ASSEMBLING OF A BREEDING DATABASE USING PEDIMAP: AS A TOOL OF EASY BREEDING DECISION-MAKING IN RICE BREEDING PROGRAMS IN SRI LANKA

By

P.G.R.G. RATHNAYAKE

A PROJECT PROPOSAL

In partial fulfillment of the requirements of the Degree of Bachelor of Science Honors in Molecular Biology and Biotechnology

Of

UNIVERSITY OF PERADENIYA
SRI LANKA
2019

Summary

Rice is the second most grown cereal crop in the world, and in Sri Lanka, more than 1.8 million rural families are employed in rice cultivation. While the present demand for rice is very high, the local rice production is insufficient to fulfill the annual rice requirement of the country due to the influences of several biotic and abiotic stresses. The varietal improvement with desirable and defensive traits is essential for sustainable rice production. If an organized breeding database for rice is available, the flow of phenotypic and genetic traits can be easily identified. Based on them, better breeding decisions can be taken, and the most efficient and effective breeding programs can be planned. This study will focus on the assembly of a breeding database for rice breeding programs in Sri Lanka, by using Pedimap, which is the genetic and phenotypic data visualization software. All the available phenotypic data, including average yield, reaction with the pest and diseases, the period of maturity will be collected and recorded into the database. The genetic information, including DNA marker scores and identical-by-descent probability values, will be evaluated. The pedigree-based visualizations for the flow of the traits, through the pedigrees and allelic representations in different rice varieties, will be used to determine the inheritance patterns of the trait and to select the best parent lines, to design crosses and to determine the selection method. The outcomes of this study will be a preliminary background for breeding decision-making of rice breeding programs in Sri Lanka.

Keywords: Pedimap, breeding database, breeding decision making, rice breeding programs in Sri Lanka, rice breeding

1. Introduction

1.1 Background

Rice or *Oryza sativa* is the world's second most cultivating cereal crop, and more than half of the world population consume it as the staple food (Van Nguyen and Ferrero 2006). Annual production of rice in the world is about 480 million tons, and nearly 90% of it is consumed by Asia including China and India (Muthayya et al. 2014; Adjao and Staatz 2015). Approximately 62.9% of world rice production occurs in the South, and East Asian countries, including India, Indonesia, Bangladesh, Vietnam, Thailand, Myanmar, Philippine, and Pakistan (Calpe 2006). About 1.8 million local rural families are engaged in rice production in Sri Lanka (RRDI 2019). A total of 870,000 ha farmlands is cultivated annually under tropical climatic conditions within two seasons per year as *Maha* season and the *Yala* season (Dharmarathna et al. 2012).

The annual rice production of 2.7 million tons can satisfy only 95% of the local rice demand (Central Bank of Sri Lanka 2017), so the sufficient supply of rice is highly questionable. The price volatility, lack of supply, geographic concentration, relatively low world stockholdings (Jayne 1993; Dawe 2002), socio-economic factors (Herath Banda et al. 1998) and the effect of biotic and abiotic stresses also leads to instability in the local rice market. Sri Lankan rice production declined by 46.1% in 2017 due to drought (Central Bank of Sri Lanka 2017). The country has experienced a temperature increment of 0.016 C° from the last few decades (IPCC 2007) and 3% of the total farmlands were affected by salinity (RRDI 2019). Irregular patterns in annual rainfall, nutrient deficiencies (Walisinghe et al. 2010), mineral toxicities, and inadequate irrigation facilities (Peng et al. 2009), are also causing the low productivity in Sri Lankan rice cultivation (Davis et al. 2016; Farmer 1979). As phosphorus deficiency is the major nutrient deficiency in local farmlands, Sri Lanka spends 0.3 billion USD annually to import fertilizers (Aluwihare et al. 2015,2016). Biotic stresses, including microbial infections, attacks by pests and plant diseases are of significant concern in Sri Lankan rice production (Davis et al. 2016; Wang et al. 2005). Brown planthopper outbreak destroyed thousands of acres of rice fields in the world including Sri Lanka during last few decades (Dyck and Thomas 1979; Khush et al. 1985) and in 1974, 16,200 ha was ravaged, while 2,800 ha of rice cultivation was lost in Ampara (Fernando et al. 1979). The overcoming these restraints is the challenging task today to maintain domestic rice production at a satisfactory level.

1.2 Crop improvement by breeding

The varietal improvement is the best-identified solution (Duvick 1984) to overcome biotic, and abiotic stresses and gain high yield. The introgression of the favorable traits into elite offsprings by hybridization (Xu 2010) with the use of molecular breeding techniques, helps to preserve desirable traits within the germplasm efficiently and accurately (Dhanapala 2007; Duvick 1984). In marker-assisted selection, the allelic segregation of the cross can be visualized with high accuracy by using the DNA markers (Jiang 2013a, b). Identification of the efficient parental lines (Ragot et al. 2018) and detection of the successful selection method (Acquaah 2012) are the two vital components in planning breeding programs based on breeding priorities and the phenotypic and molecular data alone with pedigree history, which are the critical factors needed in the breeding decision-making process.

Computer programs and databases are used to manipulate the pool of primitive data used in breeding decision-making today. Among those data handling platforms, Pedimap is one of the most efficient software, which is used to visualize phenotypic and genetic information in pedigrees and calculate allelic representation (Weebedda et al. 2010). Also, it is helpful to clarify the genetic structure of breeding germplasm and trace the inheritance of traits along with pedigrees based on molecular markers (Voorrips et al. 2012). Its visualizations are used to make breeding decisions about the parentage selection and in planning crosses. Pedimap-based pedigree illustrations have been mainly used to plan crosses in Rosaceae research community (Rosyara et al. 2013; Peace et al. 2014), HiDRAS project (Patocchi et al. 2009), genetic visualizations in Aluwihare et al. (2017) and Paulo et al. (2008), but not yet practiced in the rice breeding programs.

1.3 Objectives

Today there are many breeding databases available in the world; however, no organized breeding database and planning-procedure to produce new rice varieties are present in Sri Lanka. The main aim of this study will be to model a breeding database; an organized form of data collection (Hogan 2018), using Pedimap by collecting all the available pedigree history, phenotypic and genetic data. The other aim will be the assessment of the marker scores of the valuable traits-related DNA markers and calculate the identical-by-descent (IBD) probabilities for each of them.

2. Material and methods

2.1 Evaluation of the DNA markers and scoring

2.1.1 Plant material

The breeder seeds of available rice cultivars will be obtained from the Rice Research and Development Institute (RRDI), Bathalagoda, Sri Lanka. The seeds will be soaked for 48 hours, then germinate in the dark and established on Petri dishes in laboratory conditions. They will be allowed to grow 10 – 14 days (Ahmadikhah 2009).

