Appendix 1: Making selective media for nitrogen fixing organisms

⚠ **Safety precaution:** Environmental microorganisms are cultured in this research project. Laboratory safety precautions for BSL1 organisms should be followed. Among others, students should wear gloves, long pants and closed toe shoes throughout the entire program. Cultures should be kept sealed, and transferred under a flow hood. A sink for students to wash their hands prior to leaving the lab should be available.

⚠ **Safety precaution:** ensure that students use the correct chemicals to make these solutions: you can either pull them out from your chemical cabinet yourself, or have the students show you the containers prior to making the solutions → no nitrogen should go into the media.

Stock Solutions

Solution 1: $FeSO_4$-$7H_2O$ 500 mg; dissolve in 500 mLs of distilled water. Filter sterilize.

Solution 2: $NaMoO_4$-$2H_2O$ 125 mg; dissolve in 500 mLs of distilled water. Filter sterilize.

Solution A: KH_2PO_4 0.4 grams

K_2HPO_4 1.6 grams. Dissolve in 900 mLs, pour into a bottle and autoclave.

Solution B: Sucrose 100.0 grams

$MgSO_4$-$7H_2O$ 1.0 grams

$CaCl_2$-H_2O 0.46 grams

Dissolve into ca. 350 mLs of distilled water. Bring volume up to 500 mLs with a graduated cylinder, pour into a bottle and autoclave. After it cools add 25 mLs of solution 1 and 5.0 mLs of solution 2. This is 10X.

Nitrogen-free minimal media

Aseptically add 100 mLs of Solution B to 900 mLs of cooled Solution A. (We point out to students that phosphate will form an insoluble precipitate at high temperatures with magnesium or calcium, therefore we autoclave solutions A and B separately). If making solid media, 15 grams of agar are added to Solution A prior to autoclaving.

Teaching tips:

- The following should be emphasized prior to students making their solutions:
    - Zero the balance with the weigh boat in it
    - Do not put unused chemicals back in the original container
    - Use a clean weigh boat for each chemical and rinse and dry the spatula in between chemicals
    - Clean up the counter and the balance at the end
    - Beakers and bottles are not reliable ways to measure volume of liquids
    - Suggested way of making a solution:

    - Add distilled water to a beaker equipped with a magnetic stirring bar on a stirring plate. The volume of water added should be slightly less than the final volume desired.
    - Add solutes to this beaker. Once the solutes are dissolved, pour the solution into a graduated cylinder and carefully add more distilled water until the bottom of the meniscus reaches the desired volume.
    - Decant into a bottle that is labelled with the name of the solution, the name or initials of the researcher making it, and the date.

- Solutions 1 and 2 are especially required for nitrogenase. This is a good opportunity to remind students that nitrogenase is able to carry out the feat of breaking a triple covalent bond because it has the help of iron and molybdenum. Advanced students should be able to understand the concept of enzyme catalysis; in the case of nitrogenase, how the molybdenum and iron atoms transfer electrons obtained from ferredoxin, a cellular electron carrier, to reduce N_2 into two NH_3 molecules. Students will also realize that they only need to weigh small amounts of these compounds which also presents an opportunity to introduce the concept of micronutrients. We typically discuss this during the "beginning-of-lab" discussion the following day as a review of what they have done so far and why.
- Please note that this media composition is typically used to grow *Azotobacter* species. There are other media composition for other diazotrophs such as cyanobacteria or Rhizobia. In addition, the high amount of sucrose included encourages capsule synthesis: isolates should form shiny gooey colonies on agar plates.

Appendix 2: Autoclave, aseptic technique

Autoclave

Teaching tips:

- The following should be emphasized when students use the autoclave:
    - Principles of sterilization used with the autoclave: a steam of high heat (121ºC) and pressure (15 psi) that will kill nearly all microorganisms
    - ⚠ ALWAYS use autoclave gloves when taking things in and out of the autoclave and when handling solutions and glassware that have come out of the autoclave
    - ⚠ ALWAYS at the end of an autoclave cycle, carefully open the door a few inches and let the steam exit the chamber for 5-10 minutes before opening the door completely
    - ⚠ ALWAYS use the slow exhaust cycle for liquids: this releases the high pressure slowly so that the liquids do not boil over; and the fast exhaust cycle if only autoclaving glassware
    - ⚠ ALWAYS slightly loosen the lids of bottles containing liquids prior to putting them in the autoclave: this prevents potential explosions!
    - ⚠ ALWAYS put items to be autoclaved in a tray and not directly in the autoclave rack
    - Use autoclave tape if available

Aseptic technique

Teaching tips:

- The following should be emphasized when teaching aseptic technique:
  - ⚠ Do not leave the flame unattended or the gas valve on
  - ⚠ Be careful with the flame: do not lean over it or have loose hair around it
  - Flame the lip of bottles or flasks before and after a pipette or inoculating loop is inserted
  - Always use sterile pipettes or inoculating loops
  - Do not leave containers open or lids laying around: always close a container first before opening a new one
  - Gather all of the materials necessary and have a "game plan" prior to opening a sterile container

- We do an aseptic technique test that consists of students adding ca. 5 mLs of LB into two test tubes. One of the test tubes is inoculated with *E. coli* cells and the other is left uninoculated as a control. We incubate both test tubes overnight shaking at 37ºC –a chance to introduce the concept of optimal growth temperature. The next day they score their aseptic technique: the uninoculated test tube should still be transparent whereas the inoculated test tube should be turbid, and we introduce the concept of turbidity as a telltale sign of microbial growth.

Appendix 3: Sampling and handling the enrichments

Sampling

- ⚠ **Safety precaution:** students should wear gloves when handling the soil and dress appropriately for field sampling (i.e. close toe shoes, no sandals)

We make a "sampling packing list" the day before going sampling. This typically includes:

- 50 mL falcon tubes to collect soil samples
- Markers
- Lab gloves
- Sterilized spatulas (wrapped individually in foil)
- 95% ethanol in a 50 mL falcon tube (to sterilize spatulas in the field if necessary)
- Ziplock bag (to store 50 mL falcon tubes according to sites)
- Camera
- Notebook and pen

We are fortunate that our sampling Field Station has coring soil sampler tools available. Since we are sampling near the rhizosphere, we ask students to sample about 6-12 inches deep near a plant. Carolina sells cheap soil samplers (Carolina catalog item # 653250), but a small hole on the ground could be made with a shovel or a falcon tube and students could get soil from the bottom with their spatulas. If the ground is dry and students are using a coring tool, we

recommend bringing a soft headed mallet to push the corer into the ground. We take pictures of the students sampling; these pictures are fun and are used by the students in their posters and presentations later at the end of the project.

Handling the enrichments

- ⚠ **Safety precaution:** students should wear gloves and handle soil in hood as much as possible
- ⚠ **Safety precaution:** although nitrogen-fixing organisms are not pathogens, students should always handle enrichments with gloves and work in the hood when transferring cultures from flask to agar plate
- We store soil samples at 4ºC.
- For the enrichments, students add 1 gram of soil to 50 mLs of nitrogen-free media in 125 mL flasks.
- We incubate the flasks at room temperature (about 22-25ºC) with gentle shaking.
- Students are encouraged to take pictures of their enrichments on their first day and then again after a week or so.
- Students aseptically transfer 1 mL of their enrichments on selected days to cuvettes to measure optical density at 600 nM wavelength (students are instructed in the correct way to handle cuvettes, i.e. not touching the path of the light, and in the basics of spectrophotometry: i.e. the amount of light that reaches the detector is proportional to the amount of material in the solution). If cuvettes do not have lids, a small square of parafilm can be used to cover the top.

Appendix 4. Teaching tips for microscopy

There are excellent microscope guides in most introductory biology laboratory textbooks that should be consulted if the instructor is not familiar with microscope use. In general, the following concepts should be introduced when students use the microscope:

- Magnification (individual lenses and total): ocular lens has a 10X magnification; there are several objective lenses, each with a different magnification (typically 5X, 40X and 100X). Students should be told what the total magnification should be with the 5X objective (total=50X), and based on this they should be prompted to figure out total magnification with the other objective lenses.
- Preparing slides for viewing: wet mounts (such as they ones they prepare for this project) require coverslip and the total amount of sample should be such that the coverslip is not floating on top. Slides should be securely placed on the microscope stage and secured with the clip.
- ⚠ Focusing: Students should be aware of the coarse and focusing knobs for focusing and of the concept of working distance (distance between the slide and the objective lens). They should only use the coarse knob when working on the lowest magnification objective. They should start by focusing their slide on the lowest magnification objective using the coarse knob, with a touch of the fine tuning knob. A good tip is to focus on the marker writing, or on the edge of the coverslip, and then move onto the specimen. Once the sample is focused with the lowest objective, students should be instructed to use the nosepiece (most try to just push on the objective lens) to switch objectives, and to just use

the fine tuning knob. For our phase microscopes we remind students to use the correct phase for each objective.

