

## Exploring rice gene expression during stress conditions

### De Vega Group

Abiotic and biotic factors such as heat, drought and salinity contribute a major proportion of crop yield losses. They pose a serious threat to global food security and any attempts to eradicate poverty to achieve fair, sustainable societies. Rice is the most important cereal crop and staple food eaten by more than half of the world's population in every meal, particularly in Asia. Stress tolerance traits are multigenic and the candidate genes are poorly known.

The majority of transcriptome sequencing (RNA-seq) expression studies in plants remain underutilised and inaccessible due to the use of disparate references and the lack of skills and resources to analyse and visualise these data. Our group at the Earlham Institute has reanalysed over 70 RNA-seq experiments (from more than 2,000 samples) in different abiotic and biotic stress conditions in rice cultivars using common criteria and pipelines in order to visualise the result in a holistic way.

The student will be trained in R and will carry out comprehensive exploration of this large expression dataset, compare data and sort the information based on the associated literature in order to describe how different factors influence the observations and relate to each other. The student will also design a decision tree to score and identify rice genes potentially associated with adaptations to individual or combined stress in the dataset. Interest in basic plant physiology and data analysis is desirable.

This project will provide training in data science and bioinformatics. These are areas for which we experience a shortage of skills and can directly have an impact on the agricultural industry and the global need for food security.