems Center, Woods Hole, MA). Nitrogen derived from the atmosphere (NDFA) was calculated, and nitrogen production by CI050 and CM002 are shown in **Figure 7**.

Phytohormone Production Assay

[00169] The dwarf tomato (*Solanum lycopersicum*) cultivar ‘Micro-Tom’ has previously been used to study the influence of indole-3-acetic acid on fruit ripening through an *in vitro* assay (Cohen 1996; J Am Soc Hortic Sci 121: 520–524). To evaluate phytohormone production and secretion by candidate microbes, a plate-based screening assay using immature Micro-Tom fruit was developed. Twelve-well tissue culture test plates were prepared by filling wells with agar

medium, allowing it to solidify, and spotting 10 uL of overnight microbial cultures onto the agar surface, as shown in **Figure 8**. Wells with agar containing increasing amounts of gibberellic acid (GA) but no bacterial culture were used as a positive control and standards. Flowers one day post-anthesis abscised from growing Micro-Tom plants were inserted, stem-first, into the agar at the point of the bacterial spot culture. These flowers were monitored for 2-3 weeks, after which the fruits were harvested and weighed. An increase in plant fruit mass across several replicates indicates production of plant hormone by the inoculant microbe, as shown in **Figure 9**.

Example 6: Cyclical Host-Microbe Evolution

[00170] Corn plants were inoculated with CM013 and grown 4 weeks to approximately the V5 growth stage. Those demonstrating improved nitrogen accumulation from microbial sources via ^{15}N analysis were uprooted, and roots were washed using pressurized water to remove bulk soil. A 0.25g section of root was cut and rinsed in PBS solution to remove fine soil particles and non-adherent microbes. Tissue samples were homogenized using 3mm steel beads in QIAGEN TissueLyser II. The homogenate was diluted and plated on SOB agar media. Single colonies were resuspended in liquid media and subjected to PCR analysis of 16s rDNA and mutations unique to the inoculating strain. The process of microbe isolation, mutagenesis, inoculation, and re-isolation can be repeated iteratively to improve microbial traits, plant traits, and the colonization capability of the microbe.

Example 7: Compatibility Across Geography

[00171] The ability of the improved microbes to colonize an inoculated plant is critical to the success of the plant under field conditions. While the described isolation methods are designed to select from soil microbes that may have a close relationship with crop plants such as corn, many strains may not colonize effectively across a range of plant genotypes, environments, soil types, or inoculation conditions. Since colonization is a complex process requiring a range of interactions between a microbial strain and host plant, screening for colonization competence has become a central method for selecting priority strains for further development. Early efforts to assess colonization used fluorescent tagging of strains, which was effective but time-consuming and not scalable on a per-strain basis. As colonization activity is not amenable to straightforward improvement, it is imperative that potential product candidates are selected from strains that are natural colonizers.

[00172] An assay was designed to test for robust colonization of the wild-type strains in any given host plant using qPCR and primers designed to be strain-specific in a community sample.

This assay is intended to rapidly measure the colonization rate of the microbes from corn tissue samples. Initial tests using strains assessed as probable colonizers using fluorescence microscopy and plate-based techniques indicated that a qPCR approach would be both quantitative and scalable.

[00173] A typical assay is performed as follows: Plants, mostly varieties of maize and wheat, are grown in a peat potting mix in the greenhouse in replicates of six per strain. At four or five days after planting, a 1 mL drench of early stationary phase cultures of bacteria diluted to an OD₅₉₀ of 0.6-1.0 (approximately 5E+08 CFU/mL) is pipetted over the emerging coleoptile. The plants are watered with tap water only and allowed to grow for four weeks before sampling, at which time, the plants are uprooted and the roots washed thoroughly to remove most peat residues. Samples of clean root are excised and homogenized to create a slurry of plant cell debris and associated bacterial cells. We developed a high-throughput DNA extraction protocol that effectively produced a mixture of plant and bacterial DNA to use as template for qPCR. Based on bacterial cell spike-in experiments, this DNA extraction process provides a quantitative bacterial DNA sample relative to the fresh weight of the roots. Each strain is assessed using strain-specific primers designed using Primer BLAST (Ye 2012) and compared to background amplification from uninoculated plants. Since some primers exhibit off-target amplification in uninoculated plants, colonization is determined either by presence of amplification or elevated amplification of the correct product compared to the background level.

[00174] This assay was used to measure the compatibility of the microbial product across different soil geography. Field soil qualities and field conditions can have a huge influence on the effect of a microbial product. Soil pH, water retention capacity, and competitive microbes are only a few examples of factors in soil that can affect inoculum survival and colonization ability. A colonization assay was performed using three diverse soil types sampled from agricultural fields in California as the plant growth medium (**Figure 16A**). An intermediate inoculation density was used to approximate realistic agricultural conditions. Within 3 weeks, Strain 5 colonized all plants at 1E+06 to 1E+07 CFU/g FW. After 7 weeks of plant growth, an evolved version of Strain 1 exhibited high colonization rates (1E+06 CFU/g FW) in all soil types. (**Figure 16B**).

[00175] Additionally, to assessment colonization in the complexity of field conditions, a 1-acre field trial in in San Luis Obispo in June of 2015 was initiated to assess the impacts and colonization of seven of the wild-type strains in two varieties of field corn. Agronomic design and execution of the trial was performed by a contract field research organization, Pacific Ag Research. For inoculation, the same peat culture seed coating technique tested in the inoculation

methods experiment was employed. During the course of the growing season, plant samples were collected to assess for colonization in the root and stem interior. Samples were collected from three replicate plots of each treatment at four and eight weeks after planting, and from all six reps of each treatment shortly before harvest at 16 weeks. Additional samples were collected from all six replicate plots of treatments inoculated with Strain 1 and Strain 2, as well as untreated controls, at 12 weeks. Numbers of cells per gram fresh weight of washed roots were assessed as with other colonization assays with qPCR and strain-specific primers. Two strains, Strain 1 and Strain 2, showed consistent and widespread root colonization that peaked at 12 weeks and then declined precipitously (**Figure 16C**). While Strain 2 appeared to be present in numbers an order of magnitude lower than Strain 1, it was found in more consistent numbers from plant to plant. No strains appeared to effectively colonize the stem interior. In support of the qPCR colonization data, both strains were successfully re-isolated from the root samples using plating and 16S sequencing to identify isolates of matching sequence

[00176] The use of the terms “a” and “an” and “the” and similar referents in the context of describing the invention (especially in the context of the following claims) are to be construed to cover both the singular and the plural, unless otherwise indicated herein or clearly contradicted by context. The terms “comprising,” “having,” “including,” and “containing” are to be construed as open-ended terms (i.e., meaning “including, but not limited to,”) unless otherwise noted. Recitation of ranges of values herein are merely intended to serve as a shorthand method of referring individually to each separate value falling within the range, unless otherwise indicated herein, and each separate value is incorporated into the specification as if it were individually recited herein. For example, if the range 10-15 is disclosed, then 11, 12, 13, and 14 are also disclosed. All methods described herein can be performed in any suitable order unless otherwise indicated herein or otherwise clearly contradicted by context. The use of any and all examples, or exemplary language (e.g., “such as”) provided herein, is intended merely to better illuminate the invention and does not pose a limitation on the scope of the invention unless otherwise claimed. No language in the specification should be construed as indicating any non-claimed element as essential to the practice of the invention.

[00177] While preferred embodiments of the present invention have been shown and described herein, it will be obvious to those skilled in the art that such embodiments are provided by way of example only. Numerous variations, changes, and substitutions will now occur to those skilled in the art without departing from the invention. It should be understood that various alternatives to the embodiments of the invention described herein may be employed in practicing

the invention. It is intended that the following claims define the scope of the invention and that methods and structures within the scope of these claims and their equivalents be covered thereby.

[00178] Table 3A

Name	Minimum	Maximum	Length	Direction
murein lipoprotein CDS	2,929,898	2,930,134	237	forward
membrane protein CDS	5,217,517	5,217,843	327	forward
zinc/cadmium-binding protein CDS	3,479,979	3,480,626	648	forward
acyl carrier protein CDS	4,563,344	4,563,580	237	reverse
ompX CDS	4,251,002	4,251,514	513	forward
DNA-binding protein HU-beta CDS	375,156	375,428	273	forward
sspA CDS	629,998	630,636	639	reverse
tatE CDS	3,199,435	3,199,638	204	reverse
LexA repressor CDS	1,850,457	1,851,065	609	forward
hisS CDS	<3999979	4,001,223	>1245	forward

[00179] Table 3B

Name	Differential Expression Absolute Confidence	Differential Expression Ratio	RNASeq_nifL - Raw Read Count	RNASeq_nifL - Raw Transcript Count	RNASeq_WT - Raw Read Count	RNASeq_WT - Raw Transcript Count
murein lipoprotein CDS	1000	-1.8	12950.5	10078.9	5151.5	4106.8
membrane protein CDS	1000	-1.3	9522.5	5371.3	5400	3120
zinc/cadmium-binding protein CDS	3.3	1.1	6461	1839.1	5318	1550.6
acyl carrier protein CDS	25.6	1.6	1230.5	957.6	1473.5	1174.7
ompX CDS	1.7	1.1	2042	734.2	1687.5	621.5
DNA-binding protein HU-beta CDS	6.9	-1.3	1305	881.7	725	501.8
sspA CDS	0.2	1	654	188.8	504.5	149.2
tatE CDS	1.4	1.3	131	118.4	125	115.8
LexA repressor CDS	0.1	-1.1	248	75.1	164	50.9
hisS CDS	0	-1.1	467	69.2	325	49.3

[00180] Table 3C

Name	Prm (In Forward direction, -250 to +10 region)	Expressed Sequence	Neighbor Sequence
murein lipoprotein CDS	GCCTCTCGGGGCGC TTTTTTTATTCCGG CACTAGCCGCTATT AATAAAAATGCAA ATCGGAATTTACTA TTTAACGCGAGATT ATCTAAGATGAATC CGATGGAAGCGCG CTGTTTTCACTCGC CTTTTTAAAGTTAC GTGATGATTTTCGAT GCTTCTTTGAGCGA ACGATCAAAAATA AGCGTATTCAGGTA AAAAAATATTCTCA TCACAAAAAAGTTT GTGTAATACTTGTA ACGCT--- ACATGGAGATTAA CTC	ATGAATCGTACTAAA CTGGTACTGGGCGCG GTAATCCTGGGTTCT ACTCTGCTGGCTGGT TGCTCCAGCAATGCT AAAATCGATCAGCTG TCTTCTGACGTTTCA ACTCTGAACGCTAAA GTTGACCAGCTGAGC AACGACGTGAACGC AATGCGTTCCGACGT TCAGGCTGCTAAAGA TGACGCAGCTCGCGC TAACCAGCGTCTGGA CAACGCAGCTACTAA ATACCGTAAGTAA	ATGAAAAAGACCAAAAT TGTTTGCACCATCGGTCC GAAAACCGAATCCGAAG AGATGTTGACCAAAATGC TGGACGCGGGCATGAAC GTTATGCGTCTGAACTTC TCTCACGGTGACTATGCG GAACACGGTCAGCGCATC CAGAATCTGCGCAATGTG ATGAGTAAAACCGGTAA GAAAGCGGCAATCCTGCT GGACACCAAAGGTCCGG AAATCCGTACCATTAAGC TGGAAGGCGGCAACGAC GTCTCCCTGAAAGCGGGC CAGACCTTCACCTTCACC ACCGATAAATCCGTTGTC GGTAATAACGAAATCGTT GCGGTGACCTATGAAGGC TTCACCAGCGACCTGAGC GTTGGCAACACGGTACTG GTTGACGATGGTCTGATC GGTATGGAAGTGACCGCT ATCGAAGGCAACAAAGTT GTTTGTAAGTGCTGAAC AACGGCGACCTCGGCGA GAACAAAGGCGTTAACCT GCCGGGCGTATCTATCGC GCTGCCGGCGCTGGCTGA AAAAGACAAACAGGATC TGATCTTCGGTTGCGAAC AGGGCGTTGACTTTGTTG CGGCATCCTTTATCCGTA AGCGTTCTGACGTTGTTG AAATCCGTGAGCACCTGA AAGCCCACGGCGGCGAG AAGATCCAGATCATCTCC AAAATCGAAAACCAGGA AGGCCTGAACAACTTCGA CGAAATCCTCGAAGCCTC TGACGGCATCATGGTAGC CCGTGGCGACCTGGGCGT TGAAATCCCGGTTGAAGA AGTTATCTTCGCGCAGAA GATGATGATCGAGAAATG TATCCGCGCGCGTAAAGT CGTTATCACCGCGACCCA GATGCTGGATTCCATGAT

			CAAAAACCCGCGTCCGAC CCGTGCGGAAGCAGGCG ACGTGGCCAACGCCATCC TCGACGGCACCGACGCA GTTATGCTGTCCGGCGAA TCCGCGAAAGGTAAATAC CCGCTGGAAGCGGTCACC ATCATGGCGACCATCTGC GAACGTACCGACCGCGTC ATGACCAGCCGTCTTGAG TACAACAACGACAACCGT AAGCTGCGCATCACCGAA GCGGTGTGCCGCGGTGCG GTAGAAACGGCTGAAAA ACTGGAAGCGCCGCTGAT CGTTGTGGCAACCCAGGG CGGTAAATCCGCGCGCGC CGTACGTAAATACTTCCC GGATGCCACTATCCTGGC GCTGACCACCAACGAAA CCACCGCGCGTCAGCTGG TGCTGAGCAAAGGCGTTG TGGCACAGCTGGTTGAAG ATATCTCCTCTACCGATG CGTTCTACATCCAGGGTA AAGAACTGGCGCTGCAG AGCGGTCTGGCGCGTAA GCGACGTGGTTGTTATG GTTTCGGCGCGTTAGTC CCGAGCGGAACCACCAA TACCGCTTCCGTGCACGT GCTGTAA
membrane protein CDS	GGTTACATAAACA TAATTATCGCCACG GCGATAGCCGTAC GCTTTTTGCGTCAC AACATCCATGGTGA AGCCGGCTTTTTCA AGAACACGCGCCA CCTCATCGGGTCTT AAATACATACTCAT TCCTCATTATCTTTT ACCGCACGTTAACC TTACCTTATTCATT AAAGGCAACGCTTT CGGAATATTCCATA AAGGGCTATTTACA GCATAATTCAAAAT CTTGTCCTACACTT ATAGACTCAATGG AATTAAGGGA	ATGGCCAACCGAGCA AACCGCAACAACGTA GAAGAGAGCGCTGA AGATATCCATAACGA TGTCAGCCAATTAGC GGATACGCTGGAAG AGGTGCTGAAATCGT GGGGCAGCGACGCC AAAGACGAAGCGGA GGCCGCGCGCAAAA AAGCGCAGGCGCTGC TGAAAGAGACCCGC GCCCGGCTTAACGGC AACAACCGCGTCCAG CAGGCGGCGTGCGAC GCCATGGGCTGCGCT GACAGCTACGTGCGC GACAAACCGTGGA AAGCGTCGGCGCCGC AGCAGCCGTTGGGGT ATTTATTGGCGTATT ACTGAATTTACGTCG ATAA	ATGTATTTAAGACCCGAT GAGGTGGCGCGTGTCTT GAAAAAGCCGGCTTCACC ATGGATGTTGTGACGCAA AAAGCGTACGGCTATCGC CGTGGCGATAATTATGTT TATGTGAACCGTGAAGCT CGTATGGGGCGTACCGCG TTAATTATTCATCCGGCTT TAAAAGAGCGCAGCACA ACGCTTGCGGAGCCCGCG TCGGATATCAAAACCTGC GATCATTATGAGCAGTTC CCGCTCTATTTAGCGGGG GATGCTCAACAGCATTAT GGTATTCCACACGGGTTC AGTTCGGAATGGCGCTT GAGCGTTTTCTGAGTGGC CTGTTTGGCGAAACGCAG TATAGCTGA

zinc/cadmium-binding protein CDS	GCGCGGAAAATCG ACGCATAGCGCATT CTCAGAAGCCGGC CTGGTCTCGGTGGA AAAGCGAATCTTTC CCACGACCGCCGG GCCTTTAACAAAAG AATCAATGACCTGA TTAATGTCGCTATC CATTCTCTCTCCGC GTAATGCGATCTTT TTTCATCATACCTA ACAAACTGGCAGA GGGAAAAGCCGCG CGGTTTTTCTGCGA AGTGTATTGTAAGA TTTGTTTGATATGT TATATCGTAACATA TTATTGCAAACAT	ATGACCAAAAAGATT TCCGCCCTAGCGTTT GGCATTGGCATGGTA ATGGCGAGCAGCCA GGCTTTTGCCACGG TCACCATAGTCATGG CCCGGCGCTGACCGA AGCGGAACAAAAGG CGAGTGAAGGCATTT TTGCTGACCAGGACG TAAAGGACAGGGCG CTGAGCGACTGGGAG GGGATCTGGCAGTCG GTTAACCCCTATCTG CTGAACGGGGATTTA GATCCGGTTCTGGAG CAGAAGGCCAAAAA GGCCGGTAAAAGCGT GGCGGAATATCGGG AATATTATAAGAAGG GCTACGCTACCGATG TCGACCAGATTGGTA TCGAGGATAACGTCA TGGAGTTTCACGTCG GGAAAACCGTCAAC GCCTGTAAGTACAGC TATTCCGGTTACAAA ATTCTGACCTACGCA TCCGGTAAAAAAGGC GTGCGCTACCTGTTT GAATGCCAGCAGGC GGATTCAAAAGCGCC GAAGTTTGTTTCAGTT TAGCGATCACACCAT CGCGCCACGCAAGTC CCAGCATTTCCACAT CTTTATGGGCAATGA GTCCCAGGAAGCGCT GCTGAAAGAGATGG ATAACTGGCCAACT ACTATCCTTATGCGC TGCATAAAGAGCAG ATTGTCGACGAAATG CTGCACCACTAA	ATGGATAGCGACATTAAT CAGGTCATTGATTCTTTT GTAAAGGCCCGGCGGTC GTGGGAAAGATTCTGCTTT TCCACCGAGACCAGGCCG GCTTCTGAGAATGCGCTA TGCGTCGATTTTCCGCGC CTCGAAATCATGCTTGCG GGTCAGCTTCACGATCCG GCGATTAAAGCCGATCGC GCCCAGCTCATGCCGCAC GATGTGCTGTATATTCCC GCTGGCGGATGGAATGAC CCGCAATGGCTGGCGCCC TCCACTCTGCTCACTATC TTATTTGGTAAACAGCAG CTGGAATTCGTCTTGCGC CACTGGGACGGCAGCGC GCTTAACGTGCTGGATAA ACAGCAGGTTCCGCGCCC CGGTCCCCGGGTCTGGCTC TTTTCTGCTGCAGGCGCT GAATGAAATGCAGATGC AGCCGCGGGAGCAGCAC ACGGCCCGCTTTATTGTC ACCAGCCTGCTCAGCCAC TGTGCCGATCTGCTGGGC AGCCAGGTACAAACCTCA TCGCGCAGCCAGGCGCTT TTTGAAGCGATTTCGTAAG CATATTGACGCCCACTTT GCCGACCCGTAAACCCGG GAGTCGGTGGCGCAGGC GTTTTACCTCTCGCCAAA CTATCTATCCCACCTGTT CCAGAAATGCGGGCCAA TGGGCTTTAACGAGTATC TGAATCACATCCGCCTGG AGCAGGCCAGAATGCTGT TAAAAGGCCACGATATGA AAGTGAAAGATATCGCCC ACGCCTGCGGTTTCGCCG ACAGCAACTACTTCTGCC GCCTGTTTTGCAAAAACA CCGAACGCTCGCCGTCGG AGTATCGCCGTCAATATC ACAGCCAGCTGACGGAA AAAACAGCCCCGGCAAA AAACTAG
acyl carrier protein CDS	CTGACGAAGCGAG TTACATCACCGGTG AAACTCTGCACGTC AACGGCGGAATGT ATATGGTCTGACCG	ATGAGCACTATCGAA GAACGCGTTAAGAA AATTATCGGCGAACA GCTGGGCGTTAAGCA GGAAGAAGTTACCA	ATGAGTTTTGAAGGAAAA ATCGCGCTGGTTACCGGT GCAAGTCGCGGGGATTGGC CGCGCAATCGCTGAAACG CTCGTTGCCCGTGGCGCG

	AGATTTGCGCAAA ACGCTCAGGAACC GCGCAGTCTGTGCG GTTCACTGTAATGT TTTGTACAAAATGA TTTGCGTTATGAGG GCAAACAGCCGCA AAATAGCGTAAAA TCGTGGTAAGACCT GCCGGGATTTAGTT GCAAATTTTTCAAC ATTTTATACACTAC GAAAACCATCGCG AAAGCGAGTTTTGA	ACAATGCTTCCTTCG TTGAAGACCTGGGCG CTGATTCTCTTGACA CCGTTGAGCTGGTAA TGGCTCTGGAAGAAG AGTTTGATACTGAGA TTCCGGACGAAGAAG CTGAGAAAATCACTA CTGTTTCAGGCTGCCA TTGATTACATCAACG GCCACCAGGCGTAA	AAAGTTATCGGGACTGCG ACCAGCGAAAGCGGCGC GCAGGCGATCAGCGATTA TTTAGGTGCTAACGGTAA AGGTCTGCTGCTGAATGT GACCGATCCTGCATCTAT TGAATCTGTTCTGGGAAA TATTCGCGCAGAATTTGG TGAAGTTGATATCCTGGT GAACAATGCCGGGATCA CTCGTGATAACCTGTAA TGCGCATGAAAGATGATG AGTGGAACGATATTATCG AAACCAACCTGTCATCTG TTTTCCGTCTGTCAAAAG CGGTAATGCGCGCTATGA TGAAAAAGCGTCATGGAC GTATTATCACTATCGGTT CTGTGGTTGGTACCATGG GAAATGCGGGTCAGGCC AACTACGCTGCGGCGAA AGCGGGTCTGATTGGCTT CAGTAAATCACTGGCTCG CGAAGTTGCGTCCCGCGG TATTACTGTAAACGTTGT TGCTCCGGGCTTTATTGA AACGGACATGACGCGTG CGCTGACCGATGAGCAGC GTGCGGGTACGCTGGCGG CAGTTCCTGCGGGGCGCC TCGGCTCTCCAAATGAAA TCGCCAGTGCGGTGGCAT TTTTAGCCTCTGACGAAG CGAGTTACATCACCGGTG AAACTCTGCACGTCAACG GCGGAATGTATATGGTCT GA
ompX CDS	ACGCCTGGGGCGC CGACCAGCGGGAA GAGTGATTTGGCCA ACGAGGCGCCGCT CTGAATGGAAATC ATGGCGATTAAAAT AACCAGTATCGGC AACCATGCCGGTAC CTTACGAGACGAG CCGGGCATCCTTTC TCCTGTCAATTTTG TCAAATGCGGTAA AGGTTCCAGTGTA TTGAATTACCCCGC GCCGGTTGAGCTAA TGTTGAAAAAAG GGTCTTAAAAGCA GTACAATAGGGCG	ATGAATAAAATTGCA CGTTTTTCAGCACTG GCCGTTGTTCTGGCT GCATCCGTAGGTACC ACTGCTTTCGCTGCG ACTTCTACCGTTACC GGTGGCTACGCGCAG AGCGACATGCAGGGT GAAGCGAACAAAGC TGGCGGTTTCAACCT GAAGTACCGCTACGA GCAAGACAACAACC CGCTGGGTGTTATCG GTTCTTTCACCTACA CCGAAAAAGATCGTT CTGAATCTGGCGTTT ACAAAAAAGGCCAG TACTACGGCATCACC	ATGCCCGGCTCGTCTCGT AAGGTACCGGCATGGTTG CCGATACTGGTTATTTTA ATCGCCATGATTTCCAT

	GGTCTGAAGATAAT TTCA	GCAGGTCCGGCTTAC CGTCTGAACGACTGG GCTAGCATCTACGGC GTAGTGGGTGTTGGT TACGGTAAATTCCAG GACAACAGCTACCCG AACAAATCTGATATG AGCGACTACGGTTTC TCTTACGGCGCTGGT CTGCAGTTCAACCCG ATCGAAAACGTTGCC CTGGACTTCTCCTAC GAGCAGTCTCGCATT CGTAACGTTGACGTT GGCACCTGGATTGCT GGCGTAGGTTACCGC TTCTAA	
DNA-binding protein HU- beta CDS	TCTGATTCCCTGATG AAAATAAACGCGA CCTTGAAGAAATTC CGGATAACGTTATC GCCGATTTAGATAT CCATCCGGTGAAAC GAATCGAGGAAGT TCTGGCACTTGCGC TACAGAACGAACC GTTTGGAATGGAA GTCGTCACGGCAA AATAGTGATTTTCG GCAAATAGCGCTA AGAAAAATAGGGC TGGTAAGTAAATTC GTACTTGCCAGCCT TTTTTTGTGTAGCT AACTTAGATCGCTG GCAGGGGGGTCAA TT	GTGAATAAATCTCAA CTGATTGACAAAATT GCTGCCGGTGCGGAC ATTTCTAAAGCCGCA GCTGGACGTGCGTTA GATGCTTTAATCGCT TCTGTTACTGAATCT CTGCAGGCTGGAGAT GACGTTGCGCTGGTA GGGTTTGGTACTTTT GCTGTAAAGAGCGC GCTGCCCGTACTGGT CGCAATCCGCAAACA GGCAAAGAAATCAC CATTGCTGCTGCTAA AGTTCCGGGTTTCCG CGCAGGTAAAGCGCT GAAAGACGCGGTAA ACTGA	ATGAATCCTGAGCGTTCT GAACGCATTGAAATCCCC GTATTGCCGTTGCGCGAT GTGGTGGTTTATCCGCAC ATGGTCATACCCCTGTTT GTAGGGCGGGAAAAATC TATCCGTTGTCTCGAAGC AGCCATGGACCATGATAA AAAAATCATGCTGGTTGC GCAGAAAGAAGCCTCGA CGGATGAGCCGGGTGTAA ACGATCTTTTCACCGTCG GGACCGTGGCGTCTATTT TGCAAATGCTGAAGCTAC CGGACGGTACTGTTAAAG TGCTGGTCTGAAGTTTGC AGCGCGCGCGCATCTCTG CGCTGTCTGATAATGGCG AACATTTTTTCGGCGAAGG CGGAATACCTTGAATCGC CGGCGATTGACGAACGC GAGCAGGAAGTGCTGGTT CGTACCGCTATCAGCCAG TTTGAAGGCTACATCAAG CTGAACAAAAAATCCCT CCGGAAGTGCTGACGTCG CTGAATAGCATCGACGAT CCGGCGCGTCTGGCGGAT ACCATCGCTGCGCATATG CCGCTGAAGCTGGCGGAC AAACAGTCCGTGCTGGAG ATGTCCGACGTTAACGAG CGTCTGGAATATCTGATG GCGATGATGGAGTCGGA AATCGATCTGCTGCAGGT GGAGAAGCGTATTCGCAA CCGCGTGAAAAAGCAGA TGGAGAAATCTCAGCGCG