- ⚠ 100X Oil-objective: Students should be aware that only the 100X objective can handle oil, and that the 40X objective in particular should never touch a slide with oil. Students should be instructed to use the nosepiece and pause half way between the 40X and 100X objectives, add a drop of microscope oil, and then continue moving the nosepiece until the 100X is in place.
- ⚠ Clean up: Students should be instructed to use lens paper to wipe oil from the 100X objective, turn the microscope off and cover it to avoid dust accumulation.

Appendix 5. Extracting soil DNA and Nanodrop measurement

Extracting soil DNA

Students are introduced to many routine molecular biology laboratory skills through this procedure. We use MoBio PowerSoil DNA extraction kit and follow manufacturer's instructions except for we use a bead beater/homogenizer instead of a vortex adapter for cell lysis, and we ask students to elute with 50 μL instead of 100 μL to obtain more concentrated DNA. Students are asked to read the kit's detailed instructions prior to using the kit the first time.

Students are instructed on how to use micropipettes and sterile tips (usually practice with a beaker with distilled water before they "graduate" to using the kit). They are given an overall description of DNA extraction: briefly, they are told that we start with DNA inside of cells which are also full of other materials such as carbohydrates and proteins, and that the end goal is to have a solution of just DNA. To achieve that goal we mechanically break cell walls, use "soap" to dissolve cell membranes and centrifugation to eliminate cellular components. Students are also introduced to the ubiquitous spin filter, component of nearly all DNA and RNA extraction kits. Additionally, this is also the first time students will use a table top centrifuge.

Nanodrop measurement

To verify the success of the soil DNA extraction we instruct students on the use of a Nanodrop spectrophotometer. This step is optional since the first PCR they carry out with their samples is also meant to be a diagnostic tool (see appendix 5). Concentrations range from 5 to 30 ng/μL and the DNA is not always clean (based on 260/230 and 260/280 ratios) but amplicons can be generated even with this low concentrations.

Appendix 6: PCR and gel electrophoresis

Polymerase chain Reaction, PCR

For all of our PCRs we use Promega's GoTaq mix, a 2X mastermix containing Taq polymerase, buffer and nucleotides. A typical 50 µL PCR reaction consisted of 5 µL of extracted DNA, 1µL each of 10 µM forward and reverse primers, 18 µl of sterile distilled water and 25 µl of GoTaq master mix.

For the PCR reactions amplifying the 16S rRNA gene we use the following universal bacterial primers:

16s27F: AGAGTTTGATCMTGGCTCAG

16S1492R: TACGGYTACCTTGTTACGACTT

And the following program:

- One cycle of 5 minutes at 94ºC for initial denaturation
- 30 cycles of 1 minute at 94ºC, 1minute at 55ºC and 2 minutes at 72ºC
- One cycle of 72ºC for 10 minutes for final elongation

The amplicon size is 1425 base pairs

For the PCR reactions amplifying the *nifH* gene we use the following universal *nifH* primers:

nifHF: TAYGGIAARGGIGGIATYGGIAARTC

nifHR: TCIGGIGARATGATGGC

And the following program:

- One cycle of 5 minutes at 94ºC for initial denaturation

- 30 cycles of 1 minute at 94ºC, 1minute at 50ºC and 1 minute at 72ºC
- One cycle of 72ºC for 10 minutes for final elongation

The amplicon size is 470 base pairs

For identifying microorganisms from isolated colonies generated by streak plating we use the same *16S rRNA* gene amplification protocol as above but instead of using extracted DNA we do the following:

- Pick a colony with a sterile toothpick and resuspend it in 10 µL of 1X Taq polymerase buffer in a PCR tube (NOTE: this is just 1X buffer –from Promega's Taq Polymerase, and not 1X GoTaq master mix); distilled water would also work.
- Lyse the cells in the Thermocycler with the following program:
    - 5 minutes at 96ºC
    - 1.5 minutes at 50ºC
    - 1.5 minutes at 96ºC
    - 1.5 minutes at 45ºC
- Use 5 µL of these lysed cells as template PCR and proceeded with *16S rRNA* gene amplification as above

Teaching Tips

We introduce the Polymerase Chain Reaction in four different concepts: (1) Heat; (2) DNA polymerase (3) primers and (4) annealing and elongation temperatures.

(1) Most students are already familiar with the concept of DNA copying through unzipping of the double helix; in PCR we are copying DNA inside of a tube and we unzip the helix

by applying heat that breaks the hydrogen bonds holding the two strands of DNA together

(2) Most students understand base pairing; the enzyme responsible for adding those nucleotides –"those letters"- is DNA polymerase, therefore we add DNA polymerase to our reaction mix

(3) DNA polymerase needs to know what gene to copy and amplify, hence the need for primers –"gene flags."

(4) DNA polymerase and primers cannot do their job at the high temperatures that we employ to separate the DNA strands, therefore we use a thermocycler instrument that changes temperature to match the activity needed to amplify DNA

Because here in Montana we live so close to the Hot Springs of Yellowstone National Park from where *Thermus aquaticus* was first isolated, we also explain to them how prior to the discovery of thermophiles, scientists had to add DNA polymerase after every cycle since it would denature when the temperature was increased to separate the strands.

Gel electrophoresis

⚠ **Safety precaution: If staining gels with ethidium bromide, students need close supervision and to be trained with the proper procedure on handling equipment and disposing of materials that are contaminated with ethidium bromide**

We use 50 mL 0.8% agarose gels made with 1X Tris-acetate-EDTA (TAE) buffer. Students use our lab 50X stock (In 1L: Tris base, 242 g, Acetate, 57.1 mL and EDTA, 100 mL of 0.5 sodium EDTA) and 1X stocks. Students are given detailed instructions on how to prepare the gels:

- Weigh out 0.4 g of agarose and put it in a 250 mL flask
- Add 50 mLs of 1X TAE
- Microwave for 1 minute
- With a hot glove take the flask out of the microwave and wait 10 minutes
- Carefully add 3 µl of 10 mg/mL ethidium bromide stock solution
- Pour into the casting tray containing the combs
- Wait for at least 15 minutes
- Carefully take gel out of casting tray and insert into electrophoresis chamber
- Mix 10 µL of DNA sample with 3 µL of 6X loading buffer (we tend to err on the side of extra loading buffer)
- Load 5 µl of DNA ladder (we use Fischer exACT 1 Kb gene ladder) into the first well
- Load 10 µL of your DNA-loading buffer mixture to subsequent lanes
- Turn on the power supply (ours is error-proof, in that the positive and negative electrodes are set, but students should be instructed on connecting the positive electrode to the bottom of the gel)
- After 30-40 minutes, stop the power supply, take the gel out and use the transilluminator to visualize the DNA

Teaching Tips

At the end of their PCR reactions we instruct students to evaluate if their PCR worked just by looking at their PCR reaction test tubes. Of course this is not possible to do and we introduce the need for gel electrophoresis as a tool to visualize the gene we have just amplified.

We use the analogy of two people, one carrying a meter long stick and the other carrying a ten meter long stick –horizontally-, trying to make it to the other side of a thick forest. Students quickly determine that carrying a shorter stick will allow you to maneuver the forest easier and move along faster, just like a shorter piece of DNA will also maneuver the pores in the agarose easier and move along faster.

Appendix 7. DNA sequencing, BLAST and 16*S rRNA* tree using the Ribosomal Database Project

We use Promega's Wizard SV Gel and PCR Clean-Up System to prepare PCR products for sequencing, and Functional Biosciences, Inc. to carry out the actual sequencing (results are typically available within 48 hours).

When students receive their sequences we introduce the concepts central to Sanger sequencing, that is: (1) the use of chain-terminating fluorescent dideoxynucleotides to generate amplicons of different sizes (with advanced students we take this opportunity to remind them that nucleotides in DNA already lack a hydroxyl group in their ribose, and that these dideoxynucleotides lack both hydroxyl groups); and (2) the use of capillaries to separate the different sized fragments. Students, at the very minimum, can correlate the capillaries with the gel electrophoresis experiments they have carried out and that the fluorescent bases correspond to the peaks in the electropherograms.

For BLAST analyses we simply introduce the concept of databases and of hundreds of thousands of scientists over time entering sequences into these databases. We also introduce the power of bioinformatics as it would be nearly impossible for us to search through these databases, but software like BLAST can quickly do it for us. We briefly let them know that there are specific scoring rules so that BLAST can assign a final score to each match ("hit") that it finds.