2.1.2 DNA extraction and PCR

Immature young leaf samples of rice cultivars will be collected and subjected to extract DNA using Dneasy® plant mini kit (Qiagen, Solna, Sweden). The isolated DNA of each cultivar will be stored in TE buffer at -40 C°.

2.1.3 PCR and gel electrophoresis

The extracted DNA for each rice variety will be subjected to PCR amplification with the selected markers, which are associated with favorable qualitative and quantitative traits (Table 1). The PCR conditions will be programmed in the Thermal Cycler (Takara, Japan) as follows. Initial denaturation at 94 °C for 5 minutes, then 35 cycles of 94 °C for 30 seconds for denaturation, primer annealing temperature (Ta) (Table 1) for 90 seconds, and 72 °C for 2 minutes, finally extension at 72 °C for 10 minutes. The PCR products will be size-separated using ethidium bromide stained 2.5% agarose gel with Quick-Load Purple 50 bp DNA Ladder. Then the sizes of the bands will be evaluated, and the marker scores will be defined.

Table 1 The details of the selected DNA markers will use in genetic analysis

Marker	Forward and reverse primer sequence	Ta /C°	Reference
K29	CCATAGTAGCACAAGAAACCGACA GCTTCAATGAGCCCAGATTACGAA	55	Chin *et al.*, 2010
C3-14	GGCAAAATTAGACGGCACG GAATATGCATTTTGTTTGGAG	55	Hu *et al*., (2015)
RM463	TTCCCCTCCTTTTATGGTGC TGTTCTCCTCAGTCACTGCG	55	Li-Hong *et al*., (2006)

2.2 Data analysis in Pedimap

All the available phenotypic data including pedigree data, average yield, maturity, plant height, leaf sheath color, recommended land type, reaction against pest and diseases, brown rice recovery, milling recovery, head rice recovery, amylose content, gelatinization temperature, weight of 1000 grains, grain shape, pericarp color, bushel weight, availability of buff color and molecular marker data for locally cultivated rice varieties will be collected from RRDI, published scientific literature and other publicly available databases including NCBI. The data will be entered according to the format described in figure 1 and figure 2.

Pedimap input file will be created in Microsoft Excel 2019 by using collected data, and the file will be exported as a tab-delimited text file, which is the input data file of the Pedimap. The input data file contains four main sections: header, pedigree, marker data, and IBD probabilities section. The header section (Figure 1 - A) and the red highlighted columns in the pedigree section in Figure 1 - B are the essential elements in the input data file. The rest of the sections in B (Figure 1 -B), and C and D are additional data inputs.

The name of the target population, symbols for unknown values including null homozygous and confirm null homozygous alleles, the ploidy level of the target population and the number of total founder alleles will be entered under header section as shown in figure 2-A. Pedigree information of the varieties, qualitative and quantitative trait information will be entered into pedigree section (Figure 2-B). The marker information including linkage group, the chromosomal location of the marker, marker name, allelic scores of the varieties will be added next, under the marker data section in figure 2-C. Then the IBD probability values will be calculated with the use of FlexQTL (Bink et al. 2008) and the calculated values are inserted under IBD probabilities section (Figure 2-D) (Van de Weg 2014).

POPULATION	=	Sri_Lanka_Rice_Germplasm
UNKNOWN	=	-
NULLHOMOZ	=	$
CONFIRMEDNULL	=	$$
PLOIDY	=	2
NALLELES	=	6

PEDIGREE

NAME	PARENT1	PARENT2	Yeild	Maturity	Leaf_color	BPH	MG	BL
Bg941	-	-	-	-	-			
Pokkali	-	-	-	-	-			
At354	Bg941	Pokkali	6.5	95	Green			
At401	Bg941	Pokkali	5	115	Green			

LINKAGEGROUP 12					
MAP					
RM101	48.2				
RM277	62				
LOCUS	RM101				
ALLELENAMES	110 115 120 125				
FOUNDERALLELES	110 110 110 110 115 120				
LOCUS	RM277				
ALLELENAMES	200 225 250 275 300				
FOUNDERALLELES	200 200 200 200 300 275				
IBDPOSITIONS	48.2 62				
ALLELES	RM323				
Bg941	140	0	140	0	
Pokkali	160	0	180	0	

IBDPOSITION	48.6								
Bg941	1	0	0	0	0	0	0	1	
Pokkali	0	0	1	0	0	0	0	0	
At354	0	0	0	0	0	0	1	0	
At401	0	0	0	0	0	0	0	0	
IBDPOSITION	62.0								

Figure 1 The input data file structure of the Pedimap database; The input file will be created as Microsoft Excel worksheet, contains four main sections. A: Header, B: Pedigree, C: Marker data, D: IBD probabilities. A: In the header section, essential elements are highlighted in blue color, which contains primary data and abbreviation for the components in the database. B: Pedigree section contains the pedigree data of each individual. Essentials are highlighted in red. Qualitative and quantitative data will be added into transections as additional data. C: Marker data section contains linkage map-information for all available markers and marker scores for each was cited. D: Elements in IBD probability section is highlighted in yellow color. IBD probability for each allele combination will be separately calculated and entered here. (Qualitative and quantitative information in section B, Section C, and D were not compulsory for the input data file). The final file will be obtained as tab-delimited text (.txt) file.

