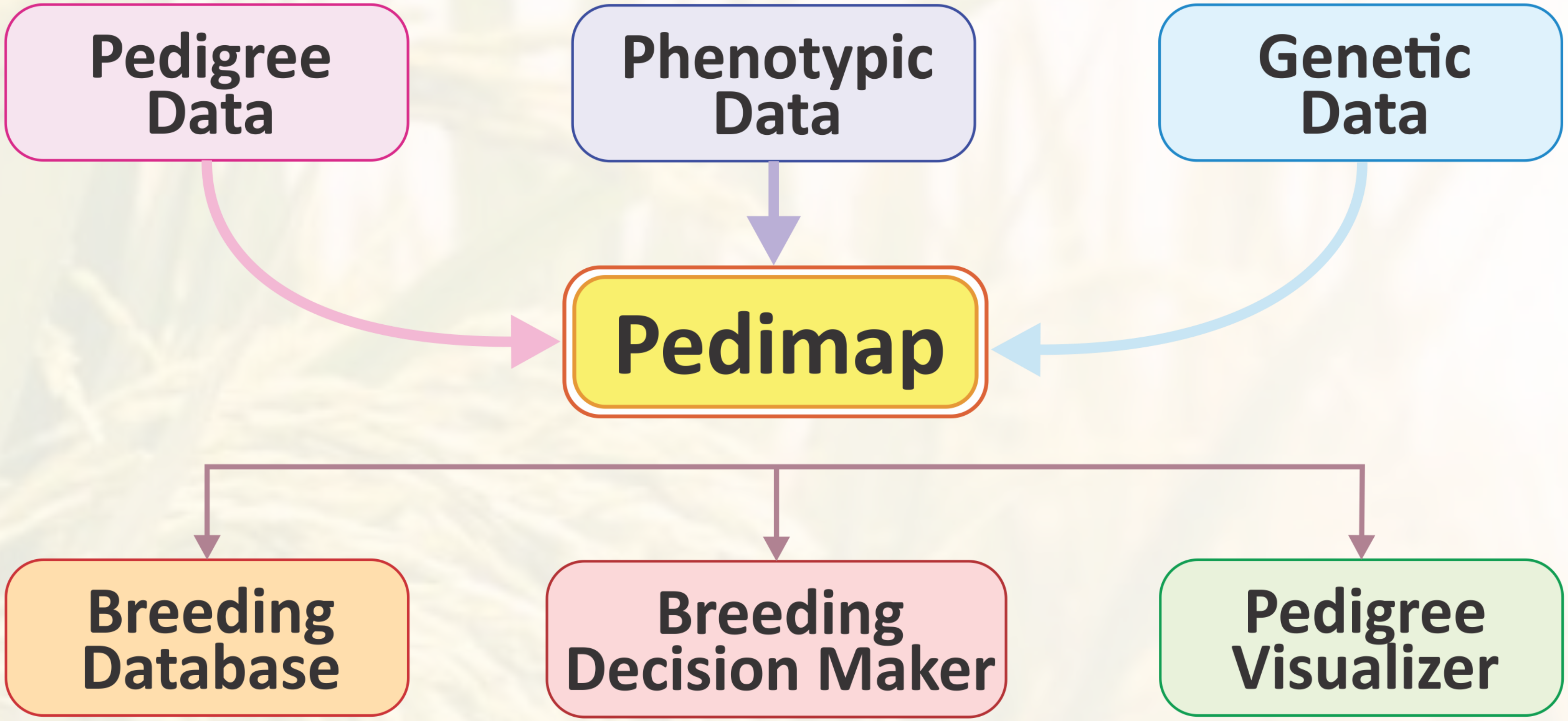
ORGANIZATION OF THE PHENOTYPIC AND GENETIC INFORMATION OF RICE BREEDING GERMPLASM IN SRI LANKA USING PEDIMAP TO FACILITATE THE DECISION-MAKING PROCESS IN VARIETAL IMPROVEMENT

P.G.R.G. Rathnayake (S/14/029)
Supervisor: Prof. S.D.S.S. Sooriyapathirana

Outline



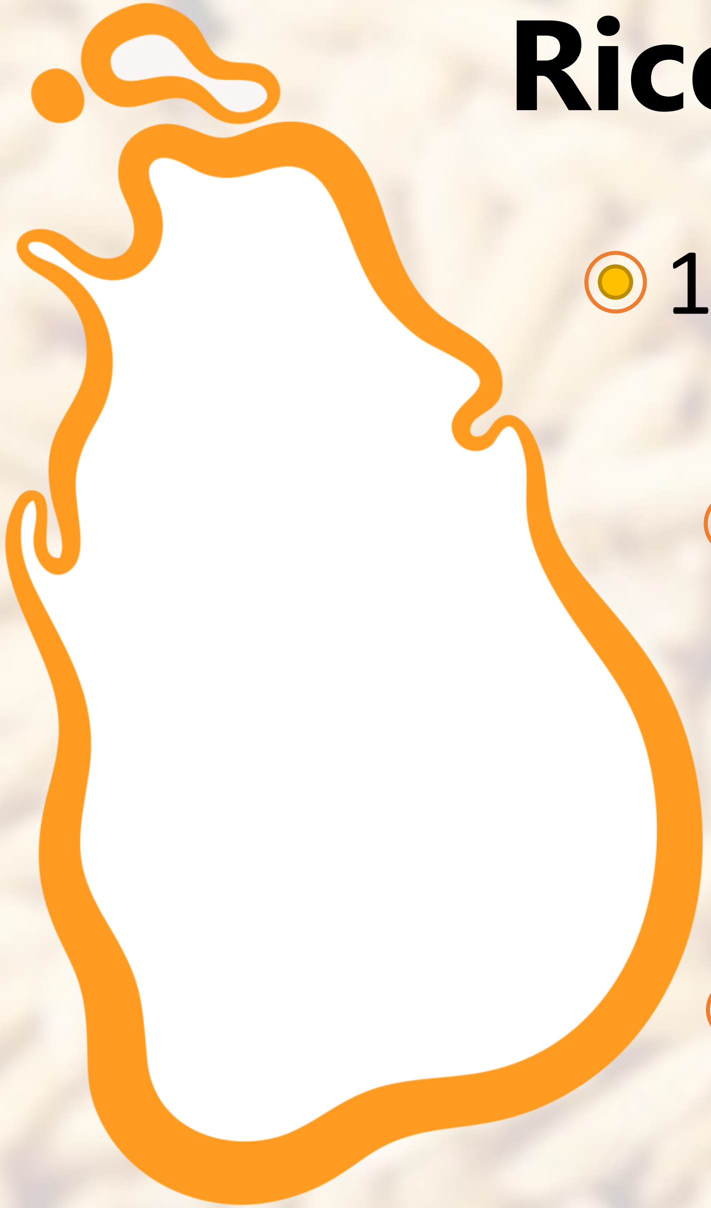
In Sri Lankan rice breeding programs

Introduction

- Half of the world consume RICE as staple food
- World annual rice production is over 700 Million metric tons
- Demand increasing

