Australian Journal of Crop Science AJCS

AJCS 14(04):597-604 (2020) ISSN:1835-2707
doi: 10.21475/ajcs.20.14.04.p2092

# Assessment of genetic diversity in Thai upland rice varieties using SSR markers

**Somrudee Nilthong*, Ekachai Chukeatirote, Rungrote Nilthong**

**School of Science, Mae Fah Luang University, Chiang Rai 57100, Thailand**

***Corresponding author: somrudee@mfu.ac.th**

## Abstract

Upland rice (*Oryza sativa* L.) is precious genetic resource containing some valuable alleles not common in modern germplasm. In this study, genetic diversity and population structure of 98 upland rice varieties from northern part of Thailand were examined using nine simple sequence repeat markers. Number of alleles detected by the above primers was 50 with a minimum and maximum frequency of 2 to 10 alleles per locus, respectively. The polymorphic information content (PIC) values ranged from 0.375 to 0.714 with an average of 0.605 for the primers RM164 and RM1, respectively. Dendrogram cluster analysis of the SSR data distinctly classified all genotypes into three major groups (I, II and III), which corresponded to their places of collection. Population structure divided these genotypes into two distinct subpopulations. Subpopulation 1 consisted of upland rice varieties that collected from Chiang Rai province while the majority of subpopulation 2 were collected from Phayao and Phitsanulok provinces. Analysis of molecular variance revealed 68% variance among two subpopulations and 32% variance within subpopulations, suggesting a high genetic differentiation between the two subpopulations. The huge genetic variability of upland rice in northern part of Thailand can be used to complement the gene pool of modern genotypes in rice breeding program.

**Keywords:** *Oryza sativa* L. SSR markers upland rice genetic diversity.

## Introduction

Rice (*Oryza sativa* L.), one of the most important crops, is globally cultivated and feeds over all of the population in the world (Mohanty 2013). Especially in Asia, rice is a staple crop and often considered as a cash crop because of its potential for export. Thailand in particular is well known as the world's largest rice exporter and is also among the world's largest rice producers (Milovanovic and Smutka 2017). Some critical issues occur however, regarding the rice production. For example, unforeseen climatic changes (i.e., drought and flooding) affect directly the rice yield. On the other hand, the trend of the consumers has changed, especially for those with high incomes which focus on a premium quality rice. As a result, several breeding programs have been undertaken in order to improve rice with desired characteristics, corresponding to this tentative situation.

It has been suggested that rice was domesticated between 8,000 – 10,000 years ago from its wild ancestor, *Oryza rufipogon* (Oka 1988). Since then, the domestication process involving strong screening for desirable traits, causes in precipitous loss of the genetic diversity (Londo et al. 2006). Rice in particular is a good example for this situation. Modern rice varieties have been bred for high yield as well as high quality. These rice varieties being bred for such purposes tend to loss genetic diversity and thus this may have a series of effects from susceptibility to epidemic diseases or even cause a serious threat (i.e. rice extinction). Therefore, knowledge of genetic diversity in the genepool of rice is crucial considering that such information can be used efficiently in the rice breeding program generating new varieties suitable to changing cultivated conditions.

Upland rice cultivars have been traditionally cultivated by minority people mainly in mountain areas of Southeast Asia (Oka 1988; Sato 1987, 1991). Upland rice is grown during rainy season without irrigation, depending only on rain. Most grains are consumed by the farmer family and the rest is sown in the next year without severe selection. Upland rice is considered as an important gene source for the resistance of insects, pathogens and abiotic stress (Ishikawa et al. 2006).

Recently, DNA technology has been successfully applied in the plant breeding program. One of the major applications is to introduce the DNA markers specific for the desirable traits of the plant cultivars allowing direct detection of these 'desired' plants in the breeding program. In rice, there are a large number of microsatellite markers with different simple sequence repeat (SSR) motifs available on databases (Akagi et al. 1996; Chen et al. 1997; Panaud et al. 1996; Temnykh et al. 1999; Wu and Tanksley 1993). The microsatellite markers are distributed uniformly throughout the genome and can detect a high level of allelic diversity in cultivated varieties and distantly related species that made it possible to investigate the incidence and variability of simple sequence repeats at the whole-genome level (Cho et al. 2000; McCouch et al. 1997). Many studies have used SSR markers to investigate the genetic diversity and population structure within rice (Pusadee et al. 2009; Salgotra et al. 2015;

Vilayheuang et al. 2016; Wunna et al. 2016). For example, Wunna et al. (2016) examined genetic variation of rice (*Oryza sativa* L.) germplasm in Myanmar using SSR markers and found that rice germplasm in Myanmar has high genetic diversity among ecosystems and areas. Furthermore, microsatellite DNA markers were used to study genetic diversity and population structure of 'Khao Kai Noi', a landrace rice, in Laos. The result showed that genetic variation was largest among accessions and smallest within accessions. 'Khao Kai Noi' accessions were clustered into three different genetic backgrounds (Vilayheuang et al. 2016). However, there was no report of genetic diversity of Thai upland rice. Therefore the genetic diversity of 98 upland rice varieties collected from northern of Thailand were analyzed using nine SSR markers to understand the present genetic diversity in Thailand and to be utilized in rice breeding program in the future.

## Results and Discussion

### *Genetic diversity values among 98 upland rice varieties*

A total of 50 alleles from 9 SSR primer pairs were detected across all 98 upland rice varieties in northern part of Thailand. The number of alleles per primer pair (locus) detected by microsatellite primers varied from 2 to 10 with an average of 5.556 alleles per locus with 33.33% and 22.22% of the loci having five and four alleles, respectively (Table 1). The average numbers of alleles per locus observed in this study correspond well to Cho et al. (2000) who reported that the average alleles per locus for various classes of microsatellites in rice germplasm were 2.0 - 5.5 alleles per locus. However, the mean of alleles per locus is in agreement with Brondani et al. (2006) who detected an average of 5.4 alleles per locus when 25 SSR markers were used to distinguish 20 and 10 cultivars of upland rice and commercial rice from Brasil, respectively. In addition, these results were similar to the previous report of Vilayheuang et al. (2016) who calculated an average of 5.7 alleles per locus among 70 accessions of Khao Kai Noi (Lao rice) from Laos.

