PhD Project Advertisement

Project title: New software to detect horizontal gene transfer in microbiomes: from forage to the rumen
Project No: FBS2023-58-Swain-aq
Lead supervisor: Martin Swain, Life Sciences, Aberystwyth University
Email: mts11@aber.ac.uk
Co-supervisors:
Chris Creevey, Queen's University Belfast; Kerrie Farrar, Aberystwyth University

Project description:

Horizontal gene transfer or HGT is a key process that helps organisms adapt to new environments. HGT involves the sharing of genetic material between organisms that are not in a vertical relationship of inheritance, such as parents and offspring. For instance, HGT allows diverse species of bacteria to exchange genetic materials with each other, allowing them to develop resistance to antibiotics.

While HGT has been studied in bacteria for some time, until recently it was thought to be unimportant for the evolution of higher organisms. Recent advances in genome sequencing, however, have resulted in large genome collections of plants and animals, and the analysis of these data sets is changing how we see the role of HGT in higher organisms. Increasing evidence is being collected to show that HGT can be very important for understanding evolution across the plant, animal and microbial kingdoms.

There is a need to develop better computational methods to detect HGT, especially in higher organisms. Approaches to detecting HGT events can be divided into two. The first involves the construction of phylogenetic trees using sequence alignment, to detect differences between the evolutionary history of a species and the evolutionary history of individual gene sequences. The second approach involves the analysis of properties intrinsic to genome sequences i.e. the patterns of order, disorder, entropy and information content. These approaches can be described as alignment-free sequence analysis. We have recently published an article that demonstrates how alignment-free methods can be improved through calibration using training data sets, and this approach can be used to improve software for HGT detection. Using this methodology, the main aim of the PhD project will be to develop novel software for detecting horizontal gene transfer that may be applied to existing collections of genome sequence. We expect that our novel process of calibrating alignment-free methods will be instrumental in giving significant performance gains over existing approaches.

Typically HGTs occur between species that live in biological promiscuity, such as parasites, pathogens (including viruses), symbionts, and their plant or animal hosts. Through the analysis of HGT in bacterial genomes, we will gain insight into the genetic structure of bacteria known as endophytes that live within plants, including forage crops such as the grasses. In particular we will study their acquisition (or loss) of genes that are fundamental to this symbiotic relationship. It is becoming increasingly clear that plant-endophyte associations have important roles in maintaining ecological balance. In addition they have potential for biotechnological applications. Endophyte relationships have been engineered to improve phytoremediation i.e. the use of plants to clean up contaminated soil, air, or water; or to introduce novel traits into crops so that they can thrive in challenging environments like salt-rich soils.

We also plan to apply our software to HGT in eukaryotic systems. Eukaryotes are organisms with a nucleus in their cells and include microbes such as ciliates, which are found in the digestive system of sheep and cattle. Ciliates are interesting organisms to study with regard to the invasion and defence against transposable elements. Transposable elements or "jumping genes" are DNA sequences that have the ability to change their position within a genome, and horizontal transposable element transfer (or HTT) refers to the exchange of transposable elements between different species. An interesting challenge will be to apply our computational methodology to HTT.

Our improved approach to HGT and HTT detection will generate novel insights into systems where these processes are hypothesised to play an important role. This will help to better understand the developmental mechanisms for symbiosis and pathogenicity, two important types of relationship microorganisms can have with forage crops and the livestock that feed on those crops.

Training opportunities:
The project is within the area of genomic technologies, including data science and bioinformatics, and involves interdisciplinary approaches. The student will develop agricultural informatics approaches, which will develop their knowledge of the management of crops, farmed animals and their microbiomes.

We appreciate that some capable students will lack experience in aspects of these skills, and so we will provide extensive training in computing skills and bioinformatics. This will include genome assembly and analysis, alignment and alignment-free approaches to sequence analysis, phylogenetic tree building, programming skills in R and Python, and the use of Linux. The project will make extensive use of high performance computing clusters.

Budget will be available for the student to attend at least one relevant Summer School to develop their computational skills. They will also be encouraged to develop communication skills through conference presentations, and postgraduate teaching opportunities, which can lead to Associate Fellow status in the Higher Education Academy.

The bioinformatics software to be developed is based on pattern recognition and statistical analysis – and is therefore fundamental to machine learning and artificial intelligence research. The project will result in a valuable set of transferable computational skills for data-intensive life sciences, highly sought after by employers.

Student profile:
The project would be suitable for students with genuine interests in performing cross-disciplinary research at the interface between the computational and life sciences. We are open to applicants from diverse backgrounds, including those with strong academic records in a numerate, biological, agricultural or otherwise relevant science, and with a clear motivation in developing the expertise required to fulfil this exciting research program. The ability to learn new skills quickly would be highly advantageous as well as a demonstrable interest in data intensive science and software development.

Stipend (Salary):
FoodBioSystems DTP students receive an annual tax free stipend (salary) that is paid in instalments throughout the year. For 2022/23 this will be £17,668 and this 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:

- Swain, M. T., & Vickers, M. (2022). Interpreting alignment-free sequence comparison: what makes a score a good score? NAR genomics and bioinformatics, 4(3), lqac062.
- Van Etten, J., & Bhattacharya, D. (2020). Horizontal gene transfer in eukaryotes: not if, but how much?. Trends in Genetics, 36(12), 915-925.
- Tiwari, P., & Bae, H. (2020). Horizontal gene transfer and endophytes: An implication for the acquisition of novel traits. Plants, 9(3), 305.
- Aubin, E., El Baidouri, M., & Panaud, O. (2021). Horizontal gene transfers in plants. Life, 11(8), 857
- Allen, S. E., & Nowacki, M. (2017). Necessity is the mother of invention: ciliates, transposons, and transgenerational inheritance. Trends in genetics, 33(3), 197-207.

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