CHAPTER 3

Genetic Diversity and Population Structure of a Landrace of Thai Rice (*Oryza sativa*)

3.1 Introduction

Cultivated rice, *Oryza sativa*, feeds more people than any other species and is a staple crop for nearly all of Asia. In spite of its central importance for the world's food supply, many aspects of its origin, domestication, and evolutionary genetics remain enigmatic. Rice was domesticated between 8,000 and 10,000 years ago from its wild ancestor, *Oryza rufipogon*, a broadly distributed native species of Asia (Oka, 1988). Domestication of rice appears to have occurred at least twice, once in the region south of the Himalayan mountains of eastern India, Myanmar, and Thailand and again in Southern China (Londo *et al*., 2006). The process of domestication in most crops involves strong selection and genetic bottlenecks, both of which can result for in a precipitous loss of the genetic diversity that is found within the wild ancestor. Rice is no exception; during domestication genetic diversity of cultivated rice was reduced up to 80% from the wild ancestor (Londo *et al*., 2006). The most extreme lost of diversity is seen in modern, high yield rice varieties that are often invariant for many genetic makers. Such loss of diversity can have serious consequences for the crop from susceptibility to epidemic disease to the lack of evolutionary potential for adaptation to novel environments.

In contrast, landraces of rice are thought to be an intermediate stage in the domestication process from wild ancestor to cultivated rice. Landraces are defined as "geographically or ecologically distinctive populations, which are conspicuously diverse in their genetic composition both between landraces and within them," (Brown, 1978) and they are each identifiable by their unique morphologies and well-established local names (Harlan, 1992). Landraces are the traditional varieties of rice, grown by local farmers, which are passed down from generation to generation. Landraces represent a unique and critical source of genetically variable traits that can serve as a resource for future rice improvement. Genetically diverse landraces that are grown and kept by farmers specifically in the vicinity of the center of diversity and domestication of crops are among the world's most important natural resources, resources that are rapidly diminishing. Over the past 40 years, local landraces of rice have been largely replaced by genetically uniform modern varieties over in many parts of Asia including vast regions of China and Vietnam (Pingali and Rajaram, 1998). Thailand, which lies partly in the center of diversity and domestication of rice, is among the exceptions. Local varieties are still grown in some 20% of the country's cultivated rice land, 1.75 m. ha, at the turn of the millennium (OAE, 1998).

While landraces may under some conditions have lower yields than modern varieties, farmers in many regions of the world may favor landraces because they are better adapted to specific local conditions and they are developed for regional uses of rice (Parzies et al., 2004). The genetic variability found within landraces affords the possibility of genetic flexibility; landraces have the potential to adapt to local field conditions and they can adapt to changing environments, farming practices, and specific uses such as animal versus human consumption (McCouch, 2004).

Moreover, the genetic diversity of traditional landrace varieties is the most immediately useful and economically valuable component of rice biodiversity (Wood and Lenne, 1997). In order to efficiently conserve, manage, and use such germplasm resources, an understanding of structure, apportionment and dynamics of local landrace variation is required. Several studies have examined genetic variation and differentiation *among* rice landrace varieties (Li, 2002; Neeraja *et al*., 2005; Fukuoka *et al*., 2006; Bajracharya *et al*., 2006). However, little to no information is available on how genetic diversity is structured within a given landrace. Nor do we know if the genetic variability within landraces, in fact, provides the evolutionary flexibility necessary for local adaptation.

The genetic structure of landrace populations is affected by processes that are usually insignificant in modern high yield crops such as genetic drift, selection, and the influence of mating system. In addition to these "natural" process, the management practices of local farmers strongly influence the distribution and apportionment of variation through the process of seed selection and exchange (Parzies *et al*., 2004; Balma *et al*., 2000). For example, if farmers use seed from only a few plants to establish next years crop, genetic drift due to a bottleneck may reduce genetic diversity (Grayuer *et al*., 2005) and increase genetic differentiation among fields. In contrast, seed exchange among farmers may enhance diversity of local germplasm and increase the genetic similarity between fields. Consequently, farmer's management may affect the dynamic of crop genetics either increasing or decreasing diversity and thus influence the evolutionary dynamics of local crop varieties.

How do these various factors interplay to affect the evolutionary potential of a landrace? In this study we examine genetic diversity within a local variety of Thai

rice, Bue Chomee. Bue Chomee is a traditional landrace of rice used by the Karen, a minority group who populate the mid altitudes of the hills of western and northern Thailand. Bue Chomee is adapted to upland paddy cultivation within an altitudinal range of 600-700 m. (Meesin, 2003). The average annual temperature varies from 25 to 26°C with the annual range of 8 °C. Soil fertility varies among soil types, acrisols (poor fertility with low phosphorus) on slopes of an inclination of 20° to 40°, cambisols on steeper slopes, and ferralsols on more gentle slopes (Schmidt-Vogt, 2001). Bue Chomee is grown outside of the habitat of its wild ancestor, *Oryza rufipogon*.

Here we show that a single landrace of rice is a genetically dynamic system. Specifically the rice variety Bue Chomee (i) is genetically variable with most genetic diversity represented within single field; (ii) genetic variation among fields shows an isolation by distance pattern of genetic differentiation; and (iii) the key factors that shape the structure of landraces are farmers' management and selection of seed along with potential environmental differentiation among regions.

3.2 Materials and Methods

3.2.1 Sample collections

Thirty-three fields, representing 33 subpopulations of the local rice variety, Bue Chomee, were collected from 33 farmers in thirteen villages in Chiang Mai and Mae Hong Son Province of northern Thailand. Abbreviation of each field were designated by their village of origin and the number identifies each farmer who provided the seed, from his/her storage (see Table 3.2.1 and Fig. 3.2.1). Seeds were germinated in petri dishes for 5 days then transferred to 30 cm diameter pots, 10 plants per pot. At the tillering stage leaves from10 to 20 plants of each sample were collected and silica-dried and stored at -20^{o}C. DNA was extracted using a modified cetyltrimethyl ammonium bromide extraction method. The relative purity and concentration of extracted DNA was estimated by ethidium bromide staining on agarose gels compared with known DNA concentration markers.