Students are instructed to

(1) Choose nucleotide BLAST since we are comparing nucleotides (https://blast.ncbi.nlm.nih.gov/Blast.cgi)
(2) Paste their sequences into the Enter Query Sequence Window

(3) Submit their sequence by clicking on the BLAST icon at the bottom of the page

Once the sequence window appears, students are instructed to

(1) Write down the number of hits
(2) Scroll down to the Descriptions of the sequences producing significant alignments
(3) Write down the description of the top 3 organisms and the % identity.
(4) Click on the Accession number for those organisms to access the GenBank file.
(5) Investigate the file to find out the environment from which this organism was isolated (all of them so far from the rhizosphere of some plant, not necessarily a crop)

16*S rRNA* tree using the Ribosomal Database Project

Instructions on how to build a phylogenetic tree using the Ribosomal Database Project are available on their website: https://rdp.cme.msu.edu/

NOTE: a text editor such as the free Notepad plus is useful and you will need to generate an RDP account (all of this is free).

First you generate a file with your "anchor" sequences:

- Log in to myRDP
- Go to browsers by clicking on the Browsers icon on the top menu
- Select organisms from representative clades (I typically use a few known diazotrophic Proteobacteria, Cyanobacteria, Firmicutes and include the sequence from a methanogen as root).

- Save those and open that file with Notepad → Reference sequence file. This reference file can be generated by students if there is time; we generate it ourselves and make it available to them

Upload your sequences (that you obtained from sequencing the *16S rRNA* gene from your isolates) by clicking upload in the menu on the right size. Sequences will take a few minutes to align. Once they are aligned, select them by clicking on the + icon.

- Go to Seqcart and down load those → open with Notepad plus

Now you will generate a file that combines both your reference sequences with your isolate sequences:

- ❖ Copy and paste your sequence IDs (line might look like this U010402912 Name of sequence) at the end of the reference sequence file
- ❖ Save this new file as a text file (i.e. Reference_plus_isolates)
- ❖ In RDP go to seqcart and upload this file (Reference_plus_isolates)
- ❖ Click on Tree Builder and Click CREATE tree

Because the RDP server is sometimes slow, we have also created trees by aligning the following "anchor" sequences with ClustalX, using their Build tree option to generate a simple phylogenetic tree, which we visualize with FigTree (both of these software are available to download for free)

>Bradyrhizobium_japonicum
CTGGCTCAGAGCGAACGCTGGCGGCAGGCTTAACACATGCAAGTCGAGCG
GGCGTAGCAATACGTCAGCGGCAGACGGGTGAGTAACGCGTGGGAACGTA
CCTTTTGGTTCGGAACAACACAGGGAAACTTGTGCTAATACCGGATAAGC

CCTTACGGGGAAAGATTTATCGCCGAAAGATCGGCCCGCGTCTGATTAGC
TAGTTGGTAGGGTAACGGCCTACCAAGGCGACGATCAGTAGCTGGTCTGA
GAGGATGATCAGCCACATTGGGACTGAGACACGGCCCAAACTCCTACGGG
AGGCAGCAGTGGGGAATATTGGACAATGGGGGCAACCCTGATCCAGCCAT
GCCGCGTGAGTGATGAAGGCCCTAGGGTTGTAAAGCTCTTTTGTGCGGGA
AGATAATGACGGTACCGCAAGAATAAGCCCCGGCTAACTTCGTGCCAGCA
GCCGCGGTAATACGAAGGGGGCTAGCGTTGCTCGGAATCACTGGGCGTAA
AGGGTGCGTAGGCGGGTCTTTAAGTCAGGGGTGAAATCCTGGAGCTCAAC
TCCAGAACTGCCTTTGATACTGAAGATCTTGAGTTCGGGAGAGGTGAGTG
GAACTGCGAGTGTAGAGGTGAAATTCGTAGATATTCGCAAGAACACCAGT
GGCGAAGGCGGCTCACTGGCCCGATACTGACGCTGAGGCACGAAAGCGTG
GGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAA
TGCCAGCCGTTAGTGGGTTTACTCACTAGTGGCGCAGCTAACGCTTTAAG
CATTCCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAAGGAATTGAC
GGGGGCCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGACGCAACGCGC
AGAACCTTACCAGCCCTTGACATGTCCAGGACCGGTCGCAGAGATGTGAC
CTTCTCTTCGGAGCCTGGAACACAGGTGCTGCATGGCTGTCGTCAGCTCG
TGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCCCGTCCTT
AGTTGCTACCATTTAGTTGAGCACTCTAAGGAGACTGCCGGTGATAAGCC
GCGAGGAAGGTGGGGATGACGTCAAGTCCTCATGGCCCTTACGGGCTGGG
CTACACACGTGCTACAATGGCGGTGACAATGGGATGCTAAGGGGCGACCC
TTCGCAAATCTCAAAAAGCCGTCTCAGTTCGGATTGGGCTCTGCAACTCG
AGCCCATGAAGTTGGAATCGCTAGTAATCGTGGATCAGCACGCCACGGTG
AATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAGTTGG
CTTTACCTGAAGACGGTGC
>Oscillatoria_PCC10802
CTGGCTCAGGATGAACGCTGGCGGTCTGCTTAACACATGCAAGTCGAACG
GGAGGAGAAATCCTCTAGTGGCGGACGGGTGAGTAACGCGTGAGAATCTG
CCTTCAGGTCGGGGACAACAGCTGGAAACGGCTGCTAAGACCCGATGTGC
CGAGAGGTGAAAGATTTATCGCCTGAAGATGAGCTCGCGTCCGATTAGCT
AGTTGGTAGGGTTATAGCCTACCAAGGCGACGATCGGTAACTGGTCTGAG
AGGACGACCAGTCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGA
GGCAGCAGTGGGGAATTTTCCGCAATGGGCGAAAGCCTGACGGAGCAAGA
CCGCGTGAGGGACGAAGGCTCTTGGGTTGTAAACCTCTTTTCTCAGGGAA
GAAACCTGACGGTACCTGAGGAATAAGCATCGGCTAACTCCGTGCCAGCA

GCCGCGGTAAGACGGAGGATGCAAGCGTTATCCGGAATGATTGGGCGTAA
AGCGTTCGCAGGTGGCTGTTCAAGTCGGCTGTCAAAGACTGGGGCTTAAC
CCCGGAAAGGCAGTGGAAACTGAACGGCTAGAGTGCATTAGGGGTAGAGG
GAATTCCCAGTGTAGCGGTGAAATGCGTAGAGATTGGGAAGAACACCGGT
GGCGAAAGCGCTCTGCTGGACTGCAACTGACACTGAGGGACGAAAGCTAG
GGGAGCGAAAGGGATTAGATACCCCTGTAGTCCTAGCCGTAAACGATGGA
TACTAGGCGTTGCCGGGCTCGACCCTGGCAGTGCCGGAGCTAACGCGTTA
AGTATCCCGCCTGGGAAGTACGCACGCAAGTGTGAAACTCAAAGGAATTG
ACGGGGGCCCGCACAAGCGGTGGAGTATGTGGTTTAATTCGATGCAACGC
GAAGAACCTTACCAGGGCTTGACATGTCGCGAACCCCGCTGAAAGGTGGG
GGTGCCTTAGGGAGCGCGAACACAGGTGGTGCATGGCTGTCGTCAGCTCG
TGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCTCGTGTTC
AGTTGCCATCATTAAGTTGGGAACTCTGAACAGACTGCCGGTGACAAACC
GGAGGAAGGTGGGGATGACGTCAAGTCAGCATGCCCCTTACGTCCTGGGC
AACACACGTACTACAATGCTGCGGACAGAGGGTAGCGAGCCGGTGACGGC
AAGCCAATCCCGCAAACCGTGGCTCAGTTCAGATCGCAGGCTGCAACTCG
CCTGCGTGAAGGCGGAATCGCTAGTAATCGCAGGTCAGCATACTGCGGTG
AATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAGCTGG
TTTAGCCCGAAGTCGTTA