		AGTACTATCTGAATGAGC AAATGAAAGCCATTCAAA AAGAGCTCGGCGAGATG GACGACGCCCCGGACGA GAACGAAGCGCTGAAGC GTAAGATCGACGCGGCG AAAATGCCGAAAGAGGC AAAAGAGAAAACCGAAG CGGAAGTGCAAAACTG AAAATGATGTCCCCGATG TCGGCGGAAGCGACCGTC GTTTCGCGGCTACATCGAC TGGATGGTGCAGGTACCG TGGAACGCTCGCAGCAA GGTTAAAAAAGACCTGCG TCAGGCTCAGGAGATCCT CGATACCGATCACTACGG CCTTGAGCGCGTGAAGGA TCGCATTCTTGAGTACCT CGCGGTGCAGAGCCGTGT TAACAAGCTCAAAGGGC CGATCCTGTGCCTGGTTG GGCCTCCGGGGGTAGGTA AAACCTCTCTCGGCCAAT CCATCGCCAAAGCAACTG GACGCAAATATGTGCGTA TGGCGCTGGGCGGCGTGC GTGATGAAGCGGAAATCC GCGGTCACCGCCGTACCT ATATTGGCTCAATGCCGG GCAAAGTATCCAGAAA ATGGCTAAAGTGGGCGTT AAAAACCCGCTGTTCTTG CTGGATGAGATCGACAAG ATGTCTTCTGACATGCGC GGCGATCCGGCCTCGGCG CTGCTGGAGGTGTTGGAT CCGGAACAGAACGTGGC CTTTAACGACCACTATCT GGAAGTGGATTACGATCT CAGCGACGTGATGTTTCGT TGCGACCTCTAACTCCAT GAACATCCCGGCGCCGCT GCTGGATCGTATGGAAGT GATCCGCTCTCCGGCTA TACCGAAGATGAGAAGCT AAACATCGCCAAACGCC ATCTGCTGTCAAAACAGA TTGAGCGTAACGCGCTCA AGAAAGGCGAGCTGACG GTGGATGACAGCGCGATT ATCGGCATCATTCGCTAC TACACCCGTGAAGCAGGC GTGCGTGGTCTGGAGCGT GAAATCTCGAAACTGTGC
--	--	--

			CGCAAAGCGGTGAAACA GCTGCTGCTGGATAAGTC GCTGAAACACATCGAGAT TAACGGCGACAACCTGCA CGATTTCTTGGCGTGCA GCGCTACGACTATGGTCG TGC GGATAGCGAAAACC GCGTAGGTCAGGTGACCG GACTGGCGTGGACGGAA GTGGGCGGGGATCTGCTG ACCATTGAAACCGCCTGC GTTCCGGGTAAAGGCAAA CTGACCTACACCGGTTCA CTGGGTGAAGTCATGCAG GAATCCATCCAGGCGGCG CTGACGGTGGTTTCGTTCA CGTGCGGATAAGCTGGGT ATTA ACTCAGACTTTTAC GAAAAACGTGATATTCAC GTTACGTGCCGGAAGGC GCGACGCCGAAGGATGG TCCAAGCGCCGGTATCGC GATGTGCACCGCGCTGGT TTCCTGTCTGACGGGTAA TCCGGTACGCGCCGACGT GGCGATGACCGGTGAGAT TACCCTCCGTGGCCAGGT ATTGCCGATTGGTGGTCT GAAGGAAAAACTGTTGG CCGCGCATCGCGGCGGCA TTAAGACTGTTCTGATTC CTGATGAAAATAAACGCG ACCTTGAAGAAATTCCGG ATAACGTTATCGCCGATT TAGATATCCATCCGGTGA AACGAATCGAGGAAGTTC TGGCACTTGCGCTACAGA ACGAACCGTTTGGAATGG AAGTCGTCACGGCAAAAT AG
sspA CDS	GTAAGAAAGTCGG CCTGCGTAAAGCAC GTCGTCGTCCTCAG TTCTCCAAACGTTA ATTGTTTTCTGCTC ACGCAGAACAATTT GCGAAAAAACCCG CTTCGGCGGGTTTT TTTATGGATAAATT TGCCATTTTCCCTC TACAAACGCCCCAT TGTTACCACTTTTT CAGCATTTCCAGAA TCCCCTCACCACAA CGTCTTCAAAATCT	ATGGCTGTCGCTGCC AACAAACGTTTCGGTA ATGACGCTGTTTTCT GGTCCTACTGACATC TATAGCCATCAGGTC CGCATCGTGCTGGCC GAAAAAGGTGTTAGT TTTGAGATAGAGCAC GTGGAGAAGGACAA CCCGCCTCAGGATCT GATTGACCTCAACCC GAATCAAAGCGTACC GACGCTTGTTGGATCG TGAGCTCACTCTGTG GGAATCTCGCATCAT	ATGGCTGAAAATCAATAC TACGGCACCGGTGCGCCGC AAAAGTTCCGCAGCTCGC GTTTTTCATCAAACCGGGC AACGGTAAAATCGTTATC AACCAGCGTTCTCTGGAA CAGTACTTCGGTCGTGAA ACTGCCCCGATGGTAGTT CGTCAGCCGCTGGAAC TG GTCGACATGGTTGAGAAA TTAGATCTGTACATCACC GTAAAGGTGGTGGTATC TCTGGTCAGGCTGGTGCG ATCCGTACGGTATCACC CGCGCTCTGATGGAGTAC

	GGTAAACTATCATC CAATTTTCTGCCA AATGCAGGTGATTG TTCATTTTT	TATGGAATATCTGGA TGAGCGTTTCCCGCA TCCGCCGCTCATGCC GGTTTACCCGGTGGC GCGTGGGGAAAGCC GTCTGTATATGCAGC GTATCGAAAAGGACT GGTATTCGTTGATGA ATACCATTGACACCG GTACCGCTGCGCAGG CTGATACTGCGCGTA AGCAGCTGCGTGAAG AACTACAGGCGATTG CGCCAGTTTTCAACC AGAAGCCCTACTTCC TGAGCGATGAGTTCA GCCTGGTGGACTGCT ACCTGGCACCCTGCT TGTGGCGTCTGCCGG TTCTCGGCGTAGAGC TGGTCGGCGCTGGCG CGAAAGAGCTTAAA GGCTATATGACTCGC GTATTTGAGCGCGAC TCTTTCCTCGCTTCTT TAACTGAAGCCGAAC GTGAAATGCGTCTCG GTCGGGGCTAA	GACGAGTCCCTGCGTGGC GAACTGCGTAAAGCTGGT TTCGTTACTCGTGATGCT CGTCAGGTTGAACGTAAG AAAGTCGGCCTGCGTAAA GCACGTCGTCGTCCTCAG TTCTCCAAACGTAA
tatE CDS	GTCAAAGCCGTATT ATCGACCCCTTAGG GACAACGCTTGCCG GGGCGGGAGAGCG GCCGCAGTTGATTT TTGCCGAACCTTCA GCTGATTATATTCA GCAGGTACGCGAG CGCCTGCCGGTGTT GCGCAATCGCCGCT TTGCGCCACCGCAA TTATTATGACGTTT TTTTAAACAAGGCT TGATTCACCTTGTT ACAGATTGCTATTG TGTCCGCGCGTCAA ATAGCCGTTAATTG TATGCGTGTATGAT GGCGTATTCG	ATGGGTGAGATTAGT ATTACCAAAGTCTG GTAGTCGCAGCGCTG ATTATCCTGGTGT GGTACCAAAAAGTTA CGCACGCTGGGTGGA GACCTGGGCTCGGCT ATCAAAGGCTTTAAA AAAGCCATGAGCGAT GACGATGACAGTGCG AAGAAGACCAGTGCT GAAGAAGCGCCGGC ACAGAAGCTCTCTCA TAAAGAGTAA	ATGTTTGTGCTGCCGGA CAATTTGCCGTAACGCCG GACTGGACGGGAAACGC GCAGACCTGCGTCAGCAT GATGCGCCAGGCCGCGG AGCGGGGGGCGTCGCTTC TGGTTCTGCCTGAGGCGT TGCTGGCGCGAGACGATA ACGATGCGGATTTATCGG TTAAATCCGCCACGAGC TGGATGGCGGCTTCTTAC AGCTCTTGCTGGCGGAGA GCGAAAACAGCGCTTTGA CGACGGTGCTGACCCTGC ATATCCCTTCCGGCGAAG GTCGAGCGACGAATACG CTGGTGGCCCTGCGTCAG GGGAAGATTGTGGCGCA ATATCAGAACTGCATCT CTATGATGCGTTCAATAT CCAGGAATCCAGGCTGGT CGATGCCGGGCGGCAAA TTCCGCCGCTGATCGAAG TCGACGGGATGCGCGTGC GGCTGATGACCTGCTACG ATTTACGTTTCCCTGAGC TGGCGCTGTCGTTAGCGC

			TCAGCGGCGCGCAGCTCA TAGTGTTGCCTGCCGCGT GGGTAAAAGGGCCGCTG AAGGAACATCACTGGGC GACGCTGCTGGCGGCGCG GGCGCTGGATACAACCTG CTATATTGTCGCCGAGG AGAGTGCGGGACGCGTA ATATCGGTCAAAGCCGTA TTATCGACCCCTTAGGGA CAACGCTTGCCGGGGCGG GAGAGCGGCCGAGTTG ATTTTGGCCGAACTTTCA GCTGATTATATTCAGCAG GTACGCGAGCGCCTGCCG GTGTTGCGCAATCGCCGC TTTGCGCCACCGCAATTA TTATGA
LexA repressor CDS	GAGGCGGTGGTTG ACCGTATCGGTCCC GAGCATCATGAGCT TTCGGGGCGAGCG AAAGATATGGGAT CGGCGGCGGTACT GCTGGCGATTATCA TCGCGCTGATCGCG TGGGGAACGCTGCT GTGGGCGAACTAC CGCTAAGTCTTGTC GTAGCTGCTCGCAA AACGGAAAGAAAC TCCTGATTTTTGTG TGAAATGTGGTTCC AAAATCACCGTTAG CTGTATATACTCAC AGCATAACTGTATA TACACCCAGGGGG C	ATGAAAGCGTTAACG ACCAGGCAGCAAGA GGTGTGTTGATCTCAT TCGGGATCATATCAG CCAGACGGGCATGCC GCCGACGCGTGCGGA GATTGCTCAGCGCTT GGGGTTTCGCTCCCC AAACGCGGCGGAAG AGCATCTGAAAGCGC TGGCGCGTAAAGGCG CAATCGAGATCGTTT CCGGCGCCTCCCGCG GTATTCGTCTGCTGA CGGAAGAAGAAACC GGTCTGCCGCTTATT GGCCGCGTCGCGGCA GGTGAGCCGCTGCTA GCGCAGCAGCACATT GAAGGCCACTACCAG GTGGACCCGGCCATG TTTAAGCCGAACGCC GATTTTCTGCTGCGT GTTAGCGGTATGTCG ATGAAGGATATCGGT ATTCTCGATGGCGAC CTGCTGGCTGTCCAT AAAACGCAGGATGT GCGCAATGGTCAGGT GGTTGTGGCGCGTAT CGACGAAGAAGTGA CCGTGAAGCGTCTGA AAAAACAGGGTAAC GTCGTGGAATTGCTG CCGGAAAACAGCGA ATTCTCGCCGATCGT GGTCGACCTTCGCGA	ATGGCCAATAATACCACT GGGTAAACCCGAATTATT AAAGCGGCCCGGTATTCC TGGAAAGGATTCCGTGCG GCGTGGGTCAATGAGGCC GCATTTCTGTCAGGAAGGC ATCGCGGCCGTTATTGCC GTGGCGATCGCCTGCTGG TTGGACGTCGATGCCATC ACGCGGGTGCTGCTCATT AGCTCGGTCTGTAGTG ATGATAGTTGAAATTATC AATAGCGCGATTGAGGCG GTGGTTGACCGTATCGGT CCCGAGCATCATGAGCTT TCGGGGCGAGCGAAAGA TATGGGATCGGCGGCGGT ACTGCTGGCGATTATCAT CGCGCTGATCGCGTGGGG AACGCTGCTGTGGGCGAA CTACCGCTAA

		ACAAAGCTTTACTAT TGAAGGCCTGGCCGT CGGCGTTATCCGCAA CGGCAACTGGCAATA A	
hisS CDS	TAAGAAAAGCGGC CTGTACGAAGACG GCGTACGTAAAGA CAGGCTGGATAAC GACGATATGATCG ATCAGCTGGAAGC GCGTATTCGCGCTA AAGCATCGATGCTG GATGAGGCGCGTC GTATCGATATCCAG CAGGTTGAAGCGA AATAACGTGTTGGG AAGCGATACGCTTC CCGTGTATGATTGA ACCTGCGGGCGCG AGGCGCCGGGGTT CATTTTTGTATATA TAAAGAGAATAAA CGTGGCAAAGAAC ATTCAA	...ATGAACGATTATCT GCCGGGCGAAACCG CTCTCTGGCAGCGCA TTGAAGGCTCACTGA AGCAGGTGCTTGGTA GCTACGGTTACAGCG AAATCCGTTTGCCGA TTGTAGAGCAGACCC CGTTATTCAAACGCG CTATCGGCGAAGTGA CCGACGTGGTTGAAA AAGAGATGTACACCT TTGAGGACCGTAACG GCGATAGCCTGACTC TACGTCCGGAAGGCA CGGCTGGCTGCGTAC GCGCCGGTATCGAAC ATGGTCTCCTGTACA ATCAAGAACAGCGCC TGTGGTACATTGGGC CGATGTTCCGCCACG AACGTCCGCAAAAA GGCCGCTACCGTCAG TTCCACCAGATTGGC GCCGAAGCGTTTGGC CTGCAGGGGCCGGAT ATCGATGCCGAGCTG ATTATGCTGACCGCC CGCTGGTGGCGCGAG CTGGGCATCTCCGGC CACGTTGCGCTGGAG CTGAACTCTATCGGT TCGCTGGAGGCTCGC GCTAACTATCGCGAC GCGCTGGTGGCCTAT CTTGAGCAGTTTAAA GATAAGCTGGACGA AGACTGCAAACGCCG CATGTACACCAACCC GCTGCGCGTGCTGGA TTCTAAAAACCCGGA CGTCCAGGCGCTGCT GAACGACGCCCCGAC GCTGGGCGACTATCT TGATGAAGAGTCCAA AACGCATTTTGCCGG GCTGTGCGCGCTGCT GGATGATGCCGGTAT TCGCTATACCGTGAA TCAGCGTCTGGTACG	ATGCATAACCAGGCTCCG ATTCAACGTAGAAAATCA AAACGAATTTACGTTGGG AATGTGCCGATTGGCGAT GGCGCCCCATCGCCGTA CAGTCGATGACAAACAC GCGCACCAACCGATGTGGC GGCGACGGTAAATCAAAT TAAAGCCCTCGAGCGCGT TGGCGCGGATATCGTGCG CGTTTCGGTGCCGACGAT GGATGCGGCGGAAGCGTT CAAACCTTATCAAACAGCA GGTTAACGTCCCGCTGGT TGCCGATATCCACTTCGA TTACCGCATTGCGCTGAA GGTAGCGGAATACGGCGT TGATTGCCTGCGTATTAA CCCGGGCAATATCGGCAA CGAAGAGCGTATCCGCAT GGTGGTGGACTGCGCTCG CGATAAAAATATTTCCTAT CCGTATCGGGGTAAACGC CGGTTCTCTGGAAAAAGA TCTCCAGGAAAAATACGG CGAACCGACTCCGCAGGC GCTGCTGGAATCGGCAAT GCGCCATGTTGATCATCT CGATCGTCTCAACTTCGA TCAGTTTAAAGTCAGCGT AAAAGCCTCCGATGTGTT CCTCGCGGTTGAATCCTA TCGCCTGTTGGCGAAACA GATCGATCAGCCTCTGCA CCTCGGGATCACCGAAGC GGGCGGCGCGCGCAGCG GCGCGGTGAAGTCCGCG ATCGGCCTCGGCCTGCTG CTGTCTGAAGGGATTGGC GATACGCTGCGCGTCTCT CTGGCGGCGGATCCCGTT GAAGAGATCAAAGTGGG CTTCGATATTCTCAAGTC GCTGCGTATTGCTCTCG CGGGATCAACTTTATTGC CTGCCCCACCTGTTACG TCAGGAGTTTGACGTTAT CGGTACCGTTAACGCGCT GGAGCAGCGCCTGGAAG ATATCATTACGCCGATGG

	CGGTCTCGACTACTA CAACCGCACCGTGTT TGAGTGGGTCACCAC CAGCCTCGGTTCCCA GGGCACCGTCTGCGC CGGAGGCCGTTACGA TGGTCTGGTTGAGCA GCTTGGCGGTTCGCGC TACCCCTGGCGTCGG CTTTGCGATGGGGCT GGAACGTCTTGTTTT ACTGGTTCAGGCAGT GAATCCGGAATTTAA AGCCGATCCTGTTGT CGATATATACCTGGT AGCCTCCGGAACCTGA CACCCAGTCCGCAGC AATGCGTCTGGCTGA ACAGGTACGCGATGC GTTACCCGGCGTTAA GCTGATGACCAACCA TGGCGGCGGCAACTT TAAGAAGCAGTTTGC GCGCGCTGATAAATG GGGCGCTCGCGTTGC GCTGGTGCTGGGCGA ATCAGAAATCGCCGA CGGAAACGTGGTAGT GAAAGATTTACGCTC AGGTGAGCAAACCTAC CGTAACGCAGGATAG CGTTGCTGCGCATTT GCGCACACTTCTGGG TTAA	ATATTTTCGATCATTGGCT GCGTGGTAAACGGTCCCG GCGAGGCGCTGGTTTCCA CCCTCGGCGTAACCGGCG GCAATAAGAAAAGCGGC CTGTACGAAGACGGCGTA CGTAAAGACAGGCTGGAT AACGACGATATGATCGAT CAGCTGGAAGCGCGTATT CGCGCTAAAGCATCGATG CTGGATGAGGCGCGTCGT ATCGATATCCAGCAGGTT GAAGCGAAATAA
--	--	---

[00181] Table of Strains

Sort	First Reference	Current Name	Universal Name	Lineage	Mutagenic DNA Description	Genotype	Gene 1 mutation
1	Application text	CI006	CI006	Isolated strain from Enterobacter genera	None	WT	
2	Application text	CI008	CI008	Isolated strain from Burkholderia genera	None	WT	
3	Application text	CI010	CI010	Isolated strain from Klebsiella genera	None	WT	
4	Application text	CI019	CI019	Isolated strain from Rahnella genera	None	WT	
5	Application text	CI028	CI028	Isolated strain from Enterobacter genera	None	WT	
6	Application text	CI050	CI050	Isolated strain from Klebsiella genera	None	WT	
7	Application text	CM002	CM002	Mutant of CI050	Disruption of nifL gene with a kanamycin resistance expression cassette (KanR) encoding the aminoglycoside O-phosphotransferase gene aphI inserted.	Δ nifL::KanR	ATGAGCCATATTCAACGGGAAACGTC TTGCTCCAGGCCGCGATTAAATTCCA ACATGGATGCTGATTTATATGGGTAT AAATGGGCTCGCGATAATGTCGGGCA ATCAGGTGCGACAATCTATCGATTGT ATGGGAAGCCCGATGCGCCAGAGTTG TTTCTGAAACATGGCAAAGGTAGCGT TGCCAATGATGTTACAGATGAGATGG TCAGACTAAACTGGCTGACGGAATTT ATGCCTCTTCCGACCATCAAGCATTTT ATCCGTACTCCTGATGATGCATGGTT ACTACCACTGCGATCCCCGGGAAAA CAGCATTCCAGGTATTAGAAGAATAT CCTGATTCAGGTGAAAATATTGTTGA TGCGCTGGCAGTGTTCTGCGCCGGT TGCATTGATTCCTGTTTGTAATTGTC CTTTTAACAGCGATCGCGTATTTTCGTC TCGCTCAGGCGCAATCACGAATGAAT AACGGTTTGGTTGATGCGAGTGATTT TGATGACGAGCGTAATGGCTGGCCTG TTGAACAAGTCTGGAAAGAAATGCAT AAGCTTTTGCCATTCTCACCGGATTCA GTCGTCACCTCATGGTGATTCTCACTT GATAACCTTATTTTTGACGAGGGGAA ATTAATAGGTTGTATTGATGTTGGAC GAGTCGGAATCGCAGACCGATAACCAG GATCTTGCCATCCTATGGAACTGCCT CGGTGAGTTTTCTCCTTCATTACAGAA ACGGCTTTTTCAAAAATATGGTATTG ATAATCCTGATATGAATAAATTGCAG TTTCATTTGATGCTCGATGAGTTTTTC TAATAAGCCTGCCTGGTTCTGCGTTTC CCGCTCTTAATACCCCTGACCGGAGG TGAGCAATGA

Sort	First Reference	Current Name	Universal Name	Lineage	Mutagenic DNA Description	Genotype	Gene 1 mutation
8	Application text	CM011	CM011	Mutant of CI019	Disruption of nifL gene with a spectinomycin resistance expression cassette (SpecR) encoding the streptomycin 3"-O-adenylyltransferase gene aadA inserted.	Δ nifL::SpeR	ATGAGCATCACGGCGTTATCAGCATC ATTTCCTGAGGGGAATATCGCCAGCC GCTTGCTGCTGCAACATCCTTCACTGT TTTATAACCGTGTTGAACAATCTTCG GTGGCGAGCGTGTGAGTCATCCTGA CTAGCTGAGATGAGGGCTCGCCCCCT CGTCCCGACACTTCCAGATCGCCATA GCGCACAGCGCCTCGAGCGGTGGTAA CGGCGCAGTGGCGGTTTTCATGGCTT GTTATGACTGTTTTTTGGGGTACAGT CTATGCCTCGGGCATCCAAGCAGCAA GCGCGTTACGCCGTGGGTGATGTTT GATGTTATGGAGCAGCAACGATGTTA CGCAGCAGGGCAGTCGCCCTAAAAACA AAGTTAAACATCATGAGGGAAGCGGT GATCGCCGAAGTATCGACTCAACTAT CAGAGGTAGTTGGCGTCATCGAGCGC CATCTCGAACCAGCTTGCTGGCCGT ACATTTGTACGGCTCCGAGTGGATG GCGGCCTGAAGCCACACAGTGATATT GATTTGCTGTTACGGTGACCGTAAG GCTTGATGAAACAACGCGGCGAGCTT TGATCAACGACCTTTTGAAAACCTTCG GCTTCCCCTGGAGAGAGCGAGATTCT CCGCGCTGTAGAAGTCACCATTTGTTG TGCACGACGACATCATTCCGTGGCGT TATCCAGCTAAGCGCGAACTGCAATT TGGAGAATGGCAGCGCAATGACATTC TTGCAGGTATCTTCGAGCCAGCCACG ATCGACATTGATCTGGCTATCTTGCTG ACAAAAGCAAGAGAACATAGCGTTG CCTTGGTAGGTCCAGCGGCGGAGGAA CTCTTTGATCCGGTTCCTGAACAGGA TCTATTTGAGGCGCTAAATGAAACCT TAACGCTATGGAACTCGCCGCCCGAC TGGGCTGGCGATGAGCGAAATGTAGT GCTTACGTTGTCCCGCATTTGGTACA GCGCAGTAACCGGCAAAATCGCGCCG AAGGATGTCGCTGCCGACTGGGCAAT GGAGCGCCTGCCGGCCAGTATCAGC CCGTCATACTTGAAGCTAGACAGGCT TATCTTGGACAAGAAGAAGATCGCTT GGCCTCGCGCGCAGATCAGTTGGAAG AATTTGTCCACTACGTGAAAGGCGAG ATCACCAGGTAGTCGGCAAATAATG TCTAACAATTCTGTTCAAGCCGACGCC GCTTCGCGGCGCGGCTTAACCTCAAGC GTTAGATGCACTAAGCACATAATTGC TCACAGCCAACTATCAGGTCAAGTC TGCTTTTATTATTTTAAAGCGTGATA ATAAGCCCTACACAAATGGTACCCGA CCGGTGGTGAATTTAATCTCGCTGAC GTGTAGACATTCCCTTATCCAGACGC TGATCGCCCATCATCGCGGTTCTTTAG ATCTCTCGGTCCGCCCTGATGGCGGC ACCTTGCTGACGTTACGCCTGCCGGT ACAGCAGGTTATCACCGGAGGCTTAA AATGA
9	Application text	CM013	CM013	Mutant of CI006	Disruption of nifL gene with a kanamycin resistance expression	Δ nifL::KanR	CTGATCCTTCAACTCAGCAAAAGTTT GATTTATTCAACAAAGCCACGTTGTG TCTCAAAATCTCTGATGTTACATTGCA CAAGATAAAAATATATCATCATGAAC AATAAACTGTCTGCTTACATAAACA GTAATACAAGGGGTGTTATGAGCCAT

