



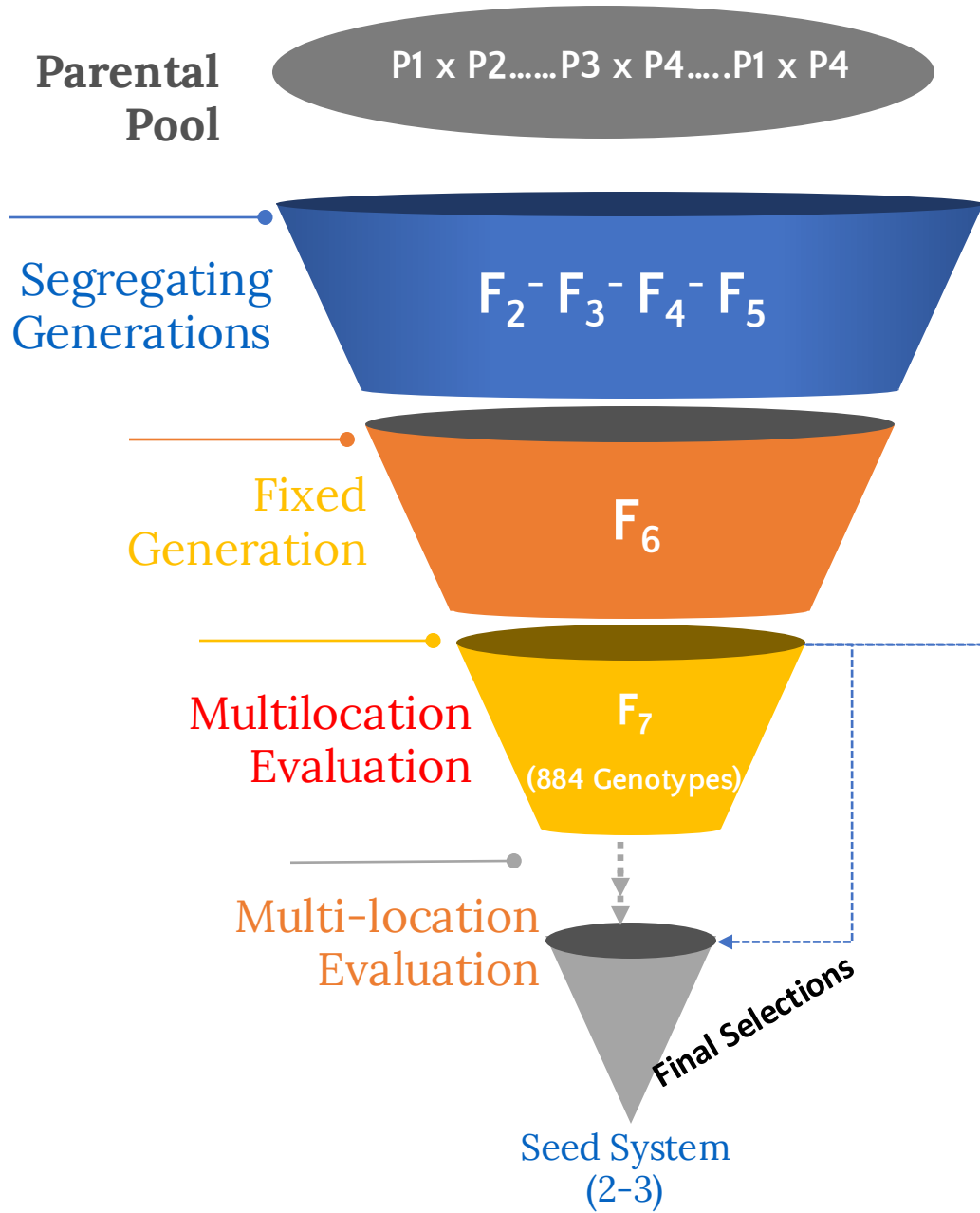
# Fundamentals of Genomic Prediction and Data-Driven Crop Breeding (August 4-8, 2025)