(Ali 2006; FAO, 2018; Muthayya et al. 2014)

Rice cultivation in Sri Lanka



- 1.8 Million farming families
- Annual demand is 3 million metric tons
- Production is 2.3 million metric tons
- USD 400 million import rice

(Central Bank, 2017, 2018; RRDI 2919)

Stresses on RICE cultivation

Biotic Stresses

Pest and diseases
Blast disease
Brown plant hopper
Bacteria leaf blight

Abiotic Stresses

Drought
Irregular rain fall
Climatic changes
Salinity

(RRDI 2019, Dharmarathna et al. 2014, Dhanapala 2007, Mackill 1996)

Breeding is crucial for high yield and stress resilient rice variety production

Varietal Improvement by breeding

Conventional Breeding methods

Molecular Breeding methods

Marker-assisted breeding | Marker- assisted selection

(Duvick 1984)

Most important step of a breeding program is,

Breeding Decision-Making

- Identify breeding priorities (Traits)
- **Parental varieties**
- **Selection methods**
- **Number of parents**
- Cycles of replication
- Economic feasibility
- Pre-breeding methods

Based on

Pedigree

Phenotypic

Genetic

(Acquaah 2012,Ragot et al. 2018)

The current practices on **parental selection and Selection method identification are,**

- Using haphazardly collected data in field note books
- Depends on breeder's desire
- Conducting in disorganized manner
- Lot of subjectivism

To organize all information in breeding decisioning needs data organization and visualization tool

It needs,

Organized Database + **Data Visualizer**

- **Data sharing**
- **Data mining**
- **Data retrieval**
- **Data Handling**

Pedimap Able to access the large pool of genetic and phenotypic data quickly and generate pedigrees

(Voorips *et al*. 2012)

Pedimap

- Data and pedigree visualization software
- Also can be used as database
- **RosBREED** project/ **HIDRAS** project previously used this

Can be used to visualize:

- Pedigree data/ Parentage
- Qualitative and quantitative data
- Marker alleles
- Calculate IBD (identity by descent) probabilities

(Peace et al. 2014; Rosyara et al, 2013; Evans et al, 2012; Voorips *et al*. 2012)

Objectives

- To organize the information of the released varieties and the parental genotypes of RRDI breeding programs as a Pedimap database: as a speedy breeding decision-making tool in Rice breeding programs
- To identify the DNA marker polymorphism to use in marker-assisted selections (MAS)

Materials and Methods

1. Collection of Plant materials
2. DNA Extraction
3. PCR and DNA marker analysis using *K29*, *Seq7-8* and *RM463*
4. Data curation
5. Pedimap Procedure
6. Illustration of the applicability of Pedimap in breeding decision-making process

1. Collection of Plant materials

- All available Landraces
- Traditional varieties
- Improved varieties
- Foreign varieties

2. DNA extraction

Using Dneasy® plant mini kit (Qiagen, Solna, Sweden)

3. PCR and DNA marker analysis

K29 ***Seq 7-8*** ***RM 463***

(Chin *et al.* 2010; Hu et al. 2015; Lu et al. 2012)

4. Data curation

- Parentage
- Qualitative and quantitative data
- Marker alleles
- Calculated IBD (identity by descent) probabilities

5. Pedimap Procedure

- A Pedimap input data file is created in MS Excel (2019). The input file contains four main subdivisions; header, pedigree, marker data, and IBD probability section

(Voorips 2007 and Voorips et al. 2012)

Pedimap procedure cont.

Pedimap in-put file structure

		A	B	C	D	E	F	G	H	I	J
A	1										
	2	POPULATION	=	Sri_Lanka_Rice_Germplasm							
	3	UNKNOWN	=	-							
	4	NULLHOMOZ	=	$							
	5	CONFIRMEDNULL	=	$$	i						
	6	PLOIDY	=	2							
	7	NALLELES	=	6	ii						
	8										
B	9	PEDIGREE					iv				
	10	NAME	PARENT1	PARENT2	Yiled	Maturity	Leaf_color	BPH	MG	BL	
	11	Bg941	-	-	-	-	-				
	12	Pokkali	-	- iii	-	-	-				
	13	At354	Bg941	Pokkali	6.5	95	Green				
	14	At401	Bg941	Pokkali	5	115	Green				
	15										
	16										

Figure 1 The input data file structure of the Pedimap database; The input file was created as an MS Excel worksheet, contains four main sections. A: Header, B: Pedigree, C: Marker data, D: IBD probabilities. A: Essential elements in header are highlighted in blue color (i) abbreviations for missing data (i.e., unknown). (ii) Total number of Null-alleles. B: The essentials are highlighted in orange. (iii) Missing values were accepted. (vi) Qualitative and quantitative data.

(Voorips 2007 and Voorips et al. 2012)

Pedimap procedure cont.

Figure 1 cont.

C	18	LINKAGEGROUP 12									
	19	MAP									
	20	RM101	48.2								
	21	RM277	62								
	22										
	23	LOCUS	RM101								
	24	ALLELENAMES	110	115	120	125					
	25	FOUNDERALLELES	110	110	110	110	115	120			
	26										
	27	LOCUS	RM277								
	28	ALLELENAMES	200	225	250	275	300				
	29	FOUNDERALLELES	200	200	200	200	300	275			
	30	IBDPOSITIONS	48.2	62							
	31										
	32	ALLELES	RM323								
	33	Bg941	140	0	140	0					
	34	Pokkali	160	0	180	0					
	35										
D	36	IBDPOSITION	48.6								
	37										
	38	Bg941	1	0	0	0		0	0	0	1
	39	Pokkali	0	0	1	0		0	0	0	0
	40	At354	0	0	0	0		0	0	1	0
	41	At401	0	0	0	0		0	0	0	0
	42										
	43	IBDPOSITION	62.0								

Annotations: (v) rows 23–25; (vi) arrow to row 30; (vii) rows 33–34; (viii) and (ix) row 37; Define Linkage group (rows 18–30).

Figure 1 cont. C: Marker data section contains linkage map information for all available markers and marker scores. (v) Each linkage group should be cited separately in ascending order. (vi) IBD probabilities for each locus (vii) The additional column for color specifications. D: IBD probabilities for each allele is entered separately. (viii) Maternal probability and (ix) paternal probability are entered and separated by a column. The final file must be saved as a tab-delimited text (.txt) file.

