An ecogeographic analysis of *Oryza* series *Sativae* in Asia and the Pacific

Maria Celeste N. Banaticla-Hilario

An ecogeographic analysis of *Oryza* series *Sativae* in Asia and the Pacific Maria Celeste N. Banaticla-Hilario

An ecogeographic analysis of *Oryza* series *Sativae* in Asia and the Pacific

Maria Celeste N. Banaticla-Hilario

Thesis committee

Promoter
Prof.dr. M.S.M. Sosef
Professor of Biosystematics
Wageningen University / Naturalis Biodiversity Center

Co-promoters
Dr. R.G. van den Berg
Associate Professor, Biosystematics Group
Wageningen University

Dr. K.L. McNally
Senior Scientist, T.T. Chang Genetic Resources Center
International Rice Research Institute, Los Baños, Philippines

Other members
Prof.dr. B.J. Zwaan, Wageningen University
Dr.ir. T.J.L. van Hintum, Centre for Genetic Resources, Wageningen
Dr. M.J.M. Smulders, Plant Research International, Wageningen
Dr. H. Duistermaat, Leiden University

This research was conducted under the auspices of the Graduate School of Experimental Plant Sciences.

An ecogeographic analysis of *Oryza* series *Sativae* in Asia and the Pacific

Maria Celeste N. Banaticla-Hilario

Thesis
submitted in fulfillment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus
Prof. dr. M.J. Kropff,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Wednesday 24 October 2012
at 11 a.m. in the Aula.

Maria Celeste N. Banaticla-Hilario
An ecogeographic analysis of *Oryza* series *Sativae* in Asia and the Pacific

PhD thesis, Wageningen University, NL (2012)
With references, with summaries in English, Dutch and Filipino

ISBN 978-94-6173-378-8

CONTENTS

Chapter 1:	General Introduction	7
Chapter 2:	Morphological variation patterns between and within *Oryza nivara* and *O. rufipogon*	31
Chapter 3:	Local differentiation amidst extensive allele sharing in *Oryza nivara* and *O. rufipogon*	79
Chapter 4:	Crossability patterns in Asia Pacific *Oryza* series *Sativae*	149
Chapter 5:	Resolving the taxonomic ambiguity of *Oryza nivara*	189
Chapter 6:	General Discussion	207
Summary		221
Samenvatting (*Summary in Dutch*)		225
Kabuuran (*Summary in Filipino*)		229
Acknowledgements		233
Curriculum vitae		235
Education statement		237

CHAPTER 1
General Introduction

The genus *Oryza* L.

It was Linnaeus who, back in 1753, first formally published the genus name *Oryza* L. in his famous Species Plantarum. More than 250 years later, a wealth of information about this grass genus has been and is still being discovered.

Oryza is defined by its oblique spikelets having two rudimentary glumes whose function is taken over by two small sterile lemmas that are acuminate to setiform, and the single fertile lemma being generally awned and crustaceous or rarely chartaceous-coriaceous (Tateoka 1964; Duistermaat 1987). Since its first publication, the circumscription of the genus *Oryza* expanded from a single species *Oryza sativa* (Linnaeus 1753), to more than 20 species found in different regions of the tropics.

Tateoka (1962) classified the genus into species complexes: groups of closely related species. Vaughan (1989) adapted Tateoka`s system and modified it by adding *O. glumaepatula*, *O. granulata*, *O. meridionalis*, *O. nivara* and *O. rhizomatis* to the species list, and removing members that are now relegated to the genus *Leersia*.

Lu (1999) reviewed the taxonomic history of *Oryza* and compared the different subgeneric classifications published by Roschevics (1931), Chevalier (1932), Ghose et al. (1965), Sharma and Shastry (1965, 1972), Tateoka (1963), Oka (1988) and Vaughan (1989). He then proposed his own scheme applying the section and series categories used by Sharma and Shastry (1965). In 2001, Lu et al. modified this classification by combining *O. meyeriana* with *O. granulata* and reviving the species status of *O. malampuzhaensis* and *O. schweinfurthiana*. Vaughan et al. (2003) also revised the species list, synonymizing *O. nivara* with *O. rufipogon* and recognizing *O. malampuzhaensis* as a distinct species.

The most current taxonomic systems of Lu et al. (2001) and Vaughan et al. (2003) are presented in Table 1, together with the classification scheme used in the taxonomic database of the Germplasm Resources Information Network (GRIN Taxonomy) (USDA, ARS, National Genetic Resources Program 2012). GRIN Taxonomy is an online system that provides a standard reference for the classification and nomenclature of economically important plants and their wild and weedy relatives. It has been widely adopted by the genebank community. Taxonomic information in this database is derived from standard botanical literature, recent taxonomic revisions and specialist consultation (USDA, ARS, National Genetic Resources Program 2012).

Table 1. Species of genus *Oryza* as classified by Lu et al. (2001), Vaughan et al. (2003) and the Germplasm Resources Information Network (USDA, ARS, National Genetic Resources Program 2012).

Lu et al. (2001)	Vaughan et al. (2003)	GRIN (2012)	Genome group	Distribution
I. Section *Padia* (Zoll. & Moritzi) Baill.		I. Section *Padia* (Zoll. & Moritzi) Baill.		
1. Series *Meyerianae* Sharma & Shastry	*O. granulata* complex	1. Series *Meyerianae* Sharma & Shastry		
O. granulata Nees & Arn ex Wall. (syn: *O. meyeriana*)	*O. granulata* Nees & Arn ex Wall.	*O. meyeriana* (Zoll. & Moritzi) Baill. var. *granulata* (Nees & Arn ex Wall.) Duist.	GG	South and Southeast Asia
	O. meyeriana (Zoll. & Moritzi) Baill. (syn: *O. neocaledonica*)	*O. meyeriana* var. *meyeriana* (Zoll. & Moritzi) Baill.	GG	Southeast Asia
		O. meyeriana (Zoll. & Moritzi) Baill. var. *inandamanica* (J.L. Ellis) Veldkamp	GG	Andaman and Nicobar Isl.
O. neocaledonica Morat		*O. neocaledonica* Morat	GG	New Caledonia
2. Series *Ridleyanae* Sharma & Shastry	*O. ridleyi* complex	2. Series *Ridleyanae* Sharma & Shastry		
O. longiglumis Jansen	*O. longiglumis* Jansen	*O. longiglumis* Jansen	HHJJ	New Guinea

Table 1. (Continued) Species of genus *Oryza* as classified by Lu et al. (2001), Vaughan et al. (2003) and the Germplasm Resources Information Network (USDA, ARS, National Genetic Resources Program 2012).

Lu et al. (2001)	Vaughan et al. (2003)	GRIN (2012)	Genome group	Distribution
O. ridleyi Hook.f.	*O. ridleyi* Hook.f.	*O. ridleyi* Hook.f.	HHJJ	South Asia
3. Series *Schlechterianae* Sharma & Shastry		3. Series *Schlechterianae* Sharma & Shastry		
O. schlechteri Pilger	*O. schlechteri* Pilger	*O. schlechteri* Pilger	Unknown	New Guinea
II. Section *Brachyantha* B.R.Lu		II. Section *Brachyantha* B.R.Lu		
4. Series *Brachyanthae* Sharma & Shastry		4. Series *Brachyanthae* Sharma & Shastry		
O. brachyantha Chev. & Roehr.	*O. brachyantha* Chev. & Roehr.	*O. brachyantha* Chev. & Roehr.	FF	Africa
III. Section *Oryza*		III. Section *Oryza*		
5. Series *Sativae* Sharma & Shastry	*O. sativa* complex	5. Series *Oryza*		
O. barthii A.Chev.	*O. barthii* A.Chev.	*O. barthii* A.Chev.	AA	Africa
O. glaberrima Steud.	*O. glaberrima* Steud.	*O. glaberrima* Steud.	AA	West Africa

Table 1. (Continued) Species of genus *Oryza* as classified by Lu et al. (2001), Vaughan et al. (2003) and the Germplasm Resources Information Network (USDA, ARS, National Genetic Resources Program 2012).

Lu et al. (2001)	Vaughan et al. (2003)	GRIN (2012)	Genome group	Distribution
O. glumaepatula Steud.	*O. glumaepatula* Steud.	*O. glumipatula* Steud.	AA	South and Central America
O. longistaminata Chev. & Roehr.	*O. longistaminata* Chev. & Roehr.	*O. longistaminata* Chev. & Roehr.	AA	Africa
O. meridionalis N.Q.Ng	*O. meridionalis* N.Q.Ng	*O. meridionalis* N.Q.Ng	AA	New Guinea and tropical Australia
O. sativa L.	*O. sativa* L.	*O. sativa* L.	AA	Cultivated worldwide
O. rufipogon Griff.	*O. rufipogon* Griff. (syn: *O. nivara*)	*O. rufipogon* Griff.	AA	Tropical and subtropical Asia, tropical Australia
O. nivara Sharma & Shastry		*O. nivara* Sharma & Shastry	AA	Tropical and subtropical Asia
6. Series *Latifoliae* Sharma & Shastry	*O. officinalis* complex	6. Series *Latifoliae* Sharma & Shastry		
O. alta Swallen	*O. alta* Swallen	*O. alta* Swallen	CCDD	South and Central America

Table 1. (Continued) Species of genus *Oryza* as classified by Lu et al. (2001), Vaughan et al. (2003) and the Germplasm Resources Information Network (USDA, ARS, National Genetic Resources Program 2012).

Lu et al. (2001)	Vaughan et al. (2003)	GRIN (2012)	Genome group	Distribution
O. eichingeri Peter	*O. eichingeri* Peter	*O. eichingeri* Peter	CC	South Asia and East Africa
O. grandiglumis (Döll) Prodoehl.	*O. grandiglumis* (Döll) Prodoehl.	*O. grandiglumis* (Döll) Prodoehl.	CCDD	South and Central America
O. latifolia Desv.	*O. latifolia* Desv.	*O. latifolia* Desv.	CCDD	South and Central America
O. malampuzhaensis Krish. & Chandr.	*O. malampuzhaensis* Krish. & Chandr.	*O. malampuzhaensis* Krish. & Chandr.	BBCC	India
O. minuta J.Presl & C.Presl	*O. minuta* J.Presl & C.Presl	*O. minuta* J.Presl & C.Presl	BBCC	Philippines and Papua New Guinea
O. officinalis Wall. ex Watt	*O. officinalis* Wall. ex Watt	*O. officinalis* Wall. ex Watt	CC	Tropical and subtropical Asia, tropical Australia
O. rhizomatis Vaughan	*O. rhizomatis* Vaughan	*O. rhizomatis* Vaughan	CC	Sri Lanka
O. punctata Kotschy ex Steud.	*O. punctata* Kotschy ex Steud. (syn: *O. schweinfurthiana*)	*O. punctata* Kotschy ex Steud.	BB, BBCC	Africa
O. schweinfurthiana Prodoehl		*O. schweinfurthiana* Prodoehl	BBCC	Africa

Table 1. (Continued) Species of genus *Oryza* as classified by Lu et al. (2001), Vaughan et al. (2003) and the Germplasm Resources Information Network (USDA, ARS, National Genetic Resources Program 2012).