Nine SSR primer pairs used in this study could generate polymorphic bands and the polymorphic information content (PIC) values that reflected allele diversity and frequency among the upland rice varieties. The PIC values are a good indication of the usefulness of markers for linkage analysis when defining the inheritance between offspring and parental genotypes (Shete et al. 2000). In this study, the PIC values ranged from 0.375 in RM164 to 0.714 in RM1 with an average of 0.605 (Table 1). Botstein et al. (1980) reported that the PIC value > 0.5 meaning the locus was high diversity. If the PIC value was between 0.25 and 0.50 meaning, the locus was intermediate diversity when PIC value < 0.25, the locus was low diversity. Our study showed that the PIC values for almost all the SSR markers (excepting RM164) were higher than 0.5 indicating that all the SSR markers were considered high informative markers. Similar results were also found in 175 accessions of upland and lowland rice in Myanmar, Thailand and Yunan in China, which had 0.75 PIC values (Wunna et al. 2016). Shannon's Information index (*I*) averaged 1.266 and ranged from 0.693 to 1.723. Expected heterozygosity (*He*) in the population varied from 0.5 (RM164) to 0.74 (RM1) with an average of 0.66, while observed heterozygosity (*Ho*) ranged from 0.000 (RM253) to 1.000 (RM22) with a mean of 0.472.

### *Genetic relationship among upland rice varieties*

All 50 SSR alleles scored were used to calculate the genetic similarity which used to determine the level of relatedness among the upland rice varieties. The Dice's similarity coefficients among upland rice varieties ranged from 0.1554 to 0.8000, indicating a high genetic diversity among the 98 upland rice varieties. This is in agreement with Wunna et al. (2016) who studied the genetic variation of rice germplasm in Myanmar, including landraces and improved types from upland and lowland rice using SSR markers and the results show that rice germplasm in Myanmar has high genetic diversity. In addition, Thailand, Myanmar, Laos and other countries in Southeast Asia are located at the center of diversity for rice (Nakagahra and Hayashi 1977).

Genetic similarity values among the upland rice varieties were then used to group the varieties and to construct a dendrogram based on the UPGMA cluster analysis using the R program. In the dendrogram (Fig. 1), all genotypes of upland rice varieties were distinctly separated into three major groups, designated as I, II and III. Group I contained 34 varieties (34.69% in total) at a similarity coefficient of 20%. This group could be further sub-clustered into 4 subgroups with varying levels of similarity coefficients. The major subgroup A at a similarity coefficient of about 24% comprised of 15 varieties that received from Phayao province. The remaining varieties from Phayao province were clustered in the subgroup D at a similarity coefficient of about 34%. Similarly, collections from Chiang Rai province were grouped into the subgroup B and C at similarity coefficients of about 28% and 31%, respectively (Fig. 1). Group II, the smallest group, comprised of eight varieties mainly received from Phitsanulok province which clustered at similarity coefficient of 30%. Group III contained 56 varieties (57.14% in total), most of them collected from Chiang Rai province. The similarity coefficients of this group ranged from 44% to 100%. This cluster could be divided into two groups. One major group with similarity coefficient of about 38%, comprising of 10 varieties, was mostly collected from the same village. The second group at the similarity coefficient of about 39% contained 46 varieties collected from many villages in Chiang Rai province. Based on the dendrogram, our data showed that upland rice varieties were well clustered with respect to their places/ geographic area of collection and the genetic diversity among upland rice varieties from the three areas in Thailand has high genetic diversity.

### *Population structure analysis*

The Bayesian model-based structure analysis was carried out by K values from 1 to 10 with 10 iterations using all 98 genotypes. In order to find the optimal K-value, the possible cluster numbers (K-value) were plotted against $\Delta K$ which showed a clear maximum peak at K = 2 (Fig. 2A). A continuous gradual increase was observed in the log likelihood with the increased of K (Fig. 2B). The optimal K-value stratified that two subpopulations assigned to the subpopulation 1 and 2 showed the highest probability for population clustering. The subpopulation 1 (orange color,

**Table 1.** Nine SSR primer pairs information and the information of polymorphism obtained from 98 upland rice varieties. information content.

| SSR primers | Chr. | SSR motif | Primers sequences (5'→3') | Number of alleles | *He* | *Ho* | *I* | PIC |
|---|---|---|---|---|---|---|---|---|
| RM1 | 1 | $(GA)_{26}$ | F: GCG AAA ACA CAA TGC AAA AA<br>R: GCG TTG GTT GGA CCT GAC | 10 | 0.740 | 0.500 | 1.723 | 0.714 |
| RM10 | 2 | $(GA)_{15}$ | F:TTG TCA AGA GGA GGC ATC G<br>R: CAG AAT GGG AAA TGG GTC C | 5 | 0.717 | 0.133 | 1.342 | 0.658 |
| RM19 | 4 | $(ATC)_{10}$ | F: CAA AAA CAG AGC AGA TGA C<br>R: CTC AAG ATG GAC GCC AAG A | 4 | 0.611 | 0.480 | 1.125 | 0.556 |
| RM22 | 3 | $(GA)_{22}$ | F: GGT TTG GGA GCC CAT AAT CT<br>R: CTG GGC TTC TTT CAC TCG TC | 4 | 0.688 | 1.000 | 1.125 | 0.627 |
| RM164 | 5 | $(GT)_{16}TT(GT)_4$ | F: TCT TGC CCG TCA CTG CAG ATA TCC<br>R: GCA GCC CTA ATG CTA CAA TTC TTC | 2 | 0.502 | 0.122 | 0.693 | 0.375 |
| RM241 | 4 | $(CT)_{31}$ | F: GAG CCA AAT AAG ATC GCT GA<br>R: TGC AAG CAG CAG ATT TAG TG | 9 | 0.742 | 0.980 | 1.552 | 0.698 |
| RM252 | 4 | $(GA)_{19}$ | F: TTC GCT GAC GTG ATA GGT TG<br>R: ATG ACT TGA TCC CGA GAA CG | 5 | 0.662 | 0.929 | 1.252 | 0.597 |
| RM253 | 6 | $(GA)_{25}$ | F: TCC TTC AAG AGT GCA AAA CC<br>R: GCA TTG TCA TGT CGA AGC C | 6 | 0.672 | 0.000 | 1.295 | 0.621 |
| OSR28 | 9 | $(AGA)_n$ | F: AGC AGC TAT AGC TTA GCT GG<br>R: ACT GCA CAT GAG CAG AGA CA | 5 | 0.641 | 0.102 | 1.287 | 0.599 |
| Total | | | | 50 | | | | |
| Average | | | | 5.556 | 0.664 | 0.472 | 1.266 | 0.605 |

