CHAPTER 2

Genetic Diversity and Population Structure of *Oryza sativa* Primary Gene Pool in Thailand

2.1 Introduction

In Asia, *Oryza sativa* primary gene pool consisted of 3 components; wild ancestor (*Oryza rufipogon* Griff.), weedy form (*Oryza sativa* f spontanea) and cultivated rice (*Oryza sativa* L.) (Chang, 1976). During domestication ca. 8,000 - 10,000 years ago, genetic diversity of cultivated rice (*Oryza sativa* L.) was reduced more than 50% of its wild ancestor (*Oryza rufipogon* Griff.) due to drift, bottleneck and selection including natural or human selection (Oka 1988; Londo *et al*., 2006). Wild ancestor of Asian cultivated rice, common wild rice (*O. rufipogon* Griff.), is an important potential reservoir of genetic variation for crop improvement.

Common wild rice is widespread in a number of Asian countries including China, Laos, Thailand, Myanmar, Nepal, Bangladesh, Taiwan, Vietnam, Cambodia, Indonesia and India, keep differentiation in local populations (Vaughan *et al*., 2008a). In some areas, common wild rice differentiated into two types, perennial type and annual type. The major differences between these two types are life-history trait and mating system (Sano and Morishima 1982; Barbier 1989a and 1989b).

Cultivated rice (*O. sativa* L.), the domestication race derived from common wild rice illustrated the reduction of genetic variation due to domestication bottleneck (Doebley *et al*., 2006). As rice is the agricultural crop therefore, it was faced with the

strong selections of both natural and farmers preferences lead to severe genetic variation reduction (Doebley *et al*., 2006) and differentiation into several different types e.g. *indica* and *japonica* subspecies, glutinous and non-glutinous types, upland, paddy and submergence rice types. Even though cultivated rice is autogamous, however, low but significant degrees of outcrossing rate are found and affected genetic variation of cultivated rice (Niruntrayakul, 2008). The present chapter focusing on 2 types of cultivated rice systems, modern and improved traditional varieties as most rice cultivation in Thailand farmers cultivate modern variety in the irrigation system while improved traditional varieties were grown mostly in rain-fed system.

Weedy form, the intermediated form between wild ancestors and domesticated species, originated from three principle evidences (de Wet and Harlan, 1975); 1) from wild colonizers through selection towards adaptation to continuous habitat disturbances; 2) as derivatives of hybridization between wild and cultivated species; 3) from abandoned domesticates through selection towards a less intimate association with man. However, weedy form of Asian cultivated rice called weedy rice has been proved to originate from interspecific hybridization between common wild rice (*O. rufipogon* Griff.) and cultivated rice (*O. sativa* L.) (Niruntrayakul, 2008). Therefore, weedy rice usually and frequently found in the area where wild and cultivated rice are co-exist, even in the area where absence of wild populations (Oka, 1988).

However, in the past decade weedy rice have been reported to invade in many rice cultivation areas in Thailand (Maneechote, 2004) and became a serious weed problem where first found in Central plain. Recently, weedy rices have been increasingly reported widespread in many directed-seed paddy fields in Lower North

and Northeast. The 2 hypotheses on origin of widespread invasive weedy rice in various areas in Thailand will be tested. First, weedy rice was spread from Central region to Lower North or Northeast paddy fields by seed flow. Seed flow/seed migration was the result of weedy rice seed contamination in seed stock from farmer-to-farmer or by harvesting machine from invaded paddy fields to the clean fields. The second is weedy rice in various areas originated from independent hybridization between local common wild rice and popular cultivated rice in each area. For this reason, common wild rice and cultivated rice should share the same habitat and have overlap flowering period.

Jarvis and Hodgkin (1999) suggested that hybridization or introgression between wild relatives and its cultivar may cause the increasing or decreasing of genetic diversity of both components. As the results of such hybridization population genetic structure of the cultivar, wild relatives and weedy form are supposed to be the ongoing evolutionary processes concurrently with rice evolution. For farmers, weedy rice is the aggressive weed that reduced crop yield, and inferior crop quality. For plant breeder, on the other hand, weedy rice is an alternative resource of genetic variation for selection of benefit traits in breeding programs.

Therefore, the objectives of the present study are as followed: i) To evaluate population structure and genetic diversity of common wild rice, weedy rice and cultivated rice; ii) to determine evolutionary factors influenced population genetic structure of *O. sativa* primary gene pool and iii) to trace the origin of widespread weedy rice in various areas.

2.2 Materials and Methods

2.2.1 Genetic materials

Common wild rice (Table 2.2.1) and weedy rice (Table 2.2.2) samples were surveyed and collected in Upper North, Lower North, Central and Northeast of Thailand during 2002 to 2006. The optimum time for surveying the distribution of common wild rice in Thailand was in October until December while weedy rice was depending on cultivated rice cultivation. During the survey, the following data were recorded; geographical reference (UTM), habitat condition, and life-history trait type (see Table 2.2.4 for the classification between perennial and annual types) for wild rice. For weedy rice, companion cultivated rice were recorded. Weedy rice samples were partially kindly provided by Dr. Chanya Maneechote, Plant Protection Research and Development Office, Department of Agriculture. Leaves of individual plant of common wild rice and weedy rice were collected in the fields, 10 plants per population. Leaves were silica-dried following the method desCr1bed by Chase and Hill (1991). Dried leaf samples were subsequently assessed for genetic diversity and population structure using 12 microsatellite loci.

Cultivated rice consisted of 37 varieties including, 23 modern varieties (MV) and 14 improved traditional varieties (ITV) were kindly provided from Thai Rice Research institute, Department of Agriculture (Table 2.2.3). Seeds were germinated in petri dishes for 5 days then transferred to 30 cm diameter pots, 10 plants per pot at Department of Plant Science and Natural Resources, Faculty of Agriculture, Chiang Mai University. At the tillering, leaves of 10 individuals each variety were collected and silica-dried following the method desCr1bed by Chase and Hill (1991) and stored

at -20^{o}C until used. Dried leaf samples were DNA extracted and subsequently assessed genetic diversity and population structure using 12 microsatellite loci.

Table 2.2.1 Common wild rice accession, location and habitat in this study

No.	Accession No.	Region	Province	Habitat condition	Type*	Year collection	UTM
1	UN1	Upper North	Chiang Rai	ditch near rice fields	P	2006	
2	UN2		Chiang Rai	deep marsh	P	2006	
3	UN3		Chiang Mai	irrigation canal	P	2003	
4	UN4		Lumphun	small marsh in rice field	P	2005	N 18.57297 E 098.98835
5	UN5		Lumphun	canal in longgan orchard	P	2003	N 18.57335 E 098.99501
6	LN1	Lower North	Phitsanulok	abandon fields, marsh	P	2004	N 16.56934 E 100.20102
7	LN2		Phitsanulok	marsh	P	2004	N 16.48659 E 100.21312
8	LN3		Phichit	marsh	P	2004	N 16.42549 E 100.31940
9	LN4		Sukhothai	abandoned field near rice field	P	2004	N 16.87597 E 099.79168
10	LN5		Phetchabun	abandon fields	P	2004	
11	CT1	Central	Ang Thong	ditch near rice field	P	2006	
12	CT2		Phra Nakhon Si Ayutthaya	swamp	P	2006	N 14.50177 E100.39212
13	CT3		Phra Nakhon Si Ayutthaya	abandoned rice field	P	2006	N 14.48885 E 100.47448
14	CT4		Nakhon Nayok	abandoned field near rice field	P	2006	N 14.06215 E 101.18994
15	CT5		Sing Buri	abandoned fields	P	2006	
16	CT6		Suphan Buri	swamp	P	2006	
17	CT7		Kanchanaburi	small canal near rice fields	P	2002	
18	CT8		Kanchanaburi	abandoned field near rice fields	P	2002	
19	CT9		Prachin Buri	deep swamp, *in situ* conservation area	P	2003	
20	NE1	Northeast	Sa Kaeo	abandoned fields	S	2005	N 14.09166 E 102.69646
21	NE2		Buri Rum	small ditch	P	2004	
22	NE3		Buri Rum	abandoned fields	P	2005	N 14.68020 E 102.55420
23	NE4		Buri Rum	shallow ditch	A	2005	N 14.66321 E 102.61402
24	NE5		Buri Rum	swamp	P	2005	N 14.62856 E 103.06692
25	NE6		Buri Rum	near rice field	A	2005	N 14.62856 E 103.24828
26	NE7		Nakhon Ratchasima	shallow ditch	A	2004	
27	NE8		RoiEt	ditch near rice fields	P	2004	
28	NE9		RoiEt	shallow ditch near rice fields	A	2005	N 16.07418 E 103.61746
29	NE10		RoiEt	shallow ditch near rice fields	A	2005	N 16.03048 E 103.86951
30	NE11		Udon Thani	shallow field near rice fields	A	2005	N 17.36042 E 102.96401
31	NE12		Udon Thani	shallow swamp near rice fields	A	2005	N 14.65276 E 102.85075
32	NE13		Surin	shallow swamp near rice fields	A	2004	
33	NE14		Surin	abandon field	A	2005	N 14.62291 E 103.89379
34	NE15		Ubon Ratchathani	shallow ditch near rice fields	A	2004	
35	NE16		Yasothon	ditch near rice fields	P	2005	N 15.94862 E 104.04889
36	NE17		Yasothon	ditch near rice field	P	2005	N 15.79287 E 104.17415
37	NE18		Maha Sarakham	marsh	P	2005	N 16.19982 E 103.26114
38	NE19		Maha Sarakham	shallow marsh	A	2005	N 16.25317 E 103.03255
39	NE20		Sakon Nakhon	deep swamp, *in situ* conservation area	P	2003	
40	NE21		Si Saket	shallow ditch	A	2004	
41	NE22		Si Saket	ditch	P	2005	N 15.09111 E 104.33101
42	NE23		Nong Khai	ditch	P	2004	

***** Life-history trait type classification sees **Table 2.2.4**

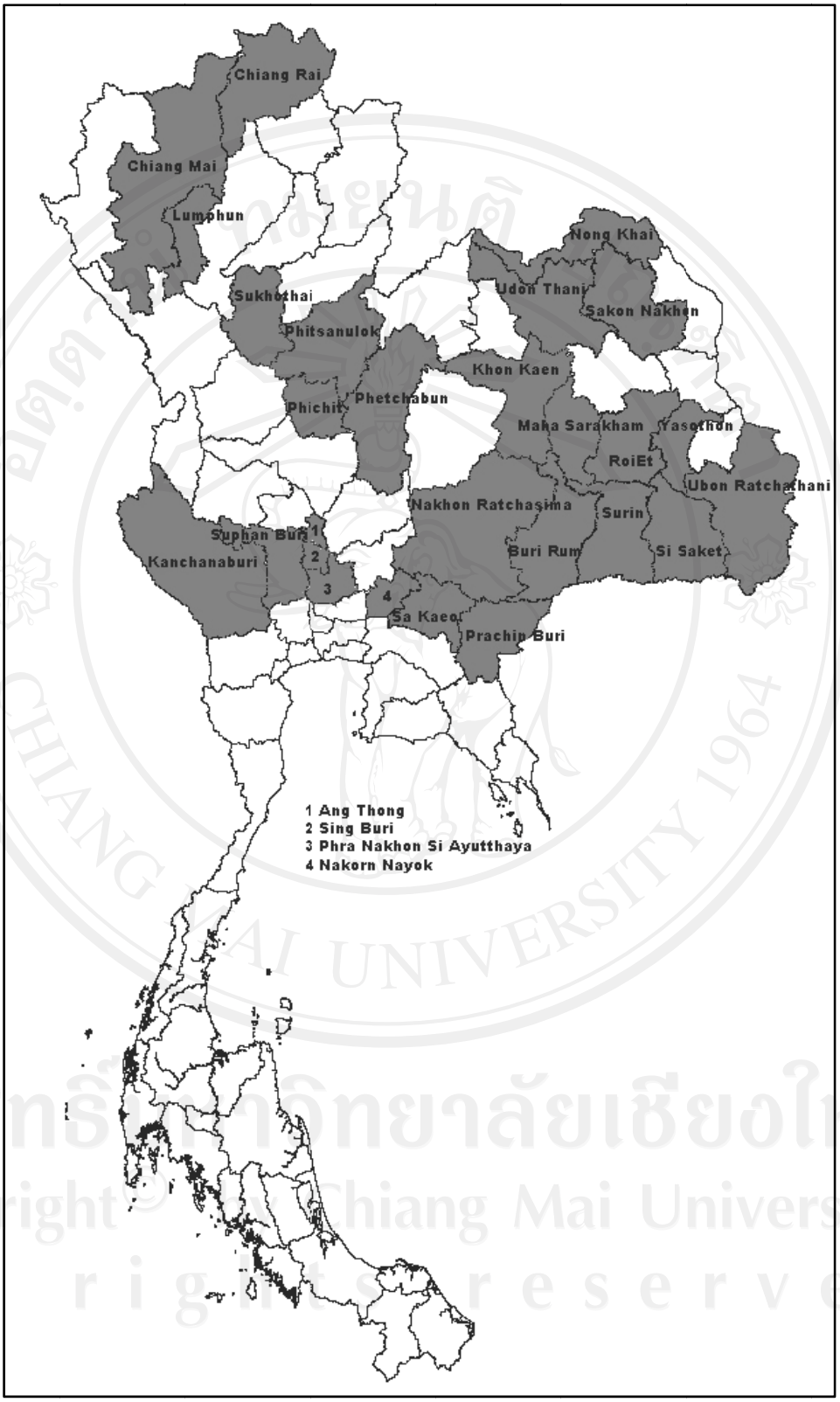

Figure 2.3.1 Location of common wild rice sampling in Thailand of the present study

Table 2.2.2 Location and description of weedy rice populations

No.	Weedy rice Accession No.	Region	Province	Rice variety in invaded fields*	Year collection	UTM
1	LNWE1	Lower North	Phitsanulok	ITV/ MV	2006	
2	LNWE2		Phichit	MV	2006	
3	CTWE1	Central	Kanchanaburi	MV	2002	
4	CTWE2		Kanchanaburi	MV	2002	
5	CTWE3		Suphan Buri	MV	2003	
6	CTWE4		Suphan Buri	MV	2003	
7	CTWE5		Suphan Buri	MV	2003	
8	NEWE1	Northeast	Yasothon	ITV	2005	N 15.72180 E 104.24088
9	NEWE2		Surin	ITV	2005	N 14.78988 E 103.83912
10	NEWE3		Si Saket	ITV	2006	N 14.95359 E 104.20601
11	NEWE4		Buri Rum	ITV/ MV	2005	N 14.87909 E 103.30641
12	NEWE5		Buri Rum	ITV	2005	N 14.62856 E 103.24280

* **MV**=Modern variety; **ITV**=Improved traditional variety

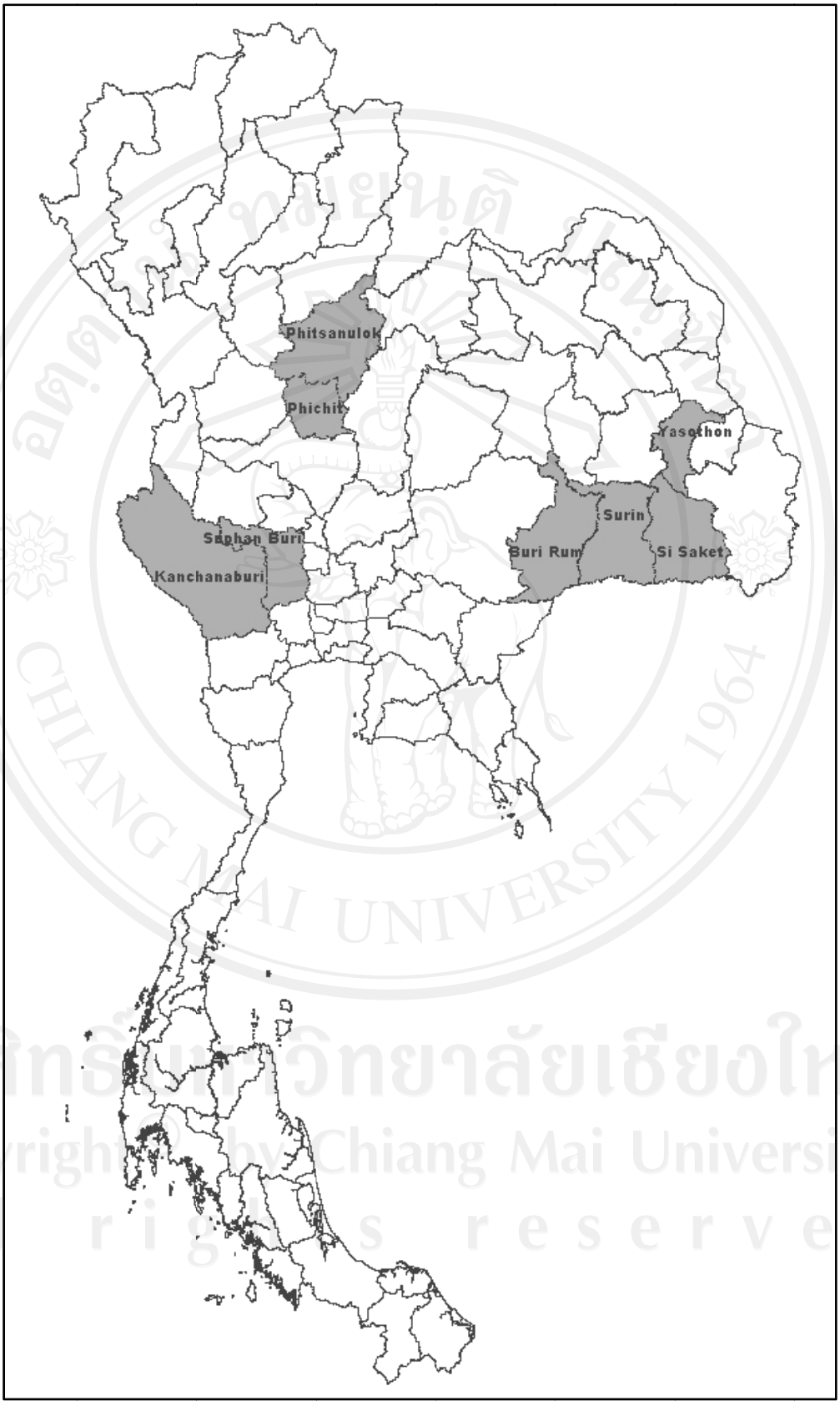

Figure 2.3.2 Location of weedy rice sampling in Thailand of the present study

Table 2.2.3 Name, type, pedigree description and year released of cultivated rice in the present study

No.	Variety	Abbreviation	Thai name	Type*	Type**	DesCr1ption	Year released
1	Doh Hawm26	DH26	ดอหอม26	NG	ITV	Pure line selection from traditional variety	
2	Dawk Mali3	DML3	ดอกมะลิ3	NG	ITV	Pure line selection from traditional variety	
3	Kum Doi Sa Ket	KDK	กำดอยสะเก็ด	G	ITV	Pure line selection from traditional variety	2005
4	Khao Dawk Mali 105	KDML105	ขาวดอกมะลิ105	NG	ITV	Pure line selection from traditional variety	2502
5	Gam Pai	KP41	กำผาย41	--	ITV	Pure line selection from traditional variety	
6	Khao Pahk Maw 148	KPM148	ขาวปากหม้อ148	NG	ITV	Pure line selection from traditional variety	2508
7	Gow Ruang 88	KR88	เก้ารวง88	G	ITV	Pure line selection from traditional variety	2505
8	Muey Nawng 62M	MN62M	เหมยนอง62M	G	ITV	Pure line selection from traditional variety	2502
9	Nahng Mon S-4	NMS-4	นางมล เอส-4	NG	ITV	Pure line selection from traditional variety	2508
10	Nam Sa-gui 19	NSG19	น้ำสะกุย19	NG	ITV	Pure line selection from traditional variety	2511
11	Pra Deu Dang	PDD	ประดู่แดง	--	ITV	Pure line selection from traditional variety	
12	R258	R258	อาร์ 258	G	ITV	Pure line selection from traditional variety	2530
13	RD10	RD10	กข10	G	ITV	Improved by mutation breeding from irradiated RD1	2524
14	RD6	RD6	กข6	NG	ITV	Improved by mutation breeding from irradiated KDML105	
15	Chai Nat1	CNT1	ชัยนาท1	NG	MV	IR13146-158-1/IR15314-43-2-3-3/BKN6995-16-1-1-2	2536
16	Chai Nat2	CNT2	ชัยนาท2	NG	MV	Hawm Phama (GS. No. 3780)/IR11418-19-2-3	2547
17	Chai Nat3	CNT3	ชัยนาท3	NG	MV	---	
18	Chai Nat80	CNT80	ชัยนาท80	NG	MV	Derived from F_1-SPR60/IR29692-99-3-2-1/IR11418-19-2-3	2550
19	Hawm Khlong Luang 1	HKL1	หอมคลองหลวง1	NG	MV	Nahng Mon S-4/IR841-85-1-1-2	2540
20	Hawm Suphan Buri	HSP	หอมสุพรรณ	NG	MV	SPR84177-8-2-2-2-1/SPR85091-13-1-1-4//KDML105	2540
21	IR68144	IR68144	IR68144	NG	MV	---	
22	Phitsanulok 1	PSL1	พิษณุโลก1	NG	MV	KDML105/LA29'72NFU-14-3-1-1//IR58	2541
23	Phitsanulok 2	PSL2	พิษณุโลก2	NG	MV	CNTLR81122-PSL-37-2-1/SPRLR81041-195-2-1//IR56	2543
24	Phitsanulok 3	PSL3	พิษณุโลก3	NG	MV	RD27 / LA29'73-NF1U-14-13-1-1	
25	Pathum Thani 1	PTT1	ปทุมธานี1	NG	MV	BKNA6-18-3-2/PTT85061-86-3-2-1	2543
26	Pathum Thani 2	PTT2	ปทุมธานี2	NG	MV	---	
27	RD1	RD1	กข1	NG	MV	Leuang Tawng/IR8	2512
28	RD23	RD23	กข23	NG	MV	RD7/IR32//RD1	
29	RD29	RD29	กข29	NG	MV	Chai Nat 80	
30	RD4	RD4	กข4	G	MV	LueangTawng/IR8//W1252///RD2	2516
31	RD7	RD7	กข7	NG	MV	C4-63/Gow Ruang 88//Sigadis	2518
32	Sakon Nakhon	SKN	สกลนคร	NG	MV	Hawm Ohm/RD10	2543
33	Suphan Buri 1	SPR1	สุพรรณบุรี1	NG	MV	IR25393-57-2-3/RD23//IR27316-96-3-2-2///SRPLR77205-3-2-1-1/SPRLR79134-51-2-2	2537
34	Suphan Buri 2	SPR2	สุพรรณบุรี2	NG	MV	RD23/IR60	2537
35	Suphan Buri 60	SPR60	สุพรรณบุรี60	NG	MV	Leuang Tawng/C4-63//IR48	2530
36	Suphan Buri 90	SPR90	สุพรรณบุรี90	NG	MV	RD21/IR4422-98-3-6-1//RD11/RD23	2534
37	San-pah-tawng1	SPT1	สันป่าตอง1	G	MV	BKNLR75001-B_3-CNT-B_4-RST-36-2/RD2	2543