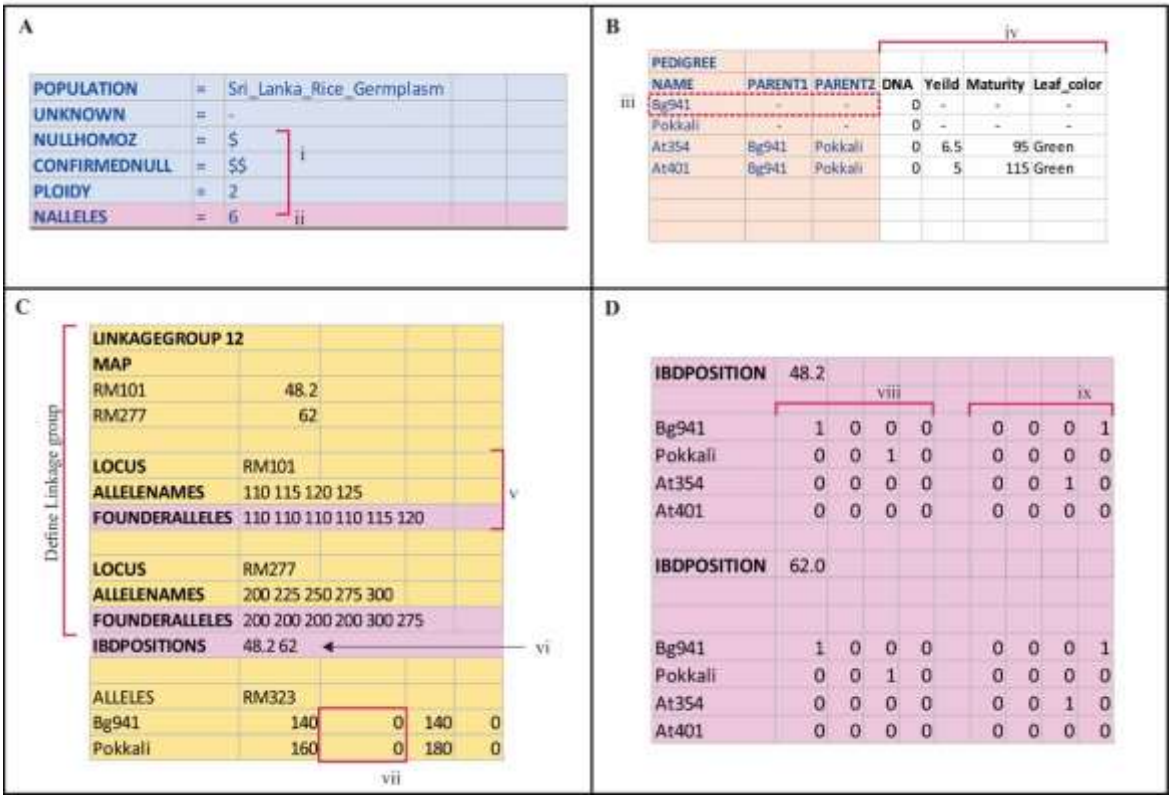

Figure 2 Main sections of the Pedimap input file. A: Header, B: Pedigree, C: Marker data, D: IBD probabilities. A: Header contains basic information about the input population data. (i) abbreviations for missing data, null alleles, confirmed null alleles, and ploidy level will be inserted respectively. (ii) "NALLELE" section was only necessary if the IBD probabilities will be used. A number of total founder alleles will be mentioned here. Each text must be entered without using spaces, and the abbreviations should not be used with other texts. B: Pedigree information is compulsory here. (iii) Initial parents must be entered, and missing values were accepted if it is properly mentioned. (iv) Qualitative and Quantitative data about the QTL will be entered in these columns. The average yield, period of maturity and the basal leaf color were entered here as phenotypic data C:(v) Each linkage group should be cited separately in ascending order. After defining the linkage group, alleles (markers) list with the recombination frequency in cM will be entered. Then each locus was defined with details. Allele score for all observed allele types was entered and, all founder allelic state will be entered which was necessary to work with IDB probabilities. (vi) At the end of each linkage group, IBD probabilities for each locus was cited, and multiple positions were separated by spaces. After defining all linkage groups, allele information for all the mentioned alleles must be entered according to the above order in the linkage group section. (vii) The additional column can be used to assign specific colors for each locus as needed. D: IBD probabilities for each allele will be entered separately. (vii) Maternal probability and (ix) paternal probability were entered and separated by a column. All the IBD probabilities will be calculated by using FlexQTL™.

3. Expected results

3.1 Identifying the marker polymorphism and marker scores for improved rice varieties in Sri Lanka.

The gel electrophoresis will be revealing the allelic scores for selected markers with the use of the appropriate ladder. The available marker details, including linkage map positions and scores, will be added to Pedimap.

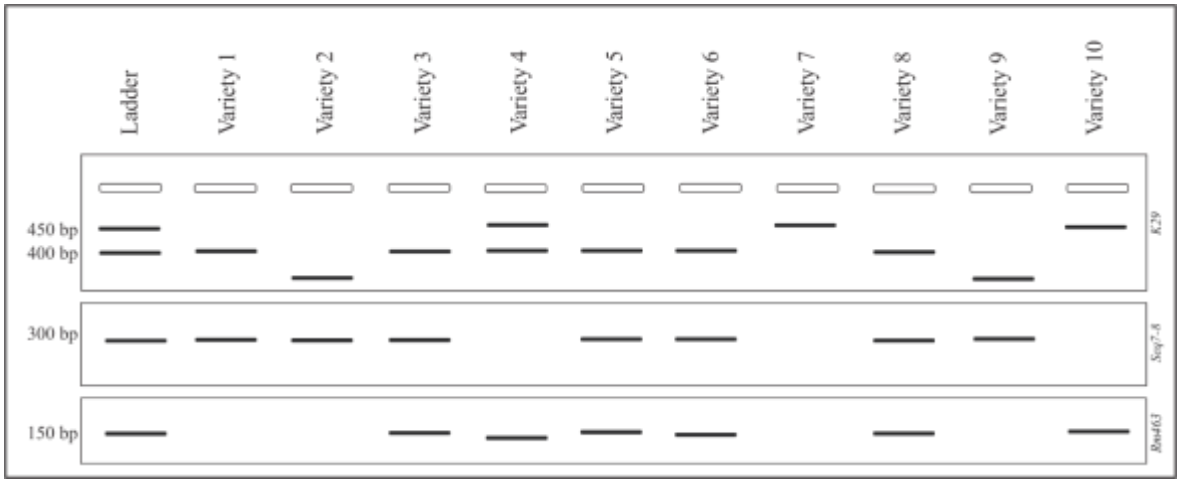

Figure 3 The expected composite gel image of *K29, Seq7-8,* and *RM463* markers of selected improved rice varieties in Sri Lanka. The ladder will be used to evaluate the allelic score for each gel band. The markers are co-dominant, and high allelic polymorphism will be expected.

3.2 Use Pedimap in breeding decision-making.

The main objective of this study is to assemble a breeding database containing data of phenotypic traits and genetic information of Sri Lankan rice varieties. The trait-related co-dominant markers will be used for genetic analysis, and based on the analysis, the marker scores will be entered according to chromosomal position and linkage groups. This evaluated marker scores will be helpful in marker-assisted selection and determination of the allelic flow throughout the cross. The calculated IBD probabilities are the probability of acquiring identical allele from their descent. The inheritance pattern of a particular trait can be identified by using the pedigree visualizations. Based on the identified inheritance patterns for different traits, multi-criteria decision-making (MCDM) (Kariuki et al. 2017; Velasquez and Hesterc 2013) can be performed in a breeding program. Through the application of rice phenotyping protocol for the selected parents (Evans et al. 2011), the most suitable parental lines will be identified, and the type of cross, the method of selection and appropriate markers will be identified.