Lu et al. (2001)	Vaughan et al. (2003)	GRIN (2012)	Genome group	Distribution
7. Series *Australiensis* Tateoka ex Sharma & Shastry		IV. Section *Australiensis*		
O. australiensis Domin.	*O. australiensis* Domin.	*O. australiensis* Domin.	EE	Tropical Australia
		O. coarctata Roxb.	Unknown	Tropical Asia

Several taxa were treated differently in the three schemes (Table 1). Lu et al. (2001) synonymized *O. meyeriana* with *O. granulata* and recognized *O. neocaledonica*, while Vaughan et al. (2003) recognized *O. meyeriana* and placed *O. neocaledonica* under it. In GRIN Taxonomy, *O. meyeriana* is classified into three varieties including *O. granulata* (sensu Vaughan) and variety *inandamanica* (USDA, ARS, National Genetic Resources Program 2012). *O. schweinfurthiana* and *O. nivara* were treated by Vaughan et al. (2003) as the tetraploid form of *O. punctata* and the annual ecotype of *O. rufipogon,* respectively. *O. glumaepatula* is listed as *O. glumipatula* and Series *Sativae* is renamed as Series *Oryza* in GRIN Taxonomy (USDA, ARS, National Genetic Resources Program 2012).

Oryza coarctata Roxb. (synonym of *Porteresia coarctata* (Roxb.) Tateoka) was not included in the treatments of Lu and Vaughan but both accepted its inclusion in the genus (Lu and Ge 2003; Vaughan et al. 2008a) after crossing experiments (Sarker et al. 1993; Farooq et al. 1996) and phylogenetic studies (Ge et al. 2002; Guo and Ge 2005) showed its closer affinity to *Oryza* species. *O. coarctata* is recognized in GRIN Taxonomy but is not designated to any section or series (USDA, ARS, National Genetic Resources Program 2012).

The series in Lu`s classification more or less correspond to the species complexes of Vaughan (Table 1). *O. australiensis,* the only species of series *Australiensis* was included in the *O. officinalis* complex. *O. brachyantha* and *O. schlechteri* were not assigned to any species complex and could represent earlier diverged lineages in the genus (Vaughan et al. 2003, 2005). The scheme used by GRIN Taxonomy is analogous to Lu's except that series *Australiensis* (sensu Lu) was promoted into section *Australiensis*.

Members of a series/species complex possess the same genome type, share certain morphological features, exhibit interfertility to a certain degree, and often occupy similar habitats. Vaughan et al. (2005, 2008a) discussed the habitats, distribution range and dispersal of each species complex in detail.

Phylogenetic analyses using RFLP (Wang et al. 1992), AFLP (Aggarwal et al. 1999), ISSR, (Joshi et al. 2000), MITE-AFLP (Park et al. 2003), SSR and flanking regions (Nishikawa et al. 2005) and DNA sequences (Iwamoto et al. 1999; Takahashi et al. 2008) confirmed the grouping of *Oryza* species into series as proposed by Lu et al. (2001) (Table 1) and identified series *Meyerianae* and *Ridleyanae* as basal groups in the genus.

Series *Sativae*

Also known as the AA genome group, the pantropical series *Sativae* contains 8 diploid species, including two cultigens: *O. sativa* (Asian rice) and *O. glaberrima* (African rice) (Table 1). The eight members of the series will be discussed briefly, arranged according to their geographic distribution.

Africa

There are three species of series *Sativae* in Africa. The annual *O. barthii* is believed to be the ancestral species of the cultigen *O. glaberrima*. These two species consistently show strong affinity in various molecular studies (Wang et al. 1992; Ishii et al. 1996; Kwon et al. 2006; Bautista et al. 2001; Duan et al. 2007). *O. longistaminata* has been reported as being the perennial ancestor of *O. barthii* and *O. glaberrima* (Khush 1997). Although DNA sequence analyses indicate the close relationship of the three African species (Kwon et al. 2006; Duan et al. 2007), data based on isozyme (Second 1985), MITE-AFLP (Park et al. 2003), RAPD, RFLP and SSLP (Bautista et al. 2001) suggest that *O. longistaminata* is genetically distinct from *O. barthii* and *O. glaberrima.* In several phylogenetic studies, *O. longistaminata* was either the earliest (Aggarwal et al. 1999; Iwamoto et al. 1999; Ren et al. 2003) or the second earliest (Wang et al 1992; Park et al. 2003) species to diverge from the rest of the series. In addition to genetic evidence, incongruent molecular dates of divergence within the series suggests the presence of two separate lineages, one for *O. barthii* and one for *O. longistaminata* (Vaughan et al. 2008a).

Americas

O. glumipatula, the only American AA genome species is phylogenetically closer to the African members of the series *Sativae* (Aggarwal et al. 1999; Vaughan and Morishima 2003; Zhu and Ge 2005; Kwon et al. 2006; Takahashi et al. 2008). However, isozyme (Second 1985) and *SPW1* gene (Teranishi et al. 2008) studies suggest its closer relationship to the Asian species. *O. glumipatula* could have evolved from an AA genome ancestor introduced to tropical America (Vaughan et al. 2005). Its taxonomic position as a distinct species is supported by both morphological (Juliano et al. 1998) and genetic data (Zhu and Ge 2005; Duan et al. 2007).

Asia and Pacific

Four series *Sativae* species can be found in Asia and the Pacific. *O. meridionalis* represents the oldest lineage in the series (Second 1985; Wang et al. 1992; Kwon et al. 2006; Park et al. 2003; Zhu and Ge 2005; Duan et al. 2007) that emerged 2 mya

(Zhu and Ge 2005). Data on morphology (Ng et al. 1981a; Ng et al. 1981b), hybridization (Juliano et al. 2005) and DNA (Wang et al. 1992; Martin et al. 1997; Bautista et al. 2001; Park et al. 2003; Juliano et al. 2005; Xu et al. 2005; Kwon et al. 2006) reinforces its status as a distinct species. It is reported to be genetically closer to the African species despite its geographical proximity to the other Asian series *Sativae* members (Duan et al. 2007).

The other three species, *O. rufipogon, O. nivara* and *O. sativa*, exhibit close genetic relationships (Ishii et al. 1996; Aggarwal et al. 1999; Iwamoto et al. 1999; Bautista et al. 2001; Ren et al. 2003; Kwon et al. 2006) and compose the primary gene pool of Asian rice (Vaughan and Morishima 2003).

O. nivara and *O. rufipogon*

O. *nivara* and *O. rufipogon* differ in life cycle, habitat, breeding habit, phenology, and certain morphological traits (Table 2). Morphometric analyses separate the two taxa from each other (Ng et al. 1981a; Uga et al. 2003). Data from AFLPs (Aggarwal et al. 1999), SSRs (Kuroda et al. 2007), SNPs (Xu et al. 2012) and combined sequences from chloroplast, mitochondrial and nuclear DNA (Duan et al. 2007) also produce a clear separation of the two species. In addition, Takahashi et al. (2008) reported that a 69-bp deletion in the ORF100 sequence discriminates between O. *nivara* and *O. rufipogon.*

Table 2. Comparison of *O. nivara* and *O. rufipogon* features based on data from Sharma and Shastry (1965), Duistermaat (1987), Hiroi et al. (1990), Uga et al. (2003) and Vaughan and Morishima (2003).

Feature	*O. nivara*	*O. rufipogon*
Life cycle	Annual	Perennial
Habitat	Seasonally dry areas	Deepwater areas
Breeding habit	Self-pollinating	Predominant cross-pollinating
Photoperiod sensitivity (flowering time under Philippine conditions)	Not photoperiod sensitive (flowers throughout the year)	Photoperiod sensitive (flowers mainly from September to December)
Presence of stolons	Absent	Present

Table 2. (Continued) Comparison of *O. nivara* and *O. rufipogon* features based on data from Sharma and Shastry (1965), Duistermaat (1987), Hiroi et al. (1990), Uga et al. (2003) and Vaughan and Morishima (2003).

Feature	*O. nivara*	*O. rufipogon*
Panicle exsertion	Partially inserted	Exserted
Panicle density	Semi-compact	Open
Panicle branching	Less developed	Developed
Length of panicle axis and branches	Short	Long
Anther length	Shorter than 3.5 mm	Longer than 3.5 mm
Stigma length	Short (1.41 mm)	Long (1.67 mm)
Grain shape	Relatively broad (length to width ratio < 3)	Relatively narrow (length to width ratio > 3)
Seed set	High	Low

However, because of continuous morphological variation linking the two taxa and the lack of a complete reproductive isolation, some rice scientists demote *O. nivara* to an ecotype (Tateoka 1963; Oka 1988; Vaughan et al. 2003) or subspecies (Vaughan and Morishima 2003) of *O. rufipogon*. Pronounced genetic differentiation between *O. nivara* and *O. rufipogon* was not established by studies based on isozymes (Second 1985), RAPDs (Martin et al. 1997; Ren et al. 2003), *Rim2*/*Hipa* - transposon display markers (Kwon et al. 2006), tourist sequences (Iwamoto et al. 1999), SSRs (Ren et al. 2003), MITE-AFLPs (Park et al. 2003), allozymes and RFLPs (Cai et al. 2004), MITE insertions and nuclear genes sequences (Zhu and Ge 2005) and sequences from seven chloroplast and nuclear loci (Zheng and Ge, 2010). In all, the status of the two taxa still remains a debate.

Unfortunately, some studies have not discussed the actual contribution of the annual species to the variation within the Asian AA gene pool because *O. nivara* was either poorly represented in the sampling materials (Wang et al. 1992; Park et al. 2003; Nishikawa et al. 2005; Teranishi et al. 2008) or not included in the study at all (Ishii et al. 1996; Bautista et al. 2001; Londo et al. 2006; Rakshit et al. 2007).

Asian rice domestication

The existence of two types of Asian rice was recognized as early as 100 A.D., during the Han Dynasty in China (Matsuo et al. 1997). These two major groups are presently recognized as `indica' and `japonica', arbitrarily treated as races (Chang 1985), subspecies (Oka 1988; Vaughan and Morishima 2003; Sweeney and McCouch 2007) or varietal groups (Vaughan et al 2008b, 2008c) of *O. sativa*. The recent years have witnessed a burst of renewed interest in the origin of rice as scientists analyze genetic evidence and publish their own interpretation of the domestication process (Kovach et al. 2007; Sang and Ge 2007a, 2007b; Sweeney and McCouch 2007; Tang and Shi 2007; Vaughan et al. 2008b, 2008c; Glemin and Bataillon 2009; Molina et al. 2011).

