CROP BREEDING AND APPLIED BIOTECHNOLOGY

ARTICLE

Population structure and diversity of Southeast Asian rice varieties

Muhamad Sabran[1*], Puji Lestari[1], Dwinita Wikan Utami[1], Karden Mulya[1], Dani Satyawan[1], Susilawati[1], Rerenstradika Nizar Terryana[1], Kristianto Nugroho[1], Marilyn Ferrer[2], Rahiniza Kamaruzaman[3] and Koukham Vilayheuang[4]

Crop Breeding and Applied Biotechnology
23(4): e461423414, 2023
Brazilian Society of Plant Breeding.
Printed in Brazil
http://dx.doi.org/10.1590/1984-70332023v23n4a49

Abstract: *This study assessed the structure and genetic diversity of rice populations of Southeast Asian varieties, based on quantitative morphological and molecular traits. Population structure analysis revealed four distinct populations as ancestral origin of the varieties in the collection. Some traditional varieties from different countries share the same ancestry, while on the other hand, admixture was observed in the ancestry of some varieties. High diversity in quantitative morphological traits was confirmed in the rice collection. Spikelet fertility and plant height contributed significantly to the diversity.*

Keywords: *Rice, SNP chip, traditional variety, improved variety*

INTRODUCTION

Rice is a staple food in Southeast Asia, where the daily diet of approximately 650 million people depends heavily on this crop. Consequently, rice production is a high priority in the agricultural development plans of Southeast Asian governments. This region accounts for approximately 25% of the global production of the cereal (Mutert and Fairhurst 2002).

Southeast Asia, the secondary center of origin of rice, played an important role in the domestication of the crop (Sweeney and McCouch 2007, Bellwood 2011). Rice was brought from China to Laos and Bhutan between 2500 and 200 BC and via maritime routes to the Philippines, Malaysia and Indonesia. Genomic studies have indicated that the initial type of rice dispersed in Southeast Asia was ssp. *japonica* of domesticated *Oryza sativa*. In addition, the tropical line of this subpopulation in the Malay Archipelago differs from that in Laos and Bhutan (Gutaker et al. 2020).

While most rice in Southeast Asia belongs to the species *Oryza sativa*, long domestication, breeding and cultivation periods have resulted in significant regional genetic variation. Traditional varieties or landraces have evolved in specific areas over several hundred years of cultivation and selection, without crossbreeding with other varieties. Accordingly, their characteristics vary across different locations (Casañas et al. 2017). Similarly, new varieties developed for various purposes using multiple gene sources have resulted in a wide range of improved varieties. The gene sources of the improved varieties released in different countries may be similar, as has been shown by population structure and diversity analyses of traditional and improved varieties (Gutaker et al. 2020, Hour et al. 2020).

***Corresponding author:**
E-mail: muha329@brin.go.id
ORCID: 0000-0002-2542-4059

Received: 13 July 2023
Accepted: 01 October 2023
Published: 25 October 2023

[1] Agriculture and Food Research Organization, National Research and Innovation Agency, Cibinong Science Center, Jl Jakarta-Bogor no 147, Cibinong, Indonesia
[2] Philippines Rice Research Center, Maligaya, Science City of Munoz, 3119 Nueva Ecija, Philippines
[3] Malaysian Agriculture Research and Development Institute Kuala Selangor, Peti Surat 12301, Pejabat Pos Besar, 50774, Kuala Lumpur, WP Kuala Lumpur, Malaysia
[4] National Agriculture and Forestry Research Institute (NAFRI), Nongviengkham Village, Xaythany District, Vientiane Capital, Lao P.D.R., Laos

Genetic variation is crucial for the success of breeding programs in developing new rice varieties. Assessing and effectively exploiting genetic diversity in breeding programs is critical for sustainable genetic improvement and rapid adaptation to the changing breeding objectives (Apraku et al. 2021). Understanding the population structure can help assess systematic differences in allele frequencies between subpopulations and thus be useful for breeding purposes. Molecular markers such as single nucleotide polymorphisms (SNPs) have been used to study population structure. Simultaneously, genetic diversity in rice can also be assessed based on SNP analysis and quantitative morphological traits (Nachimuthu et al. 2015). The population structure and diversity of various sets of varieties or genotypes, including rice populations at the national level (Thomson et al. 2007, Lestari et al. 2017, Hour et al. 2020) or specific types of rice have been analyzed (Lahkar and Tanti 2017, Islam et al. 2018, Rashid et al. 2018, Epe et al. 2021). However, rice populations from multiple countries of a specific region have yet to be examined. In this study, we analyzed the population structure and diversity of a total of 92 rice varieties, consisting of both traditional and improved varieties from four countries in Southeast Asia.