Table 3.2.1 Locations and sample sizes for 13 populations, representing 13 villages of *Bue Chomee* local rice variety in Chiang Mai (CM) and Mae Hong Son (MHS) province northern Thailand.

Population abbreviation	Location*	No. of subpopulation	No. of individual	Altitude (msl)
HEC	Huai-e-cang, Maewang, CM	13	164	970
NT	Nong-tao, Maewang, CM	4	57	1110
PLR	Pong-lom-rang, Maewang, CM	2	40	1120
PK	Pa-kloy, Maewang, CM	2	40	1130
MLC	Mae-lan-come, Samerng, CM	3	41	730
GSM	Gue-sere, Samerng, CM	1	20	950
WH	Wieng-hang, Wieng Hang, CM	1	20	730
NL	Nhong-lom, Chom Thong, CM	2	40	1070
MG	Muang-glang, Chom Thong, CM	1	20	310
MJ	Mae-cham-Luang, Mae Cham, CM	1	20	940
MT	Mae-tho, Hod, CM	1	20	1190
KSN	Khun-sa-nai, Pai, MHS	1	20	740
HN	Huai-na, Mae Sariang, MHS	1	20	1060

* villages name in Thai see Appendix B-1

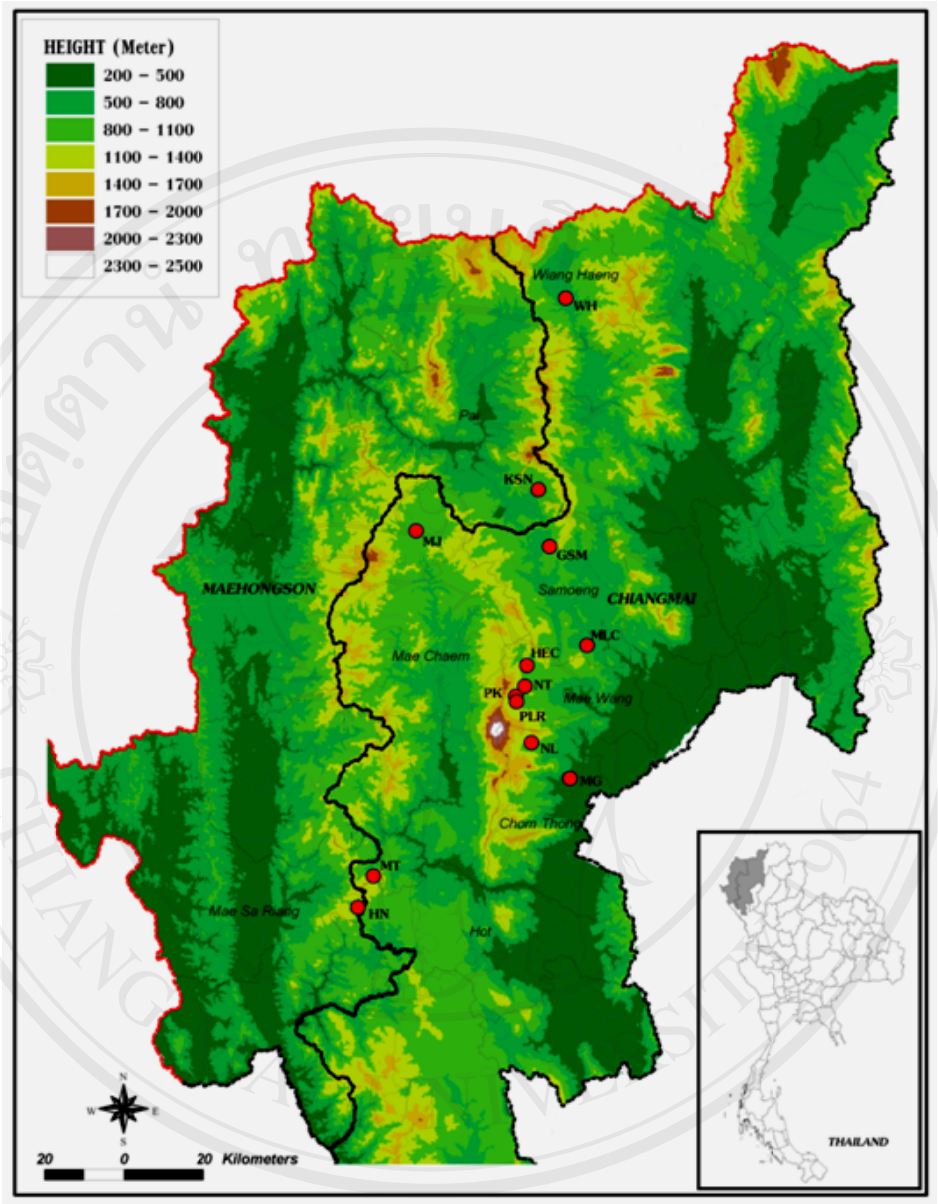

Figure 3.2.1 Sampling locations of 33 subpopulations of *Bue Chomee* local rice (*Oryza sativa* L.) variety from 13 villages, 8 districts in Chiang Mai (CM) and Mae Hong Son (MHS) province, northern Thailand.

3.2.2 Microsatellite analysis

A total of six microsatellite primer pairs that were previously characterized and mapped on the rice chromosomes (Akagi *et al*., 1996; Panaud *et al*., 1996; Chen *et al*., 1997) were chosen for this study, RM1, RM22, RM164, RM241, RM253 and OSR28. Microsatellite polymorphism was analyzed by polymerase chain reaction (Panaud *et al*., 1996). Amplification of DNA was performed in 20 µl reactions consisting of 20-50 ng DNA, 0.25 mM of each dNTP, 0.2 µM of each primers and 0.5 unit of Taq DNA polymerase (Invitrogen©). Amplified products were mixed with loading dye and were separated in 10% polyacrylamide gels by electrophoresis. Gels were stained with ethidium bromide and photographed under UV light.