Chr.= chromosomes; *He*= expected heterozygosity; *Ho*= observed heterozygosity; *I*= Shanon information index; PIC= polymorphism.

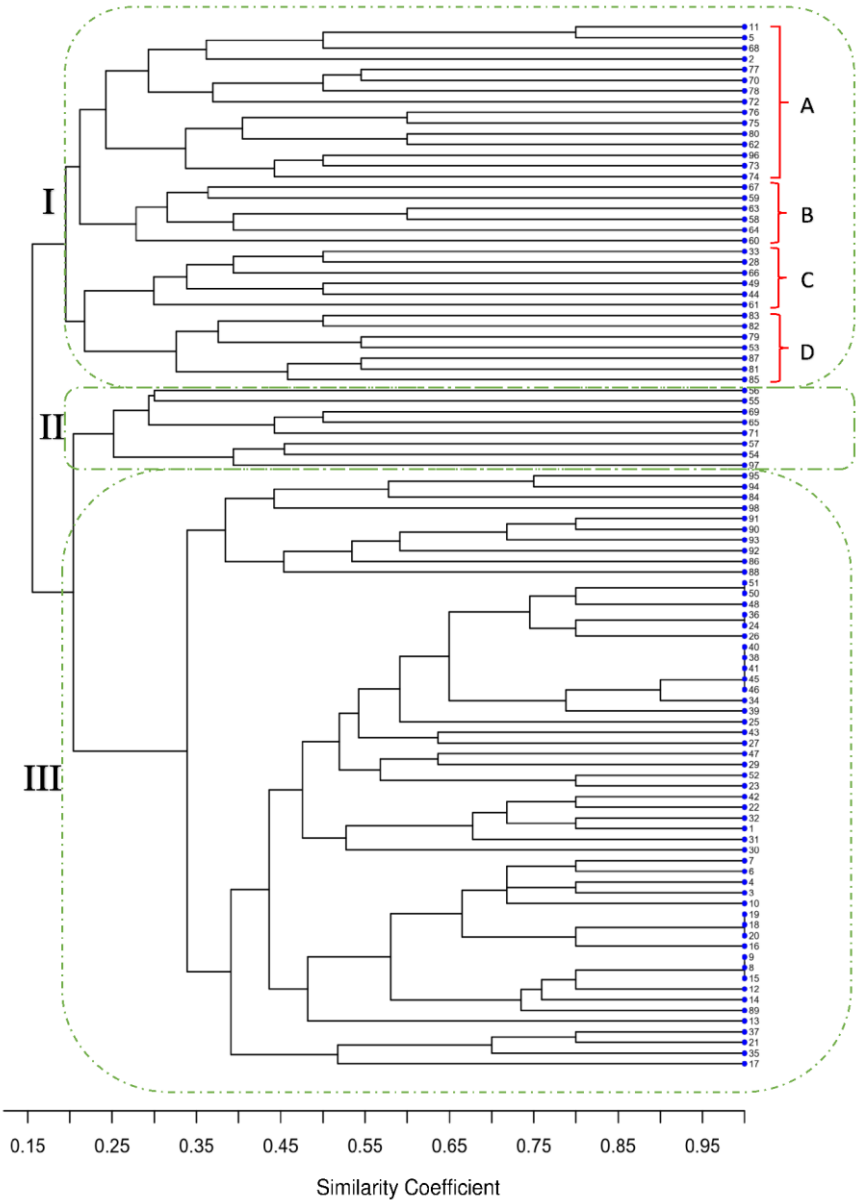


**Fig 1.** UPGMA dendrogram showing three clusters (I, II and III) of all 98 upland rice varieties.

**Table 2.** Population structure results of 98 upland rice varieties for the fixation index (Fst), expected heterozygosity (*He*), number of genotypes in each subpopulation and inferred subpopulation.

| Subpopulation | Fst | *He* | No. of genotypes | Inferred subpopulation |
|---|---|---|---|---|
| Subpopulation 1 | 0.4456 | 0.4222 | 45 | 0.481 |
| Subpopulation 2 | 0.0280 | 0.6672 | 53 | 0.518 |

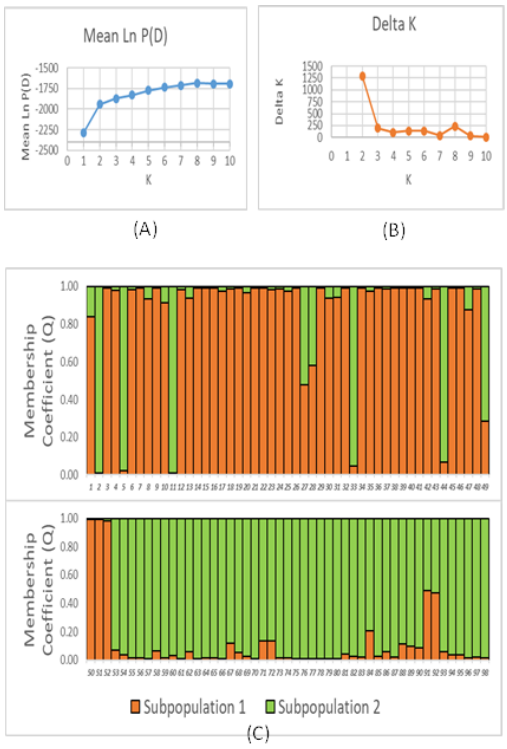


**Fig 2.** Population structure of 98 varieties of upland rice. (A) The relationship between ΔK and K showing the maximum peak at K = 2 (B) The average log-likelihood of K-value against the number of K (C) The population structure of 98 upland rice varieties on K = 2. Varieties in orange color clustered into subpopulation 1 and the ones in green grouped into subpopulation 2.