Source: Department of Agriculture

*NG=non-glutinous rice, G=Glutinous rice,

**MV=modern variety, ITV=improved traditional variety

Table 2.2.4 Typical characteristic description of common wild rice for life-history trait type classification according to Wongtamee (2008)

Life-history trait	Description
Perennial type (P)	Exist in deep swamp (up to 50 cm. in depth), big and prostrate plant type, big clum with ~ 2 m. in length, broad and long leaf, big and open panicle, seeds mostly sterile
Annual type (A)	Exist in shallow swamp (parched in dry season), small and spreading plant type, small clum with ~50 cm. in length, narrow and short leaf, small and open panicle, fertile seeds

2.2.2 DNA analysis

Microsatellite markers analysis

Total DNA were extracted from dried leaf individually using modified CTAB (Cetyltrimethyl Ammonium Bromide) method (Doyle and Doyle 1987). Sixteen microsatellite loci distributed on 12 rice chromosomes were chosen (Londo *et al*., 2007). Of these, 12 markers (Table 2.2.5) were consistently amplified and scored (Appendix A). Polymerase chain reactions (PCR) were performed in a total volume of 10 µl containing 20 ng of template DNA, 2 pM fluorescence-labeled (6FAM, HEX, NED,VIC; Applied Biosystems) forward primer and 2 pM unlabelled reverse primer, 1 µl *Taq* $MgCl_2$-free 10x buffer, 250 µM deoxyribonucleotides (dNTP), 2.0 mM $MgCl_2$, and 0.05 unit of *Taq* polymerase (Promega). Reaction were performed to denaturing at 94ºC (4 min) followed by 40 cycles of 94ºC denaturing, 55 ºC annealing, 72 ºC extension, each for 30 sec intervals followed by a final 72 ºC extension for 5 min and a 4 ºC hold. The loci RM109 required annealing temperatures of 67ºC (McCouch *et al*., 2002). Amplified products of different size and contrasting fluorescent labels were multiplexed and run on capillary electrophoresis sequencer/genotyper, ABI 3130 Genetic Analyser (Applied Biosystems) using GENESCAN 400HD ROX as an internal size standard.

Table 2.2.5 Microsatellite markers in the present study

Locus	Chromosome number	Primer sequence*	Tm (°C)	No. of Alleles	Expected size (bp)	Reference
RM1	1	F 5'- GCG AAA ACA CAA TGC AAA AA-3' R 5'- GCGTTGGTTGGACCTGAC-3'	55	31	50-150	McCouch *et al*., 1997
RM259	1	F 5'- TGGAGTTTGAGAGGAGGG-3' R 5'- CTTGTTGCATGGTGCCATGT-3'	55	33	155-176	McCouch *et al*., 1997
RM109	2	F 5'- GCCGCCGGAGAGGGAGAGAGAG-3' R 5'- CCCCGACGGGATCTCCATCGTC-3'	67	22	86-101	Temnykh *et al*., 2000
RM211	2	F 5'- CCGATCTCATCAACCAACTG-3' R 5'- CTTCACGAGGATCTCAAAGG-3'	55	16	135-163	McCouch *et al*., 1997
RM251	3	F 5'- GAATGGCAATGGCGCTAG-3' R 5'- ATGCGGTTCAAGATTCGATC-3'	55	37	95-154	McCouch *et al*., 1997
RM133	6	F 5'- TTGGATTGTTTTGCTGGCTCGC-3' R 5'- GGAACACGGGGTCGGAAGCGAC-3'	55	11	229-237	Temnykh *et al*., 2000
RM481	7	F 5'- TAGCTAGCCGATTGAATGGC-3' R 5'- CTCCACCTCCTATGTTGTTG-3'	55	37	133-176	McCouch *et al*., 2000
RM234	7	F 5'- ACAGTATCCAAGGCCCTGG-3' R 5'- CACGTGAGACAAAGACGGAG-3'	55	28	109-164	McCouch *et al*., 2000
RM477	8	F 5'- TCTCGCGGTATAGTTTGTGC-3' R 5'- ACCACTACCAGCAGCCTCTG-3'	55	8	218-227	McCouch *et al*., 2000
RM316	9	F 5'- CTA GTT GGG CAT ACG ATG GC-3' R 5'- ACG CTT ATA TGT TAC GTC AAC-3'	55	37	159-251	Temnykh *et al*., 2000
RM206	11	F 5'- CCCATGCGTTTAACTATTCT-3' R 5'- CGTTCCATCGATCCGTATGG-3'	55	37	118-189	McCouch *et al*., 1997
RM247	12	F 5'- TAGTGCCGATCGATGTAACG-3' R 5'- CATATGGTTTTGACAAAGCG-3'	55	21	119-173	McCouch *et al*., 1997

* F-forward, R-reverse

2.2.3 Data analysis

Genetic diversity analysis

Microsatellite alleles were scored using GENEMAPPER version 3.7 software (Applied Biosystems). Number of allele (A), number of allele per population per locus (A_T), observed heterozygosity (Ho), average gene diversity (H_S), total gene diversity (H_T), degree of gene differentiation among and between populations (F_{ST}) and Wright's inbreeding coefficient (F_{IS}), were calculated using FSTAT version 2.9.3 (Goudet, 2002). In addition, outcrossing rate of each population was estimated using the following equation; $t=(1-F_{IS})/(1+F_{IS})$ (Weir, 1990) where *F* is Wright's inbreeding coefficient. Standard measures of genetic diversity were calculated for the estimate of unbiased Nei's (1973) gene diversity (h) using POPGENE version 1.32 (Yeh *et al*., 1999).

Genetic structure analysis

Genetic structure was analyzed by hierarchical analysis of molecular variance (AMOVA) implemented in the software of GeneAlEx version 6.1 (Peakall and Smouse 2006). In addition, AMOVA yielded statistic analogous to Weir and Cockerham (1984) unbiased F_{ST} estimator, to partition genetic variation into components. Total genetic diversity variance was partitioned into the following components: between and within cultivated rice varieties, weedy rice, and common wild rice populations. The significant of *F*-statistics was tested by permutation, with the probability of non-differentiation for 10000 randomizations.

Population structure analysis

A Bayesian-clustering program utilizing a Markov Chain Monte Carlo (MCMC) approach, STRUCTURE version 2.2 (Pritchard *et al*., 2000), was used to

elucidated the structuring of genetic variation and identify the number of genetically distinct clusters or gene pool of rice populations. STRUCTURE identifies *K* clusters of individuals based on the multilocus genotypes of all individuals in the study and differences in allele frequencies at tested loci (Pritchard *et al*., 2000). STRUCTURE was run five independent times for each value of *K* ranging from 1 to 5 using a burn-in period of 100000 generations and 100000 Markov chain Monte Carlo replications and using a model allowing for admixture and correlated allele frequencies. Parameters were set to their default values, following the documentation of STRUCTURE (Pritchard and Wen 2004). The probability of how the data best fit into each number of assumed clusters (*K*) was estimated by ln probability of the data (ln *L*), so that individuals were assigned to a cluster based on their multilocus genotypic profile. Each test yielded a log-likelihood value of the data (ln *L*), with the highest indicating which test was closest to the actual number of genetically distinct clusters (*K*). Individuals were assigned probabilistically to a population (called inferred populations) or to multiple populations if their genotype profile indicated admixture. In this study, an accession was assigned to a cluster if at least 95% of the given value (*Q*) was estimated to belong to that cluster (*K*). In addition, to trace the origin of widespread weedy rice in various regions proportion of admixture ancestral genotype, different common wild rice genotypes and different cultivated rice genotypes, of each weedy rice population was observed.

Principal coordinate analysis (PCA) was conducted, on the basis of genetic similarity using EIGEN procedure in GeneAlEx 6.1 (Peakall and Smouse 2006) to observed the distribution of the rice populations. The PCA is a method to reduce original total variance among the individuals and to clarify the relationship between

two or more characters into limited number of uncorrelated new variables (Wiley, 1981). This will allow visualization of the differences among the individuals and identify possible groups or clusters (Mohammadi and Prasanna, 2003).

To illustrating genetic relationships among rice populations, Neighbour-joining (NJ) tree was constructed using MEGA version 4 (Tamura *et al*., 2007) based on C.S. chord genetic distance (Cavalli-Sforza and Edwards 1967) obtained by POWERMARKER version 3.0 (Liu and Muse 2005) on the basis of 12 microsatellite loci.

Genetic diversity and population structure of wild, cultivated, and weedy rice in Thailand

Common wild rice, weedy rice collected from the survey and 37 Thai cultivated rice varieties including modern variety (MV) and improved traditional variety (ITV) were evaluated for genetic variation, genetic structure and population structure using 12 microsatellite loci (Table 2.2.5). Genetic diversity parameters were assessed using FSTAT version 2.9.3 (Goudet, 2002), and POPGENE version 1.32 (Yeh *et al*., 1996). Analysis of molecular variance (AMOVA) was used to analyze the hierarchical genetic structure by the apportionment of genetic variation implemented in GeneAlEx version 6.1 (Peakall and Smouse 2006). Population structure was assessed based on three different analysis methods. Bayesian clustering analysis was undertaken using the program STRUCTURE version 2.2 (Pritchard *et al*., 2000). The principal coordinate analysis (PCA) was conducted on the basis of EIGEN method in GeneAlEx version 6.1 (Peakall and Smouse, 2006). In addition, AMOVA was used to estimate statistic analogous to Weir and Cockerham (1984) unbiased F_{ST} estimator, to partition genetic variation into components. Cavalli-Sforza

Chord genetic distance was obtained by POWERMARKER version 3.0 (Liu and Muse, 2005) and subsequently visualized the Neighbor-joining (NJ) tree using MEGA version 4 (Tamura et al., 2007). (The detail of each analysis is desCr1bed in 2.2.3 *Data analysis*).

Tracing of widespread invasive weedy rice in various areas in Thailand

Weedy rice, the weedy form of cultivated rice, was the consequence of gene flow between common wild rice (*O. rufipogon* Griff.) and cultivated rice (*O. sativa* L.) especially in the areas where they are co-existed (Niruntrayakul, 2008). Therefore, 11 common wild rice populations, 12 weedy rice populations and 11 popular cultivated rice varieties (Table 2.2.6) were used to examine the origin of widespread weedy rice in various areas. For cultivated varieties, 11 popular varieties including 7 modern varieties those are grown in irrigated areas particularly in Lower North and Central. Two modern varieties, SPT1 and SKN, and 2 improved traditional varieties, KDML105 and RD6, are grown in rainfed system in Upper North and Northeast.

To identify origin of widespread weedy rice, 12 microsatellite loci were analyzed among 3 rice groups (Tabel 2.2.6): common wild rice (*O. rufipogon* Griff.), cultivated rice (*O. sativa* L.) and weedy form (*O. sativa* f *spontanea*). Genetic relationship among weedy rice and common wild rice and cultivated rice were evaluated based on EIGEN procedure of the principal coordinate analysis (PCA) using GeneAlEx version 6.1 (Peakall and Smouse 2007). Bayesian clustering using Markov Chain Monte Carlo (MCMC) method for population assignment was obtained by STRUCTURE version 2.1 (Pritchard *et al*., 2000).

Table 2.2.6 Popular cultivated rice varieties, weedy rice and common wild rice populations in each rice growing region

Region	Cultivated rice		Weedy rice	Wild rice	Location
	Name	Abbreviation			
Lower North	Chai-nat1	CNT1	LNWE1	LN1	Phitsanulok
	Suphan Buri1	SPR1		LN2	Phitsanulok
	Phitsanulok1	PSL1	LNWE2	LN3	Phichit
	Phitsanulok2	PSL2			
	Phitsanulok3	PSL3			
	Pathum Thani1	PTT1			
	Pathum Thani2	PTT2			
	Khao Dawk Mali105	KDML105			
	RD6	RD6			
Central	Chai-nat1	CNT1	CTWE1	CT1	Kanchanaburi
	Suphan Buri1	SPR1	CTWE2	CT2	
	Phitsanulok1	PSL1	CTWE3	CT3	Suphan Buri
	Phitsanulok2	PSL2	CTWE4		
	Phitsanulok3	PSL3	CTWE5		
	Pathum Thani1	PTT1			
	Pathum Thani2	PTT2			
Northeast	Khao Dawk Mali 105	KDML105	NEWE1	NE1	Yasothon
	RD6	RD6		NE2	
	Sakon Nakhon	SNK	NEWE2	NE3	Surin
	Sanpathong1	SPT1		NE4	
			NEWE3	NE5	Si Saket
				NE6	
			NEWE4	NE7	Buri Rum
			NEWE5	NE8	
				NE9	
				NE10	
				NE11	

2.3 Results

Population genetic structure of common wild rice (*Oryza rufipogon* Griff.)

Forty-two populations were found and collected from 4 geographical regions including, Upper North (5 populations), Lower North (5), Central (9) and Northeast (23) (Table 2.2.1 and Figure 2.3.1). Most of them (45%) were found to scatter in lake or ditches or swamp which inundated all year round. The rest were found inside or near rice field and abandon field (Table 2.3.1). All populations from the north and central regions were perennial type. In contrast, 10 perennial and 13 annual types were found in the northeast region.

Allele diversity

High level of genetic variation was found in common wild rice from all regions (Table 2.3.2). Total number of polymorphic alleles detected with 12 SSR loci was 295 with an average of 24.5 alleles per locus. For each locus, number of alleles ranged from 8 alleles in RM477 to 35 alleles in RM206 with an average of 4.1 alleles per locus.

For each population, total number of alleles were in the range of 21 alleles in NE10 (RoiEt) to 78 alleles in CT2 (Phra Nakhon Si Ayutthaya) with 49.5 alleles in average. For each locus, number of alleles per population were between 1 to 12 alleles per locus with an average for all 12 SSR loci per population ranging from 1.9 alleles in NE10 (RoiEt) to 6.3 alleles in CT2 (Phra Nakhon Si Ayutthaya) (Table 2.3.2),

At the region level, the highest number of allele was observed from Northeast (251 alleles), whereas the lowest was observed in the Lower North (138 alleles) populations. Average number of alleles per region revealed similar level in the Upper

North, Lower North and Central with 57.2, 59.6 and 59.1 alleles, respectively, whereas those from the Northeast had the lowest average number of alleles (41.4 alleles). For each region, number of allele per locus of Northeast was the highest (16.4 alleles) following with Central (13.8 alleles), Upper North (12.4 alleles) and the lowest in Lower North populations (11.5 alleles) (Table 2.3.2).

Wild rice populations with different life-history also exhibited different levels of allele diversity (Table 2.3.3). Total number of alleles of all 12 SSR loci of annual (209 alleles) was lower than that of perennial (253 alleles) types. For each locus, number of alleles of perennial type at RM1, RM109, RM234 and RM316 were much higher than those of annual type. Those of the remaining eight loci were similar.

Table 2.3.1 Distribution of 42 common wild rice populations from 4 regions determined by life-history traits, habitat type, and water conditions

Region	Life-history trait type*		Habitat type				Water condition	
	P	A	Abandoned field	Near rice field	Inside rice field	Lake/Ditch /Swamp	Parched in dry season	Inundate
Upper North (n=5)	5			1	1	3	1	4
Lower North (n=5)	5		1	1		3	1	4
Central (n=9)	9		2	3		4	4	5
Northeast (n=23)	10	13	3	10		10	13	10
Total	**29** (69%)	**13** (31%)	**6** (14%)	**15** (36%)	**1** (2%)	**20** (48%)	**19** (31%)	**23** (69%)

* P-perennial, A-annual

Table 2.3.2 Allele diversity of 42 common wild rice populations collected from Upper North (UP), Lower North (LN), Central (CT) and Northeast (NE) using 12 SSR markers.

Population	Type[†]	Number of alleles (A)												Total no. of allele	Average no. of allele
		RM1	RM206	RM109	RM247	RM211	RM477	RM251	RM234	RM259	RM133	RM316	RM481		
UN1	P	7	5	3	4	3	2	5	4	4	4	8	4	53	4.9
UN2	P	6	4	8	7	3	2	8	4	7	4	5	4	62	4.9
UN3	P	7	3	8	9	4	2	5	6	4	5	2	6	61	4.3
UN4	P	6	2	5	4	4	1	4	7	4	5	5	6	53	4.9
UN5	P	8	6	7	4	4	2	3	4	6	4	5	4	57	4.1
Total UN		18	14	16	13	8	2	16	13	16	8	13	12	149[¶] 57.2[††]	12.4[§]
LN1	P	7	4	7	8	2	3	4	5	6	5	7	2	60	5.1
LN2	P	8	5	6	5	6	2	7	7	6	7	7	5	71	5.9
LN3	P	7	4	7	5	5	1	8	6	5	6	6	7	67	5.5
LN4	P	4	2	2	2	4	1	4	3	3	2	4	4	35	3.1
LN5	P	7	4	7	6	5	2	6	6	6	5	5	6	65	5.4
Total LN		18	9	13	12	8	3	13	11	16	8	15	12	138[¶] 59.6[††]	11.5[§]
CT1	P	4	4	4	3	2	1	4	3	4	2	4	2	37	3.0
CT2	P	7	11	4	5	5	1	5	12	5	8	8	7	78	6.3
CT3	P	8	6	7	3	4	1	8	8	6	3	5	8	67	5.4
CT4	P	10	8	6	5	4	2	6	8	6	4	5	7	71	5.6
CT5	P	4	3	5	6	3	2	4	5	5	3	4	3	47	3.6
CT6	P	4	3	6	8	2	3	5	3	3	4	4	3	48	3.7
CT7	P	3	1	8	5	1	2	3	5	6	2	1	5	42	3.3
CT8	P	5	8	6	9	7	2	4	6	6	3	3	11	70	5.6
CT9	P	4	8	7	6	6	2	9	8	7	4	5	6	72	5.7
Total CT		22	24	16	13	11	4	18	19	17	8	15	19	186[¶] 59.1[††]	13.8[§]
NE1	A	3	6	4	4	1	1	3	3	5	3	6	8	47	3.8
NE2	A	5	4	3	3	4	2	3	4	4	3	4	6	45	3.8
NE3	A	4	4	6	4	2	1	2	5	4	4	6	7	49	2.3
NE4	A	2	4	1	1	4	1	2	2	4	2	2	5	30	4.9
NE5	P	3	2	1	3	2	1	3	3	3	4	2	3	30	5.8
NE6	A	2	2	2	3	2	1	4	1	4	2	2	3	28	3.6
NE7	P	6	8	4	4	2	1	5	6	5	6	6	7	60	3.9
NE8	A	4	5	4	5	3	2	6	9	6	2	7	7	60	1.9
NE9	A	4	2	1	2	3	1	2	6	4	2	4	6	37	2.6
NE10	P	2	1	2	1	1	1	2	2	3	2	1	3	21	3.6
NE11	A	7	5	4	2	3	1	3	4	5	1	4	6	45	4.9
NE12	A	3	2	3	1	2	1	2	2	3	2	1	3	25	3.0
NE13	P	6	6	6	3	5	2	3	9	4	4	5	3	56	2.2
NE14	A	2	2	1	3	2	1	2	2	2	1	3	3	24	4.0
NE15	A	3	2	3	4	3	2	3	3	3	2	4	5	37	2.1
NE16	A	6	5	4	3	3	1	5	5	4	3	3	6	48	3.1
NE17	A	2	4	1	2	2	2	4	4	4	2	1	3	31	2.6
NE18	A	2	2	1	3	1	1	1	2	6	2	2	3	26	2.0
NE19	A	2	1	2	1	3	1	2	2	2	2	2	3	23	2.4
NE20	P	6	6	8	8	5	2	6	9	6	5	6	6	73	4.3
NE21	A	5	4	3	4	3	1	5	7	7	4	3	4	50	4.7
NE22	P	5	4	3	2	2	3	3	3	3	4	8	4	44	5.1
NE23	P	5	4	7	8	3	3	6	9	5	4	6	4	64	2.0
Total NE		25	33	20	18	11	8	24	23	26	7	27	29	251[¶] 41.4[††]	16.2[§]
Average		4.8	4.2	4.5	4.2	3.2	1.6	4.3	5.1	4.6	3.5	4.3	4.9	49.5	4.1[*§]
Range		2-10	1-11	1-8	1-9	1-6	1-3	1-9	1-12	2-7	1-7	1-7	2-11	21-78	
Total		30	35	22	20	14	8	33	26	33	9	34	31	295	24.5[‡]

[†]P-perennial, A-annual
[¶]Number of alleles per region
[††]Average number of alleles per region
[§]Number of alleles per region
[*§]Average number of alleles per population
[‡]Average number of alleles per locus

Table 2.3.3 Allele diversity of common wild rice for 2 life-history trait types

Life-history trait	Number of allele (A)												Total no. of allele
	RM1	RM206	RM109	RM247	RM211	RM477	RM251	RM234	RM259	RM133	RM316	RM481	
Perennial type													
Range	2-10	1-11	1-8	1-9	1-7	1-3	1-9	1-12	1-7	1-7	1-8	1-11	
Average	5.7	4.7	5.5	5.1	3.6	1.8	5.0	5.8	4.9	4.2	4.9	5.0	4.7
Total Perennial (n=29)	28	27	21	16	14	5	26	24	29	9	29	25	253 56.3†
Annual type													
Range	1-7	1-6	1-6	1-4	1-4	1-2	1-6	1-9	1-7	1-4	1-7	1-8	
Average	3.5	3.4	2.7	2.8	2.6	1.3	3.1	3.8	4.2	2.3	3.4	4.9	3.2
Total Annual (n=13)	22	29	14	18	11	5	22	17	22	5	20	24	209 37.8†

†Average number of alleles

Genetic diversity

At the whole population level of 420 individuals, high diversity of common wild rice was detected as displayed by high average gene diversity (H_S=0.568) and total gene diversity (H_T=0.827) (Table 2.3.4). Total observed heterozygosity (H_O) was lower than Nei's (1973) gene diversity expected at HWE (h). Low level of inbreeding coefficient (0.197) was observed together with high level of total outcrossing rate (0.671).