3.2.3 Data analysis

Standard measures of genetic diversity were calculated, including the effective number of allele (ne), Shannon's information index (I) and estimate of unbiased Nei's (1973) gene diversity (h) using POPGENE 1.32 (Yeh *et al*., 1999) while allelic richness (Rs), number of total alleles (A), inbreeding coefficients (F_{IS}), within population gene diversity (H_S), overall gene diversity (H_T), amount of gene diversity among populations (D_{ST}) and pairwise genetic differentiation (F_{ST}) between villages were calculated using FSTAT 2.9.3 (Goudet, 2001). Genetic parameters of Bue Chomee local rice variety were compared with 4 modern varieties, Suphan Buri 1 (SPR1), Chainat 1 (CNT1), Khao Dawk Mali 105 (KDML105) and RD6.

Wright's coefficient (F_{IS}) (Wright, 1990) was calculated according to the methods of Weir and Cockerham (Weir and Cockerham, 1984). F_{IS} is the mean reduction in heterozygosity of an individual due to non-random mating within a subpopulation. The significance of F_{IS} departures from zero was evaluated using

permutation tests after standard Bonferroni corrections using the computer program FSTAT.

Genetic structure was analyzed by hierarchical analysis of molecular variance (AMOVA) (Excoffier *et al*., 1992) implemented in the software of GeneAlEx6 (Peakall and Smouse, 2006). In addition, we used AMOVA yielded statistic analogous to Weir and Cockerham (Weir and Cockerham, 1984) unbiased F_{ST} estimator, to partition genetic variation into components attributable to differences among villages hierarchical group (F_{VT}), among fields within village hierarchical group (F_{FV}) and among fields across the entire study area (F_{FT}). The significant of F-statistics was tested by permutation, with the probability of non-differentiation for 10000 randomizations.

To illustrate genetic relationships among fields and villages based on their pairwise genetic distances, UPGMA clustering were constructed using C.S. chord genetic distance (Cavalli-Sforza and Edwards, 1967) obtained by PowerMarker V3.0 (Liu and Muse, 2005). MEGA 2 (Kumar *et al*., 2001) was used to construct the dendrograms. Test for isolation by distance was evaluated by assessing the correlation matrix between pairwise geographical distances and C.S. chord distance between villages matrices using a Mantel's test in the program IBD (Bohonak, 2002). A total of 100000 random permutations were performed. Mantel's tests evaluate the significance of correlations between two or more matrices using a permutation procedure that accounts for the autocorrelations of the elements in the matrix. In addition, genetic diversity was tested against geographical longitude by a Spearman' rank correlation coefficient to determine if there was a north-south pattern of differentiation. Furthermore, Spearman' rank correlation coefficient was also used to

test whether genetic diversity increase with the increasing or decreasing of the altitude.

3.3 Results

Six microsatellite loci were analyzed to reveal genetic diversity and structure of a landrace rice variety “Bue Chomee” collected randomly from 33 fields in 13 villages. Abbreviation of each field were designated by their village of origin and the number identifies each farmer who provided the seed, from his/her storage (Table. 3.2.1 and Fig. 3.2.1). Genetic diversity within and between fields and villages were assessed. Genetic relationship among fields and villages were examined. In addition, correlation between genetic distance and geographic distance were tested for isolation by distance structure.

The rice landrace Bue Chomee is genetically variable at all six microsatellite loci. In contrast, no variation was found at these same loci for the improved or modern varieties rice grown in Thailand. A total of 21 alleles at six loci were detected in the 525 individuals of Bue Chomee surveyed from 33 fields in 13 villages. The number of alleles varied by locus with a maximum of six alleles at RM1 to only two alleles at RM22 and OSR 28 (Table 3.3.1). Genetic diversity ranged from 0.493 at RM1 to 0.087 at the RM22 locus (Table 3.3.2). Genetic variation occurred at three hierarchical levels: within fields, between fields within a village, and between villages.

3.3.1 Genetic diversity within farmer’s fields

Individual plants within fields of Bue Chomee are genetically variable with the levels of variation varying by locus and by location. Some fields (WH and KSN, see Materials and Methods) were polymorphic for all 6 loci whereas others fields

contained only 2 polymorphic loci. The effective number of alleles (n_e), Shannon's information index (I), Nei's gene diversity (Nei, 1973) (h), total number of allele (A) and allelic richness (R_S) per field are given in Table 3.3.3. All of the fields are genetically diverse, with Nei's gene diversity (h) ranging from 0.174 in MLC2 to 0.433 in MJ, with an average gene diversity across all fields of 0.435. Effective number of allele (n_e) ranges from 1.280 in MLC2 to 2.069 in MJ, with an average of 1.963. Shannon's information index (I) shows the same trend as Nei's gene diversity, with MJ again being the most diverse (0.696) and with MLC2 showing the least diversity (0.264) with an average of 0.765. Total number of alleles, A, was highest in KSN containing 15 of the 21 total alleles, while MLC2, HEC6 and HEC11 all contained only 9 of the 21 alleles. Allelic richness (Rs) measures the number of alleles independent of population size and allows for comparisons across populations. Average allelic richness (Rs) for the 33 fields is 2.636 with KSN having the highest allelic richness, R = 2.469 while modern high yield varieties were monomorphic for a single allele at each locus. Based on all of these measures of diversity, the field MLC2 in Samerng district, Chiang Mai Province is the least variable of all fields, whereas the fields of MJ, Mae Cham district, Chiang Mai Province, showed the highest level of diversity.

$F_{IS,}$ a measure of heterozygote deviation from Hardy-Weinberg equilibrium, was also calculated for each field. F_{IS} ranged 0.859 to 1, indicating that individuals within fields are mostly homozygous, as would be expected for this inbreeding plant.