(Voorips 2007 and Voorips et al. 2012)

6. Illustration of the applicability of Pedimap in the breeding decision-making process

Example 1

Select parents for,

- White pericarp
- Yield > 3.5mt/ha
- BPH resistance
- Maturity of <135 days of
- Intermediate bold grains

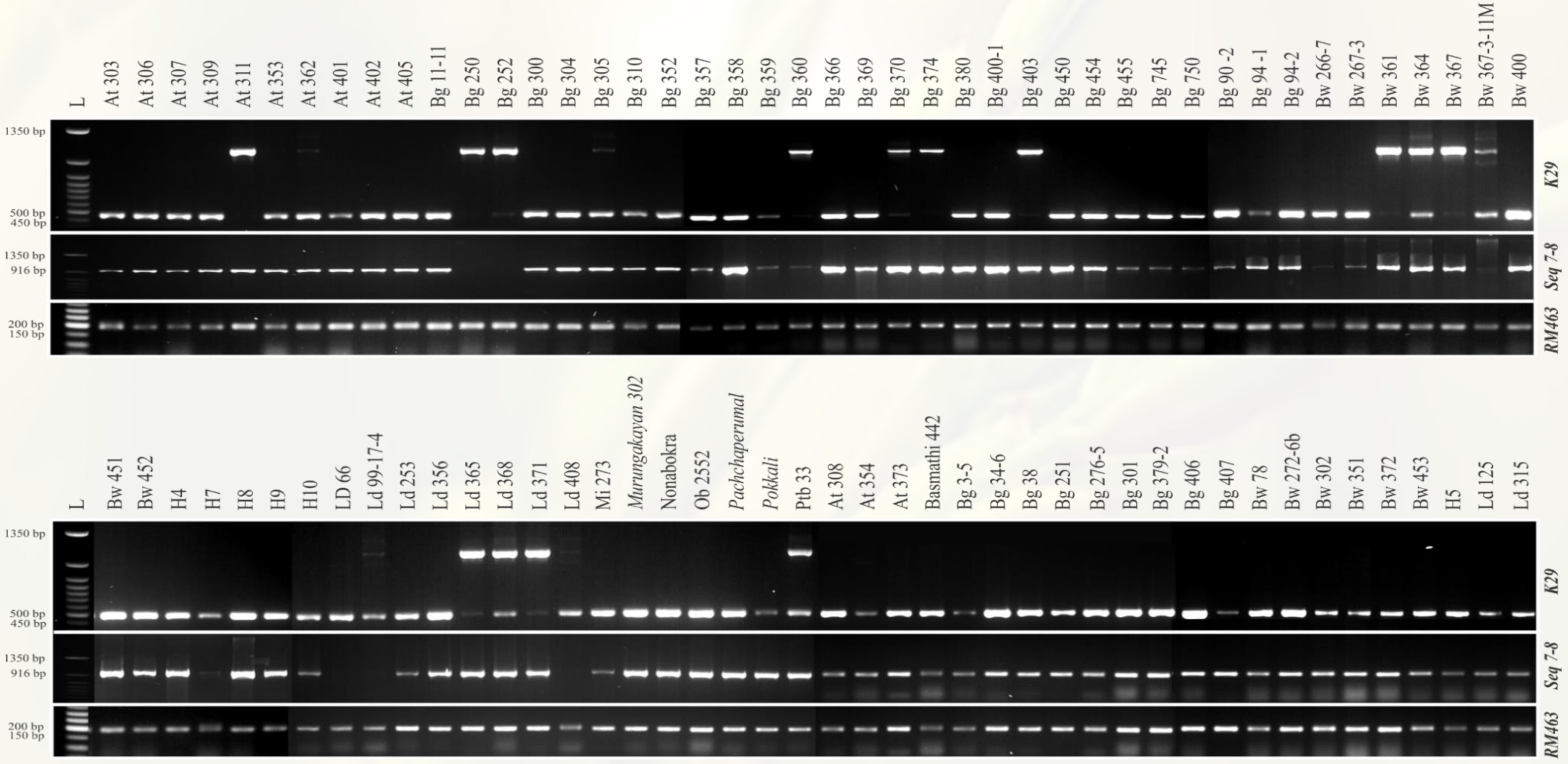
Example 2

Select parents and plan cross for,

- Phosphorus deficiency tolerant (PDT)
- Yield > 5mt/ha
- BPH and Blast resistance
- Maturity of 90 -150 days
- Higher amylose content

Results and Discussion

- DNA Marker analysis



***K29* and *Seq 7-8* shows marker polymorphism, while *RM463* is monomorphic for all the rice varieties**

Figure 2 The polymorphism of three co-dominant DNA markers; *K29*, *Seq 7-8,* and *RM463* in 90 rice cultivars. The band sizes are indicated at the left side of the Figure, and the DNA marker names are mentioned on the right side. The cultivar names are given at the top.

Pedimap procedure

- Data for all 86 RRDI released rice varieties were collected
- Pedimap input file was constructed with,

