JARQ 57 (1), 21-35 (2023) https://www.jircas.go.jp

Diversity of Traits Related to Panicle Architecture and Grain Size in Cambodian Rice Germplasm and Newly Developed Mini-core Collection

Hiroki SAITO[1]*, Chhourn ORN[2], Vathany THUN[2], Makara OUK[3],
Asami TOMITA[1†], Kazuhiro SASAKI[4], Mitsuhiro OBARA[4],
Nobuya KOBAYASHI[1] and Yoshimichi FUKUTA[5]

[1] Tropical Agriculture Research Front, Japan International Research Center for Agricultural Sciences, Ishigaki, Japan
[2] Plant Breeding Division, Cambodian Agricultural Research and Development Institute (CARDI), Phnom Penh, Cambodia
[3] Plant Breeder, Former Director of CARDI
[4] Biological Resources and Post-harvest Division, Japan International Research Center for Agricultural Sciences, Tsukuba, Japan
[5] Research Planning and Partnership Division, Japan International Research Center for Agricultural Sciences, Tsukuba, Japan

Abstract

Cambodia has a wealth of rice landraces (traditional varieties) adapted to different ecosystems. These are expected to include diverse genetic resources that could be utilized in the genetic improvement of yield. A detailed investigation of panicle architecture and grain size in 173 Cambodian accessions revealed a wide diversity in these traits. The diversity in improved cultivars (grown mostly in irrigated fields) was clearly reduced in most traits. Although diversity was reduced in the improved cultivars, the superior traits had contributed increasing productivity. On the other hand, the landraces conserved wide diversities. More than 80% of rice cultivation in Cambodia is still rainfed, with uncontrolled and limited access to water and nutrients. Therefore, it is difficult for uniform improved cultivars to adapt to such heterogeneous growing conditions. Although landraces may be inferior to improved cultivars in overall yield performance under favorable conditions, many are superior to improved cultivars in various characteristics. This indicates their value as resources for breeding for an increased yield. To efficiently identify and evaluate them, we developed a mini-core collection of 53 accessions that maintain the diversity of all accessions. The mini-core collection will be valuable for further detail genetic analysis and genetic yield improvement.

Discipline: Crop Science
Additional key words: breeding, improved variety, landrace, yield

Introduction

The grain yield of rice per unit of cultivated area is a complex trait mainly determined by four components: grain number per panicle, panicle number per unit of area, grain weight, and spikelet fertility. Of these, the grain number per panicle is determined by the length of the panicle, number of rachis branches, and number of grains formed on the branches (collectively referred to as panicle architecture). Grain weight is mainly determined by grain length, width, and thickness (collectively referred to as grain size). Recent molecular genetic studies have demonstrated many genes that contribute to panicle architecture (Yin et al. 2021) or grain size (Li et al. 2018). Because many of these genes have been identified from studies using populations based on

---

Present address:
[†] Faculty of Agriculture, Okayama University, Okayama, Japan
*Corresponding author: hirokisaito@affrc.go.jp
Received 22 February 2022; accepted 26 April 2022.

biparental crossing, these results only partially explain the natural variation in panicle architecture and grain size. Therefore, it is necessary to explore diverse genetic resources for the genetic improvement of yield.

In Cambodia, rice is the staple food, providing 65% of the population's total caloric intake (Maltsoglou et al. 2010). Not only it is directly important to the Cambodian people, but it is also indirectly an important crop that supports the rural economy, as rice agriculture employed about 37% of the total labor force in 2017 (MAFF 2019) and contributed about 22.8% to the GDP in 2020 (MAFF 2021). Therefore, stable rice production and increased yield are crucial for food security in Cambodia.

Cambodia's rainfall distribution, flooding patterns, topography, and soil types create very diverse rice-growing environments, which can be classified into four ecosystems: irrigated, upland, rainfed lowland, and deepwater (Ouk et al. 2001, 2017). In addition, being close to the center of origin of rice, Cambodia has a wealth of landraces (traditional varieties) adapted to different ecosystems. For hundreds of years, natural selection pressures, such as drought, submergence, flooding, nutrient stresses, and biotic stresses, in addition to artificial selection, have significantly contributed to the evolution of landraces adapted to different environments; over 2000 landraces have been identified as unique to Cambodia (Nesbitt 1997). However, the civil war in Cambodia from 1970 to 1979 severely damaged national agriculture, and much rice germplasm was lost during that time (Javier 1997).

To recover the lost genetic diversity, the Cambodian government retrieved 756 Cambodian accessions from the International Rice Research Institute (IRRI) Genebank (Sahai et al. 1992a), began the collection and evaluation of remaining landraces, and imported breeding lines from IRRI, Thailand, Vietnam, India, and Africa. Orn et al. (2020) collected a wide and diverse set of Cambodian germplasms corresponding to the various rice ecosystems in Cambodia and identified diversity in blast resistance and heading date. The resultant genetic variation in this germplasm set is similar to that in Bangladesh and higher than that in West Africa (Orn et al. 2020). Although the Cambodian germplasm is suitable for assessing a wide range of characteristics adapted to diverse environments, further analyses of yield-related traits, such as panicle architecture and grain size, have not been conducted.

An early study of Cambodian rice germplasm analyzed mainly seed morphology (Hamada 1965). Lando & Mak (1994) broadly described the general characteristics and grouped accessions by date of flowering or harvest. Other studies conducted later by the Cambodia–IRRI–Australia Project (CIAP) and the Cambodian Agriculture Research and Development Institute (CARDI) characterized heading dates and several morphological traits in 2842 accessions (Sahai et al. 1992a, b, Javier et al. 1999, Ouk et al. 2017). Although several studies have evaluated morphological characteristics, no studies have focused on panicle architecture or grain size, which are highly influenced by environmental factors (e.g., climatic conditions, soil texture and fertility, and human management). Choi et al. (2013) reported that temporary lack of sunshine (10-14 days) harms panicle morphogenesis (differentiated spikelet number, branch number, and panicle length). Adriani et al. (2016) reported that the panicle architecture of a large-panicle genotype is prone to genotype × environment interaction. Therefore, multi-site trials using large-scale genetic resources are necessary to search for useful genetic resources and to elucidate their genetic factors. However, such trials are impractical as they require a great deal of labor.

To efficiently identify and evaluate large-scale valuable genetic resources, Frankel (1984) proposed the concept of a core collection, defined as a limited set of accessions representing, with a minimum of duplication, the genetic diversity of a crop species and its wild relatives. The development of a core collection can facilitate the characterization and evaluation of genetic resources, reduce costs, and make resources available for other activities. To date, various types of core collections have been developed in rice on the basis of DNA polymorphisms (and phenotypic variations), such as World Rice Core Collection (Kojima et al. 2005), 3k Rice Genome (Kumar et al. 2020), USDA rice germplasm (Yan et al. 2007), wild rice (*Oryza rufipogon* Griff.) (Liu et al. 2015), Chinese accessions (Li et al. 2011), and *Oryza glaberrima* (Ndjiondjop et al. 2017). However, core collections are still too large for practical assessment. Upadhyaya and Ortiz (2001) proposed a smaller cultivar set, called a "mini-core collection" or "micro-core collection," to practically evaluate the variation. Ebana et al. (2008) developed a mini-core collection of Japanese rice landraces, which has been used for screening for different phenotypes and allele mining for target genes, such as genes involved in boron toxicity (Ochiai et al. 2011), heading date (Fujino et al. 2013), and eating quality (Iijima et al. 2019).

Here, we performed a detailed investigation of panicle architecture and grain size in 173 Cambodian rice accessions widely collected from diverse geographical regions, covering both the landraces and modern improved cultivars. Through genotyping by simple sequence repeat (SSR) markers, we selected a small number of accessions ("Cambodian mini-core collection") with most of the overall genetic diversity. We discuss the diversity of rice genetic resources in Cambodia for panicle architecture and grain size and the potential utility of the mini-core collection for breeding new high-yielding cultivars.

Materials and methods

1. Plant materials

We used 173 accessions collected by Orn et al. (2020) (Tables 1 and S1). A total of 15 improved cultivars (I) and 158 landraces (L) were collected from three ecosystems (rainfed lowland, RL; upland, UL; and irrigated lowland, IL). Their genotypes were previously analyzed by 63 SSR markers (Orn et al. 2020). As external controls, we used two representative cultivars in the tropic Asia, IR 64 (improved variety) and Kasalath (local landrace). All were grown at the Tropical Agricultural Research Front, Japan International Research Center for Agricultural Sciences (JIRCAS), Ishigaki, Japan (24°38′N, 124°20′E), in the second season from July to November 2017. Ten 28-day-old seedlings per line were transplanted, 1 per hill, 18 cm apart in rows 30 cm apart. An organic slow-release fertilizer (5.2 g N, 5.2 g P_2O_5, 5.2 g K_2O per m^2) was applied as a basal fertilizer.

Table 1. Accessions of Cambodian germplasm used in this study and their categorizations

Cultivar type	Ecosystem			Total
	Irrigated (IL)	Rainfed lowland (RL)	Upland (UL)	
Local	0	150	8	158
Improved	12	3	0	15
Total	12	153	8	173

2. Investigation of panicle architecture and grain size

Three average individuals were selected from each line, and a single panicle was collected from the main culm of each. We recorded the panicle length (PL), number of primary branches (PB), number of secondary branches (SB), and number of grains per panicle (GN) as traits related to panicle architecture. We assessed the grain size using the SmartGrain software (Tanabata et al. 2012) with default settings, measuring grain length (GL), and grain width (GWd) of each filled grain, from which we calculated the ratio of GL to GWd (LWR = GL/GWd). We also measured 1,000-filled-grain weight (GWt) as a trait related to grain size. To determine whether bias among cultivar types or ecosystem categories was observed in each trait, four parameters (mean difference [MD%], variance difference [VD%], coincidence rate of the range of each trait [CR%], and variable rate of the coefficient of variance [VR%]) against all accessions (Hu et al. 2000) were calculated and compared after classification based on cultivar types and ecosystem categories. Principal component analysis was conducted using the "prcomp" function in "R v. 4.0.1."