**Table 3.** Analysis of molecular variance (AMOVA) among and within two subpopulations of 98 upland rice varieties.

| Source of variance | *df* | SS | MS | Variance components | Variation (%) |
|---|---|---|---|---|---|
| Among subpopulation | 1 | 215864.919 | 215864.919 | 4392.818 | 68 |
| Within subpopulation | 96 | 196919.816 | 2051.248 | 2051.248 | 32 |
| Total | 97 | 412784.735 | | 6444.066 | 100 |

*df*: degrees of freedom; SS: Sums of squares, MS: Means squares.

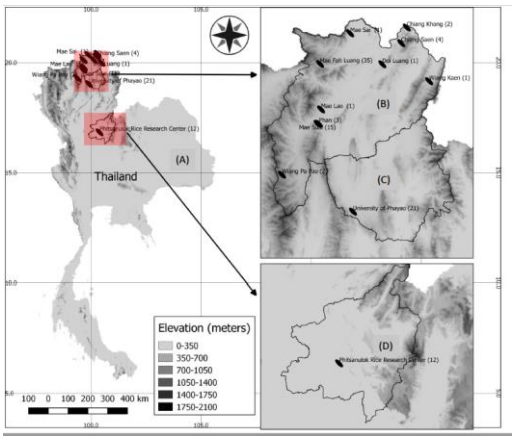


**Fig 3.** Sampling locations of upland rice varieties from 3 provinces in Thailand. (A) Map of Thailand (B) Enlarged view of upland rice growing districts in Chiang Rai province (C) and (D) Maps of University of Phayao and Phitsanulok Rice Research Center, respectively. The numbers in parentheses stand for the total numbers of upland rice varieties collected from that place.

**Table 4.** Name, code, origin, latitude and longitude of provinces in Thailand and cluster based on the UPGMA clustering of the 98 upland rice varieties.