The application of the program in breeding decision-making will be explained by using the following three cases. Each case will be selected based on the widespread outbreaks in local rice production.

3.2.1 Case 1: Planning of new rice variety with high yield.

The Pedimap gives pedigree relations of the whole available varieties as the basic output of the program. The output of the study will be customized, according to Voorrips (2007) and represent the flow of the average yield of the improved rice varieties in Sri Lanka, as shown in figure 4. High yielding varieties, which can be determined based on the phenotypic flow, will be able to be used as the parents in the new cross.

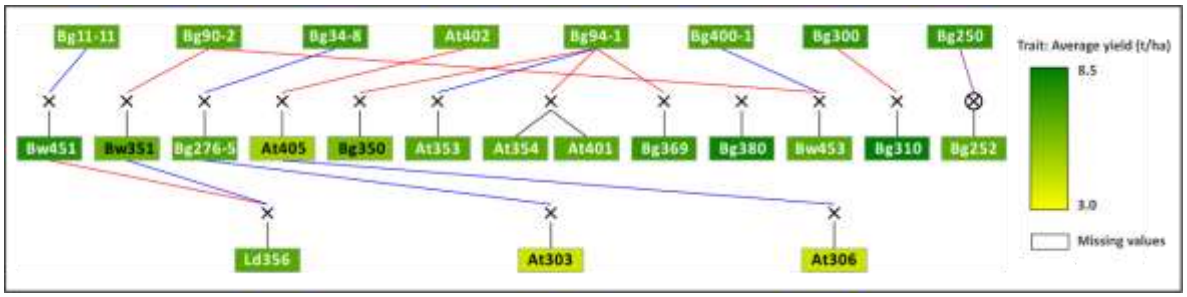

Figure 4 The expected pedigree visualization of the average yield of selected improved rice varieties in Sri Lanka. Female parentage will be indicated by red lines and male parentage by blue lines. The cross mark will indicate the cross between two varieties and the cross inside the circle represents the selection. The background colors will indicate the average yield of the variety, and the white color is for unavailable values. The legend of color indications will be mentioned in the right corner.

3.2.2 Case 2: Use of marker data and IBD probabilities for the development of high yield, Phosphorus deficiency tolerance rice variety.

The visualizations of the phenotype-associated DNA marker data (Figure 6 -A) and IBD probabilities (Figure 6- B) through the pedigree, will be used to trace the allelic segregation of desirable alleles and determine the probability of acquiring them in the progeny. It will illustrate the chromosomal pattern of progressive recombination, through the generations. The parental selection, cross designing, and the identification of the selection method will be done based on the genetic architecture and zygosity of the alleles. The crossing order of the multiple parents will be based on the IBD probability values. The parental lines with the high IBD probability values will be crossed first, and then, the parents with low IBD probabilities will be subjected to subsequent cross.

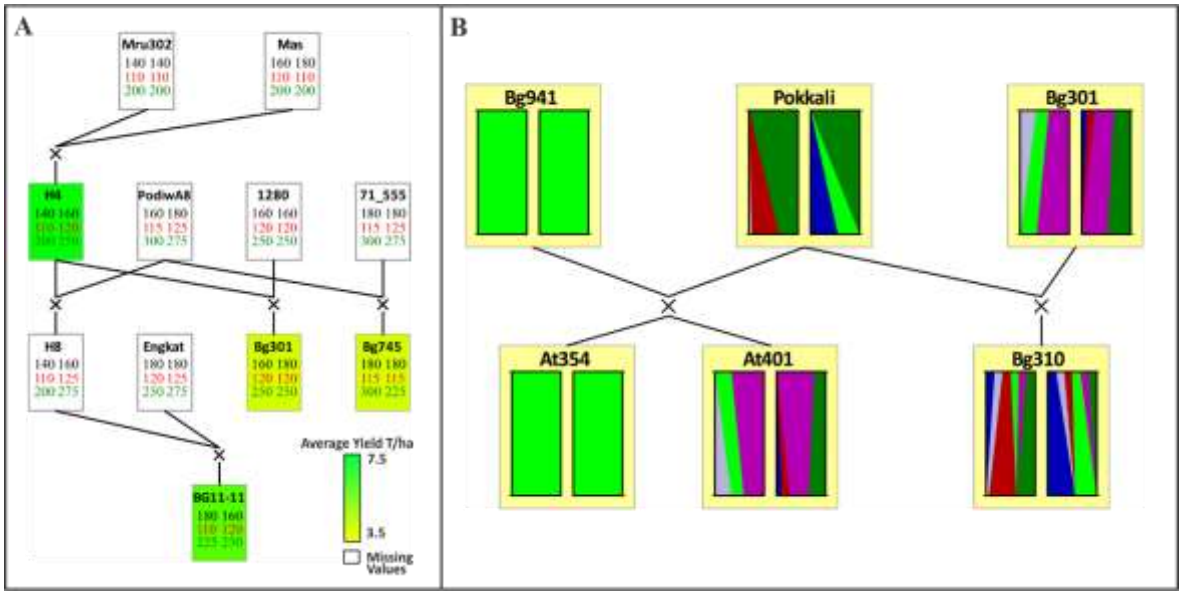

Figure 5 The expected visualization of the inheritance of the phosphorus deficiency tolerance associated markers in the pedigree. The inheritance of the phosphorus deficiency tolerance is tested with *K29* (black), *K41* (red) and *K46* (green) DNA markers in the order of the marker score. The inheritances of the yield based on quantitative data are also combined with the pedigree.

3.2.3 Case 3: Multiple traits integration in crop improvement.

Multiple trait integration is one of the novel approaches in breeding, rather than planning on improving a single trait. The Pedimap visualizations will be used with MCDM in rice breeding planning to identify multiple parental lineages with desirable traits. Along with the high yield (Figure 4), BPH resistance (Figure 6 – A) and the period of maturity (Figure 6 – B) will be considered. All the possible parental lines will be selected, and if there are multiple parents, they will be weighted according to the phenotyping protocol (Evans 2011). Then, the most efficient parental lineages can be identified, and the order of crossing will be considered on the allelic segregation frequencies.