Sort	First Reference	Current Name	Universal Name	Lineage	Mutagenic DNA Description	Genotype	Gene 1 mutation
					cassette (KanR) encoding the aminoglycoside O-phosphotransferase gene aph1 inserted.		ATTCAACGGGAAACGTCTTGCTCCAG GCCGCGATTAAATCCAACATGGATG CTGATTTATATGGGTATAAATGGGCT CGCGATAATGTCGGGCAATCAGGTGC GACAATCTATCGATTGTATGGGAAGC CCGATGCGCCAGAGTTGTTTCTGAAA CATGGCAAAGGTAGCGTTGCCAATGA TGTTACAGATGAGATGGTCAGACTAA ACTGGCTGACGGAATTTATGCCTCTT CCGACCATCAAGCATTTTATCCGTAC TCCTGATGATGCATGGTTACTACCA CTGCGATCCCCGGGAAAACAGCATT CAGGTATTAGAAGAATATCCTGATT AGGTGAAAATATTGTTGATGCGCTGG CAGTGTTCTGCGCCGGTTGCATTTCG ATTCTGTTTGTAAATTGTCCTTTTAAC AGCGATCGCGTATTTCTCGTCTCA GGCGCAATCACGAATGAATAACGGTT TGGTTGATGCGAGTGATTTTGTATGAC GAGCGTAATGGCTGGCCTGTTGAACA AGTCTGGAAAGAAATGCATAAGCTTT TGCCATTCTCACCGGATTCAGTCGTC ACTCATGGTGATTCTCACTTGATAAC CTTATTTTTTGACGAGGGAAATTAAT AGGTTGTATTGATGTTGGACGAGTCG GAATCGCAGACCGATACCAGGATCTT GCCATCCTATGGAAGTGCCTCGGTGA GTTTTCTCCTTCATTACAGAAACGGCT TTTTCAAAAATATGGTATTGATAATC CTGATATGAATAAATTGCAGTTTCAT TTGATGCTCGATGAGTTTCTAATAA GCCTTGACCCTACGATTCCCGCTATTT CATTCCTGACCGGAGGTTCAAAATG A
10	Figure 4A	CM004	CM004	Mutant of CI010	Disruption of amtB gene with a kanamycin resistance expression cassette (KanR) encoding the aminoglycoside O-phosphotransferase gene aph1 inserted.	Δ amtB::KanR	ATGAAGATAGCAACAATGAAAACAG GTCTGGGAGCGTTGGCTCTTCTCCCT GATCCTTCAACTCAGCAAAAGTTTCA TTTATTCAACAAAGCCACGTTGTGTCT CAAAATCTCTGATGTTACATTGCACA AGATAAAAATATATCATCATGAACAA TAAACTGTCTGCTTACATAAACAGT AATACAAGGGGTGTTATGAGCCATAT TCAACGGGAAACGTCTTGCTCCCGTC CGCGCTTAACTCCAACATGGACGCT GATTTATATGGGTATAAATGGGCTCG CGATAATGTCGGGCAATCAGGTGCGA CAATCTATCGCTTGTATGGGAAGCCC GATGCGCCAGAGTTGTTTCTGAAACA TGGCAAAGGTAGCGTTGCCAATGATG TTACAGATGAGATGGTCCGTCTCAAC TGGCTGACGGAGTTTATGCCTCTCCC GACCATCAAGCATTTTATCCGTACTC CTGATGATGCGTGGTTACTCACCACC GCGATTCTGCGGAAAACAGCCTTCCA GGTATTAGAAGAATATCCTGATTTCAG GTGAAAATATTGTTGATGCGCTGGCC GTGTTCTGCGCCGGTTACATTTCGATT CCTGTTTGTAAATTGTCCTTTTAACAGC GATCGTGATTTTCGTCTTGCTCAGGCG CAATCACGCATGAATAACGGTTTGGT TGATGCGAGTGATTTTGTATGACGAGC GTAATGGCTGGCCTGTTGAACAAGTC TGGAAGAAATGCACAAGCTCTTGCC ATTCTCACCGGATTCAGTCGTCCTC

Sort	First Reference	Current Name	Universal Name	Lineage	Mutagenic DNA Description	Genotype	Gene 1 mutation
							ATGGTGATTTCCTCACTTGATAACCTTA TTTTTGACGAGGGGAAATTAATAGGT TGTATTGATGTTGGACGGGTCGGAAT CGCAGACCGTTACCAGGACCTTGCCA TTCTTTGGAAGTGCCTCGGTGAGTTTT CTCCTTCATTACAGAAACGGCTTTTTTC AAAAATATGGTATTGATAATCCTGAT ATGAATAAATTGCAGTTTCATTTGAT GCTCGATGAGTTTTTCTAATAAGCCT GTGAAGGGCTGGACGTAAACAGCCA CGGCGAAAACGCCTACAACGCCTGA
11	Figure 4A	CM005	CM005	Mutant of CI010	Disruption of nifL gene with a kanamycin resistance expression cassette (KanR) encoding the aminoglycoside O-phosphotransferase gene aph1 inserted.	$\Delta nifL::KanR$	ATGACCCTGAATATGATGCTCGATAA CGCCGTACCCGAGGCGATTGCCGGCT GATCCTTCAACTCAGCAAAAGTTTCA TTTATTCAACAAAGCCACGTTGTGTCT CAAAATCTCTGATGTTACATTGCACA AGATAAAAATATATCATCATGAACAA TAAAACTGTCTGCTTACATAAACAGT AATACAAGGGGTGTTATGAGCCATAT TCAACGGGAAACGTCTTGCTCCCGTC CGCGCTTAAACTCCAACATGGACGCT GATTTATATGGGTATAAATGGGCTCG CGATAATGTCGGGCAATCAGGTGCGA CAATCTATCGCTTGATGCGAAGCCC GATGCGCCAGAGTTGTTTCTGAAACA TGGCAAAGGTAGCGTTGCCAATGATG TTACAGATGAGATGGTCCGTCTCAAC TGGCTGACGGAGTTTATGCCTCTCCC GACCATCAAGCATTTCCTGCTACTC CTGATGATGCGTGGTTACTCACCACC GCGATTCTGGGAAAACAGCCTTCCA GGTATTAGAAGAATATCCTGATTGAG GTGAAAATATTGTTGATGCGCTGGCC GTGTTCTGCGCCGGTTACATTGATT CCTGTTTGTAATTGTCCTTTTAACAGC GATCGTGTATTTCGCTTTGCTCAGGCG CAATCACGCATGAATAACGGTTTGGT TGATGCGAGTGATTTTGATGACGAGC GTAATGGCTGGCCTGTTGAACAAGTC TGGAAAGAAATGCACAAGCTCTTGCC ATTCTCACCGGATTCAGTCGTCACCTC ATGGTGATTTCCTCACTTGATAACCTTA TTTTTGACGAGGGGAAATTAATAGGT TGTATTGATGTTGGACGGGTCGGAAT CGCAGACCGTTACCAGGACCTTGCCA TTCTTTGGAAGTGCCTCGGTGAGTTTT CTCCTTCATTACAGAAACGGCTTTTTTC AAAAATATGGTATTGATAATCCTGAT ATGAATAAATTGCAGTTTCATTTGAT GCTCGATGAGTTTTTCTAATAAGCCTT GGTTCTGCGTTTCCCCTCTTTAATAC CCTGACCGGAGGTGAGCAATGA
12	Figure 4B	CM015	CM015	Mutant of CI006	Disruption of nifL gene with a fragment of the region upstream of the ompX gene inserted (Prm5).	$\Delta nifL::Prm5$	ATGACCCTGAATATGATGATGGATGC CGGCGGACATCATCGCGACAAACAAT ATTAATACCGGCAACCACACCGGCAA TTTACGAGACTGCGCAGGCATCCTTT CTCCCGTCAATTTCTGTCAATAAAG TAAAAGAGGCAGTCTACTTGAATTAC CCCCGGCTGGTTGAGCGTTTGTGAA AAAAAGTAACTGAAAAATCCGTAGA ATAGCGCCACTCTGATGGTTAATTAA CCTATTCAATTAAGAATTATCTGGAT GAATGTGCCATTAAATGCGCAGCATA ATGGTGCGTTGTGCGGGAAAACCTGCT

Sort	First Reference	Current Name	Universal Name	Lineage	Mutagenic DNA Description	Genotype	Gene 1 mutation
							TTTTTTTGAAGGGTTGGTCAGTAGC GGAAACAACCTCACTTCACACCCCGAA GGGGGAAGTTGCCTGACCCTACGATT CCCGCTATTTCACTTCACTGACCGGAG GTTCAAAATGA
13	Figure 4B	CM021	CM021	Mutant of CI006	Disruption of nifL gene with a fragment of the region upstream of an unannotated gene and the first 73bp of that gene inserted (Prm2).	Δ nifL::Prm2	ATGACCCTGAATATGATGATGGATGC CGGCTCACCACGGCGATAACCATAGG TTTTCGGCGTGGCCACATCCATGGTG AATCCCACTTTTCCAGCACGCGCGC CACTTCATCGGGCTTAAATACATAG ATTTTCCTCGTCATCTTTCCAAAGCCT CGCCACCTTACATGACTGAGCATGGA CCGTGACTCAGAAAATTCCACAAACG AACCTGAAAGGCGTGATTGCCGTCTG GCCTTAAAAATTATGGTCTAAACTAA AATTTACATCGAAAACGAGGGAGGAT CCTATGTTTAACAAACCGAATCGCCG TGACGTAGATGAAGGTGTTGAGGATA TTAACCACGATGTTAACCAGCTCGAA CTCACTTCACACCCCGAAGGGGGAAG TTGCCTGACCCTACGATTCCCGCTATT TCATTCCTGACCGGAGGTTCAAAAT GA
14	Figure 4B	CM023	CM023	Mutant of CI006	Disruption of nifL gene with a fragment of the region upstream of the acpP gene and the first 12bp of the acpP gene inserted (Prm4).	Δ nifL::Prm4	ATGACCCTGAATATGATGATGGATGC CGGCTGACGAGGCAGGTTACATCACT GGTGAAACCCTGCACGTCAATGGCGG AATGTATATGGTTTAACCACGATGAA AATTATTTGCGTTATTAGGGCGAAAG GCCTCAAAATAGCGTAAAATCGTGGT AAGAACTGCCGGGATTTAGTTGCAAA TTTTTCAACATTTTATACACTACGAAA ACCATCGCGAAAGCGAGTTTGTAGAG GAAATTTAAGAGTATGAGCACTATCG AAGAACGCGTTAAGAAAATTATCGGC GAACAGCTGGGCGTTAAGCAGGAAG AAGTTACCAACAATGCTTCCTTCGTT GAAGACCTGGGCGCTGATTCTCTTGA CACCGAACCTCACTTCACACCCCGAAG GGGGAAGTTGCCTGACCCTACGATT CCGCTATTTCACTTCACTGACCGGAGG TTCAAAATGA
15	Figure 10A	CM014	CM014	Mutant of CI006	Disruption of nifL gene with a fragment of the region upstream of the lpp gene and the first 29bp of the lpp gene inserted (Prm1).	Δ nifL::Prm1	ATGACCCTGAATATGATGATGGATGC CGGCCGTCTGTAATAATAACCGGAC AATTCGGACTGATTAATAAAGCGCCC TTGTGGCGCTTTTTTATATTTCCGCG TCCATTTAAAAATAAAAAATCCAATCG GATTTCACTATTTAAACTGGCCATTAT CTAAGATGAATCCGATGGAAGCTCGC TGTTTTAACACGCGTTTTTAACCTTT TATTGAAAGTCGGTGCTTCTTTGAGC GAACGATCAAATTTAAGTGGATTCCC ATCAAAAAAATATTCTCAACCTAAAA AAGTTGTGTAATACTTGTAAACGCTA CATGGAGATTAACTCAATCTAGAGGG TATTAATAATGAATCGTACTAAACTG GTACTGGGCGCAACTCACTTCACACC CCGAAGGGGAAGTTGCCTGACCCTA CGATTCCTCGCTATTTCACTTCACTGACC GGAGGTTCAAAATGA
16	Figure 10A	CM016	CM016	Mutant of CI006	Disruption of nifL gene with a fragment of the region	Δ nifL::Prm9	ATGACCCTGAATATGATGATGGATGC CGGCATATTGACACCATGACGCGCGT AATGCTGATTGGTTCTGTGACGCTGG TAATGATTGTCGAAATTCTGAACAGT GCCATCGAAGCCGTAGTAGACCGTAT

Sort	First Reference	Current Name	Universal Name	Lineage	Mutagenic DNA Description	Genotype	Gene 1 mutation
					upstream of the <i>lexA</i> 3 gene and the first 21bp of the <i>lexA</i> 3 gene inserted (Prm9).		TGGTGCAGAATTCCATGAACCTTTCCG GGCGGGCGAAGGATATGGGGTCGGC GGCGGTGCTGATGTCCATCCTGCTGG CGATGTTTACCTGGATCGCATTACTCT GGTCACATTTTCGATAACGCTTCCAG AATTCGATAACGCCCTGGTTTTTTGCT TAAATTTGGTTCCAAAATCGCCTTTA GCTGTATATACTACAGCATAACTGT ATATACACCCAGGGGGCGGGATGAA AGCATTAACGGCCAGGAACCTCACTTC ACACCCCGAAGGGGGAAGTTGCCTGA CCCTACGATTCCCGCTATTTTATTAC TGACCGGAGGTTCAAAATGA
17	Figure 10A	CM022	CM022	Mutant of CI006	Disruption of <i>nifL</i> gene with a fragment of the region upstream of the <i>mntP</i> 1 gene and the first 53bp of the <i>mntP</i> 1 gene inserted (Prm3).	$\Delta nifL::Prm3$	ATGACCCTGAATATGATGATGGATGC CGGCATCATATTGCGCTCCCTGGTTAT CATTTGTTACTAAATGAAATGTTATA ATATAACAATTATAAATACCACATCG CTTTCAATTACACAGCCAAATGAGAG GAGCGCCGTCTGACATAGCCAGCGCT ATAAAACATAGCATTATCTATATGTT TATGATTAATAACTGATTTTTCGCTTT TGGATTTGGCTGTGGCATCCTTGCCG CTCTTTTCGACGCTCTGCGTTTTTGC CCTCCGGTCAGGGCATTAAAGGGTCA GCAATGAGTTTTTACGCAATTACGAT TCTTGCTTCGGCATGTCGATGGATG CTTTAACTCACTTCACACCCCGAAGG GGGAAGTTGCCTGACCCTACGATTCC CGCTATTTTATTCACTGACCGGAGGT TCAAAATGA
18	Figure 10A	CM024	CM024	Mutant of CI006	Disruption of <i>nifL</i> gene with a fragment of the region upstream of the <i>sspA</i> gene inserted (Prm7).	$\Delta nifL::Prm7$	ATGACCCTGAATATGATGATGGATGC CGGCCGCGTCAGGTTGAACGTAAAAA AGTCGGTCTGCGCAAAGCACGTCGTC GTCCGCAGTTCTCCAAACGTTAATTG GTTTCTGCTTCGGCAGAACGATTGGC GAAAAAACCCGGTGCGAACCGGGTTT TTTTATGGATAAAGATCGTGTTATCC ACAGCAATCCATTGATTATCTCTTCTT TTTCAGCATTTCAGAAATCCCTCACC ACAAAGCCCGCAAATCTGGTAAACT ATCATCCAATTTTCTGCCCAAATGGCT GGGATTGTTCAATTTTTGTTTGCCTTA CAACGAGAGTGACAGTACGCGCGGG TAGTTAACTCAACATCTGACCGGTG ATAACTCACTTCACACCCCGAAGGGG GAAGTTGCCTGACCCTACGATTCCCG CTATTTTATTCACTGACCGGAGGTTT AAAATGA
19	Figure 10A	CM025	CM025	Mutant of CI006	Disruption of <i>nifL</i> gene with a fragment of the region upstream of the <i>hisS</i> gene and the first 52bp of the <i>hisS</i> gene inserted (Prm10).	$\Delta nifL::Prm10$	ATGACCCTGAATATGATGATGGATGC CGGCCCTGTATGAAGATGGCGTGCGC AAAGATCGCCTGGATAACAGCGATAT GATTAGCCAGCTTGAAGCCCGCATTC GCGCGAAAGCGTCAATGCTGGACGA AGCGCGTCGTATCGATGTGCAACAGG TAGAAAAATAAGGTTGCTGGGAAGC GGCAGGCTTCCCGTGTATGATGAACC CGCCCGGCGCGACCCGTTGTTTCGTCG CGGCCCCGAGGGTTCATTTTTTGTATT AATAAAGAGAATAAACGTGGCAAAA AATATTCAAGCCATTGCGGCGATGAA CGATTATCTGCCTGGCGAACTCACTT CACACCCCGAAGGGGGAAGTTGCCTG ACCCTACGATTCCCGCTATTTTATTCA CTGACCGGAGGTTCAAAATGA

Sort	First Reference	Current Name	Universal Name	Lineage	Mutagenic DNA Description	Genotype	Gene 1 mutation
20	Figure 10B	CM006	CM006	Mutant of CI010	Disruption of glnB gene with a kanamycin resistance expression cassette (KanR) encoding the aminoglycoside O-phosphotransferase gene aph1 inserted.	Δ glnB::KanR	ATGAAAAAGATTGATGCGATTATTAA ACCTTTCAAACCTGGATGACGTGCGCT GATCCTTCAAACCTCAGCAAAAGTTTCA TTTATTCAACAAAGCCACGTTGTGTCT CAAAATCTCTGATGTTACATTGCACA AGATAAAAATATATCATCATGAACAA TAAAACTGTCTGCTTACATAAACAGT AATACAAGGGGTGTTATGAGCCATAT TCAACGGGAAACGTCTTGTCTCCGTC CGCGCTTAAACTCCAACATGGACGCT GATTTATATGGGTATAAATGGGCTCG CGATAATGTCGGGCAATCAGGTGCGA CAATCTATCGCTTGTATGGGAAGCCC GATGCGCCAGAGTTGTTTCTGAAACA TGGCAAAGGTAGCGTTGCCAATGATG TTACAGATGAGATGGTCCGTCTCAAC TGGCTGACGGAGTTTATGCCTCTCCC GACCATCAAGCATTTTATCCGTACTC CTGATGATGCGTGGTTACTACCACC GCGATTCTTGGGAAAACAGCCTTCCA GGTATTAGAAGAATATCCTGATTAG GTGAAAATATTGTTGATGCGCTGGCC GTGTTCTGCGCCGGTTACATTGATT CCTGTTTGTAATTGTCCTTTTAACAGC GATCGTGTATTTCTGCTTGTCTCAGGCG CAATCACGCATGAATAACGGTTTGGT TGATGCGAGTGATTTTGTATGACGAGC GTAATGGCTGGCCTGTTGAACAAGTC TGGAAAGAAATGCACAAGCTCTTGCC ATTCTCACCGGATTACGTCGTCACCTC ATGGTGATTCTCACTTGATAACCTTA TTTTGTACGAGGGGAAATTAATAGGT TGTATTGATGTTGGACGGGTGCGAAT CGCAGACCGTTACCAGGACCTTGCCA TTCTTTGGAAGTGCCTCGGTGAGTTT CTCCTTCATTACAGAAACGGCTTTTTC AAAAATATGGTATTGATAATCCTGAT ATGAATAAATTGCAGTTTCATTGAT GCTCGATGAGTTTCTAATAAGCCTC GCGCGTGATTTCGTATCCGCACCGGCG AAGAAGACGACGCGCGATTAA
21	Figure 10C	CI028 nifL::KanR	CM017	Mutant of CI028	Disruption of nifL gene with a kanamycin resistance expression cassette (KanR) encoding the aminoglycoside O-phosphotransferase gene aph1 inserted.	Δ nifL::KanR	ATGACCATGAACCTGATGACGGATGT CGTCTCAGCCACCGGGATCGCCGGGT TGCTTTCACGACAACACCCGACGCTG TTTTTTACACTAATTGAACAGGCCCCC GTGGCGATCAGCTGACGGATACCGC TGCCCGCATTGTCTATGCCAACCCGG GCGTGTGAGTCATCTGACTAGCTG AGATGAGGGCTCGCCTGATCCTTCAA CTCAGCAAAAGTTTCGATTTATTCAAC AAAGCCACGTTGTGTCTCAAATCTC TGATGTTACATTGCACAAGATAAAAA TATATCATCATGAACAATAAACTGT CTGCTTACATAAACAGTAATACAAGG GGTGTTATGAGCCATATTCAACGGGA AACGTCTTGCTCCAGGCCGCGATTAA ATTCCAACATGGATGCTGATTTATAT GGGTATAAATGGGCTCGCGATAATGT CGGGCAATCAGGTGCGACAATCTATC GATTGTATGGGAAGCCGATGCGCCA GAGTTGTTTCTGAAACATGGCAAAGG TAGCGTTGCCAATGATGTTACAGATG AGATGGTCAGACTAACTGGCTGACG GAATTTATGCCTCTTCCGACCATCAA

Sort	First Reference	Current Name	Universal Name	Lineage	Mutagenic DNA Description	Genotype	Gene 1 mutation
							GCATTTTATCCGTACTCCTGATGATGC ATGGTTACTCACCCTGCGATCCCCG GGAAAACAGCATTCCAGGTATTAGAA GAATATCCTGATTGAGGTGAAAATAT TGTTGATGCGCTGGCAGTGTTCCTGC GCCGGTTCATTTCGATTCCCTGTTTGT ATTGTCCTTTTAAACAGCGATCGCGTAT TTCGTCTCGCTCAGGCGCAATCACGA ATGAATAACGGTTTGGTTGATGCGAG TGATTTTGATGACGAGCGTAATGGCT GGCCTGTTGAACAAGTCTGGAAGAA ATGCATAAGCTTTTGCCATTCTCACCG GATTCAGTCGTCACCTCATGGTATTC TCACTTGATAACCTTATTTTTCGAG GGGAAATTAATAGGTTGTATTGATGT TGGACGAGTCGGAATCGCAGACCGAT ACCAGGATCTTGCCATCCTATGGAAC TGCTCGGTGAGTTTCTCCTTCATTA CAGAAACGGCTTTTCAAAAATATGG TATTGATAATCCTGATATGAATAAAT TGCAGTTTCATTTGATGCTCGATGAGT TTTTCTAATAAGCCTGACCGTGGTG AATTAAATCTCGCTGACGTGTAGACA TTCATCGATCTGCATCCACGGTCCGG CGGCGGTACCTGCCTGACGCTACGTT TACCGCTCTTTATGAACTGACCGGA GGCCCAAGATGA
22	Figure 10C	CI019 nifL::SpecR	CM011	Mutant of CI019	Disruption of nifL gene with a spectinomycin in resistance expression cassette (SpecR) encoding the streptomycin 3"-O- adenylyltran sferase gene aadA inserted.	Δ nifL::SpeR	ATGAGCATCACGGCGTTATCAGCATC ATTTCTGAGGGGAATATCGCCAGCC GCTTGTCGCTGCAACATCCTTCACTGT TTTATACCGTGGTTGAACAATCTTCG GTGGCGAGCGTGTGAGTCATCCTGA CTAGCTGAGATGAGGGCTCGCCCCCT CGTCCCGACACTTCCAGATCGCCATA GCGCACAGCGCCTCGAGCGGTGGTAA CGGCGCAGTGGCGGTTTTCATGGCTT GTTATGACTGTTTTTTGGGGTACAGT CTATGCCTCGGGCATCCAAGCAGCAA GCGCGTTACGCGGTGGGTCTGATGTTT GATGTTATGGAGCAGCAACGATGTTA CGCAGCAGGGCAGTCGCCCTAAAACA AAGTTAAACATCATGAGGGAAGCGGT GATCGCCGAAGTATCGACTCAACTAT CAGAGGTAGTTGGCGTCATCGAGCGC CATCTCGAACCGACGTTGCTGGCCGT ACATTTGTACGGCTCCGAGTGGATG GCGGCCTGAAGCCACACAGTGATATT GATTTGCTGGTTACGGTGACCGTAAG GCTTGATGAAACAACGCGGCGAGCTT TGATCAACGACCTTTTGGAACTTCG GCTTCCCCTGGAGAGAGCGAGATTCT CCGCGCTGTAGAAGTCACCATTGTTG TGCACGACGACATCATTCCTGGCGT TATCCAGCTAAGCGCAACTGCAATT TGGAGAATGGCAGCGCAATGACATT TTGCAGGTATCTTCGAGCCAGCCACG ATCGACATTGATCTGGCTATCTTGCTG ACAAAAGCAAGAGAACATAGCGTTG CCTTGGTAGGTCCAGCGGCGGAGGAA CTCTTTGATCCGGTTCCTGAACAGGA TCTATTTGAGGCGCTAAATGAAACCT TAACGCTATGGAACGCGCCGCCGAC TGGGCTGGCGATGAGCGAAATGTAGT GCTTACGTTGTCCCGCATTTGGTACA