## Genomic Selection Workflow: Example and Implementation

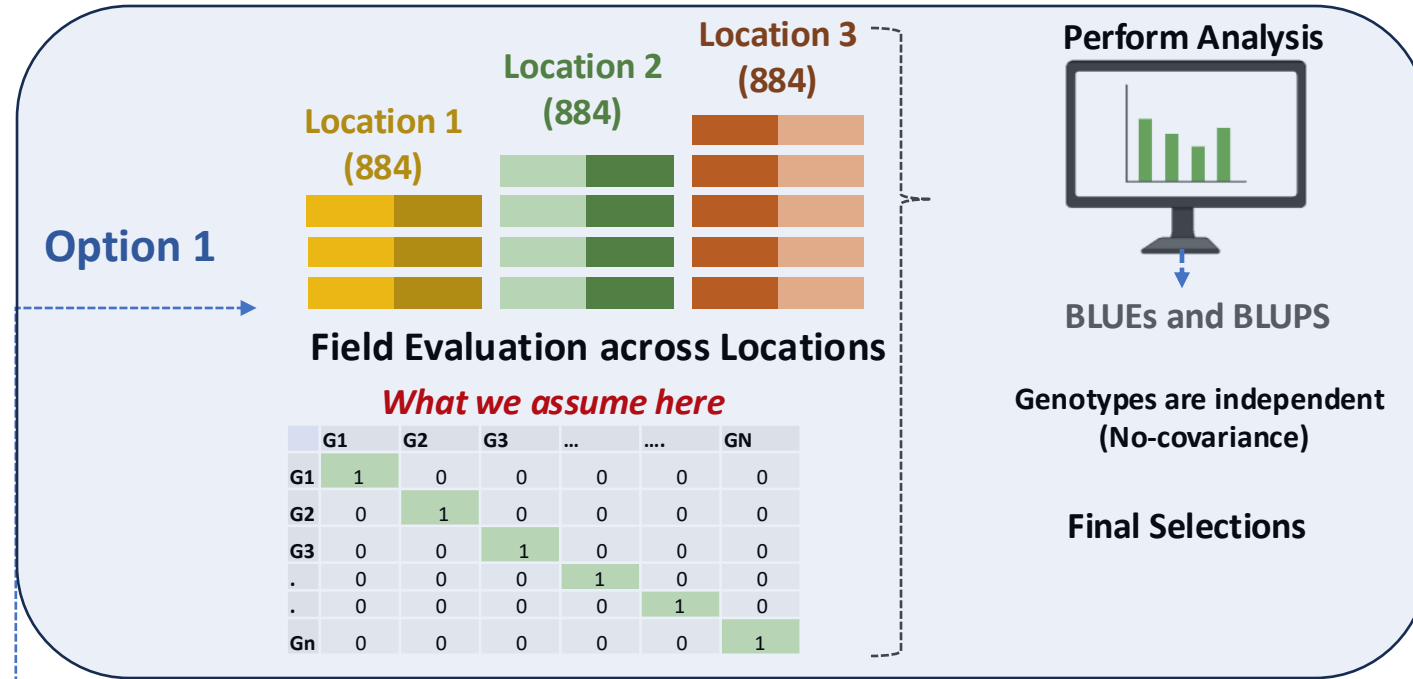
**Module 4**  
**August 6, 2025**

**Waseem Hussain and Mahender Anumalla**  
**Rice Breeding Innovations Platform**  
**IRRI**

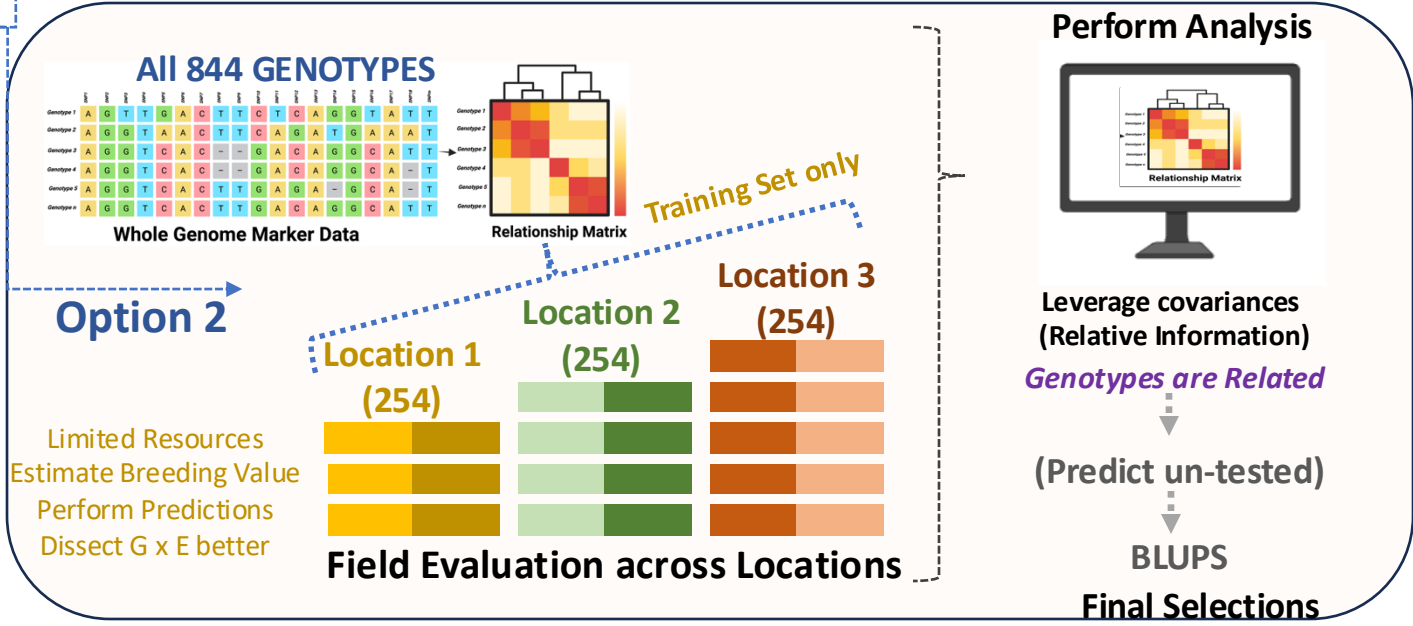
# General Breeding Pipeline



## Phenotypic Selection

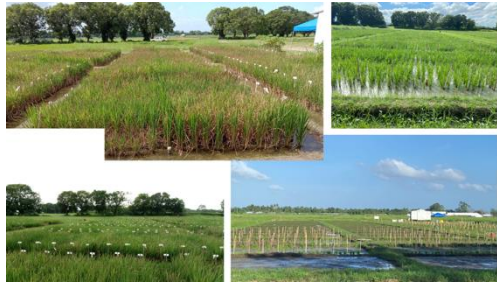


## Genomic Selection





# Steps of Implementing Genomic Selection



Whole F<sub>7</sub> Population



Genotype all Population

Step 1



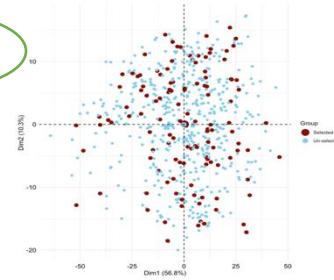
Un-tested Population

Pre-processing and Quality Check

- ✓ Calling rate (genotype and marker)
- ✓ Minor allele frequency
- ✓ Heterozygosity

Create a Training Population

Step 2



Phenotype Training Population

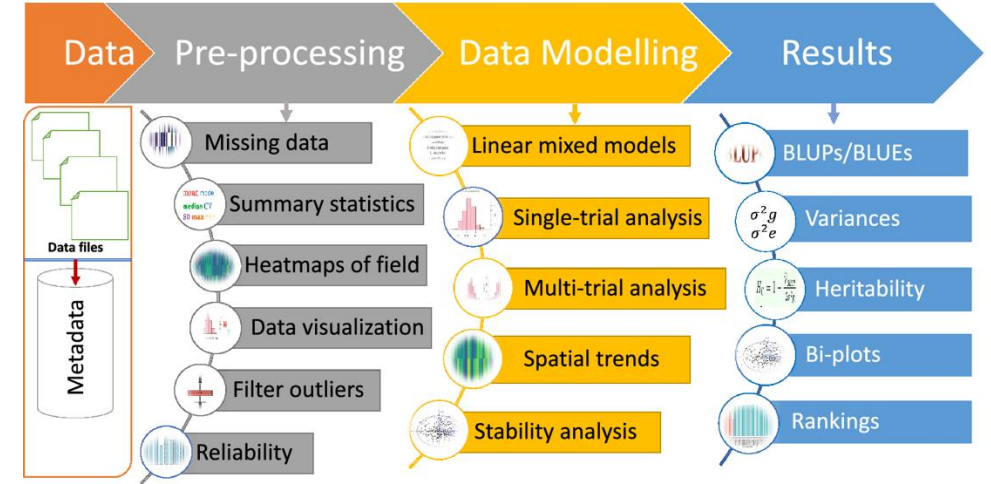
Step 3

Step 5

Prediction Equation

$$y = Xb + Zu + e$$

Predicted Values  
(BLUPs)



Pre-processing and Data Analysis Step 4

Step-Wise Analysis



Extract BLUES Per Environment

Model G x E  
also