3. Selection and evaluation of mini-core collection

The R package Core Hunter v. 3.2.1 (de Beukelaer et al. 2018) was used to select a subset of accessions from the 173 accessions that broadly represents the genetic diversity held in all accessions (Orn et al. 2020) using genotyping data of 63 SSR markers. A homogeneity test (F-test) of variance and a t-test of means ($\alpha = 0.05$) were conducted to identify the trait differences between the mini-core collection and all accessions. We further assessed their homogeneity by using Nei's gene diversity (H) for each molecular marker (Nei 1972), mean difference (MD%), variance difference (VD%), variable rate of the coefficient of variance (VR%), and coincidence rate of the range (CR%) (Hu et al. 2000). Chi-squared (χ^2) tests were employed to evaluate the similarity between the mini-core collection and all accessions originating from five geographic regions (Table S1). The genetic population structure of the mini-core collection was estimated in the structure software v. 2.3.4 (Pritchard et al. 2000). Ten runs were performed for each number of populations (K) set from 2 to 10. Each run used a burn-in of 50,000 cycles and Markov chain Monte Carlo replication number of 100,000. The most probable K-value was determined from the log probability of the data (ln $P(X|K)$) and delta K (ΔK) based on the rate of change in ln $P(D)$ between successive K-values.

Results

1. Characteristics of panicle architecture and grain size in Cambodian accessions

Eight traits related to panicle architecture and grain size showed wide continuous distributions (Fig. 1). Many accessions were superior to IR 64 or Kasalath in PB, SB, and GN, some in GWt, but few in PL (Fig. 1; Table 2). GN was highly positively correlated with SB (Table 3). Furthermore, GN was correlated positively with PB and negatively with GWt, GL, and GWd. GWt was also correlated negatively with SB but positively with GL and GWd.

No significant differences were observed between the landraces and all accessions in the means of any trait,

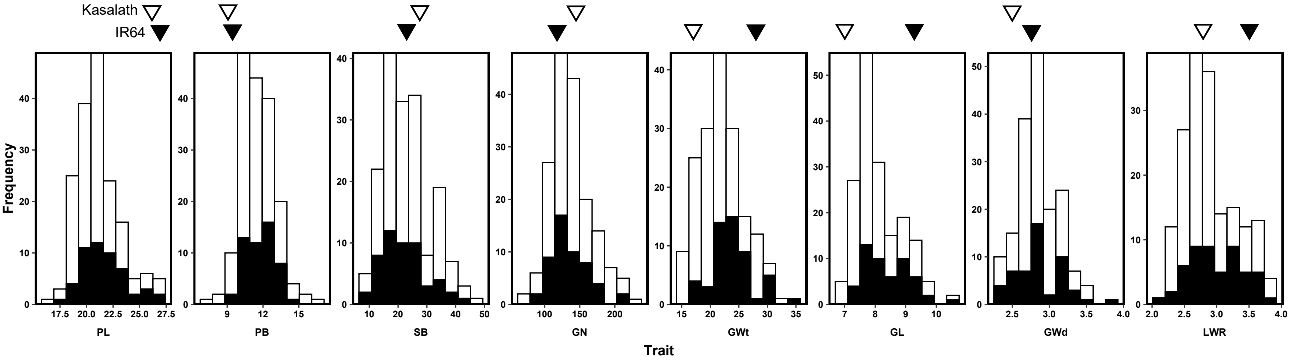

Fig. 1. Distribution of Cambodian accessions in eight traits related to panicle architecture and grain size
PL, panicle length (cm); PB, number of primary branches; SB, number of secondary branches; GN, number of total grains per panicle; GWt, grain weight (g); GL, grain length (cm); GWd, grain width (mm); LWR, ratio of GL to GWd. The white and black bars indicate all accessions and mini-core collections, respectively.

Table 2. Means and standard deviations of eight traits in Cambodian accessions (n = 173), IR64, and Kasalath

Line	Panicle architecture				Grain size			
	PL	PB	SB	GN	GWt	GL	GWd	LWR
Cambodian germplasm	21.1 ± 2.0	11.5 ± 1.4	22.7 ± 8.0	137.5 ± 29.6	22.1 ± 3.8	8.1 ± 0.7	2.8 ± 0.3	2.9 ± 0.4
IR64	27.3 ± 2.6	9.7 ± 1.2	22.7 ± 5.1	124.0 ± 18.2	27.3 ± 0.2	9.3 ± 0.5	2.7 ± 0.1	3.5 ± 0.2
Kasalath	26.2 ± 0.8	9.3 ± 1.2	27.0 ± 4.0	140.7 ± 17.6	17.1 ± 1.0	7.0 ± 0.5	2.5 ± 0.2	2.8 ± 0.2

Table 3. Pearson's correlation coefficients among eight traits related to panicle architecture and grain size

		Panicle architecture				Grain size			
		PL	PB	SB	GN	GWt	GL	GWd	LWR
Panicle architecture	PL	-							
	PB	0.034	-						
	SB	0.175 *	0.331 ***	-					
	GN	0.217 *	0.560 ***	0.936 ***	-				
Grain size	GWt	0.131	–0.057	–0.386 ***	–0.358 ***	-			
	GL	0.326 ***	–0.132	–0.218 **	–0.205 **	0.599 ***	-		
	GWd	–0.276 **	–0.021	–0.364 ***	–0.324 ***	0.345 ***	–0.073	-	
	LWR	0.445 ***	–0.071	0.107	0.089	0.170 *	0.722 ***	–0.734 ***	-

PL, panicle length; PB, number of primary branches; SB, number of secondary branches; GN, number of grains per panicle; GWt, 1,000-filled-grain weight; GL, grain length; GWd, grain width; LWR, ratio of GL to GWd.
*, **, and *** indicate the significant correlation at the 5%, 1%, and 0.1% levels, respectively.

because the landraces comprised more than 90% of all accessions (Table 4). Partially, the variance difference (VD%) of PL were only decreased in the landraces, whereas those of other seven traits and other three parameters (MD%, CR%, and VR%) were almost similar to that of all accessions (Table 4). On the other hand, improved cultivars had greater PL, SB, GN, GL, and LWR and narrower GWd than all accessions or the landraces (Fig. 2a; Table 4). However, the improved cultivars did not exhibit significant differences in PB or GWt between all accessions or landraces. Interestingly, the VD%, CR%, and VR% in the improved cultivars decreased in all traits, except VD% in SB, which means that the variance and range in all eight traits of improved cultivars were significantly reduced (Table 4).

Differences were also found among ecosystem origins. There was no significant difference between the rainfed lowland cultivars (RL) and all accessions in the means of any trait because the RL comprised more than 90% of all accessions (Table 5). Furthermore, the

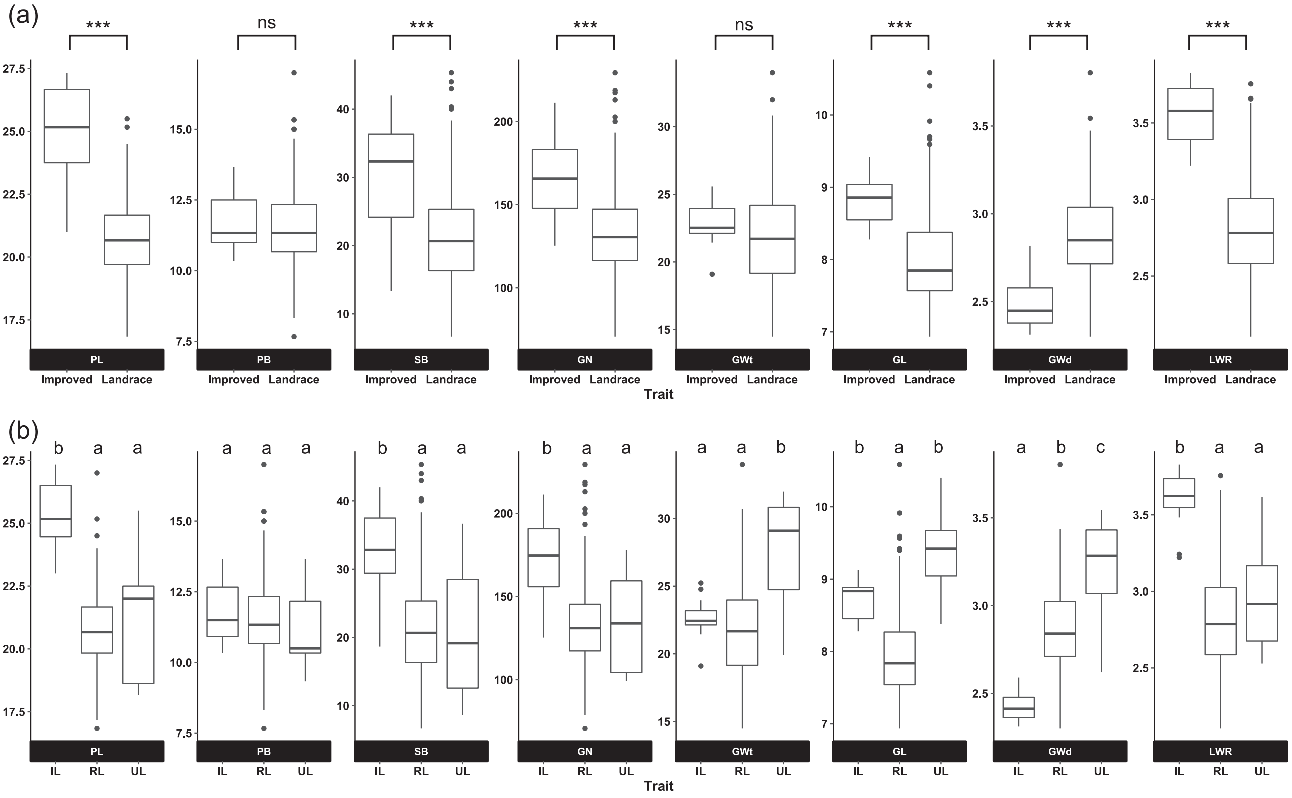

Fig. 2. Box plots of eight traits related to panicle architecture and grain size

(a) Comparisons between improved cultivars and landraces. Differences are significant at the ***0.1% level by Student's *t*-test; "ns," no significant difference

(b) Comparisons among irrigated lowland (IL), rainfed lowland (RL), and upland (UL) cultivars. Boxplots with the same letter are not significantly different at the 5% level by the Kruskal–Wallis multiple comparison test.