Table 3.3.1 Number of alleles per locus and per population

Population	RM1	RM164	RM241	RM253	RM22	OSR28
HEC1	3	2	2	1	2	2
HEC2	3	2	2	1	2	2
HEC3	3	2	2	2	2	1
HEC4	3	1	2	2	2	2
HEC5	1	1	3	2	1	2
HEC6	3	1	2	1	1	1
HEC7	2	2	2	1	1	2
HEC8	2	3	2	1	2	2
HEC9	3	1	2	1	1	2
HEC10	3	2	2	2	2	2
HEC11	1	2	2	1	1	2
HEC12	3	2	2	1	2	2
HEC13	3	2	2	1	1	2
NT1	2	2	3	2	1	2
NT2	3	2	3	2	1	2
NT3	3	2	2	3	1	2
NT4	3	2	2	2	1	2
PK1	3	2	2	1	1	2
PK2	4	1	2	2	1	2
PLR1	3	3	3	1	1	1
PLR2	4	2	3	2	1	2
MLC1	2	3	1	2	1	1
MLC2	2	2	1	2	1	1
MLC3	3	2	1	2	1	2
GSM1	2	2	2	1	1	2
WH	2	4	2	2	2	2
NL1	2	4	3	1	1	2
NL2	3	4	3	1	1	2
MG	4	4	2	1	1	2
MJ	4	3	2	2	1	2
MT	3	3	2	2	1	2
KSN	5	2	2	2	2	2
HN	1	2	1	2	2	2
Mean	2.8	2.2	2.1	1.6	1.3	1.8
Total	6	5	3	3	2	2

Table 3.3.2 Estimates genetic diversity per locus and per field using an unbiased estimator (Nei's 1973)

Population	RM1	RM164	RM241	RM253	RM22	OSR28
HEC1	0.533	0.129	0.533	0	0.248	0.133
HEC2	0.615	0.154	0.462	0	0.282	0.154
HEC3	0.648	0.467	0.363	0.44	0.143	0
HEC4	0.564	0	0.462	0.154	0.282	0.154
HEC5	0	0	0.621	0.167	0	0.545
HEC6	0.709	0	0.509	0	0	0
HEC7	0.409	0.402	0.409	0	0	0.167
HEC8	0.356	0.378	0.533	0	0.2	0.533
HEC9	0.682	0	0.545	0	0	0.409
HEC10	0.564	0.154	0.538	0.154	0.282	0.462
HEC11	0	0.303	0.485	0	0	0.485
HEC12	0.692	0.44	0.527	0	0.143	0.363
HEC13	0.654	0.077	0.154	0	0	0.282
NT1	0.467	0.467	0.733	0.356	0	0.2
NT2	0.529	0.454	0.700	0.233	0	0.5
NT3	0.626	0.357	0.538	0.275	0	0.495
NT4	0.563	0.442	0.521	0.442	0	0.337
PK1	0.426	0.1	0.521	0	0	0.337
PK2	0.553	0	0.521	0.1	0	0.442
PLR1	0.668	0.353	0.616	0	0	0
PLR2	0.705	0.189	0.647	0.1	0	0.395
MLC1	0.44	0.275	0	0.44	0	0
MLC2	0.363	0.264	0	0.495	0	0
MLC3	0.615	0.513	0	0.513	0	0.538
GSM1	0.409	0.526	0.281	0	0	0.105
WH	0.257	0.771	0.514	0.095	0.257	0.181
NL1	0.337	0.716	0.647	0	0	0.521
NL2	0.338	0.717	0.552	0	0	0.495
MG	0.593	0.693	0.485	0	0	0.485
MJ	0.699	0.691	0.529	0.441	0	0.382
MT	0.484	0.626	0.505	0.442	0	0.442
KSN	0.779	0.312	0.368	0.247	0.519	0.455
HN	0	0.209	0	0.209	0.503	0.503
Mean	0.493	0.339	0.449	0.161	0.087	0.318

Table 3.3.3 Genetic parameters of *Bue Chomee* landrace from 33 fields and 4 elite rice varieties (*)

Population	N	n_e	I	h	Rs	A	F_{IS}
HEC1	15	1.428	0.406	0.246	1.963	12	0.915
HEC2	13	1.459	0.425	0.257	1.984	12	0.954
HEC3	14	1.606	0.499	0.319	1.987	12	0.965
HEC4	13	1.419	0.408	0.249	1.977	12	1
HEC5	12	1.417	0.316	0.204	1.659	10	1
HEC6	11	1.446	0.287	0.185	1.500	9	1
HEC7	12	1.330	0.329	0.213	1.663	10	0.88
HEC8	10	1.509	0.468	0.300	2.000	12	1
HEC9	12	1.544	0.381	0.250	1.667	10	1
HEC10	13	1.583	0.523	0.331	2.144	13	1
HEC11	12	1.331	0.287	0.194	1.500	9	1
HEC12	14	1.685	0.520	0.335	1.988	12	1
HEC13	13	1.353	0.312	0.180	1.787	11	0.934
NT1	10	1.680	0.523	0.333	2.000	12	1
NT2	16	1.803	0.598	0.379	2.157	13	0.922
NT3	14	1.682	0.559	0.355	2.142	13	0.938
NT4	20	1.675	0.543	0.365	1.959	12	1
PK1	20	1.373	0.353	0.219	1.782	11	1
PK2	20	1.486	0.413	0.256	1.910	12	1
PLR1	20	1.609	0.435	0.259	1.949	12	1
PLR2	20	1.759	0.550	0.323	2.242	14	1
MLC1	14	1.287	0.284	0.179	1.642	10	1
MLC2	14	1.280	0.264	0.174	1.500	9	1
MLC3	13	1.684	0.489	0.335	1.826	11	1
GSM1	20	1.338	0.312	0.203	1.629	10	1
WH	20	1.788	0.538	0.321	2.272	14	1
NL1	20	1.871	0.571	0.353	2.124	13	0.91
NL2	20	1.772	0.552	0.332	2.231	14	0.887
MG	20	1.743	0.520	0.337	2.187	14	0.859
MJ	20	2.069	0.696	0.433	2.306	14	1
MT	20	1.782	0.616	0.396	2.157	13	1
KSN	20	1.887	0.646	0.407	2.469	15	1
HN	20	1.366	0.328	0.216	1.656	10	1
Total	525	1.963	0.765	0.435	2.972	21	0.970
KDML105*	5	1	0	0	1	6	1
RD6*	5	1	0	0	1	6	1
SPR1*	5	1	0	0	1	6	1
CNT1*	5	1	0	0	1	6	1