Two versions of a single origin of rice were prevalent a few decades ago. Chang (1976) proposed that *O. sativa* evolved from *O. nivara,* and that *O. nivara* evolved from *O. rufipogon.* On the other hand, Oka (1974) proposed that both *O. nivara* and *O. sativa* were independently derived from *O. rufipogon*.

However, phylogenetic data revealed further genetic differences between and within indica and japonica, suggesting multiple domestication events, with indica and japonica arising from different populations of *O. rufipogon* (Second 1985; Wang et al. 1992; Bautista et al. 2001; Park et al. 2003; Zhu and Ge 2005; Londo et al. 2006; Kawakami et al. 2007; Kovach et al. 2007; Rakshit et al. 2007; Sweeney and McCouch 2007; Teranishi et al. 2008) or with *O. nivara* as the putative ancestor of indica and *O. rufipogon* of japonica (Cheng et al. 2003; Yamanaka et al. 2003; Ohtsubo et al. 2004; Xu et al. 2007).

Nonetheless, molecular studies based on the *sh4* (Li et al. 2006) and *prog1* (Tan et al. 2008) genes, microsatellites (Gao and Innan 2008), nuclear, mitochondrial and chloroplast DNA sequences (Duan et al. 2007) and single nucleotide polymorphisms (Molina et al. 2011) show marked genetic similarities between indica and japonica, indicating the two varietal groups originate from a single common ancestor. Vaughan et al. (2008b; 2008c) argue that Asian rice was domesticated once, as indica and japonica share the same key domestication allele for non-shattering (of the grain) and that the current genetic diversity of *O. sativa* was a product of gene flow (among different wild, cultivated and intermediate populations) and farmers` selection that succeeded the first domestication event.

The need for wild rice conservation

Oryza series *Sativae* is an invaluable source of genes that can be used to improve cultivated material of *O. sativa.* Some useful traits derived from AA genome species are: grassy stunt virus resistance from *O. nivara*; bacterial blight resistance from *O. longistaminata*; and tungro virus resistance and acid sulfate tolerance from *O. rufipogon* (Brar and Khush 2000). However, the wild species in this series, along with the rest of the genus, are threatened by habitat loss mainly due to various agricultural activities, urbanization and climate change. Local extinction and genetic erosion have been reported in populations of *O. rufipogon* in China (Gao et al. 2000, Song et al. 2005), Thailand (Akimoto et al. 1999) and Irian Jaya (Silitonga, 1999) as well as in populations of *O. barthii* and *O. longistaminata* in eastern and southern Africa (Kiambi et al. 2005). In tropical Northern Australia, *O. meridionalis* grasslands are being displaced by the invasive grass *Urochloa mutica* (Ferdinands et al. 2005). This shows that in order to conserve wild rice diversity effective conservation measures must be implemented.

Rationale of the study

Knowledge on the population genetic structure of wild rice is essential to: 1) design sampling strategies for germplasm collecting; 2) prioritize highly diverse populations for field maintenance and gene bank safekeeping; and 3) improve management practices in both *in situ* and *ex situ* conservation programs. Resolving the taxonomic ambiguities related to *O. nivara* can guide gene banks in correctly classifying and labeling their Asian wild rice accessions and thus get a better insight in which germplasm material they actually own and which is still missing.

As mentioned earlier, *O. nivara* was either barely (Wang et al. 1992; Park et al. 2003; Nishikawa et al. 2005; Teranishi et al. 2008) or not included (Ishii et al. 1996; Bautista et al. 2001; Londo et al. 2006; Rakshit et al. 2007) in genetic studies; hence its contribution to the variation within the Asian AA genome gene pool remains unclear. Incorporating sufficient samples of *O. nivara* in future phylogenetic research can give better insights in its role in the diversification of the genus *Oryza* and probably even in the domestication of *O. sativa.*

Research questions

How do we explain the pattern of morphological and molecular variation within and between *O. meridionalis*, *O. nivara* and *O. rufipogon*?

How do ecology and geography influence the current diversity of *Oryza* series *Sativae* species in Asia and Australia? How do interactions between the sympatric species differ across their geographic range?

Are the distinguishing characters used in practice (Table 2) to separate *O. nivara* and O. *rufipogon* effective enough? In general, there is an obvious need to re-examine the different classification schemes and select or propose a new one that best reflects the actual morphological, genetic and ecogeographical variation in the two taxa.

Aim and contents of the thesis

This study aims to establish inter- and intra-specific variation patterns of *Oryza* series *Sativae* in Asia and the Pacific based on several sources of biosystematic evidence.

This chapter (Chapter 1) provides background taxonomic information about *Oryza* series *Sativae* and introduces the purpose and essence of the thesis.

Chapter 2 looks into the morphological diversity within and among the species of *Oryza* series *Sativae* in Asia and the Pacific and evaluates and refines the existing set of distinguishing features that can clearly delineate the three species morphologically and eco-geographically.

Chapter 3 examines intra- and inter specific genetic diversity produced by SSR genotyping of *O. nivara* and *O. rufipogon* and evaluates the genetic differentiation between them at different spatial scales.

Chapter 4 determines the extent of crossability between and within species of *Oryza* series *Sativae* in Asia and the Pacific in relation to the spatial distance between populations.

Chapter 5 ascertains the taxonomic status of *O. nivara* and explains the variation within and between *O. meridionalis, O. nivara* and *O. rufipogon* based on morphological, genetic and hybridization data.

Chapter 6 discusses the implications of the thesis results in the context of wild rice research and conservation and Asian rice domestication.

References

Aggarwal, R. K., Brar, D. S., Nandi, S., Huang, N. & Khush, G. S. (1999). Phylogenetic relationships among *Oryza* species revealed by AFLP markers. *Theoretical and Applied Genetics* **98,** 1320-1328.

Akimoto, M., Shimamoto, Y. & Morishima, H. (1999). The extinction of genetic resources of Asian wild rice, *Oryza rufipogon* Griff.: a case study in Thailand. *Genetic Resources and Crop Evolution* **46,** 419-425.

Bautista, N. S., Solis, R., Kamijima, O. & Ishii, T. (2001). RAPD, RFLP and SSLP analyses of phylogenetic relationships between cultivated and wild species of rice. *Genes & Genetic Systems* **76,** 71-79.

Brar, D. S. & Khush, G. S. (2002). Transferring genes from wild species into rice. In Kang, M.S. (ed) *Quantitative Genetics, Genomics and Plant Breeding*, pp. 197-217. CAB International, Wallingford (UK).

Cai, H. W., Wang, X. K. & Morishima, H. (2004). Comparison of population genetic structures of common wild rice (*Oryza rufipogon* Griff.), as revealed by analyses of quantitative traits, allozymes, and RFLPs. *Heredity* **92,** 409-417.

Chang, T. T. (1976). The origin, evolution, cultivation, dissemination and diversification of Asian and African rices. *Euphytica* **25,** 425-441.

Chang, T. T. (1985). Crop history and genetic conservation: rice - a case study. *Iowa State Journal of Research* **59,** 425-455.

Cheng, C., Motohashi, R., Tsuchimoto, S., Fukuta, Y., Ohtsubo, H. & Ohtsubo, E. (2003). Polyphyletic origin of cultivated rice: based on the interspersion pattern of SINEs. *Molecular Biology and Evolution* **20,** 67-75.

Chevalier, A. (1932). Nouvelle contribution a l' etude systematique des Oryza. *Revue de Botanique Appliquee et d' Agriculture Tropicale* **12,** 101-1032.

Duan, S., Lu, B., Li, Z., Tong, J., Kong, J., Yao, W., Li, S., Zhu, Y. (2007). Phylogenetic analysis of AA-genome *Oryza* species (Poaceae) based on chloroplast, mitochondrial, and nuclear DNA sequences. *Biochemical Genetics* **45,** 113-129.

Duistermaat, H. (1987). A revision of *Oryza* (Gramineae) in Malesia and Australia. *Blumea* **32,** 157-193.

Farooq, S., Iqbal, N., Shah, T. M. & Asghar, M. (1996). Problems and prospects for utilizing *Porteresia coarctata* in rice breeding programs. *Cereal Research Communications* **24,** 41-47.

Ferdinands, K., Beggs, K. & Whitehead, P. (2005). Biodiversity and invasive grass species: multiple-use or monoculture? *Wildlife Research* **32,** 447-457.

Gao, L., Chen, W., Jiang, W., Ge, S., Hong, D. & Wang, X. (2000). Genetic erosion in Northern marginal population of the common wild rice *Oryza rufipogon* Griff. and its conservation, revealed by the change of population genetic structure. *Hereditas* **133,** 47-53.

Gao, L. & Innan, H. (2008). Nonindependent domestication of the two rice subspecies, *Oryza sativa* ssp. *indica* and ssp. *japonica*, demonstrated by multilocus microsatellites. *Genetics* **179,** 965-976.

Ge, S., Li, A., Lu, B.-R., Zhang, S.-Z. & Hong, D.-Y. (2002). A phylogeny of the rice tribe Oryzeae (Poaceae) based on matK sequence data. *American Journal of Botany* **89,** 1967-1972.

Ghose, R. L. M., Ghatge, M. B. & Subramanyan, V. (1965). *Rice in India.* Indian Council of Agricultural Research, New Delhi.

Glemin, S. & Bataillon, T. (2009). A comparative view of the evolution of grasses under domestication. *New Phytologist* **183,** 273-290.

Guo, Y.-L. & Ge, S. (2005). Molecular phylogeny of Oryzeae (Poaceae) based on DNA sequences from chloroplast, mitochondrial, and nuclear genomes. *American Journal of Botany* **92,** 1548-1558.

Hiroi, K., Al Mamun, A., Wada, T. & Takeoka, Y. (1990). A study on the interspecific variation of panicle structure in the genus *Oryza*. *Japan Journal of Crop Science* **59,** 815 - 823.

Ishii, T., Nakano, T., Maeda, H. & Kamijima, O. (1996). Phylogenetic relationships in A-genome species of rice as revealed by RAPD analysis. *Genes & Genetic*

Systems **71,** 195-210.

Iwamoto, M., Nagashima, H., Nagamine, T., Higo, H. & Higo, K. (1999). A Tourist element in the flanking region of the catalase gene CatA reveals evolutionary relationships among Oryza species with various genome types. *Molecular Genetics and Genomics* **262,** 493-500.