PL, panicle length; PB, number of primary branches; SB, number of secondary branches; GN, number of total grains per panicle; GWt, grain weight; GL, grain length; GWd, grain width; LWR, ratio of GL to GWd

variances of six traits except PL and GWd were not different from all accessions, because three parameters of each trait in the RL, except VD% in PL and GWd, were around 100% (Table 5). The irrigated lowland cultivars (IL) had significantly greater PL, SB, GN, GL, and LWR and narrower GWd than all accessions or RL cultivars, although there were no significant differences in PB and GWt (Fig. 2b; Table 5). The IL cultivars also had significantly greater PL, SB, GN, and LWR and lower GWd and GWt than upland cultivars (UL), although there were no significant differences in PB (Fig. 2b). As in the improved cultivars, the VD%, CR%, and VR% in the IL decreased in all traits, which means that the variances and ranges were significantly reduced. On the other hand, UL had unique characteristics of grain size, with greater GWt, GL, and GWd (Fig. 2b; Table 5). Furthermore, UL exhibited wide variations in PL, SB, and GN because of over 100% value in the VD% and VR% (Table 5).

To comprehensively evaluate the relationship among the eight traits, we conducted principal component analysis. Principal component 1 (PC1) explained 34.7% and PC2 explained 30.3% of the variance, 65% in total (Fig. 3). We interpreted the biological significance of the two PCs on the basis of their factor loadings. PC1 was highly positively correlated with PB, SB, and GN (Table 6). This means that PC1 represents traits related to the panicle architecture. PC2 was correlated highly positively with GWd and highly negatively with PL, GWt, GL, and LWR (Table 6). This means that PC2 represents traits related to grain size. The improved cultivars formed clusters in the bottom-right quadrant with positive values on PC1 and negative values on PC2, whereas the landraces were widely distributed in all quadrants (Fig. 3). These results indicate that the improved cultivars featured improved GN and GWt, leading to higher yield.

2. Development of mini-core collection

We selected 53 accessions as the mini-core collection (Table S1). There was no bias in origins between all accessions and mini-core collections (χ^2 test, data not shown). The means of Nei's genetic diversity showed no difference between all accessions and mini-core

Table 4. Basic statistics of Cambodian accessions summarized in 15 improved cultivars and 158 landraces in 8 traits

	Trait	Cultivar type	Mean	Variance	Standard deviation	Coefficient of variance	Maximum	Minimum	Range	Mean difference (MD%)$^{a, b}$	Variance difference (VD%)^{a}	Coincidence rate (CR%)^{a}	Variance rate (VR%)^{a}
Panicle architecture	PL	All accessions	21.1	4.1	2.0	0.10	27.3	16.8	16.8-27.3				
		Improved	25.0	3.5	1.9	0.08	27.3	21.0	21-27.3	118.2***	86.6	60.3	79.0
		Landraces	20.8	2.6	1.6	0.08	25.5	16.8	16.8-25.5	98.3^{ns}	64.3	82.5	81.8
	PB	All accessions	11.5	2.0	1.4	0.12	17.0	7.7	7.7-17				
		Improved	11.7	1.0	1.0	0.08	13.7	10.3	10.3-13.7	101.8^{ns}	47.9	35.7	68.2
		Landraces	11.5	2.1	1.5	0.13	17.0	7.7	7.7-17	99.8^{ns}	105.1	100.0	103.0
	SB	All accessions	22.7	64.4	8.0	0.35	45.3	6.7	6.7-45.3				
		Improved	30.1	69.5	8.3	0.28	42.0	13.3	13.3-42	132.8**	107.9	74.1	78.5
		Landraces	22.0	58.6	7.7	0.35	45.3	6.7	6.7-45.3	96.9^{ns}	91.0	100.0	98.7
	GN	All accessions	137.5	879.2	29.6	0.22	229.3	70.7	70.7-229.3				
		Improved	167.1	717.8	26.8	0.16	211.3	125.3	125.3-211.3	121.5***	81.6	54.2	74.6
		Landraces	134.7	807.5	28.4	0.21	229.3	70.7	70.7-229.3	98.0^{ns}	91.8	100.0	98.1
Grain size	GWt	All accessions	22.1	14.8	3.8	0.17	34.0	14.5	14.5-34				
		Improved	22.9	2.7	1.6	0.07	25.6	19.1	19.1-25.6	103.7^{ns}	18.4	33.3	41.4
		Landraces	22.0	15.9	4.0	0.18	34.0	14.5	14.5-34	99.6^{ns}	107.4	100.0	104.3
	GL	All accessions	8.1	0.5	0.7	0.09	10.6	6.9	6.9-10.6				
		Improved	8.8	0.1	0.4	0.04	9.4	8.3	8.3-9.4	108.7***	26.0	31.3	47.1
		Landraces	8.1	0.5	0.7	0.09	10.6	6.9	6.9-10.6	99.2^{ns}	97.4	100.0	99.8
	GWd	All accessions	2.8	0.1	0.3	0.09	3.8	2.3	2.3-3.8				
		Improved	2.5	0.0	0.2	0.07	2.8	2.3	2.3-2.8	87.9***	43.8	33.7	75.5
		Landraces	2.9	0.1	0.2	0.09	3.8	2.3	2.3-3.8	101.1^{ns}	87.8	100.0	92.9
	LWR	All accessions	2.9	0.2	0.4	0.14	3.8	2.1	2.1-3.8				
		Improved	3.5	0.0	0.2	0.06	3.8	3.2	3.2-3.8	123.0***	28.3	35.2	43.4
		Landraces	2.8	0.1	0.3	0.12	3.8	2.1	2.1-3.8	97.8^{ns}	77.3	95.8	90.2

^{a} Percentages indicate proportions against all accessions.

^{b} Differences are significant at the *5%, **1%, and ***0.1% levels by Dunnett's test between all accessions and either improved cultivars or landraces; "ns," no significant difference.

collections (Table 7). There were significant differences in mean GWt and GL between all accessions and mini-core collections but no difference in the variance of any traits (Table 8).

We employed model-based clustering to investigate genetic kinship within the mini-core collection and to reveal population structure. The maximum likelihood method indicated the optimal number of populations (K) to be 3 (Fig. 4a); the probability that $K = 3$ was by far the highest among the models that assumed $K = 2$-10. Thus, the mini-core collection comprised three subgroups (Fig. 4b). Group 1 (blue) comprised two upland cultivars (Rimke and Chkort) and one rainfed lowland cultivar (Romeas Lourt). Group 2 (red) included improved cultivars and accessions from IRRI and other countries, as well as several landraces. Group 3 (green) was largely landraces.

Discussion

The environmental diversity, topographic variation, and soil types in Cambodia have contributed to the development of diverse rice-growing systems and a wealth of landraces (Nesbitt 1997). The diversity of Cambodian accessions was clearly reflected in the traits related to panicle architecture and grain size, which showed a wide range (Fig. 1). In addition, the Cambodian accessions hold a wider and heavier GWt, 14.5 g-34.0 g (Fig. 1; Table 2), which is higher than in Bangladeshi cultivars at 7.7 g-28.3 g (Islam et al. 2016). This could be attributed to the suitability of tropical Japonica Group cultivars to upland culture in the mountains of Cambodia, unlike the predominance of lowland cultivars in Bangladesh. Therefore, the Cambodian accessions used in this study seem to be a valuable population that has retained its topographical and morphological diversities.

Cambodia was an advanced producer of rice until the 1960s, but the civil war of the 1970s and the political

Table 5. Basic statistics of Cambodian accessions based on the ecosystems, irrigated lowland (IL, n = 12), rainfed lowland (RL, n = 153), and upland (UL, n = 8) in eight traits

	Trait	Ecosystem	Mean	Variance	Standard deviation	Coefficient of variance	Maximum	Minimum	Range	Mean difference (MD%)[a, b]	Variance difference (VD%)[a]	Coincidence rate (CR%)[a]	Variance rate (VR%)[a]
Panicle architecture	PL	All accessions	21.1	4.1	2.0	0.10	27.3	16.8	16.8-27.3				
		IL	25.3	2.1	1.4	0.06	27.3	23.0	23-27.3	119.8***	51.2	41.3	59.9
		RL	20.8	2.7	1.6	0.08	27.0	16.8	16.8-27	98.4^{ns}	65.5	96.8	82.5
		UL	21.2	6.6	2.6	0.12	25.5	18.2	18.2-25.5	100.5^{ns}	161.6	69.8	126.8
	PB	All accessions	11.5	2.0	1.4	0.12	17.0	7.7	7.7-17				
		IL	11.8	1.2	1.1	0.09	13.7	10.3	10.3-13.7	101.9^{ns}	57.8	35.7	74.8
		RL	11.5	2.1	1.5	0.13	17.0	7.7	7.7-17	100.0^{ns}	103.6	100.0	102.1
		UL	11.2	2.1	1.5	0.13	13.7	9.3	9.3-13.7	96.9^{ns}	104.8	46.4	106.0
	SB	All accessions	22.7	64.4	8.0	0.35	45.3	6.7	6.7-45.3				
		IL	32.5	50.1	7.1	0.22	42.0	18.7	18.7-42	143.3***	77.7	60.3	61.7
		RL	22.0	55.8	7.5	0.34	45.3	6.7	6.7-45.3	97.0^{ns}	86.7	100.0	96.3
		UL	21.2	112.9	10.6	0.50	36.7	8.7	8.7-36.7	93.3^{ns}	175.3	72.4	142.4
	GN	All accessions	137.5	879.2	29.6	0.22	229.3	70.7	70.7-229.3				
		IL	172.7	679.2	26.1	0.15	211.3	125.3	125.3-211.3	125.6***	77.3	54.2	70.2
		RL	134.9	791.1	28.1	0.21	229.3	70.7	70.7-229.3	98.1^{ns}	90.0	100.0	97.0
		UL	134.6	1,078.7	32.8	0.24	178.0	99.3	99.3-178	97.9^{ns}	122.7	49.6	113.5
Grain size	GWt	All accessions	22.1	14.8	3.8	0.17	34.0	14.5	14.5-34				
		IL	22.6	2.5	1.6	0.07	25.2	19.1	19.1-25.2	102.6^{ns}	17.2	31.5	40.5
		RL	21.7	14.0	3.7	0.17	34.0	14.5	14.5-34	98.5^{ns}	94.4	100.0	99.0
		UL	27.6	18.1	4.2	0.15	32.0	19.9	19.9-32	125.1***	122.2	61.9	88.6
	GL	All accessions	8.1	0.5	0.7	0.09	10.6	6.9	6.9-10.6				
		IL	8.7	0.1	0.3	0.03	9.1	8.3	8.3-9.1	107.1*	16.0	23.2	37.5
		RL	8.0	0.5	0.7	0.08	10.6	6.9	6.9-10.6	98.6^{ns}	85.8	100.0	94.2
		UL	9.4	0.4	0.6	0.06	10.4	8.4	8.4-10.4	115.5***	67.9	55.3	71.5
	GWd	All accessions	2.8	0.1	0.3	0.09	3.8	2.3	2.3-3.8				
		IL	2.4	0.0	0.1	0.04	2.6	2.3	2.3-2.6	85.3***	13.6	18.5	43.4
		RL	2.9	0.1	0.2	0.08	3.8	2.3	2.3-3.8	100.5^{ns}	75.9	100.0	86.9
		UL	3.2	0.1	0.3	0.10	3.5	2.6	2.6-3.5	112.6***	141.3	61.5	105.9
	LWR	All accessions	2.9	0.2	0.4	0.14	3.8	2.1	2.1-3.8				
		IL	3.6	0.0	0.2	0.06	3.8	3.2	3.2-3.8	124.7***	26.5	35.2	41.4
		RL	2.8	0.1	0.3	0.12	3.8	2.1	2.1-3.8	97.9^{ns}	78.4	95.8	90.7
		UL	3.0	0.1	0.4	0.13	3.6	2.5	2.5-3.6	102.7^{ns}	94.3	63.2	94.8

