PhD Project Advertisement

Project title: Multi-omics analysis of nitrogen metabolism by the soil microbial community
Project No: FBS2024-047-Larionov-cr

Lead supervisor: Alexey Larionov, Bioinformatics group at the School of Water, Energy and Environment, Cranfield University

Email: alexey.larionov@cranfield.ac.uk
Co-supervisor:
Tom Sizmur, University of Reading

Project description:

This is a bioinformatics project aiming to develop new tools to analyse multi-omics long reads sequencing data, for a better understanding of nitrogen metabolism in soil.

Nitrogen is an essential nutrient for plants. Although gaseous nitrogen (N_2) is the most abundant component of the atmosphere, it is not bioavailable for plants in this form. It must be converted into ammonia (NH_3) or nitrate (NH_3^-) to be used by plants. Converting N_2 to ammonia is called “nitrogen fixation”, and it is done by diazotrophic microorganisms in soil. Then soil microorganisms may further convert ammonia to nitrites and nitrates NO_2^-/NO_3^- (the process called “nitrification”). Finally, denitrifying bacteria reduce nitrites and nitrates to NO, N_2O or N_2 (“denitrification”), completing the nitrogen cycle in soil (Kuypers,2018). The microbial nitrogen cycle is the only natural source of nitrogen for plants; the other source is from the nitrogen fertilisers. However, often less than 50% of bio-available N is taken up by agricultural crops (Lehnert,2018). The remaining nitrogen is either leached to the aquatic environment as a nitrate (a major contributor to eutrophication), or emitted to the atmosphere as N_2O (a potent greenhouse gas) or ammonia (a contributor to poor air quality). Therefore, understanding the biology of the soil N-cycle in soil is essential for both improving sustainable food production and protecting the environment.

The soil N cycle has been extensively studied (e.g. see Kuypers,2018, Lehnert,2018, Ouyang,2018, Sun 2022). Examples of the key microbial genes involved in N-metabolism include nitrogenase (nifH), ammonia-monooxygenases (amoA, archaeal and bacterial), nitrite and nitrous-oxide reductases (nirK/S,nosZ), urease (ureC) and others. Each of the N-cycle steps can be performed by multiple microbial species, for instance: denitrification genes are present in tens of genera. However, so far most studies measure the abundance of N-cycle genes in soli either by (i) qPCR simultaneously reporting the gene abundancies of multiple species, or (ii) by assessing the composition of soil microbial community, including N-metabolising taxa, using 16S amplicon sequencing, which does not measure the actual species’ contributions to the N-cycle. These conventional approaches provide no information about species-specific gene expression at RNA level, and how the RNA expression is regulated.

The goal of this project is to leverage on the recent development of long-read sequencing technologies (PacBio and Nanopore) to develop the experimental approches and bioinformatic pipelines and resources necessary to study these currently unknown aspects of N-cycling in soil.

PacBio and Nanopore sequencing technologies generate fragments spanning an entire gene (or multiple genes) in one read, and can report epigenetic marks at the same time. The differences between homologous genes in different microbial taxa should allow us to distinguish the taxa of origin for each DNA fragment containing an N-metabolising gene. For the same reason, long-read RNA-sequencing will allow us to measure species-specific gene expression. This would not be possible using conventional qPCR or short-read technologies (such as Illumina sequencing).

You will be based in the Cranfield University Bioinformatics Group. The project will start by reviewing currently available bioinformatics resources for N-cycle genes in soil, and the recent developments of long read sequencing technologies relevant to the project (e.g. PacBio Revio and MAS-seq). This will allow you to develop bespoke bioinformatics pipelines

for detecting species-specific abundance, expression and epigenetic modifications (methylation) of N-cycle genes from long-read DNA- and RNA-sequencing using data already available via online databases.

The pipelines you develop will be applied to study soil samples you will collect from field experiments on the University of Reading farm. These will include experiments established by LEGUMINOSE project, which studies new approaches to regenerative agriculture where cereal and (nitrogen fixing) legume crops are grown alongside each other in intercrop mixtures, and different plots receive different rates of inorganic N-fertilisers. At the moment, it is planned to use PacBio CCS (Revio) for both shot-gun DNA and RNA sequencing of 12-24 samples from the LEGUMINOSE field experiment, which would allow you to evaluate the genes involved in all phases of the N-cycle. However, depending on the preliminary results, the project may focus on targeted sequencing of selected denitrification genes.