| Code | Name | Place of collection | Latitude and longitude of provinces in Thailand | Height (m) | Cluster |
|---|---|---|---|---|---|
| 1 | Lap Chang | Mae Sai district, Chiang Rai | N 19 40 50.5 E 99 37 59.6 | 467 | III |
| 2 | Khao Sim Khao | Phan district district, Chiang Rai | N 19 40 54.8 E 99 38 37.1 | 467 | I (A) |
| 3 | Beu Mheu | Wiang Par Pao district, Chiang Rai | N 19 18 10.62 E 99 22 13.91 | 1097.83 | III |
| 4 | Pi Ai Zoo | Wiang Par Pao district, Chiang Rai | N 19 18 10.62 E 99 22 13.91 | 1097.83 | III |
| 5 | Hom Doi | Chiang Khong district, Chiang Rai | N 20 23 44.1 E 100 17 43.9 | 382 | I (A) |
| 6 | Khao Kam09 | Chiang Khong district, Chiang Rai | N 20 23 44.1 E 100 17 43.9 | 382 | III |
| 7 | Khao Daeng | Chiang Saen district, Chiang Rai | N 20 16 48.7 E 100 15 15.3 | 440 | III |
| 8 | Khao Kam012 | Chiang Saen district, Chiang Rai | N 20 16 48.7 E 100 15 15.3 | 440 | III |
| 9 | Khao Kam013 | Chiang Saen district, Chiang Rai | N 20 16 48.7 E 100 15 15.3 | 440 | III |
| 10 | Khao' Pleuak Kheaw | Chiang Saen district, Chiang Rai | N 20 16 48.7 E 100 15 15.3 | 440 | III |
| 11 | Unknown015 | Mae Fah Luang district, Chiang Rai | N 20 17 35.1 E 99 48 55.8 | 940 | I (A) |
| 12 | Khao Jao Doi016 | Mae Suai district, Chiang Rai | N 19 40 48.6 E 99 39 14.1 | 453 | III |
| 13 | Khao' Jao Doi017 | Mae Suai district, Chiang Rai | N 19 68 018 E 99 65 393 | 453 | III |
| 14 | Unknown 018 | Doi Luang district, Chiang Rai | N 20 7 22.0 E 100 6 51.0 | 379 | III |
| 15 | A-Kha Ja Bue | Mae Fah Luang district, Chiang Rai | N 20 10 41.5 E 99 42 18.5 | 737 | III |
| 16 | La Hae020 | Mae Fah Luang district, Chiang Rai | N 20 10 41.5 E 99 42 18.5 | 737 | III |
| 17 | Unknown021 | Mae Fah Luang district, Chiang Rai | N 20 10 41.5 E 99 42 18.5 | 737 | III |
| 18 | Chaw Miae Chae | Mae Fah Luang district, Chiang Rai | N 20 10 41.5 E 99 42 18.5 | 737 | III |
| 19 | La Hae023 | Mae Fah Luang district, Chiang Rai | N 20 10 41.5 E 99 42 18.5 | 737 | III |
| 20 | Unknown024 | Mae Fah Luang district, Chiang Rai | N 20 10 41.5 E 99 42 18.5 | 737 | III |
| 21 | Khao Khao Chae Bah | Mae Fah Luang district, Chiang Rai | N 20 14 42.6 E 99 33 35.9 | 1028 | III |
| 22 | Chae Mew | Mae Fah Luang district, Chiang Rai | N 20 14 42.6 E 99 33 35.9 | 1028 | III |
| 23 | Chae Yah Yaw Ti | Mae Fah Luang district, Chiang Rai | N 20 14 42.6 E 99 33 35.9 | 1028 | III |
| 24 | Chae Yah Yaw Heu | Mae Fah Luang district, Chiang Rai | N 20 14 42.6 E 99 33 35.9 | 1028 | III |
| 25 | Jar Lo Mah | Mae Fah Luang district, Chiang Rai | N 20 14 42.6 E 99 33 35.9 | 1028 | III |
| 26 | Kha Pah Chae Ne | Mae Fah Luang district, Chiang Rai | N 20 14 42.6 E 99 33 35.9 | 1028 | III |
| 27 | Daw Choo | Mae Fah Luang district, Chiang Rai | N 20 14 42.6 E 99 33 35.9 | 1028 | III |
| 28 | Chair Miaw Rae | Mae Fah Luang district, Chiang Rai | N 20 14 42.6 E 99 33 35.9 | 1028 | I (C) |
| 29 | Unknown034 | Wiang Kan district, Chiang Rai | N 19 59 47 E 100 27 44 | 446 | III |
| 30 | Unknown035 | Mae Sai district, Chiang Rai | N 20 35 198 E 99 876 | 400 | III |
| 31 | Ja Naw Vuey | Mae Fah Luang district, Chiang Rai | N 20 7 30.9 E 99 39 7.6 | 894 | III |
| 32 | Kaw Rue Sue | Mae Fah Luang district, Chiang Rai | N 20 7 30.9 E 99 39 7.6 | 894 | III |
| 33 | Kaw Hom | Mae Fah Luang district, Chiang Rai | N 20 7 30.9 E 99 39 7.6 | 894 | I (C) |
| 34 | Ja Beu Mah | Mae Fah Luang district, Chiang Rai | N 20 7 30.9 E 99 39 7.6 | 894 | III |
| 35 | Khaw Mah Hah040 | Mae Fah Luang district, Chiang Rai | N 20 7 30.9 E 99 39 7.6 | 894 | III |
| 36 | Ja Hae | Mae Fah Luang district, Chiang Rai | N 20 7 30.9 E 99 39 7.6 | 894 | III |
| 37 | Ja Seu Hae | Mae Fah Luang district, Chiang Rai | N 20 7 30.9 E 99 39 7.6 | 894 | III |
| 38 | Ja Bi Ger or Ja Ber Ger | Mae Fah Luang district, Chiang Rai | N 20 7 30.9 E 99 39 7.6 | 894 | III |
| 39 | Pae Hah Ja Naw | Mae Fah Luang district, Chiang Rai | N 20 7 30.9 E 99 39 7.6 | 894 | III |
| 40 | Kaw Mah Hah045 | Mae Fah Luang district, Chiang Rai | N 20 7 30.9 E 99 39 7.6 | 894 | III |
| 41 | Ja Sue Mah | Mae Fah Luang district, Chiang Rai | N 20 7 30.9 E 99 39 7.6 | 894 | III |
| 42 | Ja Na Gui | Mae Fah Luang district, Chiang Rai | N 20 7 30.9 E 99 39 7.6 | 894 | III |
| 43 | Kaw Mah Hah Ja Chi | Mae Fah Luang district, Chiang Rai | N 20 7 30.9 E 99 39 7.6 | 894 | III |
| 44 | O-Sa | Mae Fah Luang district, Chiang Rai | N 20 7 58.2 E 99 38 3.1 | 1107 | I (C) |
| 45 | Khao Maw | Mae Fah Luang district, Chiang Rai | N 20 7 58.2 E 99 38 3.1 | 1107 | III |
| 46 | Che Ba Ma | Mae Fah Luang district, Chiang Rai | N 20 7 58.2 E 99 38 3.1 | 1107 | III |
| 47 | U-Mah Na | Mae Fah Luang district, Chiang Rai | N 20 7 58.2 E 99 38 3.1 | 1107 | III |
| 48 | Che Bah Jui | Mae Fah Luang district, Chiang Rai | N 20 7 58.2 E 99 38 3.1 | 1107 | III |
| 49 | Chae Sa | Mae Fah Luang district, Chiang Rai | N 20 7 58.2 E 99 38 3.1 | 1107 | I (C) |
| 50 | Ka Moo | Mae Fah Luang district, Chiang Rai | N 20 7 58.2 E 99 38 3.1 | 1107 | III |
| 51 | Unknown056 | Mae Lao district, Chiang Rai | N 19 47 15.0 E 99 39 36.0 | 489 | III |
| 52 | Khao Kum057 | Phan district, Chiang Rai | N 19 40 54.8 E 99 38 37.1 | 467 | III |
| 53 | Khao Sim Khao053 | Phan district, Chiang Rai | N 19 40 54.8 E 99 38 37.1 | 467 | I (D) |
| 54 | Chil Mae Jan | Wang Thong district, Phitsanulok | N 16 50 19.0 E 100 22 41.0 | 46 | II |
| 55 | Jaow Num Roo | Wang Thong district, Phitsanulok | N 16 50 19.0 E 100 22 41.0 | 46 | II |
| 56 | Law Take | Wang Thong district, Phitsanulok | N 16 50 19.0 E 100 22 41.0 | 46 | II |
| 57 | Blae Klur | Wang Thong district, Phitsanulok | N 16 50 19.0 E 100 22 41.0 | 46 | II |
| 58 | Ber Por Lo | Wang Thong district, Phitsanulok | N 16 50 19.0 E 100 22 41.0 | 46 | I (B) |
| 59 | San Par Tong | Wang Thong district, Phitsanulok | N 16 50 19.0 E 100 22 41.0 | 46 | I (B) |
| 60 | Bar Nhi | Wang Thong district, Phitsanulok | N 16 50 19.0 E 100 22 41.0 | 46 | I (B) |
| 61 | Pa Ya Lurm Kang | Wang Thong district, Phitsanulok | N 16 50 19.0 E 100 22 41.0 | 46 | I (C) |
| 62 | La Oup | Wang Thong district, Phitsanulok | N 16 50 19.0 E 100 22 41.0 | 46 | I (A) |
| 63 | Khaow Tar Hong | Wang Thong district, Phitsanulok | N 16 50 19.0 E 100 22 41.0 | 46 | I (B) |
| 64 | Hang Pla Lhai | Wang Thong district, Phitsanulok | N 16 50 19.0 E 100 22 41.0 | 46 | I (B) |
| 65 | Mon Pu | Wang Thong district, Phitsanulok | N 16 50 19.0 E 100 22 41.0 | 46 | II |
| 66 | Situ Patenggang | Muang Phayao district, Phayao | N 19 1 43.0 E 99 53 47.0 | 494 | I (C) |
| 67 | Bue Nue Mu | Muang Phayao district, Phayao | N 19 1 43.0 E 99 53 47.0 | 494 | I (B) |