[a] Percentages indicate proportions against all accessions.

[b] Differences are significant at the ***0.1% level by Dunnett's test between all accessions and each ecosystem type; "ns," no significant difference.

upheaval that followed devastated the rural areas, leading to the loss of many local traditional landraces and destruction of production systems. In 1987, CIAP was launched to revive rice cultivation and to implement new breeding programs, resulting in the development of many high-yielding improved cultivars (IRRI–Cambodia Project 1990, Ouk et al. 2009). The improved cultivars had greater PL and GN than landraces due to a greater SB (Fig. 2a; Table 4). They also had characteristics of long-grain cultivars with narrower GWd and longer GL (Fig. 2a; Table 4). It is noteworthy that the variances and ranges (VD%, CR%, and VR%) in all eight traits were much less in the improved cultivars than in the landraces (Table 4), and wide variations in seven traits except PL were certainly retained in the landraces (Table 4). These results indicate that the improved cultivars lost diversity in these traits while upgrading superior traits, such as large panicles and many long, narrow grains per panicle. The wide diversity in the seven traits conserved in the landraces indicates the value of the landraces as genetic resource for breeding for higher yields.

The ecosystem characteristics also demonstrated interesting tendencies. UL cultivars had very different grain size characteristics from other ecosystem cultivars

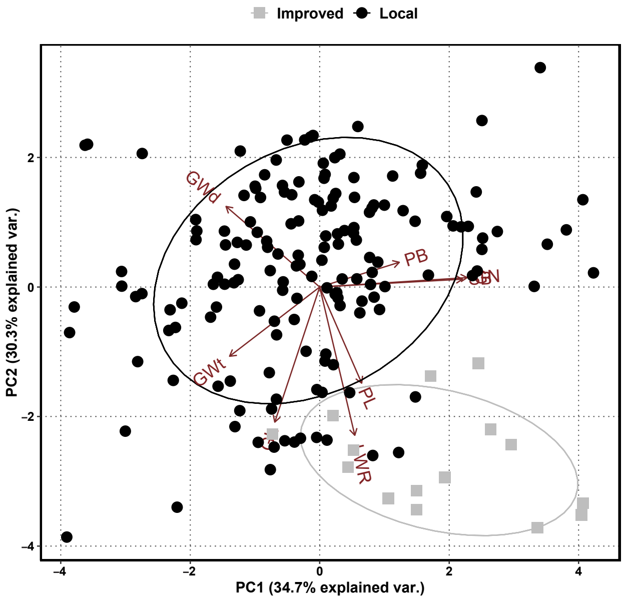

Fig. 3. Distribution of the Cambodian accessions based on the PCA of data collected from eight traits related to panicle architecture and grain size

Arrows indicate the relationship with PC1 and PC2 of eight traits: PL, panicle length; PB, number of primary branches; SB, number of secondary branches; GN, number of total grains per panicle; GWt, grain weight; GL, grain length; GWd, grain width; LWR, ratio between GL and GWd. Plots include 80% of variance among ● landraces and ■ improved cultivars.

(Fig. 2b; Table 5). Furthermore, although there were no significant differences in the means of PL, SB, and GN between UL cultivars and all accessions, the variances (VD% and VR%) of these traits in UL cultivars were wide, indicating that significant diversity is maintained in the UL cultivars, which are grown in heterogeneous and challenging environments. We consider that these curious characteristics are necessary for adaptation to challenging environments.

IL cultivars exhibited similar trends to those of the improved cultivars, with greater PL, SB, GN, GL, and LWR than all accessions (Tables 4, 5). The variances and ranges (VD%, CR%, and VR%) in all eight traits of IL cultivars were also significantly reduced (Table 5). These results are consistent with the use of these improved cultivars in irrigated cultivation. In fact, most of these improved cultivars are suitable for double cropping (wet and dry seasons) owing to their weak photoperiod sensitivity and early maturity (Orn et al. 2020). The construction of irrigation facilities to promote the utilization of unused natural resources, including water from the Pursat River and a large area of uncultivated land for the upland crops, stimulated the increase in rice production (Yagura 2021). Therefore, the strategic use of breeding for high yields in rice with infrastructure development contributed to the recovery of rice production in Cambodia.

The genetic diversity in Cambodian modern cultivars is lower than that in landraces (Orn et al. 2015). Here the variance in most traits related to panicle architecture and grain size was clearly narrow in IL cultivars (mostly improved cultivars) (Fig. 2a; Table 4). Furthermore, more than 80% of rice cultivation in Cambodia is still rainfed, with uncontrolled and limited access to water and nutrients. Therefore, it is difficult for uniform improved cultivars to adapt to such heterogeneous growing conditions. Although landraces may be inferior to improved cultivars in overall yield performance under favorable conditions, many are superior to improved cultivars in various characteristics. PB, which contributes to GN, was higher in some landraces than in improved cultivars (Fig. 2a; Table 4). In addition, many landraces had much heavier GWt, which is directly related to yield (Fig. 2a; Table 4). This indicates that many landraces have been selected and maintained with superior traits that are adapted to diverse rainfed environments.

The genetic diversity of the mini-core collection was representative of all accessions (Table 7). There were significant differences in the mean of GWt and GL from all accessions, but the variance of the eight traits related to panicle architecture and grain size was clearly similar (Table 8). The mean value of the VR% of the eight traits was 102.5%, and the CR% was 87.1% (Table 8). As a representative selection from all accessions, Hu et al.

Table 6. Pearson's correlation coefficients between principal components 1 (PC1) and 2 (PC2) and eight traits related to panicle architecture and grain size

	Panicle architecture				Grain size			
	PL	PB	SB	GN	GWt	GL	GWd	LWR
PC1	0.264**	0.501***	0.897***	0.927***	–0.566***	–0.282**	–0.588***	0.221
PC2	–0.604***	0.155	0.049	0.059	–0.435***	–0.848***	0.506***	–0.931***

PL, panicle length; PB, number of primary branches; SB, number of secondary branches; GN, number of grains per panicle; GWt, 1,000-filled-grain weight; GL, grain length; GWd, grain width; LWR, ratio of GL to GWd
*, **, and *** indicate the significant correlation at the 5%, 1%, and 0.1% levels, respectively.

Table 7. Numbers of alleles at 63 SSR markers and values of Nei's genetic diversity index in all accessions and the selected mini-core collection

Chromosome	Marker	Number of alleles		Nei's genetic diversity index	
		All accessions	Mini-core	All accessions	Mini-core
1	RM495	3	3	0.468	0.581
	RM3601	3	3	0.221	0.350
	RM1	5	5	0.534	0.673
	RM8111	3	2	0.135	0.229
	RM259	3	3	0.554	0.605
	RM6840	3	2	0.153	0.282
	RM8137	4	4	0.369	0.479
2	RM3865	4	4	0.653	0.667
	RM1347	4	4	0.634	0.699
	RM6378	6	6	0.664	0.720
	RM324	5	5	0.664	0.703
	RM262	5	5	0.753	0.778
	RM3874	6	6	0.747	0.755
	RM1367	4	4	0.160	0.296
	RM240	5	5	0.619	0.655
	RM406	3	2	0.033	0.073
3	RM6959	7	6	0.810	0.821
	RM8208	5	4	0.346	0.425
	RM168	3	3	0.549	0.582
	RM8203	4	4	0.694	0.716
	RM7000	6	6	0.724	0.735
	RM7389	4	4	0.703	0.723
4	RM8213	6	6	0.778	0.759
	RM3317A	3	3	0.404	0.466
	RM5586	5	5	0.680	0.691
	RM3524	5	4	0.721	0.710
	RM3367	5	5	0.712	0.714
	RM3836	5	5	0.592	0.616
5	RM405	2	2	0.043	0.073
	RM267	3	3	0.065	0.108
	RM413	6	6	0.726	0.739
	RM1089	4	4	0.701	0.708
	RM3663	5	5	0.354	0.512
	RM3467	2	1	0.022	0.000
	RM3790	2	1	0.011	0.000
	RM6313	4	4	0.232	0.356
6	RM508	3	3	0.352	0.412
	RM510	5	5	0.750	0.740
	RM276	2	2	0.214	0.201
	RM162	4	4	0.519	0.548
7	RM1134	2	2	0.222	0.350
	RM7121	3	3	0.138	0.270
	RM11	4	4	0.495	0.491
	RM234	4	4	0.562	0.588