Joshi, S. P., Gupta, V. S., Aggarwal, R. K., Rakenjar, P. K. & Brar, D. S. (2000). Genetic diversity and phylogenetic relationship as revealed by inter simple sequence repeat (ISSR) polymorphism in the genus *Oryza*. *Theoretical and Applied Genetics* **100,** 1311-1320.

Juliano, A. B., Naredo, M. E. B. & Jackson, M. T. (1998). Taxonomic status of *Oryza glumaepatula* Steud. I. Comparative morphological studies of New World diploids and Asian AA genome species. *Genetic Resources and Crop Evolution* **45,** 197-203.

Juliano, A. B., Naredo, M. E. B., Lu, B.-R. & Jackson, M. T. (2005). Genetic differentiation in *Oryza meridionalis* Ng based on molecular and crossability analyses. *Genetic Resources and Crop Evolution* **52,** 435-445.

Kawakami, S., Ebana, K., Nishikawa, T., Sato, Y., Vaughan, D. A. & Kadowaki, K. (2007). Genetic variation in the chloroplast genome suggests multiple domestication of cultivated Asian rice (*Oryza sativa* L.). *Genome* **50,** 180-187.

Khush, G. S. (1997). Origin, dispersal, cultivation and variation of rice. *Plant Molecular Biology* **35,** 25-34.

Kiambi, D. K., Ford-Lloyd, B. V., Jackson, M. T., Guarino, L., N., M. & Newbury, H. J. (2005). Collection of wild rice (*Oryza* L.) in east and southern Africa in response to genetic erosion. *PGR Newsletter* **142,** 10-20.

Kovach, M. J., Sweeney, M. T. & McCouch, S. R. (2007). New insights into the history of rice domestication. *Trends in Genetics* **23,** 578-587.

Kuroda, Y., Sato, Y.-I., Bounphanousay, C., Kono, Y. & Tanaka, K. (2007). Genetic structure of three Oryza AA genome species (O. rufipogon, O. nivara and O. sativa) as assessed by SSR analysis on the Vientiane Plain of Laos. *Conservation Genetics* **8,** 149-158.

Kwon, S.-J., Lee, J. K., Hong, S.-W., Park, Y.-J., McNally, K. L. & Kim, N.-S. (2006). Genetic diversity and phylogenetic relationship in AA Oryza species as revealed by Rim2/Hipa CACTA transposon display. *Genes & Genetic Systems* **81,** 93-101.

Li, C., Zhou, A. & Sang, T. (2006). Rice domestication by reducing shattering. *Science* **311,** 1936-1939.

Linnaeus, C. (1753). *Species Plantarum*. Stockholm.

Londo, J. P., Chiang, Y. C., Hung, K. H., Chiang, T. Y. & Schaal, B. A. (2006). Phylogeography of Asian wild rice, *Oryza rufipogon*, reveals multiple independent domestications of cultivated rice, *Oryza sativa*. *Proceedings of the National Academy of Sciences of the United States of America* **103,** 9578-9583.

Lu, B.-R. (1999). Taxonomy of the genus *Oryza* (Poaceae): historical perspective and current status. *International Rice Research Notes* **24,** 4-8.

Lu, B.-R.Ge, S. (2003). *Oryza coarctata*: the name that best reflects the relationships of *Porteresia coarctata* (Poaceae: Oryzeae). *Nordic Journal of Botany* **23,** 555-558.

Lu, B.-R., Ge, S., Sang, T., Chen, J.-K. & Hong, D.-Y. (2001). The current taxonomy and perplexity of the genus *Oryza* (Poaceae). *Acta Phytotaxonomica Sinica* **39,** 373-388.

Martin, C., Juliano, A., Newbury, H. J., Lu, B. R., Jackson, M. T. & Ford-Lloyd, B. V. (1997). The use of RAPD markers to facilitate the identification of *Oryza* species within a germplasm collection. *Genetic Resources and Crop Evolution* **44,** 175-183.

Matsuo, T., Futsuhara, Y. & Yamaguchi, H. (1997). *Science of the rice plant*. Food and Agriculture Policy Research Center, Tokyo.

Molina, J., Sikora, M., Garud, N., Flowers, J. M., Rubinstein, S., Reynolds, A., Huang, P., Jackson, S., Schaal, B. A., Bustamante, C. D., Boyko, A. R., & Purugganan, M. D. (2011). Molecular evidence for a single evolutionary origin of domesticated rice. *Proceedings of the National Academy of Sciences of*

the United States of America **108,** 8351-8356.

Ng, N. Q., Chang, T. T., Williams, J. T. & Hawkes, J. G. (1981a). Morphological studies of Asian rice and its related wild species and the recognition of a new Australian taxon. *Biological Journal of the Linnean Society* **16,** 303-313.

Ng, N. Q., Hawkes, J. G., Williams, J. T. & Chang, T. T. (1981b). The recognition of a new species of rice (Oryza) from Australia. *Botanical Journal of the Linnean Society* **82,** 327-330.

Nishikawa, T., Vaughan, D. A. & Kadowaki, K.-I. (2005). Phylogenetic analysis of *Oryza* species, based on simple sequence repeats and their flanking nucleotide sequences from the mitochondrial and chloroplast genomes. *Theoretical and Applied Genetics* **110,** 696-705.

Ohtsubo, H., Cheng, C., Ohsawa, I., Tsuchimoto, S. & Ohtsubo, E. (2004). Rice retroposon p-SINE1 and origin of cultivated rice. *Breeding Science* **54,** 1-11.

Oka, H. I. (1974). Experimental studies on the origin of cultivated rice. *Genetics* **78,** 475-486.

Oka, H. I. (1988). *Origin of cultivated rice.* Japan Science Society Press, Tokyo.

Park, K. C., Kim, N. H., Cho, Y. S., Kang, K. H., Lee, J. K. & Kim, N. S. (2003). Genetic variations of AA genome *Oryza* species measured by MITE-AFLP. *TAG Theoretical and Applied Genetics* **107,** 203-209.

Rakshit, S., Rakshit, A., Matsumura, H., Takahashi, Y., Hasegawa, Y., Ito, A., Ishii, T., Miyashita, N., & Terauchi, R. (2007). Large-scale DNA polymorphism study of *Oryza sativa* and *O. rufipogon* reveals the origin and divergence of Asian rice. *TAG Theoretical and Applied Genetics* **114,** 731-743.

Ren, F., Lu, B.-R., Li, S., Huang, J. & Zhu, Y. (2003). A comparative study of genetic relationships among the AA-genome Oryza species using RAPD and SSR markers. *TAG Theoretical and Applied Genetics* **108,** 113-120.

Roschevics, R. (1931). A contribution to the study of rice. *Trudy po Prikladnoj Botanike Genetike i Selekcii.* **27,** 3-133.

Sang, T., & Ge, S. (2007a). Genetics and phylogenetics of rice domestication. *Current Opinion in Genetics & Development* **17,** 533-538.

Sang, T. & Ge, S. (2007b). The puzzle of rice domestication. *Journal of Integrative Plant Biology* **49,** 760-768.

Sarker, R. H., Samad, M. A., Seraj, Z. I., Hoque, M. I. & Islam, A. S. (1993). Pollen tube growth in crosses between *Porteresia coarctata* and *Oryza sativa*. *Euphytica* **69,** 129-134.

Second, G. (1985). Evolutionary relationships in the Sativa group of *Oryza* based on isozyme data. *Genetics Selection Evolution* **17,** 89-114.

Sharma, S. D. & Shastry, S. V. S. (1965). Taxonomic studies in the genus *Oryza* L. III. *O. rufipogon* Griff. sensu stricto and *O. nivara* Sharma et Shastry *nom. nov*. *Indian Journal of Genetics and Plant Breeding* **25,** 157-167.

Sharma, S. D. & Shastry, S. V. S. (1972). Evolution in genus *Oryza,* in *Advancing Frontiers in Cytogenetics – Proceedings of the National Seminar,* 1972 New Delhi, India, p. 5 - 20.

Silitonga, T. S. (1999). *Exploration and collection for wild rice species in Irian Jaya, Indonesia (C. I. Cooperative, Ed.)*. Research Institute for Food Crop Biotechnology and Central Research Institute for Food Crop, Indonesia.

Song, Z., Li, B., Chen, J. & Lu, B.-R. (2005). Genetic diversity and conservation of common wild rice (*Oryza rufipogon*) in China. *Plant Species Biology* **20,** 83-92.

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

Takahashi, H., Sato, Y.-I. & Nakamura, I. (2008). Evolutionary analysis of two plastid DNA sequences in cultivated and wild species of *Oryza*. *Breeding Science* **58,** 225-233.

Tan, L., Li, X., Liu, F., Sun, X., Li, C., Zhu, Z., Fu, Y., Cai, H., Wang, X., Xie, D. & Sun, C. (2008). Control of a key transition from prostrate to erect growth in rice domestication. *Nature Genetics* **40,** 1360-1364.

Tang, T. & Shi, S. (2007). Molecular population genetics of rice domestication. *Journal of Integrative Plant Biology* **49,** 769-775.

Tateoka, T. (1962). Taxonomic studies of *Oryza* II. Several species complex. *Botanical Magazine Tokyo* **75,** 418-427

Tateoka, T. (1963). Taxonomic studies of *Oryza*. III. Key to the species and their enumeration. *Botanical Magazine Tokyo* **76,** 165-173.

Tateoka, T. (1964). Notes on some grasses. XVI. Embryo structure of the genus *Oryza* in relation to their systematics. *American Journal of Botany* **51,** 539-543.

Teranishi, C., Yoshida, K. & Miyashita, N. T. (2008). DNA polymorphism in the SUPERWOMAN1 (SPW1) locus of the wild rice *Oryza rufipogon* and its related species. *Genes & Genetic Systems* **83,** 403-415.

Uga, Y., Fukuta, Y., Ohsawa, R. & Fujimura, T. (2003). Variations of floral traits in Asian cultivated rice (*Oryza sativa* L.) and its wild relatives (*O. rufipogon* Griff.). *Breeding Science* **53,** 345-352.

USDA, ARS, National Genetic Resources Program. (2012). Germplasm Resources Information Network - (GRIN) [Online Database]. National Germplasm Resources Laboratory, Beltsville, Maryland. URL: http://www.ars-grin.gov/cgi-bin/npgs/html/index.pl (12 April 2012)

Vaughan, D. A. (1989). The genus *Oryza* L. Current status of taxonomy. *IRRI Research Paper Series* **138**, 1-21.

Vaughan, D. A., Kadowaki, K.-I., Kaga, A. & Tomooka, N. (2005). On the phylogeny and biogeography of the genus *Oryza*. *Breeding Science* **55,** 113-122.