**Table 4** Continued.

| Code | Name | Place of collection | Latitude and longitude of provinces in Thailand | Height (m) | Cluster |
|---|---|---|---|---|---|
| 68 | Khao Lueng Hom | Muang Phayao district, Phayao | N 19 1 43.0 E 99 53 47.0 | 494 | I (A) |
| 69 | Mali Nam Nao | Muang Phayao district, Phayao | N 19 1 43.0 E 99 53 47.0 | 494 | II |
| 70 | CPAC060014 | Muang Phayao district, Phayao | N 19 1 43.0 E 99 53 47.0 | 494 | I (A) |
| 71 | CPAC08043 | Muang Phayao district, Phayao | N 19 1 43.0 E 99 53 47.0 | 494 | II |
| 72 | Nam Ru | Muang Phayao district, Phayao | N 19 1 43.0 E 99 53 47.0 | 494 | I (A) |
| 73 | IR78914-B-22-B-B-B | Muang Phayao district, Phayao | N 19 1 43.0 E 99 53 47.0 | 494 | I (A) |
| 74 | IR81423-B-B-111-3 | Muang Phayao district, Phayao | N 19 1 43.0 E 99 53 47.0 | 494 | I (A) |
| 75 | IR7887-048-B-B-2 | Muang Phayao district, Phayao | N 19 1 43.0 E 99 53 47.0 | 494 | I (A) |
| 76 | IR71700-247-1-1-2 | Muang Phayao district, Phayao | N 19 1 43.0 E 99 53 47.0 | 494 | I (A) |
| 77 | PSL85051-14-2-1-2 | Muang Phayao district, Phayao | N 19 1 43.0 E 99 53 47.0 | 494 | I (A) |
| 78 | CNT86095-42-2-3 | Muang Phayao district, Phayao | N 19 1 43.0 E 99 53 47.0 | 494 | I (A) |
| 79 | Unknown UP-53 | Muang Phayao district, Phayao | N 19 1 43.0 E 99 53 47.0 | 494 | I (D) |
| 80 | IR13240-108-2-2-3 | Muang Phayao district, Phayao | N 19 1 43.0 E 99 53 47.0 | 494 | I (A) |
| 81 | IR15675-81-2-3 | Muang Phayao district, Phayao | N 19 1 43.0 E 99 53 47.0 | 494 | I (D) |
| 82 | IR15795-199-3-3 | Muang Phayao district, Phayao | N 19 1 43.0 E 99 53 47.0 | 494 | I (D) |
| 83 | Bue Wa | Muang Phayao district, Phayao | N 19 1 43.0 E 99 53 47.0 | 494 | I (D) |
| 84 | Nam Ru | Muang Phayao district, Phayao | N 19 1 43.0 E 99 53 47.0 | 494 | III |
| 85 | 2R-43 | Muang Phayao district, Phayao | N 19 1 43.0 E 99 53 47.0 | 494 | I (D) |
| 86 | Ja Chi | Muang Phayao district, Phayao | N 19 1 43.0 E 99 53 47.0 | 494 | III |
| 87 | Khao Sill | Mae Suai district, Chiang Rai | N 19 49 23.6 E 99 33 27.4 | 1142 | I (D) |
| 88 | Jaa Ngee Si | Mae Suai district, Chiang Rai | N 19 49 23.6 E 99 33 27.4 | 1142 | III |
| 89 | Jaa Bae Bae | Mae Suai district, Chiang Rai | N 19 49 23.6 E 99 33 27.4 | 1142 | III |
| 90 | Jaa Da Mor | Mae Suai district, Chiang Rai | N 19 49 23.6 E 99 33 27.4 | 1142 | III |
| 91 | Chep Pea | Mae Suai district, Chiang Rai | N 19 49 23.6 E 99 33 27.4 | 1142 | III |
| 92 | Char-Ku-Lae | Mae Suai district, Chiang Rai | N 19 49 23.6 E 99 33 27.4 | 1142 | III |
| 93 | Ta-Tae-Maa-Cha | Mae Suai district, Chiang Rai | N 19 49 23.6 E 99 33 27.4 | 1142 | III |
| 94 | Jaa-Da-Ma | Mae Suai district, Chiang Rai | N 19 49 23.6 E 99 33 27.4 | 1142 | III |
| 95 | A-The-Ma | Mae Suai district, Chiang Rai | N 19 49 23.6 E 99 33 27.4 | 1142 | III |
| 96 | Khao-Neaw-LeeSaw | Mae Suai district, Chiang Rai | N 19 49 23.6 E 99 33 27.4 | 1142 | I (A) |
| 97 | Lee-Su-Jaa | Mae Suai district, Chiang Rai | N 19 49 23.6 E 99 33 27.4 | 1142 | II |
| 98 | Ta-The-Ma-Ja | Mae Suai district, Chiang Rai | N 19 49 23.6 E 99 33 27.4 | 1142 | III |

Fig. 2C) consisted of 45 varieties (45.92%) collected from Chiang Rai province and the subpopulation 2 (green color, Fig. 2C) included 53 (54.08%) of varieties collected from Phayao and Phitsanulok provinces. The structure analysis suggested differentiation between two subpopulations and clustered them with the geographic area. The fixation index (Fst) for each of the subpopulation was estimated the genetic variation. Genetic differentiation of subpopulation 1 was very strong differentiation (Fst = 0.4456). However, a low Fst value (0.0280) was found in the subpopulation 1 meaning little differentiation.

The subpopulations 1 and 2 had Fst values of 0.4456 and 0.0280, respectively, with an average value of 0.2368 (Table 2). suggested that there was significant divergence within the subpopulation 2.