For each population, moderate to high genetic variation was detected in 42 common wild rice populations (Table 2.3.4). Observed heterozygosity (H_O) ranged from 0.117 in NE6 (Buri Rum) to 0.800 in UN4 (Lamphun). Within population genetic diversity (h) were ranging from 0.219 in NE10 (RoiEt) to 0.756 in CT2 (Phra Nakhon Si Ayutthaya). The inbreeding coefficient (F_{IS}) varied widely among populations ranging from -0.431 in CT1 (Ang Thong) to 0.667 in NE6 (Buri Rum). Outcrossing rate (*t*) was the highest in CT1 (2.515) and the lowest in NE6 (0.199).

When partitioning populations into geographical regions (Table 2.3.5), high levels of genetic diversity were found in all 4 regions. Both H_o and gene diversity (h) of Northeast populations were lower than those of the other three regions. Upper North populations exhibited the highest average gene diversity within region (H_S=0.675) and total gene diversity (H_T=0.739), following with Lower North, Central and the lowest was found in Northeast. Genetic differentiation (F_{ST}) between populations within regions were between 0.210 (Lower North) to 0.493 (Northeast). The highest inbreeding coefficient (*F*) was found in Northeast (0.526) following with Central (0.328), Lower North (0.309) and the lowest was Upper North (0.213).

Outcrossing rates were rank in the same order as lowest in Northeast and highest in the Upper North populations.

When populations were partitioned into different life-history types, all genetic parameters, except F_{ST} and F_{IS}, of populations with annual type were lower than those of perennial type (Table 2.3.6 and Figure 2.3.3). Observed heterozygosity (H_O) of annual type populations were distributed within the range of 0.2 and 0.6 (Figure 2.3.3) with average of 0.321 whereas those of perennial type were between 0.4 and 0.8 with average of 0.537 (Table 2.3.6). Average gene diversity (h), H_S and H_T of annual were slightly lower than those of perennial. Populations of annual type were more variable than perennial type revealed by the higher level of genetic differentiation (F_{ST}) among populations (0.527), while F_{ST} of perennial populations was 0.335. For mating parameter, annual type exhibited higher inbreeding coefficient and lower outcrossing rate than perennial type.

Genetic structure

Analysis of Molecular Variance (AMOVA)

Analysis of molecular variance (AMOVA) was used to examine partitioning of total variation into components. Fifty-seven percent of total variance of common wild rice (0.827) was partitioned into among individuals within population, the rest 43% was the variation distributed among populations (Table 2.3.7).

When the genetic variation between populations (43%) was partitioned further according to geographical regions, it was found that only 2% of the 43% was distributed among regions and 41% was among populations within region. About 20 – 49% of total variation found in each region was distributed among populations within region (Table 2.3.7) with Northeast exhibited the highest.

Partitioning of total variation into different life-history traits using AMOVA illustrated that 5% was the variation between annual and perennial types and 38% was among populations within each type. Considering partitioning within each type, about 34% and 53% was distributed among populations within perennial and annual type, respectively (Table 2.3.8).

Pairwise genetic differentiation within wild rice from different regions was about ten times higher than between regions (Table 2.3.9). Similarly, pairwise genetic differentiation within each life-history trait types was also about ten times higher than between types. However, the differentiation found among regions (2%) was due to the differentiation of difference life-history traits of Northeast those mixed between perennial and annual type as displayed in Table 2.3.10.

Table 2.3.4 Genetic parameters of 42 common wild rice populations collected from Upper North (UN), Lower North (LN), Central (CT) and Northeast (NE) based on 12 SSR markers.

Accession	Region	Location	n	Ho	h	H_S	H_T	F_{ST}	F_{IS}	*t*
UN1	Upper North	Chiang Rai	10	0.642	0.682				0.074	0.862
UN2		Chiang Rai	10	0.667	0.698				0.093	0.829
UN3		Chiang Mai	10	0.500	0.632				0.223	0.635
UN4		Lumphun	10	0.800	0.634				-0.223	1.574
UN5		Lumphun	10	0.558	0.687				0.248	0.603
LN1	Lower North	Phitsanulok	10	0.442	0.629				0.332	0.502
LN2		Phitsanulok	10	0.750	0.742				-0.003	1.006
LN3		Phichit	10	0.633	0.644				0.066	0.876
LN4		Sukhothai	10	0.150	0.573				0.650	0.212
LN5		Phetchabun	10	0.583	0.693				0.186	0.686
CT1	Central	Ang Thong	10	0.675	0.479				-0.431	2.515
CT2		Phra Nakhon Si Ayutthaya	10	0.692	0.756				0.074	0.862
CT3		Phra Nakhon Si Ayutthaya	10	0.558	0.688				0.198	0.669
CT4		Nakhon Nayok	10	0.625	0.706				0.094	0.828
CT5		Sing Buri	10	0.625	0.599				-0.072	1.155
CT6		Suphan Buri	10	0.533	0.535				0.139	0.756
CT7		Kanchanaburi	10	0.258	0.421				0.501	0.332
CT8		Kanchanaburi	10	0.325	0.701				0.587	0.260
CT9		Prachin Buri	10	0.525	0.714				0.269	0.576
NE1	Northeast	Sa Kaeo	10	0.292	0.569				0.469	0.361
NE2		Buri Rum	10	0.450	0.673				0.362	0.468
NE3		Buri Rum	10	0.383	0.567				0.315	0.521
NE4		Buri Rum	10	0.275	0.336				0.258	0.589
NE5		Buri Rum	10	0.642	0.457				-0.318	1.933
NE6		Buri Rum	10	0.117	0.332				0.667	0.199
NE7		Nakhon Ratchasima	10	0.533	0.689				0.228	0.628
NE8		RoiEt	10	0.617	0.686				0.082	0.848
NE9		RoiEt	10	0.325	0.439				0.275	0.568
NE10		RoiEt	10	0.275	0.219				-0.113	1.255
NE11		Udon Thani	10	0.433	0.515				0.109	0.803
NE12		Udon Thani	10	0.217	0.272				0.260	0.587
NE13		Surin	10	0.442	0.709				0.340	0.493
NE14		Surin	10	0.142	0.233				0.444	0.385
NE15		Ubon Ratchathani	10	0.408	0.565				0.334	0.499
NE16		Yasothon	10	0.467	0.567				0.139	0.756
NE17		Yasothon	10	0.183	0.254				0.386	0.443
NE18		Maha Sarakham	10	0.192	0.239				0.368	0.462
NE19		Maha Sarakham	10	0.333	0.287				-0.071	1.153
NE20		Sakon Nakhon	10	0.575	0.746				0.232	0.623
NE21		Si Saket	10	0.350	0.638				0.490	0.342
NE22		Si Saket	10	0.450	0.627				0.303	0.535
NE23		Nong Khai	10	0.508	0.696				0.283	0.558
Total Wild rice		**42**	**420**	**0.456**	**0.560**	**0.568**	**0.827**	**0.427**	**0.197**	**0.671**

Number of individuals (n), Observed heterozygosity (Ho), Nei's (1973) gene diversity (h),
Average gene diversity (H_S), Total gene diversity (H_T), Inbreeding coefficient (F_{IS}), and Outcrossing rate (*t*)

Table 2.3.5 Summary of genetic parameters of 42 common wild rice populations based on geographical regions using12 SSR loci

Accession	Region	N	n	Ho	h	H_S	H_T	F_{ST}	F_{IS}	*t*
UN	Upper North	5	50	0.633	0.667	0.675	0.813	0.312	0.213	0.649
LN	Lower North	5	50	0.512	0.656	0.654	0.739	0.210	0.309	0.528
CT	Central	9	90	0.535	0.622	0.635	0.795	0.342	0.328	0.506
NE	Northeast	23	230	0.374	0.492	0.500	0.801	0.493	0.526	0.311
Among regions		4	420	**0.456**	**0.560**	**0.568**	**0.827**	**0.027**	**0.197**	**0.671**

Number of populations (N), Number of individuals (n), Observed heterozygosity (Ho), Nei's (1973) gene diversity (h), Average gene diversity (H_S), Total gene diversity (H_T), Inbreeding coefficient (*F*), and Outcrossing rate (*t*)

Table 2.3.6 Summary of genetic parameters of 42 common wild rice populations based on life history traits using 12 SSR loci

Life-history trait	N	n	Ho	h	H_S	H_T	F_{ST}	F_{IS}	*t*
Perennial type	26	290	0.537	0.809	0.635	0.815	0.335	0.154	0.733
Annual type	16	130	0.321	0.764	0.445	0.783	0.527	0.557	0.285
Among types	2	420	**0.456**	**0.560**	**0.568**	**0.827**	**0.049**	**0.197**	**0.671**

Number of populations (N), Number of individuals (n), Observed heterozygosity (Ho), Nei's (1973) gene diversity (h), Average gene diversity (H_S), Total gene diversity (H_T), Inbreeding coefficient (*F*), and Outcrossing rate (*t*)

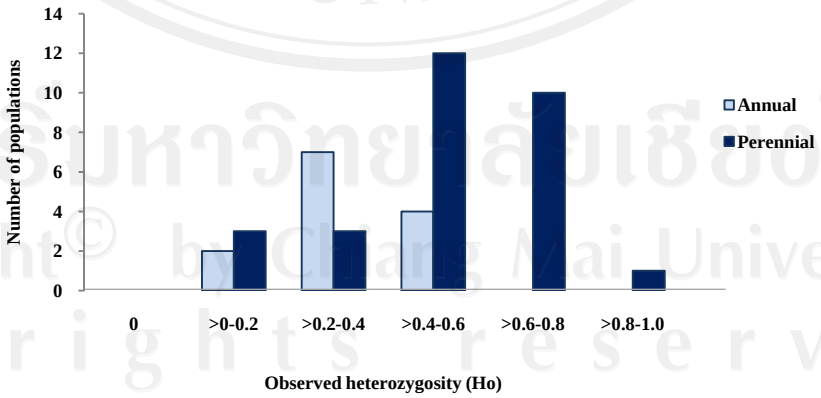

Figure 2.3.3 Distribution of observed heterozygosity (Ho) of annual and perennial type of common wild rice.

Table 2.3.7 Analysis of molecular variance (AMOVA) among populations and regions for 420 individual of 42 common wild rice populations on 12 SSR loci.

Source	d.f.	SS	Variance component	% of the total variance
Among Populations	41	2963.176	6.374	43%
Among Regions	3	304.451	0.360	2%
Populations/Region	38	2658.725	6.144	41%
Populations/Upper North	4	199.040	4.076	31%
Populations/Lower North	4	145.760	2.646	21%
Populations/Central	8	453.456	4.755	34%
Populations/Northeast	22	1860.470	7.669	49%
Individuals/Population	378	3223.800	8.529	57%

Table 2.3.8 Analysis of molecular variance (AMOVA) among populations and life-history types for 420 individual of 42 common wild rice populations on 12 SSR loci.

Source	d.f.	SS	Variance component	% of the total variance
Among Populations	41	2963.176	6.374	43%
Perennial&Annual	1	217.010	0.749	5%
Populations/Type	40	2746.166	6.013	38%
Populations/Perennial	25	1389.104	4.637	34%
Populations/Annual	15	1357.063	8.302	53%
Individuals/Population	378	3223.800	8.529	57%

Table 2.3.9 Pairwise genetic differentiation (F_{ST}) within and between common wild rice collected from 4 regions and 2 different life-history traits

Pairwise F_{ST}	Regions		Pairwise F_{ST}	Life-history trait type	
	Within	Between		Within	Between
Upper North	0.312		*Perennial*	0.335	
Lower North		0.047	Annual		0.049
Central		0.029	*Annual*	0.527	
Northeast		0.035			
Lower North	0.210				
Central		0.045			
Northeast		0.065			
Central	0.342				
Northeast		0.029			
Northeast	0.493				

Table 2.3.10 Pairwise genetic differentiation (F_{ST}) between common wild rice collected from Upper North (UN), Lower North (LN), Central (CT) and Northeast (NE, P-perennial, A- annual)

Pairwise F_{ST}	NE		
	P+A	**P**	**A**
UN	0.035	0.035	0.070
LN	0.065	0.065	0.093
CT	0.029	0.029	0.056

Population structure

Forty-two populations of common wild rice were structured into 2 inferred populations (*K*=2) (Figure 2.3.4) on the basis of 12 microsatellite loci using STRUCTURE program. In term of genome proportion (*Q*), at least 95% of the genotypes of individual population were estimated belong to their populations of origin (Figure 2.3.4 and Table 2.3.11). The first inferred cluster consisted of 26 populations and belong to the perennial type (W1) (represented by green color in Figure 2.3.4 and Table 2.3.11) including 5 populations each from the Upper and Lower North, 9 from Central and 7 from Northeast. The second inferred cluster consisted of 16 annual populations from the Northeast (W2) (represented by red color in Figure 2.3.4 and Table 2.3.11). However, there were 10 populations shared genetic constitutions between the W1 and W2 groups revealed by *Q* value less than 95%. This group consisted of one population each of the Upper and Lower North, two from Central and six from Northeast.

The assignment of perennial and annual types into different group (W1 and W2) by STRUCTURE was also observed by the PCA analysis (Figure 2.3.5). The EIGEN analysis of the pairwise Chord genetic distance measured among 42 populations explained 40.96% of the variation presented in the two life-history groups within the first and second axes. The two common wild rice groups were not completely differentiate from each other, however, there were the overlap populations fell between perennial and annual populations revealed in Figure 2.3.5. Perennial populations represented by green color in the PCA graph were clustered in the right of the graph while annual populations represented by red color were clustered in the left. Genetic relationship among common wild rice revealed by cluster analysis using NJ

method on the basis of C.S. Chord (1967) genetic distance also clearly separated the common wild rice populations into two major clusters, perennial and annual types (Figure 2.3.6).

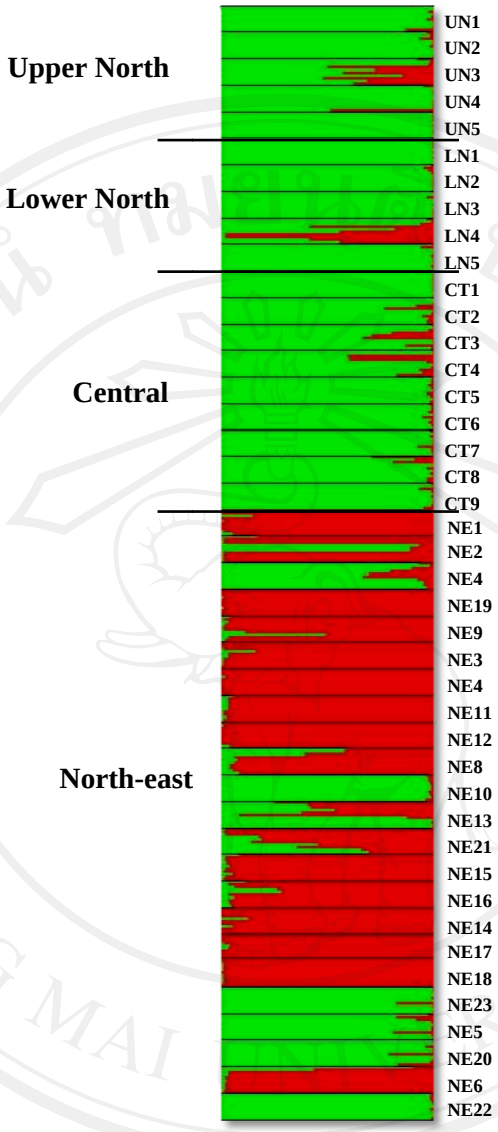

Figure 2.3.4 Population assignment of 42 common wild rice populations in natural habitat collected from 4 regions of Thailand reveal K=2. Each bar represent each common wild rice population consisted of 10 individuals. Different colors represent different inferred populations; Red-Annual type**;** Green-Perennial type

Table 2.3.11 Proportion of estimated assigned populations (*Q*) of 42 common wild rice populations (*K*=2)

No.	Population	Province	n	Given inferred clusters (Q)		Group*	Individual plants with admixture
				Perennial (W1)	Annual (W2)		
1	UN1	Chiang Rai	10	0.979	0.021	P	1
2	UN2	Chiang Rai	10	0.727	0.273	P	1
3	UN3	Chiang Mai	10	0.971	0.029	P	8
4	UN4	Lumphun	10	0.938	0.062	P	1
5	UN5	Lumphun	10	0.992	0.008	P	0
6	LN1	Phitsanulok	10	0.992	0.008	P	0
7	LN2	Phitsanulok	10	0.982	0.018	P	1
8	LN3	Phichit	10	0.989	0.011	P	0
9	LN4	Phetchabun	10	0.564	0.436	P/A	8
10	LN5	Sukhothai	10	0.983	0.017	P	1
11	CT1	Ang Thong	10	0.991	0.009	P	0
12	CT2	Phra Nakhon Si Ayutthaya	10	0.949	0.051	P	3
13	CT3	Phra Nakhon Si Ayutthaya	10	0.893	0.107	P	4
14	CT4	Nakhon Nayok	10	0.878	0.122	P	4
15	CT5	Sing Buri	10	0.979	0.021	P	0
16	CT6	Suphan Buri	10	0.976	0.024	P	1
17	CT7	Kanchanaburi	10	0.979	0.021	P	1
18	CT8	Kanchanaburi	10	0.923	0.077	P	3
19	CT9	Prachin Buri	10	0.975	0.025	P	2
20	NE1	Sa Kaeo	10	0.026	0.974	A	2
21	NE2	Buri Rum	10	0.293	0.707	A	4
22	NE3	Buri Rum	10	0.028	0.972	A	1
23	NE4	Buri Rum	10	0.007	0.993	A	0
24	NE5	Buri Rum	10	0.941	0.059	P	3
25	NE6	Buri Rum	10	0.151	0.849	A	2
26	NE7	Nakhon Ratchasima	10	0.869	0.131	P	8
27	NE8	Roi Et	10	0.196	0.804	A	8
28	NE9	Roi Et	10	0.076	0.924	A	2
29	NE10	Roi Et	10	0.971	0.029	P	0
30	NE11	Udon Thani	10	0.027	0.973	A	1
31	NE12	Udon Thani	10	0.010	0.990	A	0
32	NE13	Surin	10	0.588	0.412	P/A	10
33	NE14	Surin	10	0.023	0.977	A	2
34	NE15	Ubon Ratchathani	10	0.030	0.970	A	2
35	NE16	Yasothon	10	0.096	0.904	A	4
36	NE17	Yasothon	10	0.012	0.988	A	0
37	NE18	Maha Sarakham	10	0.008	0.992	A	0
38	NE19	Maha Sarakham	10	0.006	0.994	A	0
39	NE20	Sakon Nakhon	10	0.948	0.052	P	2
40	NE21	Si Saket	10	0.269	0.731	A	8
41	NE22	Si Saket	10	0.984	0.016	P	0
42	NE23	Nong Khai	10	0.974	0.026	P	1