After being validated using samples from the LEGUMINOSE project, the pipelines and resources you develop could be applied to study nitrogen cycle genes in soil in other contexts too, thus ensuring the project has impact.

Training opportunities:

Depending on your scientific background, your training will focus on bioinformatics or soil science.

Training in applied bioinformatics (statistics, metagenomics), programming (R, Python, Shell), IT skills (including High Performance Computing cluster), soil science and soil biogeochemistry will be provided by the project supervisors and through participation in teaching and research offered at Cranfield University (e.g. selected modules from the Cranfield Applied Bioinformatics course) and in the University of Reading (e.g. Quantitative Analysis of Environmental Data course), and by attending other relevant courses available at the time of the project (e.g. the Statistics Summer School in Cranfield). In addition, the student will elevate their math and computer programming skills by enrolling on the BBSRC SysMIC course offered by UCL.

The student will become an early career member of the British Society of Soil Science, enabling them to attend early career conferences and join a community of soil scientists at a similar career stage. This also allows access to several training opportunities such as the 'Practical Introduction to Soils' and 'Working with Soil' courses, alongside webinars and other events.

Student profile:

A student should have a degree relevant to (bio) informatics or soil sciences. The minimum entry degree is a BSc honours at Upper second class level (or equivalent). The background and/or experience in bioinformatics/metagenomics is recommended although not mandatory. Experience with sequencing data analysis (ideally: long reads sequencing) is desirable.

Background in soil biology and/or microbiology would also be appropriate, if the candidate shows interest and abilities for developing the required bioinformatics skills.

Stipend (Salary):

FoodBioSystems DTP students receive an annual tax free stipend (salary) that is paid in instalments throughout the year. For 2023/24 this is £18,622 and it will increase slightly each year at rate set by UKRI.

Equality Diversity and Inclusion:

The FoodBioSystems DTP is committed to equality, diversity and inclusion (EDI), to building a doctoral researcher(DR) and staff body that reflects the diversity of society, and to encourage applications from under-represented and disadvantaged groups. Our actions to promote diversity and inclusion are detailed on the FoodBioSystems DTP website.

In accordance with UKRI guidelines, our studentships are offered on a part time basis in addition to full time registration. The minimum registration is 50% FT and the studentship end date will be extended to reflect the part-time registration.

References:

Selected references about Nitrogen Cycle in soil:
Kuypers 2018: https://www.nature.com/articles/nrmicro.2018.9
Lehnert 2018: https://www.nature.com/articles/s41570-018-0041-7
Grzyb 2021: https://doi.org/10.3390/agronomy11071415
Sun 2022: https://www.sciencedirect.com/science/article/pii/S004896972202664X
Ouyang 2018: https://www.sciencedirect.com/science/article/abs/pii/S0038071718302864
Blow 2016: https://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1005854
Selected references about long-read sequencing technology for soil microbial analysis:
Albertsen 2023: https://www.nature.com/articles/s41592-022-01726-6
Amarasinghe 2020: https://genomebiology.biomedcentral.com/articles/10.1186/s13059-020-1935-5
Callahan 2019: https://academic.oup.com/nar/article/47/18/e103/5527971
https://www.pacb.com/products-and-services/applications/complex-populations/microbial
https://www.pacb.com/blog/shotgun-metagenomics-with-hifi-reads-now-more-affordable
https://nanoporetech.com/applications/techniques/metagenomics-and-microbiome-analysis-with-nanopore-technology
Examples of microbial genomics resources that may be used in the project:
https://www.ncbi.nlm.nih.gov/genome/microbes/
Gaby 2014: https://academic.oup.com/database/article/doi/10.1093/database/bau001/2633766
https://blogs.cornell.edu/buckley/nifh-sequence-database
Parks 2021: https://academic.oup.com/nar/article/50/D1/D785/6370255
https://gtdb.ecogenomic.org

For up to date information on funding eligibility, studentship rates and part time registration, please visit the FoodBioSystems website.