224 Inputs = 188 known Identities + 36 intermediate genotypes

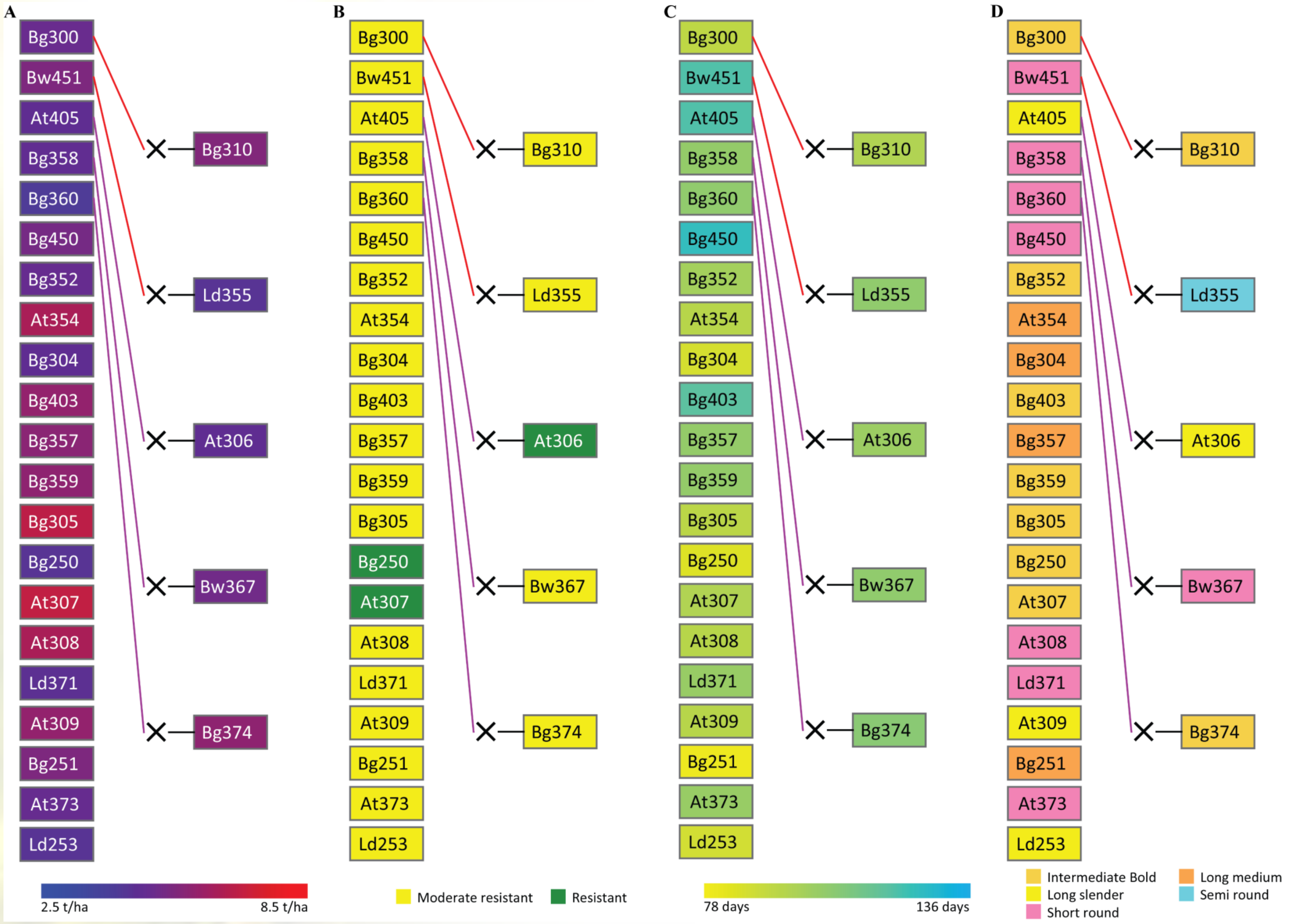
- Applicability of Pedimap in breeding decisioning is demonstrated using the example problems

Example 1

Parents for,

- **White pericarp**
- **high yield,**
- **less maturity period,**
- **BPH resistance**
- **intermediate bold grains**

Figure 3 The pedigree visualization for parental selection with white pericarp, yield ≥ 3.5 mt/ha, moderate or complete BPH resistance, maturity period ≤ 125 days, and diverse grain shapes. The selected pedigree is colored separately for four traits. A: Yield; B: Degree of resistance to brown planthopper (BPH); C: Maturity period; D: Grain shape.



Female and male parentages are indicated by red and purple lines, respectively. The symbol '×' indicates the cross between two parents. The background colors of the cultivar-name boxes indicate the trait values, as shown in the colored legends below.

Example 1 cont.

Parent varieties with White Pericarp

Higher yield (≥ 3.5 mt/ha)	BPH resistance	Maturity period (≤ 125 days)	Grain Shape (Intermediate bold)
At354	Bg255	Bg250	Bg250
Bg 305	At306	Bg251	Bg300
At307	At307	At305	Bg305
At 308		AT 305	Bg310
		At307	Bg252
		At308	Bg259
			Bg403
			Bg374
			At307

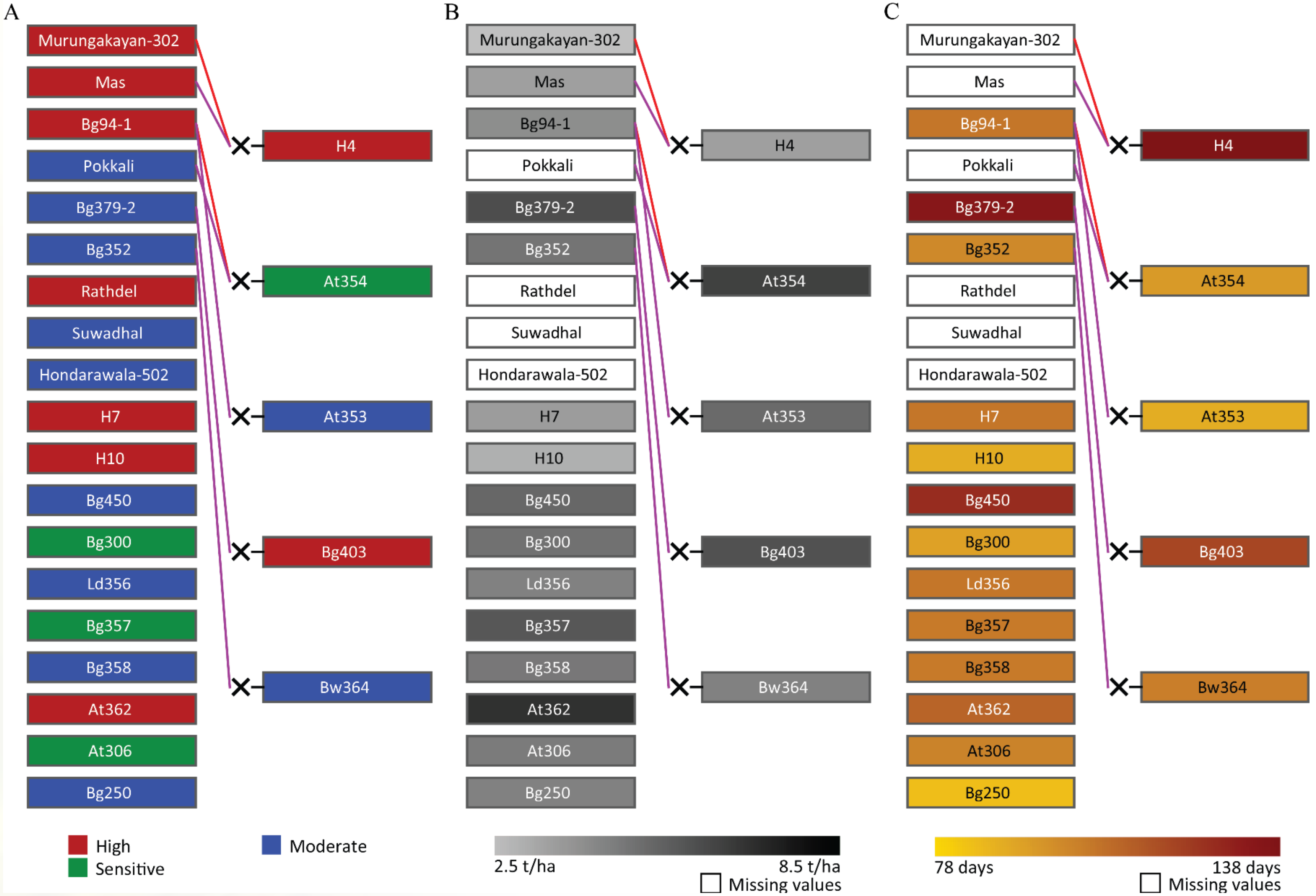
At307 can be used as the best parent for these breeding priorities