Vaughan, D. A. & Morishima, H. (2003). Biosystematics of the genus *Oryza.* In Wayne Smith, Dilday, R. H. (eds) *Rice: Origin, History, Technology and Production*, pp. 27 - 65. John Wiley and Sons, New Jersey.

Vaughan, D. A., Morishima, H. & Kadowaki, K. (2003). Diversity in the Oryza genus. *Current Opinion in Plant Biology* **6,** 139-146.

Vaughan, D. A., Song, G., Kaga, A. & Tomooka, N. (2008a). Phylogeny and Biogeography of the Genus *Oryza*. In Hirano, H. Y., Hirai, A., Sano, Y. & Sasaki, T. (eds) *Rice Biology in the Genomics Era*, pp. 219 - 234. Springer, Berlin.

Vaughan, D. A., Lu, B.-R. & Tomooka, N. (2008b). The evolving story of rice evolution. *Plant Science* **174,** 394-408.

Vaughan, D. A., Lu, B.-R. & Tomooka, N. (2008c). Was Asian rice (*Oryza sativa*) domesticated more than once? *Rice* **1,** 16-24.

Wang, Z. Y., Second, G. & Tanksley, S. D. (1992). Polymorphism and phylogenetic relationships among species in the genus *Oryza* as determined by analysis of nuclear RFLPs. *TAG Theoretical and Applied Genetics* **83,** 565-581.

Xu, J.-H., Cheng, C., Tsuchimoto, S., Ohtsubo, H. & Ohtsubo, E. (2007). Phylogenetic analysis of *Oryza rufipogon* strains and their relations to *Oryza sativa* strains by insertion polymorphism of rice SINEs. *Genes & Genetic Systems* **82,** 217-229.

Xu, J.-H., Kurata, N., Akimoto, M., Ohtsubo, H. & Ohtsubo, E. (2005). Identification and characterization of Australian wild rice strains of *Oryza meridionalis* and *Oryza rufipogon* by SINE insertion polymorphism. *Genes & Genetic Systems* **80,** 129-134.

Xu, X., Liu, X., Ge, S., Jensen, J. D., Hu, F., Li, X., Dong Y., Gutenkunst R. N., Fang, L., Huang, L. Li, J., He, W., Zhang, G., Zheng, X., Zhang, F., Li, Y., Yu, C., Kristiansen, K., Zhang, X., Wang, J., Wright, M., McCouch, S., Nielsen, R., Wang J., Wang, W. (2012). Resequencing 50 accessions of cultivated and wild rice yields markers for identifying agronomically important genes. *Nature Biotechnology* **30,** 105–111.

Yamanaka, S., Nakamura, I., Nakai, H. & Sato, Y.-I. (2003). Dual origin of the cultivated rice based on molecular markers of newly collected annual and perennial strains of wild rice species, *Oryza nivara* and *O. rufipogon*. *Genetic Resources and Crop Evolution* **50,** 529-538.

Zheng, X. & Ge, S. (2010). Ecological divergence in the presence of gene flow in two closely related *Oryza* species (*Oryza rufipogon* and *O. nivara*) *Molecular Ecology* **19,** 2439-2454.

Zhu, Q. & Ge, S. (2005). Phylogenetic relationships among A-genome species of the genus *Oryza* revealed by intron sequences of four nuclear genes. *New Phytologist* **167,** 249-265.

CHAPTER 2
Morphological variation patterns between and within *Oryza nivara* and *O. rufipogon*

Maria Celeste N. Banaticla-Hilario[1,2], Marc S.M. Sosef[2,3], Kenneth L. McNally[1], Nigel Ruaraidh Sackville Hamilton[1] and Ronald G. van den Berg[2]

[1]*T.T. Chang Genetic Resources Center, International Rice Research Institute, Los Baños, Laguna, Philippines*
[2]*Biosystematics Group, Wageningen University and Research Center, Wageningen, The Netherlands*
[3]*Naturalis Biodiversity Center, section NHN, Wageningen University, Wageningen, The Netherlands.*

Abstract

To search for variation patterns and diagnostic features, several numerical methods were applied to phenotypic data obtained from 121 accessions representing sympatric populations of *O. nivara* and *O. rufipogon* from tropical continental Asia and of *O. meridionalis* and *O. rufipogon* from Australasia, as well as *O. rufipogon* populations from maritime Southeast Asia.

Ordination and cluster analyses separate *O. rufipogon* from *O. nivara* indicating two sympatric morphological species occupying different ecological niches. Principal component analysis and hierarchical clustering place *O. nivara* and *O. meridionalis* in one cluster while k-means clustering separate *O. meridionalis* as a distinct group at k = 5. *O. meridionalis* and *O. nivara* appear to be two distinct, geographically isolated species occupying similar habitats, consequently evolving similar phenotypes.

O. nivara and *O. rufipogon* are morphologically more differentiated in South Asia than in mainland Southeast Asia implying more active gene flow among sympatric populations in the latter region.

Spatial and morphological distances were positively correlated within *O. nivara* and within *O. rufipogon* in continental Asia but not within *O. rufipogon* in Australasia and maritime Southeast Asia.

O. nivara exhibits South and Southeast Asian phenotypes. *O. rufipogon* does not display such regional differentiation although k-means analysis recognizes the partial separation of continental Asian accessions from the Australasians and maritime Southeast Asians ones at k = 5. The Australasian populations are distinct from the rest of *O. rufipogon*.

Seedling height, culm number and diameter, leaf length and width, and anther length were significantly correlated to certain geographic and climatic factors and displayed opposing correlation directions for *O. nivara* and *O. rufipogon*, implying that the two species respond differently to geographic and climatic gradients.

Anther length can readily distinguish *O. rufipogon* from the annual species and spikelet width and awn length can discriminate *O. nivara* and *O. meridionalis* from each other. Additionally, culm habit and length, leaf length, flag leaf width and attitude, panicle branching type, spikelet fertility, awn length and thickness,

spikelet width, sterile lemma width and anther length to spikelet length ratio are significantly different among the three species. Botanical descriptions are developed to delineate the three species morphologically.

Introduction

In the International Rice Genebank (IRG) at the International Rice Research Institute (IRRI), *O. nivara* Sharma & Shastry and *O. rufipogon* Griff. make up the majority of the wild rice collection. *O. nivara* is an annual self-pollinator found in seasonally dry areas of tropical continental Asia while *O. rufipogon* is a perennial cross-pollinator found in permanently wet areas (e.g. ponds and swamps) of the Asia-Pacific region (South China, South and Southeast Asia to Northern Australia). Under Philippine conditions, *O. rufipogon* flowers mainly from September to December while *O. nivara* produces inflorescences throughout the year. Morphologically distinct accessions are common but mislabelling still occurs due to the close physical resemblance of the seeds of the two species. Plant descriptions are often incomplete and the morphological distinction between these species is often not clear from the literature. A number of accessions also appear to represent intermediate forms of *O. nivara* and *O. rufipogon.*

Although these two taxa are generally recognized as different species (Sharma and Shastry 1965; Ng et al. 1981a; Ng et al. 1981b; Lu 1999; Lu et al. 2001; Aggarwal et al. 1999; Kuroda et al. 2007; Duan et al. 2007; Xu et al. 2012), some rice scientists consider *O. nivara* to be the annual ecotype or subspecies of *O. rufipogon* (with *O. rufipogon sensu stricto* as the perennial ecotype) due to their continuous (morphological and genetic) variation and inter-fertility (Tateoka 1963; Oka 1988; Hiroi et al. 1990; Morishima et al. 1992; Vaughan et al. 2003; Second 1985; Iwamoto et al, 1999; Ren et al. 2003; Park et al. 2003; Zhu and Ge 2005; Kwon et al. 2006; Zhou et al. 2008; Zheng and Ge, 2010). In this study, *O. nivara* and *O. rufipogon* are provisionally treated as separate species.

A clear perception of the morphological variation between and within *O. nivara* and *O. rufipogon* and well defined phenotype-based species delineation can help genebanks in dealing with misidentified accessions and in managing intermediate forms.

O. meridionalis Ng is another closely related annual species found in seasonally dry areas of tropical Australia and western New Guinea (Irian Jaya). Its geographic distribution lies within the range of *O. rufipogon*. However, unlike *O. nivara*, it is more widely recognized as a distinct species which is supported by morphological (Ng et al. 1981a; Ng et al. 1981b), hybridization (Juliano et al. 2005) and molecular (Wang et al. 1992; Bautista et al. 2001; Ohtsubo et al. 2004; Juliano et al. 2005; Xu et al. 2005; Duan et al. 2007) data. On account of its sympatry with *O. rufipogon* and its recognition as a distinct taxonomic species, *O. meridionalis* was included in this study to serve as reference in determining whether the differences between *O. nivara* and *O. rufipogon* is large enough to merit a species status for the former.

This chapter re-examines the morphology of the three species across their entire geographic range. Geographically overlapping populations of *O. nivara* and *O. rufipogon* and of *O. meridionalis* and *O. rufipogon* were used to determine whether sympatric populations are consistently morphologically different throughout their distribution. Geographical variation patterns within species were also examined.

The objectives were to: 1) evaluate the morphological differences between *O. nivara*, *O. rufipogon*, and *O. meridionalis*; 2) determine geographical patterns of morphological variation within and between the three species; 3) provide a set of distinguishing characters that can clearly delineate the three species morphologically.

Materials and methods

Selection and georeferencing of plant material

IRG-IRRI provided the 131 populations used in this study (Appendix 1). The collection locality data of the accessions were obtained from IRG collection reports and from the International Rice Genebank Collection Information System (IRGCIS) (http://www.irgcis.irri.org:81/grc/Irgcishome.html) and these were georeferenced using the following software:

- Biogeomancer workbench (http://www.biogeomancer.org/workbench.html)
- Biogeomancer spatial attribute lookup (http://bg.berkeley.edu/sal/)
- U.S. National Imagery and Mapping Agency database (http://earth-info.nga.mil/gns/html/index.html

- World Gazetteer (http://world-gazetteer.com/)

Location data were also validated using DIVA-GIS (Hijmans 2001) and Google earth (http://earth.google.com/).

Selected accessions represent (Figure 1):

1) 52 sympatric populations of *O. nivara* and *O. rufipogon* from tropical continental/mainland Asia (South China, Vietnam, Laos, Cambodia, Thailand, Myanmar, Bangladesh, Nepal, India and Sri Lanka) (104 accessions);
2) 5 sympatric populations of *O. meridionalis* and *O. rufipogon* from Australasia (Irian Jaya and Northern Australia) (10 accessions); and
3) 9 populations of *O. rufipogon* from insular/maritime Southeast Asia (Malaysia, the Philippines and Indonesia) where *O. nivara* and *O. meridionalis* reportedly do not exist (9 accessions).