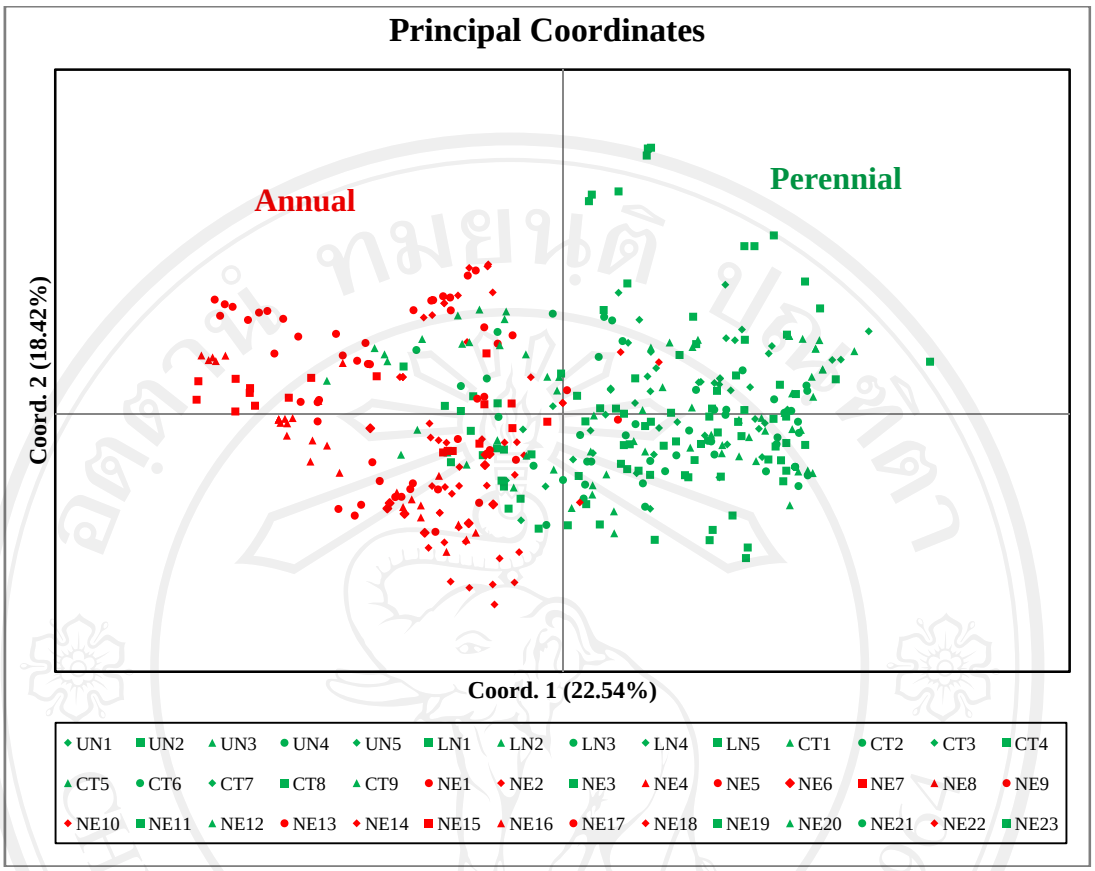

Figure 2.3.5 Distribution of 420 individuals of 42 wild rice populations. Different color represent 2 clusters referred to the assignment obtained from STRUCTURE; red - annual type and green – perennial type, formed by the principal coordinate analysis (PCA) on the basis of 12 microsatellite markers.

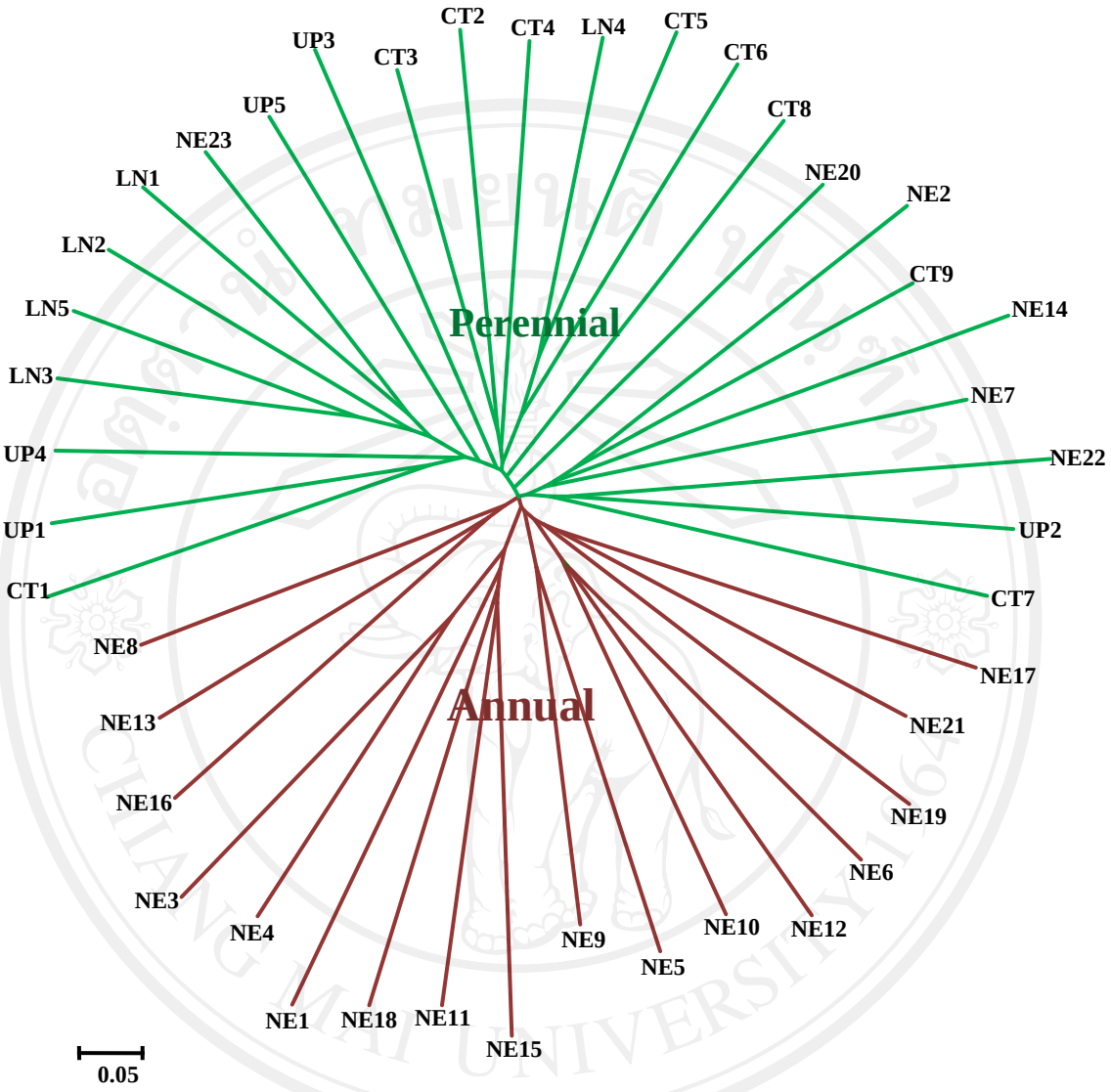

Figure 2.3.6 Cluster analysis using NJ method based on C.S. Chord (1967) genetic distance among 42 common wild rice populations in Thailand

Population genetic structure of cultivated rice (*Oryza sativa* L.)

Allele diversity

No allele diversity was detected within pure line cultivated rice. Total number of polymorphic alleles observed among 37 cultivated rice varieties in 12 SSR loci was 117 alleles with average of 9.7 alleles per locus. Number of alleles for each locus ranged from 1 allele in RM477 to 16 alleles in RM206. Each variety was fixed at one allele per locus per variety. High level of polymorphic information content (PIC) was found in all 12 SSR loci ranging from 0.669 in RM259 to 0.912 in RM206 with 0.737 in average. With the exception of RM477, monomorphic (PIC=0) was found for all 37 cultivated varieties (Table 2.3.12).

Thirty-seven cultivated rice varieties were classified into 2 groups, modern variety (MV) and improved traditional variety (ITV), according to their background pedigree (Table 2.2.3). Different rice groups illustrated different degree of allele variation found in each SSR locus. ITV had more number of private alleles (5 alleles) than MV (1 allele). Number of allele of each locus of MV were mostly lower than ITV with the exception of RM251, RM259 and RM481 which slightly higher than ITV. Both rice groups showed similar total number of alleles (83 alleles) and PIC (0.69), however, average number of alleles per locus per group was higher in ITV (6.1 alleles) than MV (5.6 alleles) (Table 2.3.12).

Table 2.3.12 Allelic richness of 37 Thai cultivated rice varieties

Locus	Average no. of allele	Total (n=37)		Modern varieties (n=23)		Improved traditional varieties (n=14)	
		Allele no.	PIC*	Allele no.	PIC*	Allele no.	PIC*
RM1	1	13	0.883	8	0.842	10	0.866
RM206	1	16	0.912	10	0.869	11	0.889
RM109	1	8	0.764	6	0.735	7	0.759
RM247	1	9	0.746	6	0.642	7	0.792
RM211	1	12	0.828	7	0.776	8	0.786
RM477	1	1	0	1	0	1	0
RM251	1	12	0.861	11	0.835	6	0.701
RM234	1	11	0.820	8	0.768	8	0.817
RM259	1	7	0.669	7	0.605	5	0.587
RM133	1	6	0.694	4	0.664	4	0.541
RM316	1	8	0.784	5	0.737	7	0.792
RM481	1	14	0.887	10	0.843	9	0.853
No. of private allele		**6**		**1**		**5**	
Average	**1**	**9.7**	**0.737**	**5.6§**	**0.693**	**6.1§**	**0.699**
Total		**117**		**83**		**83**	

*Polymorphic information content (PIC)
§Average number of allele per locus per group

Genetic diversity

High level of genetic diversity was detected between 37 cultivated varieties determined by total gene diversity (0.800). Comparison of genetic diversity in modern varieties (MV) to that of improved traditional varieties (ITV) showed that the ITV had similar level of total gene diversity within group (0.778) with MV group (0.751) (Table 2.3.13).

Genetic differentiation (F_{ST}) within MV and ITV showed similar levels with total genetic differentiation which was the consequence of the fixation of each homozygous genotype at each locus found in each variety (F_{ST}=1). However, low but significant level of genetic differentiation was detected between MV and ITV (0.094^{**}). Similar results of partitioning of genetic variation among 37 cultivated rice varieties observed using AMOVA analysis (Table 2.3.14) and was consistent with the F_{ST} revealed that most genetic variation was apportioned into among varieties (91%) than among rice groups (9%).

Table 2.3.13 Genetic parameters of 37 Thai cultivated rice varieties based on 12 SSR markers

Parameters	Total	Modern varieties (MV)	Improved traditional varieties (ITV)
No. of varieties	37	23	14
Observed heterozygosity (Ho)	0	0	0
Average gene diversity (H_S)	0	0	0
Total gene diversity (H_T)	0.800	0.751	0.778
Genetic differentiation (F_{ST})	1.00**	1.00**	1.00**
F_{ST} between MV & ITV	-	0.094**	-
Inbreeding coefficient (F)	1	1	1
Outcrossing rate (*t*)	0	0	0
No. samples/variety	10	10	10

** P<0.01

Table 2.3.14 Analysis of molecular variance (AMOVA) among varieties, between and within rice groups for 370 individuals of 37 cultivated varieties on 12 SSR loci

Source	d.f.	SS	Variance component	% of the total variance
Among Varieties	36	6724.746	186.798	100%
Among Groups	1	330.861	0.851	9%
Populations/Group	35	6393.884	18.268	91%
Individuals/Population	333	0.900	0.003	0%

Population structure

Cultivated rice varieties in the present study were separated into 3 inferred populations (*K*=3) on the basis of 12 microsatellite loci using STRUCTURE program (Figure 2.3.7). The first inferred population (Cr1) consisted of 12 modern varieties. The second inferred population (Cr2) had 10 varieties, most of which were regarded to improved traditional varieties addition with 3 modern varieties; RD1, RD4 and SPT1 due to their background pedigree shared the same traditional variety parent, Leuang Tawng. The remaining 15 varieties, which are either improved traditional varieties or modern varieties (Cr3) (Table 2.3.15). More than 95% of genome proportion of each cultivated variety was assigned into its inferred population except for CNT1, where 41% was fell into the Cr1 population (MV and ITV) and the rest 59% was fall into Cr3 population (MV).

In addition, principal coordinate analysis (PCA) showed that 37 cultivated varieties were widely distributed across the two axes. NJ clustering method based on C.S. Chord genetic distance also divided these 37 cultivated rice varieties into 3 major groups. The first group (Cr1) mostly consisted of MV. The second group (Cr2) mostly consisted of ITV and the last group (Cr3) were consisted of MV and ITV (Figure 2.3.9).

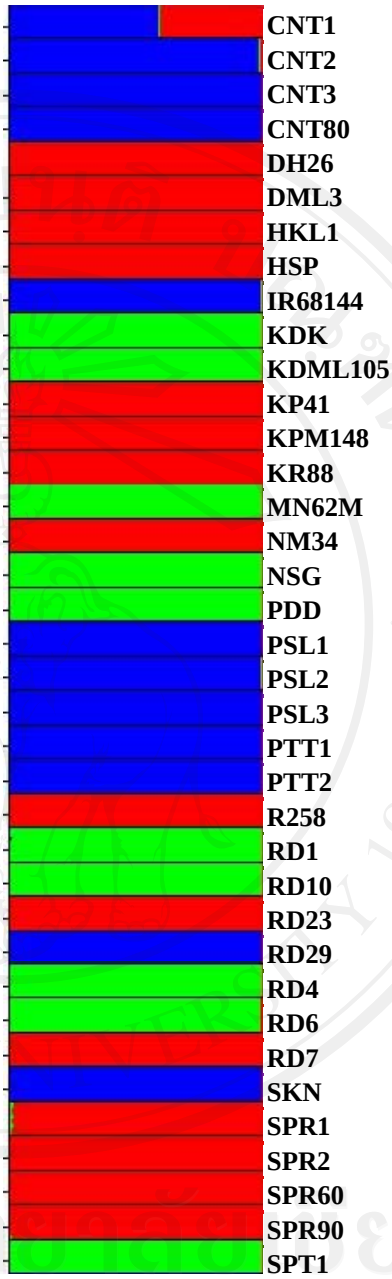

Figure 2.3.7 Population assignment of 37 cultivated rice varieties reveal *K*=3. Each bar represent each population consisted of 10 individuals. Different colors represent different inferred populations, referred to different *K*.

Table 2.3.15 Proportion of estimate assigned populations (*Q*) of 37 cultivated rice populations (K=3)

Variety abbreviation	Group	Given inferred clusters (*Q*)			Description
		Blue (Cr1)	Green (Cr2)	Red (Cr3)	
CNT1	MV	0.589	0.005	0.406	(IR13146-158-1/IR15314-43-2-3-3/BKN6995-16-1-1-2)
CNT2	MV	0.981	0.011	0.008	Hawm Phama (GS. No. 3780)/IR11418-19-2-3
CNT3	MV	0.997	0.001	0.002	
CNT80	MV	0.997	0.001	0.002	F_1-SPR60/IR29692-99-3-2-1/IR11418-19-2-3
PSL1	MV	0.997	0.002	0.001	KDML105/LA29'72NFU-14-3-1-1//IR58
PSL2	MV	0.986	0.01	0.004	CNTLR81122-PSL-37-2-1/SPRLR81041-195-2-1//IR56
PSL3	MV	0.997	0.002	0.001	RD27 / LA29'73-NF1U-14-13-1-1
PTT1	MV	0.995	0.002	0.004	BKNA6-18-3-2/PTT85061-86-3-2-1
PTT2	MV	0.996	0.002	0.001	
RD29	MV	0.997	0.001	0.001	F_1-SPR60/IR29692-99-3-2-1/IR11418-19-2-3
IR68144	MV	0.992	0.006	0.002	
SKN	MV	0.994	0.005	0.002	Hawm Ohm/RD10
KDK	ITV	0.004	0.994	0.002	Pure line selection from traditional variety
MN62M	ITV	0.002	0.996	0.002	Pure line selection from traditional variety
NSG	ITV	0.001	0.996	0.002	Pure line selection from traditional variety
PDD	ITV	0.002	0.996	0.002	Pure line selection from traditional variety
KDML105	ITV	0.007	0.991	0.002	Pure line selection from traditional variety
RD6	ITV	0.004	0.984	0.011	Improved by mutation breeding from KDML105
RD1	MV	0.007	0.991	0.002	Leuang Tawng/IR8
RD10	MV	0.005	0.989	0.006	Improved by mutation breeding from RD1
RD4	MV	0.003	0.993	0.005	Lueang Tawng/IR8//W1252///RD2
SPT1	MV	0.004	0.993	0.002	BKNLR75001-B_3-CNT-B_4-RST-36-2/RD2
DH26	ITV	0.002	0.001	0.997	Pure line selection from traditional variety
DML3	ITV	0.002	0.001	0.997	Pure line selection from traditional variety
KP41	ITV	0.001	0.002	0.997	Pure line selection from traditional variety
KPM148	ITV	0.002	0.002	0.996	Pure line selection from traditional variety
KR88	ITV	0.003	0.002	0.995	Pure line selection from traditional variety
NMS-4	ITV	0.002	0.002	0.996	Pure line selection from traditional variety
R258	ITV	0.002	0.002	0.996	Pure line selection from traditional variety
RD7	MV	0.003	0.003	0.994	C4-63/Gow Ruang 88//Sigadis
RD23	MV	0.001	0.002	0.997	RD7/IR32//RD1
HKL1	MV	0.002	0.002	0.996	Nahng Mon S-4/IR841-85-1-1-2
HSP	MV	0.002	0.002	0.997	SPR84177-8-2-2-2-1/SPR85091-13-1-1-4//KDML105
SPR1	MV	0.004	0.011	0.986	IR25393-57-2-3/RD23//IR27316-96-3-2-2///SRPLR77205-3-2-1-1/SPRLR79134-51-2-2
SPR2	MV	0.002	0.003	0.995	RD23/IR60
SPR60	MV	0.001	0.002	0.997	Leuang Tawng/C4-63//IR48
SPR90	MV	0.005	0.003	0.993	RD21/IR4422-98-3-6-1//RD11/RD23

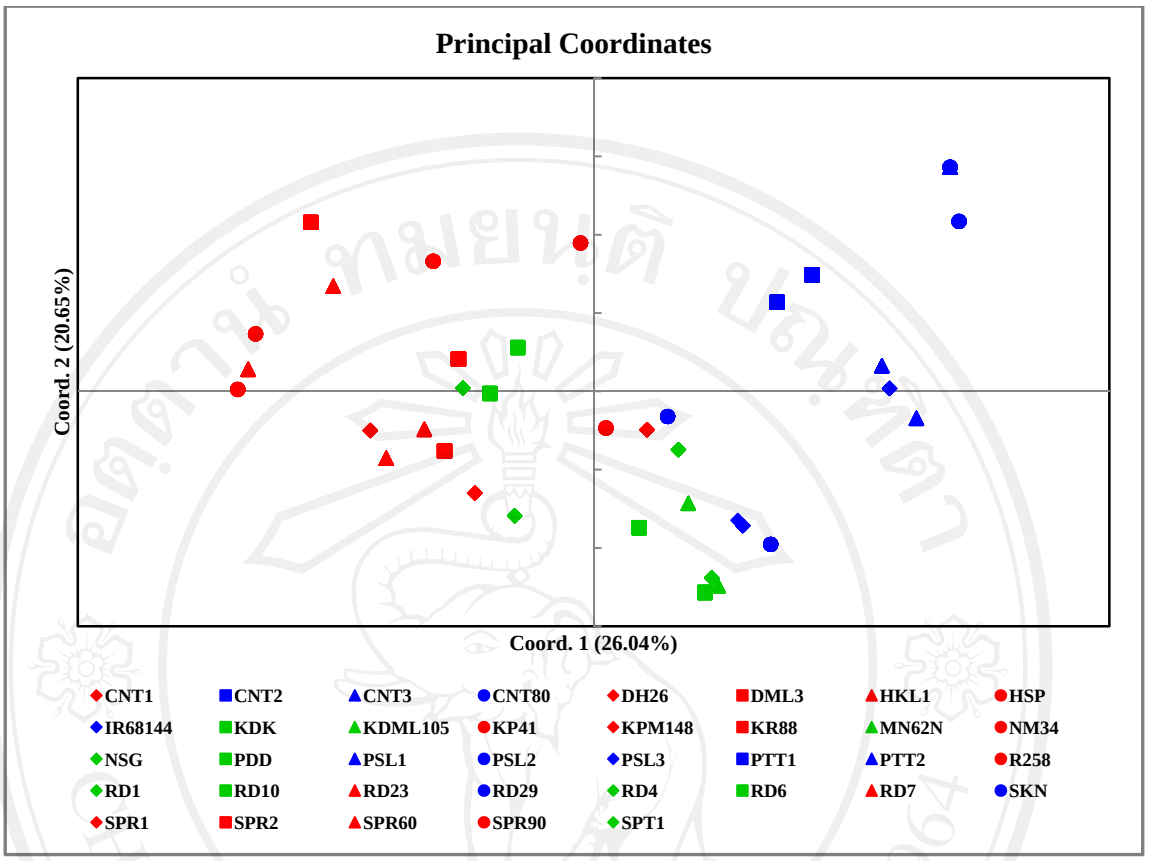

Figure 2.3.8 Distribution of 37 cultivated Thai rice varieties. Different colors represent 3 clusters referred to the assignment obtained from STRUCTURE; Blue - Cr1 gene pool (modern varieties), Green - Cr2 gene pool (improved traditional varieties), Red - Cr3 gene pool (improved traditional varieties and modern varieties) formed by the principal coordinate analysis (PCA) on the basis of 12 microsatellite markers.