Sort	First Reference	Current Name	Universal Name	Lineage	Mutagenic DNA Description	Genotype	Gene 1 mutation
							GCGCAGTAACCGGCAAAATCGCGCCG AAGGATGTCGCTGCCGACTGGGCAAT GGAGCGCCTGCCGGCCAGTATCAGC CCGTCATACTTGAAGCTAGACAGGCT TATCTTGGACAAGAAGAAGATCGCTT GGCCTCGCGCGCAGATCAGTTGGAAG AATTTGTCCACTACGTGAAAGGCGAG ATCACCAAGGTAGTCGGCAAATAATG TCTAACAATTCGTTCAAGCCGACGCC GCTTCGCGGCGCGGCTTAACTCAAGC GTTAGATGCACTAAGCACATAATTGC TCACAGCCAACTATCAGGTCAAGTC TGCTTTTATTATTTTAAAGCGTGCATA ATAAGCCCTACACAAATGGTACCCGA CCGGTGGTGAATTTAATCTCGCTGAC GTGTAGACATTCCCTTATCCAGACGC TGATCGCCCATCATCGCGTTCTTTAG ATCTCTCGGTCCGCCCTGATGGCGGC ACCTTGCTGACGTTACGCCTGCCGGT ACAGCAGGTTATCACCGGAGGCTTAA AATGA
23	Figure 10C	CI006 nifL::Kan R	CM013	Mutant of CI006	Disruption of nifL gene with a kanamycin resistance expression cassette (KanR) encoding the aminoglycos ide O- phosphotran sferase gene aphI inserted.	Δ nifL::Kan R	CTGATCCTTCAACTCAGCAAAAGTTT GATTTATTCAACAAAGCCACGTTGTG TCTCAAAATCTCTGATGTTACATTGCA CAAGATAAAAATATATCATCATGAAC AATAAACTGTCTGCTTACATAAACA GTAATACAAGGGGTGTTATGAGCCAT ATTCAACGGGAAACGTCTTGCTCCAG GCCGCGATTAAATTCACATGGATG CTGATTTATATGGGTATAAATGGGCT CGCGATAATGTCGGGCAATCAGGTGC GACAATCTATCGATTGTATGGGAAGC CCGATGCGCCAGAGTTGTTTCTGAAA CATGGCAAAGGTAGCGTTGCCAATGA TGTTACAGATGAGATGGTCAGACTAA ACTGGCTGACGGAATTTATGCCTCTT CCGACCATCAAGCATTTTATCCGTAC TCCTGATGATGCATGGTTACTCACCA CTGCGATCCCCGGGAAAACAGCATTC CAGGTATTAGAAGAATATCCTGATTC AGGTGAAAATATTGTTGATGCGCTGG CAGTGTTCTGCGCCGTTGCATTCTG ATTCTGTTTGTAAATGTCTTTTAAAC AGCGATCGCGTATTTCTGCTCGCTCA GGCGCAATCACGAATGAATAACGGTT TGTTTGATGCGAGTGATTTTGATGAC GAGCGTAATGGCTGGCCTGTTGAACA AGTCTGGAAAGAAATGCATAAGCTTT TGCCATTCTCACCGGATTCAGTCGTC ACTCATGGTGATTCTCACTTGATAAC CTTATTTTGTACGAGGGGAAATTAAT AGGTTGTATTGATGTTGGACGAGTCG GAATCGCAGACCGATACAGGATCTT GCCATCCTATGGAAGTGCCTCGGTGA GTTTTCTCTTCATTACAGAAACGGCT TTTTCAAAAATATGGTATTGATAATC CTGATATGAATAAATTGCAGTTTCAT TTGATGCTCGATGAGTTTTTCTAATAA GCCTTGACCCTACGATTCCTCGCTATTT CATTCAGTACCGGAGGTTCAAAATG A

Sort	First Reference	Current Name	Universal Name	Lineage	Mutagenic DNA Description	Genotype	Gene 1 mutation
24	Figure 10C	CI010 nifL::Kan R	CM005	Mutant of CI010	Disruption of nifL gene with a kanamycin resistance expression cassette (KanR) encoding the aminoglycos ide O- phosphotran sferase gene aph1 inserted.	Δ nifL::Ka nR	ATGACCCTGAATATGATGCTCGATAA CGCCGTACCCGAGGCGATTGCCGGCT GATCCTTCAACTCAGCAAAAGTTTCGA TTTATTCAACAAAGCCACGTTGTGTCT CAAAATCTCTGATGTTACATTGCACA AGATAAAAATATATCATCATGAACAA TAAAACTGTCTGCTTACATAAACAGT AATACAAGGGGTGTTATGAGCCATAT TCAACGGGAAACGTCCTTGCTCCCGTC CGCGCTTAAACTCCAACATGGACGCT GATTTATATGGGTATAAATGGGCTCG CGATAATGTCGGGCAATCAGGTGCGA CAATCTATCGCTTGTATGGGAAGCCC GATGCGCCAGAGTTGTTTCTGAAACA TGGCAAAGGTAGCGTTGCCAATGATG TTACAGATGAGATGGTCCGTCCTCAAC TGGCTGACGGAGTTTATGCCTCTCCC GACCATCAAGCATTTTATCCGTACTC CTGATGATGCGTGGTTACTCACCACC GCGATTCTTGGGAAAACAGCCTTCCA GGTATTAGAAGAATATCCTGATTGAG GTGAAAATATTGTTGATGCGCTGGCC GTGTTCTGCGCCGGTTACATTGCGATT CCTGTTTGTAATTGTCCTTTTAACAGC GATCGTGTATTTCGTCTTGCTCAGGCG CAATCACGCATGAATAACGGTTTGGT TGATGCGAGTGATTTTGATGACGAGC GTAATGGCTGGCCTGTTGAACAAGTC TGGAAGAAATGCACAAGCTCTTGCC ATTCTCACCGGATTGAGTCGTCACCT ATGGTGATTCTCACTTGATAACCTTA TTTTTGACGAGGGGAAATTAATAGGT TGTATTGATGTTGGACGGGTCGGAAT CGCAGACCGTTACCAGGACCTTGCCA TTCTTTGGAAGTGCCTCGGTGAGTTT CTCCTTCATTACAGAAACGGCTTTTTC AAAAATATGGTATTGATAATCCTGAT ATGAATAAATTGCAGTTTCATTGAT GCTCGATGAGTTTCTAATAAGCCTT GGTTCTGCGTTTCCCCTCTTTAATAC CCTGACCGGAGGTGAGCAATGA
25	Figure 4C	Strain 2	CI006	Isolated strain from Enterobacter genera	None	WT	
26	Figure 4C	Strain 4	CI010	Isolated strain from Klebsiella genera	None	WT	
27	Figure 4C	Strain 1	CI019	Isolated strain from Rahnella genera	None	WT	
28	Figure 4C	Strain 3	CI028	Isolated strain from Enterobacter genera	None	WT	
29	Figure 4B	Strain 2	CI006	Isolated strain from Enterobacter genera	None	WT	

Sort	First Reference	Current Name	Universal Name	Lineage	Mutagenic DNA Description	Genotype	Gene 1 mutation
30	Figure 4B	High	CM014	Mutant of CI006	Disruption of nifL gene with a fragment of the region upstream of the lpp gene and the first 29bp of the lpp gene inserted (Prm1).	$\Delta nifL::Prm1$	ATGACCCTGAATATGATGATGGATGC CGGCCGTCCTGTAATAATAACCGGAC AATTCGGACTGATTAATAAAGCGCCC TTGTGGCGCTTTTTTATATTCCCGCC TCCATTTAAAAATAAAAAATCCAATCG GATTTCACTATTTAAACTGGCCATTAT CTAAGATGAATCCGATGGAAGCTCGC TGTTTTAACACGCGTTTTTAAACCTTT TATTGAAAGTCGGTGCTTCTTTGAGC GAACGATCAAATTTAAGTGGATTCCC ATCAAAAAAATATTCTCAACCTAAAA AAGTTTGTGTAATACTTGTAACGCTA CATGGAGATTAACCTCAATCTAGAGGG TATTAATAATGAATCGTACTAAACTG GTACTGGGCGCAACTCACTTCACACC CCGAAGGGGAAGTTGCCTGACCCTA CGATTCCCGCTATTTCACTTCACTGACC GGAGGTTCAAAATGA
31	Figure 4B	Med	CM015	Mutant of CI006	Disruption of nifL gene with a fragment of the region upstream of the ompX gene inserted (Prm5).	$\Delta nifL::Prm5$	ATGACCCTGAATATGATGATGGATGC CGGCGGACATCATCGCGACAAACAAT ATTAATACCGGCAACCACACCGGCAA TTTACGAGACTGCGCAGGCATCCTTT CTCCCGTCAATTTCTGTCAAATAAAG TAAAAGAGGCAGTCTACTTGAATTAC CCCCGGCTGGTTGAGCGTTTGTTGAA AAAAAGTAAGTGAATAATCCGTAGA ATAGCGCCACTCTGATGGTTAATTAA CCTATTCAATTAAGAATTATCTGGAT GAATGTGCCATTAATGCGCAGCATA ATGGTGCGTTGTGCGGGAATACTGCT TTTTTTTGAAGGGTTGGTCAGTAGC GGAAACAACCTCACTTCACACCCCGAA GGGGGAAGTTGCCTGACCCTACGATT CCCGCTATTTCACTTCACTGACCGGAG GTTCAAAATGA
32	Figure 4B	Low	CM023	Mutant of CI006	Disruption of nifL gene with a fragment of the region upstream of the acpP gene and the first 121bp of the acpP gene inserted (Prm4).	$\Delta nifL::Prm4$	ATGACCCTGAATATGATGATGGATGC CGGCTGACGAGGCAGGTTACATCACT GGTGAAACCCTGCACGTCAATGGCGG AATGTATATGGTTAAACCACGATGAA AATTATTTGCGTTATTAGGGCGAAAG GCCTCAAAATAGCGTAAAATCGTGGT AAGAACTGCCGGGATTTAGTTGCAAA TTTTTCAACATTTTATACACTACGAAA ACCATCGCGAAAGCGAGTTTTGATAG GAAATTTAAGAGTATGAGCACTATCG AAGAACGCGTTAAGAAAAATTATCGGC GAACAGCTGGGCGTTAAGCAGGAAG AAGTTACCAACAATGCTTCCTTCGTT GAAGACCTGGGCGCTGATTCTCTTGA CACCGAACTCACTTCACACCCCGAAG GGGGAAGTTGCCTGACCCTACGATTC CCGCTATTTCACTTCACTGACCGGAGG TTCAAAATGA
33	Figure 4D	Strain 2	CI006	Isolated strain from Enterobacter genera	None	WT	

Sort	First Reference	Current Name	Universal Name	Lineage	Mutagenic DNA Description	Genotype	Gene 1 mutation
34	Figure 4D	Evolved	CM029	Mutant of CI006	Disruption of nifL gene with a fragment of the region upstream of the ompX gene inserted (Prm5) and deletion of the 1287bp after the start codon of the glnE gene containing the adenylyl-removing domain of glutamate-ammonia-ligase adenylyltransferase (Δ glnE-AR_KO1).	Δ nifL::Prm5 Δ glnE-AR_KO1	ATGACCCTGAATATGATGATGGATGC CGGCGGACATCATCGCGACAAACAAT ATTAATACCGGCAACCACACCGGCAA TTTACGAGACTGCGCAGGCATCCTTT CTCCCGTCAATTTCTGTCAAATAAAG TAAAAGAGGCAGTCTACTTGAATTAC CCCCGGCTGGTTGAGCGTTTGTGAA AAAAAGTAACTGAAAAATCCGTAGA ATAGCGCCACTCTGATGGTTAATTAA CCTATTCAATTAAGAATTATCTGGAT GAATGTGCCATTAAATGCGCAGCATA ATGGTGCGTTGTGCGGGAAAACCTGCT TTTTTTTGAAAGGGTTGGTCACTAGC GGAAACAACCTCACTTACACCCCCGAA GGGGGAAGTTGCCTGACCCTACGATT CCCGCTATTTCACTTCACTGACCGGAG GTTCAAAATGA
35	Figure 14C	Wild	CI006	Isolated strain from Enterobacter genera	None	WT	
36	Figure 14C	Evolved	CM014	Mutant of CI006	Disruption of nifL gene with a fragment of the region upstream of the lpp gene and the first 29bp of the lpp gene inserted (Prm1).	Δ nifL::Prm1	ATGACCCTGAATATGATGATGGATGC CGGCCGCTGTAAATAAACCAGGAC AATTTCGGAAGTATTAATAAAGCGCCC TTGTGGCGCTTTTTTATATTCGCCG TCCATTTAAATAAAAAATCCAATCG GATTTCACTATTTAAACTGGCCATTAT CTAAGATGAATCCGATGGAAGCTCGC TGTTTTAACACGCGTTTTTAACTTT TATTGAAAGTCGGTGCTTCTTTGAGC GAACGATCAAATTTAAGTGGATTCCC ATCAAAAAATATTCTCAACCTAAAA AAGTTTGTGTAATACTTGTAAACGCTA CATGGAGATTAACTCAATCTAGAGGG TATTAATAATGAATCGTACTAAACTG GTACTGGGCGCAACTCACTTACACCC CCGAAGGGGAAGTTGCCTGACCCTA CGATTCCCGCTATTTCACTTCACTGACC GGAGGTTCAAAATGA
37	Figure 14B	Wild	CI019	Isolated strain from Rahnella genera	None	WT	
38	Figure 14B	Evolved	CM011	Mutant of CI019	Disruption of nifL gene with a spectinomycin resistance expression cassette (SpecR) encoding the streptomycin 3"-O-adenylyltran	Δ nifL::SpecR	ATGAGCATCACGGCGTTATCAGCATC ATTTCTGAGGGGAATATCGCCAGCC GCTTGTCGCTGCAACATCCTTCACTGT TTTATACCGTGTTGAACAATCTTCG GTGGCGAGCGTGTGAGTCATCCTGA CTAGCTGAGATGAGGGCTCGCCCCCT CGTCCCGACACTTCCAGATCGCCATA GCGCACAGCGCTCGAGCGGTGGTAA CGGCGCAGTGGCGGTTTTTCATGGCTT GTTATGACTGTTTTTTTGGGGTACAGT CTATGCCTCGGGCATCCAAGCAGCAA GCGCGTTACGCCGTGGGTCGATGTTT

Sort	First Reference	Current Name	Universal Name	Lineage	Mutagenic DNA Description	Genotype	Gene 1 mutation
					sferase gene aadA inserted.		GATGTTATGGAGCAGCAACGATGTTC CGCAGCAGGGCAGTCGCCCTAAACAA AAGTTAAACATCATGAGGGAAGCGGT GATCGCCGAAGTATCGACTCAACTAT CAGAGGTAGTTGGCGTCATCGAGCGC CATCTCGAACCGACGTTGCTGGCCGT ACATTTGTACGGCTCCGCAGTGGATG GCGGCCTGAAGCCACACAGTGATATT GATTTGCTGGTTACGGTGACCGTAAG GCTTGATGAAACAACGCGGCGAGCTT TGATCAACGACCTTTTGGAACTTCG GCTTCCCCTGGAGAGAGCGAGATTCT CCGCGCTGTAGAAGTCACCATTTGTTG TGCACGACGACATCATTCCGTGGCGT TATCCAGCTAAGCGCGAACTGCAATT TGGAGAATGGCAGCGCAATGACATTC TTGCAGGTATCTTCGAGCCAGCCAG ATCGACATTGATCTGGCTATCTTGCTG ACAAAAGCAAGAGAACATAGCGTTG CCTTGGTAGGTCCAGCGGCGGAGGAA CTCTTTGATCCGGTTCCTGAACAGGA TCTATTTGAGGCGCTAAATGAAACCT TAACGCTATGGAATCGCCGCCCGAC TGGGCTGGCGATGAGCGAAATGTAGT GCTTACGTTGTCCCGCATTGTTGAC GCGCAGTAACCGGCAAAATCGCGCCG AAGGATGTCGCTGCCGACTGGGCAAT GGAGCGCCTGCCGGCCAGTATCAGC CCGTCATACTTGAAGCTAGACAGGCT TATCTTGGACAAGAAGAAGATCGCTT GGCCTCGCGCGCAGATCAGTTGGAAG AATTTGTCCACTACGTGAAAGGCGAG ATCACCAAGGTAGTCGGCAAATAATG TCTAACAATTCGTTCAAGCCGACGCC GCTTCGCGGCGCGGCTTAACCAAGC GTTAGATGCACTAAGCACATAATTGC TCACAGCCAACTATCAGGTCAAGTC TGCTTTTATTATTTTAAAGCGTGCATA ATAAGCCCTACACAAATGGTACCCGA CCGGTGGTGAATTTAATCTCGCTGAC GTGTAGACATTCCCTTATCCAGACGC TGATCGCCCATCATCGCGGTTCTTAG ATCTCTCGGTCCGCCCTGATGGCGGC ACCTTGCTGACGTTACGCTGCCGGT ACAGCAGGTTATCACCAGGAGCTTAA AATGA
39	Figure 14A	Evolved	CM011	Mutant of CI019	Disruption of nifL gene with a spectinomycin resistance expression cassette (SpecR) encoding the streptomycin 3"-O-adenylyltransferase gene aadA inserted.	Δ nifL::SpeR	ATGAGCATCACGGCGTTATCAGCATC ATTTCTGAGGGGAATATCGCCAGCC GCTTGTGCTGCAACATCCTTCACTGT TTTATAACCGTGGTTGAACAATCTTCG GTGGCGAGCGTGTGAGTCATCCTGA CTAGCTGAGATGAGGGCTCGCCCCCT CGTCCCGACACTTCCAGATCGCCATA GCGCACAGCGCCTCGAGCGGTGGTAA CGGCGCAGTGGCGGTTTTTCATGGCTT GTTATGACTGTTTTTTTGGGGTACAGT CTATGCCTCGGGCATCCAAGCAGCAA GCGCGTTACGCCGTGGGTGATGTTT GATGTTATGGAGCAGCAACGATGTTC CGCAGCAGGGCAGTCGCCCTAAACAA AAGTTAAACATCATGAGGGAAGCGGT GATCGCCGAAGTATCGACTCAACTAT CAGAGGTAGTTGGCGTCATCGAGCGC CATCTCGAACCGACGTTGCTGGCCGT

Sort	First Reference	Current Name	Universal Name	Lineage	Mutagenic DNA Description	Genotype	Gene 1 mutation
							ACATTTGTACGGCTCCGCAGTGGATG GCGGCCTGAAGCCACACAGTGATATT GATTTGCTGGTTACGGTGACCGTAAG GCTTGATGAAACAACGCGGCGAGCTT TGATCAACGACCTTTTGGAACTTCG GCTTCCCCTGGAGAGAGCGAGATTCT CCGCGCTGTAGAAGTCACCATTGTTG TGCACGACGACATCATTCCGTGGCGT TATCCAGCTAAGCGCGAACTGCAATT TGGAGAATGGCAGCGCAATGACATT TTGCAGGTATCTTCGAGCCAGCCACG ATCGACATTGATCTGGCTATCTTGCTG ACAAAAGCAAGAGAACATAGCGTTG CCTTGGTAGGTCCAGCGGCGGAGGAA CTCTTTGATCCGGTTCCTGAACAGGA TCTATTTGAGGCGCTAAATGAAACCT TAACGCTATGGAACTCGCCGCCGAC TGGGCTGGCGATGAGCGAAATGTAGT GCTTACGTTGTCCCGCATTGTTGACA GCGCAGTAACCGGCAAAATCGCGCCG AAGGATGTCGCTGCCGACTGGGCAAT GGAGCGCCTGCCGGCCAGTATCAGC CCGTCATACTTGAAGCTAGACAGGCT TATCTTGGACAAGAAGAAGATCGCTT GGCCTCGCGCGCAGATCAGTTGGAAG AATTTGTCCACTACGTGAAAGGCGAG ATCACCAAGGTAGTCGGCAAATAATG TCTAACAATTCGTTCAAGCCGACGCC GCTTCGCGGCGCGGCTTAACTCAAGC GTTAGATGCACTAAGCACATAATTGC TCACAGCCAACTATCAGGTCAAGTC TGCTTTTATTATTTTAAAGCGTGATA ATAAGCCCTACACAAATGGTACCCGA CCGGTGGTGAATTTAATCTCGCTGAC GTGTAGACATTCCCTTATCCAGACGC TGATCGCCCATCATCGCGGTTCTTTAG ATCTCTCGGTCCGCCCTGATGGCGGC ACCTTGCTGACGTTACGCTGCCGGT ACAGCAGGTTATCACCGGAGGCTTAA AATGA
40	Figure 15A	Wild	CI006	Isolated strain from Enterobacter genera	None	WT	
41	Figure 15A	Evolved	CM013	Mutant of CI006	Disruption of nifL gene with a kanamycin resistance expression cassette (KanR) encoding the aminoglycoside O-phosphotransferase gene aphI inserted.	Δ nifL::KanR	CTGATCCTTCAACTCAGCAAAAGTTT GATTTATTCAACAAAGCCACGTTGTG TCTCAAAATCTCTGATGTTACATTGCA CAAGATAAAAATATATCATCATGAAC AATAAACTGTCTGCTTACATAAACA GTAATACAAGGGGTGTTATGAGCCAT ATTCAACGGGAAACGTCTTGCTCCAG GCCGCGATTAAATCCAACATGGATG CTGATTTATATGGGTATAAATGGGCT CGCGATAATGTCGGGCAATCAGGTGC GACAATCTATCGATTGTATGGGAAGC CCGATGCGCCAGAGTTGTTTCTGAAA CATGGCAAAGGTAGCGTTGCCAATGA TGTTACAGATGAGATGGTCAGACTAA ACTGGCTGACGGAATTTATGCCTCTT CCGACCATCAAGCATTTTATCCGTAC TCCTGATGATGCATGGTTACTACCA CTGCGATCCCCGGGAAAACAGCATTC CAGGTATTAGAAGAATATCCTGATTC AGGTGAAAATATTGTTGATGCGCTGG

Sort	First Reference	Current Name	Universal Name	Lineage	Mutagenic DNA Description	Genotype	Gene 1 mutation
							CAGTGTTCTCTGCGCCGGTTGCATTTCG ATTCTGTGTTTGTAAATTGTCCTTTTAAC AGCGATCGCGTATTTTCGTCTCGCTCA GGCGCAATCACGAATGAATAACGGTT TGGTTGATGCGAGTGATTTTGATGAC GAGCGTAATGGCTGGCCTGTTGAACA AGTCTGGAAAGAAATGCATAAGCTTT TGCCATTCTCACCGGATTTCAGTCGTC ACTCATGGTGATTCTCACTTGATAAC CTTATTTTTGACGAGGGGAAATTAAT AGGTTGTATTGATGTTGGACGAGTCG GAATCGCAGACCGATAACCAGGATCTT GCCATCCTATGGAAGTGCCTCGGTGA GTTTTCTCCTTCATTACAGAAACGGCT TTTTCAAAAATATGGTATTGATAATC CTGATATGAATAAATTGCAGTTTCAT TTGATGCTCGATGAGTTTTTCTAATAA GCCTTGACCCTACGATTCCCGCTATTT CATTCACTGACCGGAGGTTCAAAATG A
42	Figure 15B	No name	CM011	Mutant of CI019	Disruption of nifL gene with a spectinomycin resistance expression cassette (SpecR) encoding the streptomycin 3"-O-adenylyltransferase gene aadA inserted.	Δ nifL::SpecR	ATGAGCATCACGGCGTTATCAGCATC ATTTCTGAGGGGAATATCGCCAGCC GCTTGTCGCTGCAACATCCTTCACTGT TTTATACCGTGGTTGAACAACTTCG GTGGCGAGCGTGTGATCATCCTGA CTAGCTGAGATGAGGGCTCGCCCCCT CGTCCCGACACTTCCAGATCGCCATA GCGCACAGCGCCTCGAGCGGTGGTAA CGGCGCAGTGGCGGTTTTTCATGGCTT GTTATGACTGTTTTTTGGGGTACAGT CTATGCCTCGGGCATCAAGCAGCAA GCGCGTTACGCCGTGGGTGATGTTT GATGTTATGGAGCAGCAACGATGTTA CGCAGCAGGGCAGTCGCCCTAAAACA AAGTTAAACATCATGAGGGAAGCGGT GATCGCCGAAGTATCGACTCAACTAT CAGAGGTAGTTGGCGTCATCGAGCGC CATCTCGAACCAGCGTTGCTGGCCGT ACATTTGTACGGCTCCGCGAGTGGATG GCGGCCTGAAGCCACACAGTGATATT GATTTGCTGGTTACGGTGACCGTAAG GCTTGATGAAACAACGCGGCGAGCTT TGATCAACGACCTTTTGAAACTTCG GCTTCCCCTGGAGAGAGCGAGATTCT CCGCGCTGTAGAAGTCACCATTGTTG TGCACGACGACATCATTCGGTGGCGT TATCCAGCTAAGCGCAACTGCAATT TGGAGAATGGCAGCGCAATGACATT TTGCAGGTATCTTCGAGCCAGCCACG ATCGACATTGATCTGGCTATCTTGCTG ACAAAAGCAAGAGAACATAGCGTTG CCTTGGTAGGTCCAGCGGCGGAGGAA CTCTTTGATCCGGTTCCTGAACAGGA TCTATTTGAGGCGCTAAATGAAACCT TAACGCTATGGAAGTCCGCCGCCGAC TGGGCTGGCGATGAGCGAAATGTAGT GCTTACGTTGTCCCGCATTTGGTACA GCGCAGTAACCGGCAAAATCGCGCCG AAGGATGTCGCTGCCGACTGGGCAAT GGAGCGCCTGCCGGCCAGTATCAGC CCGTCATACTTGAAGCTAGACAGGCT TATCTTGGACAAGAAGAAGATCGCTT GGCTCGCGCGCAGATCAGTTGGAAG AATTTGTCCACTACGTGAAAGGCGAG