Six *O. sativa* L. and two *O. officinalis* Wall. ex G.Watt accessions were also included for comparison (Appendix 1).

Altitude, mean annual temperature and mean annual precipitation data on each geographic location were extracted from the 5-arc minute grids downloaded from the WorldClim website (http://worldclim.org/) and obtained using DIVA-GIS (Hijmans 2001).

Phenotyping

Five individuals from each accession were planted in the Genetic Resources Center screenhouse at IRRI, Philippines following a randomized complete block design (Figure 2). Unfortunately, the seeds of two *O. nivara* accessions from China (N17) and Myanmar (N30) failed to germinate. Their corresponding sympatric *O. rufipogon* accessions (R17 and R30) were still included in the analysis (Appendix 1).

Phenotyping was conducted using 8 qualitative and 22 quantitative characters (Appendix 2) suggested and described in the list of descriptors for wild and cultivated rice (*Oryza* spp.) published by Bioversity International, IRRI and Africa Rice Center (2007). Four additional quantitative characters (stigma length, style length, anther length to spikelet length ratio and spikelet width to spikelet length ratio) were also obtained (Appendix 2). Herbarium voucher specimens were collected from each individual plant and deposited at the IRG Herbarium.

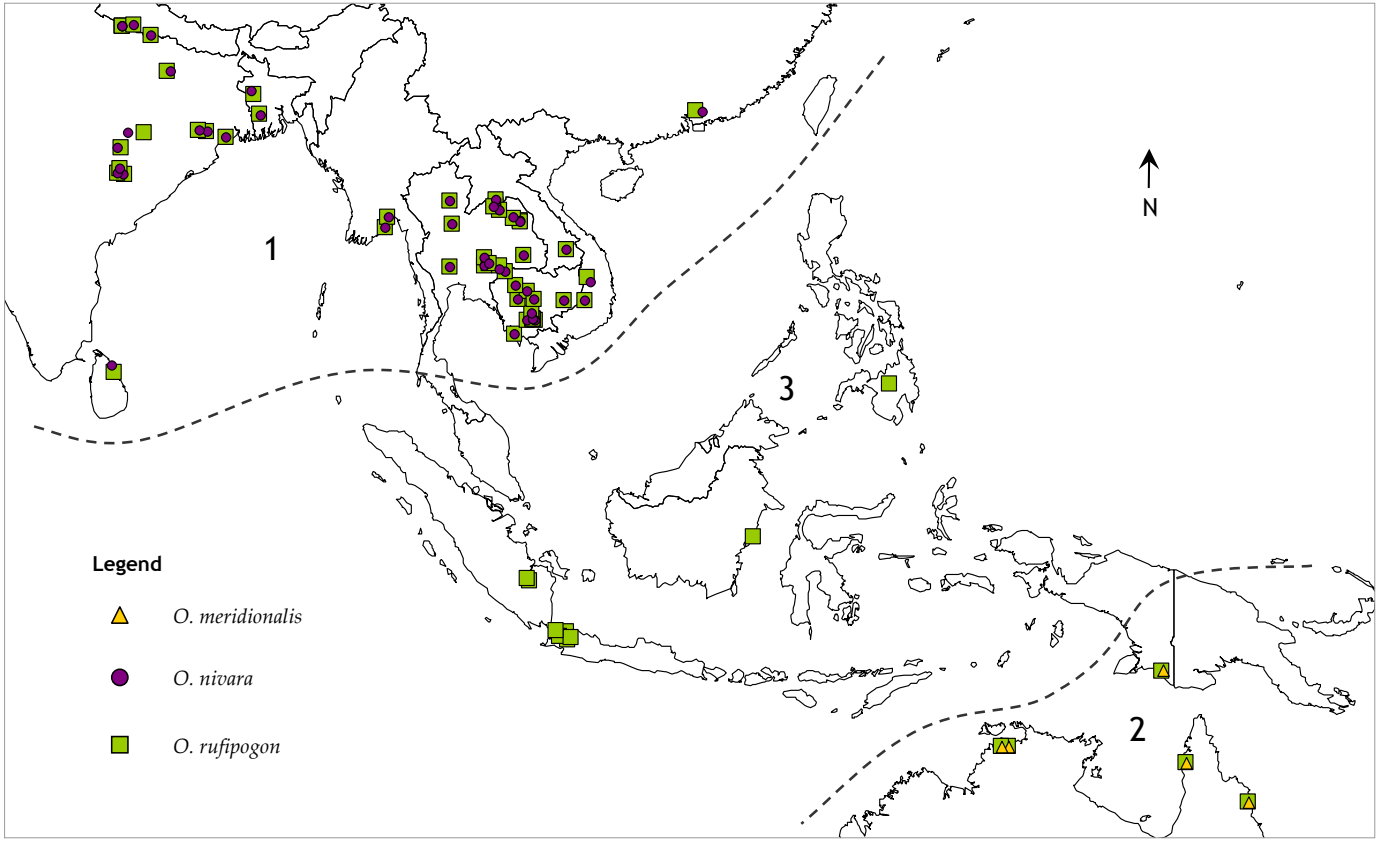


Figure 1. Geographic distribution of the studied populations. Subset 1 comprises sympatric accessions of *O. nivara* and *O. rufipogon* in South and continental Southeast Asia. Subset 2 is composed of sympatric accessions of *O. meridionalis* and *O. rufipogon* in Irian Jaya and Northern Australia while subset 3 contains *O. rufipogon* accessions that do not overlap with both *O. nivara* and *O. meridionalis.*

A

Transplanting period. In the foreground are a few pots already planted with wild rice seedlings. Photo taken on May 13, 2008.

B

Phenotyping of vegetative characters. The plants are in their late vegetative stage (48 days after transplanting). Photo taken on June 30, 2008.

C

Documenting life cycle. The plants are ratooned after maturity and those that recovered (i.e., produced new culms) are considered as perennials. Photo taken on December 21, 2008

Figure 2. The screenhouse set up at different stages of the experiment.

Data validation

Before data analysis, phenotype data and herbarium specimens were cross-referenced to detect labelling errors. In case of mixtures within accessions, individuals that did not belong to the population (e.g., an *O. sativa* or *O. nivara* plant in an *O. rufipogon* accession) were removed from the analyses while intermediate phenotypes were retained. Populations N26, R5 and R29 (Appendix 1) were recoded as subpopulations N26A and N26B, R5A and R5B, and R29A and R29B, respectively as they exhibited two distinct plant types in which one appeared to be an intermediate form.

Data analyses

All analyses were conducted using the free software R 2.10 (R Development Core Team 2009) using the base and additional packages.

Pairwise character correlation analysis was performed by the cor() function using Pearson's product-moment coefficient on interval and ratio-type data and Spearman correlation coefficient on ordinal data.

Re-scaling of quantitative characters and principal component analysis (PCA) were performed by the prcomp() function.

A hierarchical cluster analysis (HCA) of qualitative and quantitative characters was conducted with the cluster package (Maechler et al. 2005). The daisy() function was applied in constructing a dissimilarity matrix based on Gower's coefficient while the hclust() function performed agglomerative hierarchical clustering using the unweighted pair group method with arithmetic mean (UPGMA). Bootstrapping with 10,000 replicates was conducted by the consensus() function of the agricolae package (Mendinburu 2010).

K-means clustering was performed by the kmeans() function of the base package. To determine the appropriate number of clusters, the within group sum of squared error (SSE) of each cluster (from k=1 to k=15) were plotted and screened for the cluster solution where the decrease in SSE decelerated markedly (indicated by a bend in the plot). Principal components (PCs) are said to be the continuous solution of the K-means clustering membership indicators (Ding and He 2004). Cluster solutions were plotted against the first two PCs using the clusplot() function of the cluster package (Maechler et al. 2005).

To determine the correlation between geographic distance and Euclidean distance matrices, a Mantel test with 10000 permutations was conducted using the packages fields (Furrer et al. 2010) and vegan (Oksanen et al. 2010).

The cor.test() function was used to detect the correlation of latitude, longitude, altitude, annual mean temperature and annual precipitation with the Euclidean distances of sympatric *O. nivara* and *O. rufipogon* populations and also with quantitative measurements within species. There were not enough samples to analyze sympatric *O. meridionalis* and *O. rufipogon* populations.

To establish whether Euclidean distance of sympatric populations significantly differ between geographic regions, an analysis of variance (ANOVA) and Tukey's HSD test were performed using the agricolae package (Mendiburu 2010). The geographic regions compared are South Asia (Bangladesh, India, Nepal and Sri Lanka) and (continental) Southeast Asia (Cambodia, Laos, Myanmar, Thailand and Vietnam).

A multivariate analysis of variance (MANOVA) was conducted to test the overall effect of the characters on inter- and intra-specific differences, followed by a univariate ANOVA and Tukey's HSD test using the agricolae package (Mendiburu 2010) to identify specific characters that are significantly different between species and between geographical populations of each species. Box and whisker plots of the characters were constructed using the boxplot() function. The groups compared were *O. meridionalis* and the geographical populations of *O. nivara* in South Asia and (mainland) Southeast Asia, and of *O. rufipogon* in South Asia, continental Southeast Asia, insular Southeast Asia and Australasia.

Results

Character correlation

To lessen data redundancy, one character was eliminated from each strongly correlated ($|\text{correlation coefficient}| \geq 0.6$) and functionally or structurally linked pair except in cases where both characters are known to be highly discriminating for the studied taxa. Eight quantitative characters (out of 26) were excluded from the PCA, HCA and K-means analysis and one qualitative character (out of 8) was removed from the HCA (Table 1).

Table 1. Highly correlated characters and their correlation coefficient (denoted by r).

Characters		r
Anther length to spikelet length ratio*	Anther length	0.98
	Spikelet width	-0.79
	Spikelet width to spikelet length ratio*	-0.74
	Spikelet fertility*	-0.71
	Flag leaf length*	-0.69
	Awn thickness*	-0.65
	Number of days from seeding to first heading	0.61
	Leaf length	-0.60
	Leaf length	0.85
Flag leaf length*	Anther length	-0.66
	Spikelet width	-0.66
	Spikelet fertility*	0.62
Spikelet width to spikelet length ratio*	Spikelet width	0.89
	Anther length	-0.77
	Anther length	-0.68
Spikelet fertility*	Spikelet width	0.66
	Awn length	0.78
Awn thickness*	Anther length	-0.64
Ligule length*	Leaf length	0.63
	Panicle length	0.63
Anther length	Spikelet width	-0.75
	Stigma length	0.63
Flag leaf width*	Leaf width	0.93
Panicle number*	Culm number	0.80
Leaf length	Panicle length	0.64
Culm length	Number of days from seeding to first heading	0.77
Rhizome and stolon formation*	Life cycle	0.73

*excluded from principal component, k-means and hierarchical cluster analyses

Principal component analysis

The shape of the scree plot (Figure 3A), with a steep decline in % variance over the first 3 axes, indicates deep structure in two dimensions. This in turn implies three main groups, with a more diffuse multi-dimensional structure of variation within groups. The first two principal components accounted for 43.8% of the total variance. The factor loadings and proportion of variance of PC1 to PC4 are presented in Table 2.