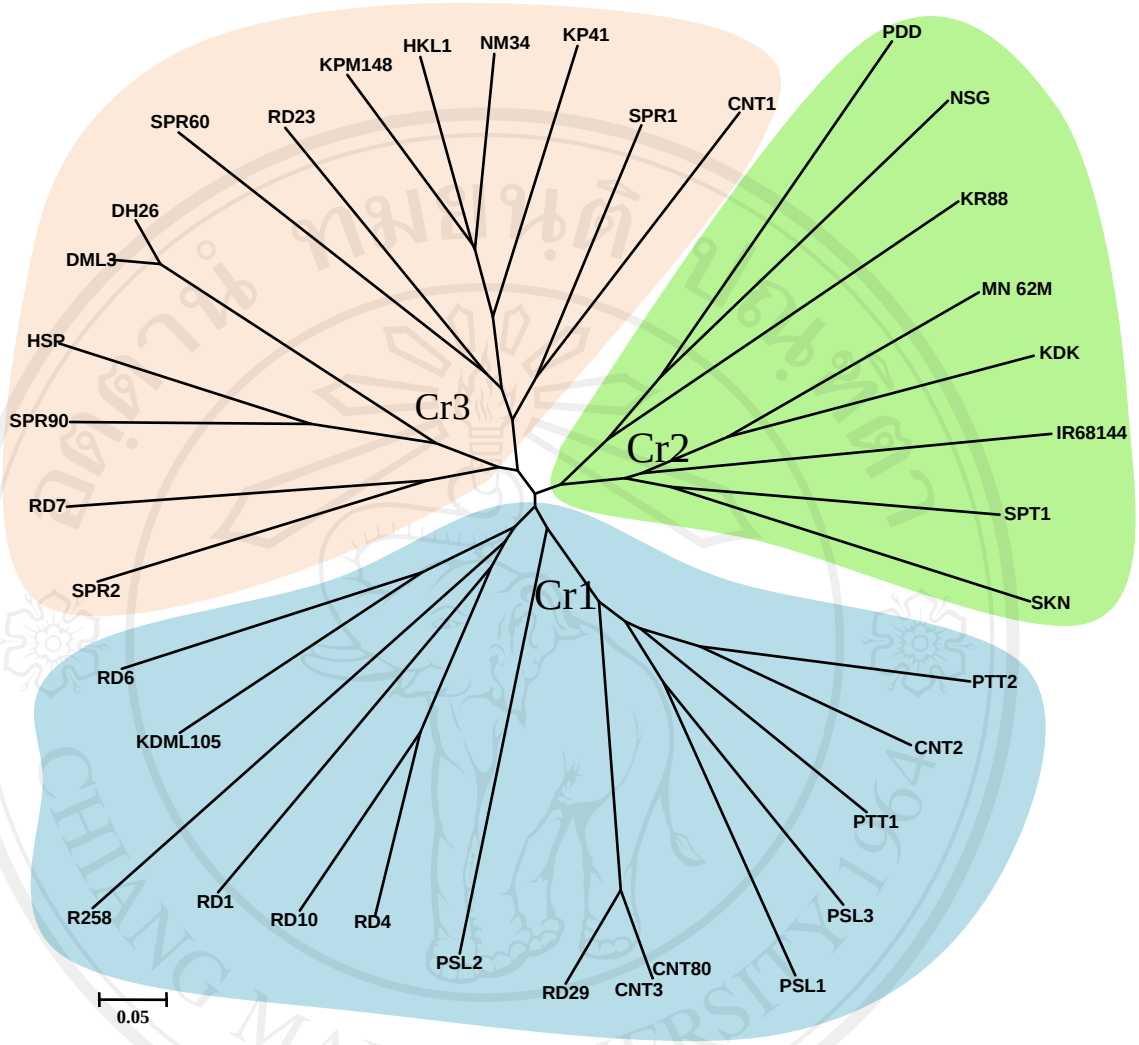

Figure 2.3.9 Cluster analysis using NJ method based on C.S. Chord (1967) genetic distance among 37 cultivated rice varieties derived from 12 SSR loci.

Population genetic structure of weedy rice (*Oryza sativa* f *spontanea*)

Allele diversity

Various levels of allele diversities of 12 weedy rice populations were presented in Table 2.3.16. Total number of 140 alleles of weedy rice was observed with 12 SSR loci with the average of 11.7 alleles. Number of alleles per locus ranged from 2 alleles in RM477 to 23 alleles in RM481. For 12 loci, average number of allele per population was 3.1 alleles.

At individual population level, total number of alleles were ranging from 21 alleles in CTWE 4 and CTWE5 (Suphan Buri) to 57 alleles in LNWE1 (Phitsanulok) and NEWE3 (Buri Rum) with 37 alleles in average (Table 2.3.16). Number of alleles per locus were ranging from 1 to 8 alleles per population with an average ranging from 1.7 allele per locus in CTWE 4 and CTWE5 (Suphan Buri) to 4.7 alleles per locus in LNWE1 (Phitsanulok) and NEWE3 (Buri Rum).

Among the regions, the highest number of alleles was observed in the Northeast totally 90 alleles, followed with 70 alleles in the Central while the Lower North was the lowest with 67 alleles. Average number of alleles per region was 43 alleles in Lower North, 39.6 in Northeast and the lowest in Central, 32 alleles. Average number of alleles per region per locus was the highest in Northeast (7.5 alleles), while Lower North and Central had similar values about 5.8 and 5.6 alleles, respectively.

Table 2.3.16 Allele diversity of 12 weedy rice populations collected from Lower North (LN), Central (CT) and Northeast (NE) using 12 SSR markers

Population	Number of allele (A)												Total no. of allele	Average no. of allele
	RM1	RM206	RM109	RM247	RM211	RM477	RM251	RM234	RM259	RM133	RM316	RM481		
LNWE1	8	8	3	6	4	1	2	7	4	2	5	7	57	4.8
LNWE1	3	3	1	3	1	1	2	4	2	3	2	4	29	2.4
Total LN	10	9	3	7	4	1	4	8	5	4	6	9	70[¶] 43.0[‡]	5.8[§]
CTWE1	5	2	2	2	1	1	3	3	3	2	5	7	36	3.0
CTWE2	2	4	2	2	2	1	3	3	3	2	5	8	37	3.1
CTWE3	5	3	2	3	4	1	4	5	4	3	3	8	45	3.7
CTWE4	1	1	1	1	1	1	3	3	2	1	1	5	21	1.7
CTWE5	2	1	2	1	2	1	2	2	1	1	2	4	21	1.7
Total CT	7	6	3	4	4	1	8	5	5	4	6	14	67[¶] 32.0[‡]	5.6[§]
NEWE1	2	3	2	2	2	1	2	2	3	2	3	5	29	2.4
NEWE2	4	3	3	2	2	1	2	4	4	3	3	7	38	3.2
NEWE3	6	6	3	5	3	2	4	7	6	2	5	8	57	4.8
NEWE4	5	4	2	4	4	2	4	6	5	2	2	5	46	3.8
NEWE5	2	2	2	2	2	1	4	4	3	1	1	4	28	2.3
Total NE	12	10	4	7	5	2	9	11	9	3	6	12	90[¶] 39.6[‡]	7.5[§]
Average	8.3	8.5	4.0	6.6	4.4	1.4	7.5	8.2	6.3	3.5	5.2	8.7	37	3.1[*§]
Range	1-8	1-8	1-3	1-6	1-4	1-2	2-4	2-7	1-6	1-3	1-5	4-7	21-57	
Total	20	15	5	10	7	2	16	14	10	4	14	23	140	11.7[§§]

[¶] Number of alleles per region
[‡]Average number of alleles per region
[§] Number of alleles per region
[§§]Average number of alleles per locus
[*§]Average number of alleles per population

Genetic diversity

Weedy rice exhibited high level of genetic diversity demonstrated by average gene diversity (0.457) and total gene diversity (0.736) (Table 2.3.17). Total observed heterozygosity (0.238) was two times lower than Nei's (1973) gene diversity (0.457) indicated the deviation from Hardy-Weinberg equilibrium. The results of F_{ST} illustrated that more than half (52%) of genetic variation of those found in weedy rice was the differences between 12 weedy rice populations (Table 2.3.17). Moderate level of total inbreeding coefficient (0.601) was detected, lead to the moderate to high level of total outcrossing rate (0.249).

For each population, twelve weedy rice populations exhibited broad level of genetic diversity (Table 2.3.17). Population of NEWE3 (Si Saket) showed the highest observed heterozygosity (0.567) whereas CTWE5 (Suphan Buri) had the lowest (0.083) with 0.244 in average. Similarly, the highest Nei's (1973) gene diversity (h) was also found in NEWE3 (0.706) and the lowest was in CTWE5 (0.131) with 0.457 in average. Weedy rice displayed various levels of inbreeding coefficient (F_{IS}). The highest level was found in CTWE2 (Kanchanaburi) and the lowest level was NEWE4 (Buri Rum). On the other hand, population of NEWE4 (Buri Rum) showed the highest level of outcrossing rate (0.671) while CTWE2 (Kanchanaburi) showed the lowest (0.109).

For each region (Table 2.3.18), weedy rice collected from all 3 regions showed lower level of observed heterozygosity (H_O) than Nei's (1973) gene diversity (h). Weedy rice from the Northeast exhibited the highest genetic variation including observed heterozygosity (0.337), Nei's gene diversity (0.666) and average gene diversity (0.539) while Central was the lowest. All weedy rice from 3 regions

exhibited similar total gene diversity (H_T), 0.695 in Northeast, 0.693 in Lower North and 0.600 in Central. The highest level of inbreeding coefficient (F_{IS}) was found in Central (0.761) following with Lower North (0.569) and the lowest was Northeast (0.495). Conversely, the Northeast weedy rice exhibited the highest level of outcrossing rate (0.338) following with Lower North (0.275) and the lowest in Central (0.275).

Genetic structure

Analysis of Molecular Variance (AMOVA)

Analysis of molecular variance (AMOVA) was used to examine the partitioning of total variation in weedy rice (Table 2.3.19). About 52% of total variance (0.736, Table 2.3.17) was distributed among 12 populations of weedy rice while the rest 48% was distributed between 120 individuals. Partitioning among regions showed that 11% of the total genetic variation was distributed among regions. For within each region, about 39%, 43% and 57% was the distribution of genetic variation among populations within Lower North, Central and Northeast, respectively (Table 2.3.19).

In term of genetic differentiation (F_{ST}), weedy rice exhibited higher level of F_{ST} within region than between regions. Weedy rice populations of Northeast illustrated the highest level of genetic differentiation within populations while Lower North was the lowest. Considering pairwise genetic differentiation between regions, weedy rice of Lower North to Northeast (0.122) were more similar than those weed rice of Northeast to Central (0.189) and Lower North to Central (0.222) (Table 2.3.20).

Table 2.3.17 Genetic parameters of 12 weedy rice populations collected from Lower North (LN), Central (CT) and Northeast (NE) based on 12 SSR markers.

Population	Location	n	Ho	h	H_S	H_T	F_{ST}	F_{IS}	*t*
LNWE1	Phitsanulok	10	0.417	0.584				0.287	0.554
LNWE2	Phichit	10	0.092	0.362				0.747	0.145
CTWE1	Kanchanaburi	10	0.142	0.463				0.694	0.181
CTWE2	Kanchanaburi	10	0.100	0.508				0.803	0.109
CTWE3	Suphan Buri	10	0.233	0.595				0.608	0.244
CTWE4	Suphan Buri	10	0.108	0.150				0.276	0.567
CTWE5	Suphan Buri	10	0.083	0.131				0.364	0.466
NEWE1	Yasothon	10	0.308	0.482				0.342	0.490
NEWE2	Surin	10	0.125	0.590				0.675	0.194
NEWE3	Si Saket	10	0.567	0.706				0.457	0.373
NEWE4	Buri Rum	10	0.550	0.561				0.197	0.671
NEWE5	Buri Rum	10	0.208	0.356				0.415	0.413
Total	**12 populations**	**120**	**0.238**	**0.457**	**0.457**	**0.736**	**0.524**	**0.601**	**0.249**

Number of individuals (n), Observed heterozygosity (Ho), Nei's (1973) gene diversity (h), Inbreeding coefficient (F_{IS}), and Outcrossing rate (*t*)

Table 2.3.18 Summary of genetic parameters of weedy rice collected from 3 geographical regions based on 12 SSR loci

Accession	Region	N	n	Ho	h	H_S	H_T	F_{ST}	F_{IS}	*t*
LN	Lower North	2	20	0.254	0.589	0.473	0.693	0.389	0.569	0.275
CT	Central	5	50	0.133	0.557	0.369	0.600	0.433	0.761	0.136
NE	Northeast	5	50	0.337	0.666	0.539	0.695	0.568	0.495	0.338
Among region	**3**	**12**	**120**	**0.238**	**0.457**	**0.457**	**0.736**	**0.110**	**0.601**	**0.249**

Number of populations (N), Number of individuals (n), Observed heterozygosity (Ho), Nei's (1973) gene diversity (h), Average gene diversity (H_S), Total gene diversity (H_T), Inbreeding coefficient (F_{IS}), and Outcrossing rate (*t*)

Table 2.3.19 Analysis of molecular variance (AMOVA) among populations and 3 regions for 120 individuals of 12 weedy rice populations on 12 SSR loci

Source	d.f.	SS	Variance component	% of the total variance
Among Populations	11	898.80	7.492	52%
Among Regions	2	263.130	1.625	11%
Populations/region	9	635.670	6.384	43%
Populations/Lower North	1	61.15	5.284	39%
Populations/Central	4	250.62	5.539	43%
Populations/Northeast	4	323.90	7.526	57%
Individuals/Population	108	733.70	6.794	48%

Table 2.3.20 Pairwise of genetic differentiation (F_{ST}) within and between 3 regions of weedy rice

Pairwise F_{ST}	Regions	
	Within	Between
Lower North	0.389	
Central		0.222
Northeast		0.122
Central	0.433	
Northeast		0.189
Northeast	0.568	

Population structure

Twelve weedy rice populations were structured into 2 inferred populations (K=2) on the basis of 12 microsatellite loci using STRUCTURE program (Figure 2.3.9). The first inferred population consisted of weedy rice populations of LNWE2, CTWE1, CTWE 2, CTWE4, CTWE5 and NEWE4 which were collected from modern variety cultivation fields (Table 2.2.6). Whereas the second inferred population consisted of weedy rice populations of LNWE1, CTWE3, NEWE1, NEWE2, NEWE3, and NEWE5 which were collected from the fields with improved traditional varieties particularly KDML105 or RD6 (Table 2.2.6). In term of genome proportions (*Q*), the admixtures of 2 inferred populations, were found in 4 weedy populations (CTWE1, CTWE2, CTWE3andNEWE2) (Table 2.3.21).

Principal coordinate analysis (PCA) clustered the 12 weedy rice populations into 2 groups. The EIGEN analysis of the pairwise Chord genetic distance measures for the among 12 common wild rice populations explained 53.88% of the variation presented in the two weedy rice groups within the first and second axes (Figure 2.3.11). NJ clustering method base on C.S. Chord genetic distance also divided 12 weedy rice populations into 2 major clusters (Figure 2.3.12). The first cluster consisted of weedy rice collected from modern varieties fields particularly CNT1 or SPR1 variety (Table 2.2.6). The second cluster consisted of weedy rice collected from improved traditional varieties cultivation fields especially KDML105 or RD6 variety (Table 2.2.6).

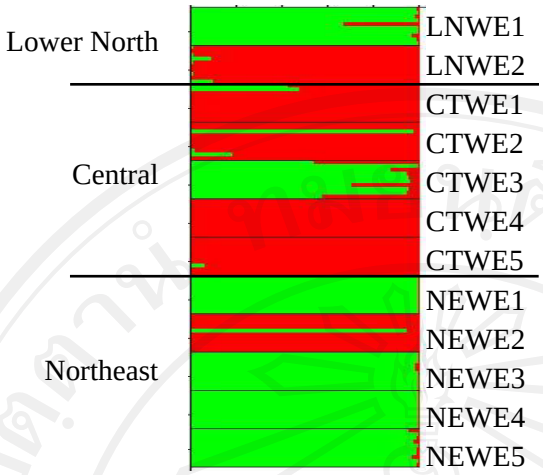

Figure 2.3.10 Population assignment of 12 weedy rice varieties reveal *K*=2. Each bar represent each population consisted of 10 individuals. Different colors represent different inferred populations, referred to different *K*.

Table 2.3.21 Proportion of estimated assign populations (*Q*) of 12 weedy rice populations (*K*=2)

No.	Accession	Province	Rice variety in invaded fields*	n	Given inferred clusters (*Q*)	
					Pop1 (Red)	Pop2 (Green)
1	LNWE1	Phitsanulok	ITV/MV	10	0.042	0.958
2	LNWE2	Phichit	MV	10	0.976	0.024
3	CTWE1	Kanchanaburi	MV	10	0.906	0.094
4	CTWE2	Kanchanaburi	MV	10	0.879	0.121
5	CTWE3	Suphan Buri	MV	10	0.149	0.851
6	CTWE4	Suphan Buri	MV	10	0.997	0.003
7	CTWE5	Suphan Buri	MV	10	0.991	0.009
8	NEWE1	Yasothon	ITV	10	0.004	0.996
9	NEWE2	Surin	ITV	10	0.903	0.097
10	NEWE3	Si Saket	ITV	10	0.006	0.994
11	NEWE4	Buri Rum	ITV/ MV	10	0.003	0.997
12	NEWE5	Buri Rum	ITV	10	0.016	0.984

*__ITV__=improved traditional variety; __MV__=modern variety

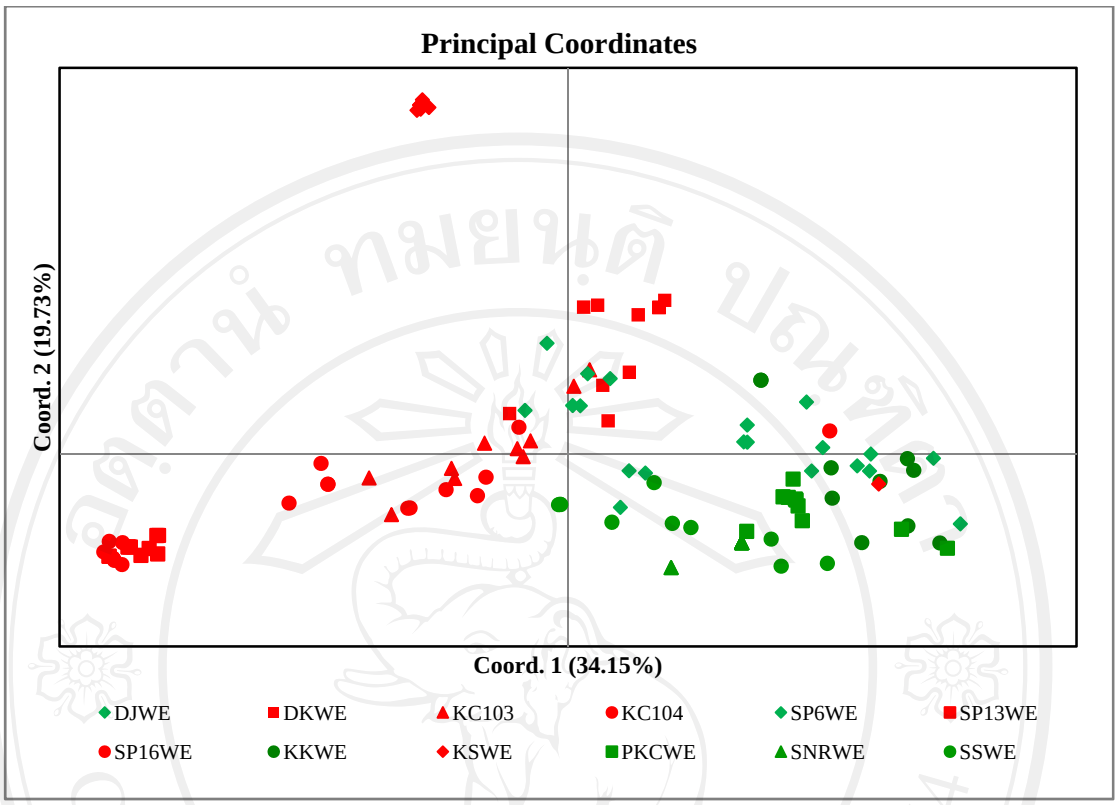

Figure 2.3.11 Distribution of 120 individuals of 12 weedy rice populations. Different color represent 2 clusters referred to the assignment obtained from STRUCTURE: Red-weedy rice collected from modern variety (MV) cultivation fields and Green-weedy rice populations collected from improved traditional variety (ITV) cultivation fields, formed by the principal coordinate analysis (PCA) on the basis of 12 microsatellite markers.