Example 2

Parents for,

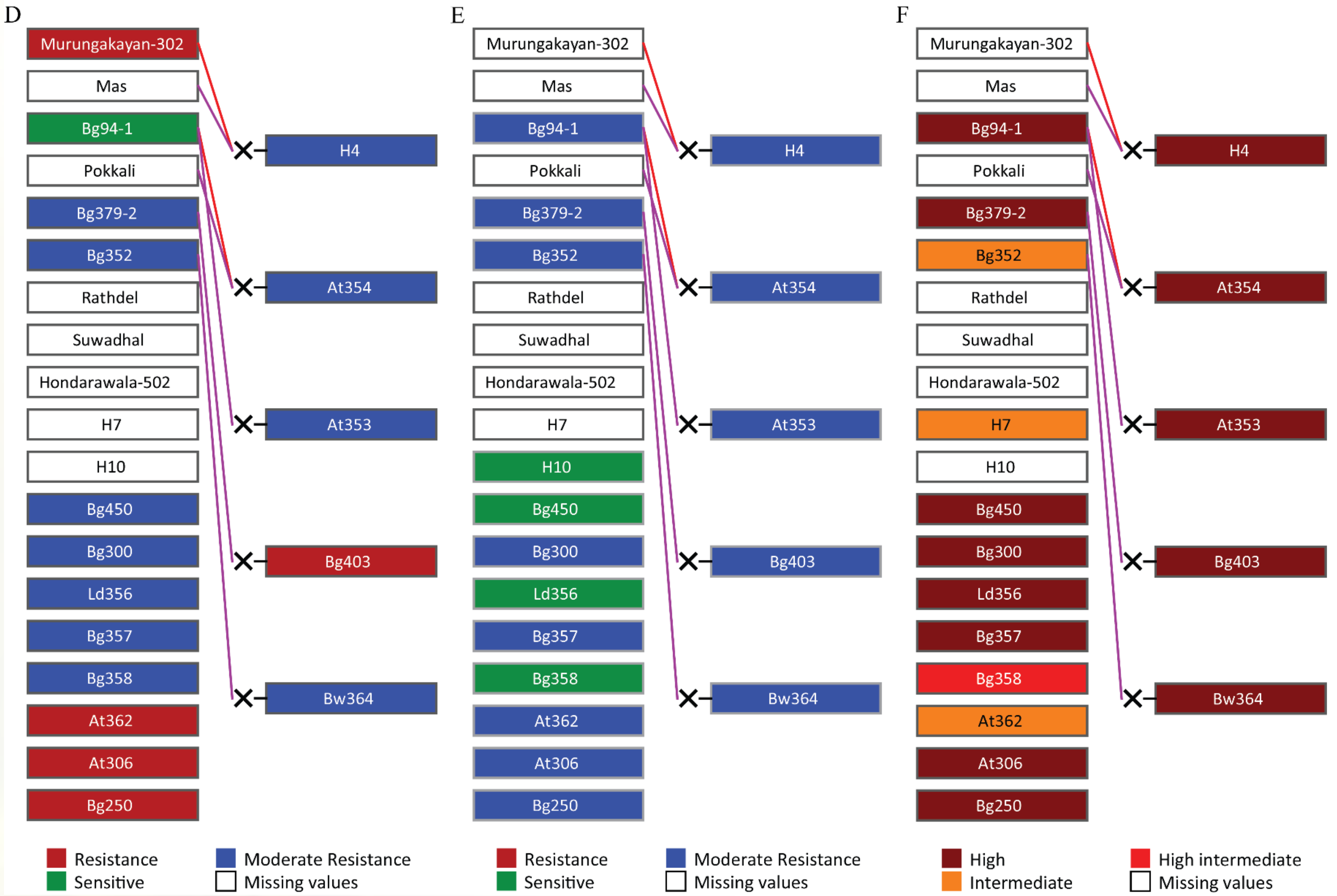
- **PDT**
- **high yield,**
- **less maturity period,**
- **BPH resistance**
- **Blast Resistance**
- **intermediate bold grains**

Figure 4 The pedigree visualization for select parent rice varieties with phosphorus deficiency tolerance, yield ≥ 5 mt/ha, moderate or complete BPH and Blast resistance, maturity period between 90-120 days, and grains with high amylose content). The selected pedigree is colored separately for six traits. A: PDT; B: Yield; C: Maturity period; D: Degree of resistance to BPH; E: Degree of resistance to BLAST; F: Amylose content.



Female and male parentages are indicated by red and purple lines, respectively. The symbol '×' indicates the cross between two parents. The background colors of the cultivar-name boxes indicate the trait values, as shown in the colored legends below.

Example 2 cont.



Female and male parentages are indicated by red and purple lines, respectively. The symbol '×' indicates the cross between two parents. The background colors of the cultivar-name boxes indicate the trait values, as shown in the colored legends below.

Figure 4 cont. The pedigree visualization for the selection of parent rice varieties with phosphorus deficiency tolerance, yield ≥ 5 mt/ha, moderate or complete BPH and Blast resistance, maturity period between 90-120 days, and grains with high amylose content). The selected pedigree is colored separately for six traits. A: PDT; B: Yield; C: Maturity period; D: Degree of resistance to BPH; E: Degree of resistance to BLAST; F: Amylose content.

Example 2 cont.

Among 24 selected varieties, high trait values were recorded for At362 and Bg250

At362	Bg250
High PDT	Moderate PDT
High yield of 7 t/ha	High yield of 4.5 t/ha
The maturity period of 110 days	The maturity period of 85 days
Resistant to BPH	Resistant to BPH
Moderate resistant to Blast	Moderate resistant to Blast
Intermediate level of Amylose content	High level of Amylose content

But still, the alleles for complete Blast resistance are not available with the progeny.

To introgress alleles for complete Blast resistant : Backcrossing with Bg252

Example 2 cont.

Phase 1:

Planning cross between At362 and Bg250.