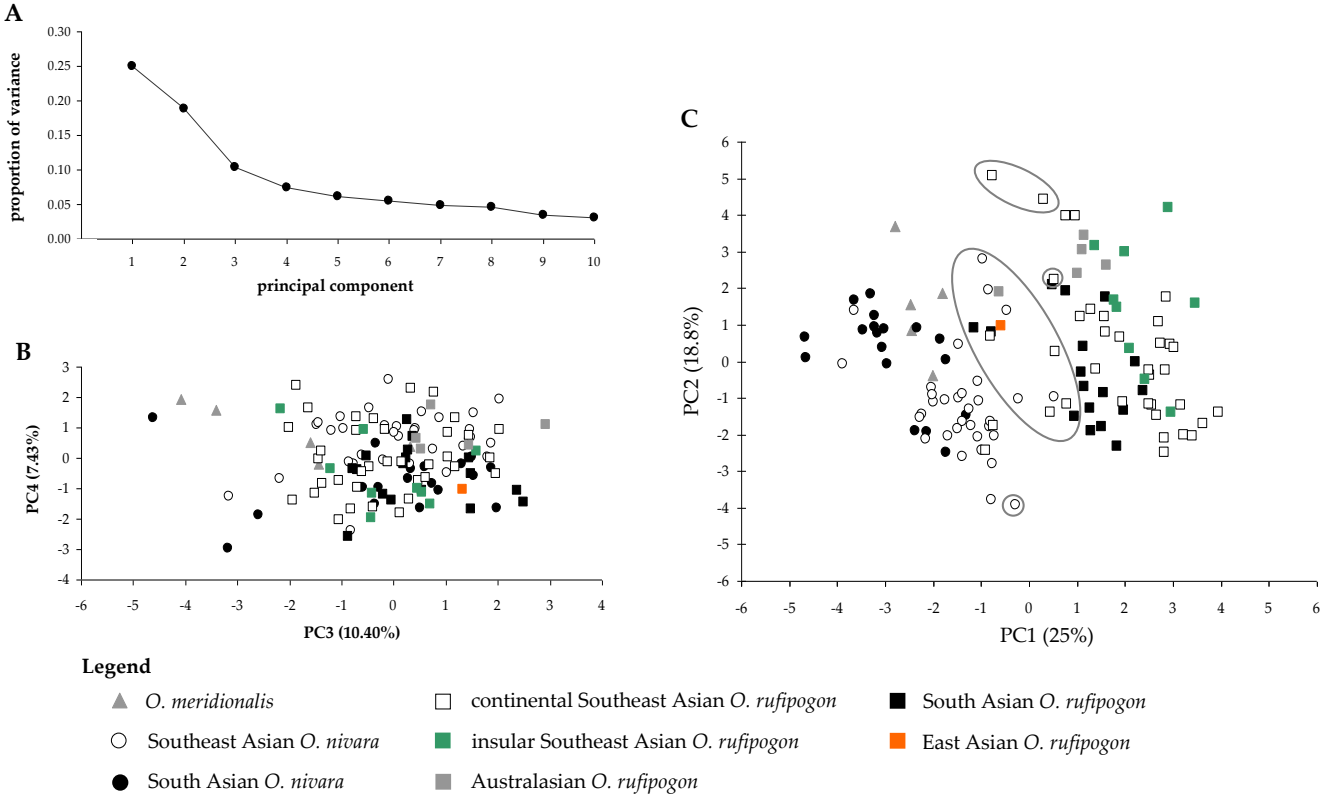


Figure 3. PCA results obtained from 18 quantitative characters. **A,** screeplot; **B,** plot of the scores of the third and fourth principal components; **C,** plot of the scores of the first two principal components. The encircled populations in **C** were tentatively identified as intermediate forms (i.e., intermediate between *O. nivara* and *O. rufipogon, O. nivara* and *O. sativa* or *O. rufipogon* and *O. sativa*) based on screenhouse and herbarium observations.

Table 2. Factor loadings and proportion of variance of the first four principal components. Components providing factor loadings above 0.25 are highlighted.

Characters	Principal component			
	1	2	3	4
seedling height	-0.1967	0.1895	0.1936	-0.0405
culm number	0.0291	-0.3818	0.0679	0.3519
culm diameter	-0.1196	0.3835	-0.3510	-0.1543
culm length	0.2845	0.2973	-0.1364	-0.0013
leaf length	-0.3090	0.2424	-0.1807	0.2949
leaf width	-0.2107	0.2253	-0.1963	-0.3520
number of basal primary branches of the panicle	0.0461	0.1599	-0.3880	0.1231
panicle length	-0.1327	0.3745	-0.0464	0.4217
distance from panicle base to lowest spikelet insertion	0.2347	0.3006	0.0959	0.2752
awn length	-0.2193	-0.0606	0.0164	0.5372
spikelet length	-0.1695	0.2479	0.3493	-0.0600
spikelet width	-0.4144	-0.0016	0.0771	-0.0973
sterile lemma length	-0.1659	0.1037	0.2956	-0.2069
sterile lemma width	-0.2115	0.0630	0.2574	-0.0132
stigma length	0.2228	0.2204	0.3424	0.1444
style length	0.0351	0.2209	0.3807	0.0002
anther length	0.3979	0.1452	0.1706	-0.0730
number of days from seeding to first heading	0.3472	0.1376	-0.0991	-0.0191
Standard deviation	2.1200	1.8380	1.3680	1.1563
Proportion of variance	0.2500	0.1880	0.1040	0.0743
Cumulative proportion	0.2500	0.4370	0.5410	0.6156

PC1 explains 25% of the total variance and separates *O. nivara* and *O. meridionalis* from *O. rufipogon* as shown in the scatter plot (Figure 3C). The eighteen accessions (including R5A and R29A) that fall between the two main clusters are considered as intermediate forms as confirmed by preliminary field and herbarium observations. The two *O. rufipogon* accessions (R27 and R45) included in the *O. meridionalis - O. nivara* cluster were also confirmed to represent misidentified *O. nivara* populations (Figure 3C). These mislabeled and intermediate accessions were not included in the succeeding analyses. PC1 also divides *O. nivara* in two sub-groups, with a predominantly Southeast Asian cluster and a South Asian cluster.

Characters with the highest factor loadings on PC1 are spikelet width, anther length, leaf length and number of days from seeding to first heading (Table 2).

PC2 explains 18.8% of the variation and does not show any clear separation between the annual species (*O. meridionalis* and *O. nivara*) and *O. rufipogon*. However it produces geographical clustering patterns within species (Figure 3C). The majority of South Asian *O. nivara* accessions are separated from their Southeast Asian counterparts. Within *O. rufipogon*, most of the South Asian accessions cluster together as well as the Australasian populations while the (insular and continental) Southeast Asian populations do not exhibit a distinct clustering pattern.

The pairs of sympatric species are separated differently by PC2. Australasian *O. rufipogon* populations are separated from the *O. meridionalis* populations except for accession M4. *O. rufipogon* and *O. nivara* from South Asia occupy separate positions while those from mainland Southeast Asia do not show a distinct pattern as *O. rufipogon* populations are scattered over the PC2 axis. *O. rufipogon* from maritime Southeast Asia (Indonesia and Philippines) seems separated from *O. nivara* of mainland Southeast Asia (Figure 3C). Characters with the highest factor loadings on PC2 are culm number, culm diameter, culm length, distance from panicle base to lowest spikelet insertion and panicle length (Table 2).

PC3 accounts for 10.4% of the total variance and does not produce any distinct clustering pattern between and within *O. nivara* and *O. rufipogon* but separates the *O. meridionalis* accessions (except M2) from Australasian *O. rufipogon* (Figure 3B). Characters with the highest factor loadings on PC3 are culm diameter, number of basal branches of the panicle, spikelet length, stigma length and style length (Table 2).

PC4 (7.4% of the total variance) separates the Australasian *O. rufipogon*, *O. meridionalis* and Southeast Asian *O. nivara* from the South Asian *O. nivara* (Figure 3B). Awn length, panicle length, leaf width and culm number are the characters with the highest factor loadings on PC4 (Table 2).

Hierarchical cluster analysis

The UPGMA dendrogram produces clusters that more or less correspond to species and species groups (Figure 4). The *O. officinalis* accession O2 diverges first

from the rest of the populations. *O. rufipogon* forms a distinct cluster (with 63% bootstrap value). The accessions of the annual species *O. nivara, O. meridionalis* and *O. sativa* join in one cluster (albeit with a bootstrap support less than 50%). The annual species cluster can be divided in two subgroups: one composed of all *O. sativa* accessions (except J2) and the other composed of all *O. nivara* and *O. meridionalis* populations (Figure 4).

Within the *O. nivara* – *O. meridionalis* sub-cluster, *O. sativa* accession J2, *O. nivara* accession N45, and the three *O. meridionalis* accessions M3, M4 and M5 are the first to diverge from the rest of the group while *O. meridionalis* accessions M1 and M2, and all South Asian populations (except accessions N19, N21, N23, N24 and N37) are separated from the Southeast Asian accessions (except N52) (Figure 4). Subpopulations N26A and N26B both join the South Asian *O. nivara* group. In the *O. rufipogon* cluster, accessions from the same geographic region are not consistently grouped together. Grouping of populations according to sympatry and non-sympatry with *O. nivara* and *O. meridionalis* is also not observed in the *O. rufipogon* cluster.

K-means analysis

K-means converged after 3 iterations. A five-cluster solution best fitted the data as the decrease in SSE started to become gradual at k = 5 (Appendix 3). Appendix 4 shows how the accessions were partitioned in the analyses run with k = 2 to k = 5. At k = 5, *O. meridionalis* accessions are recognized as a distinct group (Appendix 4, cluster 1 in Figure 5A). *O. nivara* is separated into South Asian and Southeast Asian clusters (Appendix 4, in Figure 5A clusters 2 and 3, respectively) with the exception of six accessions that do not cluster with their supposed geographic group. The geographic division of *O. nivara* is similarly depicted in the PCA and HCA results. *O. rufipogon* is also divided into two groups. Although the geographical grouping is not as well defined as in *O. nivara*, it can be ascertained that the majority of the continental (South and Southeast) Asian populations (Appendix 4, cluster 4 in Figure 5A) separate from most of the insular Southeast Asian and all of the Australasian populations (Appendix 4, cluster 5 in Figure 5A). The same grouping pattern can be observed at k = 4 but with *O. meridionalis* clustering with South Asian *O. nivara* populations (Appendix 4, cluster 1 in Figure 5B).