MATERIAL AND METHODS

Plant material and experimental area

This study collected 73 traditional and 19 improved varieties from four Southeast Asian countries: Indonesia, Malaysia, Lao PDR, and the Philippines. Field experiments were conducted at four sites: the experimental stations of the Philippine Rice Research Institute, the National Agriculture and Forestry Research Institute (Lao PDR), the Malaysian Agricultural Research and Development Institute, and a farmer's field in Indonesia.

Experimental design

The field experiments were conducted during the dry season under well-irrigated conditions at representative sites of the four countries. A completely randomized block design with three replications was used to evaluate the 92 varieties in each country. The plot size was 0.5 x 1 m, with row and in-row plant spacing of 20 cm. Three-week-old rice seedlings were transplanted from the nursery to the field. Agronomic techniques followed local recommendations and varied among countries. For each variety, 10 randomly selected plants at various growth stages were used to measure quantitative morphological traits, according to the standard descriptors for rice (Bioversity International 2007). The following morphological features were determined (means per plant): caryopsis length (CL), caryopsis width (CW), culm diameter (CD), culm number (CN), plant height (PH), spikelet fertility (F), flag-leaf length (FLL), flag-leaf width (FLW), grain length (GL), grain width (GW), 100-grain weight (GWT), number of filled grains per panicle (GF), leaf blade length (LBL), leaf blade width (LBW), panicle length (PL), panicle number (PN), and total grain weight per plant (TWP).

Genotyping and population structure analysis

Leaf samples of each plant were diluted to a concentration of 50 ng μL^{-1} for DNA extraction and SNP genotyping. The DNA quality and quantity were evaluated by spectrophotometry. Genotyping was performed at laboratories of the International Rice Research Institute (IRRI) using the 7K SNP BeadChip (Morales et al. 2020), according to the manufacturer's instructions (Illumina, USA). The resulting SNP data were transformed into PLINK format and implemented in the Admixture software (Apraku et al. 2021) to determine the population structure. Cross-validation techniques were applied at various K values ranging from K=2 to K=8, and cross-validation error estimates were monitored to determine the actual genetic population (K) number. The SNP genotyping data were also analyzed in TASSEL (Bradbury et al. 2007), to determine the phylogenetic relationships within the rice collection by the neighbor-joining method.

Phenotypic data analysis

Of the17 traits, 15 were consistently analyzed in replication in each country. Incomplete data of number of filled grains per panicle and total grain weight per plant were excluded from further analysis. The following model was used to study the variance of each trait:

$$Y_{ijk} = \mu + C_i + \beta C_{k(i)} + V_j + CV_{ij} + \varepsilon_{ijk} \qquad (01)$$

where C_i is the effect of the i^{th} country; $\beta C_{k(i)}$ the k^{th} block effect within the i^{th} country; V_j the effect of the j^{th} variety; CV_{ij}

the effect of interaction between the i^{th} country and the j^{th} variety; and ε_{ijk} and Y_{ijk} are the error term and the observed value, respectively, of a morphological trait in the i^{th} country of the j^{th} variety at the k^{th} block (i=1,2,3,4; j=1,2,…89; k=1,2,3). The genotypic variance (σ_g^2) and phenotypic variance (σ_p^2) were estimated by the following formula:

$$\sigma_g^2 = \frac{MSV - MSE}{3 \times 4} \quad (2)$$

$$\sigma_p^2 = \sigma_g^2 + \sigma_e^2 \quad (3)$$

where MSV and MSE, respectively, are the mean square variety and mean square error of the analyses of variance based on model (1).

The genotypic coefficient of variation (GCV) and phenotypic coefficient of variation (PCV) of trait X were calculated as follows:

$$GCV\ (\%) = \left(\frac{\sqrt{\sigma_g^2}}{\bar{X}}\right) x100 \quad (4)$$

$$PCV\ (\%) = \left(\frac{\sqrt{\sigma_g^2}}{\bar{X}}\right) x100 \quad (5)$$

where $\bar{X}$ is the mean value of X traits. Furthermore, heritability in the broad sense (H^2) and selection gain (SG) as percent of mean were calculated by the following formula (Falconer 1986):

$$H^2 = \frac{\sigma_g^2}{\sigma_p^2} \quad (6)$$

$$SG\ (\%) = K \times H\frac{\sqrt{\sigma_p^2}}{\bar{X}} x100 \quad (7)$$

where K is a constant whose value depends on the proportion of the population included in the selected group (K =2.063, at 5% selection intensity).

The contribution of each character to divergence was computed by considering all combinations of varieties (Zaman et al. 2005). Principal component and cluster analysis were conducted based on the mean data of the measured morphological traits averaged across the four countries. The clusters were determined by the average hierarchical clustering method, which was selected due to its high goodness of fit in generating a dendrogram (Saracli et al. 2013). All phenotypic data analyses were performed using R Statistical software version 4.2.2.