Sort	First Reference	Current Name	Universal Name	Lineage	Mutagenic DNA Description	Genotype	Gene 1 mutation
							ATCACCAAGGTAGTCGGCAAATAATG TCTAACAATTCGTTCAAGCCGACGCC GCTTCGCGGCGCGGCTTAACTCAAGC GTTAGATGCACTAAGCACATAATTGC TCACAGCCAACTATCAGGTCAAGTC TGCTTTTATTATTTTAAAGCGTGCATA ATAAGCCCTACACAAATGGTACCCGA CCGGTGGTGAATTTAATCTCGCTGAC GTGTAGACATTCCCTTATCCAGACGC TGATCGCCCATCATCGCGGTTCTTAG ATCTCTCGGTCCGCCCTGATGGCGGC ACCTTGCTGACGTTACGCCTGCCGGT ACAGCAGGTTATCACCGGAGGCTTAA AATGA
43	Figure 16B	Strain 5	CI008	Isolated strain from Burkholderia genera	None	WT	
44	Figure 16B	Strain 1	CM011	Mutant of CI019	Disruption of nifL gene with a spectinomycin resistance expression cassette (SpecR) encoding the streptomycin 3"-O-adenylyltransferase gene aadA inserted.	Δ nifL::SpeR	ATGAGCATCACGGCGTTATCAGCATC ATTTCTGAGGGGAATATCGCCAGCC GCTTGTCGCTGCAACATCCTTCACTGT TTTATACCGTGTTGAACAATCTTCG GTGGCGAGCGTGTGAGTCATCCTGA CTAGCTGAGATGAGGGCTCGCCCCCT CGTCCCGACACTTCCAGATCGCCATA GCGCACAGCGCCTCGAGCGGTGGTAA CGGCGCAGTGGCGGTTTTCATGGCTT GTTATGACTGTTTTTTTGGGGTACAGT CTATGCCTCGGGCATCCAAGCAGCAA GCGCGTTACGCCGTGGGTGATGTTT GATGTTATGGAGCAGCAACGATGTTA CGCAGCAGGGCAGTCGCCCTAAAACA AAGTTAAACATCATGAGGGAAGCGGT GATCGCCGAAGTATCGACTCAACTAT CAGAGGTAGTTGGCGTCATCGAGCGC CATCTCGAACCAGCGTTGCTGGCCGT ACATTTGTACGGCTCCGCAGTGGATG GCGGCCTGAAGCCACACAGTGATATT GATTTGCTGGTTACGGTGACCGTAAG GCTTGATGAAACAACGCGGCGAGCTT TGATCAACGACCTTTTGAAACTTCG GCTTCCCCTGGAGAGAGCGAGATTCT CCGCGCTGTAGAAGTACCATTGTTG TGCACGACGACATATTCCGTGGCGT TATCCAGCTAAGCGGCAACTGCAATT TGGAGAATGGCAGCGCAATGACATTC TTGCAGGTATCTTCGAGCCAGCCACG ATCGACATTGATCTGGCTATCTTGCTG ACAAAAGCAAGAGAACATAGCGTTG CCTTGGTAGGTCCAGCGGCGGAGGAA CTCTTTGATCCGGTTCCTGAACAGGA TCTATTTGAGGCGCTAAATGAAACCT TAACGCTATGGAAGTCCGCGCCGAC TGGGCTGGCGATGAGCGAAATGTAAGT GCTTACGTTGTCCCGCATTTGGTACA GCGCAGTAACCGGCAAAATCGCGCCG AAGGATGTCGCTGCCGACTGGGCAAT GGAGCGCCTGCCGCGCCAGTATCAGC CCGTCATACTTGAAGCTAGACAGGCT TATCTTGGACAAGAAGAAGATCGCTT GGCTCGCGCGCAGATCAGTTGGAAG AATTTGTCCACTACGTGAAAGGCGAG ATCACCAAGGTAGTCGGCAAATAATG TCTAACAATTCGTTCAAGCCGACGCC

Sort	First Reference	Current Name	Universal Name	Lineage	Mutagenic DNA Description	Genotype	Gene 1 mutation
							GCTTCGCGGCGCGGGCTTAACTCAAGC GTTAGATGCACTAAGCACATAATTGC TCACAGCCAAACTATCAGGTCAAGTC TGCTTTTATTATTTTAAAGCGTGCATA ATAAGCCCTACACAAATGGTACCCGA CCGGTGGTGAATTTAATCTCGCTGAC GTGTAGACATTCCCTTATCCAGACGC TGATCGCCCATCATCGCGGTTCTTTAG ATCTCTCGGTCCGCCCTGATGGCGGC ACCTTGCTGACGTTACGCCTGCCGGT ACAGCAGGTTATCACCGGAGGCTTAA AATGA

Table of Strains (cont'd)

Sort	First Reference	Current Name	Universal Name	Lineage	Mutagenic DNA Description	Genotype	Gene 2 mutation
34	Figure 4D	Evolved	CM029	Mutant of CI006	Disruption of nifL gene with a fragment of the region upstream of the ompX gene inserted (Prm5) and deletion of the 1287bp after the start codon of the glnE gene containing the adenylyl-removing domain of glutamate-ammonia-ligase adenylyltransferase (Δ glnE-AR_KO1).	Δ nifL::Prm5 Δ glnE-AR_KO1	ATGTTTAACGATCTGATTGGCGATGATGAAACGGATTGCGCGGAAGATGCGCTTTCTGAGAGCTGGCGCGAATTGTGGCAGGATGCGTTGCAGGAGGAGGATTCCACGCCCCGTGCTGGCGCATCTCTCAGAGGACGATCGCCGCCGCTGGTGGCGCTGATTGCCGATTTTCGCAAAGAGTTGGATAAACGCACCATTTGGCCCCGCGAGGGCGGCAGGTACTCGATCACTTAATGCCGATCTGCTCAGCGATGTATGCTCGCGGACGATGCGCCAGTACCGCTGTACGCCTGACGCCGCTGCTACCCGGAATTATTACCCGCACCACTTACCTTGAAGCTGCTAAGTGAATTTCCCGGCGCACGTGAAACACCTCATTTCCCTGTGTGCCGCGTCGCCGATGGTTGCCAGTCAGCTGGCGCGCTACCCGATCCTGCTTGATGAATTGCTCGACCCGAATACGCTCTATCAACCGACGGCGATGAATGCCTATCGCGATGAGCTGCGCCAATACCTGCTGCGGTGCCGGAAGATGATGAAGAGCAACAGCTTGAGGCGCTGCGGCAGTTTAAAGCAGGCGCAGTTGCTGCGCGTGGCGGCGGCGGATATTGCCGGTACGTTGCCAGTAATGAAAGTGAGCGATCACTTAACTGGCTGGCGGAAGCGATTATTGATGCGGTGGTGCAGCAAGCCTGGGGGAGATGGTGGCGCGTTATGGCCAGCCAACGCATCTGCACGATCGCGAAGGGCGCGGTTTTGCGGTGGTTCGGTTATGGCAAGCTGGGCGGCTGGGAGCTGGGTACAGCTCCGATCTGGATCTGGTATTCTGCACGACTGCCGATGGATGTGATGACCGATGGCGAGCGTGAAATCGATGGTCGCCAGTTCTATTTGCGTCTCGCGCAGCGGTGATGCACCTGTTTAGCACGCGCACGTCGTCGGCATCCTTTATGAAGTTGATGCGCGTCTGCGTCCATCTGGCGCTGCGGGGATGCTGGTCACTACTACGGAATCGTTCGCCGATTACCAGCAAACGAAGCCTGGACGTGGGAACATCAGGCGCTGGCCCGTGCGCGCGTGGGTGACGGCGATCCGCAACTGACCGCCGAATTTGACGCCATTTCGCCGCGATATCTGATGACGCCTCGCGACGGCGCAACGCTGCAAACCGACGTGCGAGAAATGCGGAGAAAATGCGTGCCCATCTTGGCAACAAGCATAAAGACCGCTTCGATCTGAAAGCCGATGAAGGCGGTATCACCAGCATCGAGTTTATCGCCCAATATCTGGTGCTGCGCTTTGCCCATGACAAGCCGAAACTGACGCGCTGGTTCGGATAATGTGCGCATTCTCGAAGGGCTGGCGCAAAACGGCATCATGGAGGAGCAGGAAAGCGCAGGCATTGACGCTGGCGTACACCACATTGCGTGATGAGCTGCACCACCTGGCGCTGCAAGAGTTGCCGGGACATGTGGCGCTCTCCTGTTTTGTGCGCGACGCTGCGCTTATTAAAACAGCTGGGACAAGTGGCTGGTGAACCGTGCGCCCGGCGTAA

Notwithstanding the appended claims, the disclosure set forth herein is also defined by the following clauses:

1. A method of producing one or more bacteria, comprising:
 - (a) isolating bacteria from tissue or soil of a first plant;
 - (b) introducing genetic variation into one or more of the bacteria to produce one or more variant bacteria;
 - (c) exposing a plurality of plants to the variant bacteria;
 - (d) isolating bacteria from tissue or soil of one of the plurality of plants, wherein the plant from which the bacteria is isolated has an improved trait relative to other plants in the plurality of plants; and
 - (e) repeating steps (b) to (d) with bacteria isolated in step (d).
2. The method of clause 1, wherein the improved trait is enhanced nitrogen fixation in the plant from which bacteria are isolated.
3. The method of clause 1, wherein the genetic variation is a variation in a gene selected from the group consisting of: *nifA*, *nifL*, *ntrB*, *ntrC*, *glnA*, *glnB*, *glnK*, *draT*, *amtB*, *glnD*, *glnE*, *nifJ*, *nifH*, *nifD*, *nifK*, *nifY*, *nifE*, *nifN*, *nifU*, *nifS*, *nifV*, *nifW*, *nifZ*, *nifM*, *nifF*, *nifB*, and *nifQ*.
4. The method of clause 1, wherein the genetic variation is a variation in a gene encoding a protein with functionality selected from the group consisting of: glutamine synthetase, glutaminase, glutamine synthetase adenylyltransferase, transcriptional activator, anti-transcriptional activator, pyruvate flavodoxin oxidoreductase, flavodoxin, or NAD⁺-dinitrogen-reductase ADP-D-ribosyltransferase.
5. The method of clause 1, wherein the genetic variation is a mutation that results in one or more of: increased expression or activity of *NifA* or glutaminase; decreased expression or activity of *NifL*, *NtrB*, glutamine synthetase, *GlnB*, *GlnK*, *DraT*, *AmtB*; decreased adenylyl-removing activity of *GlnE*; or decreased uridylyl-removing activity of *GlnD*.
6. The method of clause 1, wherein the genetic variation is a knock-out mutation.
7. The method of clause 1, wherein the genetic variation results in elimination or abolishment of activity of a protein domain.
8. The method of clause 1, wherein the genetic variation alters or abolishes a regulatory sequence of a target gene.
9. The method of clause 1, wherein the genetic variation comprises insertion of a heterologous regulatory sequence.

10. The method of clause 1, wherein the genetic variation comprises insertion of a regulatory sequence found within a genome of a bacterial species or genus corresponding to the bacteria into which the genetic variation is introduced.

11. The method of clause 10, wherein the regulatory sequence is selected based on expression level of a gene in a bacterial culture or within plant tissue.

12. The method of clause 1, wherein the genetic variation is produced by chemical mutagenesis.

13. The method of clause 1, wherein step (c) further comprises exposing the plants to biotic or abiotic stressors.

14. The method of clause 2, wherein bacteria isolated after repeating steps (b) to (d) one or more times produce 1% or more of nitrogen in a second plant of the same type as the first plant.

15. The method of clause 2, wherein bacteria isolated after repeating steps (b) to (d) one or more times exhibit at least a 2-fold increase in nitrogen fixation as compared to bacteria isolated from the first plant.

16. The method of clause 14, wherein the second plant is grown in the presence of fertilizer supplemented with glutamine, ammonia, or other chemical source of nitrogen.

17. The method of clause 1, wherein the first plant is an agricultural crop plant.

18. The method of clause 17, wherein the agricultural crop plant is selected from barley, rice, maize, wheat, sorghum, sweet corn, sugar cane, onions, tomatoes, strawberries, or asparagus.

19. The method of clause 1, wherein the first or plants in the plurality of plants are a model plant.

20. The method of clause 19, wherein the model plant is selected from *Setaria*, *Brachypodium*, or *Arabidopsis*.

21. The method of clause 1, wherein the genetic variation is a pre-determined genetic variation that is specifically introduced to a target site.

22. The method of clause 1, wherein the genetic variation is a random mutation within the target site.

23. The method of clause 1, wherein step (a) further comprises performing genetic analysis of isolated bacteria.

24. The method of clause 1, wherein step (b) further comprises applying a selection pressure to enrich for bacteria comprising the genetic variation.

25. The method of clause 24, wherein the selection pressure comprises cleaving genomes lacking the genetic variation introduced to a target site, wherein cleavage occurs within 100 nucleotides of the target site.

26. The method of clause 24, further comprising isolating bacteria that survive the selection pressure.

27. The method of clause 25, wherein cleavage is directed by a site-specific nuclease selected from the group consisting of a Zinc Finger nuclease, a CRISPR nuclease, a TALE nuclease, or a meganuclease.

28. The method of clause 27, wherein the site-specific nuclease is a CRISPR nuclease.

29. The method of clause 1, wherein the genetic variation is an insertion or deletion of one or more nucleotides.

30. The method of clause 1, wherein bacteria isolated after repeating steps (b) to (d) one or more times are endophytic, epiphytic, or rhizospheric.

31. The method of clause 1, wherein bacteria isolated after repeating steps (b) to (d) one or more times comprise a plurality of different bacterial taxa.

32. The method of clause 1, wherein the bacteria are isolated from plant tissue.

33. The method of clause 1, wherein isolating bacteria in step (a) comprises isolating bacteria from a seed of the first plant.

34. A method of increasing nitrogen fixation in a plant, comprising exposing the plant to bacteria comprising one or more genetic variations introduced into one or more genes regulating nitrogen fixation, wherein the bacteria produce 1% or more of nitrogen in the plant.

35. The method of clause 34, wherein the bacteria produce 5% or more of nitrogen in the plant.

36. The method of clause 34, wherein the bacteria produce 10% or more of nitrogen in the plant.

37. The method of clause 34, wherein the bacteria produce the nitrogen in the presence of fertilizer supplemented with glutamine, ammonia, or other chemical source of supplemental nitrogen.

38. The method of clause 34, wherein the genetic variation is a variation in a gene selected from the group consisting of: *nifA*, *nifL*, *ntrB*, *ntrC*, glutamine synthetase, *glnA*, *glnB*, *glnK*, *draT*, *amtB*, glutaminase, *glnD*, *glnE*, *nifJ*, *nifH*, *nifD*, *nifK*, *nifY*, *nifE*, *nifN*, *nifU*, *nifS*, *nifV*, *nifW*, *nifZ*, *nifM*, *nifF*, *nifB*, and *nifQ*.

39. The method of clause 34, wherein the genetic variation is a mutation that results in one or more of: increased expression or activity of *nifA* or glutaminase; decreased expression or

activity of nifL, ntrB, glutamine synthetase, glnB, glnK, draT, amtB; decreased adenylyl-removing activity of GlnE; or decreased uridylyl-removing activity of GlnD.

40. The method of clause 34, wherein the genetic variation (a) is a knock-out mutation; (b) alters or abolishes a regulatory sequence of a target gene; or (c) comprises insertion of a heterologous regulatory sequence.

41. The method of clause 34, wherein the bacteria are of the genus *Enterobacter*.

42. The method of clause 34, wherein the bacteria are of the genus *Rahnella*.

43. The method of claim 34, wherein the bacteria are endophytic, epiphytic, or rhizospheric.

44. The method of clause 34, wherein the bacteria comprise a plurality of different bacterial taxa.

45. The method of clause 34, wherein the plant is an agricultural crop plant.

46. The method of any one of clauses 34-45, wherein the plant is a non-leguminous plant.

47. The method of clause 45, wherein the agricultural crop plant is selected from sorghum, canola, tomato, strawberry, barley, rice, maize, and wheat.

48. The method of clause 45, wherein the plant is a genetically modified organism (GMO).

49. The method of clause 45, wherein the plant is not a genetically modified organism (GMO).

50. The method of clause 45, wherein the plant has been genetically engineered or bred for efficient nitrogen use.

51. A bacterial population comprising bacteria comprising one or more genetic variations introduced into one or more genes regulating nitrogen fixation, wherein the bacteria produce 1% or more of nitrogen in a plant grown in the presence of the population of bacteria.

52. The bacterial population of clause 51, wherein the bacteria produce the nitrogen in the presence of fertilizer supplemented with glutamine, ammonia, or other chemical source of supplemental nitrogen.

53. The bacterial population of clause 51, wherein the genetic variation is a variation in a gene selected from the group consisting of: nifA, nifL, ntrB, ntrC, glutamine synthetase, glnA, glnB, glnK, draT, amtB, glutaminase, glnD, glnE, nifJ, nifH, nifD, nifK, nifY, nifE, nifN, nifU, nifS, nifV, nifW, nifZ, nifM, nifF, nifB, and nifQ.

54. The bacterial population of clause 51, wherein the genetic variation is a mutation that results in one or more of: increased expression of nifA or glutaminase; decreased expression of

nifL, ntrB, glutamine synthetase, glnB, glnK, draT, amtB; decreased adenylyl-removing activity of GlnE; or decreased uridylyl-removing activity of GlnD.

55. The bacterial population of clause 51, wherein the genetic variation (a) is a knock-out mutation; (b) alters or abolishes a regulatory sequence of a target gene; or (c) comprises insertion of a heterologous regulatory sequence.

56. The bacterial population of clause 51, wherein the bacteria are Enterobacter.

57. The bacterial population of clause 51, wherein the bacteria are Rahnella.

58. The bacterial population of clause 51, wherein the bacteria are endophytic, epiphytic, or rhizospheric.

59. The bacterial population of clause 51, wherein bacteria comprise a plurality of different bacterial taxa.

60. A composition comprising the bacterial population of any one of clauses 51-59.

61. The composition of clause 60, wherein the composition comprises the bacterial population coated on a surface of a seed.

62. The composition of clause 60, wherein the composition is formulated as a liquid or powder.

63. A bacterium having an ATCC deposit number of PTA-122293 or PTA-122294.

CLAIMS

WHAT IS CLAIMED IS:

1. A method of increasing nitrogen fixation in a non-leguminous plant, comprising:
exposing the plant to a plurality of bacteria, each member of the plurality comprising one or more genetic variations introduced into one or more genes or non-coding polynucleotides of the bacteria's nitrogen fixation or assimilation genetic regulatory network, such that the bacteria are capable of fixing atmospheric nitrogen in the presence of exogenous nitrogen; wherein the bacteria are not intergeneric microorganisms; and wherein the bacteria, *in planta*, produce 1% or more of the fixed nitrogen in the plant.
2. The method of claim 1, wherein the one or more genetic variations comprise an introduced control sequence operably linked to said one or more genes of the nitrogen fixation or assimilation genetic regulatory network.
3. The method of claim 2, wherein the control sequence is a promoter.
4. The method of claim 3, wherein the promoter is an inducible promoter.
5. The method of claim 1, wherein the bacteria do not comprise a constitutive promoter operably linked to a gene of the nitrogen fixation or assimilation genetic regulatory network.
6. The method of claim 1, wherein the bacteria do not comprise a constitutive promoter operably linked to a gene in the *nif* gene cluster.
7. The method of claim 1, wherein the bacteria, *in planta*, excrete the nitrogen-containing products of nitrogen fixation.
8. The method of claim 1, wherein the plurality of bacteria exposed to the plant do not stimulate an increase in the uptake of exogenous non-atmospheric nitrogen.
9. The method of claim 1, wherein the plant is grown in soil from a field which has been administered at least 50 lbs of nitrogen-containing fertilizer per acre, and wherein the nitrogen-containing fertilizer comprises at least 5% nitrogen by weight.

10. The method of claim 9, wherein the nitrogen-containing fertilizer comprises ammonium or an ammonium containing molecule.
11. The method of claim 1, wherein the plurality of bacteria comprise at least two different species of bacteria.
12. The method of claim 1, wherein the plurality of bacteria comprise at least two different strains of the same species of bacteria.
13. The method of claim 1, wherein the exogenous nitrogen is selected from fertilizer comprising one or more of glutamine, ammonia, ammonium, urea, nitrate, nitrite, ammonium-containing molecules, nitrate-containing molecules, and nitrite-containing molecules.
14. The method of claim 1, wherein the bacteria, *in planta*, produce 5% or more of the fixed nitrogen in the plant.
15. The method of claim 1, wherein the bacteria, *in planta*, produce 10% or more of the fixed nitrogen in the plant.
16. The method of claim 1, wherein the one or more genes or non-coding polynucleotides of the bacteria's nitrogen fixation or assimilation genetic regulatory network are selected from the group consisting of: *nifA*, *nifL*, *ntrB*, *ntrC*, polynucleotide encoding glutamine synthetase, *glnA*, *glnB*, *glnK*, *drat*, *amtB*, polynucleotide encoding glutaminase, *glnD*, *glnE*, *nifJ*, *nifH*, *nifD*, *nifK*, *nifY*, *nifE*, *nifN*, *nifU*, *nifS*, *nifV*, *nifW*, *nifZ*, *nifM*, *nifF*, *nifB*, and *nifQ*.
17. The method of claim 1, wherein the one or more genetic variations is a mutation that results in one or more of: increased expression or activity of NifA or glutaminase; decreased expression or activity of NifL, NtrB, glutamine synthetase, GlnB, GlnK, DraT, AmtB; decreased adenylyl-removing activity of GlnE; or decreased uridylyl-removing activity of GlnD.
18. The method of claim 1, wherein the one or more genetic variations is (A) a knock-out mutation; (B) alters or abolishes a regulatory sequence of a target gene; or (C) comprises the insertion of a heterologous regulatory sequence.

19. The method of claim 1, wherein the plurality of bacteria are of a genus comprising *Enterobacter*, *Rahnella*, *Kosakonia*, *Burkholderia* or *Klebsiella*.
20. The method of claim 1, wherein the plurality of bacteria are endophytic, epiphytic, or rhizospheric.
21. The method of claim 1, wherein the plurality of bacteria colonize at least the plant roots such that the bacteria are present in the plant in an amount of at least about 10^5 cfu per gram fresh weight of tissue.
22. The method of claim 1, wherein the plant is an agricultural crop plant.
23. The method of claim 22, wherein the agricultural crop plant is selected from sorghum, canola, tomato, strawberry, barley, rice, maize, and wheat.
24. The method of claim 22, wherein the plant is a genetically modified organism.
25. The method of claim 22, wherein the plant is not a genetically modified organism.
26. The method of claim 22, wherein the plant has been genetically engineered or bred for efficient nitrogen use.
27. A bacterial population comprising bacteria comprising one or more genetic variations introduced into one or more genes or non-coding polynucleotides of the bacteria's nitrogen fixation or assimilation genetic regulatory network, such that the bacteria are capable of fixing atmospheric nitrogen in the presence of exogenous nitrogen; wherein the bacteria are not intergeneric microorganisms; and wherein the bacteria, *in planta*, produce 1% or more of the fixed nitrogen in a plant grown in the presence of the population of bacteria.
28. The bacterial population of claim 27, wherein the one or more genetic variations comprise an introduced control sequence operably linked to said one or more genes of the nitrogen fixation or assimilation genetic regulatory network.

29. The bacterial population of claim 28, wherein the control sequence is a promoter.
30. The bacterial population of claim 29, wherein the promoter is an inducible promoter.
31. The bacterial population of claim 27, wherein the bacteria do not comprise a constitutive promoter operably linked to a gene of the nitrogen fixation or assimilation genetic regulatory network.
32. The bacterial population of claim 27, wherein the bacteria do not comprise a constitutive promoter operably linked to a gene in the *nif* gene cluster.
33. The bacterial population of claim 27, wherein the bacteria, *in planta*, excrete the nitrogen-containing products of nitrogen fixation.
34. The bacterial population of claim 27, wherein the bacteria exposed to the plant do not stimulate an increase in the uptake of exogenous non-atmospheric nitrogen.
35. The bacterial population of claim 27, wherein the bacterial population comprise at least two different species of bacteria.
36. The bacterial population of claim 27, wherein the bacterial population comprise at least two different strains of the same species of bacteria.
37. The bacterial population of claim 27, wherein the exogenous nitrogen is selected from fertilizer comprising one or more of glutamine, ammonia, ammonium, urea, nitrate, nitrite, ammonium-containing molecules, nitrate-containing molecules, and nitrite-containing molecules.
38. The bacterial population of claim 27, wherein the bacteria, *in planta*, produce 5% or more of the fixed nitrogen in the plant.
39. The bacterial population of claim 27, wherein the bacteria, *in planta*, produce 10% or more of the fixed nitrogen in the plant.