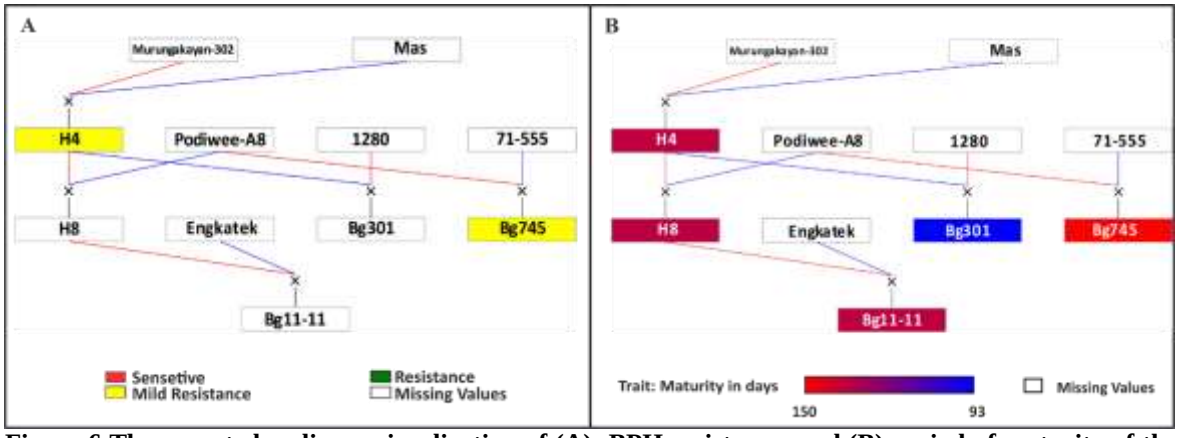

Figure 6 The expected pedigree visualization of (A): BPH resistance, and (B) period of maturity of the selected improved rice varieties in Sri Lanka. Female parentage will be indicated by red lines and male parentage by blue lines. The cross mark indicates the cross between two varieties. The legend indicates the range of the qualitative traits and the different type of quantitative traits.

4. SMART and SWOT analysis

4.1 SMART Analysis

4.1.1 Specific

The specific objective of the study is focused on assembling a breeding database for rice breeding programs in Sri Lanka. The end result of this study will be a breeding database with phenotypic and genetic data, which will allow efficient breeding decision-making such as parental selection, planning crosses, identifying the most suitable selection method and DNA markers for marker-assisted selection by identifying the inheritance patterns of the alleles.

4.1.2 Measurable

Pedigree history, average yield, maturity, plant height, leaf sheath color, recommended land type, the reaction against pest and diseases, amylose content, gelatinization temperature, gain related parameters and molecular marker data are measured to assess qualitative, quantitative traits and marker scores. An appropriate rice phenotyping protocol can be used to confirm and validate the selected parent lines for multiple traits.

4.1.3 Achievable

This study can be achieved via assembling a breeding database for Sri Lankan rice breeding programs and use it in breeding decision-making process. Feasibility-wise, this study is achievable due to the availability of samples, available laboratory facilities, resources, standardized procedures, and available primitive data.

4.1.4 Realistic/ Relevant

Despite conventional breeding approaches, breeding database with phenotypic and genetic data for breeding programs in Sri Lanka have not been created yet. Large numbers of farmlands are destroyed due to severe biotic and abiotic stresses during last few decades and improved crop varieties with defensive traits, could decrease the severity and gain high yield.

4.1.5 Time-bound

The extent of the study is from February to September 2019. Sampling will take about a month, and the database construction will take another two months. Genetic analysis will take two to three months and finalizing will be spent two to three weeks.

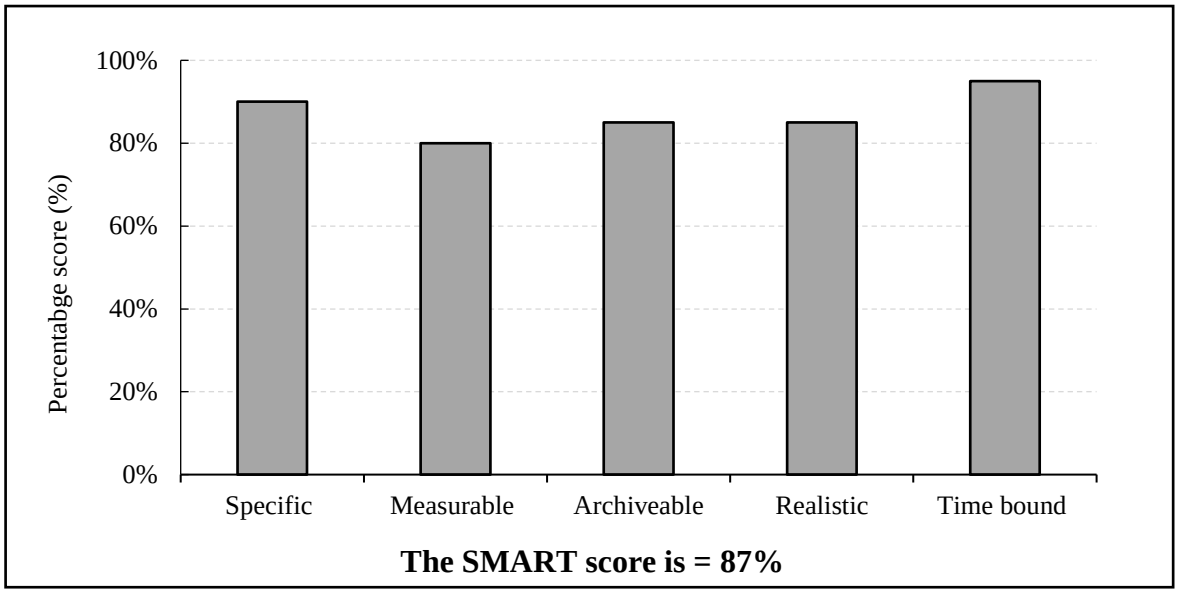

Figure 7 The graph represents the SMART score calculation. The peaks of the graph indicate the components of SMART analysis. The total SMART score was calculated by using the average value for all peaks.

4.2 SWOT analysis

4.2.1 Strengths

Strengths of this study are mainly the guidance, expertise, and experience of a supervisor, a co-supervisor, and a research assistant, along with the availability of samples at RRDI, Bathalegoda, and the availability of well-tested, standardized protocols to conduct lab work. The primers are readily available, and the preliminary data for the database development is also available.

4.2.2. Weaknesses

A few weaknesses exist including the high level of data heterogeneity. All the necessary information for all varieties is unavailable at the moment. Conducting DNA marker analysis for only three markers is insufficient, and the accuracy of the result slightly decrease, as a single QTL is associated with many DNA markers rather than one. Furthermore, the limited time, limited availability of some chemicals and consumables at the laboratory are some of the other weaknesses of the project.

