9th – Genomic Selection in Breeding
A hands-on short course in R

20th - 24th MAY 2024
Technical University of Madrid (UPM) , SPAIN

Instructors:
Dr. Julio Isidro Sanchez
Dr. Deniz Akdemir
Dr. Roberto Fritsche-Neto

Contents

1   Course Description. Why a course in Genomic Selection?.

The rapid decline in the costs of genotypic profiling per individual has led to a great explosion of many forms of data in the phenotypic-genomic spectrum in recent years in the forms of marker and sequence information, gene expression phenotypes, chromatin structure, RNA expression, protein, biochemical traits, physiological traits, metabolomics and microbiome data enriched with annotations; which, in turn, is leading to the development of new genetic tools that are transforming the way we understand biology as a whole.

Genomic-wide marker selection aka Genomic Selection (GS) is a new breeding approach that uses genome-wide molecular markers in prediction of breeding values for multi-genic quantitative traits. Many public and private breeding efforts have adopted GS as a new breeding tool.

The aim of this course is **to provide a basic quantitative and statistical framework to apply GS in a routine manner**. In this sense, the course is focusing on the application of plant breeding concepts **through practical exercises in R**. The course will provide participants with the relevant theory of GS models, as well as with hands-on experience with relevant GS techniques. The teaching assessment is a problem-based learning approach. Problems and solutions will be handed to the participants.
The learning outcomes upon course completion are:

- To have the fundamental knowledge to build GS models from the scratch.

- To have a basic understanding of the main statistical concepts of GS methods.

- To learn how to apply GxE models, Multi-trait analysis and Hybrid prediction.

- To learn about optimal parental contributions and mating designs for multi-criteria breeding.

A certificate of attendance will be provided at the end of the course. This course will be equivalent to 5 credits. For those participants who wish to obtain a certificate of attendance of the equivalence of five credit module, a series of reading and exercises will be provided before, during and after the course to assure the student achieves the learning outcomes.

2   Target Audience and Prerequisites

The **course is aimed at plant breeder scientists, graduate students, postdocs and professionals in the field of plant-crop production** who are interested to learn concepts in a problem-based learning approach.

2.1   Prerequisites

Although, the course is not aimed at researchers with advanced statistical skills. Participants should be familiar with plant breeding and statistical concepts. Course tutorials will use statistical packages in R, and experience is R is recommended but it is not essential.

3   Program

This is the tentative schedule of the course. Small changes might apply.

3.1   Day 1-May 20th (09:00 to 17:00).

Quantitative Genetics/R programming/

- Review of Quantitative genetics/ Quantitative trait loci

- Sources of quantitative trait variation

- Breeding Values and Heritabilities

- Response from selection

- Resemblance among relatives

- Pedigree vs Kinship matrix

3.2   Day 2-May 21st (9:00 to 17:00).

Introduction to Genomic Selection / Linear models/ BLUEs / BLUPs.

- Big Picture of Genomic selection/ Machine learning approach

- Factors affecting Genomic Selection.

- Optimization of GS.

- Linear Models/ANOVA/GLM

- Fixed-Random Effects

3.3   Day 3-May 22 nd (9:00 to 17:00).

Factors affecting GS/ optimization/Statistical concepts for Genomic Selection analysis in R

- Best Linear Unbiased Estimator

- Best Linear Unbiased Predictor

- Mixed Models

- Statistical Models

- Imputation

3.4   Day 4-May 23 th (9:00 to 17:00).

Onve vs Two step GS approach

- One-step model

- Two-Step models

- Cross-Validations

- Hybrid prediction

- Genotype × Environment interaction

3.5   Day 5-May 24 th (9:00 to 13:00).

Genotype × Environment interaction

- Genotype × Environment interaction

- Simulations.

- Future applications of GS

4   Costs

- Regular fee: €950

- Student fee: €550

These fees include, morning and afternoon breaks. Accommodations are not included.

5   Computers

Class notes will be distributed during the course, and a Dropbox/Google drive folder will be used to share R code, lectures and exercises. Please bring your personal. R and Rstudio should be installed in your computer to run the analysis.

6   Location

The tentative main hub for the course will be the Technical University of Madrid (UPM) at the Agronomy school. https://goo.gl/maps/F49jFi63iVM2tnHGA.

7 Online registration

Space is limited to 25 participants. All the information can be found at `www.gscourse.com`. **The registration link is at the bottom of the website**. Please click your option, and enter your details. Registration is complete when you pay the fee for the course. Deadline for registration is 20th April 2024.

8 Accommodations

Participants are responsible for their own accommodation. There are many hotel options in the area, please check here `https://www.google.com/maps/search/Hotels/@40.4419654,-3.7254641,16z/data=!4m4!2m3!5m2!5m1!1s2021-11-12`.

9 Instructors

Instructors have an extensive experience working on statistical genomics, quantitative genetics and plant breeding.
Julio Isidro y Sánchez: `https://scholar.google.com/citations?user=intwMKoAAAAJ&hl=en`
Deniz Akdemir: `https://scholar.google.com/citations?user=O_xha20AAAAJ&hl=en`
Roberto Fritsche-Neto: `https://scholar.google.com/citations?user=mQCVtVOAAAAJ&hl=en`

10 Contact

For further information do not hesitate to contact at. j.isidro@upm.es. Looking forward to seeing you in Madrid.

Note: Organizers could cancel the course due to unforeseen circumstances or if the number of participants in less than 10 (very unlikely).