40. The bacterial population of claim 27, wherein the one or more genes or non-coding polynucleotides of the bacteria's nitrogen fixation or assimilation genetic regulatory network are selected from the group consisting of: *nifA*, *nifL*, *ntrB*, *ntrC*, polynucleotide encoding glutamine synthetase, *glnA*, *glnB*, *glnK*, *drat*, *amtB*, polynucleotide encoding glutaminase, *glnD*, *glnE*, *nifJ*, *nifH*, *nifD*, *nifK*, *nifY*, *nifE*, *nifN*, *nifU*, *nifS*, *nifV*, *nifW*, *nifZ*, *nifM*, *nifF*, *nifB*, and *nifQ*.
41. The bacterial population of claim 27, wherein the one or more genetic variations is a mutation that results in one or more of: increased expression or activity of NifA or glutaminase; decreased expression or activity of NifL, NtrB, glutamine synthetase, GlnB, GlnK, DraT, AmtB; decreased adenylyl-removing activity of GlnE; or decreased uridylyl-removing activity of GlnD.
42. The bacterial population of claim 27, wherein the one or more genetic variations is (A) a knock-out mutation; (B) alters or abolishes a regulatory sequence of a target gene; or (C) comprises the insertion of a heterologous regulatory sequence.
43. The bacterial population of claim 27, wherein the plurality of bacteria are of a genus comprising *Enterobacter*, *Rahnella*, *Kosakonia*, *Burkholderia* or *Klebsiella*.
44. The bacterial population of claim 27, wherein the plurality of bacteria are endophytic, epiphytic, or rhizospheric.
45. The bacterial population of claim 27, wherein the plant is an agricultural crop plant.
46. The bacteria population of claim 45, wherein the agricultural crop plant is selected from sorghum, canola, tomato, strawberry, barley, rice, maize, and wheat.
47. The bacterial population of claim 45, wherein the plant is a genetically modified organism.
48. The bacterial population of claim 45, wherein the plant is not a genetically modified organism.
49. The bacterial population of claim 45, wherein the plant has been genetically engineered or bred for efficient nitrogen use.

50. The bacterial population of claim 27, wherein the plurality of bacteria colonize the plant such that the bacteria are present in the plant in an amount of at least about 10^5 cfu per gram of fresh weight of the plant.
51. A composition comprising the bacterial population of any one of claims 27-50.
52. The composition of claim 51, wherein the composition comprises the bacterial population coated on a surface of a seed.
53. The composition of claim 51, wherein the composition is formulated as a liquid or powder.
54. An isolated bacterium deposited as ATCC Accession Deposit No. PTA-122293 or PTA-122294.
55. A non-intergeneric bacterium comprising one or more genetic variations introduced into one or more genes or non-coding polynucleotides of the bacteria's nitrogen fixation or assimilation genetic regulatory network, such that the bacterium is capable of fixing atmospheric nitrogen in the presence of exogenous nitrogen.
56. The bacterium of claim 55, wherein the one or more genetic variations comprise an introduced control sequence operably linked to said one or more genes of the nitrogen fixation or assimilation genetic regulatory network.
57. The bacterium of claim 56, wherein the control sequence is a promoter.
58. The bacterium of claim 57, wherein the promoter is an inducible promoter.
59. The bacterium of claim 55, wherein the bacterium does not comprise a constitutive promoter operably linked to a gene of the nitrogen fixation or assimilation genetic regulatory network.

60. The bacterium of claim 55, wherein the bacterium does not comprise a constitutive promoter operably linked to a gene in the *nif* gene cluster.

61. The bacterium of claim 55, wherein the one or more genes or non-coding polynucleotides of the bacterium's nitrogen fixation or assimilation genetic regulatory network are selected from the group consisting of: *nifA*, *nifL*, *ntrB*, *ntrC*, polynucleotide encoding glutamine synthetase, *glnA*, *glnB*, *glnK*, *drat*, *amtB*, polynucleotide encoding glutaminase, *glnD*, *glnE*, *nifJ*, *nifH*, *nifD*, *nifK*, *nifY*, *nifE*, *nifN*, *nifU*, *nifS*, *nifV*, *nifW*, *nifZ*, *nifM*, *nifF*, *nifB*, and *nifQ*.

62. The bacterium of claim 55, wherein the one or more genetic variations is a mutation that results in one or more of: increased expression or activity of NifA or glutaminase; decreased expression or activity of NifL, NtrB, glutamine synthetase, GlnB, GlnK, DraT, AmtB; decreased adenylyl-removing activity of GlnE; or decreased uridylyl-removing activity of GlnD.

63. The bacterium of claim 55, wherein the one or more genetic variations is (A) a knock-out mutation; (B) alters or abolishes a regulatory sequence of a target gene; or (C) comprises the insertion of a heterologous regulatory sequence.

64. The bacterium of claim 55, wherein the bacterium is from a genus comprising *Enterobacter*, *Rahnella*, *Kosakonia*, *Burkholderia* or *Klebsiella*.

65. The bacterium of claim 55, wherein the bacterium is endophytic, epiphytic, or rhizospheric.

66. A method of producing one or more bacteria, comprising:

- (a) isolating bacteria from tissue or soil of a first plant;
- (b) introducing genetic variation into one or more of the bacteria to produce one or more variant bacteria;
- (c) exposing a plurality of plants to the variant bacteria;
- (d) isolating bacteria from tissue or soil of one of the plurality of plants, wherein the plant from which the bacteria is isolated has an improved trait relative to other plants in the plurality of plants; and
- (e) repeating steps (b) to (d) with bacteria isolated in step (d).

67. The method of claim 66, wherein the improved trait is enhanced nitrogen fixation in the plant from which bacteria are isolated.
68. The method of claim 66, wherein the genetic variation is a variation in a gene selected from the group consisting of: *nifA*, *nifL*, *ntrB*, *ntrC*, *glnA*, *glnB*, *glnK*, *draT*, *amtB*, *glnD*, *glnE*, *nifJ*, *nifH*, *nifD*, *nifK*, *nifY*, *nifE*, *nifN*, *nifU*, *nifS*, *nifV*, *nifW*, *nifZ*, *nifM*, *nifF*, *nifB*, and *nifQ*.
69. The method of claim 66, wherein the genetic variation is a variation in a gene encoding a protein with functionality selected from the group consisting of: glutamine synthetase, glutaminase, glutamine synthetase adenylyltransferase, transcriptional activator, anti-transcriptional activator, pyruvate flavodoxin oxidoreductase, flavodoxin, or NAD⁺-dinitrogen-reductase ADP-D-ribosyltransferase.
70. The method of claim 66, wherein the genetic variation is a mutation that results in one or more of: increased expression or activity of *NifA* or glutaminase; decreased expression or activity of *NifL*, *NtrB*, glutamine synthetase, *GlnB*, *GlnK*, *DraT*, *AmtB*; decreased adenylyl-removing activity of *GlnE*; or decreased uridylyl-removing activity of *GlnD*.
71. The method of claim 66, wherein the genetic variation is a knock-out mutation.
72. The method of claim 66, wherein the genetic variation results in elimination or abolishment of activity of a protein domain.
73. The method of claim 66, wherein the genetic variation alters or abolishes a regulatory sequence of a target gene.
74. The method of claim 66, wherein the genetic variation comprises insertion of a heterologous regulatory sequence.
75. The method of claim 66, wherein the genetic variation comprises insertion of a regulatory sequence found within a genome of a bacterial species or genus corresponding to the bacteria into which the genetic variation is introduced.

76. The method of claim 75, wherein the regulatory sequence is selected based on expression level of a gene in a bacterial culture or within plant tissue.
77. The method of claim 66, wherein the genetic variation is produced by chemical mutagenesis.
78. The method of claim 66, wherein step (c) further comprises exposing the plants to biotic or abiotic stressors.
79. The method of claim 67, wherein bacteria isolated after repeating steps (b) to (d) one or more times produce 1% or more of nitrogen in a second plant of the same type as the first plant.
80. The method of claim 67, wherein bacteria isolated after repeating steps (b) to (d) one or more times exhibit at least a 2-fold increase in nitrogen fixation as compared to bacteria isolated from the first plant.
81. The method of claim 80, wherein the second plant is grown in the presence of fertilizer supplemented with glutamine, ammonia, or other chemical source of nitrogen.
82. The method of claim 66, wherein the first plant is an agricultural crop plant.
83. The method of claim 82, wherein the agricultural crop plant is selected from barley, rice, maize, wheat, sorghum, sweet corn, sugar cane, onions, tomatoes, strawberries, or asparagus.
84. The method of claim 66, wherein the first or plants in the plurality of plants are a model plant.
85. The method of claim 84, wherein the model plant is selected from *Setaria*, *Brachypodium*, or *Arabidopsis*.
86. The method of claim 66, wherein the genetic variation is a pre-determined genetic variation that is specifically introduced to a target site.

87. The method of claim 66, wherein the genetic variation is a random mutation within the target site.
88. The method of claim 66, wherein step (a) further comprises performing genetic analysis of isolated bacteria.
89. The method of claim 66, wherein step (b) further comprises applying a selection pressure to enrich for bacteria comprising the genetic variation.
90. The method of claim 89, wherein the selection pressure comprises cleaving genomes lacking the genetic variation introduced to a target site, wherein cleavage occurs within 100 nucleotides of the target site.
91. The method of claim 89, further comprising isolating bacteria that survive the selection pressure.
92. The method of claim 91, wherein cleavage is directed by a site-specific nuclease selected from the group consisting of a counterselectable marker, Zinc Finger nuclease, a CRISPR nuclease, a TALE nuclease, or a meganuclease.
93. The method of claim 92, wherein the site-specific nuclease is a CRISPR nuclease.
94. The method of claim 66, wherein the genetic variation is an insertion or deletion of one or more nucleotides.
95. The method of claim 66, wherein bacteria isolated after repeating steps (b) to (d) one or more times are endophytic, epiphytic, or rhizospheric.
96. The method of claim 66, wherein bacteria isolated after repeating steps (b) to (d) one or more times comprise a plurality of different bacterial taxa.
97. The method of claim 66, wherein the bacteria are isolated from plant tissue.

98. The method of claim 66, wherein isolating bacteria in step (a) comprises isolating bacteria from a seed of the first plant.

99. A method of modifying a genome of a cell of a bacterial species, the method comprising:

(a) providing a polynucleotide comprising a subsequence consisting of sequences from the bacterial species, wherein the subsequence comprises in the 5' to 3' direction a first homologous sequence, a promoter, and a second homologous sequence; and

(b) inducing homologous recombination between a target locus of the genome and the polynucleotide to produce a recombined sequence, wherein (i) the target locus comprises a target sequence flanked by the first and second homologous sequences, and (ii) enhances expression of one or more second genes adjacent to the target locus.

100. The method of claim 99, further comprising:

(b)(iii) homologous recombination disrupts expression of a first gene comprising the target sequence.

101. The method of claim 99, further comprising:

(c) isolating bacterial cells comprising the recombined sequence.

102. The method of claim 99, wherein the promoter is selected based on expression activity above a threshold level under an environmental condition.

103. The method of claim 102, wherein the environmental condition is selected from the group consisting of: nutrient availability, nutrient deficiency, nitrogen stress, increased nitrogen exposure, heat, cold, osmotic stress, drought, flooding, salinity, presence or absence of interspecies signaling compounds, presence or absence of pathogens, and presence or absence of pesticide, herbicide, insecticide, nematocide, fungicide, or bacteriocide, wherein increased or decreased exposure is relative to a reference condition.

104. The method of claim 103, wherein selecting the promoter comprises: (a) exposing cells of the bacterial species to the environmental condition, (b) measuring expression levels of transcripts of the cells, and (c) identifying the promoter driving expression of a transcript with a level above a threshold level.

105. The method of claim 99, wherein the subsequence does not encode a protein of the bacterial species.
106. The method of claim 99, wherein the subsequence does not encode a selectable marker.
107. The method of claim 101, wherein isolating bacterial cells comprises applying a negative selection against cells lacking the recombined sequence.
108. The method of claim 107, wherein negative selection comprises cleaving genomes lacking the recombined sequence, wherein cleavage occurs within 100 nucleotides of the target sequence.
109. The method of claim 108, wherein the cleavage is directed by a site-specific nuclease selected from the group consisting of a counterselectable marker, Zinc Finger nuclease, a CRISPR nuclease, a TALE nuclease, or a meganuclease.
110. The method of claim 109, wherein the site-specific nuclease is a CRISPR nuclease.
111. The method of claim 99, wherein the first gene is a negative regulator of the one or more second genes.
112. The method of claim 99, wherein the first gene, the one or more second genes, or both are members of the nitrogen fixation pathway.
113. The method of claim 99, wherein the first gene is selected from the group consisting of: NifL, NtrB, glutamine synthetase, GlnB, GlnK, DraT, and AmtB.
114. The method of claim 99, wherein the one or more second genes are selected from the group consisting of: NifA and glutaminase.
115. The method of claim 99, wherein the isolated bacterial cells are endophytic, epiphytic, or rhizospheric.

116. The method of claim 115, wherein the isolated bacterial cells produce 5% or more of nitrogen in cells of a host plant.

117. The method of claim 116, wherein the isolated bacteria produce the nitrogen in the presence of fertilizer supplemented with glutamine, ammonia, or other chemical source of supplemental nitrogen.

118. A polynucleotide for disrupting a target gene in a cell of a bacterial species, the polynucleotide comprising a subsequence consisting of sequences derived from the bacterial species, wherein:

(a) the subsequence comprises in the 5' to 3' direction a first homologous sequence, a promoter, and a second homologous sequence;

(b) the first and second homologous sequences correspond to sequences flanking a target sequence that is deleted upon homologous recombination between the subsequence and a target locus comprising the target sequence;

(c) the target sequence comprises at least a first protein-encoding portion of a first gene; and

(d) the subsequence does not encode a protein of the bacterial species.

119. The polynucleotide of claim 118, wherein the subsequence does not encode a selectable marker.

120. The polynucleotide of claim 118, wherein the promoter is a promoter of a gene having an expression activity above a threshold level under an environmental condition.

121. The polynucleotide of claim 120, wherein the environmental condition is increased nitrogen exposure, and wherein increased exposure is relative to a reference condition.

122. The polynucleotide of claim 118, wherein the target sequence is adjacent to one or more second genes.

123. The polynucleotide of claim 122, wherein the first gene, the one or more second genes, or both are members of the nitrogen fixation pathway.

124. The polynucleotide of claim 118, wherein the first gene is selected from the group consisting of: NifL, NtrB, glutamine synthetase, GlnB, GlnK, DraT, and AmtB.
125. The polynucleotide of claim 118, wherein the cell is a bacterium that is endophytic, epiphytic, or rhizospheric.
126. An expression vector comprising the polynucleotide of any of claims 118-125.
127. A cell comprising the expression of claim 126.
128. The method of claim 1, wherein the plurality of bacteria are of the genus *Enterobacter*.
129. The method of claim 1, wherein the plurality of bacteria are of the genus *Rahnella*.
130. The method of claim 1, wherein the plurality of bacteria are of the genus *Kosakonia*.
131. The method of claim 1, wherein the plurality of bacteria are of the genus *Burkholderia*.
132. The method of claim 1, wherein the plurality of bacteria are of the genus *Klebsiella*.
133. The bacterial population of claim 27, wherein the plurality of bacteria are of the genus *Enterobacter*.
134. The bacterial population of claim 27, wherein the plurality of bacteria are of the genus *Rahnella*.
135. The bacterial population of claim 27, wherein the plurality of bacteria are of the genus *Kosakonia*.
136. The bacterial population of claim 27, wherein the plurality of bacteria are of the genus *Burkholderia*.

137. The bacterial population of claim 27, wherein the plurality of bacteria are of the genus *Klebsiella*.
138. The bacterium of claim 55, wherein the bacterium is from genus *Enterobacter*.
139. The bacterium of claim 55, wherein the bacterium is from genus *Rahnella*.
140. The bacterium of claim 55, wherein the bacterium is from genus *Kosakonia*.
141. The bacterium of claim 55, wherein the bacterium is from genus *Burkholderia*.
142. The bacterium of claim 55, wherein the bacterium is from genus *Klebsiella*.
143. The method of claim 99, further comprising:
- (b)(iii) homologous recombination disrupts expression of a first gene comprising the target sequence; and
 - (c) isolating bacterial cells comprising the recombined sequence.

1/17

FIGURE 1A

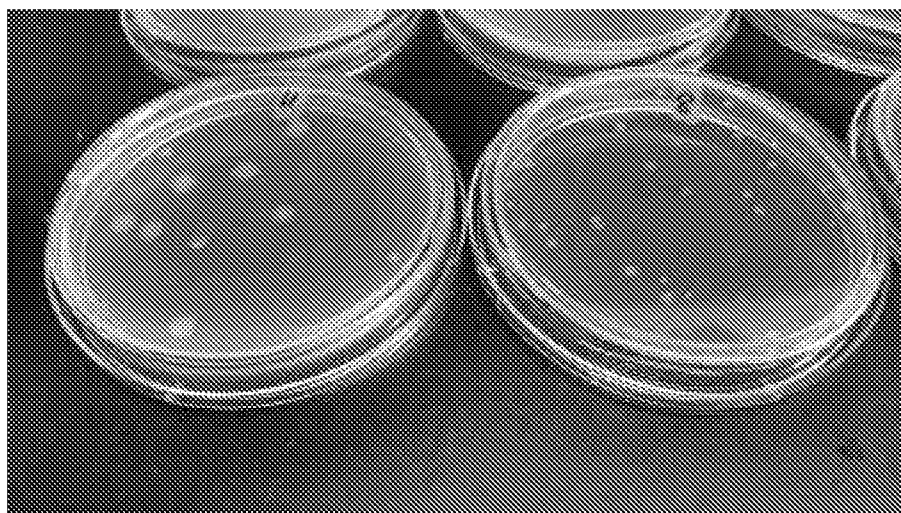


FIGURE 1B

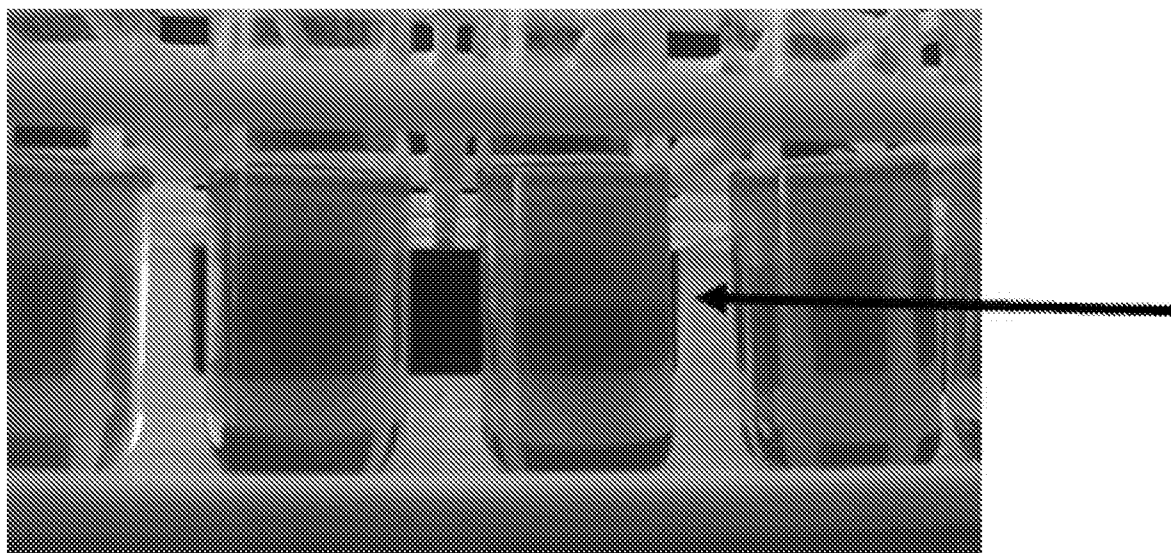


FIGURE 2

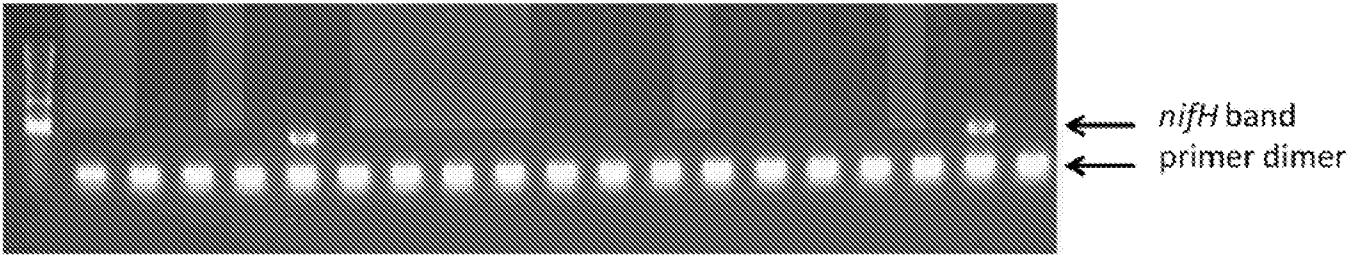
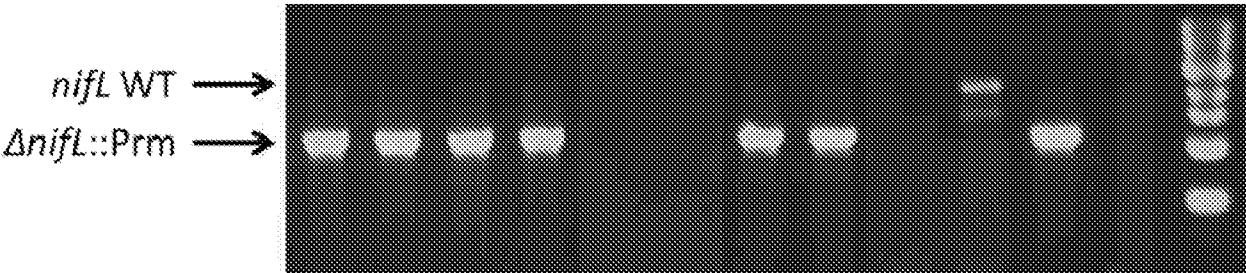


