**RRES Press Release 22nd April 2025** [**New Analysis Sheds Light on Specific Genes Found in All Grasses**](https://www.rothamsted.ac.uk/news/new-analysis-sheds-light-specific-genes-found-all-grasses)

*Research identifies over 13,000 groups of genes highly conserved in all grasses*

A new computational biology pipeline has mapped out over 13,000 groups of protein-coding genes conserved across grasses, offering a powerful tool for researchers investigating gene function in these economically and ecologically vital species.

Drawing on genomic data from 16 fully sequenced grass species within [Ensembl Plants](https://plants.ensembl.org/index.html), the study identified 13,312 highly conserved "universal" groups of grass genes. These gene groups  are present across all studied grasses and the genes within groups are highly similar suggesting they are responsible for vital functions in all grasses.

Crucially, the pipeline’s findings held up under scrutiny: 98.8% of these groups were also detected in newly sequenced genomes from two major grass groups (clades) not included in the original analysis, underscoring the robustness of the approach.

I hope the research community will find it useful to accelerate progress in grass genetics, including efforts to improve yield, stress resistance, and nutrient use in cereals like rice, wheat, and maize.

The study also identified 4,609 gene groups likely involved in functions specific to monocots, commelinids, or grasses — a significant step toward untangling the evolution of traits that led to the evolutionary success of the grasses.

What sets this study apart is its use of a statistical technique known as the  [Hidden Markov Model](https://en.wikipedia.org/wiki/Hidden_Markov_model) (HMM), which emphasises the conserved parts of genes which are important for function rather than the whole sequence. This technique outperformed a simpler approach based on percentage identity of sequence in distinguishing between known lineage-specific and non-specific genes.

Researchers working on gene discovery such as QTL analysis in grasses can now consult the newly released **[universal\_grass\_peps](https://data.rothamsted.ac.uk/it/dataset/universal_grass_peps)** database to determine whether their genes of interest are conserved across the grass family and whether they are potentially linked to lineage-specific adaptations.

“The database offers a new source of information for grass genes of unknown function, conveniently identifying those that are common to all grasses and how grass-specific their function is likely to be.” said the study’s author, Rothamsted’s Dr Rowan Mitchell. “I hope the research community will find it useful to accelerate progress in grass genetics, including efforts to improve yield, stress resistance, and nutrient use in cereals like rice, wheat, and maize.”

**Publication**

***Rowan A C Mitchell*, Identification of universal grass genes and estimates of their monocot- / commelinid- /grass-specificity, *Bioinformatics Advances*, 2025; vbaf079,**[**https://doi.org/10.1093/bioadv/vbaf079**](https://doi.org/10.1093/bioadv/vbaf079)