RESULTS AND DISCUSSION

Population structure

Population structure analysis using Admixture software indicated a high probability that four populations can be differentiated in this rice collection (hereafter population Q1, Q2, Q3, and Q4), since the cross-validation error dropped to its lowest level at K=4. The circular bar plots (Figure 1) illustrate the proportion of genetic admixture between the four populations Q1, Q2, Q3, and Q4, in the genetic composition of each variety. The admixture pattern aligns well with the neighbor-joining phylogenetic tree of the 92 rice varieties.

The geographical origin of the ancestral populations could not be clearly indicated, as each inferred population contained rice varieties from multiple countries. However, a pure (proportion 100%) Q1 ancestry was confirmed in *japonica* rice varieties from the Philippines, Indonesia, and Malaysia. Varieties with pure Q2 ancestry are *indica* rice varieties from Indonesia, Malaysia, and the Philippines. Pure Q3 ancestry was confirmed in *indica* rice varieties from Malaysia and Indonesia, while those from Laos represent predominantly pure Q4 ancestry.

It is worth noting that some varieties with a 100% proportion of a specific ancestral population may originate from countries other than the country of collection. For example, in *Ketan Maronto*, a variety sampled in Indonesia, 100% of the ancestral population was found to be identical to *indica* rice from Laos. *Ketan Maronto* is a sticky rice variety, similar to many local varieties found in Laos. This suggests that *Ketan Maronto* is native to Laos and was later introduced to Indonesia, where it became a traditional variety after centuries of domestication. Another possibility is that *Ketan*

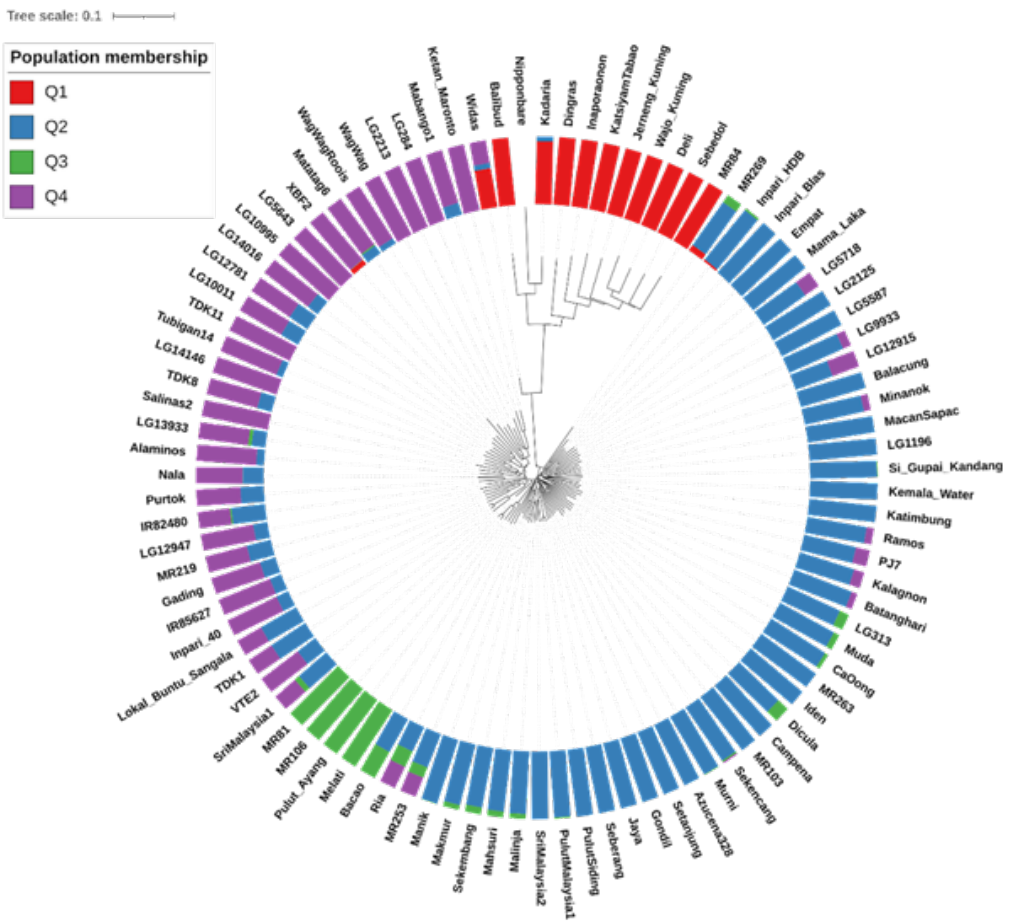

Figure 1. Model-based ancestry estimate of 92 rice varieties (circular bar plots) arranged according to phylogenetic relationships with other varieties.