Progeny contains expected traits except complete Blast resistance

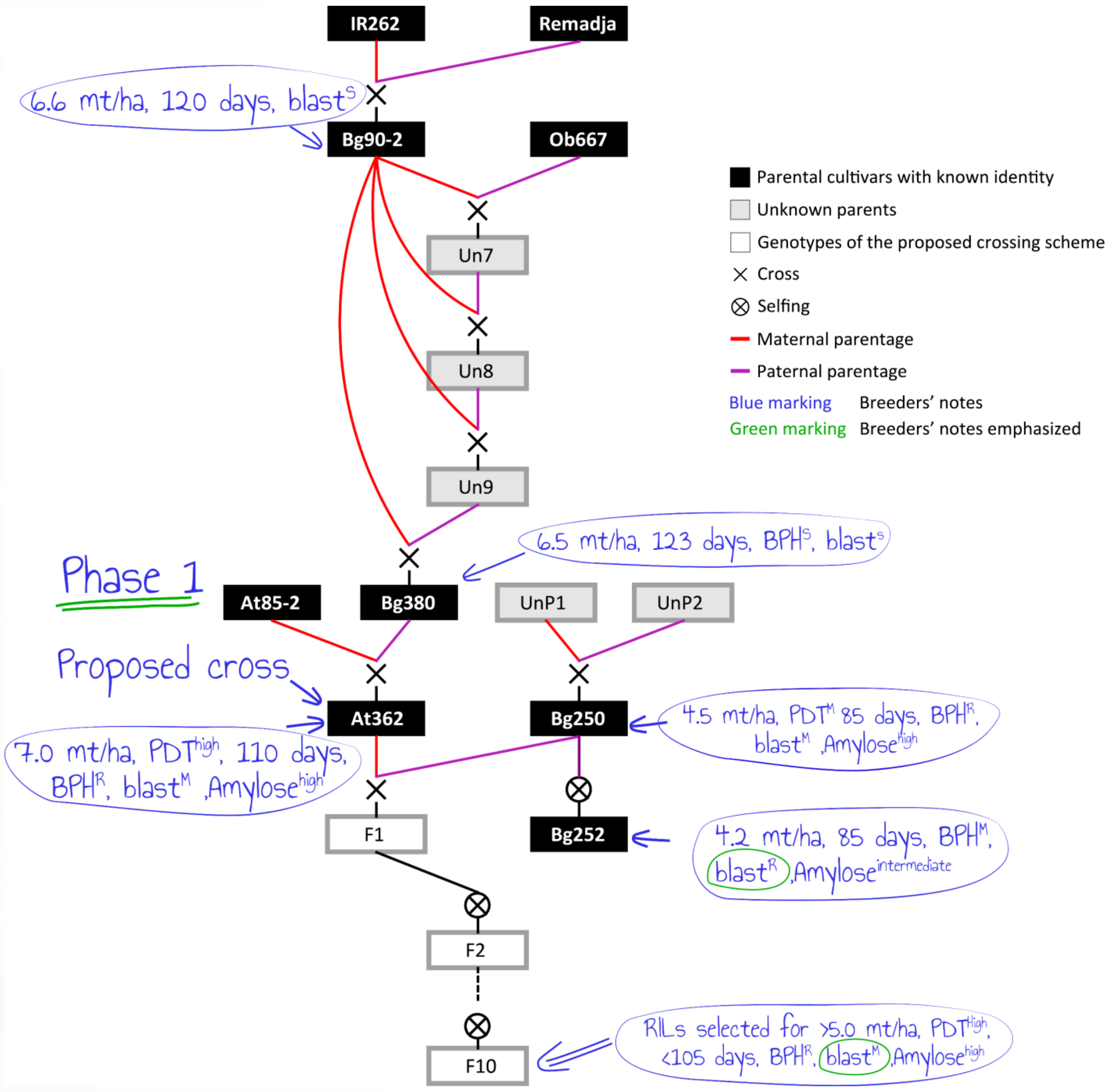


Figure 5 The pedigree visualization for planning a crossing scheme. Phase 1: Initial crossing of At362 and Bg250 and pedigree selection to obtain RILs with ≥5.0 mt/ha of mean yield, ≤105 day of maturity period, resistant to BPH, moderately resistant to blast and high level of amylose content

Example 2 cont.

Phase 2:

Cross between At362 x Bg250 progeny and Bg252.

The obtained progeny from phase 1 is subjected to backcrossing with Bg252 to introgress alleles for complete Blast resistance.