Effective number of alleles (n_e), Shannon's Information index (I), unbiased Nei's (1973) gene diversity (h), allelic richness (Rs), total number of allele over all 6 loci (A) and inbreeding coefficient (F_{IS}) were used to showed genetic diversity .

3.3.2 Genetic diversity within and among villages

As expected from the within field measures, rice plants in all 13 villages have moderately high levels of genetic variation; total gene diversity (H_T) is 0.435. The highest average gene diversity within a village (H_S) was in MJ (0.403) and the lowest was GSM (0.203) with an average of 0.332. The village of HEC had the highest number of alleles, 16, while the lowest, 10, is found in the village of GSM and HN (Table 3.3.4). Correspondingly, allelic richness (R_T) is the highest in HEC (2.394), whereas the villages of GSM and HN have the lowest value observed (1.667). In addition, a few rare alleles were detected only in single villages (Table 3.3.4).

3.3.3 Population genetic structure

Total genetic diversity of the Bue Chomee landrace was apportioned into 3 components (Table 3.3.4); genetic diversity among individuals (H_T=0.435), among fields (H_S=0.332) and among villages (D_{ST}=0.103). *F*-statistics were used to apportion the total diversity within and between fields and villages; 24.8% of total variation was due to differentiation among villages, 8.7% is due to differentiation among fields of a villages and 31.4% of the genetic diversity was due to differentiation among individuals (Table 3.3.5). The results of the *F*-statistic analysis confirmed that most genetic diversity of Bue Chomee was due to differentiation among homozygous individuals within fields, with low differentiation among the fields of a village. These results are supported by an AMOVA which indicated that 68% of total variation was placed among individuals, 25% between villages and only 7% was the variation found among fields within village.

Pairwise genetic differentiation (F_{ST}) among the 13 villages ranged from 0.079 to 0.682 (Table 3.3.6). The majority of tests for pairwise genetic differentiation (F_{ST})

between 13 villages were significant except between PLK and PL villages. The village of HN, Mae Sarieng district in Mae Hong Son province, showed high genetic differentiation with all other villages with an average F_{ST} of 0.535 and a range of 0.405 to 0.682. The village of MJ, Mae Cham district in Chiang Mai province had the lowest average differentiation; F_{ST} = 0.196.

Table 3.3.4 Population structure and genetic diversity of *Bue Chomee* among 13 villages

Population	n	N	n_e	A	R_T	R_A	H_S	H_T	D_{ST}	F_{IS}
HEC	164	13	1.620	16	2.394	-	0.251	0.322	0.071	0.978
NT	57	4	1.773	14	2.259	-	0.371	0.386	0.015	0.965
PK	40	2	1.431	13	2.043	1	0.238	0.240	0.002	1.000
PLR	40	2	1.684	15	2.376	1	0.291	0.298	0.007	1.000
MLC	41	3	1.427	12	1.914	-	0.229	0.253	0.024	1.000
GSM	20	1	1.338	10	1.667	-	0.203	0.203	0	1.000
WH	20	1	1.788	13	2.167	2	0.321	0.321	0	1.000
NL	40	2	1.871	14	2.322	1	0.342	0.354	0.012	0.897
MG	20	1	1.743	12	2.000	1	0.337	0.337	0	0.835
MJ	20	1	2.069	14	2.333	1	0.403	0.403	0	1.000
MT	20	1	1.782	13	2.167	1	0.396	0.396	0	1.000
KSN	20	1	1.887	14	2.333	2	0.341	0.341	0	1.000
HN	20	1	1.366	10	1.667	-	0.216	0.216	0	1.000
Total	**525**	**13**	**1.963**	**21**	**2.972**	**5**	**0.332**	**0.435**	**0.103**	**0.970**

n=number of individual, N=number of field
Effective number of alleles (n_e), total number of allele over all 6 loci (A), average allelic richness per population over 6 loci (R_T), number of rare allele (R_A), average gene diversity within the populations (H_S), total gene diversity (H_T), genetic diversity among population (D_{ST}) and inbreeding coefficient (F_{IS})

Table 3.3.5 Analysis of molecular variance (AMOVA) for 525 individual plants from 33 fields in 13 villages

Source	d.f.	SS	Variance component	% of the total variance	*F*-statistics	P-value
Among villages	12	723.301	60.275	24%	F_{VT}=0.248	<0.001
Among fields/village	20	173.473	8.674	7%	F_{FV}=0.087	<0.001
Within field	492	1817.197	3.693	68%	F_{FT}=0.314	<0.001
Total	524	2713.971	72.642			

F_{FT} refer to among fields to total; F_{FV} refer to among fields to village; F_{VT} refer to among villages to total