4.2.3 Opportunities

The significant opportunity of this study will be the creation of a breeding database for improved rice varieties in Sri Lanka, which helps to formulate efficient rice breeding programs locally. It will accelerate breeding decision-making action with high accuracy and

efficiency. Moreover, this application can be used to predict the properties of the particular progeny and plan new breeding strategies. By developing applications compatible with other platforms including mobile systems with attractive graphical user interface as a handheld application could enhance the usage of this technique. Besides rice breeding, the same approach can be developed into all other breeding programs as well.

4.2.4 Threats

The development of other software having the same features will increase the competition with the created database. Further addition of the user-friendly, handheld, extended features to other breeding programming software will reduce the uniqueness of the findings. The same research concept can be implemented by some other research group also make conflict with the ultimate presentation of the product.

	Strengths	Score	**Weaknesses**	Score
Internal origin	• Samples are available at the RRDI and PGRC	3	• Testing for 2 - 3 DNA markers are not enough to make efficient decisions	-3
	• Transport facilities are available	3		
	• Sample preparation for DNA extraction is easy	3	• Data heterogeneity	-2
	• Sufficient laboratory facilities are available	3		
	• Verified protocols are available	3		
	• SNP data are publicly available at the databases.	1		
	• Necessary DNA primers are currently available	3		
	• Guidelines from supervisors and co supervisors	3		
	• Support from the research assistants	3		
	The score for total Strengths	**25**	The score for total Weaknesses	**-5**
External Origin	**Opportunities**	Score	**Threats**	Score
	• Can be used to create the breeding database	3	• Other software can be developed with more features	-3
	• Easy to formulate selection method and appropriate markers in MAS	3		
	• Can be developed as a public database, that anyone can enter the data.	2		
	• Can develop software to support by all flatforms	2		
	• Can develop a mobile app	3		
	• Can be plan new crosses without more troubles	2		
	• Can be predicted a progeny of the cross	2		
	• Can be used to formulate proper breeding decision- making procedure	3		
	The score for total Opportunities	**20**	The score for total Treats	**-3**
Total SWOT score = 61.7%				

Figure 8 The graphical representation of the SWOT analysis and score calculation. All the factors mentioned under four sections were scored by using three preference levels; 1: Low level of preference, 2: Moderate level of preference, 3: High level of preference. The respective score for each factor was indicated next to the factors. Strengths and opportunities were considered as positive values, and the weaknesses and the treats were considered as negative values. The total score was taken as a percentage from the fraction of the sums for all four sections, from the total available score.

5. References

Acquaah G (2012) Principles of plant genetics and breeding, 2nd edn. Wiley-Blackwell, Oxford, pp 281-334

Adjao RT, Staatz JM (2015) Asian rice economy changes and implications for sub-Saharan Africa. Glob Food Sec 5:50-55. https://doi.org/10.1016/j.gfs.2014.11.002

Ahmadikhah A (2009) Rapid mini-prep DNA extraction method in rice (*Oryza sativa*). African Journal of Biotechnology 8(2):323-327.

Aluwihare YC, Chamikara MDM, Dissanayake DRRP, Dissanayake MDMIM, Karannagoda NNH, Dayananda AGMLK, Sirisena DN, Samarasinghe WLG, Rajapakse RGSC, Sooriyapathirana SDSS (2017) DNA sequence polymorphism of *Pup1* linked *K20-1* STS region can be effectively used in molecular breeding of rice for phosphorus deficiency tolerance. J Natn Sci Foundation Sri Lanka 45(4):413-425. https://dx.doi.org /10.4038/jnsfsr.v45i4.8235

Aluwihare YC, Chamikara MDM, Dissanayake DRRP, Karannagoda NNH, Sirisena DN, Samarasinghe WLG, Rajapakse S, Sooriyapathirana SDSS (2015) Validation of *K46*, a *Pup1*-linked marker, using a selection of Sri Lankan rice (*Oryza sativa* L.) germplasm for marker assisted selection towards phosphorous deficiency tolerance. Ceylon J Sci (Bio Sci) 44(2):45-54. https://doi.org/10.4038/cjsbs.v44i2.7349

Aluwihare YC, Ishan M, Chamikara MDM, Weebadde CK, Sirisena DN, Samarasinghe WLG, Sooriyapathirana SDSS (2016) Characterization and selection of phosphorus deficiency tolerant rice genotypes in Sri Lanka. Rice Sci 23(4):184-195. https://doi.org /10.1016/j.rsci.2015.10.001

Bink MCAM, Boer MP, Ter Braak CJF, Jansen J, Voorrips RE, Van de Weg WE (2008) Bayesian analysis of complex traits in pedigreed plant populations. Euphytica 161(1-2):85-96. https://doi.org/10.1007/s10681-007-9516-1

Calpe C (2006) Rice international commodity profile. Rome: Food and Agricultural Organization of the United Nations.

Central Bank of Sri Lanka (2017) Annual Report. https://www.cbsl.gov.lk/. Accessed 15 September 2019

Chin JH, Lu X, Haefele SM, Gamuyao R, Ismail A, Wissuwa M, Heuer S (2010) Development and application of gene-based markers for the major rice QTL Phosphorus uptake 1. Theor Appl Genet 120:1073-1086.

Davis KF, Gephart JA, Gunda T (2016) Sustaining food self-sufficiency of a nation: The case of Sri Lankan rice production and related water and fertilizer demands. Ambio 45(3):302-312. https://doi.org/10.1007/s13280-015-0720-2

Dawe D (2002) The changing structure of the world rice market, 1950–2000. Food Policy 27(4): 355-370.

Dhanapala MP (2007) Bridging the rice yield gap in Sri Lanka. Proceedings of the symposium on "Expert Consultation on Bridging the Rice Yield Gap in the Asia-Pacific Region", Bangkok, Thailand, 5–8th October 1999, pp 135–145.

Dharmarathna WRSS, Herath S, Weerakoon SB (2012) Changing the planting date as a climate change adaptation strategy for rice production in Kurunegala district, Sri Lanka. Sustain Sci 9(1):103-111. https://doi.org/10.1007/s11625-012-0192-2

Duvick DN (1984) Progress in conventional plant breeding. In: Gene manipulation in plant improvement. Springer, Boston, MA, pp 17-31

Dyck VA, Thomas B (1979) The brown planthopper problem. In: Brown planthopper: threat to rice production in Asia. International Rice Research Institute, Los Banos, Philippines, pp 3-17.