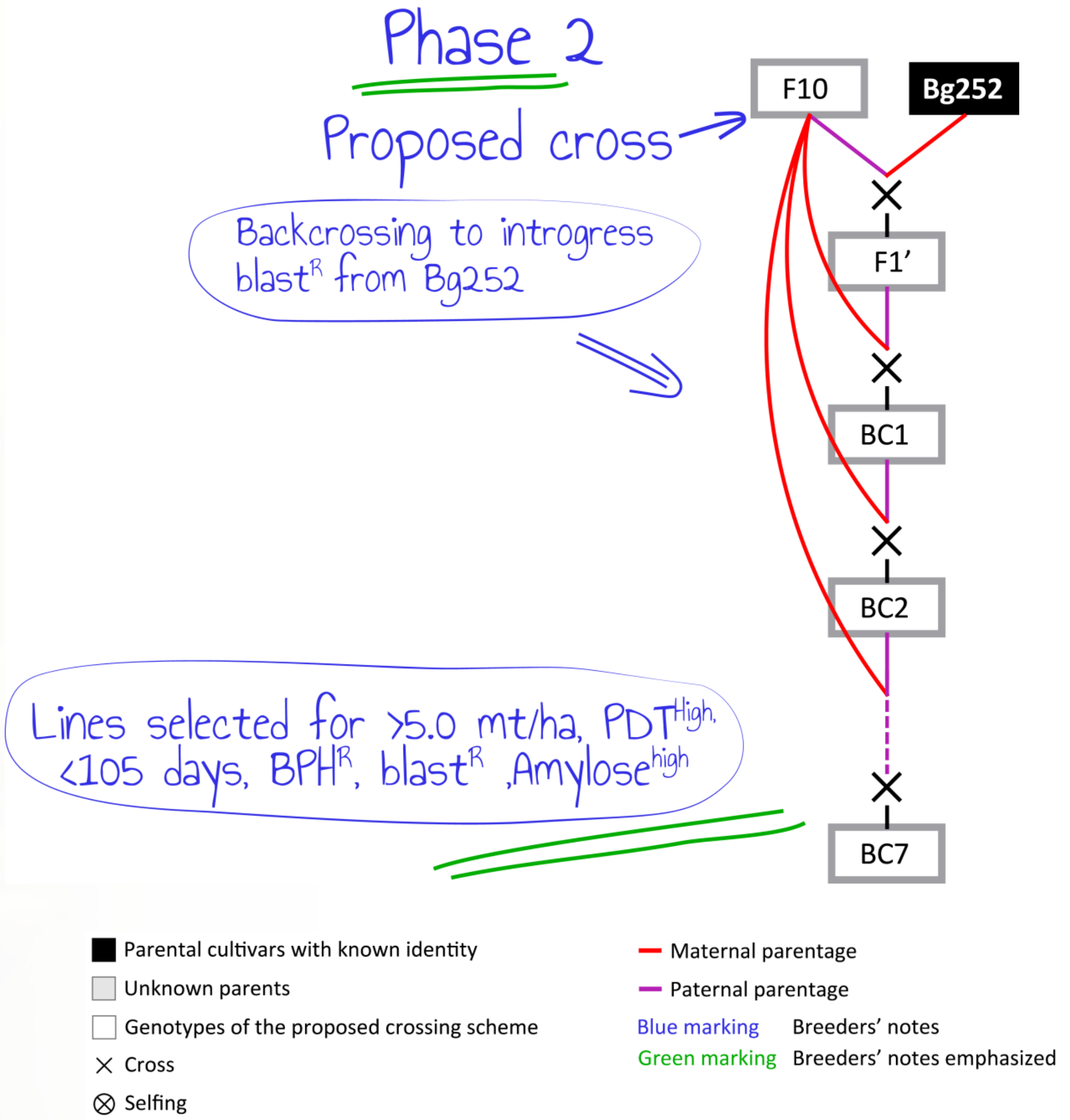


Figure 5 Cont. The pedigree visualization for planning a crossing scheme. Phase 2: Then backcrossing with Bg252 as the donor parent to introgress the blast resistance in to the phase 1 progeny

Example 2 cont.

IBD probabilities and DNA marker scores are also can be incorporated into the breeding decision making process.

IBD probabilities

The order of crossing and number of selfing cycles needed, can be predicted based on the rate of allelic segregation

DNA marker scores

These data values can be used to dentify genetic architecture of the cross and select suitable DNA markers for MAS

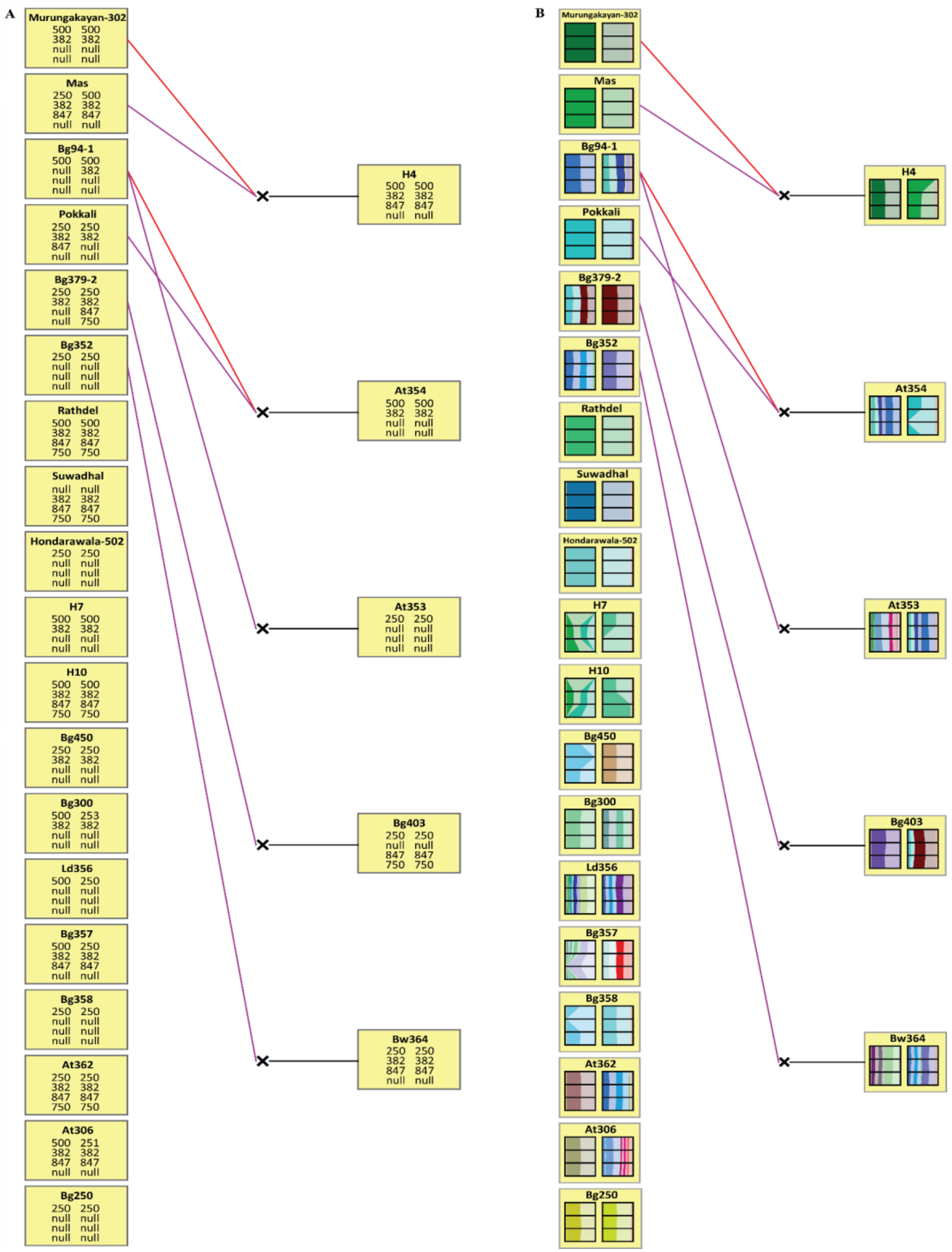


Figure 6 Visualization of selected marker genotypes and Identical by Descent (IBD). A: Marker alleles. The alleles of the DNA markers *K29-N*, *K41*, *K48*, and *K5-N* are given in vertical order.; B: IBD probabilities of four *Pup1* linked markers (on chromosome 12 at about 55 cM). Since the cultivar linkage maps are not available, we assumed 0.1 cM gap between adjacent markers for the representation of IBD values.

The results highlight that,

- Pedimap like tools are essential to work with a large mapping population
- Sequencing and SNP data also can also be include
- The successful, efficient and easy understandable substituent for highly complex tables for the breeders
- Easy to prepare the database
- Multiple traits introgression can be planned

The results highlight that, (Cont.)