FIGURE 3



3/17

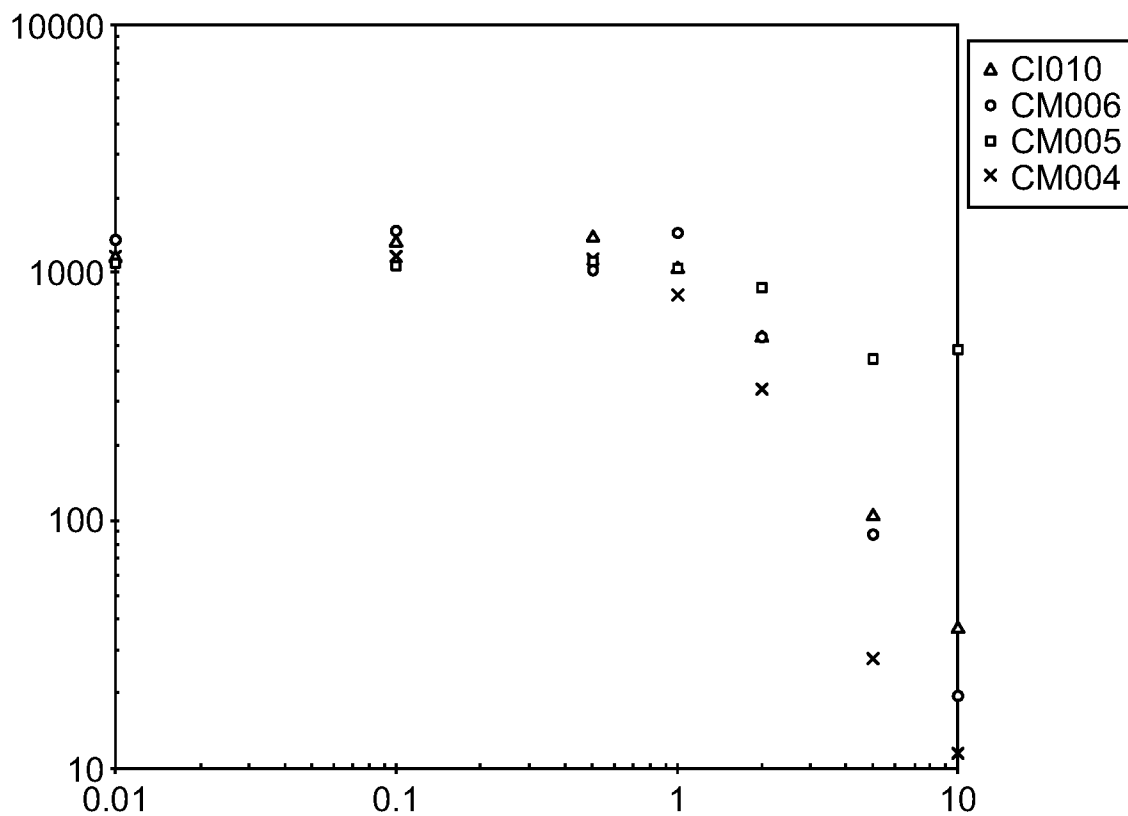


FIG. 4A

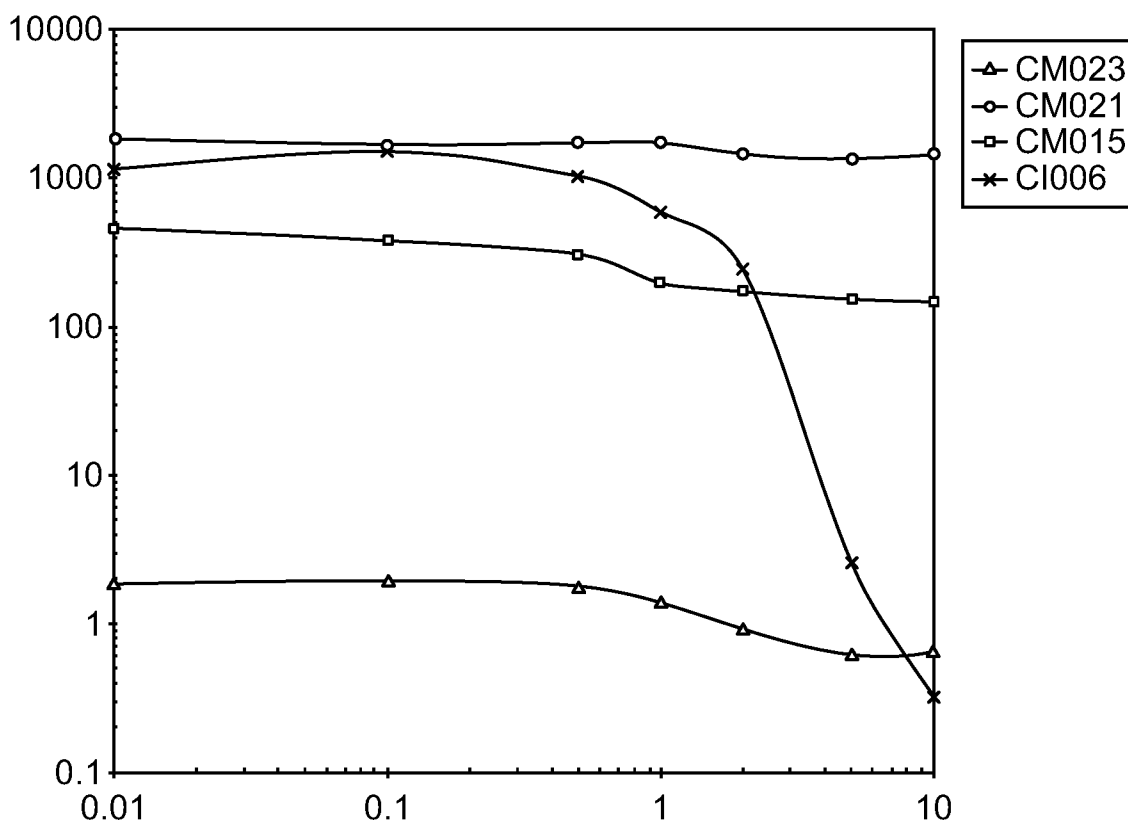


FIG. 4B

4/17

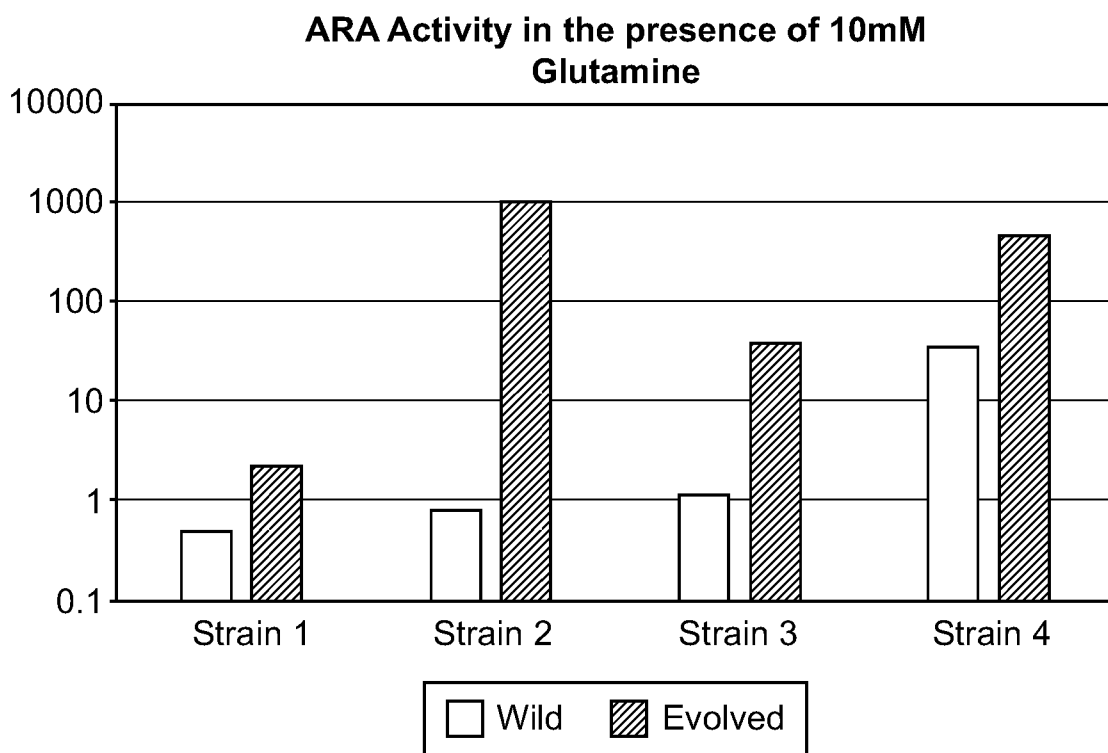


FIG. 4C

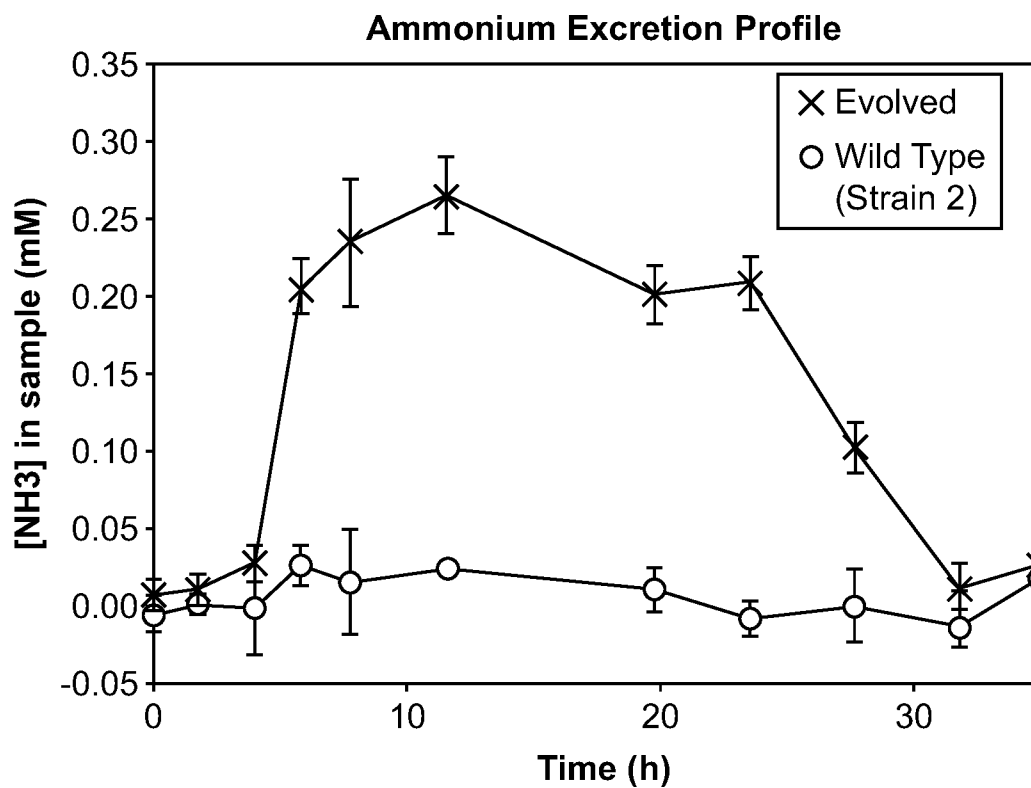


FIG. 4D

5/17

	no glutamine	1 mM glutamine	10 mM glutamine	
amtB	716462	175150	1045	0% air
galK	15	405	814	
glnB	8025	10275	7493	
glnK	752360	183994	320	
nifA	306663	92963	194	
nifH	12387186	3599183	161	
nifL	226368	42825	123	
ntrB	50439	25236	1081	
ntrC	78056	35760	1216	
amtB	241247	139599	1207	10% air
galK	404	770	1012	
glnB	8296	6899	9376	
glnK	241645	158973	288	
nifA	237483	115545	197	
nifH	4702957	2448758	108	
nifL	173765	66818	75	
ntrB	25676	19630	1118	
ntrC	40312	30703	1295	
amtB	160293	167736	1353	20% air
galK	1311	976	1200	
glnB	8522	8185	9445	
glnK	166653	191992	366	
nifA	200774	164973	198	
nifH	862984	2337297	80	
nifL	129054	99096	80	
ntrB	17326	21370	1146	
ntrC	24115	31446	1370	

