Summer Institute in Plant Breeding

20–24 June 2022

Saint Paul, Minnesota, USA

Sponsored by the

Plant Breeding Center at the University of Minnesota

Modules, Instructors, and Schedule

The *2022 Summer Institute in Plant Breeding* will have six 2.5-day modules, with triplets of modules taught concurrently. Each participant can register in either one or two modules (one module from the Monday–Wednesday triplet and a second module from the Wednesday–Friday triplet).

Module	Instructor	Start and end
Plant Breeding 101	Dr. James A. Anderson (University of Minnesota)	Mon, June 20 at 8:00am to Wed, June 22 at noon
Data Bootcamp for Genomic Prediction in Plant Breeding	Dr. Aaron J. Lorenz (University of Minnesota)	
Machine Learning Applied to Plant Science	Dr. Tolutola Oyetunde (Takeda)	
Polyploid Genomics	Dr. Laura M. Shannon (University of Minnesota)	Wed, June 22 at 1:00pm to Friday, June 24 at 4:00pm
Applied Plant Genomics and Bioinformatics	Dr. Candice N. Hirsch (University of Minnesota)	
Genomewide Markers in Plant Breeding	Dr. Rex Bernardo (University of Minnesota)	

Target Audience

Plant Breeding 101 is intended for those who seek a course in the fundamentals of plant breeding (e.g., technicians, data scientists, managers, non-plant breeding graduate students, etc.).

The five other modules are intended for M.S. and Ph.D. students, postdocs, and industry scientists.

PLANT BREEDING CENTER

UNIVERSITY OF MINNESOTA
Driven to Discover®

Costs (per module)

$475 for graduate students, postdocs, and faculty
$625 for industry participants
These fees include breakfast, lunch, and morning and afternoon breaks. Dinners and accommodations are not included. We regret that scholarships or reduced fees are unavailable.

Online Registration

Please register via Eventbrite:

https://summerinstituteplantbreeding2022.eventbrite.com

Remember that registration is limited to **one module in the first set** (Monday to Wednesday noon) and **one module in the second set** (Wednesday afternoon to Friday).
Payment is by credit card only (VISA, MasterCard, Discover, or American Express).

If the Pandemic Worsens

The *Summer Institute* will switch to delivery via Zoom if the pandemic worsens (we certainly hope not) and travel and meeting in-person become unsafe. Should this happen, participants can be in only one module (rather than two). Participants can choose to:

- Cancel their entire registration, with a full refund of fees paid; or
- Cancel one in-person registration (with a full refund) and participate in one module delivered via Zoom (with a partial refund to reflect lower fees for a Zoom course).

The *Summer Institute* will not be responsible for any monetary losses that participants incur due to a switch to Zoom delivery. Please bear this in mind as you make travel arrangements. Travel insurance is suggested.

Accommodations

Participants are responsible for making their own lodging arrangements.

Questions?

Contact Rex Bernardo by email (bernardo@umn.edu).

Module Descriptions (see next page)

Plant Breeding 101 (Dr. Jim Anderson)

This course is for those who seek to learn the fundamentals of plant breeding (e.g., technicians, data scientists, managers, non-plant breeding graduate students, etc.). Each session will include lecture/discussion, demonstrations, and problem-solving with sample data.

- Session 1: Genetic variation, qualitative and quantitative inheritance, heritability
- Session 2: Phenotyping, experimental design, genotype x environment interaction, data analysis and summary
- Session 3: Modes of reproduction, types of cultivars, parent selection, breeding methodologies
- Session 4: Genetic mapping, breeding using major genes/markers, genomic prediction
- Session 5: Polyploidy, breeding for pest resistance, breeding asexually propagated species, intellectual property protection

Data Bootcamp for Genomic Prediction in Plant Breeding (Dr. Aaron Lorenz)

This course will cover common data structures, analysis techniques, and tools used for genomic selection in plant breeding. This course will include lectures and hands-on activities. (Prerequisite: basic knowledge of programming in R)