Maronto and sticky rice from Laos share the same ancestry, and due to the low level of cross-pollination in rice, a relatively high degree of genetic purity was maintained, even after long-standing periods of cultivation in the respective countries. This highlights the potential influence of migration and genetic exchange on the distribution and diversity of rice varieties across different regions.

In certain varieties of the collection, varying admixture levels of the four ancestral populations were detected. Notably, in improved varieties with complex pedigrees, e.g., *Ria* and *MR253,* the genetic admixture from different ancestral populations was, as expected, complex as well. Surprisingly, genetic admixtures were also observed in a few traditional varieties such as *Gading* and *Lokal Buntu Sangala*, which contain genetic contributions from distinct groups of *indica* rice. This finding suggests the possibility of natural cross-pollination or deliberate hybridization by traditional farmers. However, it must be emphasized that *indica* has rarely been crossed with *japonica* rice to generate widely adopted new rice varieties, as the limited number of varieties with genetic influence of the Q1 population clearly shows.

The varieties were also grouped by the average hierarchical clustering method based on the Euclidean distance of their quantitative traits. According to the elbow method proposed by Shi et al. (2021), the optimal number of clusters for the given dataset was four, and the clustering results were represented in a circular dendrogram (Figure 2). Cluster II, the largest of the four, comprises 72 traditional varieties, followed by Cluster IV, with 18 varieties. Cluster I comprises only *Kalagnon*, a traditional variety from the Philippines, and Cluster III only *Matagtag*, an improved variety from the same region.

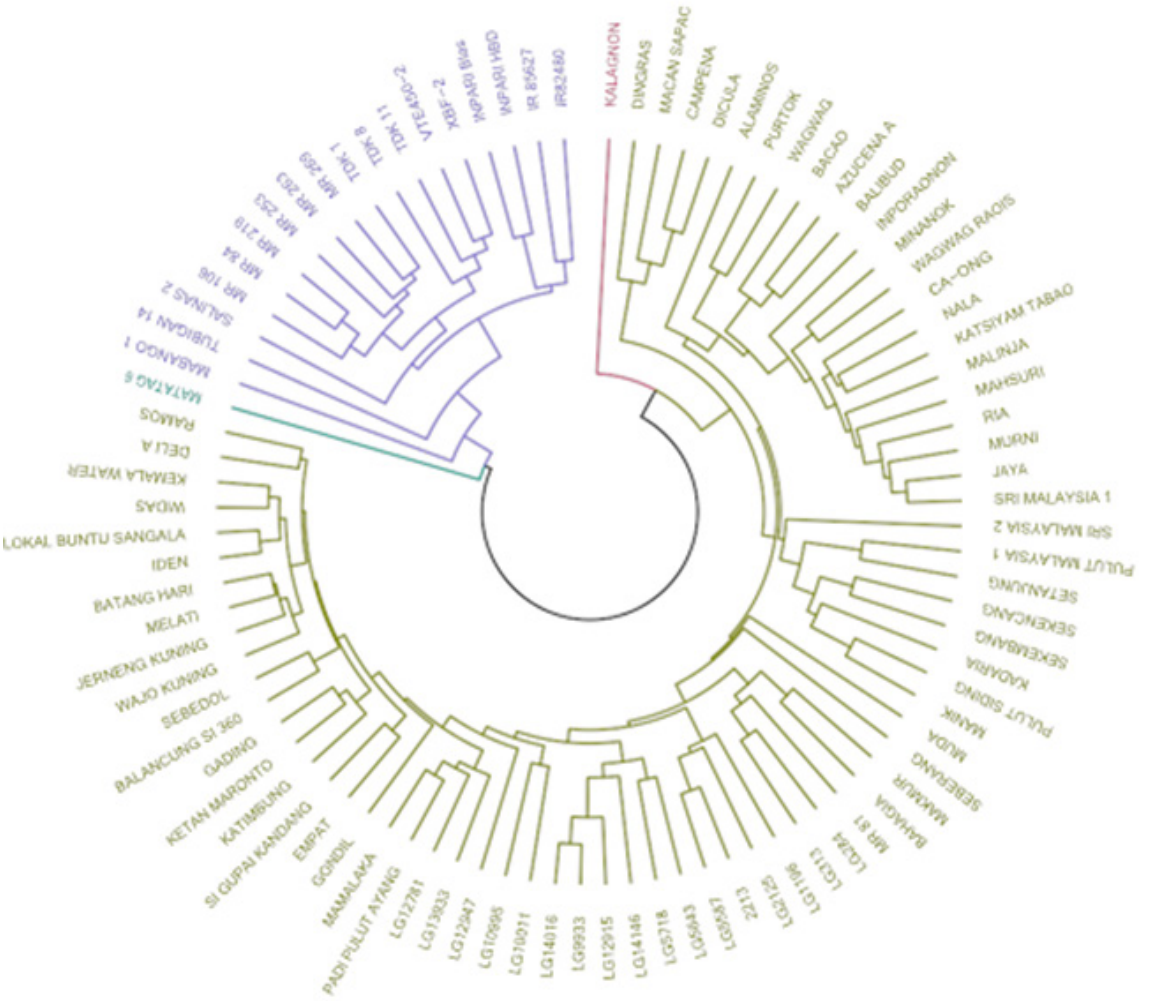

Figure 2. Circular dendrogram of clusters based on the Euclidean distance of the quantitative morphology traits of rice varieties.