Table 3.3.6 Pairwise genetic differentiation (F_{ST}) among 13 villages

	HEC	NT	PK	PLR	MLC	GSM	WH	NL	MG	MJ	MT	KSN
HEC												
NT	0.122**											
PK	0.036**	0.213**										
PLR	0.062**	0.083**	0.079ns									
MLC	0.298**	0.360**	0.384**	0.371**								
GSM	0.442**	0.385**	0.543**	0.447**	0.371**							
WH	0.241**	0.127**	0.370**	0.198**	0.360**	0.419**						
NL	0.120**	0.197**	0.170**	0.207**	0.226**	0.472**	0.264**					
MG	0.204**	0.214**	0.212**	0.209**	0.374**	0.507**	0.297**	0.076*				
MJ	0.188**	0.118**	0.207**	0.090**	0.277**	0.312**	0.180**	0.172**	0.132*			
MT	0.308**	0.210**	0.332**	0.220**	0.363**	0.162**	0.265**	0.322**	0.253**	0.090*		
KSN	0.313**	0.300**	0.336**	0.294**	0.454**	0.552**	0.368**	0.238**	0.225**	0.176**	0.348**	
HN	0.569**	0.443**	0.642**	0.542**	0.682**	0.654**	0.558**	0.562**	0.518**	0.405**	0.441**	0.407**

Significance after standard Bonferroni corrections NS: non-significant; *P<0.05; **P<0.01

3.3.4 Cluster analysis and Isolation by distance

UPGMA clustering based on the genetic distances among fields (Fig. 3.3.2) and villages (Fig. 3.3.3), shows a spatial pattern that corresponds to geographic location. Bue Chomee of HEC, NT, PK and PLR villages are grouped into the same cluster (Fig. 3.3.2). These populations are in the same geographical region, the Maewang District of Chiang Mai Province. Likewise, rice from NL and MG villages, located in the Chom Thong District of Chiang Mai Province and rice from MLC and GSM villages located in the Samerng District of Chiang Mai Province were grouped into the same cluster. These results suggest that there maybe a geographical structure to genetic variation, specifically an isolation by distance relationship between villages. Testing of isolation by distance was done by a Mantel's test in the IBD program (Bohonak, 2002) and indicated a significant correlation between genetic distance and geographic distance ($r = 0.599$) (Fig 3.3.4) as well as significant isolation by distance based on the Mantel's test ($p<0.005$ from 100000 randomizations). A Spearman rank correlation coefficient did not detect an overall north-south pattern of genetic differentiation ($r_s=0.109$). We note that the village (MJ) located near the geographical middle of our study area and which is surrounded by 6 other villages, has the highest genetic diversity but shows the lowest average genetic differentiation (F_{ST}) between other villages. In addition, there was very low positive correlation between genetic diversity and the altitude ($r_s=0.139$).

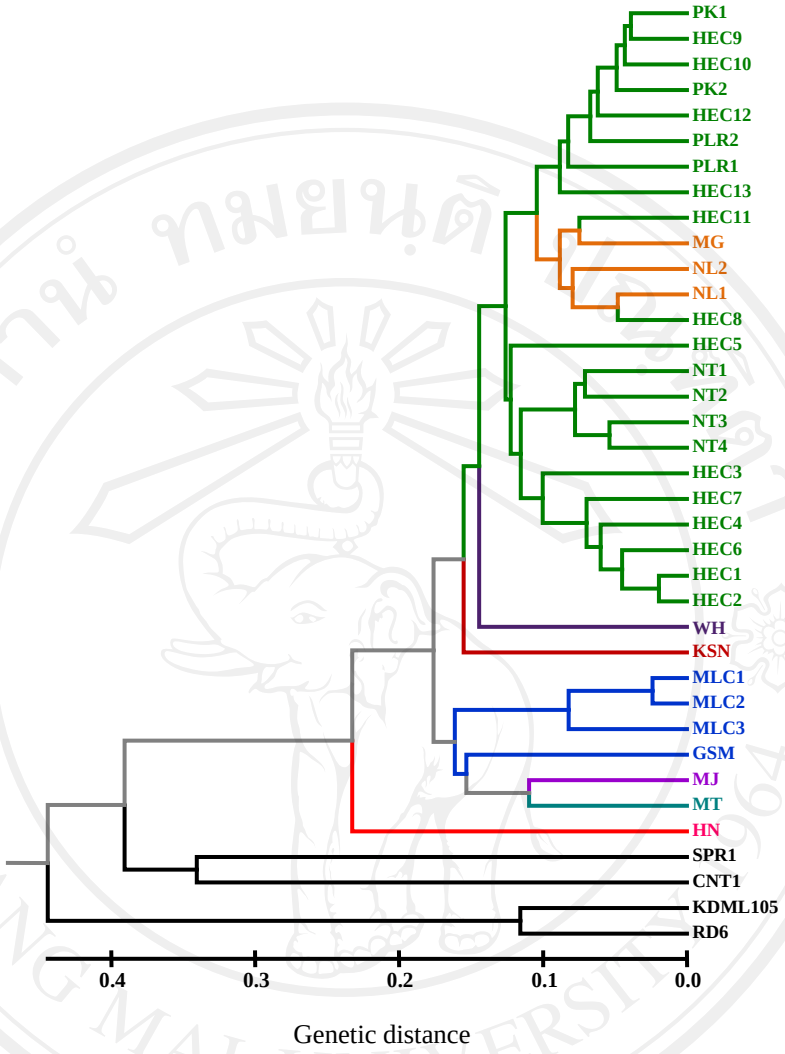

Figure 3.3.2 Dendrogram based on C.S. Chord (1967) genetic distance clustering by UPGMA methods showing genetic relationship among 33 seed lots of *Bue Chomee* from 13 villages and 4 elite rice varieties (KDML105, RD6, CNT1 and SPR1).