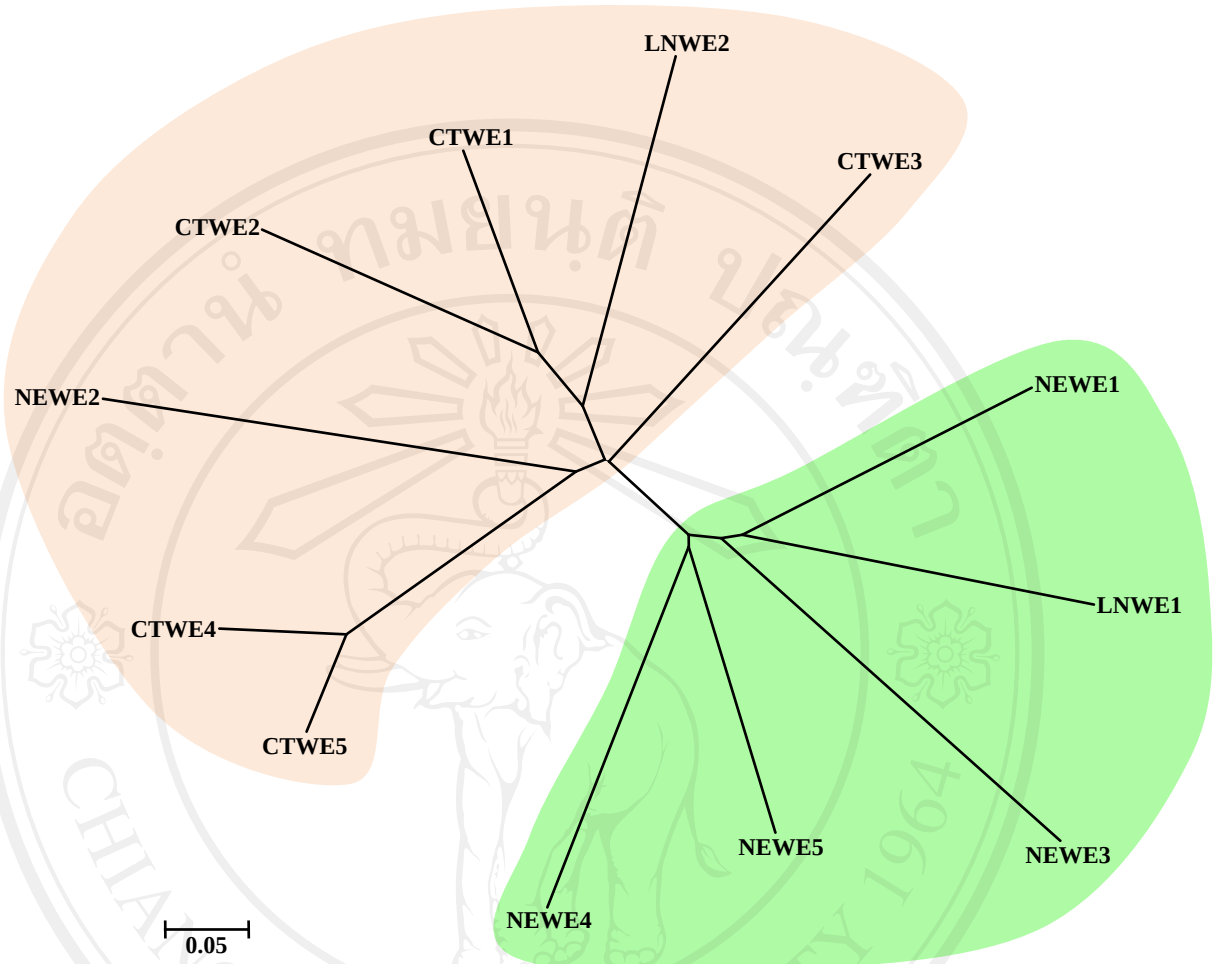

Figure 2.3.12 Cluster analysis using NJ method based on C.S. Chord (1967) genetic distance among 42 common wild rice populations in Thailand

Genetic structure and relationship among wild, cultivated and weedy rice

Genetic diversity

When data from all wild, cultivated and weedy races were pooled and analyzed together, high variation in SSR loci was found among and within the *Oryza sativa* primary gene pool in Thailand (Table 2.3.22). Three hundred and eighteen polymorphic alleles were detected in 420 individuals of common wild rice, 120 individuals of weedy rice and 37 cultivated rice varieties at 12 SSR loci. An average of 26.5 alleles per locus was observed, with 8 alleles in RM477 to 37 alleles in RM206, RM251, RM316 and RM481. Common wild rice illustrated the highest total number of allele, 295 alleles where mostly partitioned to perennial type. Cultivated rice showed the lowest total number of allele, 117 alleles. While number of alleles of weedy rice were between common wild rice and cultivated rice, 140 alleles.

In term of genetic diversity by comparison among rice group (Table 2.3.23), common wild rice showed the highest level of average and total genetic diversity where most variation found within common wild rice was apportioned into perennial group. In contrast, no variation was detected within variety of cultivated rice but high level of total genetic diversity was found. Twelve weedy rice populations showed intermediate value of genetic diversity fall between forty-two common wild rice populations and thirty-seven cultivated rice varieties but has the lowest value of total genetic diversity.

Considering mating system, inbreeding coefficient was the lowest in common wild rice, on the other hand, outcrossing rate (*t*) was the highest. Within common wild rice, perennial type displayed higher out crossing rate but lower inbreeding level than annual type. In contrast, cultivated rice was autogamy illustrated by F_{IS}=1.0, and

t=0.0. Whereas weedy rice showed lower level of outcrossing compared with common wild rice (Table 2.3.23). Analysis of molecular variance (AMOVA) was conducted to investigate the overall distribution of genetic diversity among rice groups (Table 2.3.24). About 6% of total variance was partitioned among rice groups. In addition, pairwise genetic differentiation between and within 3 rice groups was investigated (Table 2.3.25). Degree of genetic differentiation (F_{ST}) from the highest to the lowest were cultivated rice, weedy rice and wild rice, respectively. Pairwised F_{ST} was highest between cultivated vs weedy rice, followed by cultivated vs wild rice and lowest in wild rice vs weedy rice (Table 2.3.25).

Table 2.3.22 Summary of allele diversity of 42 common wild rice populations , 37 cultivated varieties and 12 weedy rice populations using 12 SSR markers.

Population	Number of alleles (A)												Total no. of allele	Average no. of allele
	RM1	RM206	RM109	RM247	RM211	RM477	RM251	RM234	RM259	RM133	RM316	RM481		
Common wild rice														
Range	2-10	1-11	1-8	1-9	1-6	1-3	1-9	1-12	2-7	1-7	1-7	2-11	21-78	
Average	4.8	4.2	4.5	4.2	3.2	1.6	4.3	5.1	4.6	3.5	4.3	4.9	49.5	4.1*§
Total	30	35	22	20	14	8	33	26	33	9	34	31	295	24.5§§
Cultivated rice														
Range	1	1	1	1	1	1	1	1	1	1	1	1	1	
Average	1	1	1	1	1	1	1	1	1	1	1	1	12	1
Total	13	16	8	9	12	1	12	11	7	6	8	14	117	9.7§§
Weedy rice														
Range	1-8	1-8	1-3	1-6	1-4	1-2	2-4	2-7	1-6	1-3	1-5	4-7	21-57	
Average	8.3	8.5	4.0	6.6	4.4	1.4	7.5	8.2	6.3	3.5	5.2	8.7	37	3.1*§
Total	20	15	5	10	7	2	16	14	10	4	14	23	140	11.7§§
Total														
Range	1-10	1-11	1-8	1-9	1-7	1-3	1-9	1-12	1-7	1-8	1-8	1-11	12-78	
Average	3.1	2.7	2.7	2.7	2.2	1.3	2.7	3.2	2.9	2.3	2.8	3.4	32	2.7*§
Total	31	37	22	21	16	8	37	28	33	11	37	37	318	26.5§§

§§Average number of alleles per locus
*§Average number of alleles per population

Table 2.3.23 Summary of genetic parameters of common wild rice, weedy rice, and cultivated rice based on 12 SSR markers

Population	N	n	A	Ho	h	H_S	H_T	F_{ST}	F_{IS}	*t*
Common wild rice	42	420	295	0.456	0.560	0.563	0.827	0.428	0.197	0.671
Perennial	26	290	253	0.537	0.809	0.635	0.815	0.335	0.154	0.733
Annual	16	130	209	0.321	0.764	0.445	0.783	0.527	0.557	0.285
Weed rice	12	120	140	0.238	0.457	0.457	0.736	0.527	0.601	0.249
Cultivated rice	37	390	117	0	0	0.000	0.780	1.000	1.000	0.000

Number of populations (N), Number of individuals (n), Number of alleles (A), Observed heterozygosity (Ho), Nei's (1973) gene diversity (h), Average gene diversity (H_S), Total gene diversity (H_T), Degree of genetic differentiation (F_{ST}), Inbreeding coefficient (F_{IS}) and Outcrossing rate (*t*)

Table 2.3.24 Analysis of molecular variance (AMOVA) for among and within 3 rice groups; common wild rice, cultivated rice and weedy rice on 12 SSR loci

Source	d.f.	SS	Variance component	% of the total variance
Among rice groups	2	799.431	1.024	6%
Populations/Group	88	10405.117	11.358	67%
Populations/Common wild rice	41	2963.176	6.374	43%
Populations/Cultivated rice	36	6724.746	186.798	100%
Populations/Weedy rice	11	898.800	7.492	52%
Individuals/Population	819	3819.000	4.663	27%

Table 2.3.25 Average pairwise genetic differentiation (F_{ST}) within and between 3 rice groups

Rice group	Genetic differentiation (F_{ST})	
	Within	Between
Wild	0.305	
Cultivated		0.657
Weed		0.371
Cultivated	0.998	
Weed		0.699
Weed	0.352	

Population structure

As gene flow is expected to be the major ongoing evolutionary process of rice evolution among common wild rice, Asian cultivated rice and weedy rice, the admixture assignment among 42 common wild rice populations, 12 weedy rice populations and 11 popular cultivated rice varieties (Table 2.2.6) were assessed. Those three rice groups were evaluated together using correlated allele frequency and admixture model in STRUCTURE in a single analysis that assumed 5 clusters; 2 wild clusters and 3 cultivated clusters. As expected, after the burn-in period of 100000 generations and 100000 Markov Chain Monte Carlo replications, the maximum inferred population was *K*=5 represent 5 distinct clusters, included 2 clusters for wild rice (W1 and W2) and 3 clusters for cultivated rice (Cr1, Cr2 and Cr3). While weedy rice populations showed admixture between cultivated and wild rice signatures and were separated into 2 groups according to its companion cultivated variety (Figure 2.3.13 and Table 2.3.26).

The genome proportions (*Q*) was used to determine the membership of rice populations. Cultivated rice varieties CNT1, SPR1 and PSL2 were assigned to the member of Crop1 gene pool (Cr1), PSL1, PSL3, PTT1 and PTT2 to Crop2 gene pool (Cr2) and KDML105, RD6, SPT1 and SKN to Crop3 gene pool (Cr3. While common wild rice populations were structured into 2 gene pools, perennial (W1) and annual (W2) gene pools (Figure 2.3.13 and Table 2.3.26).

Weedy rice was the admixture between wild and cultivated genotypes but more close to cultivated genotype (Figure 2.3.13 and Table 2.3.26). Consequently, weedy rice populations were clustered into two different patterns according to their companion cultivated rice varieties. The genome assignments of LNWE1, CTWE3,

NEWE1, NEWE2, NEWE3, and NEWE5 contained cultivated KDML105 and RD6 signatures represent by the presence of blue color (the genotype of KDML105 and RD6 cluster, Cr3) and common wild rice genotype represent by the presence of red and/or yellow color (the genotype of common wild rice cluster). These cultivated varieties are commonly found cultivation in the North and Northeast paddy fields. Whereas LNWE2, CTWE1, CTWE2, CTWE4, and CTWE5 were consisted of cultivated CNT1 and SPR1 genotype represented by the presence of pink color (the genotype of CNT1 and SPR1 cluster, Cr1) and common wild rice genotype, represent by the presence of red and/or yellow color (the genotype of common wild rice perennial type and annual type, respectively). With exception, the weedy rice population of NEWE2 where genome proportion was assigned mostly to Cr2 gene pool (PSL1, 2, 3 and PTT1, 2 varieties) represent by the presence of green color in Figure 2.3.13. These cultivated varieties are commonly cultivated in Lower North and Central paddy fields.

In addition, the introgression of cultivated gene pool into common wild rice populations were detected indicated by the presence of cultivated genotypes: Cr1 (pink) in 18 wild populations , Cr2 (green) in 10 wild populations or Cr3 (blue) in 23 population with different proportions of the admixtures range from 0.011 to 0.861 (Figure 2.3.13 and Table 2.3.26).

The relationship between wild, weedy and cultivated rice was analyzed using principal coordinate analysis (PCA) displayed that cultivated rice and common wild rice were widely distributed across the graph. Whereas weedy rice populations were distributed among cultivated rice and common wild rice (Figure 2.3.14) particularly distributed within annual type of common wild rice cluster. Neighbor-join tree also

revealed relationship among rice groups illustrated that common wild rice populations were clustered into 2 groups, perennial and annual type whereas weedy rice were mixed within the cultivated cluster (Figure 2.3.15).

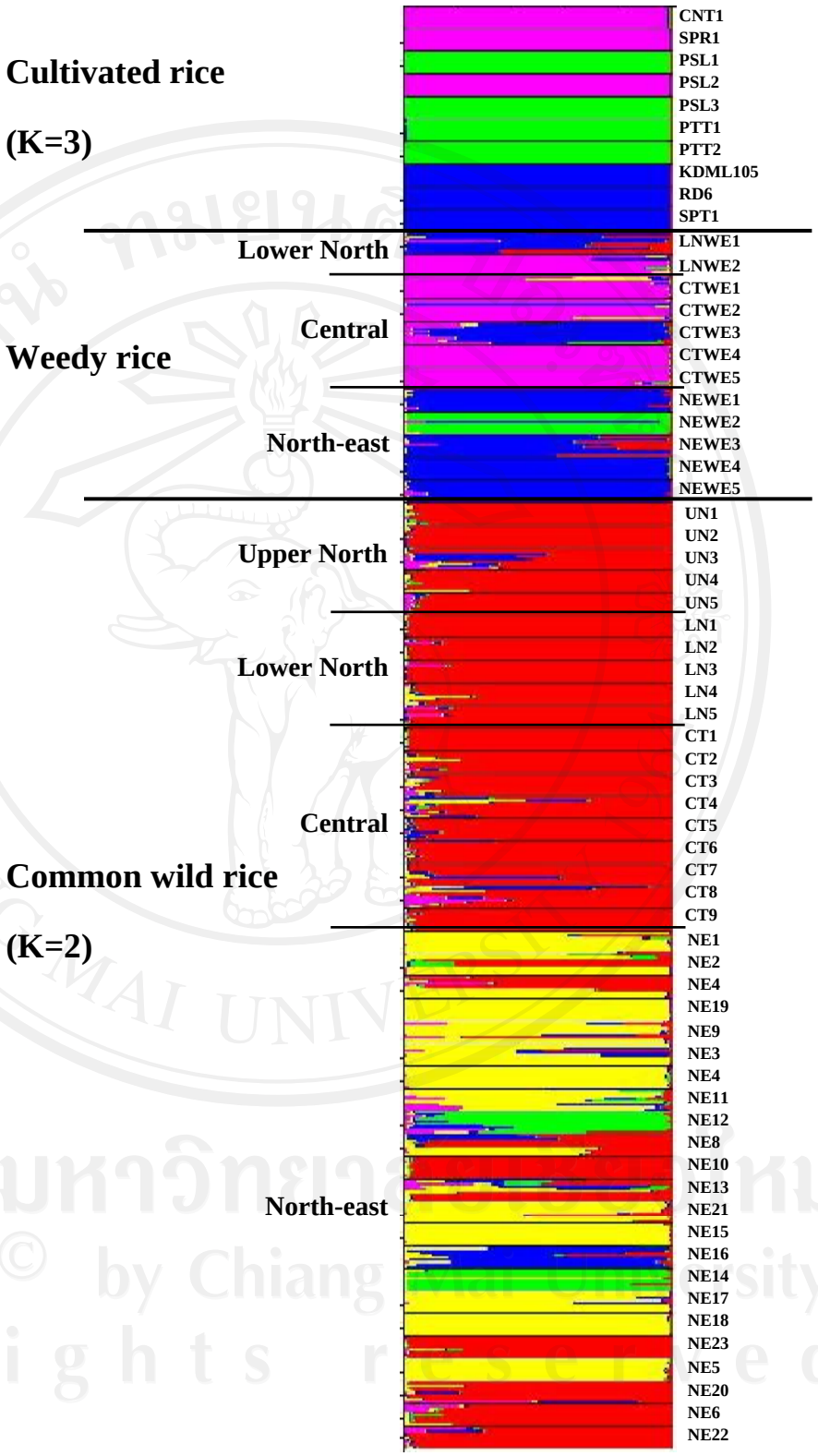

Figure 2.3.13 Population assignment of cultivated, weed, and wild populations illustrated that weedy rice genotypes were the mixture of cultivated rice (K=3; Pink, Green and Blue) and common wild rice (K=2; Red and Yellow). Each bar represent each population consisted of 10 individuals. Different colors represent different inferred populations.

