Blueprint for Seedling Select: A Web-Based Rosaceae Marker-Assisted Seedling Selection Decision Support Tool



Introduction

Marker-assisted seedling selection (MASS) uses DNA marker information to assist conventional seedling selection (CSS), which is based only on phenotypic and pedigree information, in order to enhance its accuracy and efficiency (Collard et al., 2005, Peace and Norelli, 2009).

The routine application of MASS in Rosaceae breeding programs has been limited by a lack of accurate genomic information, access to high-throughput molecular biotechnology and software tools that assist in decision making. Here we report on development of *Seedling Select*, a web-based breeding decision-support tool being developed to assist MASS in Rosaceae breeding by helping in answering following questions:

• What is the most cost-efficient selection scheme, in what way and to

1.2a Cost-efficiency comparison model:

Cost-efficiency comparisons among different selection schemes are based on the comparison of costefficiency (CE) values of each scheme. The calculation of CE requires the knowing of total cost without MASS (CC), total cost with MASS (CM), genetic gain from selection without MASS (GC), and genetic gain from selection without MASS (GM). As an alternative, relative saving by MASS (RS) can also be used as the indicator of cost-efficiency. The less the CE (or the greater the RS), the more cost-efficient the scheme is. In the primary model of *Seedling Select*, cost-efficiency comparisons are conducted for cost comparisons which assume the genetic gains for schemes with and without MASS are the same.

> SM/GC CC/GM

Cost-efficiency comparison parameters:

$$CE = \frac{CC/GC}{CM/GM} \qquad RS =$$
Output
CE: cost-efficiency RS: relative saving by MASS

- what degree should MASS be adopted to make the selection scheme more accurate and efficient? – **Module 1**
- Which seedlings to save/cull? Module 2

MASS in a general breeding scheme



CC: total cost without MASS, CC = C (when PSGi = 0, 1 = < i <=SM) CM: total cost with MASS, CM = C (when at least one PSGi = 0, 1 = < i <=SMGC : genetic gain from selection without MASS GM: genetic gain from selection with MASS SM: total savings with MASS, SM = CM - CC*Note: It is assumed that GC = GM in the primary model

1.2b Cost evaluation model:

The calculation of CC, CM in the cost-efficiency comparison model (1.2a) is described in the following equation (**1.2b**). In this model, total monetary cost of a selection scheme can be calculated by summing up total genotyping cost, phenotyping cost and raising cost for the entire seedling selection process.

 $C = \sum_{i=1}^{SM} GCi \times PSGi + \sum_{i=1}^{SM} (RCM_i + PCM_i) \times PSM_i$

(**1.2b**)

WASHINGTON STATE

I JNIVERSITY

World Class. Face to Face.

Output

Input

C: total monetary cost of a selection scheme

Input SM: number of breeding stages in MASS GCi: genotyping cost for each plant in the

GCi: genotyping cost for each plant in the *i*th stage of MASS PSGi: genotyping population size in the *i*th stage of MASS RCMi: raising cost per plant in the *i*th stage of MASS PCMi: phenotyping cost per plant in the *i*th stage of MASS

PSMi: population size in the *i*th stage of MASS

Module 2: Selection decision support

To assist in making selection decisions, Module 2 will incorporate functionalities to calculate which seedlings to cull or keep based on DNA marker information and selection thresholds provided by the user. Recommendations as well as DNA marker information will be displayed in a user-friendly way to help users trace DNA marker information as well as review selection recommendations efficiently (Fig. 4.)

Blueprint for Seedling Select Module 1: Selection scheme design

In Module 1, both CSS seedling selection scheme and available DNA tests serve as background information for logistical MASS scheme design. Cost-efficiency values will be calculated for each scheme according to the cost-efficiency comparison model (**1.2a**). Comparisons of cost-efficiency between different schemes will be conducted for the software to provide recommendations for the most cost-efficient and logistically reasonable seedling selection schemes (**Fig. 3**).