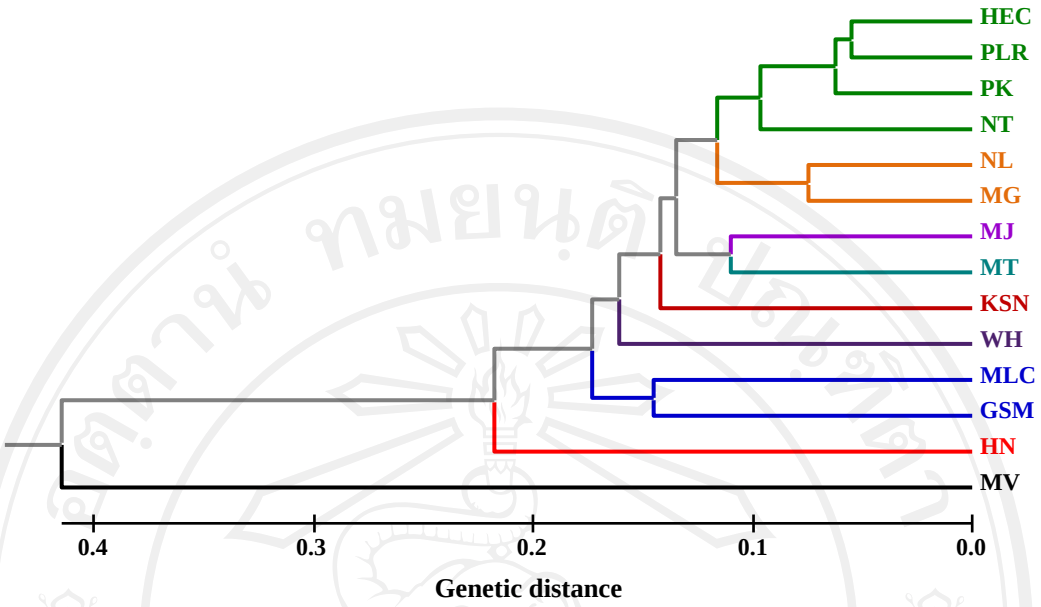

Figure 3.3.3 Dendrogram constructed based on C.S. chords (1967) genetic distance showed genetic relation among 13 Bue Chomee populations from 13 villages in 8 districts indicated by different color in 2 mountainous provinces Northern Thailand and elite rice varieties (MV).

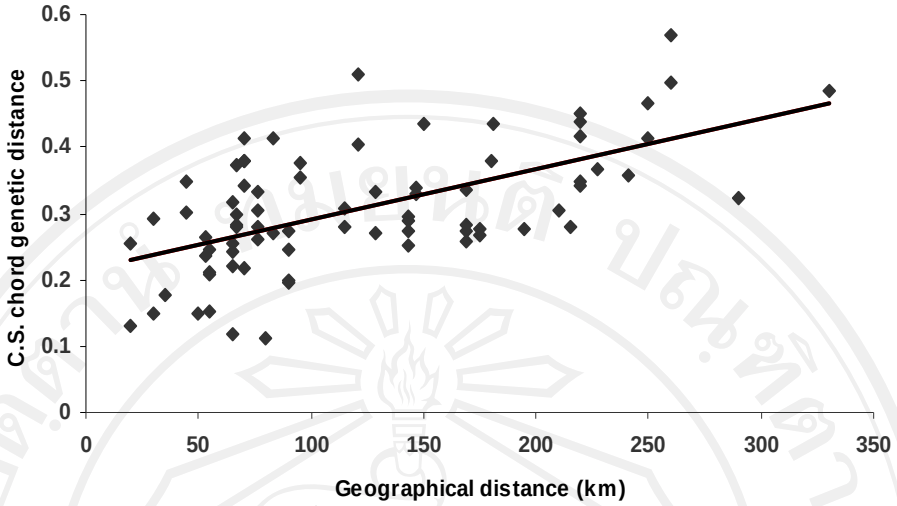

Figure 3.3.4 A Mantel's test for correlation between C.S. chords (1967) genetic distance and geographic distance (km) showed high correlation, r=0.599 p<0.0058 from 100000 randomizations.

3.4 Discussion

Domestication of plants and animals and their husbandry by agriculture is one humankind's greatest scientific and technological achievements. Domestication is an evolutionary process where strong selections for specific traits, combined with a series of population bottlenecks, greatly alter the genetic structure of populations and the underlying genetic architecture of phenotypic traits. Modern, elite varieties of many crops have low to no genetic variation within cultivars, and often only 20% of the total diversity contained within the wild ancestor is maintained through domestication e.g. cassava, (Olsen and Schaal, 2001); soybeans, (Hyten *et al*., 2006); and rice, (Londo *et al*., 2006). This loss of genetic diversity has profound consequences for agriculture. Monocultures of genetically similar individuals are susceptible to epidemic outbreaks of disease such as the Southern corn blight of the US during the 1970's which caused an estimated loss of over a billon dollars in 1970 (Ullstrup, 1972). Moreover, loss of genetic diversity has profound implications for crop improvement. As a crop passes through the domestication bottleneck, potentially useful traits may be lost, thereby reducing the set of phenotypes that can subsequently be used by plant breeders for crop improvement. The frequent paucity of new traits in crop germplasm collections has lead to increased efforts to conserve wild relatives of domesticated plants, as a reservoir of phenotypes for future crop improvement.

Landraces provide another reservoir of useful traits. Domestication must necessarily be a gradual process, occurring over hundreds and even thousands of years. Landraces represent an intermediate stage of domestication between the wild ancestor and modern elite varieties. Crop landraces of grasses can maintain genetic

diversity, for example, in maize (*Zea mays* ssp. *mays*) (Buckler *et al*., 2001), millet (Adoukonou-Sagbadja *et al*., 2007), cowpea (Tosti and Negri, 2005), common bean (Gomez *et al*., 2004), and sweetpotato (He *et al*., 2006). Because Bue Chomee is genetically diverse, and it is grown under variable environmental and agricultural conditions, Bue Chomee has the potential be a dynamic system that can undergo genetic changes in response to evolutionary forces.