***Analysis of molecular variance***

The two subpopulations generated from population structural analysis were also determined using analysis of molecular variance (AMOVA) to estimate the percentage of variation among subpopulation and within subpopulation of 98 upland rice varieties. The majority of the genetic variation in upland rice varieties based on structure was due to among subpopulation variation (68%) and the remaining 32% was attributed to individual differences within subpopulation (Table 3), indicating high genetic differentiation between the two subpopulation.

## Materials and Methods

***Plant materials***

A total of 98 upland rice varieties were used in this study (Table 4). Sixty five varieties of upland rice were collected from farmers in 10 districts in Chiang Rai province, Thailand and 21 and 12 varieties were obtained from Dr. Vaiphot Kunjoo, University of Phayao and Phitsanulok Rice Research Center, Thailand, respectively (Fig. 3). Seeds of 98 varieties were planted on cultural tray filled with soil and grown at 25°C for two weeks.

***Genomic DNA extraction***

Genomic DNA was extracted from bulk 14-day-old seeding leaves of each upland rice variety using the Cetyl Trimethyl Ammonium Bromide (CTAB) method previously described by Dolye and Dolye (1987). DNA was quantified by Nano-Drop 1000 spetrophotometer (Thermo Scientific, USA). Final concentration was adjusted to 50 ng/µl for SSR analysis.

***PCR assay***

Nine SSR primer pairs (RM1, RM10, RM19, RM22, RM164, RM241, RM252, RM253 and ORS28) with relatively high polymorphism and distributed across the rice genome were selected for genetic diversity analysis on the basis of

published rice microsattelites. The chromosome positions, repeat motifs and primer sequences for these markers can be found in the rice genome database (http://www.Gramene.org) (Table 1). The polymerase chain reaction (PCR) was conducted in a total volume of 20 µl containing 50 ng of DNA template, 2 µl of 10x PCR buffer, 0.2 mM dNTPs, 2.5 mM MgCl2, 0.2 µM each primer, 0.5 units of *Taq* DNA polymerase (Vivantis, Malaysia). PCR reactions were carried out in Eppendorf Mastercycler Nexus Gradient GSX1 Thermal Cycler (USA). Thermal cycling program involved an initial denaturation at 94ºC for 5 min, followed by 35 cycles of denaturation at 94ºC for 1 min, annealing at 55 ºC for 1 min (60ºC for RM164) primer extension at 72ºC for 30 sec, followed by a final extension at 72ºC for 5 min. The PCR products were separated by electrophoresis in 6% (w/v) denaturing polyacrylamide gels in 1X Tris-borate-EDTA (TBE) buffer at 150 volts for 1 to 2 hours depending on the size of the PCR products. Gels were stained with RedSafe (iNtRON Biotechnology, USA) and visualized under UV light of the Gel document system. Allele sizes were estimated in comparison with 25 bp DNA ladder (Invitrogen, USA).

***Data analysis***

The most intensively amplified bands for each SSR marker were scored. All upland rice varieties were scored for the presence (score '1') or absence (score '0') of the SSR band.
Polymorphic information content (PIC), a measure or the allelic diversity at a locus, was calculated according to Anderson et al. (1993) using the following equation: $PIC_i = 1 - \sum_j {f_{ij}}^2$ where $f_{ij}$ is the frequency of the $j^{th}$ pattern (present and absent) of the $i^{th}$ band. Next, the *PIC* of each primer was calculated as: $PIC = (\sum_{i=1}^{n} PIC_i)/n$ where *n* is the number of bands. Shanon information index (*I*), expected heterozygosity (*He*) and observed heterozygosity (*Ho*) of each loci were calculated in GenAlEx 6.502 software (Peakall and Smouse, 2006, 2012) and the Excel Microsatellite Toolkit (Park 2008), respectively. Genetic similarity among varieties was measured from the matrix of binary data using Jaccard coefficient. A dendrogram was constructed based on the resulting similarity coefficients using the unweighted pair-group method with the arithmetic averages (UPGMA) in the R program (Team, 2015). Analysis of molecular variance (AMOVA) was used to estimate variance among and within populations using GenAlEx 6.502 software (Peakall and Smouse, 2006, 2012). Significance of variance was tested after 999 permutations. From AMOVA, the fixation index (Fst) within the population obtained from GenAlEx 6.502 software (Peakall and Smouse, 2006, 2012). Fst measures the amount of genetic variance. The Fst value of 0 indicates no differentiation between the subpopulation while the Fst value of 1 indicates complete differentiation (Bird et al. 2007). Populations were considered to have very strong differentiation when Fst values were greater than 0.25, strong differentiation when Fst values were between 0.15 and 0.25, moderate differentiation when Fst values were between 0.05 and 0.15 and little differentiation when Fst values were less than 0.05 (Hartl 1980; Mohammadi and Prasanna 2003).
The Bayesian model-based clustering analysis was performed to infer genetic structure and to determine the optimal number of genetic clusters found among upland rice varieties using the software STRUCTURE version 2.3.4 (Pritchard et al. 2000). The number of cluster (K) was set from 1 to 10 and the analysis was repeated 10 times. The burn-in period was 100,000 interactions for each group number K and 100,000 Monte Carlo Markov Chain replications. The optimum value of K value which indicates the number of genetically distinct clusters in the data was obtained by calculating the $\Delta$k value. The $\Delta$k value was calculated based on the change in the log probability of the data between successive K values (Evanno et al. 2005).

## Conclusion

From the similarity coefficient distribution, dendrogram and population structure analysis showed that upland rice varieties in Thailand showed great genetic diversity. This knowledge of genetic diversity and population structure is important in terms of agriculture as they can be potential especially for using these upland rice varieties as a germplasm for the breeding program.


## Acknowledgements

For providing seed samples, we thank Dr. Vaiphot Kunjoo, University of Phayao and Phitsanulok Rice Research Center, Thailand. We are grateful to the farmers from 10 distincts in Chiang Rai province and members of the upland rice lab Mae Fah Luang University for collecting rice seed. This research was financially supported by a research funding project for young scientist researchers, NSTDA and Mae Fah Luang University, Thailand.