>Burkholderia_cepacia

CTGGCTCAGATTGAACGCTGGCGGCATGCCTTACACATGCAAGTCGAACG
GCAGCACGGGTGCTTGCACCTGGTGGCGAGTGGCGAACGGGTGAGTAATA
CATCGGAACATGTCCTGTAGTGGGGGATAGCCCGGCGAAAGCCGGATTAA
TACCGCATACGATCTACGGATGAAAGCGGGGGACCTTCGGGCCTCGCGCT
ATAGGGTTGGCCGATGGCTGATTAGCTAGTTGGTGGGGTAAAGGCCTACC
AAGGCGACGATCAGTAGCTGGTCTGAGAGGACGACCAGCCACACTGGGAC
TGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAATTTTGGAC
AATGGGCGAAAGCCTGATCCAGCAATGCCGCGTGTGTGAAGAAGGCCTTC
GGGTTGTAAAGCACTTTTGTCCGGAAAGAAATCCCTGGTTCTAATATAGC
CGGGGGATGACGGTACCGGAAGAATAAGCACCGGCTAACTACGTGCCAGC
AGCCGCGGTAATACGTAGGGTGCGAGCGTTAATCGGAATTACTGGGCGTA
AAGCGTGCGCAGGCGGTTTGCTAAGACCGATGTGAAATCCCCGGGCTCAA
CCTGGGAACTGCATTGGTGACTGGCAGGCTAGAGTATGGCAGAGGGGGTA
GAATTCCACGTGTAGCAGTGAAATGCGTAGAGATGTGGAGGAATACCGAT
GGCGAAGGCAGCCCCCTGGGCCAATACTGACGCTCATGCACGAAAGCGTG

GGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCCTAAACGATGTC
AACTAGTTGTTGGGGATTCATTTCCTTAGTAACGTAGCTAACGCGTGAAG
TTGACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAAGGAATTGAC
GGGGACCCGCACAAGCGGTGGATGATGTGGATTAATTCGATGCAACGCGA
AAAACCTTACCTACCCTTGACATGGTCGGAATCCTGCTGAGAGGTGGGAG
TGCTCGAAAGAGAACCGGCGCACAGGTGCTGCATGGCTGTCGTCAGCTCG
TGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCTTGTCCTT
AGTTGCTACGCAAGAGCACTCTAAGGAGACTGCCGGTGACAAACCGGAGG
AAGGTGGGGATGACGTCAAGTCCTCATGGCCCTTATGGGTAGGGCTTCAC
ACGTCATACAATGGTCGGAACAGAGGGTTGCCAACCCGCGAGGGGGAGCT
AATCCCAGAAAACCGATCGTAGTCCGGATTGCACTCTGCAACTCGAGTGC
ATGAAGCTGGAATCGCTAGTAATCGCGGATCAGCATGCCGCGGTGAATAC
GTTCCCGGGTCTTGTACACACCGCCCGTCACACCATGGGAGTGGGTTTTA
CCAGAAGTGGCTAG
>Ruminococcus_albus
AGAGTTTGATCCTGGCTCAGGACGAACGCTGGCGGCACGCTTAACACATG
CAAGTCGAACGAGCGAAAGAGTGCTTGCACTCTCTAGCTAGTGGCGGACG
GGTGAGTAACACGTGAGCAATCTGCCTTTCGGAGAGGGATACCAATTGGA
AACGATTGTTAATACCTCATAACATAACGAAGCCGCATGACTTTGTTATC
AAATGAATTTCGCCGAAAGATGAGCTCGCGTCTGATTAGGTAGTTGGTGA
GGTAACGGCCCACCAAGCCGACGATCAGTAGCCGGACTGAGAGGTTGAAC
GGCCACATTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGT
GGGGAATATTGCACAATGGGCGAAAGCCTGATGCAGCGATGCCGCGTGAG
GGAAGAAGGTTTTAGGATTGTAAACCTCTGTCTTTGGGGACGATAATGAC
GGTACCCAAGGAGGAAGCTCCGGCTAACTACGTGCCAGCAGCCGCGGTAA
TACGTAGGGAGCGAGCGTTGTCCGGAATTACTGGGTGTAAAGGGAGCGTA
GGCGGGATTGCAAGTCAGGTGTGAAATTTAGGGGCTTAACCCCTGAACTG
CACTTGAAACTGTAGTTCTTGAGTGAAGTAGAGGTAAGCGGAATTCCTAG
TGTAGCGGTGAAATGCGTAGATATTAGGAGGAACATCAGTGGCGAAGGCG
GCTTACTGGGCTTTAACTGACGCTGAGGCTCGAAAGCGTGGGGAGCAAAC
AGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGATTACTAGGTGT
GGGGGGACTGACCCCTTCCGTGCCGCAGTTAACACAATAAGTAATCCACC
TGGGGAGTACGGCCGCAAGGCTGAAACTCAAAGGAATTGACGGGGACCCG
CACAAGCAGTGGAGTATGTGGTTTAATTCGAAGCAACGCGAAGAACCTTA
CCAGGTCTTGACATCGTACGCATAGCATAGAGATATGTGAAATCCCTTCG

GGGACGTATAGACAGGTGGTGCATGGTTGTCGTCAGCTCGTGTCGTGAGA
TGTTGGGGTTAAGTCCCGCAACGAGCGCAACCCTTACTGTTAGTTGCTAC
GCAAGAGCACTCTAGCAGGACTGCCGTTGACAAAACGGAGGAAGGTGGGG
ATGACGTCAAATCATCATGCCCCTTATGACCTGGGCTACACACGTACTAC
AATGGCTGTTAACAGAGGGAAGCAAAACAGTGATGTGGAGCAAAACCCTA
AAAGCAGTCTTAGTTCGGATTGTAGGCTGCAACCCGCCTACATGAAGTCG
GAATTGCTAGTAATCGCGGATCAGCATGCCGCGGTGAATACGTTCCCGGG
CCTTGTACACACCGCCCGTCACGCCATGGGAGTCGGTAACACCCGAAGCC
TGTGTTCTAACCGCAAGGAGGAAGCAGTCGAAGGTGGGATTGATGACTGG
GGTGAAGTCGTAACAAGGTAGCCGTATCGGAAGGTGCGGCTGGATCACCT
>Cupriavidus_taiwanensis
AGATTGAACTGAAGAGTTTGATCCTGGCTCAGATTGAACGCTGGCGGCAT
GCCTTACACATGCAAGTCGAACGGCAGCGCGGGCTTCGGCCTGGCGGCGA
GTGGCGAACGGGTGAGTAATACATCGGAACGTGCCCTGTCGTGGGGGATA
ACTAGTCGAAAGATTAGCTAATACCGCATACGACCTGAGGGTGAAAGCGG
GGGACCGCAAGGCCTCGCGCGATAGGAGCGGCCGATGTCTGATTAGCTAG
TTGGTGGGGTAAAGGCCTACCAAGGCGACGATCAGTAGCTGGTCTGAGAG
GACGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGG
CAGCAGTGGGGAATTTTGGACAATGGGGGCAACCCTGATCCAGCAATGCC
GCGTGTGTGAAGAAGGCCTTCGGGTTGTAAAGCACTTTTGTCCGGAAAGA
AATGGCCCTGGCTAATACCCGGGGTCGATGACGGTACCGGAAGAATAAGC
ACCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGTGCGAGCGT
TAATCGGAATTACTGGGCGTAAAGCGTGCGCAGGCGGTTTGATAAGACAG
GCGTGAAATCCCCGGGCTCAACCTGGGAATGGCGCTTGTGACTGTCAGGC
TAGAGTGCGTCAGAGGGGGGTAGAATTCCACGTGTAGCAGTGAAATGCGT
AGAGATGTGGAGGAATACCGATGGCGAAGGCAGCCCCCTGGGACGTGACT
GACGCTCATGCACGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGT
AGTCCACGCCCTAAACGATGTCAACTAGTTGTTGGGGATTCATTTCTTCA
GTAACGTAGCTAACGCGTGAAGTTGACCGCCTGGGGAGTACGGTCGCAAG
ATTAAAACTCAAAGGAATTGACGGGGACCCGCACAAGCGGTGGATGATGT
GGATTAATTCGATGCAACGCGAAAAACCTTACCTACCCTTGACATGCCAC
TAACGAAGCAGAGATGCATTAGGTGCCCGAAAGGGAAAGTGGACACAGGT
GCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCG
CAACGAGCGCAACCCTTGTCTCTAGTTGCTACGCAAGAGCACTCTAGAGA
GACTGCCGGTGACAAACCGGAGGAAGGTGGGGATGACGTCAAGTCCTCAT