What factors may influence the distribution of genetic variation within landrace rice? Natural processes, such as drift, selection, gene flow and hybridization will affect variation. But, because rice is a domesticated species, we expect that cultivation practices such as seed exchange and selection by farmers will play a predominant role. Evidence for these processes is reflected in the lack of genetic uniformity among Karen villages. While most genetic variation is apportioned among individuals and little variation was due to differences among the fields of a village, significant genetic differentiation occurs among the villages that are located at varying altitudes and on different soil types. Most surprising is the significant correlation between genetic distance and geographical distance (Fig 3.3.4). This correlation and the significant Mantel's test reflects an isolation by distance population structure where populations at greater distances are more genetically dissimilar than those populations that are geographically close. Genetic isolation by distance is a dynamic process in space and time that produces changing genetic composition (Epperson and Li, 1996). These data suggest that Karen landrace rice is a dynamic, evolving genetic system, rather that the static set of genotypes found in a modern variety.

How does isolation by distance occur in a cultivated plant? In nature isolation by distance is the result of limited gene flow, where the probably of gene flow between two populations is a function of the geographical distance between them (Slatkin, 1993). For inbred cultivated rice where little to no pollen flow occurs, gene flow must occur by seed movement, specifically seed exchange between farmers. A key aspect of traditional agricultural systems throughout the world is the frequent exchange of seeds by farmers (Zeven, 1999). In these Karen villages, seed exchange among local farmers is also frequent. Sirabanchongkran *et al*. (2004) have analyzed the patterns of seed exchange and social networks within these Karen villages. The social structure of the community, such as marriage patterns or kinship relationships, play a role in farmer's seed exchange preferences whether within or between villages. Seeds are exchanged more frequently among the farmers within a village, resulting in high within field genetic diversity. Seed exchange also occurs among villages although at a lower frequency. Exchange among villages is most often between neighboring villages. The genetic data for Bue Chomee conform to these patterns of seed exchange. Little to no differentiation occurs among the fields within a village, reflecting the frequent exchange of seed between village farmers. Differentiation occurs between villages, reflecting more limited seed exchange. Interestingly, we find that the village located in the geographical center of the study area (MJ, Mae Cham District, Chiang Mai Province) has the highest genetic diversity with the lowest average genetic differentiation (F_{ST}) from other villages. We expect that these patterns of genetic differentiation, based on seed exchange networks, are widespread in traditional agricultural systems. Similar results are found in maize where seed

exchange between close communities plays a role in reducing maize populations structure (Perales *et al*., 2005).

Next, we consider the potential role of selection in population differentiation within Bue Chomee. The levels of variation and population structure observed for the cultivated Bue Chomee are surprisingly similar to many native species of plants such as wild lima bean (*Phaseolus lunatus* L.) (Martinez-Castillo *et al*., 2006), wild relatives of tomato (*Solanum pimpinellifoliu*) (Caiceda and Schaal, 2004), teosinte-maize wild relative (Fukunaga *et al*., 2005) and sea beet *(Beta vulgaris* ssp. *maritime*) (Cureton *et al*., 2006). Moreover, genetic variation, restricted gene flow and environmental heterogeneity, all found in Bue Chomee, may lead to genetic differentiation. In a common garden analysis of morphological differentiation of Bue Chomee from these villages, Meesin (2003) found differentiation for flowering time, suggesting selection for local environmental conditions. For example in village MLC1 and MLC2, Samerng district, Chiang Mai province located at the higher elevation, flowering time was 107 days later than for other sites. Genetic differentiation was also observed for the traits plant height, percentage of spikelet sterility, harvest index and seed characteristics.

Local adaptation plays an important role in maintaining yields in traditional agricultural systems. Selection for adaptation to each village environment by the farmer's seed selection enhances overall crop diversity and maintains evolutionary flexibility (Alvarez *et al*., 2005). Farmer selection in combination with natural selection results in landraces with high levels of adaptation to biotic and abiotic stresses and as well as for agricultural traits (Almekinders *et al*., 1994). For example, the genetic diversity of *Phaseolus vulgaris* landraces in Italy has been shaped by local

adaptation to microenvironments (Tiranti and Negri, 2007) and in wheat, selection by farmers has strongly influenced the evolution of neutral loci (Goldringer *et al*., 2001).

Our results are consistent with the importance of cultural practices for maintaining the diversity of crop germplasm. Anthropologists have long advocated that human knowledge be included as a component of plant genetic resources for species directly managed and manipulated by humans (Orlove and Brush, 1996). Four components of farmers' management have been identified effecting crops diversity, seed flow, variety selection, variety adaptation and seed selection and storage (Bellon, 1997). Farmer's seed selection is strongly influenced by local preferences, customs and culture, and allows for differentiation between varieties from the same farmer or between farmers (Pressoir and Berthaud, 2004). The importance of farmer practices, shaped by economics, culture, and in some cases religion have been documented for traditional maize varieties (Pressoir and Berthaud, 2004; Louette and Smale, 2000), pearl millet in south-western Niger (Allinne *et al*., 2007), and India (vom Brocke *et al*., 2003), cassava in South America (Elias *et al*., 2004), sorghum in Cameroon (Alvarez *et al*., 2005) and cowpea in central Italy (Tosti and Negri, 2005). In rice different ethic groups maintain and grow their own varieties with highly specific uses (brewing, animal feed, sweets). Some groups such as the Karen, have elaborate rules about how the family's rice varieties are inherited and conserved among siblings, and how they are shared or not shared. All of these practices, rooted in the varying practices and cultures of traditional farmers, results in a dynamic population genetics for these landraces.

The dynamic nature of landraces makes them particularly important sources of germplasm for breeding programs (Almekinders *et al*., 1994). For example, some

local Karen rice varieties in Thailand show variation in traits conferring adaptation to fluctuation biotic and abiotic stress, e.g. tolerance to the pest, gall midge (Supamongkol, 2006) and also are variable for traits of agricultural interest such as an observed two fold variation in iron concentration in the grain (Pintasen *et al*., 2007). Given the importance of landrace germplasm for enhancing crops and the essential role of farmer's practices in maintaining the variation within landraces, on-farm, *in situ* conservation is an essential strategy for future crop breeding efforts.