## References


Akagi H, Yokozeki Y, Inagaki A, Fujimura T (1996) Microsatellite DNA markers for rice crhomosomes. Theor Appl Genet. 93: 1071-1077.

Anderson JA, Churchill GA, Autrique JE, Tanksley SD, Sorrells ME (1993) Optimizing parental selection for genetic linkage maps. Genome. 36: 181-186.

Bird KA, An H, Gazave E, Gore MA, Pires JC, Robertson LD (2007) Population structure and phylogenetic relationships in a diverse panel of *Brassica rapa* L. Front Plant Sci.8: 321.

Botstein D, White RL, Skolinick M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet. 32(3): 314-331.

Chen X, Temnykh S, Xu Y, Cho YG, McCouch SR (1997) Development of a microsatellite framework map providing genome-wide coverage in rice (*Oryza sativa* L.). Theor Appl Genet. 95: 553-567.

Cho YG, Ishii T, Temnykh S, Chen X, Lipovich L, McCouch SR, Park WD, Ayres N, Cartinhour S (2000) Diversity of microsatellites derived from genomic libraries and GenBank sequences in rice (*Oryza sativa* L.). Theor Appl Genet. 100: 713-722.

Doyle JJ, Doyle LJ (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull. 19: 11-15.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol. 14:2611-2620.

Hartl D L (1980) Principles of Population Genetics, 1st Edn. Sunderland: Sinauer Associates.

Ishikawa R, Yamanaka S, Fukuta Y, Chitrakon S, Bounphanousay C, Kanyavong K, Tang L-H, Nakamura I, Sato T, Sato Y-I (2006) Genetic erosion from modern varieties into traditional upland rice cultivars (*Oryza sativa* L.) in northern Thailand. Genet Resour Crop Evol. 53: 245-252.

Londo J, Chiang Y, Hung K, Chiang T, Schaal B (2006) Phylogeography of Asian wild rice, *Oryza rufipogon*, reveals multiple independent domestication of cultivated rice, *Oryza sativa*. Proc Natl Acad Sci USA. 103: 9578-9583.

McCouch SR, Chen X, Panaud O, Temnykh S, Xu Y, Cho YG, Huang N, Ishii T, Blair M (1997) Microsatellite marker development, mapping and applications in rice genetics and breeding. Plant Mol Biol. 35: 89-99.

Milovanovic, V, Smutka L (2017) Asian countries in the global rice market. ACTA Universitatis agriculturae et silviculturae mendelianae Brunensis 65(2): 679-688.

Mohahty S (2013) Trends in global rice consumption. Rice Today Newsl. 12: 44-45.

Mohammadi SA, Prasanna BM (2003) Analysis of genetic diversity in crop plants—salient statistical tools and considerations. Crop Sci. 43:1235–1248.

Nakagahra M, Hayashi K (1977) Origin of cultivated rice as detected by isozyme variations. Jpn Agric Res Q. 11: 1-5.

Nei M (1973) Analysis of gene diversity in subdivided populations. Proc Natl Acad Sci USA. 70:3321-3323.

Oka HI (1988) Origin of cultivated rice. JSSP/Elsevier, Tokyo/Amsterdam, pp. 254.

Panaud O, Chen X, McCouch SR (1996) Development of microsatellite markers and characterization of simple sequence length polymorphism (SSR) in rice (*Oryza sativa* L.). Mol Gen Genet. 252: 597-607.

Park SDE (2008) Excel Microsatellite Toolkit. Computer program and documentation distributed by the author. Website http://animalgenomics.ucd.ie/sdepark/ms-toolkit/ [accessed 19 July 2019].

Peakall R, Smouse PE (2006) GENALEX6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes. 6: 288-295.

Peakall R, Smouse PE (2012) GenAlEx6.5:genetic analysis in Excel. Population genetic software for teaching and research –an update. Bioinformatics 28: 2537-2539.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus geneotype data. Genetics 155:945-959.

Pusadee T, Jamjod S, Chiang YC, Rerkasem B, Schaal BA (2015) Genetic structure and isolation by distance in a landrace of Thai rice. Proc Natl Acad Sci USA. 106(33): 13880-13885.

Salgotra RK, Gupta B., Bhat JA, Sharma S (2015) Genetic diversity and population structure of Basmati Rice (*Oryza sativa* L.) germplasm collected from north western Himalayas using trait linked SSR markers. PLoS ONE. 10(7): e0131858.

Sato YI (1987) Character association patterns among Japonica rice varieties of Asian origin. In: Proceeding of international symposium "Crop exploration and utilization of genetic resources". Taichung district agricultural improvement station, Changhua, Taiwan, pp. 21-29

Sato YI (1991) Variation in spikelet shape of Indica and Japonica rice cultivars in Asian origin. Jpn J Breed. 41: 121-134.

Shete S, Tiwari H, Elsto RC (2000) On estimating the heterozygosity and polymorphism information content value. Theor Popul Biol. 57: 265-271.

Team RC (2015) R a language and environment for statistical computing. Publishing Web. http://www.r-project.org. Accessed 15 March 2018

Temnykh S, Park W, Ayres N, Cartinhour S, Hauck N, Lipovich L, Cho YG, Ishii T, McCouch SR (1999) Mapping and genome organization of microsatellites in rice (*Oryza sativa* L.). Theor Appl Genet. 100: 698-712.

Vilayheuang K, Machida-Hirano R, Bounphanousay C, Watanabe KN (2016) Genetic diversity and population structure of 'Khao Kai Noi", a Lao rice (*Oryza sativa* L.) landrace, revealed by microsatellite DNA markers. Breed Sci. 66: 204-212.

Wu KS, Tanksley SD (1993) Abundance, polymorphism and genetic mapping of microsatellites in rice. Mol Gen Genet. 241: 225-235.

Wunna, Watanabe KN, Ohsawa R, Obara M, Yanagihara S, Aung PP, Fukuta Y (2016) Genetic variation of rice (*Oryza sativa* L.) germplasm in Myanmar based on genomic compositions of DNA markers. Breed Sci. 66: 762-767.