The intra-cluster distances of clusters I, II, III, and IV were calculated (0, 4.825, 0, and 3.547, respectively). The highest inter-cluster distance (10.665) was measured between clusters II and IV, followed by clusters II and III. Conversely, the inter-cluster distance was lowest (6.575) between cluster I and II. After examining the clustering patterns and their origin, no discernible pattern was observed, as also reported by Ranjith et al. in 2018. To achieve a substantial heterotic effect and maximize variability, selecting parents from two clusters with a more considerable inter-cluster distance is recommended, as suggested by Mishra et al. (2003) and Gour et al. (2017).

The selection and preference of parents, with a view to improving specific traits, depend on the contributions of those traits to diversity. The calculation of the contribution to diversity (Table 2), demonstrated that spikelet fertility and plant height significantly influenced the manifestation of genetic diversity. This result is in line with a related study of Ranjith et al. (2018), in which other traits contributed minimally to the overall diversity.

Genetic diversity

Analysis of variance revealed significant variation among varieties for all studied traits. To assess the genetic variability, the genotypic coefficient of variation (GCV), phenotypic coefficient of variation (PCV) and broad-sense heritability (H^2)

were estimated for each trait (Table 1). High heritability (H^2) and selection gains (SG) were found for culm diameter, flag-leaf length, plant height and leaf-blade length, assuming a proportion of 5 % of selected plants, so that K=2.063 in equation (5). The results suggested that these traits were influenced by additive gene action, in agreement with previous studies of Tandekar et al. (2010), Ranjith et al. (2018), and Lipi et al. (2020).

Principal component analysis was performed to examine the contribution of a linear combination of traits to total variation. Eigenvalues were used to determine the relevance and contribution of each component to overall variance. At the same time, the coefficients of the eigenvectors indicated the degree of contribution of each component to each trait. Significantly higher coefficients have a stronger discriminatory effect on varieties. Although there are no definitive rules or conventional tests to determine the appropriate number of components and the cut-off limit for coefficients,

Table 1. Heritability and diversity measure of the 88 chosen rice varieties[1]

Traits	**Mean**	**H^2**	**PCV** (%)	**GCV** (%)	**SG** (%)
Caryopsis length	6.167	0.519911	11.6609	84.081	12.50722
Caryopsis width	2.237	0.514804	15.0153	10.7734	15.94684
Culm diameter	5.203	0.066826	38.2476	9.8873	5.27288
Culm number	17.053	0.467155	30.1984	20.6402	29.10339
Flag leaf length	30.996	0.486567	20.7469	14.4718	20.82544
Flag leaf width	1.583	0.200773	65.9676	29.5586	27.32342
Grain length	9.043	0.255139	33.3535	16.8473	17.55566
Grain width	2.589	0.103746	33.4486	10.7736	7.15890
100-grain weight	2.493	0.116438	41.3533	14.111	9.93358
Plant height	110.576	0.813701	21.4453	19.3448	35.99943
Leaf blade length	49.46	0.615614	17.2304	13.5192	21.88282
Leaf blade width	1.278	0.105457	37.3424	12.1266	8.12411
Panicle length	26.186	0.507352	10.7556	7.6611	11.25751
Spikelet fertility	94.73	0.175659	25.5897	10.7251	9.27334
Panicle number	16.98	0.210295	30.247	13.8707	13.12234

[1] Four of the initial 92 varieties were excluded from analysis.

Table 2. The first six principal components and contribution of characters to the diversity of 92 rice varieties in this study