GGCCCTTATGGGTAGGGCTTCACACGTCATACAATGGTGCGTACAGAGGG

TTGCCAACCCGCGAGGGGGAGCTAATCCCAGAAAACGCATCGTAGTCCGG

ATCGTAGTCTGCAACTCGACTACGTGAAGCTGGAATCGCTAGTAATCGCG

GATCAGCATGCCGCGGTGAATACGTTCCCGGGTCTTGTACACACCGCCCG

TCACACCATGGGAGTGGGTTTTGCCAGAAGTAGTTAGCCTAACCGCAAGG

AGGGCGATTACCACGGCAGGGTTCATGACTGGGGTGAAGTCGTAACAAGG

TAGCCGTATCGGAAGGTGCGGCTGGATCACCTCCT

>Pseudomonas _fluorescens

TTGAACTGAAGAGTTTGATCATGGCTCAGATTGAACGCTGGCGGCAGGCC

TAACACATGCAAGTCGAGCGGTAGAGAGAAGCTTGCTTCTCTTGAGAGCG

GCGGACGGGTGAGTAATGCCTAGGAATCTGCCTGGTAGTGGGGGATAACG

TTCGGAAACGGACGCTAATACCGCATACGTCCTACGGGAGAAAGCAGGGG

ACCTTCGGGCCTTGCGCTATCAGATGAGCCTAGGTCGGATTAGCTAGTTG

GTGAGGTAATGGCTCACCAAGGCGACGATCCGTAACTGGTCTGAGAGGAT

GATCAGTCACACTGGAACTGAGACACGGTCCAGACTCCTACGGGAGGCAG

CAGTGGGGAATATTGGACAATGGGCGAAAGCCTGATCCAGCCATGCCGCG

TGTGTGAAGAAGGTCTTCGGATTGTAAAGCACTTTAAGTTGGGAGGAAGG

GTTGTAGATTAATACTCTGCAATTTTGACGTTACCGACAGAATAAGCACC

GGCTAACTCTGTGCCAGCAGCCGCGGTAATACAGAGGGTGCAAGCGTTAA

TCGGAATTACTGGGCGTAAAGCGCGCGTAGGTGGTTTGTTAAGTTGGATG

TGAAATCCCCGGGCTCAACCTGGGAACTGCATTCAAAACTGACTGACTAG

AGTATGGTAGAGGGTGGTGGAATTTCCTGTGTAGCGGTGAAATGCGTAGA

TATAGGAAGGAACACCAGTGGCGAAGGCGACCACCTGGACTAATACTGAC

ACTGAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGT

CCACGCCGTAAACGATGTCAACTAGCCGTTGGAAGCCTTGAGCTTTTAGT

GGCGCAGCTAACGCATTAAGTTGACCGCCTGGGGAGTACGGCCGCAAGGT

TAAAACTCAAATGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGG

TTTAATTCGAAGCAACGCGAAGAACCTTACCAGGCCTTGACATCCAATGA

ACTTTCTAGAGATAGATTGGTGCCTTCGGGAACATTGAGACAGGTGCTGC

ATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGTAACG

AGCGCAACCCTTGTCCTTAGTTACCAGCACGTAATGGTGGGCACTCTAAG

GAGACTGCCGGTGACAAA

>Nostoc_azollae

AGAGTTTGATCCTGGCTCAGGATGAACGCTGGCGGTATGCTTAACACATG

CAAGTCGAACGAAGTTCTTAGGAACTTAGTGGCGGACGGGTGAGTAACGC

GTGAGAATTTGGCTTTAGGTCGGGGACAACCACTGGAAACGGTGGCTAAT
ACCGGATATGCCGAGAGGTGAAAGATTTATTGCCTGAAGATAAGCTCGCG
TCTGATTAGCTAGTTGGTGGGGTAAGAGCCTACCAAGGCGACGATCAGTA
GCTGGTCTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGA
CTCCTACGGGAGGCAGCAGTGGGGAATTTTCCGCAATGGGCGAAAGCCTG
ACGGAGCAATACCGCGTGAGGGAGGAAGGCTCTTGGGTCGTAAACCTCTT
TTCTCAGGGAAGAAAAAAATGACGGTACCTGAGGAATAAGCATCGGCTAA
CTCCGTGCCAGCAGCCGCGGTAATACGGAGGATGCAAGCGTTATCCGGAA
TGATTGGGCGTAAAGGGTCCGTAGGTGGTACTGTAAGTCTGCTGTTAAAG
AGCAAGGCTCAACCTTGTAAAAGCAGTGGAAACTACAGAACTAGAGTGCG
TTCGGGGCAGAAGGGAATTCCTGGTGTAGCGGTGAAATGCGTAGATATCA
GGAAGAACACCGGTGGCGAAAGCGTTCTGCTAGGCCGCAACTGACACTGA
GGGACGAAAGCTAGGGGAGCGAATGGGATTAGATACCCCAGTAGTCCTAG
CCGTAAACGATGGATACTAGGCGTGGCTTGTATCGACCCGAGCCGTGCCG
GAGCTAACGCGTTAAGTATCCCGCCTGGGGAGTACGCACGCAAGTGTGAA
ACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGTATGTGGTTTA
ATTCGATGCAACGCGAAGAACCTTACCAAGACTTGACATGTCGCGAATTT
CGGTGAAAGCTGAGAGTGCCTTCGGGAGCGCGAACACAGGTGGTGCATGG
CTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCG
CAACCCTCGTTTTTAGTTGCCAGCATTAAGTTGGGCACTCTAAAGAGACT
GCCGGTGACAAACCGGAGGAAGGTGAGGATGACGTCAAGTCAGCATGCCC
CTTACGTTTTGGGCTACACACGTACTACAATGCTACGGACAAAGGGCAGC
AAGCTAGCAATAGCAAGCAAATCTCATAAACCGTAGCTCAGTTCAGATCG
AAGGCTGCAACTCGCCTTCGTGAAGGAGGAATCGCTAGTAATTGCAGGTC
AGCATACTGCAGTGAATTCGTTCCCGGGCCTTGTACACACCGCCCGTCAC
ACCATGGAAGTTGGTCACGCCCGAAGTCATTACCCCAACCTTTTGGAGGG
GGATGCCTAAGGTAGGACTGGTGACTGGGGTGAAGTCGTAACAAGGTAGC
CGTACCGGAAGGTGTGGCTGGATCACCT

>Rhizobium_leguminosarum

CTCAACATGAGAGTTTGATCCTGGCTCAGAACGAACGCTGGCGGCAGGCT
TAACACATGCAAGTCGAGCGCCCCGCAAGGGGAGCGGCAGACGGGTGAGT
AACGCGTGGGAATCTACCCTTGACTACGGAATAACGCAGGGAAACTTGTG
CTAATACCGTATGTGTCCTTCGGGAGAAAGATTTATCGGTCAAGGATGAG
CCCGCGTTGGATTAGCTAGTTGGTGGGGTAAAGGCCTACCAAGGCGACGA
TCCATAGCTGGTCTGAGAGGATGATCAGCCACATTGGGACTGAGACACGG

CCCAAACTCCTACGGGAGGCAGCAGTGGGGAATATTGGACAATGGGCGCA
AGCCTGATCCAGCCATGCCGCGTGAGTGATGAAGGCCCTAGGGTTGTAAA
GCTCTTTCACCGGAGAAGATAATGACGGTATCCGGAGAAGAAGCCCCGGC
TAACTTCGTGCCAGCAGCCGCGGTAATACGAAGGGGGCTAGCGTTGTTCG
GAATTACTGGGCGTAAAGCGCACGTAGGCGGATCGATCAGTCAGGGGTGA
AATCCCAGGGCTCAACCCTGGAACTGCCTTTGATACTGTCGATCTGGAGT
ATGGAAGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAATTCGTAGATAT
TCGGAGGAACACCAGTGGCGAAGGCGGCTCACTGGTCCATTACTGACGCT
GAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCA
CGCCGTAAACGATGAATGTTAGCCGTCGGGCAGTATACTGTTCGGTGGCG
CAGCTAACGCATTAAACATTCCGCCTGGGGAGTACGGTCGCAAGATTAAA
ACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGTTTA
ATTCGAAGCAACGCGCAGAACCTTACCAGCCCTTGACATGCCCGGCTACT
TGCAGAGATGCAAGGTTCCCTTCGGGGACCGGGACACAGGTGCTGCATGG
CTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCG
CAACCCTCGCCCTTAGTTGCCAGCATTCAGTTGGGCACTCTAAGGGGACT
GCCGGTGATAAGCCGAGAGGAAGGTGGGGATGACGTCAAGTCCTCATGGC
CCTTACGGGCTGGGCTACACACGTGCTACAATGGTGGTGACAGTGGGCAG
CGAGCACGCGAGTGTGAGCTAATCTCCAAAAGCCATCTCAGTTCGGATTG
CACTCTGCAACTCGAGTGCATGAAGTTGGAATCGCTAGTAATCGCGGATC
AGCATGCCGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCAC
ACCATGGGAGTTGGTTTTACCCGAAGGTAGTGCGCTAACCGCAAGGAGGC
AGCTAACCACGGTAGGGTCAGCGACTGGGGTGAAGTCGTAACAAGGTAGC
CGTAGGGGAACCTGCGGCTGGATCACCTCCTTTCT

