ISMAB: A data visualization and decision support tool for crop improvement

- Trushar Shah
Outline

- Introduction
- Functional Components
- MBDT Module
- Data Integration
- Demonstration
- Future Developments
Introduction

- Large amount of data generated in genomics and crop breeding

- As complexity and number of data points increases visual analytics provide crucial aids to decision making

- Visualization also aids in highlighting errors hence reducing the risk of failure at a later stage of the program

- Tracking and integrating data from multiple sources and of different types difficult to maintain in spreadsheets

- Data management and analysis tools being made available through the IBP
ISMAB Functional Components

- Molecular Breeding Design Tool (MBDT)
- Sample and marker tracking (LIMS)
- Molecular Selection Tool (MoSel)
- Data loading application
ISMAB Functionalities

- Track information throughout the Marker Assisted Breeding process.
- Identify ‘ideal’ parents for MAB program.
- Manage and visualize data collected over several segregating populations.
- Use the information from previous generations of the MAB program.
Marker-Assisted Breeding

**Visual selection of parents**

Donor parent \( \text{aa} \) x Recurrent parent (RP) \( \text{bb} \)

- \( F_1 \times RP \)
- \( BC_1 F_1 \times RP \)
- \( BC_2 F_1 \times RP \)
- \( BC_3 F_1 \times RP \)
- \( BC_4 F_1 \)
- \( BC_4 F_3 \) → \( BC_4 F_2 \)

**ISMAB Current Version**

**Creation of Target genotype**

M1 → QTL → M2
Use of the MBDT

- Selecting parental germplasm based on
  - presence or absence of traits in potential parents
  - level of polymorphism among the selected lines
  - analysis of background similarity via similarity matrix

- Create lists/input files for Laboratory Information Management System (LIMS)

- Design target genotype by dragging and dropping alleles from donor to target genotype

- Management and integration of information from previous generations of a backcross program
Samples from the field / lab for each Backcross

Load data sets in MBDT

Graphical Visualization

LIMS

Selection of Best individuals

1. Consistency
2. Number of QTLs
3. Flanking markers
4. Comapped QTLs
5. Polymorphism

Database

Visualization of Target Genotype (ultimate Product)

Product

MBDT Workflow
MBDT – Functions

- Read input files consisting of genotype data, map data, results of QTL mapping
- Display linkage and QTL maps

1. Accession List
2. Subset
3. QTL Maps
4. Linkage Map
5. Heat Maps
6. Creation of Target Genotype
7. Similarity Matrix
MBDT – Functions

- Display graphical genotypes of potential donors & recipients in the form of heat maps
- Choose foreground/background markers (possible regions for introgression from donor to recipient)
- Selection of flanking markers for multiple QTLs
- Sorting and similarity matrix for checking the compatibility between donor and recurrent parents
- Design target genotype from selected parents
- Integration of information from previous generations
MBDT Input Files

- **Data sources** - ICIS
- **GCP genotype data template** - data matrix sheet
- **iMAS output file** - linkage map and QTL
- **CMTV readable files** - linkage map
Data Integration

ICIS

External Databases

ISMAB

iMAS

Database

LIMS

Marker laboratory
Future Developments

- Allow display of consensus linkage maps and positioning of QTLs by marker name rather than map position
- Separate views for all the chromosomes in multiple windows
- Generation of target genotype from multiple parents
- Decision making – selection of parents, selection of markers, population size.
- Display of pedigree information
- Other MAB workflows eg: MARS and GWS
Availability

- The application is available as an open source software at http://cropforge.org/ISMAB

- Documentation being made available too.
Acknowledgements

- **Developers** – Jayashree B, Swapna N, Mallikarjun G, Sravani M, R Tuteja, Prathyusha C, A Rathore

- **Users** – D Hoisington, CT Hash, S Senthilvel, T Nepoleon, SP Deshpande, P Ramu, RK Varshney, O Riera-Lizarazu

- **Panel of Experts** – PK Gupta, HS Balyan, T Mohapatra, R Sundaram

- **Collaborators** – G Davenport, G McLaren

- **Funding** – Integrated Breeding Platform (IBP) and Department of Biotechnology (DBT-GOI)
Thank You!