	Components						**Contribution to Diversity** (%)
Parameters	**PC1**	**PC2**	**PC3**	**PC4**	**PC5**	**PC6**	
Proportion of variance	39.27	13.89	9.74	6.75	5.26	4.88	
Cumulative proportion (%)	39.27	53.16	62.90	69.65	74.91	79.79	
Traits							
Caryopsis length (CL)	0.202	-0.308	-0.428	0.018	0.128	0.122	0
Caryopsis width (CW)	-0.146	0.275	-0.026	-0.438	0.152	0.554	0
Culm diameter (SD)	-0.215	-0.316	-0.010	-0.046	0.179	0.020	0
Culm number (CN)	0.188	0.316	0.083	0.187	0.330	0.075	0.76
Flag leaf length (FLL)	-0.280	-0.078	0.191	-0.116	0.008	-0.197	0.62
Flag leaf width (FLW)	-0.073	-0.367	0.010	0.282	-0.607	-0.138	0
Grain length (GL)	0.097	-0.243	-0.500	0.074	0.218	-0.194	0.02
Grain width (GW)	0.313	-0.226	0.215	-0.117	0.028	-0.047	0.36
100-grain weight (GWT)	0.317	-0.236	0.218	-0.098	-0.030	-0.054	0.04
Plant height (PH)	-0.332	-0.071	0.013	-0.076	0.250	-0.111	47.92
Leaf blade length (LBL)	-0.341	-0.028	0.023	0.102	0.105	-0.210	0.78
Leaf blade width (LBW)	0.277	-0.210	0.329	0.003	-0.151	-0.101	0
Panicle length (PL)	0.148	-0.385	0.337	-0.052	0.315	-0.175	0.93
Spikelet fertility (F)	-0.043	-0.052	0.265	0.570	0.326	0.310	47.42
Panicle number (PN)	0.290	0.219	0.110	-0.066	0.204	-0.082	1.55

some scholars use criteria such as: eigenvalue > 1, variance contribution > 4%, or coefficients > 0.30, to discriminate among varieties (Sanni et al. 2012, Sharma et al. 2014, Sahu et al. 2021, Han et al. 2022, Almarri et al. 2023, Khan et al. 2023). Based on these criteria, six components were selected as discriminating components for the varieties, which together accounted for 84.1% of the total variation (Table 2)

The first and second components explained 52.35% of the total variation. Grain width, 100-grain weight, leaf blade length and plant height represented the first principal component (PC1), which accounted for 39.27% of the total variation. The second principal component (PC2), with caryopsis length, culm number, culm diameter, flag-leaf width and panicle length, explained 13.69% of the overall variation. The third principal component (PC3) was primarily influenced by caryopsis length, grain length, leaf-blade width and panicle length. Caryopsis width and spikelet fertility determined the fourth principal component (PC4). The fifth principal component (PC5) was distinguished by culm number, spikelet fertility, panicle length and flag-leaf width, whereas PC6 was influenced by caryopsis width and spikelet fertility. The results of principal component analysis showed that spikelet fertility and plant height were significant factors in several components, underlying the calculation of the contribution of these traits to diversity. Spikelet fertility represents reproductive traits and plant height vegetative growth traits.

The distribution of rice varieties and the extent of phenotypic variance among the 92 varieties were shown by the first and second principal components. The biplots of PC1 and PC2 for the 92 varieties indicated that improved varieties were predominantly located on the right side of the PC1 axis, while most traditional varieties were on the left (Figure 3A). Furthermore, the biplot suggested more significant variation in the improved than the traditional varieties, in traits primarily influenced by the second principal component (PC2), namely caryopsis length, culm diameter and number, flag-leaf length and panicle number. However, no clear grouping based on PC2 was observed (Figure 3B).

The population structure analysis revealed a long-standing exchange of rice genetic resources for breeding and crossbreeding among different sources. These practices can be further enhanced through collaborative development and genetic exchange of rice varieties among Southeast Asian countries. The information provided in this study could contribute to a successful joint development of new rice varieties in Southeast Asian countries. Genetic diversity plays a vital role in determining the potential for improved hybrid development and the desired frequency of recombination in subsequent generations. In addition, genetic distance plays a crucial role, as optimal parental diversity is necessary to obtain superior varieties in a segregating population. In breeding programs involving genetically diverse parents from different groups, genes of diverse nature can be combined, resulting in promising hybrid derivatives, owing to the complementary interaction of different genes in the parents.