(Continued on next page)

Table 7. Numbers of alleles at 63 SSR markers and values of Nei's genetic diversity index in all accessions and the selected mini-core collection (Continued)

8	RM408	4	4	0.697	0.696
	RM152	5	5	0.704	0.680
	RM1235	5	5	0.743	0.782
	RM3395	6	6	0.775	0.791
	RM3153A	5	4	0.677	0.691
	RM7356	2	2	0.179	0.256
	RM284	2	2	0.115	0.201
9	RM7048	4	4	0.586	0.674
	RM3164	4	4	0.572	0.629
10	RM8201	2	2	0.054	0.107
	RM258	5	4	0.677	0.717
	RM171	2	2	0.312	0.370
	RM271	2	2	0.312	0.370
11	RM552	6	6	0.739	0.728
	RM21	5	5	0.768	0.747
12	RM247	3	3	0.455	0.572
	RM7619	3	2	0.033	0.037
	RM7376	3	3	0.551	0.600
	RM17	4	4	0.579	0.591
Mean of Nei's diversity index				0.476	0.521

Table 8. Summary of eight traits related to the panicle architecture and grain size in all accessions and the selected mini-core collection

Trait		Group	Average	Range	MD	VD	VR(%)	CR(%)
			Mean ± S.D.	Min. – Max.				
Panicle architecture	PL	All accessions	21.1 ± 2.0	16.8 – 27.3	NS	NS	87.4	92.0
		Mini-core	21.7 ± 2.1	17.7 – 27.3				
	PB	All accessions	11.5 ± 1.4	7.7 – 17.0	NS	NS	103.6	60.7
		Mini-core	11.7 ± 1.2	9.3 – 15.0				
	SB	All accessions	22.7 ± 8.0	6.7 – 45.3	NS	NS	100.1	91.4
		Mini-core	21.8 ± 7.8	6.7 – 42.0				
	GN	All accessions	137.5 ± 29.6	70.7 – 229.3	NS	NS	97.9	80.5
		Mini-core	134.6 ± 27.0	83.7 – 211.3				
Grain size	GWt	All accessions	22.1 ± 3.8	14.5 – 34.0	***	NS	100.7	89.9
		Mini-core	23.8 ± 3.6	16.5 – 34.0				
	GL	All accessions	8.12 ± 0.72	6.93 – 10.58	*	NS	113.9	83.4
		Mini-core	8.40 ± 0.72	7.36 – 10.40				
	GWd	All accessions	2.85 ± 0.26	2.30 – 3.80	NS	NS	106.0	99.3
		Mini-core	2.86 ± 0.30	2.31 – 3.80				
	LWR	All accessions	2.89 ± 0.39	2.10 – 3.83	NS	NS	110.2	99.9
		Mini-core	2.98 ± 0.41	2.10 – 3.83				
Mean							102.5	87.1

PL, panicle length; PB, number of primary branches; SB, number of secondary branches; GN, number of grains per panicle; GWt, 1,000-filled-grain weight; GL, grain length; GWd, grain width; LWR, ratio of GL to GWd; S.D., standard deviation; MD, mean difference; VD, variance difference; VR, variance difference percentage (%); CR, coincidence rate of the range of each trait.
Differences are significant at the *5%, **1%, and ***0.1% levels by Student's *t*-test; "ns," no significant difference.

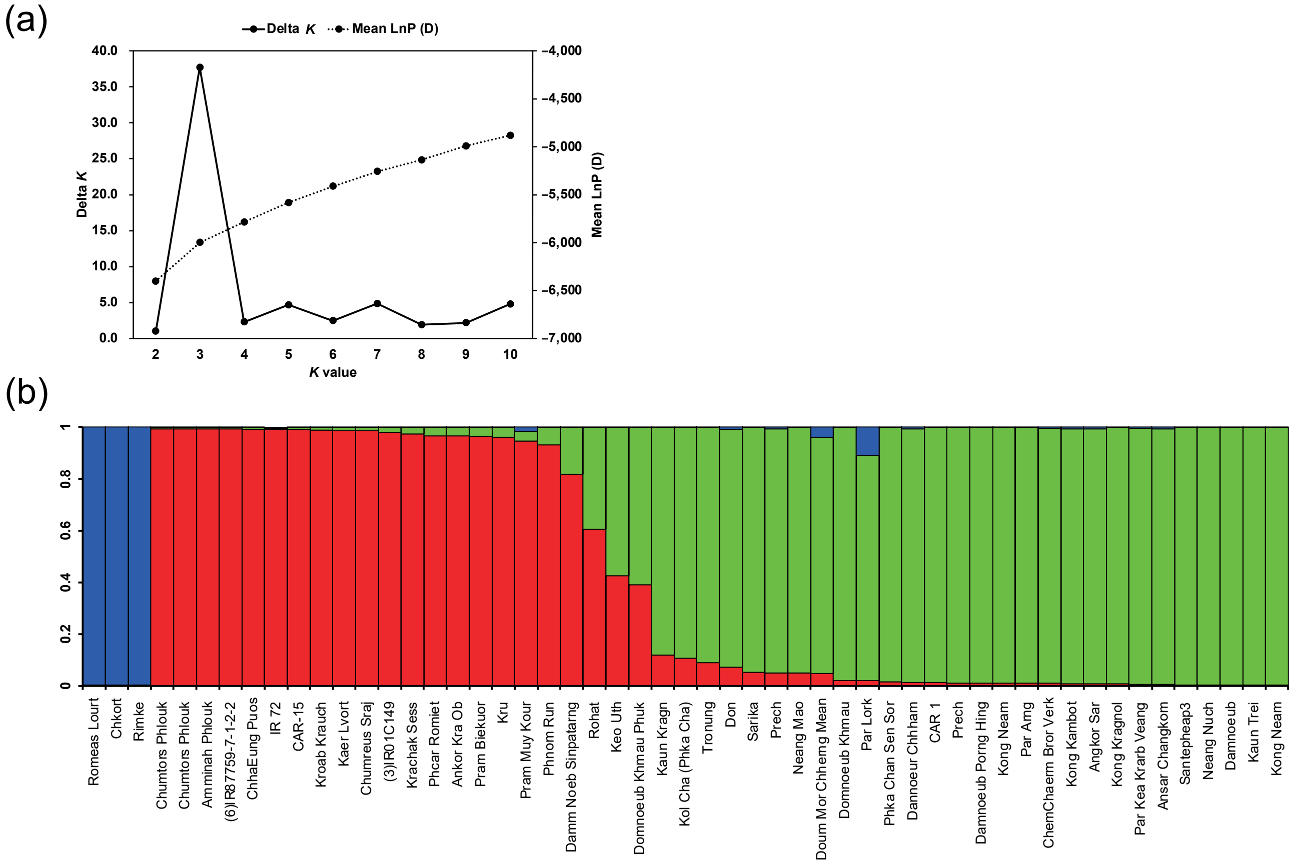

Fig. 4. Population structure of the mini-core collection based on 63 SSR markers

(a) Plot of delta *K* and ln *P*(D) against *K* = 2-10, with each *K* repeated 10 times. (b) Population structure at *K* = 3, with each accession represented by a single column partitioned into *K* colored segments, with lengths proportional to the estimated membership probability (*y*-axis)

(2000) proposed that the CR% retained by the core collection be ≥ 80%. Therefore, these results indicate that our mini-core collection maintains the diversity of all accessions.

The population structure of the mini-core collection comprised three major groups: one of UL and RL landraces derived from the tropical Japonica Group and two of cultivars derived from the Indica Group (Fig. 4; Table S1). The major factor separating the latter two groups might be heading date. It seems that the red group consists of early- to mid-heading cultivars, such as improved cultivars with photoperiod insensitivity, whereas the green group consists of mid- to late-heading cultivars grown in lowland wetlands where the water supply is reliable. Although the mini-core collection was selected on the basis of genetic diversity, it is encouraging that it also contains a wide range of variation in heading date, which we did not focus on here.

Because traits related to panicle architecture and grain size are anticipated to vary with environmental conditions, it is necessary to conduct multi-location trials to perform appropriate genetic analysis. Because handling many accessions at once requires a lot of labor and space, it is helpful to use a core or mini-core collection that preserves genetic and morphological diversities (Frankel 1984, Upadhyaya & Ortiz 2001). Thus, our mini-core collection will be helpful for genetic analysis and genetic improvement of yield.

Acknowledgement

This work was conducted as part of the research projects “Goal-oriented basic research (2016-2020)” and “Tropical crop genetic resources (2021-2025)” of the Japan International Research Center for Agricultural Sciences (JIRCAS).

References

Adriani, D. E. et al. (2016) Rice panicle plasticity in near isogenic lines carrying a QTL for larger panicle is genotype and environment dependent. *Rice* **9**, 28.

Choi, W. J. et al. (2013) How do weather extremes affect rice productivity in a changing climate? An answer to episodic lack of sunshine. *Glob. Change Biol.*, **19**, 1300-1310.

de Beukelaer, H. et al. (2018) Core Hunter 3: Flexible core subset selection. *BMC Bioinformatics*, **19**, 203.

Ebana, K. et al. (2008) Development of mini core collection of Japanese rice landrace. *Breed. Sci.*, **58**, 281-291.

Frankel, O. H. (1984). Genetic perspectives of germplasm conservation. *In* Genetic manipulation: impact on man and society. Cambridge University Press, Cambridge, pp. 161-170.

Fujino, K. et al. (2013) Roles of the *Hd5* gene controlling heading date for adaptation to the northern limits of rice cultivation. *Theor. Appl. Genet.*, **126**, 611-618.

Hamada, H. (1965) Rice in Mekong valleys. *In* Indo-China studies: Synthetic research of the culture rice-cultivating races in Southeast Asia countries (I), The Japanese Society of Ethnology, Tokyo, Japan, pp. 505-586.