- Efficient data handing mode without subjectivism, and decisioning is fully scientific
- Can edit, customize according to breeders desire
- For RRDI, this can use as database
- Retirements of breeders and their replacements are no more trouble

Conclusions

- The pedigree visualization for phenotypic and molecular data using Pedimap is user-friendly for rice breeding-decisioning with higher accuracy and resource optimization.

- The Pedimap can be applied as a decision-making tool to streamline the rice breeding programs in Sri Lanka.

- An accurate characterization of the breeding germplasm for phenotypic and molecular data is the critical prior step to harness the value of Pedimap for breeding.

Future Directions

Phenology

Need more efficient and organized **phenotyping protocol**

Genomics for Breeding

- Extend the coverage of applicable linkage maps and related DNA markers
- Use DNA sequencing and SNP data can be entered as molecular data
- Include IBD probability values

Expand the breeding germplasm: add all local varieties and international rice varieties

- Local rice varieties
- International rice varieties

Develop as online-public accessible database

Develop a mobile version as an app.

Acknowledgements

- Prof. S. D. S. S. Sooriyapathirana (Supervisor of the project)
- Dr. Kapila Sri Udawela, Rice Research and Development Institute, Bathalaoda, Sri Lanka
- Dr. W.M.W. Weerakoon, Director General, Department of Agriculture, Peradeniya
- Dr. Roeland Voorrips and Dr. Eric van de Weg at Wageningen University and Research Center Nijmegen Area, Netherlands.
- Head of the Department, Prof. R. G. S. C. Rajapakse, and all the academic members of the Department of Molecular Biology and Biotechnology
- Mr. L.T. Ranaweera, Mr. H.S.M. Jayarathne, Mrs. U.K. Wijesundara and Mr. S.M. Shahibdeen
- Non-academic staff members of Department of Molecular Biology and Biotechnology, all the demonstrators, colleagues, friends and my family members

References

- Acquaah, G. (2012) *Principles of plant genetics and breeding*, 2nd edn, pp 281-334. Oxford, UK: Wiley-Blackwell.
- Ali, A.M.S. (2006) Rice to shrimp: Land use/land cover changes and soil degradation in Southwestern Bangladesh. *Land Use Policy.* **23,** 421-435.
- Central Bank of Sri Lanka (2017) Annual Report of Central Bank of Sri Lanka in 2017. Available on-line at: https://www.cbsl.gov.lk/en/publications/economic-and-financial-reports/annual-reports/annual-report-2017.
- Central Bank of Sri Lanka (2018) Annual Report of Central Bank of Sri Lanka in 2018. Available on-line at: https://www.cbsl.gov.lk/en/publications/economic-and-financial-reports/annual-reports/annual-report-2018.
- Chin, J.H., Lu, X., Haefele, S.M., Gamuyao, R., Ismail, A., Wissuwa, M. and Heuer, S. (2010) Development and application of gene-based markers for the major rice QTL Phosphorus uptake 1. *Theor. Appl. Genet.* **120**: 1073-1086.
- Dhanapala, M.P. (2007) Bridging the rice yield gap in Sri Lanka. In *Proceedings of the symposium on Expert Consultation on Bridging the Rice Yield Gap in the Asia-Pacific Region*, 5–8 October 1999, Bangkok, Thailand. pp. 135–145.

- Dharmarathna, W.R.S.S., Herath, S. and Weerakoon, S.B. (2014) Changing the planting date as a climate change adaptation strategy for rice production in Kurunegala district, Sri Lanka. *Sustain. Sci.* **9**: 103-111.
- Duvick, D.N. (1984) Progress in conventional plant breeding. In *Gene manipulation in plant improvement*, pp. 17-31. Boston, MA: Springer.
- Evans, K., Guan, Y., Luby, J., Clark, M., Schmitz, C., Brown, S., Orcheski, B. *et al*.(2011) Large-scale standardized phenotyping of apple in RosBREED. In *Proceedings of the IV International Conference Postharvest Unlimited*, pp. 233-238.
- Food and Agriculture Organization of the United Nations (FAO) (2018) *Rice market monitor,* **21**: 1-3
- Hu, J., Xiao, C., Cheng, M., Gao, G., Zhang, Q. and He, Y. (2015) Fine mapping and pyramiding of brown planthopper resistance genes *QBph3* and *QBph4* in an introgression line from wild rice *O. officinalis*. *Mol. Breeding*, **35**:3. doi:10.1007/s11032-015-0228-2.
- Lu, L., Yan, W., Xue, W., Shao, D. and Xing, Y. (2012) Evolution and association analysis of *Ghd7* in rice. PLoS One, **7**, e34021.
- Mackill, D.J. (1996) Rainfed lowland rice improvement, 242p. Manila, Philippines: International Rice Research Institute.
- Muthayya, S., Sugimoto, J.D., Montgomery, S. and Maberly, G.F. (2014) An overview of global rice production, supply, trade, and consumption. *Ann. Ny. Acad. Sci.* **1324**, 7-14.

- Peace, C.P., Luby, J.J., Van de Wag, W.E., Bink, M.C.A.M. and Iezzoni A.F. (2014) A strategy for developing representative germplasm sets for systematic QTL validation, demonstrated for apple, peach, and sweet cherry. *Tree Genet. Genomes,* **10**, 1679-1694.
- Ragot, M., Bonierbale, M. and Weltzien, E. (2018) *From Market Demand to Breeding Decisions: A Framework. Lima (Peru).* CGIAR Gender and Breeding Initiative. GBI Working Paper. No. 2. Available on-line at: https://www.rtb.cgiar.org/gender-breeding-initiative.
- Rice Research and development Institute (RRDI) (2019, June 5) Rice Cultivation. In *Department of Agriculture, Sri Lanka*, Retrieved from http:// www.doa.gov.lk/rrdi/.
- Rosyara, U.R., Bink, M.C.A.M., Van de Weg, E., Zhang, G., Wang, D., Sebolt, A., Dirlewanger, E., *et al.* (2013) Fruit size QTL identification and the prediction of parental QTL genotypes and breeding values in multiple pedigreed populations of sweet cherry. *Mol. Breeding,* **32**, 875-887.
- Voorrips, R.E. (2007) Pedimap: Instruction manual. Netherlands: Plant Research International. Available on-line at: https://www.wur.nl/en/show/Pedimap.htm.
- Voorrips, R.E., Bink, M.C. and Van de Weg, W.E. (2012) Pedimap: software for the visualization of genetic and phenotypic data in pedigrees. *Journal of Heredity,* **103**: 903-907.

THANK YOU!

RASIKA RATHNAYAKE | S/14/029

DOI /10.1101/2020.01.20.912477

Questions & Answers

Pedimap complete germplasm ▷

Pedimap complete germplasm