Evans K, Guan Y, Luby J, Clark M, Schmitz C, Brown S, Orcheski B, Peace C, Van De Weg E, Iezzoni A (2011) Large-scale standardized phenotyping of apple in RosBREED. In: IV International Conference Postharvest Unlimited. Acta Hort 945:233-238. https://doi.org/10.17660/actahortic.2012.945.31

Farmer BH (1979) The 'green revolution' in South Asian rice fields: environment and production. The Journal of Development Studies 15(4):304-319. https://doi.org/10.1080/00220387908421735

Fernando H, Senodhera D, Elikawela Y, de Alwis HM, Kudagamage C (1979) Varietal resistance to the brown planthopper in Sri Lanka. In: Brown Planthopper: Threat to Rice Production in Asia. International Rice Research Institute, Los Banos, Philippines, pp 241-249

Herath Banda RM, Dhanapala MP, de Silva GCA, Hossain M (1998) Constraints Increasing Rice Production in Sri Lanka. Proceedings of the workshop on Prioritization of Rice Research. IRRI, Los Banõs, Laguna, Philippines, pp 20-22

Hogan R (2018) A practical guide to database design. Chapman and Hall, CRC, p 1

Hu J, Xiao C, Cheng M, Gao G, Zhang Q, 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.

IPCC (2007) Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Parry ML, Canziani OF, Palutikof JP, Van der Linden PJ, Hanson CE (eds). Cambridge University Press, Cambridge, UK, p 976

Jayne T (1993) Sources and effects of instability in the world rice market, MSU International Development Paper No. 13. Department of Agricultural Economics and Department of Economics, Michigan State University, p 104

Jiang GL (2013a) Plant Marker-Assisted Breeding and Conventional Breeding: Challenges and Perspectives. Advances in Crop Science and Technology 1:e106. https://doi.org/10.4172/2329-8863.1000e106

Jiang GL (2013b) Molecular markers and marker-assisted breeding in plants. In: Plant breeding from laboratories to fields. IntechOpen, pp 45-83 https://dx.doi.org/10.5772/52583

Kariuki CM, van Arendonk JAM, Kahi AK, Komen H (2017) Multiple criteria decision-making process to derive consensus desired genetic gains for a dairy cattle breeding objective for diverse production systems. Journal of dairy science 100(6):4671-4682. https://doi.org/10.3168/jds.2016-11454

Khush GS, Karim AR, Angeles ER (1985) Genetics of resistance of rice cultivar ARC10550 to Bangladesh brown planthopper teletype. J Genet 64(2-3):121-125. https://doi.org/10.1007/bf02931140

Li-Hong SUN, Chun-Ming WANG, Chang-Chao SU, Yu-Qiang LIU, Hu-Qu ZHAI, Jian-Min WAN (2006) Mapping and marker-assisted selection of a brown planthopper resistance gene *bph2* in rice (*Oryza sativa* L.). Acta Genetica Sinica 33:717-723.

Muthayya S, Sugimoto JD, Montgomery S, Maberly GF (2014) An overview of global rice production, supply, trade, and consumption. Ann NY Acad Sci 1324(1):7-14. https://doi.org/10.1111/nyas.12540

Patocchi A, Fernández-Fernández F, Evans K, Silfverberg-Dilworth E, Matasci CL, Gobbin D, Rezzonico F, Boudichevskaia A, Dunemann F, Stankiewicz-Kosyl M, Mathis F (2009) Development of a set of apple SSRs markers spanning the apple genome, genotyping of HiDRAS plant material and validation of genotypic data. Acta Hortic 2(814):603-608. https://doi.org/10.17660/actahortic.2009.814.102

Paulo MJ, Boer M, Huang X, Koornneef M, Van Eeuwijk F (2008) A mixed model QTL analysis for a complex cross population consisting of a half diallel of two-way hybrids in Arabidopsis thaliana: analysis of simulated data. Euphytica 161(1-2):107-114. https://doi.org/10.1007/s10681-008-9665-x

Peace CP, Luby JJ, Van de Wag WE, Bink MCAM, Iezzoni AF (2014) A strategy for developing representative germplasm sets for systematic QTL validation, demonstrated for apple, peach, and sweet cherry. Tree Genetics and Genomes 10(6):1679-1694. https://doi.org/10.1007/s11295-014-0788-z

Peng S, Tang Q, Zou Y (2009) Current status and challenges of rice production in China. Plant Prod Sci 12(1):3-8. https://dx.doi.org/10.1626/pps.12.3

Ragot M, Bonierbale M, Weltzien E (2018) From Market Demand to Breeding Decisions: A Framework. Lima (Peru). CGIAR Gender and Breeding Initiative. GBI Working Paper. No. 2. https://www.rtb.cgiar.org/gender-breeding-initiative

Rice Research and Development Institute (RRDI) (2019) Rice Cultivation. https://www.doa.gov.lk/rrdi/. Accessed 5th June 2019

Rosyara UR, Bink MCAM, Van de Weg E, Zhang G, Wang D, Sebolt A, Dirlewanger E, Quero-Garcia J, Schuster M, Iezzoni AF (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(4):875-887. https://doi.org/10.1007/s11032-013-9916-y

Van de Weg E (2014) Use of Pedimap software in Breeding. Proceeding at the workshop, Wageningen (Rhenen), The Netherlands, 12th June 2014. https://fruitbreedomics.com/index.php/media-center/learning-material-training-sessions/156-use-of-pedimap-software-in-breeding. Accessed 5 May 2019

Van Nguyen N, Ferrero A (2006) Meeting the challenges of global rice production. Paddy Water Environ 4:1–9. https://doi.org/10.1007/s10333-005-0031-5

Velasquez M, Hester PT (2013) An analysis of multi-criteria decision-making methods. International Journal of Operations Research 10(2):56-66.

Voorrips RE (2007) Pedimap: Instruction manual. The Netherlands: Plant Research International. https://www.wur.nl/en/show/Pedimap.html. Accessed 5 May 2019

Voorrips RE, Bink MC, Van de Weg WE (2012) Pedimap: software for the visualization of genetic and phenotypic data in pedigrees. Journal of Heredity 103(6):903-907. https://doi.org/10.1093/jhered/ess060

Walisinghe BR, Gunaratne LHP, Prasada DVP, Sarthaj IZ (2010) Analysis of the consumer preferences for rice-based noodles and biscuits. Annals of Sri Lanka Department of Agriculture 12: 201-211.