Hu, J. et al. (2000) Methods of constructing core collections by stepwise clustering with three. *Theor. Appl. Genet.*, **101**, 264-268.

Iijima, K. et al. (2019) Endosperm enzyme activity is responsible for texture and eating quality of cooked rice grains in Japanese cultivars. *Biosci. Biotechnol. Biochem.*, **83**, 502-510.

IRRI–Cambodia Project (1990) Rice varietal improvement. *Annual research report.* IRRI, Los Baños, Philippines, p. 15.

Islam, M. Z. et al. (2016) Variability assessment of aromatic and fine rice germplasm in Bangladesh based on quantitative traits. *Sci. World J.*, **2016**, Article ID 2796720.

Javier, E. L. (1997) Rice ecosystems and varieties. *In* Rice Production in Cambodia, IRRI, Los Baños, Philippines, pp. 31-81.

Javier, E. L. et al. (1999) *Rice germplasm catalog of Cambodia III.* Cambodia-IRRI-Australia Project. Phnom Penh, Cambodia.

Kojima, Y. et al. (2005) Development of an RFLP-based rice diversity research set of germplasm. *Breed. Sci.*, **55**, 431-440.

Kumar, A. et al. (2020) Designing a mini-core collection effectively representing 3004 diverse rice accessions. *Plant Commun.*, **1**, 100049.

Lando, R. P. & Mak, S. (1994) Rainfed lowland rice in Cambodia: a baseline survey. *IRRI Res. Pap. Ser.*, **152**.

Li, N. et al. (2018) Control of grain size in rice. *Plant Reprod.*, **31**, 237-251.

Li, X. L. et al. (2011) Strategies on sample size determination and qualitative and quantitative traits integration to construct core collection of rice (*Oryza sativa*). *Rice Sci.*, **18**, 46-55.

Liu, W. et al. (2015) Evaluation of genetic diversity and development of a core collection of wild rice (*Oryza rufipogon* Griff.) populations in China. *PLoS ONE*, **10**, e0145990.

MAFF (Ministry of Agriculture, Forestry, and Fisheries) of Cambodia. (2019). Annual Report for Agriculture, Forestry and Fisheries 2018-2019 and Directions 2019-2020. *MAFF Annual Reports*, Phnom Penh, Cambodia.

MAFF (Ministry of Agriculture, Forestry, and Fisheries) of Cambodia. (2021). Annual Report for Agriculture, Forestry and Fisheries 2020-2021 and Directions 2021-2022. *MAFF Annual Reports*, Phnom Penh, Cambodia.

Maltsoglou, I. et al. (2010) Household level impacts of increasing food prices in Cambodia. *Environment and Natural Resources Management Working Paper*, **37**, FAO, Rome, Italy.

Ndjiondjop, M. N. et al. (2017) Genetic variation and population structure of *Oryza glaberrima* and development of a mini-core collection using DArTseq. *Front. Plant Sci.*, **8**, 1748.

Nei, M. (1972) Genetic distance between populations. *Am. Nat.*, **106**, 283-292.

Nesbitt, H. (1997) *Rice Production in Cambodia*. IRRI, Los Baños, Philippines.

Ochiai, K. et al. (2011) Suppression of a NAC-like transcription factor gene improves boron-toxicity tolerance in rice. *Plant Physiol.*, **156**, 1457-1463.

Orn, C. et al. (2015) Evaluation of genetic variation among wild rice populations in Cambodia. *Breed. Sci.*, **65**, 430-437.

Orn, C. et al. (2020) Genetic variation of rice (*Oryza sativa* L.) germplasm in Cambodia. *Breed. Sci.*, **70**, 576-585.

Ouk, M. et al. (2001) Rice production systems in Cambodia. *In* Fukai, S. and Basnayake, J., eds. Increased Rainfed Lowland Rice Production in the Mekong Region. Proceedings of an International Workshop, Vientiane, Laos. *ACIAR Proceedings*, **101**, 43-52.

Ouk, M. et al. (2009) Rice breeding methods for Cambodia. *Breed. Res.*, **11**, 65-71 [In Japanese with English title].

Ouk, M. et al. (2017) *Description of crop varieties released by Cambodian Agricultural Research and Development Institute (1990-2017).* Cambodian Agricultural Research and Development Institute. Phnom Penh, Cambodia.

Pritchard, J. K. et al. (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945-959.

Sahai, V. N. et al. (1992a) *Rice Germplasm Catalog of Cambodia I.* Cambodia-IRRI-Australia Project. Phnom Penh, Cambodia.

Sahai, V. N. et al. (1992b) *Rice Germplasm Catalog of Cambodia II.* Cambodia-IRRI-Australia Project. Phnom Penh, Cambodia.

Tanabata, T. et al. (2012) SmartGrain: High-throughput phenotyping software for measuring seed shape through image analysis. *Plant Physiol.*, **160**, 1871-1880.

Upadhyaya, D. & Ortiz, R. (2001) A mini core subset for capturing diversity and promoting utilization of chickpea genetic resources in crop improvement. *Theor. Appl. Genet.*, **102**, 1292-1298.

Yagura, K. (2021) Transformation of agriculture and its mechanisms in Pursat Province, Cambodia: Utilization of untapped resources and funding from external sources. *Jpn. J. Southeast Asia Study*, **59**, 61-100 [In Japanese with English summary].

Yan, W. et al. (2007) Development and evaluation of a core subset of the USDA rice germplasm collection. *Crop Sci.*, **47**, 869-878.

Yin, C. et al. (2021) Molecular and genetic aspects of grain number determination in rice (*Oryza sativa* L.). *Int. J. Mol. Sci.*, **22**, 1-19.

Supplemental Table S1. Cambodian rice accessions used in this study: name, panicle architecture, grain size, and categories of region, cultivar type, ecosystem type, and genetic cluster