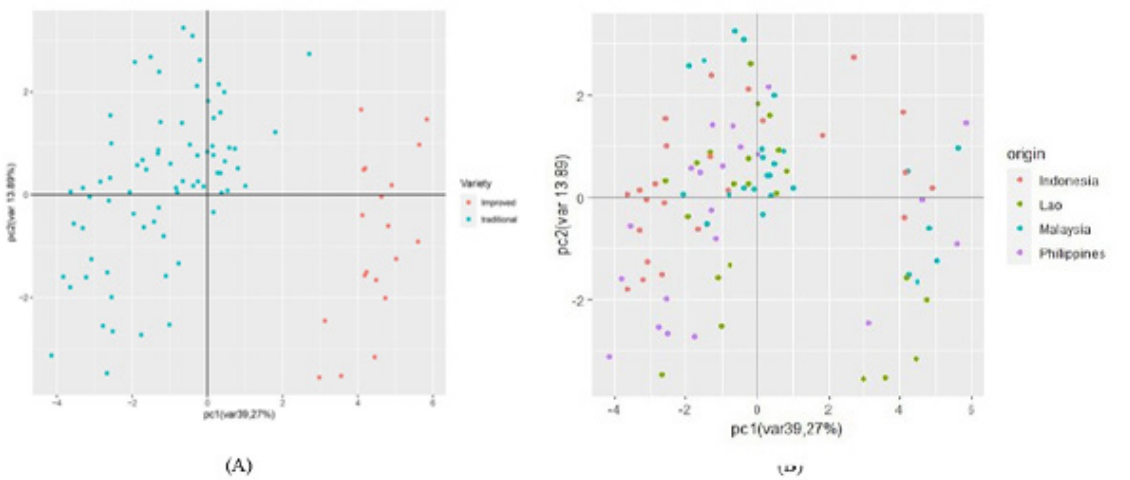

Figure 3. Biplots of the distribution of 92 rice varieties on the first and second principal components, colored according to their types (A) and country of origin (B).

CONCLUSION

The 73 traditional and 19 improved rice varieties from four Southeast Asian countries were classified into four clusters, based on SNP markers and morphological traits. Population structure analysis revealed their ancestral origins from four distinct populations. Some traditional varieties sampled from different countries shared a common ancestor, indicating a historical exchange of genetic resources. The study identified significant variation in 13 morphological variables among the varieties. Diversity analysis provided essential metrics such as broad-sense heritability, genetic coefficient of variation and phenotypic coefficient of variation. Notably, spikelet fertility and plant height contributed significantly to the observed diversity. Based on these diversity contributions and inter-cluster distances within the various clusters, several parents with desirable traits could be identified for future breeding, explicitly for the improvement of spikelet fertility and plant height.

ACKNOWLEDGMENTS

The authors would like to express their sincere gratitude to the Indonesian Center for Agricultural Biotechnology and Genetic Resources and Development (IAARD-Indonesia), the Rice Research Center of Malaysian Agricultural Research and Development Institute (MARDI-Malaysia), the Rice Research Center of National Agriculture and Forestry Research Institute (NAFRI-Lao PDR), and The Philippine Rice Research Institute (PhilRice-the Philippines) for their invaluable support in providing research facilities and human resources for this study. The authors also extend their gratitude to the Benefit Sharing Fund (BSF) of the ITPGRFA for funding this study through the BSF project no 3B-PR-08-Indonesia.

REFERENCES

Almarri NB, Alghamdi SS, ElShal MH and Afzal M (2023) Estimating genetic diversity among durum wheat (*Triticum durum* desf.) landraces using morphological and SRAP markers. Journal of the Saudi Society of Agricultural Sciences 22: 273-282.

Apraku BB, Oliveira ALG, Petroli SD, Hearne S, Adewale SA and Gedil M (2021) Genetic diversity and population structure of early and extra-early maturing maize germplasm adapted to sub-Saharan Africa. BMC Plant Biology 21: 96.

Bellwood P (2011) The checkered prehistory of rice movement southwards as a domesticated cereal from the Yangzi to the equator. Rice 4: 93-103.

Bioversity International (2007) Descriptors for wild and cultivated rice (***Oryza spp***). Bioversity International, Rome, 63p.

Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y and Buckler ES (2007) TASSEL: Software for association mapping of complex traits in diverse samples. Bioinformatics 23: 2633-2635.

Casañas F, Simó J, Casals J and Prohens J (2017) Toward an evolved concept of landrace. Frontiers in Plant Science 8: e0121381.

Epe IA, Bir MSH, Choudhury AK, Khatun A, Akhtar MM, Arefin MS, Islam MA and Park KW (2021) Genetic diversity analysis of high-yielding rice (*Oryza sativa*) varieties cultivated in Bangladesh. Korean Journal of Agricultural Science 48: 283-297.

Falconer DS (1989) Introduction to quantitative genetics. Longman, Scientific and Technical Group, Essex, 438p.

Gour L, Maurya SB, Koutu GK, Singh SK, Shukla SS and Mishra DK (2017) Characterization of rice (*Oryza* et al.) genotype using principal component analysis including scree plot and rotated component matrix. International Journal Chemistry Standard 5: 975-983.

Gutaker RM, Groen SC, Bellis ES, Choi JY, Pires IS, Bocinsky RK, Slayton ER, Wilkins O, Castillo CC, Negrão S, Oliveira MM, Fuller DQ, Guedes JAD, Lasky JR and Purugganan MD (2020) Genomic history and ecology of the geographic spread of rice. Nature Plants 6: 492-502.

Han B, Huang S, Huang G, Wu X, Jin H, Liu Y, Xiao Y and Zhou R (2022) Genetic relatedness and association mapping of horticulturally valuable traits for the Ceiba plants using ddRAD sequencing. Horticultural Plant Journal 9: 826-836.