FIG. 5

6/17

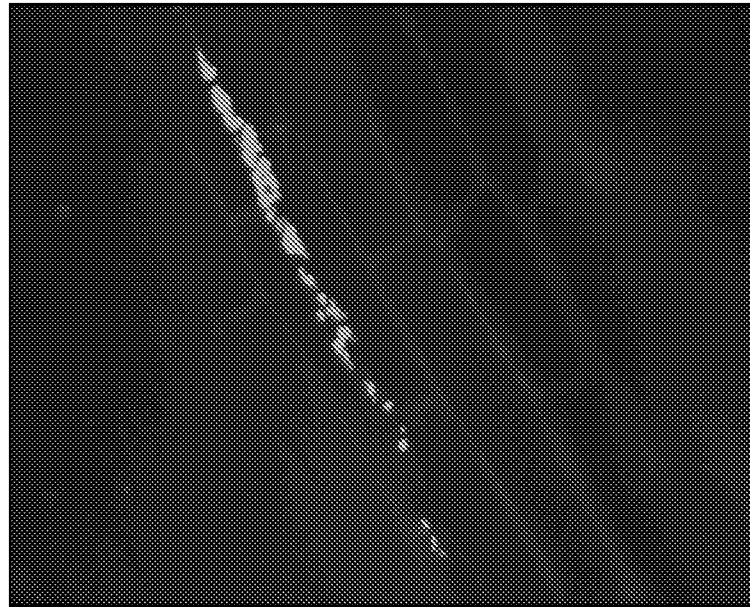


FIG. 6

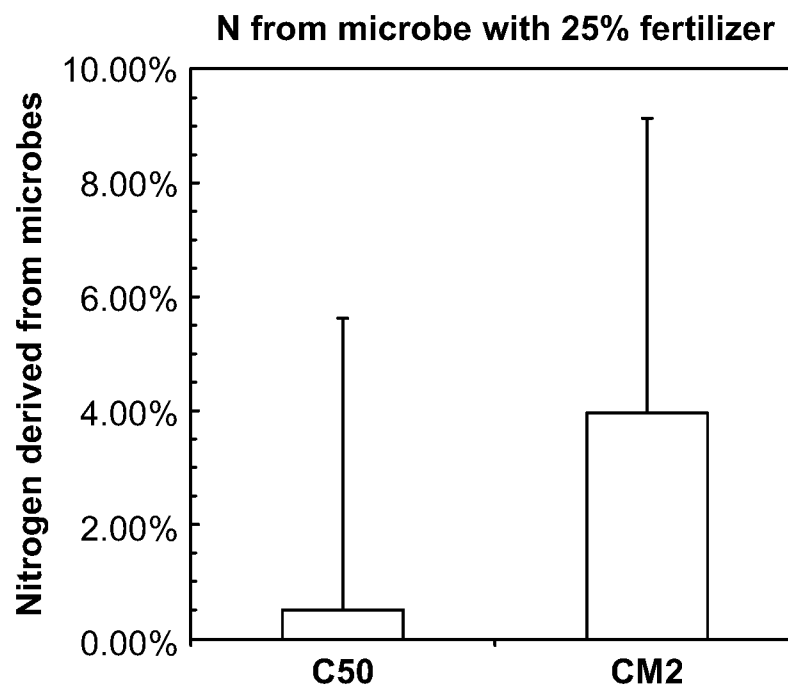


FIG. 7

7/17

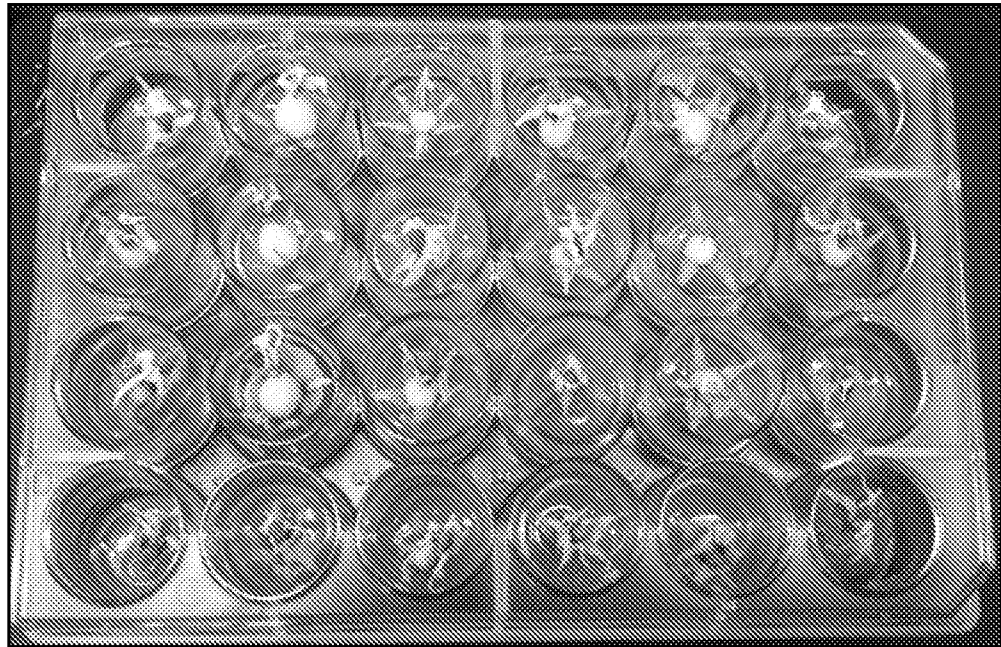


FIG. 8

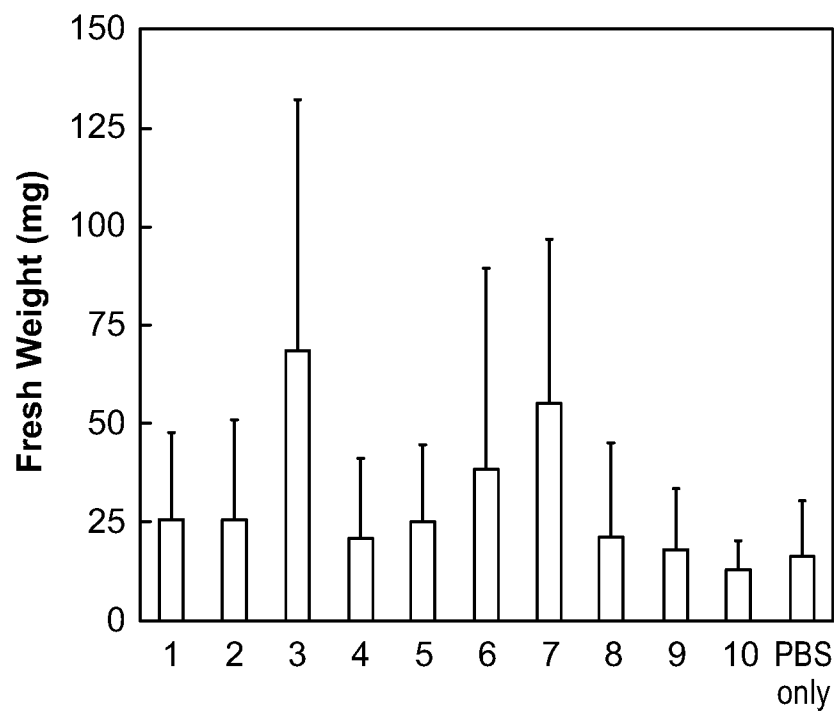


FIG. 9

8/17

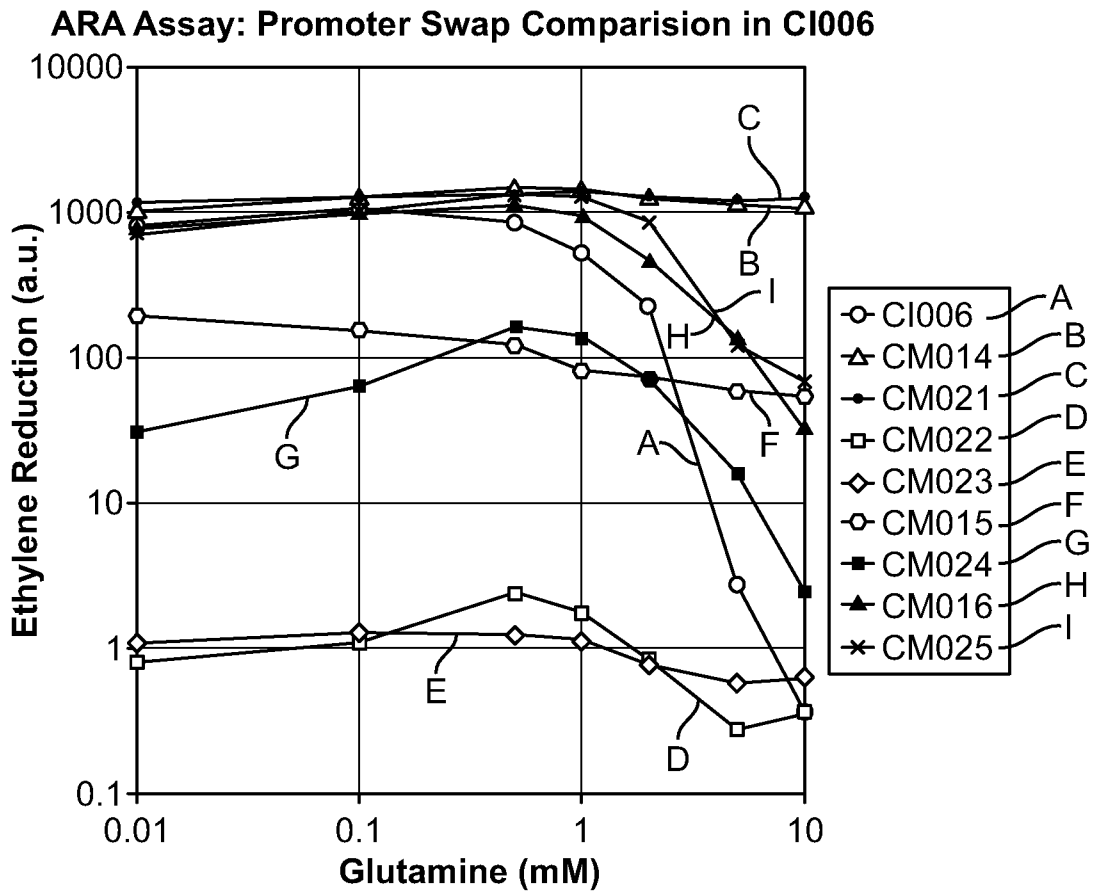


FIG. 10A

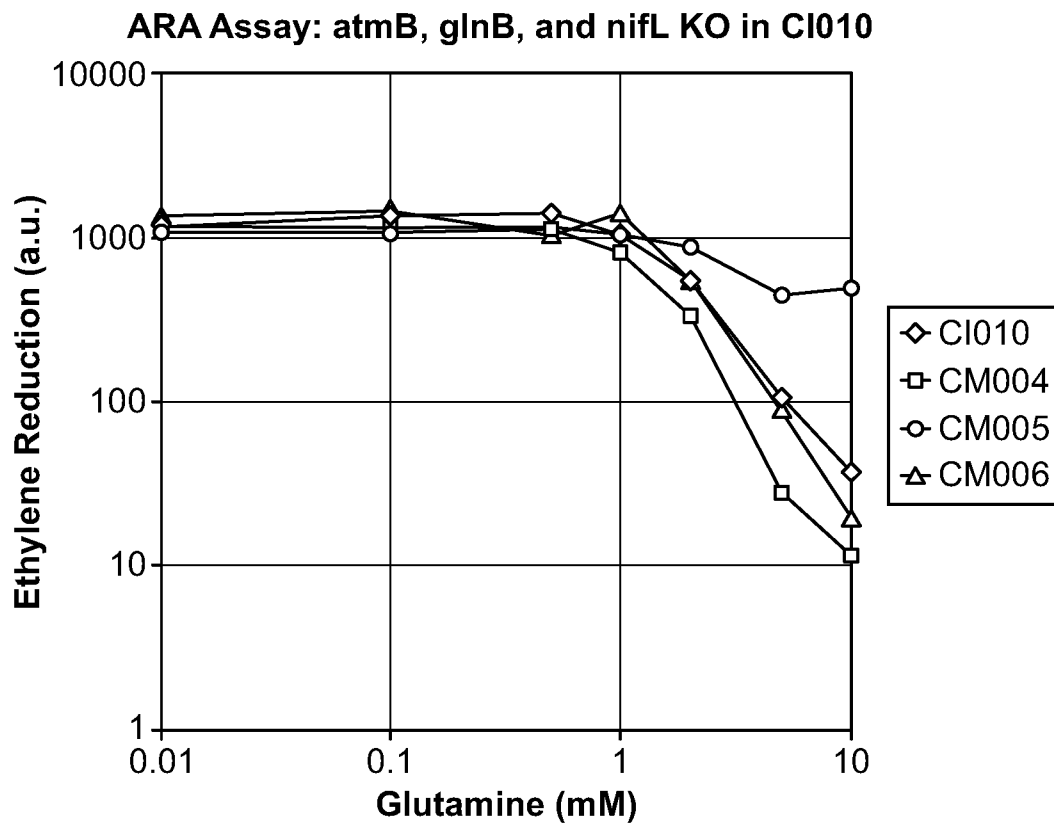


FIG. 10B

9/17

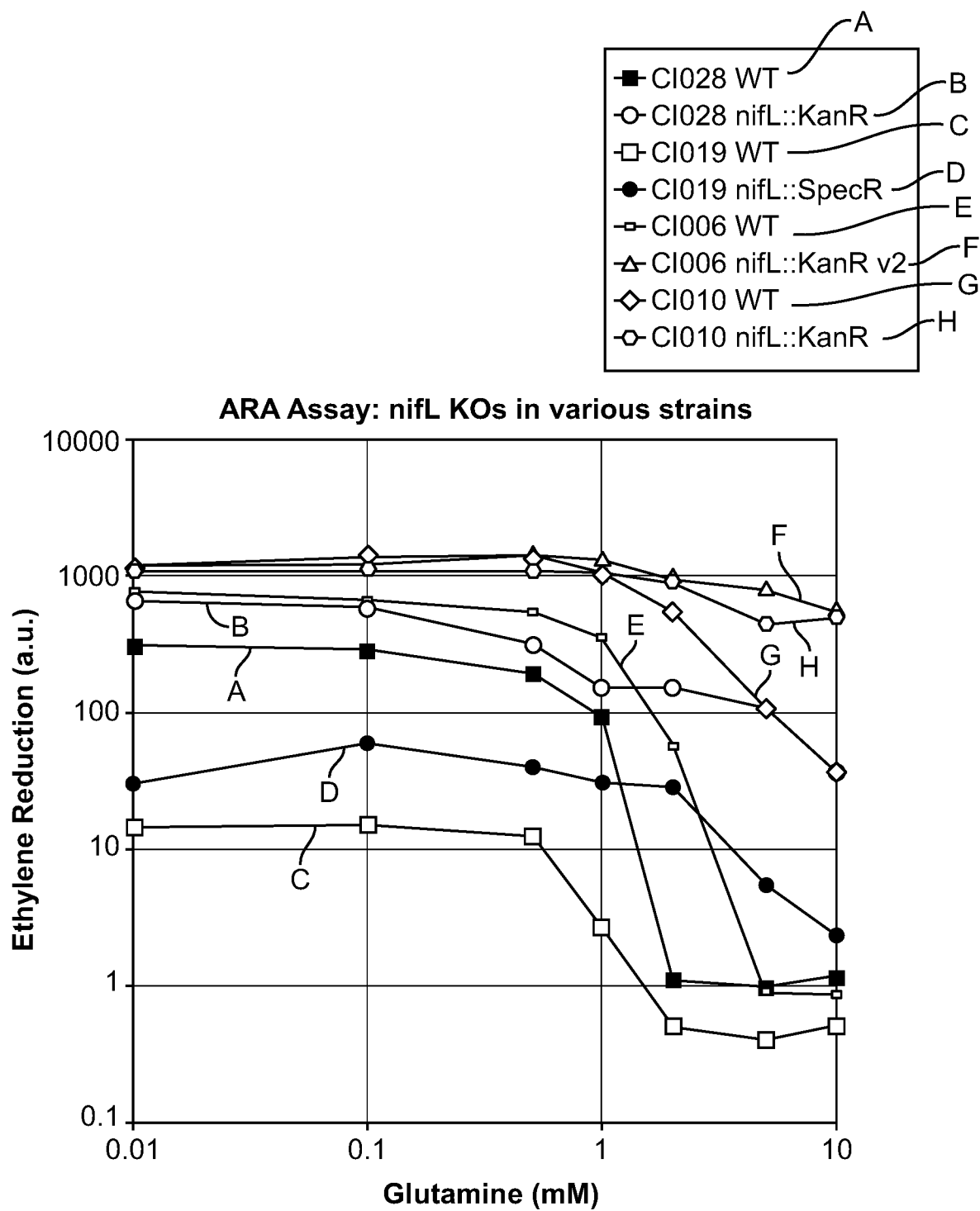


FIG. 10C

10/17

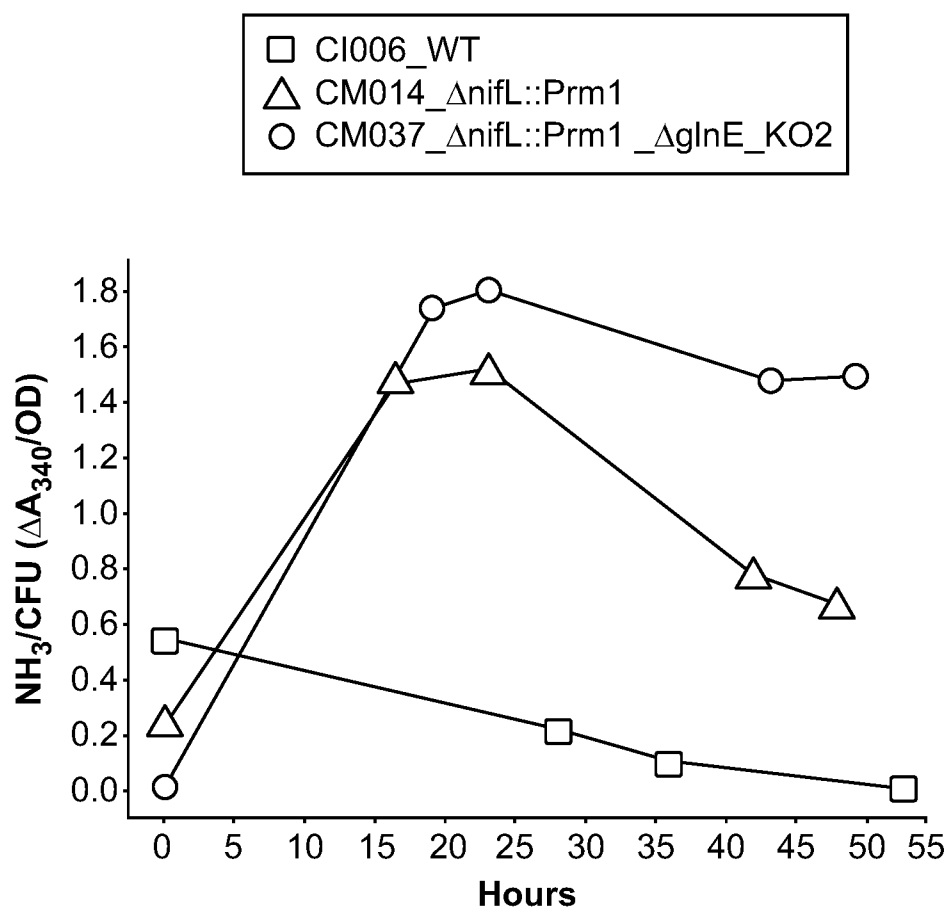
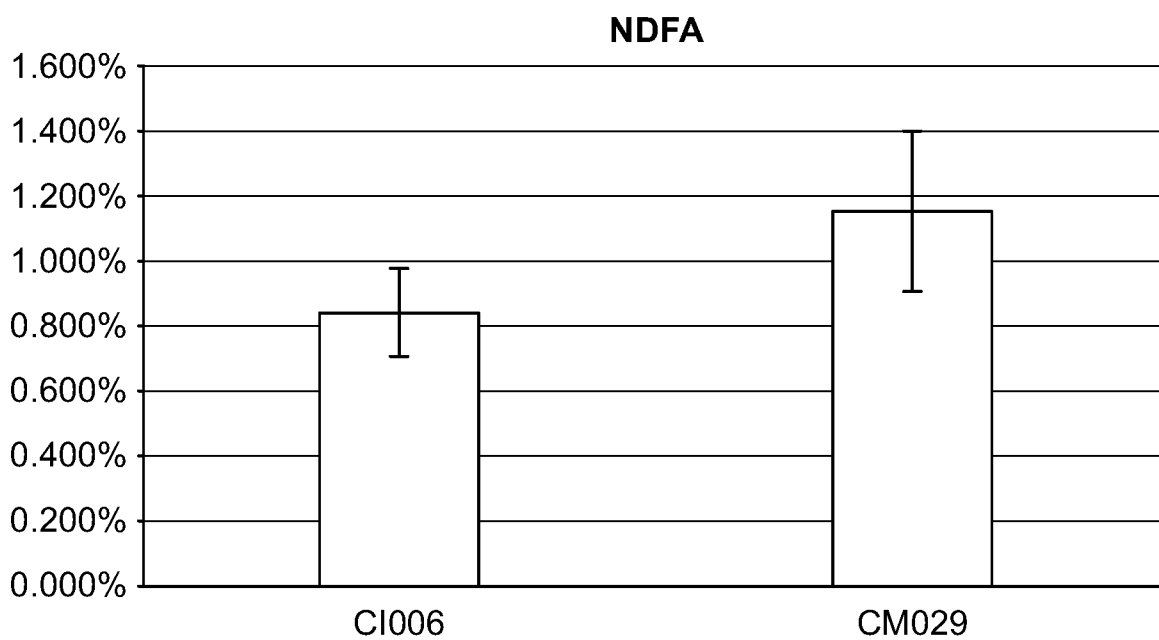


FIG. 11

11/17

**FIG. 12**

12/17

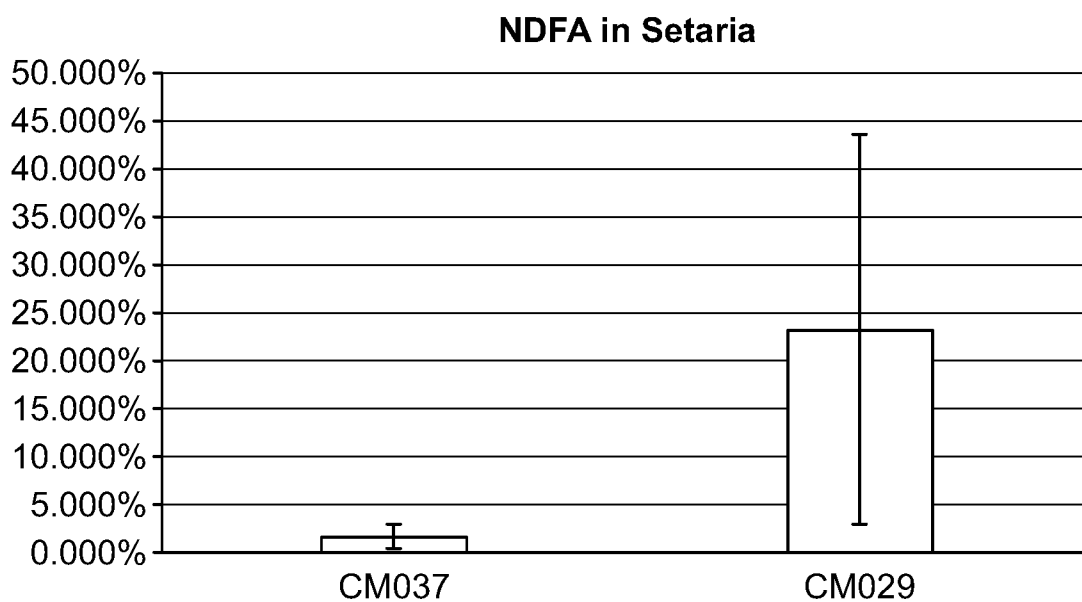


FIG. 13

13/17

FIGURE 14A

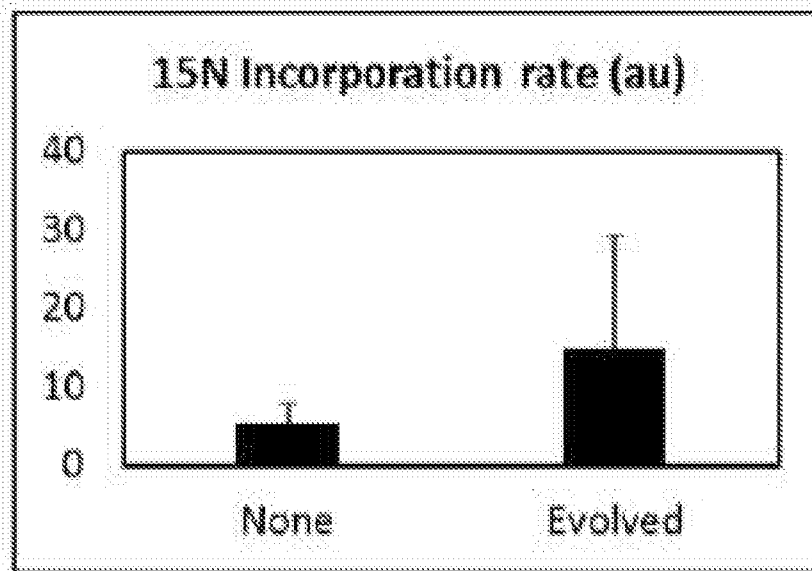
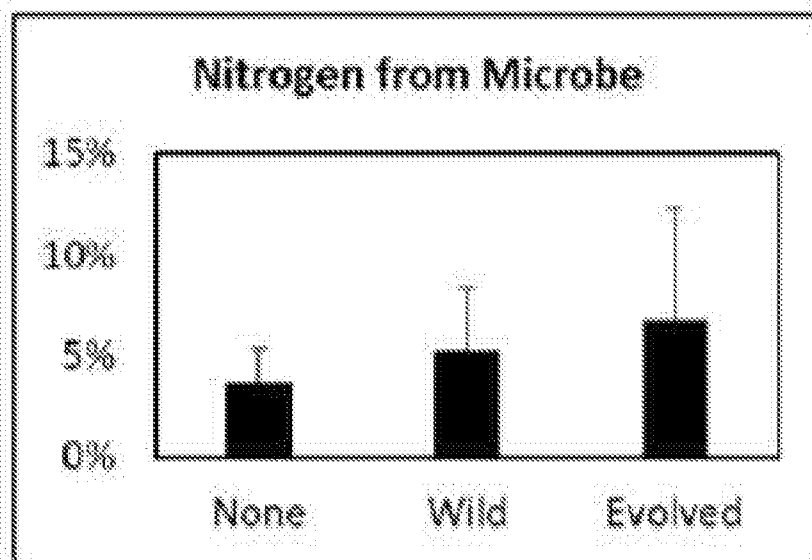


FIGURE 14B



14/17

FIGURE 14C

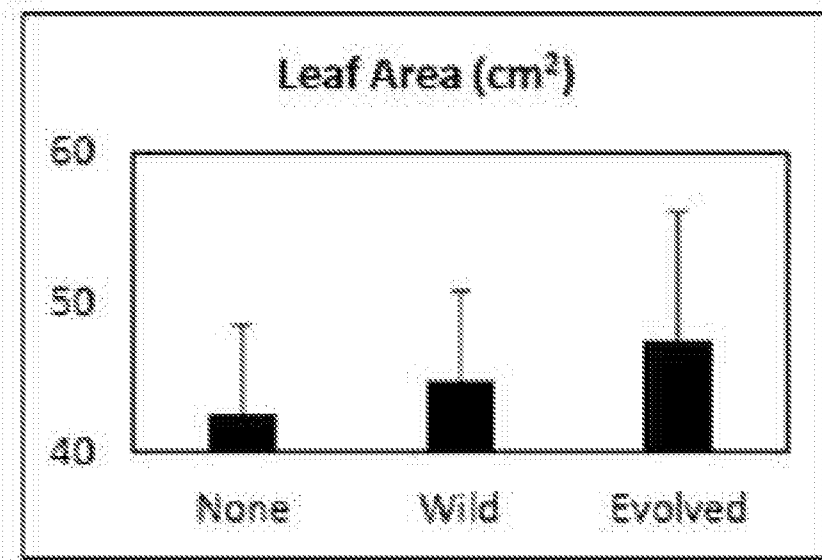
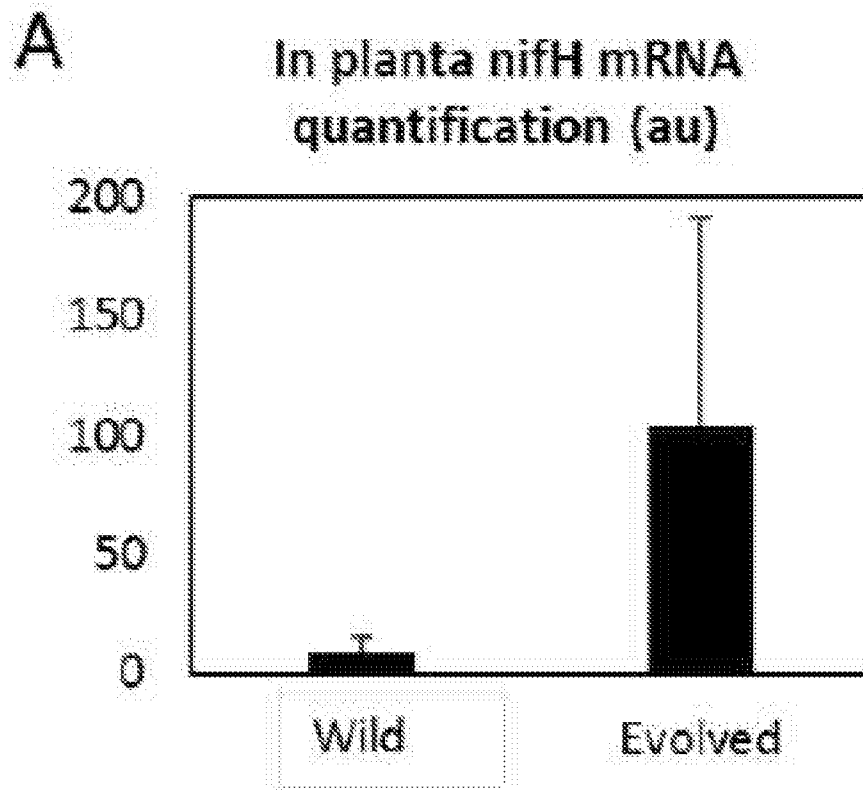
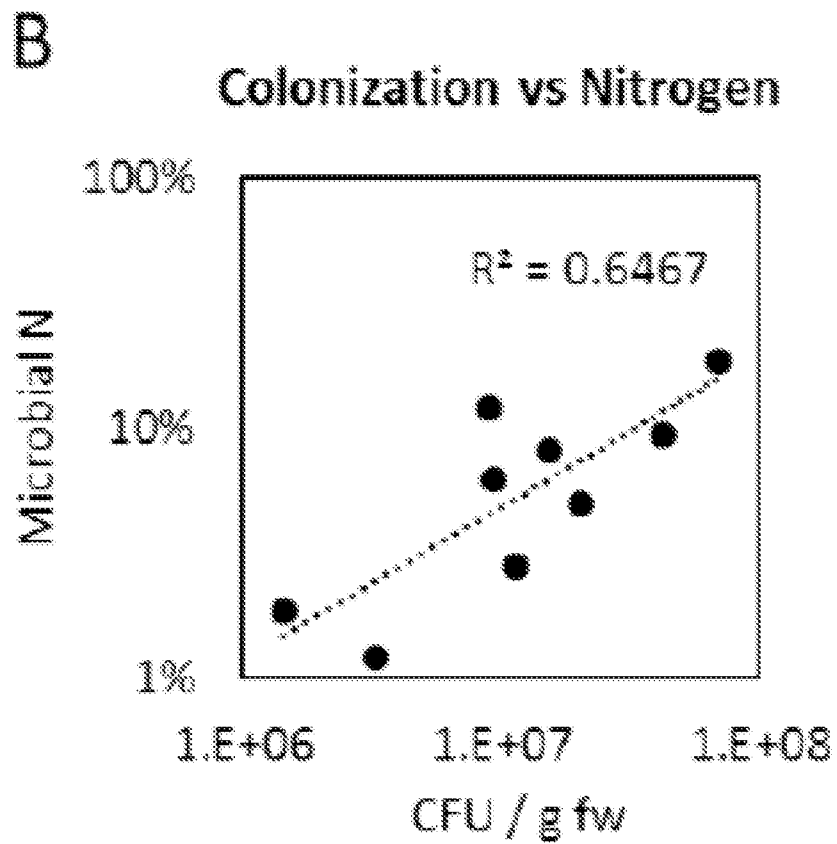


FIGURE 15A



16/17

FIGURE 15B



17/17

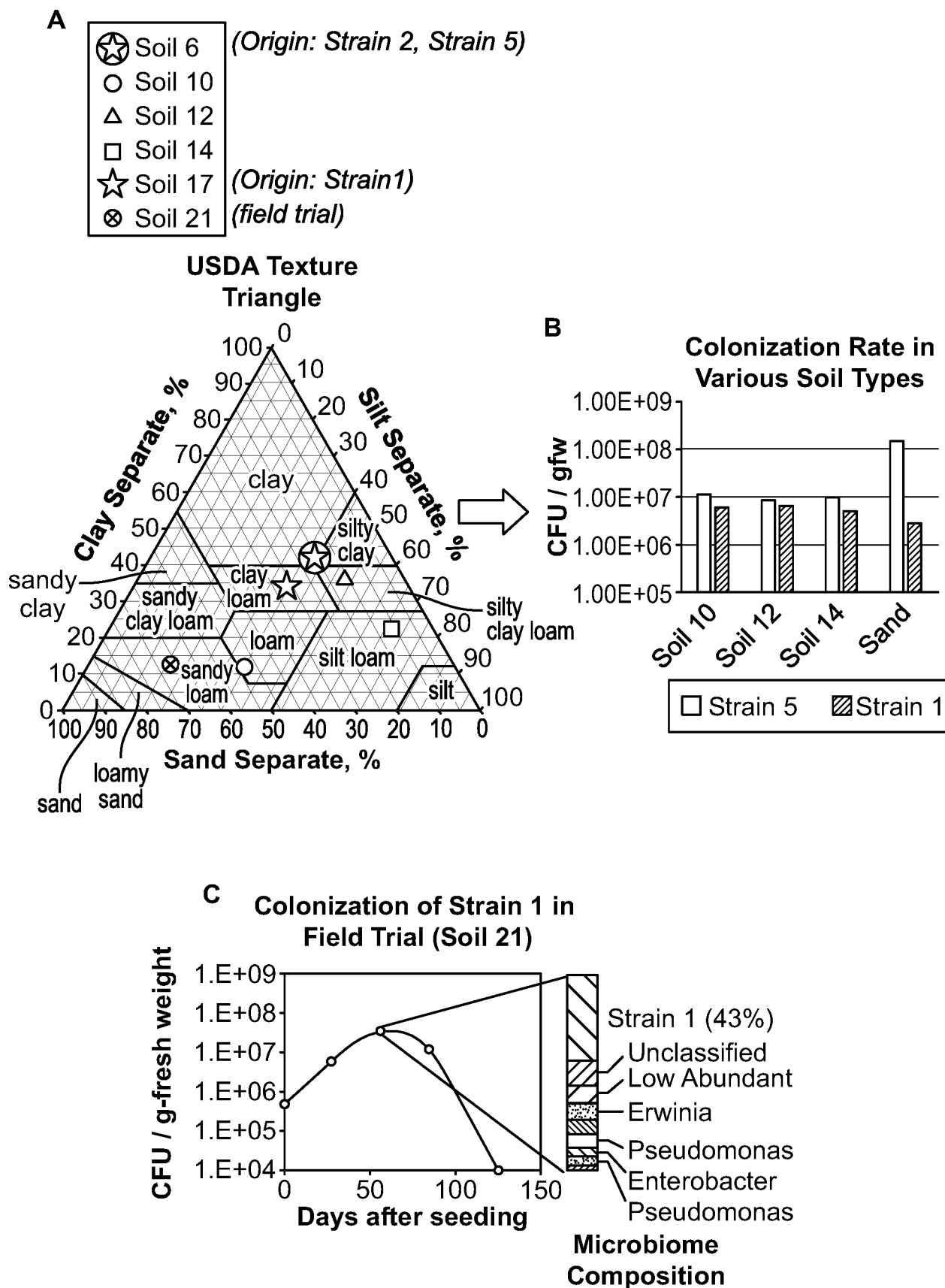


FIG. 16

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2016/042170

A. CLASSIFICATION OF SUBJECT MATTER

IPC(8) - C05F 11/08; C07K 14/195; C12N 1/20 (2016.01)

CPC - C05F 11/08; C07K 14/195; C12N 15/902; C12N 2800/90 (2016.11)

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC - C05F 11/08; C07K 14/195; C12N 1/20

CPC - C05F 11/08; C07K 14/195; C12N 15/902; C12N 2800/90

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

USPC - 435/34; 435/471; 435/252.3; 435/320.1 (keyword delimited)

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Patbase, Google Patents, PubMed, Google

Search terms used: plant bacteria nitrogen fixation ammonia fertilizer transgenic vector insert mutation NifA NifL

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2015/0128670 A1 (DEPARTMENT OF BIOTECHNOLOGY MINISTRY OF SCIENCE & TECHNOLOGY et al) 14 May 2015 (14.05.2015) entire document	55-70, 72-76, 78, 80-83, 86, 88-91, 94-108, 111-115, 118-127, 143
Y		1-53, 71, 77, 79, 84, 85, 87, 92, 93, 109, 110, 116, 117, 128-142
Y	US 2009/0137390 A1 (TRIPLETT) 28 May 2009 (28.05.2009) entire document	1-53, 79, 116, 117, 128-137, 142
Y	US 4,832,728 A (ALLAN et al) 23 May 1989 (23.05.1989) entire document	9, 10
Y	US 2012/0284813 A1 (OLIVIER JEAN PAUL et al) 08 November 2012 (08.11.2012) entire document	71, 77, 87
Y	US 2004/0235663 A1 (COCKING) 25 November 2004 (25.11.2004) entire document	84, 85
Y	US 2014/0273226 A1 (SYSTEM BIOSCIENCES, LLC) 18 September 2014 (18.09.2014) entire document	92, 93, 109, 110
Y	US 2015/0080261 A1 (BIODISCOVERY NEW ZEALAND LTD) 19 March 2015 (19.03.2015) entire document	129, 131, 134, 136, 138, 139, 141
Y	CHEN et al. "Complete genome sequence of Kosakonia sacchari type strain SP1T," Standards in Genomic Sciences, 15 June 2014 (15.06.2014), Vol. 9, No. 3, Pgs. 1311-1318. entire document	130, 135, 140



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

17 November 2016

Date of mailing of the international search report

02 DEC 2016

Name and mailing address of the ISA/US

Mail Stop PCT, Attn: ISA/US, Commissioner for Patents
P.O. Box 1450, Alexandria, VA 22313-1450

Facsimile No. 571-273-8300

Authorized officer

Blaine R. Copenheaver

PCT Helpdesk: 571-272-4300

PCT OSP: 571-272-7774

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2016/042170

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see Extra Sheet(s).

1. ☒ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☒ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2016/042170

Continued from Box No. III Observations where unity of invention is lacking

This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees need to be paid.

Group I: claims 1-26 and 128-132 are drawn to a non-leguminous plant, and methods comprising the same.

Group II claims 27-127 and 133-143 are drawn to a non-intergeneric bacterium; and a bacterial population, a composition comprising the same, and a method of producing the same bacterium.

The inventions listed in Groups I and II do not relate to a single general inventive concept under the PCT Rule 13.1, because under PCT Rule 13.2 they lack the same or corresponding special technical features for the following reasons:

The special technical features of Group I, methods of increasing nitrogen fixation in a non-leguminous plant, are not present in Groups II; the special technical features of Group II, isolated bacterium deposited as ATCC Accession Deposit No. PTA-122293 or PTA-122294, are not present in Group I.

Additionally, Groups I and II share the technical features of bacteria that are capable of fixing atmospheric nitrogen, isolated bacteria, genetic variations, exposing the plant to a plurality of bacteria, and the bacteria is isolated has an improved trait relative to other plants in the plurality of plants, and recombination between a target locus of the genome and the polynucleotide to produce a recombined sequence. However, these shared technical features do not represent a contribution over the prior art.

Specifically US 2015/0128670 A1 to Department of Biotechnology Ministry of Science & Technology et al. discloses bacteria (The recombinant *Azotobacter* microorganism is capable of fixing atmospheric nitrogen continuously, Abstract), isolated bacteria (*Azotobacter chroococcum* CBD15, a strain isolated from the soil at Indian Agricultural Research Institute, Para. [0049]), genetic variations (a process for producing a recombinant microorganism, wherein the process comprises (a) inserting a disrupted *nifL* gene and a *nifA* gene with the native promoter, Para. [0009]), exposing the plant to a plurality of bacteria (contacting seeds of crops with the transformed bacteria, Para. [0006]; Wheat seeds are soaked in a suspension of the recombinant or the non-recombinant microorganism for three hours, Para. [0105]), and the bacteria is isolated has an improved trait relative to other plants in the plurality of plants (HKD15 having Accession number MTCC 5679 is improved over that of the non-recombinant, native *Azotobacter chroococcum* CBD15, Para. [0054]; Wheat seeds are soaked in a suspension of the recombinant or the non-recombinant microorganism for three hours, Para. [0105]), and recombination between a target locus of the genome and the polynucleotide to produce a recombined sequence (inserting a disrupted *nifL* gene and a *nifA* gene with the native promoter into the chromosome of a microorganism through homologous recombination with a recombinant DNA molecule comprising a disrupted chromosomal *nifL* gene, Para. [0009]).

The inventions listed in Groups I and II therefore lack unity under Rule 13 because they do not share a same or corresponding special technical features.