Accession no.	Line name 1)	1000-grain weight (g)	Grain length (mm)	Grain width (mm)	Ratio of grain length to width	Number of primary branches	Number of secondary branches	Number of grains per panicle	Panicle length (cm)	Region 2)	Culltivar type 3)	Ecosystem 4)	Cluster by SSR marker genotyping 5)
1	Rohat	22.24	8.28	2.56	3.24	11.00	33.00	165.67	23.17	SW	I	IL	IIb
2	IR 66	21.43	8.60	2.41	3.58	11.00	39.00	189.00	24.33	SW	I	IL	IIb
3	Sen Pidao	22.85	8.95	2.45	3.67	10.67	29.67	157.00	25.17	SW	I	IL	IIb
4	Kru	24.77	8.86	2.38	3.73	10.33	32.33	152.33	25.33	SW	I	IL	IIb
5	IR 72	22.35	8.33	2.59	3.22	12.67	37.00	176.00	23.00	SW	I	IL	IIb
7	Chul'sa	19.10	8.28	2.33	3.57	10.67	32.67	177.33	24.50	SW	I	IL	IIb
8	(3)IR01C149	22.15	8.93	2.41	3.72	13.00	42.00	207.33	27.33	SW	I	IL	IIb
9	(6)IR87759-7-1-2-2	22.09	8.82	2.31	3.83	12.33	40.00	211.33	27.00	SW	I	IL	IIb
10	(8)IR 87808-21-2-2-3	23.95	8.85	2.32	3.83	12.67	35.67	196.00	27.17	SW	I	IL	IIb
11	Neang Nith	22.89	7.54	2.89	2.63	13.00	36.00	183.33	20.83	CT	L	RL	IIb
12	Ankor Kra Ob	23.66	7.91	2.84	2.79	10.67	20.33	127.67	21.83	CT	L	RL	IIb
13	Thmar Kreum	21.72	8.14	3.10	2.63	11.00	13.33	108.00	24.00	CT	L	RL	IIb
14	Damnoeub Krapoeu loun	19.93	8.60	2.30	3.76	10.00	15.33	101.33	21.17	SW	L	RL	IIb
15	Damnaoeub Tongsanke	27.69	8.52	3.20	2.67	11.00	23.33	132.33	20.83	NW	L	RL	IIb
16	Krachak Sess	21.70	8.17	2.58	3.18	9.67	22.67	120.67	21.33	SW	L	RL	IIb
17	IR-Kesar	22.93	8.49	2.45	3.48	11.33	18.67	125.33	26.33	SW	I	IL	IIb
18	CAR-14	22.52	8.87	2.37	3.75	11.67	21.67	141.33	25.17	SW	I	IL	IIb
19	CAR-15	25.23	9.13	2.57	3.58	13.67	28.67	173.33	25.17	SW	I	IL	IIb
20	Neang Koeuy	22.05	7.08	3.15	2.25	12.00	24.33	142.00	20.33	CT	L	RL	IIb
21	Phcar Slar	28.26	8.07	3.44	2.36	12.33	16.00	125.33	22.50	CT	L	RL	IIb
22	Phnom Run	21.83	7.66	2.86	2.69	12.67	23.67	150.33	24.00	CT	L	RL	IIb
23	Keo Uth	22.10	7.68	3.10	2.49	15.00	16.00	129.00	20.17	CT	L	RL	IIb
24	Phnom Run	20.66	7.76	2.91	2.67	12.33	21.67	133.67	21.33	CT	L	RL	IIb
25	Ba Rum	22.48	7.70	2.87	2.69	12.00	26.67	153.00	19.83	CT	L	RL	IIb
26	Ba Rum	17.88	7.22	2.51	2.89	11.00	32.33	170.67	23.33	CT	L	RL	IIb
27	Phdau Pen	21.17	8.04	2.59	3.11	12.00	17.00	114.67	19.17	CT	L	RL	IIb
28	Phdau Pen	27.64	8.60	2.97	2.90	10.67	25.00	137.00	19.50	CT	L	RL	IIb
29	Chumtors Phlouk	24.42	8.63	2.57	3.37	11.67	21.33	128.67	20.83	CT	L	RL	IIb
30	Ankor Kraob	22.81	7.85	2.75	2.86	11.67	25.00	140.67	21.00	CT	L	RL	IIb
31	Chumtors Phlouk	21.89	7.74	2.87	2.71	10.00	19.33	109.67	20.00	CT	L	RL	IIb
32	Neang Rith	22.19	7.71	2.90	2.66	8.67	20.67	110.33	19.00	CT	L	RL	IIb
33	Neang Mok	28.28	8.81	2.76	3.20	11.00	19.00	113.00	20.17	CT	L	RL	IIb
34	ChhaEung Puos	33.99	7.53	2.89	2.64	10.67	14.00	105.00	23.67	CT	L	RL	IIa
35	Phcar Mri	22.04	7.57	2.75	2.76	9.67	23.00	126.00	21.17	CT	L	RL	IIb
36	Sar Changkom	21.34	7.68	3.15	2.44	12.00	19.67	135.67	20.50	CT	L	RL	IIa
37	Phdau Pen	26.91	9.13	2.73	3.36	11.67	15.33	117.33	20.50	CT	L	RL	IIb
38	Sambok Angkrang	27.73	9.06	3.04	2.99	11.67	12.67	107.00	18.83	CT	L	RL	IIb
39	Pram Muy Kour	22.95	7.71	3.30	2.34	13.33	19.67	138.00	20.00	CT	L	RL	IIa
40	Damnoeub Khmao	23.19	7.42	3.31	2.25	8.33	10.67	78.67	18.17	CT	L	RL	IIa
41	Amminah Phlouk	27.03	9.10	2.82	3.23	13.00	18.00	129.33	19.83	CT	L	RL	IIa
42	Phcar Khghei	19.17	7.65	2.66	2.88	11.33	24.33	136.67	22.67	CT	L	RL	IIa
43	Pram Biekuor	23.25	8.12	2.80	2.91	13.33	14.33	123.00	20.17	CT	L	RL	IIa
44	Chrava Prum	17.34	7.14	2.84	2.52	11.33	22.33	133.67	19.00	CT	L	RL	IIa
45	Thnoat Pung	19.28	7.27	2.75	2.65	10.67	15.67	117.33	17.17	CT	L	RL	IIa
46	Chumreus Sraj	25.87	7.58	2.81	2.71	12.33	17.33	123.00	19.33	CT	L	RL	IIa
47	Phi Rum	25.91	7.86	2.85	2.77	11.33	15.67	116.33	20.67	CT	L	RL	IIb
48	Kong Neam	25.34	7.38	2.62	2.82	12.67	24.00	147.67	23.17	CT	L	RL	IIa
49	Sar Krom	29.24	7.75	2.69	2.89	12.00	27.00	153.33	20.67	CT	L	RL	IIa
50	Kroab Thnoat	29.37	8.09	3.15	2.57	14.67	18.33	134.33	19.50	CT	L	RL	IIb
51	Phcar Romiet	22.30	8.42	2.64	3.20	12.00	20.00	129.67	21.50	CT	L	RL	IIa
52	Krag Nagn	22.01	8.11	2.93	2.79	12.67	21.33	134.33	17.50	CT	L	RL	IIa
53	Kor Kanlas	21.94	7.89	2.88	2.74	10.67	20.67	124.67	20.33	CT	L	RL	IIa
54	Neang Ourk	20.05	7.30	3.03	2.41	11.00	15.67	116.67	22.33	SE	L	RL	IIa
55	Kong Gninh	19.15	7.53	2.75	2.74	10.00	25.33	133.67	18.33	SE	L	RL	IIa
56	Neang Nou	15.42	8.07	2.75	2.94	10.33	19.00	117.33	21.33	SE	L	RL	IIa
57	Neang Ourk	17.22	7.24	2.93	2.48	11.67	17.00	126.00	23.17	SE	L	RL	IIa
58	Neang Ourk	16.43	7.57	2.90	2.61	12.00	17.33	131.00	23.67	SE	L	RL	IIa
59	Neang Ourk	20.02	7.32	2.97	2.47	13.67	20.33	148.00	23.83	SE	L	RL	IIa
60	Damnoeub Porng Hing	30.04	8.21	3.33	2.47	12.00	20.67	121.33	20.33	SE	L	RL	IIa
61	Neang Sar Kragnol	21.14	8.10	2.65	3.07	13.33	20.00	145.00	18.83	SE	L	RL	IIa
62	Banla Phdau	18.39	7.83	2.81	2.79	10.67	14.33	110.33	22.17	SE	L	RL	IIa
63	Neang Khgneng	19.77	7.85	2.70	2.92	11.00	19.33	120.00	19.67	SE	L	RL	IIa
64	Kaun Srau	27.56	7.39	3.19	2.32	10.00	15.67	102.67	16.83	SE	L	RL	IIa
65	Popey Angkor Kraham	19.17	7.97	2.60	3.07	10.67	24.00	143.00	20.67	SE	L	RL	IIa
66	Angkor Sar	23.99	7.98	3.80	2.10	12.00	7.00	83.67	21.33	SE	L	RL	IIa

(Continued on next page)

Supplemental Table S1. Cambodian rice accessions used in this study: name, panicle architecture, grain size, and categories of region, cultivar type, ecosystem type, and genetic cluster (Continued 1)

67	Neang Ourk	16.02	6.93	2.92	2.38	11.00	15.67	116.33	20.17	SE	L	RL	IIa
68	Angmeas	15.04	7.68	3.00	2.57	10.00	24.33	138.33	19.33	SE	L	RL	IIa
69	Damnoeub Kraham	18.03	7.38	2.86	2.58	13.00	20.33	135.67	21.33	SE	L	RL	IIa
70	Tamten	16.08	7.16	2.90	2.48	11.67	17.33	123.67	18.67	SE	L	RL	IIa
71	Neang Sar	20.90	7.86	2.85	2.78	11.33	28.00	143.33	19.00	NW	L	RL	IIa
72	Phcar Samlei	18.27	7.73	2.68	2.90	10.67	20.67	125.00	20.33	NW	L	RL	IIa
73	Neang Sar	16.25	7.54	2.85	2.65	12.33	22.00	145.00	21.17	NW	L	RL	IIa
74	Neang Meas	17.19	8.37	2.64	3.18	10.67	27.67	143.33	20.00	SE	L	RL	IIa
75	Damnoeub	20.74	7.40	2.78	2.68	11.00	20.00	122.00	17.67	SE	L	RL	IIa
76	Damm Noeb Phka Roluosh	27.86	8.31	3.28	2.54	10.00	11.67	91.67	22.50	SE	L	RL	IIa
77	Pram Beikour	21.33	7.68	2.85	2.70	10.67	19.67	110.67	20.83	SW	L	RL	IIa
78	Pram Beikour	19.11	7.76	2.76	2.82	9.67	13.33	100.00	20.67	SW	L	RL	IIa
79	Sambokankrang	19.60	7.81	2.53	3.10	11.67	20.00	122.00	20.33	SW	L	RL	IIa
80	Sambok Angkrong	17.78	7.78	2.48	3.15	10.67	22.33	125.33	21.33	SW	L	RL	IIa
81	Dannoeur Chhham	22.03	9.22	2.76	3.35	10.33	24.00	137.00	24.50	SW	L	RL	IIb
82	Neang Sar	20.71	8.34	2.42	3.46	11.67	27.67	146.33	22.50	SW	L	RL	IIa
83	Neang Nuch	21.08	8.08	2.84	2.87	10.00	13.33	97.67	20.50	SW	L	RL	IIa
84	Phcar Tien	19.63	8.04	2.53	3.18	10.33	14.67	104.33	21.00	SW	L	RL	IIa
85	Tronung	16.45	7.85	2.56	3.09	11.00	18.33	118.33	21.00	SW	L	RL	IIa
86	Damnoeub Krapeu	28.43	10.58	2.92	3.63	7.67	9.33	70.67	21.67	SW	L	RL	IIa
87	Neang Eng	21.05	7.46	2.90	2.59	9.33	20.00	113.00	20.50	NE	L	RL	IIa
88	Damm Noeb Sinpatarng	30.69	9.19	3.43	2.69	12.33	13.67	120.00	19.50	NE	L	RL	IIa
89	Srau Kreabb Karn Tuot	24.22	8.27	3.05	2.72	10.00	19.00	114.67	20.83	NE	L	RL	IIa
90	Chiich Karng	23.88	7.58	2.93	2.60	10.67	15.67	117.33	21.17	NW	L	RL	IIa
91	Domnoeub Khmau Phuk	24.84	9.23	2.87	3.24	12.00	12.00	110.00	20.83	NW	L	RL	IIa
92	Kaer Lvort	25.28	8.23	3.10	2.67	10.00	13.33	104.33	23.17	NW	L	RL	IIa
93	Domnoeub Khmau	16.91	7.73	2.63	2.94	14.00	29.67	177.67	20.17	SW	L	RL	IIb
94	Domnoeub Kreabb Sar	31.99	9.67	3.42	2.83	10.33	13.00	99.33	18.17	NE	L	UL	IIa
95	Kol Cha (Phka Cha)	30.30	10.40	3.13	3.35	12.00	11.33	102.33	18.50	NE	L	UL	IIa
96	Neang Kang	19.41	7.60	3.04	2.51	11.00	19.00	110.33	19.83	SW	L	RL	IIa
97	CAR 1	18.99	7.76	2.46	3.17	13.00	35.33	164.33	18.67	SW	L	RL	IIa
98	CAR 11	26.04	9.91	2.73	3.66	9.33	12.33	96.00	22.17	SW	L	RL	IIb
99	CAR 2	20.24	7.22	3.07	2.36	13.00	21.33	149.00	20.83	SW	L	RL	IIb
100	CAR 3	21.93	8.02	3.18	2.53	13.67	14.00	116.67	19.67	SW	L	RL	IIb
101	Phka Chan Sen Sor	21.80	9.13	2.55	3.60	11.67	18.00	125.00	21.67	SW	L	RL	IIb
102	Phka Rumchang	22.19	9.31	2.60	3.59	10.67	14.67	110.33	21.17	SW	L	RL	IIb
103	Phka Rumchek	21.40	9.20	2.61	3.54	11.33	16.33	123.33	22.33	SW	L	RL	IIb
104	Phka Rumduol	24.04	9.57	2.62	3.66	11.33	16.00	118.67	21.50	SW	L	RL	IIb
105	Phka Romeat	22.40	9.32	2.61	3.58	12.67	23.00	145.33	23.33	SW	L	RL	IIb
108	Santepheap3	24.13	7.98	3.12	2.56	11.33	24.67	141.67	22.67	SW	L	RL	IIb
109	Sarika	29.86	8.67	2.90	3.01	10.67	13.33	100.67	22.17	SW	L	RL	IIb
110	Phnom Run	15.42	6.95	3.02	2.30	15.33	40.00	202.67	18.50	CT	L	RL	IIb
111	Kong Neam	25.43	8.86	3.14	2.83	12.00	15.67	120.33	20.83	CT	L	RL	IIb
112	Chem Chaerm Bror Verk	23.51	7.81	3.17	2.47	10.00	17.00	107.33	20.83	CT	L	RL	IIb
113	Kong Neam	22.71	7.98	2.71	2.95	13.33	32.33	166.67	20.17	NW	L	RL	IIb
114	Kong Neam	19.97	8.08	2.82	2.88	13.33	24.33	148.67	20.67	NW	L	RL	IIa
115	Kaun Kragn	23.71	8.95	2.77	3.24	10.67	19.33	124.00	22.50	CT	L	RL	IIb
116	Neang Ming	23.06	7.67	2.83	2.71	10.33	31.33	164.33	23.67	CT	L	RL	IIb
117	Kong Kambot	24.54	9.09	2.68	3.40	11.33	18.33	120.00	23.33	NW	L	RL	IIb
118	Moha Phal	21.81	7.66	2.86	2.68	12.00	24.67	152.67	20.67	CT	L	RL	IIb
119	Phcar Khgnei	22.47	7.48	3.14	2.39	10.00	23.33	125.00	20.67	CT	L	RL	IIb
120	Ansar Changkom	21.67	7.54	2.86	2.64	12.00	25.33	140.33	18.67	NW	L	RL	IIb
121	Neang Sar Kakdek	25.23	9.25	2.74	3.38	12.33	17.00	123.00	20.83	NW	L	RL	IIb
122	Phcar Daung	19.99	7.33	2.81	2.61	11.67	33.00	176.67	18.33	CT	L	RL	IIb
123	Moha Phal	23.67	8.69	2.66	3.28	12.67	22.33	137.67	22.83	CT	L	RL	IIa
124	Neang Minh	25.77	8.84	3.18	2.79	10.33	12.00	95.33	19.67	NW	L	RL	IIb
125	Neang Yuorn	27.64	8.99	2.61	3.46	10.00	18.33	115.00	20.00	CT	L	RL	IIb
126	Neang Yuorn	22.94	7.60	2.98	2.56	11.33	25.67	137.67	23.00	CT	L	RL	IIb
127	Phcar Khnei	24.45	8.72	3.10	2.83	10.67	12.00	100.33	21.00	NW	L	RL	IIb
128	Kaun Trei	23.21	7.87	2.75	2.87	12.67	21.00	136.67	22.00	NW	L	RL	IIa
129	Prech	25.73	8.45	3.20	2.66	11.33	6.67	86.67	20.50	NW	L	RL	IIb
130	Prech	22.90	8.18	2.40	3.41	11.00	18.67	121.00	21.67	CT	L	RL	IIb
131	Thnoat	23.23	7.38	3.25	2.28	12.00	17.67	125.00	20.00	CT	L	RL	IIb
132	Khpor Daung	20.31	7.70	2.71	2.85	15.00	29.33	184.00	21.50	NW	L	RL	IIb
133	Changkung Kraing	18.26	7.50	3.11	2.42	10.67	26.67	137.00	21.33	NW	L	RL	IIb
134	Khpordaung	20.29	7.63	2.80	2.74	13.33	25.33	168.67	19.33	NW	L	RL	IIb
135	Kong Kragnol	17.53	7.36	3.13	2.35	10.33	25.00	143.67	19.33	NW	L	RL	IIb