Hour AL, Hsieh WH, Chang SH, Wu YP, Chin HS and Lin YR (2020) Genetic diversity of landraces and improved rice varieties (*Oryza sativa L.*) in Taiwan. Rice 13: 82.

Islam MZ, Khalequzzaman M, Prince FRK, Siddique MA, Rashid ESMH, Ahmed MSU, Pittendrigh BR and Ali MP (2018) Diversity and population structure of red rice germplasm in Bangladesh. **Plos One 13**: e0196096.

Khan MAR, Mahmud A, Islam MN, Ghosh UK and Hossain MS (2023) Genetic variability and agronomic performances of rice genotypes in different growing seasons in Bangladesh. Journal of Agriculture and Food Research 14: 100750.

Lahkar L and Tanti B (2017) Study of morphological diversity of traditional aromatic rice landraces (*Oryza sativa* L.) collected from Assam, India. Annals of Plant Sciences 6: 1855-1861.

Lestari P, Utami DW, Rosdianti I and Sabran M (2017) Morphological variability of Indonesian rice germplasm and the associated SNP markers. Emirates Journal of Food and Agriculture 28: 660-667.

Lipi LF, Hasan MJ, Akter A, Quddus MR, Biswas PL, Ansari A and Akter S (2020) Genetic variation, heritability, and genetic advance in some promising rice hybrids. SAARC Journal of Agriculture 18: 39-49.

Mishra L, Sarawgi A and Mishra R (2003) Genetic diversity for morphological and quality traits in rice. Advance in Plant Science 16: 287-293.

Morales KY, Singh N, Perez FA, Ignacio JC, Thapa R, Arbelaez JD, Tabien RE, Famoso A, Wang DR, Septiningsih EM, Shi Y, Kretzschmar T, McCouch SR and Thomson MJ (2020) An improved 7K SNP array, the C7AIR, provides a wealth of validated SNP markers for rice breeding and genetics studies. Plos One 15: e0232479.

Mutert BR and Fairhurst TH (2002) Developments in rice production in Southeast Asia. Better Crops International 15: 12-17.

Nachimuthu VV, Muthurahan R, Duraialaguraja S, Sivakami R, Pandian BA, Ponniah G, Gunasekaran K, Swaminathan M, Suji KK and Sabariappan R (2015) Analysis of population structure and genetic diversity in rice germplasm using SSR markers: An initiative towards association mapping of agronomic traits in *Oryza sativa*. Rice 8: 30.

Ranjith P, Sahu S, Dash SK, Bastia DN and Pradhan BD (2018) Genetic diversity studies in rice (*Oryza sativa* L.). Journal of Pharmacognosy and Phytochemistry 7: 2529-2531.

Rashid M, Imran S, Islam M and Hassan L (2018) Genetic diversity analysis of rice landraces (*Oryza sativa* L.) for salt tolerance using SSR markers in Bangladesh. Fundamental and Applied Agriculture 3: 460-466.

Sahu LP, Vageeshvari, Chaudhari P and Gauraha D (2021) Diversity analysis of rice (*Oryza sativa* L.) germplasm accessions using principal component analysis. The Pharma Innovation Journal 10: 212-215.

Sanni KA, Fawole I, Ogunbayo SA, Tia DD, Somado EA, Futakuchi K, Sié M, Nwilene FE and Guei RG (2012) Multivariate analysis of the diversity of landrace rice germplasm. Crop Science 52: 494-504.

Saracli S, Doğan N and Doğan I (2013) Comparison of hierarchical cluster analysis methods by cophenetic correlation. Journal of Inequalities and Applications 203: 1-8.

Sharma SK, Singh J, Chauhan MS and Krisnamurthy SL (2014) Multivariate analysis of phenotypic diversity of rice (*Oryza sativa*) germplasm in North-West India. Indian Journal of Agricultural Sciences 84: 295-299.

Shi C, Wei B, Wei S, Wang W, Liu H and Liu J (2021) A quantitative discriminant method of elbow point for the optimal number of clusters in the clustering algorithm. EURASIP Journal on Wireless Communications and Networking 31.

Sweeney M and McCouch S (2007) The complex history of the domestication of rice. Annals of Botany 100: 951-957.

Tandekar K, Kavita A and Pushpalata T (2010) Genetic variability, heritability, and genetic advance for quantitative trait in rice (*Oryza sativa* L.) accession. Agriculture and Biology Research 26: 13-19.

Thomson MJ, Septiningsih EM, Suwardjo F, Santoso TJ, Silitonga TS and McCouch SR (2007) Genetic diversity analysis of traditional and improved Indonesian rice (*Oryza sativa* L.) germplasm using microsatellite markers. Theoretical and Applied Genetics 114: 559-568.

Zaman M, Paul D, Kabir M, Mahbub M and Bhuiya M (2005) Assessment of character contribution to divergence for some rice varieties. Asian Journal of Plant Sciences 4: 388-391.