•													
	1	2	3	4	5	6	7	8	9	10	11	12	
A	1110A01	1110A02	1110A03	1110A04	1110A05	1110A06	1110A07	1110A08	1110A09	1110A10	1110A11	1110A12	
	CULL	CULL	CULL	KEEP	KEEP	KEEP	CULL	KEEP	KEEP	CULL	CULL	CULL	
P	1110B01	1110B02	1110B03	1110B04	1110B05	1110B06	1110B07	1110B08	1110B09	1110B10	1110B11	1110B12	
D	CULL	CULL	KEEP	KEEP	KEEP	CULL	CULL	CULL	CULL	KEEP	CULL	CULL	
c	1110C01	1110C02	1110C03	1110C04	1110C05	1110C06	1110C07	1110C08	1110C09	1110C10	1110C11	1110C12	
	KEEP	KEEP	CULL	KEEP	CULL	CULL	KEEP	KEEP	KEEP	CULL	KEEP	KEEP	
D	1110D01	1110D02	1110D03	1110D04	1110D05	1110D06	1110D07	1110D08	1110D09	1110D10	1110D11	1110D12	
	KEEP	CULL	KEEP	KEEP	CULL	KEEP	KEEP	KEEP	KEEP	KEEP	CULL	KEEP	
E	1110E01	1110E02	1110E03	1110E04	Positive control	Negative control	1110E07	1110E08	1110E09	1110E10	1110E11	1110E12	
	CULL	CULL	KEEP	CULL			CULL	CULL	CULL	KEEP	CULL	CULL	
F	1110F01	1110F02	1110F03	1110F04	1110F05	1110F06	1110F07	1110F08	1110F09	1110F10	1110F11	1110F12	
	KEEP	KEEP	KEEP	CULL	KEEP	CULL	KEEP	KEEP	CULL	REVIEW	KEEP	KEEP	
G	1110G01	1110G02	1110G03	1110G04	1110G05	1110G06	1110G07	1110G08	1110G09	1110G10	1110G11	1110G12	
	KEEP	KEEP	CULL	KEEP	KEEP	KEEP	KEEP	KEEP	KEEP	CULL	REVIEW	CULL	
Η	1110H01	1110H02	1110H03	1110H04	1110H05	1110H06	1110H07	1110H08	1110H09	1110H10	1110H11	1110H12	
	CULL	REVIEW	KEEP	KEEP	KEEP	CULL	KEEP	CULL	REVIEW	REVIEW	KEEP	CULL	

Fig. 4. Selection decision results display in correspondence with locations on the genotyping plate. Green indicates keep, red cull, white further review.

Software development stages

Stage 1: Establish and test the cost-evaluation model for Module 1 based on MASS spreadsheet (Edge-Garza et al., 2010); develop software interface and functionality for cost evaluation

Stage 2: Develop selection decision support module

Stage 3: Establish and test efficiency evaluation model for Module 2; develop software interface and functionality for efficiency evaluation

Conclusions

RosBREED Enabling marker-assisted breeding in Rosaceae

Karley Contraction of the second seco

References

Collard, B., Jahufer, M., Brouwer, J. and Pang, E. (2005). An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. Euphytica 142: 169-196

Peace, C.P., Norelli, J. (2009). Genomics approaches to crop improvement in the Rosaceae. Ch 2 in: *Genetics and Genomics of Rosaceae* (Eds. K. Folta and S. Gardiner), Springer pp 19-53 Edge-Garza, D. A., Peace, C.P. and Zhu Y. (2010). Enabling marker-assisted seedling selection in the Washington Apple Breeding Program. Acta Horticulturae 859: 369-373

Seedling Select is a web-based breeding decision-support tool being developed to assist MASS in Rosaceae breeding. The aim of this software is to provide recommendations for the most cost-efficient selection scheme (focus of Module 1) and selection decisions made on seedlings (focus of Module 2).

Structures of each module are described in this poster. A primary model of cost evaluation for various selection schemes is constructed **1.2a**, **1.2b**). In this model, cost-efficiency is estimated by calculating the ratio of total cost without MASS to the total cost with MASS, where genetic gain from selection with and without MASS are assumed to be identical.

In the future, models for estimating genetic gain will be established for single trait selection, which will be further extended to multiple traits. In the meantime, integration of *Seedling Select* with other MASS decision-support tools will be addressed to collect information for more precise estimations.

Acknowledgements

This work was partially funded by USDA's National Institute of Food and Agriculture – Specialty Crop Research Initiative project, "RosBREED: Enabling marker-assisted breeding in Rosaceae" (2009-51181-05808).