Wang Y, Xue Y, Li J (2005) Towards molecular breeding and improvement of rice in China. Trends Plant Sci 10(12):610-614. https://doi.org/10.1016/j.tplants.2005.10.008

Weebadde C, Crisosto C, Fazio G, Gallardo K, Main D, Peace C, Reighard G, Sebolt A, Stone A, Van de Weg E, Voorrips R, Xu K, Iezzoni A (2010) Use of PediMap in RosBREED for Visualizing Genetic and Phenotypic Data across Pedigrees in Rosaceous Crops. HortScience: a publication of the American Society for Horticultural Science (45)8: S55-S55

Xu Y (2010) Molecular plant breeding. Cabi

6. Budget

Table 1 The estimated budget of the study

Item	Quantity	Estimated cost (SLR)
Major equipment		
Thermal cycler	1	2,000,000.00
Gel electrophoresis apparatus	1	1,200,000.00
Minor equipment		
pH meter	1	75,000.00
Micropipette	3	600,000.00
Centrifuge	1	3,000,000.00
Microwave oven	1	15,000.00
Chemical balance	1	75,000.00
Ice maker	1	1,500,000.00
Autoclave	1	1,500,000.00
Vortex machine	1	500,000.00
Refrigerators (-20 C°)	1	50,000.00
Refrigerators (-80 C°)	1	100,000.00
Mortar and pestles	5	75,000.00
Chemicals		
Taq polymerase	5 ml	40,000.00
Primers	3 pairs	9,000.00
Agarose	50 g	3,750.00
Tris base	300 g	6,000.00
EDTA disodium anhydrate	50 g	4,000.00
Ethanol	2 l	10,000.00
ETBR	100 µl	300.00
Bromophenol Blue	2.5 ml	100.00
Quick-Load Purple 50 bp DNA Ladder	1.25 ml	40,000.00
Consumable		
Pipette tips - White	2500	16,250.00
Pipette tips – Yellow	200	1,600.00
Pipette tips - Blue	150	1,600.00
Eppendorf tubes	150	1,600.00
PCR tubes	800	6,400.00
Distilled water Bottles	1	500.00
Glassware		25,000.00
Transport		
For the sample collection	4 x (110 km)	13,200.00
Labors		
Research assistance salary	1	780,000.00
Miscellaneous		
Stationaries		
Printing, aluminum foil, gloves, tissue paper, ice cube	150	5,000.00
Labels	1 pack	500.00
OPH pens	4	400.00
Cello Tapes	2	300.00
Packing Materials (polypropylene and zip lock bags)		1,500.00
Total cost		**11,657,000.00**

7. Gannt chart

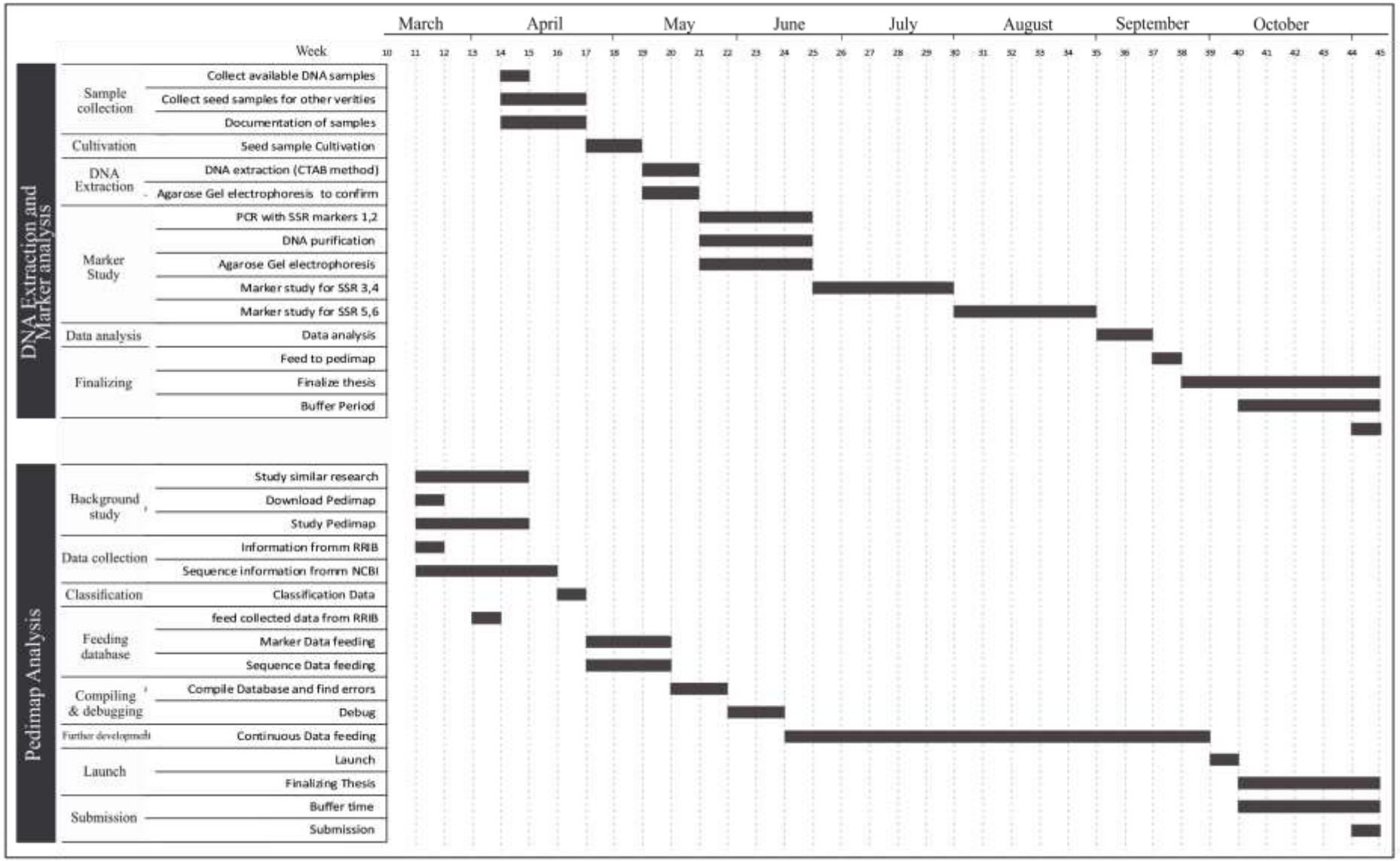

Figure 9 The Gannt chart of the project. The study will be divided into two main subsections as, DNA extraction and marker analysis, and Pedimap analysis. The expected time duration for each task is marked with solid lines.