Table 2.3.26 Inferences of cultivated, weedy and common wild rice populations to each inferred population (K=5)

No.	Accession	Population	Type	n	Given inferred clusters (*Q*)				
					Cultivated rice			Common wild rice	
					Crop1 (Cr1)	Crop2 (Cr2)	Crop3 (Cr3)	Perennial (W1)	Annual (W2)
1	CNT1	Chai Nat 1	cultivated	10	0.981	0.009	0.006	0.002	0.002
2	SPR1	Suphan Buri 1	cultivated	10	0.987	0.002	0.006	0.003	0.002
3	PSL1	Phitsanulok 1	cultivated	10	0.002	0.992	0.002	0.002	0.002
4	PSL2	Phitsanulok 2	cultivated	10	0.989	0.004	0.002	0.002	0.002
5	PSL3	Phitsanulok 3	cultivated	10	0.003	0.991	0.002	0.002	0.002
6	PTT1	Pathum Thani 1	cultivated	10	0.004	0.990	0.002	0.002	0.002
7	PTT2	Pathum Thani 2	cultivated	10	0.002	0.992	0.002	0.002	0.002
8	KDML105	Khao Dawk Mali 105	cultivated	10	0.003	0.003	0.989	0.003	0.002
9	RD6	RD 6	cultivated	10	0.004	0.003	0.988	0.003	0.003
10	SPT1	San-pah-tawng1	cultivated	10	0.003	0.004	0.987	0.003	0.002
11	LNWE1	Phitsanulok	weedy	10	0.040	0.004	**0.665**	**0.279**	0.013
12	LNWE2	Phichit	weedy	10	**0.907**	0.012	0.061	0.004	0.017
13	CTWE1	Kanchanaburi	weedy	10	**0.892**	0.003	0.021	0.005	0.080
14	CTWE2	Kanchanaburi	weedy	10	**0.849**	0.002	**0.103**	0.004	0.043
15	CTWE3	Suphan Buri	weedy	10	**0.173**	0.041	**0.742**	0.016	0.028
16	CTWE4	Suphan Buri	weedy	10	**0.990**	0.003	0.003	0.002	0.003
17	CTWE5	Suphan Buri	weedy	10	**0.978**	0.009	0.006	0.003	0.004
18	NEWE1	Yasothon	weedy	10	0.009	0.003	**0.962**	0.018	0.009
19	NEWE2	Surin	weedy	10	0.012	**0.885**	0.089	0.004	0.010
20	NEWE3	Si Saket	weedy	10	0.016	0.004	**0.793**	**0.183**	0.005
21	NEWE4	Buri Rum	weedy	10	0.003	0.007	**0.984**	0.003	0.004
22	NEWE5	Buri Rum	weedy	10	0.019	0.003	**0.965**	0.007	0.006
23	UN1	Chiang Rai	wild	10	0.006	0.008	0.007	0.964	0.013
24	UN2	Chiang Rai	wild	10	0.004	0.003	0.004	0.981	0.008
25	UN3	Chiang Mai	wild	10	0.067	0.006	0.167	0.727	0.033
26	UN4	Lumphun	wild	10	0.003	0.012	0.003	0.947	0.035
27	UN5	Lumphun	wild	10	0.027	0.007	0.012	0.950	0.004
28	LN1	Phitsanulok	wild	10	0.002	0.003	0.003	0.987	0.005
29	LN2	Phitsanulok	wild	10	0.014	0.005	0.007	0.968	0.006
30	LN3	Phichit	wild	10	0.018	0.003	0.008	0.967	0.004
31	LN4	Sukhothai	wild	10	0.006	0.005	0.015	0.921	0.053
32	LN5	Phetchabun	wild	10	0.029	0.004	0.012	0.949	0.006
33	CT1	Ang Thong	wild	10	0.004	0.006	0.005	0.982	0.005
34	CT2	Phra Nakhon Si Ayutthaya	wild	10	0.008	0.005	0.016	0.931	0.040
35	CT3	Phra Nakhon Si Ayutthaya	wild	10	0.018	0.006	0.022	0.923	0.030
36	CT4	Nakhon Nayok	wild	10	0.023	0.015	0.096	0.773	0.093
37	CT5	Sing Buri	wild	10	0.006	0.007	0.037	0.946	0.005
38	CT6	Suphan Buri	wild	10	0.017	0.004	0.035	0.935	0.009
39	CT7	Kanchanaburi	wild	10	0.005	0.005	0.064	0.904	0.023
40	CT8	Kanchanaburi	wild	10	0.098	0.006	0.150	0.678	0.068
41	CT9	Prachin Buri	wild	10	0.007	0.007	0.009	0.966	0.011
42	NE1	Sa Kaeo	wild	10	0.004	0.011	0.013	0.054	0.918
43	NE2	Buri Rum	wild	10	0.005	0.072	0.005	0.291	0.627
44	NE3	Buri Rum	wild	10	0.011	0.005	0.133	0.052	0.799
45	NE4	Buri Rum	wild	10	0.007	0.007	0.006	0.003	0.976
46	NE5	Buri Rum	wild	10	0.004	0.004	0.005	0.012	0.975
47	NE6	Buri Rum	wild	10	0.043	0.022	0.006	0.893	0.036
48	NE7	Nakhon Ratchasima	wild	10	0.059	0.006	0.007	0.543	0.386
49	NE8	RoiEt	wild	10	0.006	0.007	0.157	0.546	0.284
50	NE9	RoiEt	wild	10	0.034	0.004	0.056	0.121	0.784
51	NE10	RoiEt	wild	10	0.006	0.007	0.004	0.975	0.009
52	NE11	Udon Thani	wild	10	0.077	0.052	0.117	0.028	0.725
53	NE12	Udon Thani	wild	10	0.061	0.771	0.102	0.041	0.026
54	NE13	Surin	wild	10	0.053	0.057	0.156	0.507	0.227
55	NE14	Surin	wild	10	0.003	0.861	0.004	0.024	0.108
56	NE15	Ubon Ratchathani	wild	10	0.003	0.003	0.002	0.005	0.986
57	NE16	Yasothon	wild	10	0.006	0.020	0.814	0.072	0.088
58	NE17	Yasothon	wild	10	0.006	0.003	0.051	0.007	0.932
59	NE18	Maha Sarakham	wild	10	0.005	0.005	0.005	0.004	0.982
60	NE19	Maha Sarakham	wild	10	0.003	0.003	0.004	0.003	0.987
61	NE20	Sakon Nakhon	wild	10	0.051	0.004	0.057	0.849	0.039
62	NE21	Si Saket	wild	10	0.004	0.004	0.012	0.090	0.890
63	NE22	Si Saket	wild	10	0.039	0.005	0.031	0.919	0.007
64	NE23	Nong Khai	wild	10	0.005	0.008	0.002	0.967	0.018

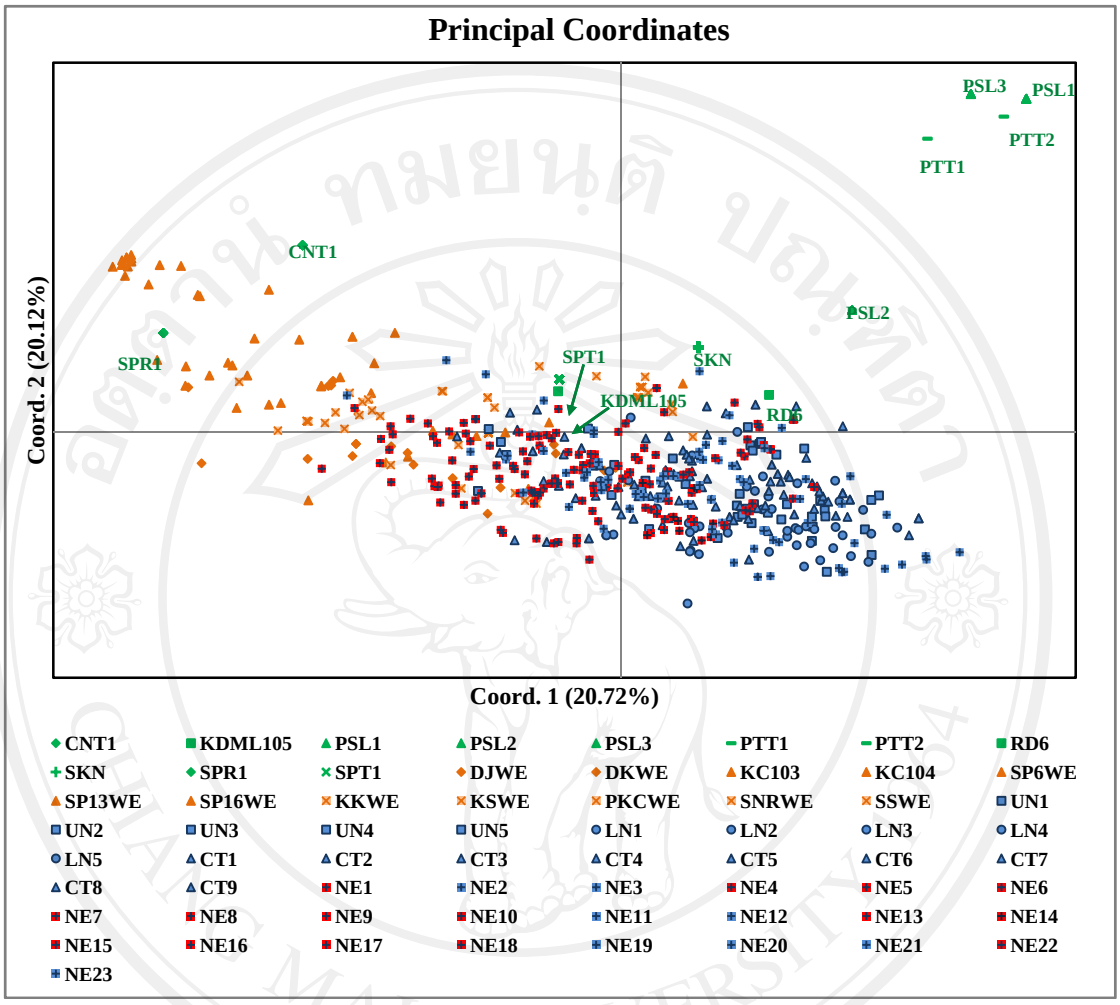

Figure 2.3.14 Distribution of *Oryza sativa* complex including, 42 common wild rice populations, 12 weedy rice populations and 11 cultivated varieties, different color represent different groups: Blue – common wild rice perennial type; Red – common wild rice annual type; Orange – weedy rice; Green-cultivated rice, formed by the principal coordinate analysis on the basis of 12 microsatellite markers.

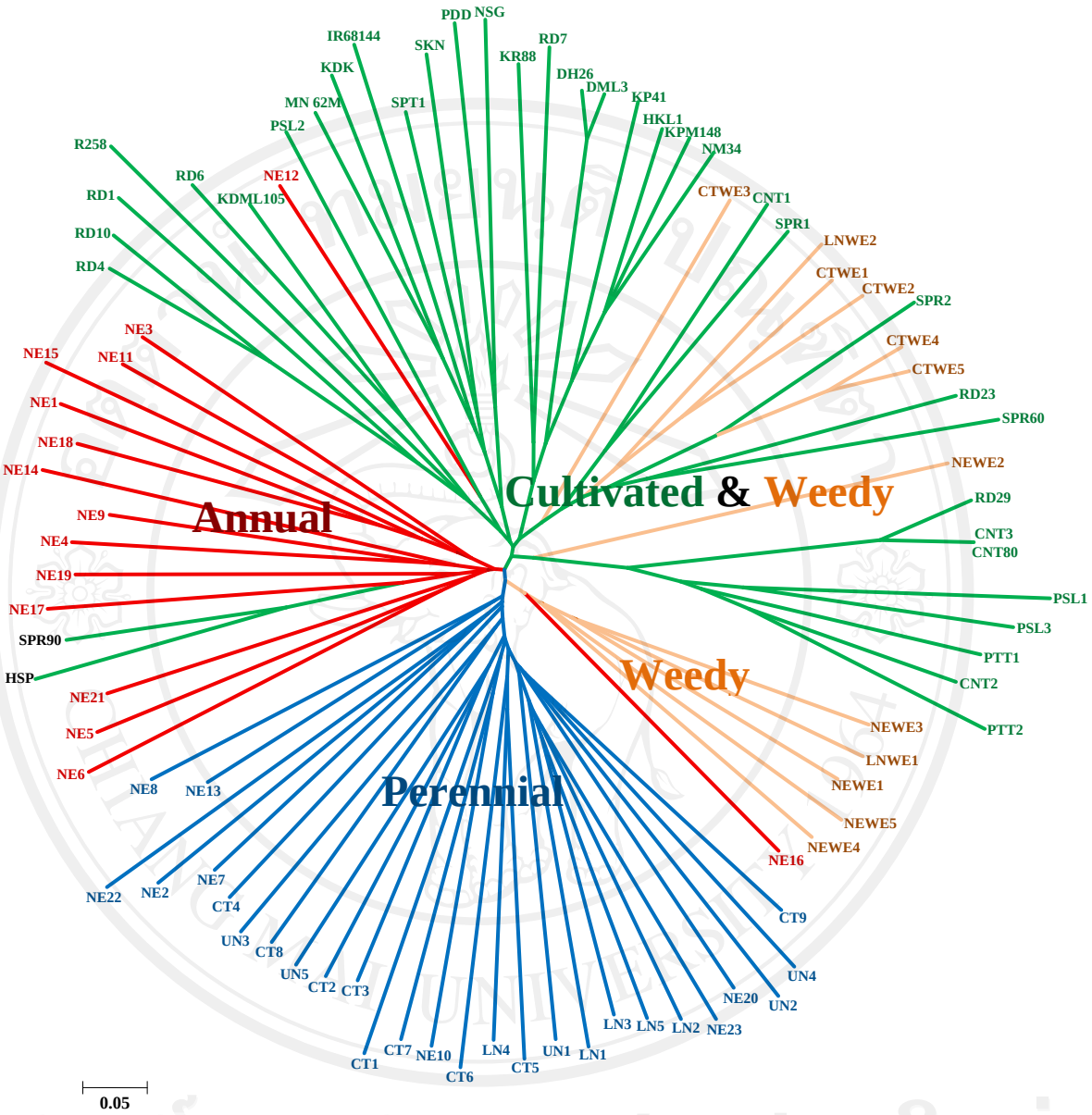

Figure 2.3.15 Cluster analysis using NJ method based on C.S. Chord (1967) genetic distance among 37 cultivated rice varieties (Green), 12 weedy rice populations (Orange) and 42 common wild rice populations (Blue and Red) in Thailand constructed based on 12 microsatellite loci.

Evidence of widespread invasive weedy rice in various areas in Thailand

To test the two hypotheses of the evidence of widespread invasive weedy rice in various areas, 1) invasive weedy rice was spread via seed flow from Central to Lower North and Northeast or 2) invasive weedy rice of each region occurred independently from interspecific hybridization between local wild populations and local popular cultivated rice varieties. The two analyses were conducted; i) principal coordinate analysis (PCA) was used to evaluate the distribution of weedy rice and to illustrate relationship among 12 weedy rice populations, and between weedy rice and cultivated rice and ii) Bayesian MCMC method was used to estimate the proportion of weedy rice's genome to the given reference populations (i.e. estimation of admixture coefficients, *Q*)

The relationship between 12 weedy rice and 11 popular cultivated rice varieties reveal by the PCA (Figure 2.3.16). As expect, twelve weedy rice populations were clustered into 2 groups according to 2 different companion cultivated rice groups. The first group consisted of weedy rice populations collected from modern varieties cultivation fields where mostly found in Lower North and Central. While the second group was weedy rice from improved traditional varieties cultivation fields where mostly found in North and Northeast. Except the population of NEWE5 (Si Saket) was distributed between CNT1 and SPR1 (Figure 2.3.16) and f LNWE1 (Phitsanulok) distributed between KDML105 and RD6.

Bayesian clustering analysis obtained by STRUCTURE program was used to examine genome assignment of weedy rice. Proportion of weedy rice genome was estimated for the assignment to the 5 inferred rice gene pools; 2 wild gene pools and 3 cultivated gene pools as the results from the previous experiment (2.3.3). As

indicated by the *Q* value, 12 weedy rice populations were roughly assigned into 3 groups (Figure 2.3.17 and Table 2.3.27). The first weedy rice group contained about 76-98% genetic constitution of Crop1 gene pool (Cr1; CNT1 and SPR1), including, LNWE2 (Phichit), CTWE1-5 and NEWE5 (Si Saket). The second group contained genetic signatures of Crop3 gene pool (Cr3; KDML105, RD6, SPT1 and SKN) with the proportion more than 90% in NEWE1 (Yasothon), and 62, 65 and 49% in NEWE3 (Buri Rum), NEWE4 (Surin) and LNWE1 (Phitsanulok), respectively. About 1-2% of Crop 2 genetic signature was found in three weedy rice populations including, LNWE2 (Phichit), CTWE3 (Suphan Buri) and NEWE2 (Buri Rum). Interestingly, Weedy rice population of NEWE5 (Si Saket) was mixed with76% of Crop1 gene pool (Cr1) and 22% of Crop3 gene pool (Cr3).

In addition, I also observed the admixtures of common wild rice gene pool in weedy rice genome (Table 2.3.27). Weedy rice population of Phitsanulok (LNWE1) was admixed with 11% of Crop1, 48% of Crop3 and 40% of perennial membership. While genetic assignment of most weedy rice populations of Northeast were between Crop3 gene pool and common wild rice of both perennial type and annual type. Weedy rice of NEWE3 (Buri Rum) was admixed with 62% of Crop3 and 34% of perennial memberships, NEWE4 (Surin) with 65% of Crop3 and 32% of annual membership and NEWE2 (Buri Rum) with 8% of Crop3 and 89% of annual membership.

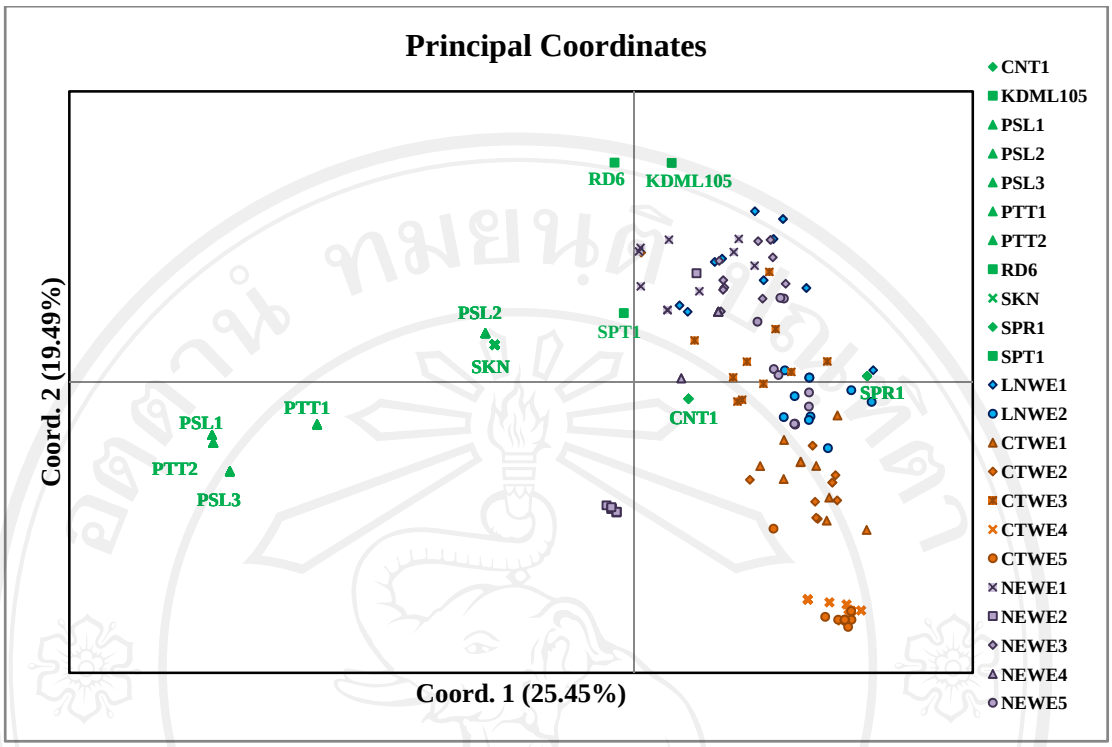

Figure 2.3.16 Distribution of 10 cultivated rice varieties and 120 individuals of 12 weedy rice populations, different color represent different rice group: Green-popular cultivated rice, Blue - Lower North, Orange - Central and Purple- North-east formed by the principal coordinate analysis on the basis of 12 microsatellite markers

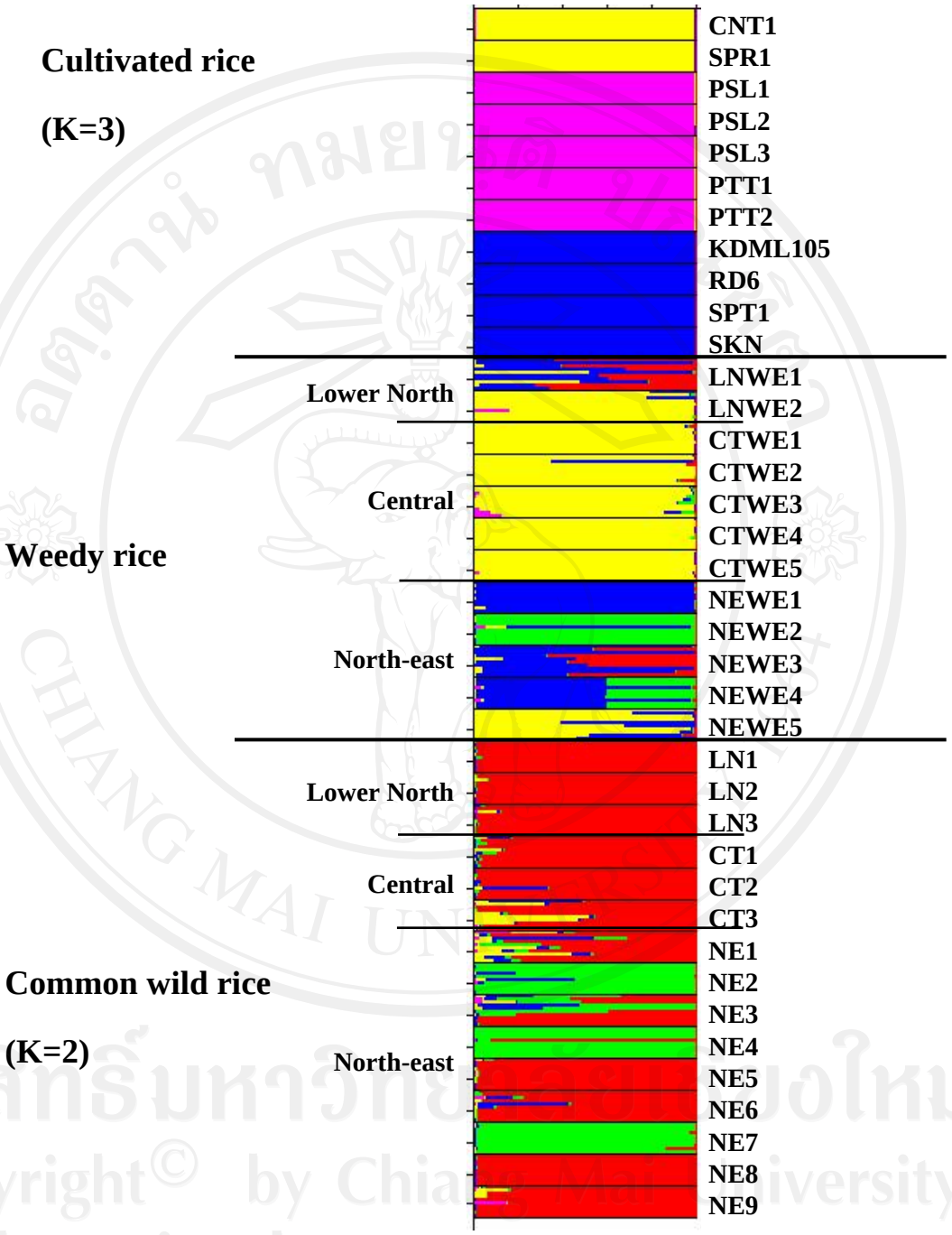

Figure 2.3.17 Population assignment of cultivated, weed, and wild populations illustrated that weedy rice genotypes were the mixture of cultivated rice (K=3; Yellow, Pink and Blue) and common wild rice (K=2; Red and Green). Each bar represent each population consisted of 10 individuals. Different colors represent different inferred populations.