# Genomic Selection Approach

## One Step Approach

$$y = X_1\beta_1 + X_2\beta_2 \dots + Z_1u_1 \dots \dots Z_2u_2 + \epsilon..$$

Fit all in one Model

(Design, Environment, and Genotype factors)

*Can We Extract BLUP in Step 1 and Fit BLUPS again in Step 2?*

<https://doi.org/10.1093/g3journal/jkae250>

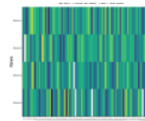
Double Shrinkage! Wrong Approach

## Two Step

### Stage 1: Extract BLUEs per Environment

General model

$$y_{ijk} = \mu + g_i + b_j + r_k + \epsilon_{ijk}$$



Row and Column

Modelling for Spatial variations

- ✓ Pre-process Data
- ✓ Trial and Data Quality check
- ✓ Filter outliers
- ✓ Correct for experimental design
- ✓ Spatial variation if coordinates available

### Stage 2: Extract the Breeding Values (BLUPs)

Model

$$B_{ij} = \mu + g_i + e_j + \epsilon_{ij}$$

$$g_i \sim N(0, A\sigma_g^2)$$

Where A is the relationship matrix

- ✓ Correcting for genetic relationship between lines
- ✓ Diverse lines based on Yield

# Example of IRRRI's Breeding Program

Let us run it in R together!