- Sessions 1-2: Formatting and quality control of genotype/phenotype data for genomic prediction, model implementation and exploration of various types of models, techniques in cross validation for assessing prediction accuracy
- Sessions 3-4: Training population optimization, multi-trait prediction, genomic prediction for genotype x environment interaction, predictions of all possible crosses
- Session 5: Other special topics in genomic prediction, tour of available tools for implementing genomic prediction

Machine Learning Applied to Plant Science (Dr. Tolutola Oyetunde)

This short course provides an introduction to core concepts in machine learning and data science. It will separate hype from fact, demystify machine learning, and critically examine prospects and limitations of data science as applied to plant science. Lectures, demonstrations, and hands-on exercises are designed to encourage the participants to think of ways of applying machine learning to solve practical problems in their current and future research. (Prerequisite: basic knowledge of programming in R)

- Session 1: Lecture - Introduction to machine learning; Lab - review of R programming and basic data analyses in R, laptop setup
- Session 2: Lecture - Fundamentals of machine learning; Code-along demo - a typical workflow for an end-to-end machine learning project; Lab - data cleaning and preprocessing
- Session 3: Lecture - Strengths and weaknesses of different machine learning models, evaluating machine algorithms, feature engineering; Lab - supervised machine learning
- Session 4: Lecture - Machine learning for quantitative genetics in plant breeding; Code-along demo - Machine learning for genomic selection; Lab - unsupervised machine learning
- Session 5: Lecture - Introduction to deep learning and reinforcement learning, computer vision in agriculture, machine learning for precision farming; Lab - short end-to-end machine learning project

Polyploid Genomics (Dr. Laura Shannon)

Polyploidy (and accompanying rediploidizaion) is a major component of plant evolution. All flowering plants have undergone at least one whole genome duplication event in their history. Furthermore, a wide variety of important crop plants are current polyploids. Despite the ubiquity of polyploidy, polyploid genomes are difficult to analyze. They complicate many of the tools and models most commonly used in genomics and breeding. This course will involve lectures, discussions, problem sets, and hands-on data analysis activities.

- Session 1: Polyploid evolution
- Sessions 2-3: Challenges presented by genotyping and sequencing polyploids; tools to overcome them; ways of determining ploidy
- Session 4-5: Population and quantitative genetic theory for autopolyploids; new tools for polyploidy research

Applied Plant Genomics and Bioinformatics (Dr. Candice Hirsch)

Principles, applications, and limitations regarding the development and use of genomics resources in plants. This course will include lectures, discussions, and hands-on activities. (Prerequisite: an undergraduate genetics course)

- Session 1: Review components of the plant genome, current sequencing technologies, and hands-on activity to access existing sequence data and assess the quality of sequence data
- Session 2: Methods and limitations for genome assembly, gene structural annotation, gene functional annotation, transposable element annotation, and hands-on resource allocation exercises
- Session 3: Genomic variation (types of variation, methods to measure variation, limitations in measurements), principles of pan-genomics and applications to crop improvement
- Session 4: Transcriptomics for estimating transcript abundance and transcriptome assembly, hands-on activity in differential expression
- Session 5: Principles of visualization for large-scale genomics data, hands-on activity to generate visualizations from large-scale genomics data

Genomewide Markers in Plant Breeding (Dr. Rex Bernardo)

Principles, concepts, and practices regarding the use of molecular markers to improve quantitative traits in plants. The course will include both theory and hands-on computer sessions in an active-learning format. (Prerequisites: a course in plant breeding and a course in statistics)

- Session 1: Review of plant breeding; overview of marker-assisted selection strategies; review of population and quantitative genetics; computer exercises on quantitative trait locus mapping
- Session 2: Linkage mapping in biparental populations; association mapping; computer exercises in linkage and association mapping
- Session 3: Marker-assisted selection for major QTL; F_2 enrichment; intro to genomewide selection
- Session 4: Theory and framework for genomewide prediction; factors affecting prediction accuracy; computer exercises in genomewide prediction
- Session 5: Best practices and challenges in integrating markers in a breeding program