Table 2.3.27 Inferences of admixed ancestors of weedy rice populations to cultivated gene pools (K=3) and common wild rice gene pools (K=2).

No.	Accession	Population	Type	n	Given inferred clusters (*Q*)				
					Cultivated rice			Common wild rice	
					Crop1 (Cr1)	Crop2 (Cr2)	Crop3 (Cr3)	Perennial (W1)	Annual (W2)
1	CNT1	Chai Nat 1	Cr1	10	**0.979**	0.011	0.005	0.002	0.003
2	SPR1	Suphan Buri 1	Cr1	10	**0.987**	0.002	0.006	0.002	0.002
3	PSL1	Phitsanulok 1	Cr2	10	0.002	**0.993**	0.002	0.002	0.002
4	PSL2	Phitsanulok 2	Cr2	10	0.003	**0.990**	0.003	0.002	0.002
5	PSL3	Phitsanulok 3	Cr2	10	0.002	**0.993**	0.002	0.002	0.002
6	PTT1	Pathum Thani 1	Cr2	10	0.002	**0.992**	0.002	0.002	0.002
7	PTT2	Pathum Thani 2	Cr2	10	0.002	**0.992**	0.002	0.002	0.002
8	KDML105	Khao Dawk Mali 105	Cr3	10	0.003	0.002	**0.991**	0.002	0.002
9	RD6	RD 6	Cr3	10	0.003	0.003	**0.990**	0.002	0.002
10	SPT1	San-pah-tawng1	Cr3	10	0.002	0.004	**0.990**	0.002	0.002
11	SKN	Sakon Nakorn	Cr3	10	0.003	0.003	**0.990**	0.002	0.002
12	LNWE1	Phitsanulok	weedy	10	**0.110**	0.002	**0.484**	**0.401**	0.003
13	LNWE2	Phichit	weedy	10	**0.928**	0.019	0.044	0.002	0.007
14	CTWE1	Kanchanaburi	weedy	10	**0.983**	0.003	0.005	0.007	0.003
15	CTWE2	Kanchanaburi	weedy	10	**0.912**	0.002	0.066	0.016	0.003
16	CTWE3	Suphan Buri	weedy	10	**0.929**	0.027	0.014	0.006	0.024
17	CTWE4	Suphan Buri	weedy	10	**0.987**	0.003	0.002	0.002	0.005
18	CTWE5	Suphan Buri	weedy	10	**0.987**	0.005	0.003	0.003	0.003
19	NEWE1	Yasothon	weedy	10	0.009	0.002	**0.980**	0.005	0.004
20	NEWE2	Buri Rum	weedy	10	0.012	0.010	0.084	0.003	**0.892**
21	NEWE3	Buri Rum	weedy	10	0.027	0.002	**0.625**	**0.342**	0.004
22	NEWE4	Surin	weedy	10	0.009	0.008	**0.653**	0.007	**0.324**
23	NEWE5	Si Saket	weedy	10	**0.759**	0.003	**0.223**	0.012	0.003
24	LN1	Phitsanulok	W1	10	0.003	0.002	0.003	**0.985**	0.006
25	LN2	Phitsanulok	W1	10	0.008	0.003	0.003	**0.983**	0.003
26	LN3	Phichit	W1	10	0.011	0.004	0.007	**0.972**	0.005
27	CT1	Suphan Buri	W1	10	0.022	0.002	0.020	**0.936**	0.020
28	CT2	Kanchanaburi	W1	10	0.007	0.002	0.036	**0.949**	0.006
29	CT3	Kanchanaburi	W1	10	0.166	0.006	0.052	**0.770**	0.006
30	NE1	Yasothon	W1	10	0.131	0.026	0.091	**0.681**	0.071
31	NE2	Yasothon	W2	10	0.007	0.004	0.063	0.004	**0.922**
32	NE3	Surin	W1	10	0.027	0.011	0.106	**0.559**	**0.297**
33	NE4	Si Saket	W2	10	0.002	0.003	0.002	0.095	**0.898**
34	NE5	Buri Rum	W1	10	0.007	0.005	0.003	**0.976**	0.009
35	NE6	Buri Rum	W1	10	0.007	0.009	0.062	**0.906**	0.016
36	NE7	Buri Rum	W2	10	0.005	0.003	0.005	0.021	**0.966**
37	NE8	Buri Rum	W1	10	0.004	0.003	0.004	**0.987**	0.003
38	NE9	Buri Rum	W1	10	0.027	0.020	0.003	**0.945**	0.004

2.4 *Discussion*

The *Oryza sativa* primary gene pool is under ongoing evolutionary. The complex consists of 3 components: common wild rice (*O. rufipogon* Griff.) - the wild ancestor of Asian cultivated rice, Asian cultivated rice (*O. sativa* L.) and weedy form (*O. sativa* f *spontanea*) of Asian cultivated rice (Chang, 1976). This chapter presented the population genetic structure of each component and the 3 components analyzed together and elucidated what evolutionary forces play a role on the dynamics of the complex. The following 4 main results were obtained (1) common wild rice retained high level of genetic variation both within and among populations and was structured based on life-history traits: perennial and annual types, (2) great total genetic diversity was observed among 37 cultivated varieties although experiencing the genetic variation reduction from wild ancestors due to the domestication bottleneck , (3) weedy rice populations displayed high total genetic diversity with various levels of genetic variation within populations and were structured based on its companion cultivated rice varieties which led to, (4) gene flow being suggested as ongoing evolutionary process that cause the increase or decrease of genetic variation of the wild and weedy rice in the rice landscape; gene flow was observed between cultivated rice vs wild rice, and cultivated rice vs weedy rice. The present results were consistent with the evidence of gene flow among rice gene pools found by several researcher: Langevin *et al*., (1990); Gealy *et al*., (2003); Song *et al*., (2003a); Zhang *et al*., (2003); Chen *et al*., (2004) and Niruntrayakul (2008).

Genetic diversity and population structure

Comparison of levels of genetic diversity within and between the 3 components of *Oryza sativa* primary gene pool based on 12 SSR markers indicated

that common wild rice was maintained the highest total genetic diversity (H_T) over cultivated and weedy rice. The estimation of genetic diversity of both within and among common wild rice populations in Thailand and the previous reports Punyalue (2006) and Wongtamee (2008) indicated high level of genetic diversity, of common wild rice in Thailand detected by SSR markers. Comparison with others studies from China (Song *et al*., 2003b; Zhou *et al*., 2003; Gao, 2004; and Zhou *et al*., 2008), Myanmar (Shishido *et al*., 2006) and Laos (Kuroda et al., 2007), suggested that Thai common wild rice in the present study illustrated the highest level of genetic diversity.

Population structure of common wild rice was predominantly structured base on life-history trait types, perennial and annual types than regional differentiation revealed by higher F_{ST} between life-history traits types than F_{ST} among regions. Wongtamee (2008) also found similarly result that common wild rice in Thailand was differentiated by life-history traits. The differentiations of these 2 types were clearly different in habitat preferences (Sano and Morishima 1982), and mating system (Barbier 1989a; Oka 1988).

Considering genetic variation of common wild rice at different life-history traits, perennial and annual types, most total variation of common wild rice was apportioned to perennial type resulting higher variation of perennial than annual types. Similar result was found in several studies of common wild rice in Thailand e.g. Morishima and Barbier (1990) using allozyme markers and Wongtamee (2008) using SSR markers. In addition, common wild rice populations of Laos (Kuroda *et al*., 2007), and China (Zhou *et al*., 2007) also illustrated similar trends.

The differentiation in habitat preference between perennial and annual types was observed. Common wild rice is widely distributed in Upper North, Lower North,

Central, and Northeast of Thailand totaling more than half of the country area. Perennial type was found widely in the Upper North, the Lower North, the Central and the Northeast while annual type mostly found in the Northeast of Thailand. Perennial type was commonly found in ditches, swamps or lakes the areas inundated with water through out the year, while annual type is usually found in abandoned fields, near rice fields, roadside or shallow swampy that become parched in dry season. The current results were consistent with Barbier (1989a) and Chitrakon (1995) that common wild rice of the Northeast in Thailand mostly characterized as annual type. The annual habit reflects a strategy for survival through the long dry summer with extremely limited water availability. In this part of the lower Makong watershed, which covers adjoining area of Cambodia, Lao and Thailand, the soil is mostly sandy with very high permeability that dried out quickly after the rain stops (Bell and Seng, 2004).

The differences between perennial and annual were subsequently the production of natural selection of allocation pattern, and such selection lead to variation in adaptive strategies (Sano and Morishima, 1982). Therefore, it would be safer for annual type in dry habitat especially Northern Thailand to produce more panicles and seeds by self-fertilization rather than vegetative reproduction (Sano *et al*., 1980) and selected primarily for greater intrinsic rate of population increasing (Sano and Morishima, 1982). In contrast, habitats of perennial type are inundated through out the year, vegetative propagation would be more advantagenous in competing with other members of the community (Sano *et al*., 1980) and help them to maintain themself in crowded habitats (Sano and Morishima, 1982).

However, perennial and annual types were not clearly separated from each others even their life-history traits and habitat types greatly differ. The present results showed that there were perennial-annual continuum illustrated by the admixtures between perennial and annual genetic constitution using STRUCTURE program (Figure 2.3.4 and Table 2.3.11), e.g. UN2 (Chiang Rai), LN4 (Sukhothai), CT4 (Nakonnayok), NE2 (Buri Rum), and NE13 (Surin). In addition, the PCA graph also showed perennial-annual intermediate populations laid between perennial and annual clusters (Figure 2.3.5). The intermediated form or perennial-annual continuum was reported to be the results of hybridization process between perennial vs annual types (Barbier 1989a) based on isozyme data, or between common wild rice of perennial or annual vs cultivated rice and subsequently adapt to habitats disturbed by human (Oka and Chang, 1961).

The differentiation of common wild rice populations by geographical distribution was not detected in this study due to less geographical differentiation of samples collection areas. Compare with other studies, significant geographical structure of common wild rice as genetic isolation by geographical distance was found. The geographical differentiation in *O. rufipogon* is accounted for geographical isolation demonstrated by Cai et *al.* (2004) found that the common wild rice of India, China, Thailand and Indonesia tend to differentiate by spatial isolation. In addition, the differentiation due to adaptation to the local environments was also observed in collection of common wild rice in China suggested by Song *et al*., (2005) and Wang *et al*., (2008).

As mating system is considered to shape genetic composition of populations reflecting population structure and genetic variation of the populations which

indicated by inbreeding coefficient (F_{IS}) and outcrossing rate (*t*) values. Common wild rice exhibited a wide range of F_{IS} (-0.431 to 0.667) suggested that common wild rice in the present study have wide range of mating system i.e. more outcrossing (F_{IS} <1) to more inbreeding (F_{IS} >0). However, negative value of F_{IS} were detected in seven common wild rice populations (Table 2.3.4) reflected the effect of small population size and inbreeding occurred less often than would be expected at random (Keller and Waller, 2002). In addition, different life-history traits reflected different mating system predominantly outbreeding plant maintaining higher levels of genetic variation than predominantly inbreeding plants (Hamrick *et al*., 1979). Mating system of perennial and annual types was also different. Perennial type is predominantly outcrossing (F_{IS} =154, *t*=0.733) while annual type is predominance of selfing (F_{IS} =0.557, *t*=0.285). The results suggested that low level of outcrossing rate of Northeast region (*t*=0.311) where most populations were classified as annual type reflected higher level of genetic diversity of perennial than annual types as mention above. Similarly, Barbier (1989b) found that estimate of outcrossing rate in perennial (44%) are higher than annual (7.9%) types of common wild rice in Thailand.

High level genetic diversity was observed in 37 Thai cultivated rice varieties was found. In contrast, many studies illustrated severe reduction of genetic variation in domesticated or cultivated races from their wild ancestors. Lower level of genetic diversity of cultivated rice than common wild rice, 80% reduction, was the subsequent of the domestication bottleneck of genetic diversity (Doebley *et al*., 2006; Burger *et al*., 2008) from it wild ancestors (*Oryza rufipogon* Griff.) (Londo *et al*., 2006; Zhu *et al*., 2007). Similar results were observed in other domesticated species such as *Zea mays* (Eyre-Walker *et al*., 1998; Fugunaka *et al*., 2005), *Vigna angularis* (Zong *et al*.,

2003), *Helianthus annuus* (Liu and Burke, 2006) and *Glycine max* (Hyten *et al*., 2006), with 50 – 80% reduction in average. Such domesticated species revealed lower degree of genetic variation reduce from their wild ancestors than cultivated rice observed in the present study. Even as domestication bottleneck causes severe reduction of genetic diversity of domesticated species, however, cultivated rice in the present study still contained high diversity revealed by high value of polymorphic information content (PIC=0.737). In comparison with other studies, Garris *et al*., (2005) illustrated that PIC of 5 groups of cultivated rice; *Aus*, *Indica*, *Aromatic*, *Temperate japonica* and *Tropical japonica* was 0.67. While Londo *et al*., (2006) reported that PIC of cultivated rice; *indica*, *japonica* and *aus* were in range of 0.487 to 0.635. The following chapter (Chapter 3) demonstrated how much genetic variation was maintained in Thai landrace rice. The results illustrated that moderate level of genetic variation was detected within one Thai landrace rice.

In addition, high variation was detected among 37 cultivated varieties, high level of F_{ST} was also found indicating that Thai cultivated rice varieties in the present study have broad genetic base. In addition, according to the analysis based on STRUCTURE program, PCA and NJ clustering 37 cultivated rice varieties were structured into 3 group according to its background pedigree; i) improved traditional varieties, ii) modern varieties and iii) mixture of improved traditional varieties and modern varieties. Differently, Garris *et al*., (2005) demonstrated that 234 accessions of cultivated rice were clustered into 5 groups corresponding to *indica*, *aus*, *aromatic*, *temperate japonica* and *tropical japonica* rice. The difference was due to large number of cultivated rice accessions and larger type of rice in the study of Garris *et al*., (2005) were very much more than the present study. Furthermore, the very large

amount of genetic variation attributed among the 37 pure line cultivated rice reflected the different background ancestors during breeding program as they were bred for different purposes. Garris *et al*. (2005) also discussed that the differentiation among rice groups is the response of structure in ancestral rice populations.

For weedy rice, various levels of variation within population were detected with slightly lower total genetic diversity than common wild rice. Low levels of observed heterozygosity were observed in six weedy rice populations; Central (4), Lower North (1) and Northeast (1). Due to the outcome of the backcrossing between weedy rice and cultivated rice at the initial stage of invasion lead to the convergence of homogeneity of weedy population toward the crop (Langevin *et al*., 1990) which confirmed by the high degree of inbreeding in these six weedy rice populations (Table 2.3.17). In addition, weedy rice populations with low heterogeneity displayed morphological characters similar to cultivated rice. While high levels of observed heterozygosity were detected in the rest six weedy rice populations was the consequence of early hybrid stage of intercross between common wild rice and cultivated rice.

Similar pattern of various levels of genetic diversity within weedy rice population in north-eastern China were found based on 20 SSR loci (Cao *et al*., 2006), but exhibited very much lower levels of overall genetic diversity comparing with weedy rice in Thailand. While overall gene diversity of weedy rice from US were analyzed based on 16 SSR loci (Londo and Schaal, 2007) showed about 10 times lower compare with weedy rice in the present study. In addition, weedy rice found in north-eastern China (Cao *et al*., 2006) and the US (Londo and Schaal, 2007) have low

level of heterozygosity indicated that they are mostly homozygous genotypes which was similar to weedy rice population of LNWE2 (Phichit) in this study.

Weedy rice in Thailand has been shown to be the result of hybridization between common wild rice and cultivated rice (Niruntrayakul, 2008). The population structure of weedy rice in the present study can be structured into 2 clusters according to its companion cultivated rice varieties. The first cluster was the weedy rice of modern rice varieties fields while the second cluster was the weedy rice of improved traditional varieties fields. However, significant differentiation among weedy rice from different regions indicated strong structure of weedy rice among regions. The results were the consequence mainly on different regions grown different cultivated varieties (Table 2.3.6).

As population structure of weedy rice in Thailand elucidated that they are genetically close to the companion cultivated variety. Therefore, tracing the origin of widespread weedy rice in Lower North or Northeast of Thailand based on their population structure whether they were the offspring of '*in situ*' hybridization, or result from another place or time.

Gene flow in the Oryza sativa complex

The origin of invasive weedy rice in many areas worldwide have been evidenced by many workers; the outcome of inter-hybridization between cultivated rice and common wild rice (Cao et al., 2006; Niruntrayakul, 2008), degenerated individuals of cultivated rice (De Wat and Harlan 1975) or in case of weedy rice in the US came from the contamination of seed stocks (Delouche *et al*., 2007) and subsequently hybridized between cultivated rice in US with 2 weedy rice types; blackhulled and strawhulled (Londo and Schaal, 2007). The present study

demonstrated that gene flow was the major evolutionary force influenced the spreading of weedy rice populations.

The spreading of weedy rice in various areas in this study was the consequence of inter-breeding between local wild rice and local popular cultivated rice. These results reflected the continuous gene flow which was the ongoing evolution process play a significant role in population dynamics among wild, cultivated and weedy form. The current results illustrated that genetic constitution of weedy rice consisted of genome those introgression from cultivated rice and common wild rice with different proportion analyzed using the STRUCTURE program (Figure 2.3.16 and Table 2.3.27). Weedy rice in Thailand were structured into 2 groups correspondence to 2 major rice cultivation systems; i) modern variety (CNT1, SPR1) and ii) improved traditional variety (KDML105, RD6).

The first weedy rice group has invaded modern varieties fields in Central and Lower North of Thailand while the second group was found to invade improved traditional varieties in Northeast of Thailand. Proportion of genome assignment for the admixture model of weedy rice demonstrated that weedy rice of different regions had different proportion of genome assignment. Weedy rice of Central and Lower North of Thailand were collected from the cultivated rice fields where modern varieties were grown, their genome proportion was mostly assigned as modern varieties genotype. Whereas weedy rice populations of Northeast were collected from improved traditional varieties rice fields, their genome was mostly assigned as improved traditional varieties. These may be caused by the subsequently backcross of weedy rice to cultivated rice in rice fields where they co-exist and share the same period of flowering time resulting the rapidly loss of wild traits in segregating

populations. However, there were 3 weedy rice populations; LNWE1, NEWE3 and NEWE4 with high genome proportion of wild rice (~50%) may by due to these weedy population were result of a recent invasion and share the same habitat with common wild rice lead to the possibility of backcross of these weedy population to common wild rice. These results were consistent with Londo and Schaal (2007) suggestion that origin of weedy red rice in the US was originated from seed contamination and subsequently hybridized with cultivated rice in the US lead to different types of weedy rice found in the US, strawhulled and blackhulled types. The strawhulled weedy rice derived via hybridization between *O. rufipogon* and *O. sativa indica*. While blackhulled weedy rice formed by either hybridization of *O. sativa Aus* with *O. rufipogon* and subsequence with loss of *O. rufipogon* alleles or the reversion of *Aus* rice to weedy form.

In conclusion, DNA evidence showed high level of genetic differentiation among weedy rice populations indicated that invasive weedy rice found in the rice fields of modern and improved traditional varieties originated from independent hybridization between local popular cultivated rice and local common wild rice of each region. Thus, weedy rice populations were differentiated into 2 groups according to their companion cultivated rice as cultivated rice varieties in this study were also found differentiate between modern and improved traditional varieties. We also found the introgression of cultivated genotype in common wild rice genome, indicating gene flow from cultivated to wild population. These finding suggest that gene flow is the major process in the evolutionary dynamics of *Oryza sativa* primary gene pool complex where they are co-existed lead to increase or decrease genetic

diversity of the wild ancestors and the intermediate form including cultivated rice in the fields.

The impact of gene flow among the *Oryza sativa* primary gene pool complex reflected the evolutionary dynamics of the complex as followed; 1) gene flow between cultivated and wild rice lead to increase or decrease variation of the wild relatives (Ellstrand *et al*., 1999), and originated noxious weedy rice invaded in rice fields, and 2) backcrossing between weedy and cultivated rice resulting crop mimicry of weed lead to the difficulty of weedy rice managements. Weedy rice in rice fields cause rice yield loss and poor quality lead to farmers economic loss, on the other hand, as weedy rice have broad genetic variation may be useful for selecting of beneficial traits as genetic resource for rice breeding. Furthermore, as wild rice is provided not only valuable resources for rice breeding, but also important materials for scientific research therefore, the impact of cultivated gene flow particularly transgenic rice into wild rice populations should be monitor for the escape of transgenes to natural wild populations (Ellstrand, 1988).