(Continued on next page)

Supplemental Table S1. Cambodian rice accessions used in this study: name, panicle architecture, grain size, and categories of region, cultivar type, ecosystem type, and genetic cluster (Continued 2)

136	Sarleak Sanleuk	18.28	7.43	2.83	2.63	12.67	22.00	142.00	18.83	NW	L	RL	IIb
137	Neang Minh	19.45	7.84	3.03	2.59	9.67	24.67	131.00	21.00	CT	L	RL	IIa
138	Kroab Krauch	17.84	7.38	3.03	2.43	12.00	20.67	137.00	23.00	NW	L	RL	IIa
139	Phcar Sla	18.84	7.16	2.56	2.81	13.33	44.00	218.67	19.50	NW	L	RL	IIa
140	Phcar Sla	20.06	7.79	3.12	2.50	12.33	24.67	140.67	21.83	NW	L	RL	IIa
141	Neang Mao	20.02	8.05	2.85	2.83	11.00	30.67	150.33	20.33	NW	L	RL	Ib
142	Chkort	30.82	9.70	3.24	3.00	10.67	17.67	112.33	18.67	NW	L	UL	IIb
143	Neang Minh	17.32	7.88	2.73	2.89	11.67	36.00	180.67	18.83	NW	L	RL	IIa
145	Chhmar Prum	16.94	7.60	2.64	2.89	12.33	36.00	186.00	20.17	SW	L	RL	IIa
146	Khnonk Srokkhlay	18.37	7.47	2.91	2.58	10.67	34.00	160.00	19.00	SW	L	RL	IIa
147	Chhmar Chang Kom	19.63	7.48	2.84	2.64	11.67	25.00	158.67	20.33	SW	L	RL	IIa
148	Phcarteen	14.48	7.92	2.77	2.87	11.00	35.00	183.33	19.83	SW	L	RL	IIa
150	Chhmar Laeth	17.76	7.50	2.78	2.70	9.67	24.33	130.00	18.50	SW	L	RL	IIa
151	Chhmar Chang Kom	18.93	7.60	2.68	2.84	9.00	14.00	96.67	18.83	SW	L	RL	IIa
152	Chumpu Pean	16.38	7.17	2.96	2.43	10.67	34.67	164.33	21.67	SW	L	RL	IIa
153	Chhmar Prum	15.79	7.90	2.75	2.87	10.67	32.67	161.67	18.67	SW	L	RL	IIa
154	Neangek	20.72	7.69	2.73	2.82	12.67	16.33	125.33	23.33	SE	L	RL	IIa
155	Neang Sar	17.54	8.14	2.55	3.20	9.00	19.33	114.33	19.83	SE	L	RL	IIa
156	Par Lork	29.21	9.59	3.09	3.12	11.00	26.00	142.33	25.17	SE	L	RL	IIa
157	Sok Soy	18.16	7.48	3.28	2.29	12.67	26.33	161.67	20.33	SE	L	RL	IIa
158	Kantuy Chachah	21.77	7.69	2.97	2.59	12.67	30.00	163.00	20.17	SE	L	RL	IIa
159	Bei Kuo	17.74	7.79	2.78	2.82	10.33	27.00	142.00	20.00	SE	L	RL	IIa
160	Buon Khaer	19.17	7.37	2.92	2.53	10.67	17.67	133.00	20.50	SW	L	RL	IIa
161	CAR-04	20.91	8.06	2.84	2.85	15.00	32.33	177.00	18.83	SW	L	RL	IIa
162	CAR-05	19.99	7.95	2.64	3.02	15.33	43.00	217.33	21.33	SW	L	RL	IIa
163	Kiev Mie	19.28	7.52	3.09	2.43	11.67	24.00	144.67	19.83	SW	L	RL	IIa
164	CAR-07	25.46	8.57	2.86	3.01	11.33	15.33	104.67	19.17	SW	L	RL	IIa
165	CAR-08	17.62	6.93	2.84	2.45	12.00	38.33	193.33	18.33	SW	L	RL	IIa
166	CAR-09	20.10	8.02	2.92	2.75	13.67	45.33	213.00	22.17	SW	L	RL	IIa
167	CAR-13	16.14	7.05	2.64	2.67	10.67	34.00	186.33	20.50	SW	L	RL	IIa
168	CAR-12	18.39	7.79	2.80	2.79	13.67	40.33	229.33	21.33	SW	L	RL	Ib
169	Romeas Lourt	24.64	8.76	3.04	2.90	10.00	26.33	138.00	21.00	SW	L	RL	IIa
170	Raing Chey	21.25	8.22	2.69	3.07	17.00	32.67	200.00	21.17	SW	L	RL	Ib
171	Par Chhung	30.83	9.08	3.54	2.58	10.33	8.67	105.00	21.67	NE	L	UL	IIa
172	Par Kea Krarb Veang	24.00	8.95	2.90	3.11	12.67	20.67	156.67	22.50	NE	L	UL	IIa
173	Doum Mor Chherng Mean	19.91	9.46	2.62	3.62	9.33	35.00	155.33	22.33	NE	L	UL	IIa
174	Doum Morb Kmou	27.89	9.38	3.47	2.71	10.33	36.67	178.00	25.50	NE	L	UL	IIa
175	Par Arng	24.99	8.38	3.33	2.53	13.67	26.33	167.67	22.50	NE	L	UL	Ib
176	Rimke	NA[6]	NA[6]	NA[6]	NA[6]	NA[6]	NA[6]	NA[6]	NA[6]	SW	I	UL	Ib
177	Khao Tah Petch	21.97	9.17	2.82	3.26	11.33	23.67	143.33	27.00	SW	I	RL	IIa
178	Don	25.58	9.42	2.80	3.38	11.33	24.67	164.67	21.00	SW	I	RL	IIa
179	Tewada	23.98	9.40	2.77	3.41	12.33	13.33	126.67	22.67	SW	I	RL	IIa

---

[1] Accessions in red were selected for the mini-core collection.

[2] NE, northeast; CT, central; SE, southeast; SW, southwest; NW, northwest

[3] I, improved cultivar; L, landrace

[4] IL, irrigated lowland; RL, rainfed lowland; UL, upland

[5] Categorized in the study by Orn et al. (2020)

[6] NA indicates “data not available.”