>Azotobacter_vinelandii

AACTGAAGAGTTTGATCATGGCTCAGATTGAACGCTGGCGGCAGGCCTAA
CACATGCAAGTCGAGCGGCAGCGGGACCTTCGGGTTGCCGGCGAGCGGCG
GACGGGTGAGTAATGCCTAGGAATCTGCCTGTTAGTGGGGGATAACGCGG
GGAAACTCGCGCTAATACCGCATACGTCCTACGGGAGAAAGTGGGGGACC
CTCGGGCCTCACGCTAACAGATGAGCCTAGGTCGGATTAGCTAGTTGGTG
GGGTAAAGGCCCACCAAGGCGACGATCCGTAACTGGTCTGAGAGGATGAT
CAGTCACACTGGAACTGAGACACGGTCCAGACTCCTACGGGAGGCAGCAG
TGGGGAATATTGGACAATGGGCGAAAGCCTGATCCAGCCATGCCGCGTGT
GTGAAGAAGGTCTTCGGATTGTAAAGCACTTTAAGTTGGGAGGAAGGGCG
CTCGGTGAATACCCAAGCGTCTTGACGTTACCGACAGAATAAGCACCGGC

TAACTTCGTGCCAGCAGCCGCGGTAATACGAAGGGTGCAAGCGTTAATCG
GAATTACTGGGCGTAAAGCGCGCGTAGGTGGTTCGGCAAGTTGGATGTGA
AAGCCCCGGGCTCAACCTGGGAACCGCATCCAAAACTACTGGGCTAGAGT
ACGGTAGAGGGTGGTGGAATTTCCTGTGTAGCGGTGAAATGCGTAGATAT
AGGAAGGAACACCAGTGGCGAAGGCGACCACCTGGACCGATACTGACACT
GAGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCA
CGCCGTAAACGATGTCGACTAGCCGTTGGGCTCCTTGAGAGCTTAGTGGC
GCAGCTAACGCATTAAGTCGACCGCCTGGGGAGTACGGCCGCAAGGTTAA
AACTCAAATGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGTTT
AATTCGAAGCAACGCGAAGAACCTTACCTGGCCTTGACATCCTGCGAACT
TTCAAGAGATTGATGGGTGCCTTCGGGAGCGCAGAGACAGGTGCTGCATG
GCTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGTAACGAGC
GCAACCCTTGTCCTTAGTTACCAGCGATTCGGTCGGGCACTCTAAGGAGA
CTGCCGGTGACAAACCGGAGGAAGGTGGGGATGACGTCAAGTCATCATGG
CCCTTACGGCCAGGGCTACACACGTGCTACAATGGTCGGTACAAAGGGTT
GCCAAGTCGCGAGGCGGAGCTAATCCCAGAAAACCGATCGTAGTCCGGAT
CGCAGTCTGCAACTCGACTGCGTGAAGTCGGAATCGCTAGTAATCGCGAA
TCAGAATGTCGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTC
ACACCATGGGAGTGGGTTGCTCCAGAAGTAGCTAGTCTAACCCTCGGGAG
GACGGTTACCACGGAGTGATTCATGACTGGGGTGAAGTCGTAACAAGGTA
GCCGTAGGGGAACCTGCGGCTGGATCACCTCCTTA

>Anabaena_variabilis

AAACGGAGAGTTTGATCCTGGCTCAGGATGAACGCTGGCGGTATGCTTAA
CACATGCAAGTCGAACGGAATCTTCGGATTTAGTGGCGGACGGGTGAGTA
ACGCGTGAGAATCTAGCTCTAGGTCGGGGACAACCACTGGAAACGGTGGC
TAATACCGGATGTGCCGAAAGGTGAAAGATTTATTGCCTAGAGATGAGCT
CGCGTCTGATTAGCTAGTTGGTGTGGTAAGAGCGCACCAAGGCGACGATC
AGTAGCTGGTCTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCC
CAGACTCCTACGGGAGGCAGCAGTGGGGAATTTTCCGCAATGGGCGAAAG
CCTGACGGAGCAATACCGCGTGAGGGAGGAAGGCTCTTGGGTTGTAAACC
TCTTTTCTCAGGGAATAAGAAAGTGAAGGTACCTGAGGAATAAGCATCGG
CTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGATGCAAGCGTTATCC
GGAATGATTGGGCGTAAAGCGTCCGCAGGTGGCACTGTAAGTCTGCTGTT
AAAGAGCAAGGCTCAACCTTGTAAAGGCAGTGGAAACTACAGAGCTAGAG
TACGTTCGGGGCAGAGGGAATTCCTGGTGTAGCGGTGAAATGCGTAGAGA

TCAGGAAGAACACCGGTGGCGAAAGCGCTCTGCTAGGCCGTAACTGACAC
TGAGGGACGAAAGCTAGGGGAGCGAATGGGATTAGATACCCCAGTAGTCC
TAGCCGTAAACGATGGATACTAGGCGTGGCTTGTATCGACCCGAGCCGTG
CCGGAGCCAACGCGTTAAGTATCCCGCCTGGGGAGTACGCACGCAAGTGT
GAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGTATGTGGT
TTAATTCGATGCAACGCGAAGAACCTTACCAAGACTTGACATGTCGCGAA
TCTTCTTGAAAGGGAAGAGTGCCTTAGGGAGCGCGAACACAGGTGGTGCA
TGGCTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGA
GCGCAACCCTCGTTTTTAGTTGCCAGCATTAAGTTGGGCACTCTAGAGAG
ACTGCCGGTGACAAACCGGAGGAAGGTGGGGATGACGTCAAGTCAGCATG
CCCCTTACGTCTTGGGCTACACACGTACTACAATGCTACGGACAGAGGGC
AGCAAGCTAGCGATAGCAAGCAAATCCCGTAAACCGTAGCTCAGTTCAGA
TCGCAGGCTGCAACTCGCCTGCGTGAAGGAGGAATCGCTAGTAATTGCAG
GTCAGCATACTGCAGTGAATTCGTTCCCGGGCCTTGTACACACCGCCCGT
CACACCATGGAAGCTGGCAACGCCCGAAGTCATTACTCCAACCATTCGTG
GGGGAGGATGCCTAAGGCAGTGCTGGTGACTGGGGTGAAGTCGTAACAAG
GTAGCCGTACCGGAAGGTGTGGCTGGATCACCTCCTTT
>Rhodospirillum_rubrum
TTTCGAGATCAACTTGAGAGTTTGATCCTGGCTCAGGACGAACGCTGGCG
GCAGGCCTAACACATGCAAGTCGAACGCATCCTTCGGGATGAGTGGCGCA
CGGGTGAGTAACACGTGGGAACGTACCTTGGAGTGCGGAATAATCTTTGG
AAACGAGGACTAATACCGCATACGCCCTTAGGGGGAAAGATTTATCGCTC
CAAGATCGGCCCGCGTCCGATTAGCTAGTTGGCGGGGTAATGGCCCACCA
AGGCGACGATCGGTAGCTGGTCTGAGAGGATGGCCAGCCACACTGGGACT
GAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAATATTGCGCA
ATGGGGGCAACCCTGACGCAGCCATGCCGCGTGAGTGAAGAAGGCCTTCG
GGTTGTAAAGCTCTTTCGGGTGTGAAGATGATGACGGTAACACCAGAAGA
AGCCCCGGCTAACTTCGTGCCAGCAGCCGCGGTAATACGAAGGGGGCAAG
CGTTGTTCGGAATTACTGGGCGTAAAGAGCGCGTAGGCGGTCTGATTAGT
CAGAGGTGAAATCCCAGAGCTCAACTTTGGAACTGCCTTTGATACTGTTA
GACTAGAATCCGTGAGAGGGTGGTGGAATTCCCAGTGTAGAGGTGAAATT
CGTAGATATTGGGAGGAACACCAGTGGCGAAGGCGGCCACCTGGCGCGGT
ATTGACGCTGAGGCGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCT
GGTAGTCCACGCCGTAAACGATGAGTGCTAGATGTCGGGGTACATGTACC
TCGGTGTCGCAGCTAACGCATTAAGCACTCCGCCTGGGGAGTACGGCCGC

AAGGTTAAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGCA
TGTGGTTTAATTCGAAGCAACGCGCAGAACCTTACCAGCCCTTGACATCC
CGTGACACTTCCAGAGATGGAAGGTTCCCTTCGGGGACACGGTGACAGGT
GCTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCG
CAACGAGCGCAACCCTCATCTTCAGTTGCCAGCAAGTAACGTTGGGCACT
CTGAAGAGACTGCCGGTGACAAGCCGGAGGAAGGTGGGGATGACGTCAAG
TCCTCATGGCCCTTACGGGCTGGGCTACACACGTGCTACAATGGCGCCTA
CAATGGGCAGCGACCTCGCGAGGGGAAGCTAATCTCCAAAAGGCGTCTCA
GTTCGGATTGCACTCTGCAACTCGGGTGCATGAAGTCGGAATCGCTAGTA
ATCGTGGATCAGCATGCCACGGTGAATACGTTCCCGGGCCTTGTACACAC
CGCCCGTCACACCATGGGAGTTGGTTCTACCCGAAGACGGTACGCTAACC
GCAAGGAGGCAGCCGGCCACGGTAGGGTCAGCGACTGGGGTGAAGTCGTA
ACAAGGTAGCCGTAGGGGAACCTGCGGCTGGATCACCTCCTTTCAAGGAT
GGT

>Synechococcus_spJA-3-3Ab

CCTGGCTCAGGATGAACGCTGGCGGTATGCCTAACACATGCAAGTCGGAC
GGACACCTTCGGGTGTCAGTGGCGGACGGGTGAGTAACGCGTAGGAATCT
ACCCCTAGGAGGGGGATAACAGCTGGAAACGGCTGCTAATACCCCATACA
GGCGCGAGCTGAAAGGAGCAATCCGCCTAGGGATGAGCCTGCGTCGGATT
AGCTAGTTGGTGGGGTAATGGCCTACCAAGGCGACGATCCGTAGCGGGTC
TGAGAGGATGGCCCGCCACACTGGGACTGAGACACGGCCCAGACTCCTAC
GGGAGGCAGCAGTGGGGAATTTTCCGCAATGGGCGAAAGCCTGACGGAGC
AATACCGCGTGAGGGAGGAAGGCCTTCGGGTTGTAAACCTCTTTTCTGGG
GGACGAAGGGGTGACGGTACCCCAGGAATAAGCCTCGGCTAACTCTGTGC
CAGCAGCCGCGGTAAGACAGAGGAGGCGAGCGTTATCCGGAGTTATTGGG
CGTAAAGGGTACGTAGGCGGCTGGTCAAGTCTGGGGTTAAAGACCACAGC
TTAACTGTGGGGGGGCCTTGGAAACTGACCGGCTGGAGTGGGGCAGGGGT
TGCGGGAATTCCCGGTGTAGCGGTGAAATGCGTAGATATCGGGAAGAACA
CCTGAGGCGAAAGCGCGCAACTGGGCCCTAACTGACGCTGAGGTACGAAA
GCTAGGGGAGCGAATGGGATTAGATACCCCAGTAGTCCTAGCCGTAAACG
ATGGACACTAGACGTTGGCCGTATCGACCCGGTCAGTGTCGTAGCTAACG
CGTTAAGTGTCCCGCCTGGGGAGTACGCACGCAAGTGTGAAACTCAAAGG
AATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGATGC
AAAGCGAAGAACCTTACCAGGGCTTGACATGCCGCGAACCCTGCGGAAAC
GTGGGGGTGCCTTCGGGAGCGCGGACACAGGTGGTGCATGGCTGTCGTCA

GCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCTCG

TCCTTAGTTGCCAGCATTGAGTTGGGCACTCTGGGGAGACTGCCGGTGAC

AAACCGGAGGAAGGTGGGGACGACGTCAAGTCAGCATGCCCCTTATGCTC

TGGGCTACACACGTGCTACAATGGCGCGGACAAAGGGCAGCGAACGCGCG

AGCGGGAGCCAATCCCAAAAACCGCGTCTCAGTTCAGATCGCAGGCTGCA

ACTCGCCTGCGTGAAGGCGGAATCGCTAGTAATCGCAGATCAGCATGCTG

CGGTGAATACGTTCCCGGGCCTTG

>Methanococcus_maripaludis

TCCGGTTGATCCCGCCGGAGGCCACTGCTATTGGGATTCGACTAAGCCAT

GCGAGTCTATGGTTTCGGCCATGGCGGACGGCTCATTAACACGTGGTTAA

CATACCCTCAGGTGGGGGATAACCTTGGGAAACTGAGGATAATACCCCAT

AGAAAAAGCAGTCTGGAATGATTCTTTTTCGAAAGCACATGCGCCTGAGG

ATTGGACTGCGCTCGATTAGGTAGTTGGTGGGGTAATGGCCCACCAAGCC

TACGATCGGTACGGGCCTTGAGAGAGGGAGCCCGGAGATGGGGACTGAGA

CACGGCCCCAGGCCCTACGGGGCGCAGCAGGCGCGAAACCTCCGCAATGC

ACGAAAGTGCGACGGGGGGATCCCAAGTGCTCATGCACAGCATGGGCTTT

TATCAAGTGTAAACAGCTTGAGGAATAAGGGCTGGGCAAGTTCGGTGCCA

GCAGCCGCGGTAACACCGACGGCCCGAGTGGTGGCCACTCTTATTGGGCC

TAAAGCGTCCGTAGCCGGTCCAGTAAGTCCCTGTTTAAATCCTGCGGCTT

AACCGCAGGACTGGCAGGGATACTGCTGGACTTGGGACCGGGAGAGGACA

AGGGTACTTCAGGGGTAGCGGTGAAATGTGTTGATCCTTGAAGGACCACC

TATGGCGAAGGCACTTGTCTGGAACGGGTCCGACGGTGAGGGACGAAAGC

CAGGGGCGCGAACCGGATTAGATACCCGGGTAGTCCTGGCCGTAAACTCT

GCGAACTAGGTGTTAGGTAGGCCCCGTGCCTATCTAGTGCCGAAGGGAAG

CCGTTAAGTTCGCCGCCTGGGGAGTACGGTCGCAAGACTGAAACTTAAAG

GAATTGGCGGGGGAGCACCACAACGGGTGGAGCCTGCGGTTTAATTGGAT

TCAACGCCGGGCATCTCACCAGGAGCGACAGCATGATGACGGCCAGGTTG

ACGACCTTGCCTGAAGCGCTGAGAGGTGGTGCATGGCCATCGTCAGCTCG

TACCGCGAGGCGTCCTGTTAAGTCAGGTAACGAGCGAGACCCGTGCCCTA

TGTTGCTACTTTTTCCTCCGGGAGAAAGGCACTCATAGGGGACCGCTGGC

GCTAAGTCAGAGGAAGGAGCGGGCAACGATAGGTCCGCATGCCCCGAATC

TCCTGGGCTACACGCGGGCTACAATGGCTAGGACAATGGGCCGCAACCCT

GAAAAGGGACGCAAATCTCCTAAACCTAGTCGTAGTTCGGATCGTGGGCT

GTAACTCGCCCACGTGAAGCTGGAATCCGTAGTAATCGCAGTTCATAATA

CTGCGGTGAATGTGTCCCTGCTCCTTGCACACACCGCCCGTCACACCACC

CGAGTTGGGTTGAAGTGAGGCCTTGGCCTTTGGCTAGGGTCGAACTTGGG
CTCAGCGAGGGGGGTGAAGTCGTAACAAGGTAGCCGTAGGGGAACCTGCG
GCTGGATCACCTCCT
>Clostridium_beijerinckii
AGAGTTTGATCCTGGCTCAGGACGAACGCTGGCGGCGTGCTTAACACATG
CAAGTCGAGCGATGAAGCTCCTTCGGGAGTGGATTAGCGGCGGACGGGTG
AGTAACACGTGGGTAACCTGCCTCATAGAGGGGAATAGCCTTTCGAAAGG
AAGATTAATACCGCATAAGATTGTAGTGCCGCATGGTATAGCAATTAAAG
GAGTAATCCGCTATGAGATGGACCCGCGTCGCATTAGCTAGTTGGTGAGG
TAACGGCTCACCAAGGCGACGATGCGTAGCCGACCTGAGAGGGTGATCGG
CCACATTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGG
GGAATATTGCACAATGGGGGAAACCCTGATGCAGCAACGCCGCGTGAGTG
ATGACGGTCTTCGGATTGTAAAGCTCTGTCTTCAGGGACGATAATGACGG
TACCTGAGGAGGAAGCCACGGCTAACTACGTGCCAGCAGCCGCGGTAATA
CGTAGGTGGCAAGCGTTGTCCGGATTTACTGGGCGTAAAGGGAGCGTAGG
TGGATATTTAAGTGGGATGTGAAATACTCGGGCTTAACCTGGGTGCTGCA
TTCCAAACTGGATATCTAGAGTGCAGGAGAGGAAAGTAGAATTCCTAGTG
TAGCGGTGAAATGCGTAGAGATTAGGAAGAATACCAGTGGCGAAGGCGAC
TTTCTGGACTGTAACTGACACTGAGGCTCGAAAGCGTGGGGAGCAAACAG
GATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAATACTAGGTGTAG
GGGTTGTCATGACCTCTGTGCCGCCGCTAACGCATTAAGTATTCCGCCTG
GGGAGTACGGTCGCAAGATTAAAACTCAAAGGAATTGACGGGGGCCCGCA
CAAGCAGCGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAACCTTACC
TAGACTTGACATCTCCTGAATTACCCTTAATCGGGGAAGCCCTTCGGGGC
AGGAAGACAGGTGGTGCATGGTTGTCGTCAGCTCGTGTCGTGAGATGTTG
GGTTAAGTCCCGCAACGAGCGCAACCCTTATTGTTAGTTGCTACCATTTA
GTTGAGCACTCTAGCGAGACTGCCTGGGTTAACCGGGAGGAAGGTGGGGA
TGACGTCAAATCATCATGCCCCTTATGTCTAGGGCTACACACGTGCTACA
ATGGCTGGTACAGAGAGATGCTAAACCGCGAGGTGGAGCCAAACTTTAAA
ACCAGTCTCAGTTCGGATTGTAGGCTGAAACTCGCCTACATGAAGCTGGA
GTTGCTAGTAATCGCGAATCAGAATGTCGCGGTGAATACGTTCCCGGGCC
TTGTACACACCGCCCGTCACACCATGAGAGTTGGCAATACCCAAAGTTCG
TGAGCTAACGCGCAAGCGAGGCAGCGACCTAAGGTAGGGTCAGCGATTGG
GGTGAAGTCGTAACAAGGTAGCCGTAGGAGAACCTGCGGCTGGATCACCT
CCTTT

Appendix 8. Suggestions for bioinformatics extension exercises to emphasize enzyme catalysis based on nitrogenase.

The huge background knowledge of the protein nitrogenase provides teachers with a priceless opportunity for working on structural bioinformatics skills in a biochemically-focused extension of the proposed microbiological research project. Students would have the chance to grasp an actual active site. For example, (1) protein alignments among nitrogenases of different organisms could be used to detect variable and invariable amino acid regions, and (2) guided-exploration of the structure of nitrogenase active site (there are several x-ray structures in Protein Data Bank). These two activities could be connected with the catalysis (and the role of metals in nitrogenase catalysis) and the chemistry of N_2 reduction to NH_3.

* For alignment exercises students could align NifH sequences from various organisms using BLOSUM62 or similar scoring matrices found in alignment programs such as Clustal-X, and search for the conserved MgATP-binding motif of GXGXXG. Or align NifD sequences and identify the conserved cysteine and histidine that coordinate FeMoco.

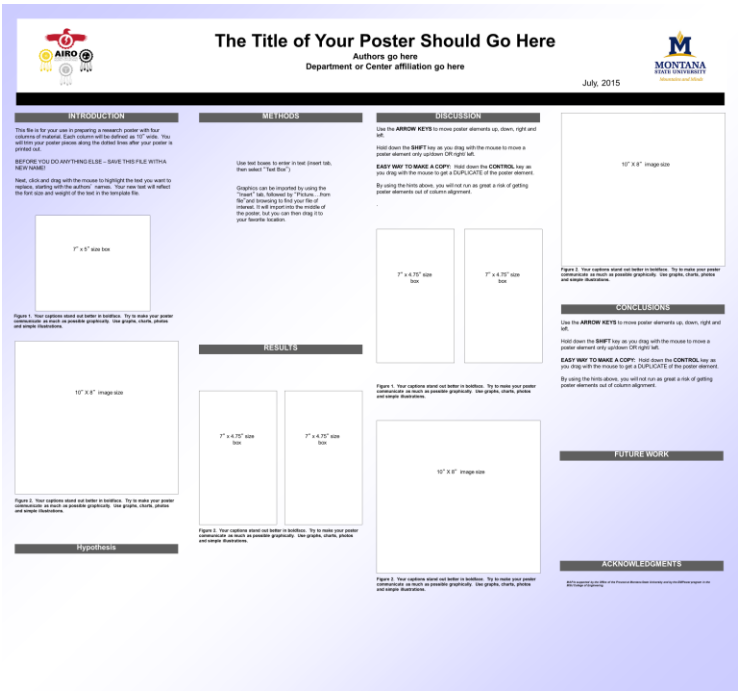
  A diversity of NifD and NifH sequences from various organism that could be used for these exercises can be found in the Protein database of NCBI (http://www.ncbi.nlm.nih.gov/). Simply select the Protein option in the database pulldown menu and type NifD or NifH. After sequences have been selected, they can be exported through the "Send to" menu at the bottom. Choose File and FASTA format. These sequences can then be aligned through alignment software such as Clustal-X

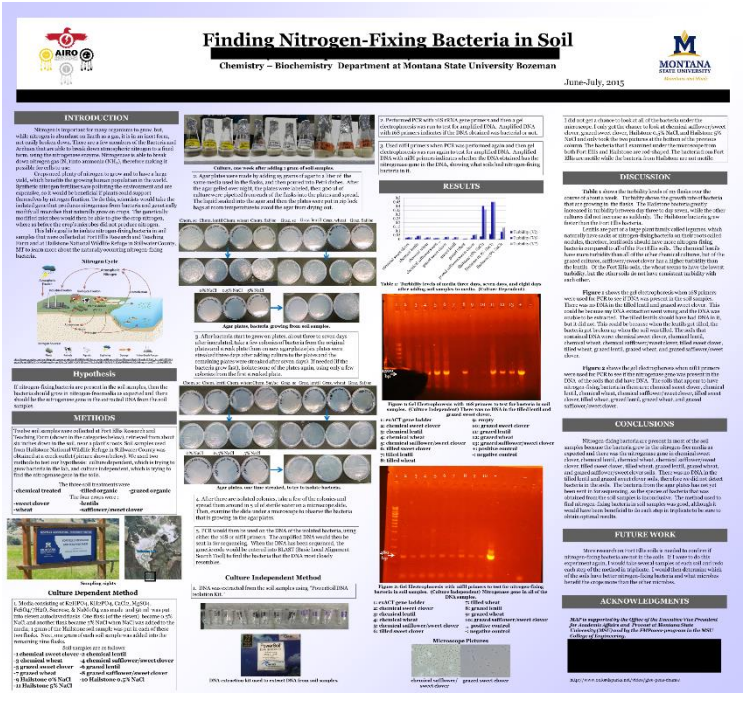
* For structural exercises students could access the Protein Data Bank (PDB), http://www.rcsb.org/pdb/home/home.do, (which could be used as an example of a similar database as those in BLAST and NCBI but this time, instead of sequences, the database is of structures). Nitrogenase structures can be accessed by typing nitrogenase, NifH, NifDK, or Nitrogenase Fe protein, etc. in the search menu, or by typing a protein database ID, such as

3U7Q, for the MoFe protein component of nitrogenase (NifDK), 1G5P for the Fe protein (NifH), or 4WZB for the entire complex. The PDB has a 3D view/JSMol viewing tool where students can choose which features to highlight in the structure. Alternatively, a viewing program such as COOT could also be used.

Appendix 9. Poster template and poster example

We take pictures of students sampling, and encourage them to photograph their liquid cultures, agar plates, and to save their micrographs and gel pictures, in addition to generating tables and graphs based on their DNA concentration and liquid culture O.D. data.





Appendix 10. Pre and post research experience survey

1. How many people live on Earth now?

2. How many people are projected to be living on Earth by the year 2050?

3. What nutrients are included in fertilizers for crops?

4. How did crops grow before fertilizers?

5. What are some pros and cons of fertilizer use?

6. What is the Haber Bosch process?

7. What is special about legumes and microorganisms that is not true for cereal plants?

8. What is nitrogenase?

9. What is nitrogen fixation?

10. Can you name microorganisms able to fix nitrogen?

11. What are redox states?

12. What are some approaches of studying microorganisms in the environment?

13. What are the most common elements in living organisms?

14. What are some essential nutrients included in microbiological media?

15. What is an enrichment culture?

16. What is the role of the autoclave?

17. What is the role of agar in microbiological media?

18. What do we use the polymerase chain reaction for?

19. What do we use gel electrophoresis for?

20. What is the 16S rRNA gene commonly used for in microbiological research?

21. What is a phylogenetic tree?