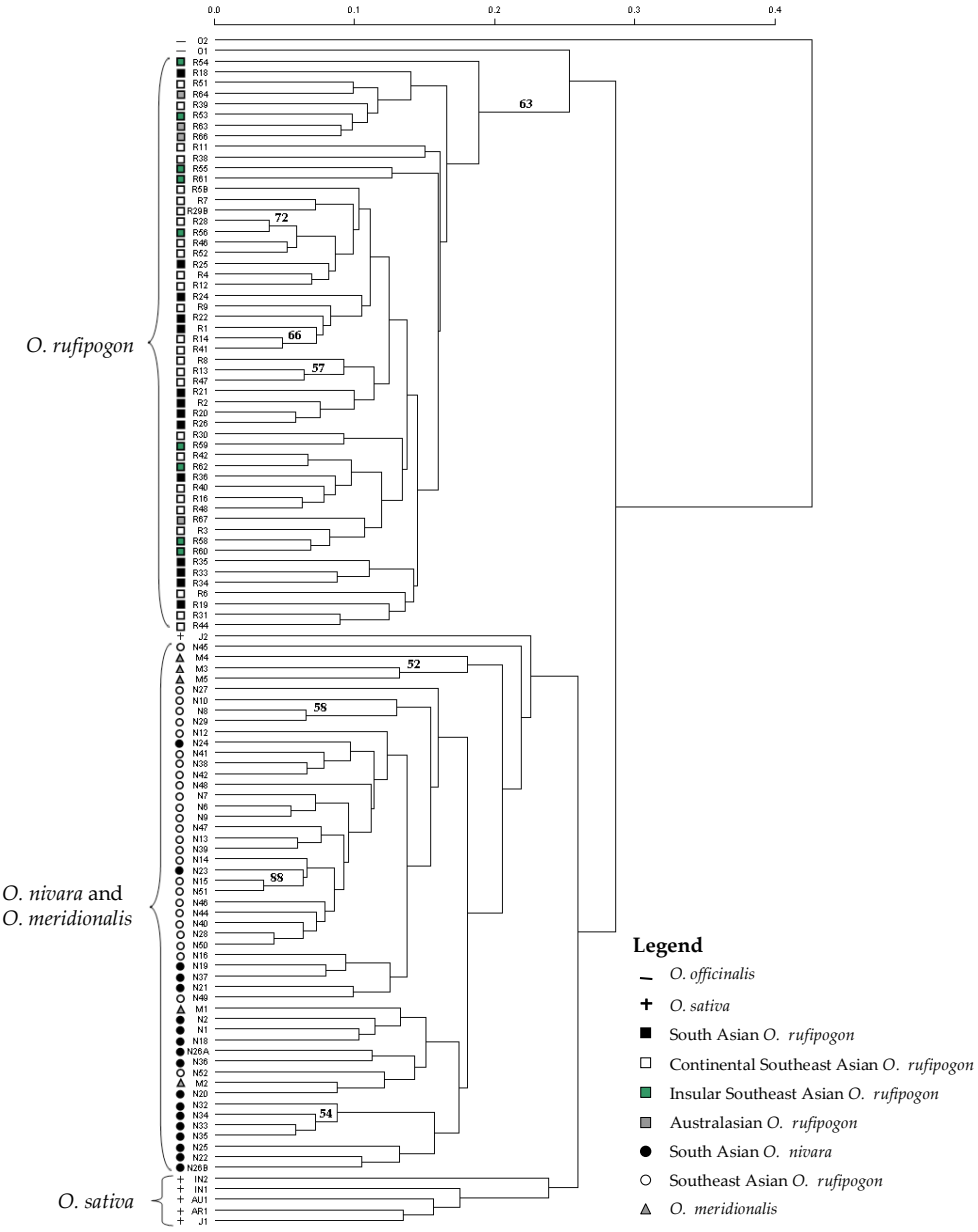


Figure 4. UPGMA dendrogram based on Gower's distance of 112 populations of *Oryza* series *Sativae* from Asia-Pacific. Bootstrap values above 50% are displayed above the branches.

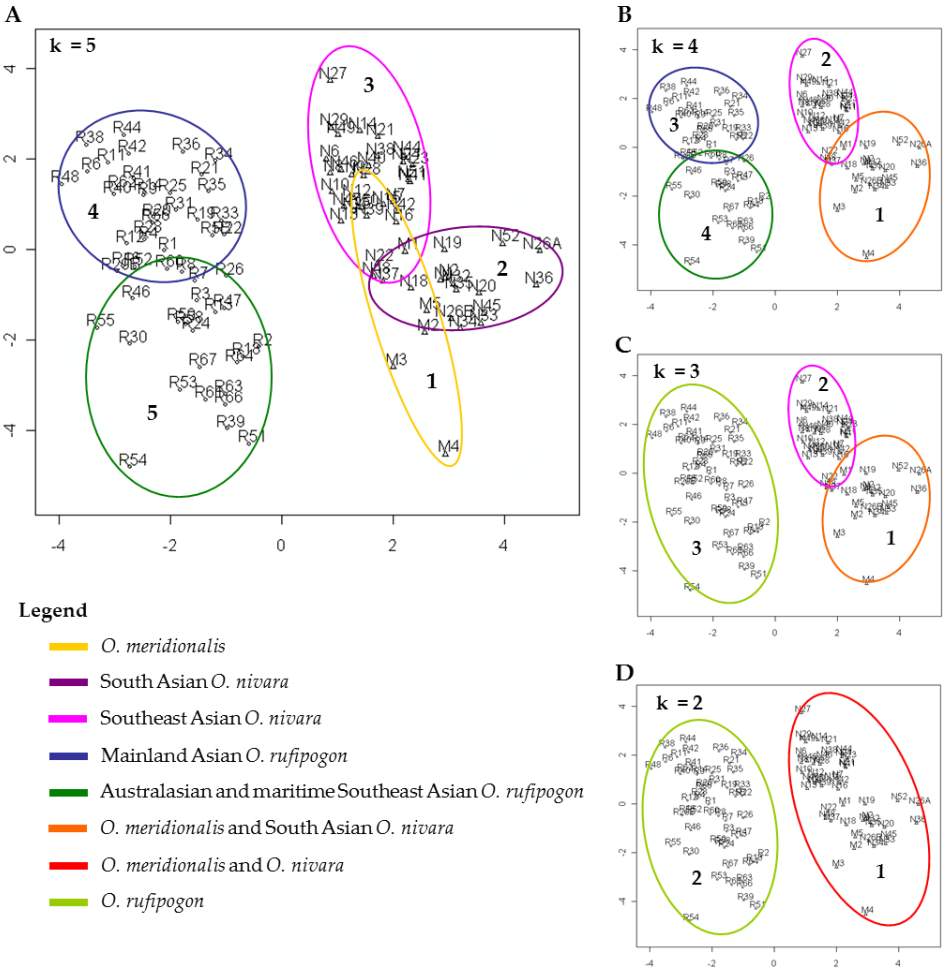


Figure 5. K-means cluster solutions plotted against the PCA scatter plot (component 1 on the x axis and component 2 on the y axis). **A**. k = 5, **B**. k = 4, **C**. k = 3, **D**. k = 2.

At k = 3, the South Asian - Southeast Asian split in *O. nivara* is retained (Appendix 4, in Figure 5C clusters 1 and 2, respectively) with *O. meridionalis* joining the South Asian populations while *O. rufipogon* is recognized as one group (Appendix 3, cluster 3 in Figure 5C). At k = 2, k-means produces a clear-cut separation of *O. rufipogon* and the annual species (Appendix 4, Figure 5D).

Correlation analyses

Euclidean and geographic distance correlations between populations and species

The Mantel test results are presented in Table 3. The Euclidean distance exhibits a weak positive correlation with geographic distance within *O. nivara* (at $p < 0.001$ level), within *O. rufipogon* (at $p < 0.001$ level), and between *O. nivara* and *O. rufipogon* across their distribution range (at $p < 0.01$ level). Similar results were obtained from the geographic populations of *O. nivara* but at lower significance levels ($p < 0.1$ for both South and Southeast Asian populations). In *O. rufipogon*, positive correlation is observed in the continental South ($p < 0.01$ level) and Southeast Asian populations (at $p < 0.05$ level); however, no correlation was detected in the insular Southeast Asian and Australasian members. Significant correlation was also observed between *O. nivara* and *O. rufipogon* in mainland Southeast Asia (at $p < 0.05$ level) but not in South Asia.

Table 3. Mantel correlation of geographic and Euclidean distances within and between the wild species of *Oryza* series *Sativae* in Asia and the Pacific. Correlation coefficient is denoted by r. Statistically significant values are highlighted.

Taxon/taxa covered	r	P-value
Across distribution		
O. meridionalis	-0.0403	0.4605
O. nivara	0.3555	0.0010
O. rufipogon	0.2958	0.0004
O. meridionalis and *O. rufipogon*	0.3048	0.1768
O. nivara and *O. rufipogon*	0.1260	0.0041
Within geographic regions		
O. nivara in South Asia	0.2418	0.0680
O. nivara in continental Southeast Asia	0.1166	0.0864
O. rufipogon in South Asia	0.2848	0.0053
O. rufipogon in continental Southeast Asia	0.2468	0.0196
O. rufipogon in insular Southeast Asia	0.0101	0.4301
O. rufipogon in Australasia	-0.3788	0.7461
O. nivara and *O. rufipogon* in South Asia	0.0475	0.3304
O. nivara and *O. rufipogon* in continental Southeast Asia	0.1780	0.0279

Correlations of geographic position and environmental factors with the Euclidean distances between sympatric populations of O. nivara and O. rufipogon

The Euclidean distances between pairs of sympatric accessions of *O. nivara* and *O. rufipogon* exhibit no correlation with annual precipitation, moderate correlation with longitude (r = -0.4224, p-value = 0.0092) and altitude (r = 0.3551, p-value = 0.0310), and a relatively strong correlation with latitude (r = 0.5804, p-value = 0.0002) and mean annual temperature (r = -0.6084, p-value = 0.0001). The correlation plots of Euclidean distance versus latitude and annual mean temperature are displayed in Figure 6.

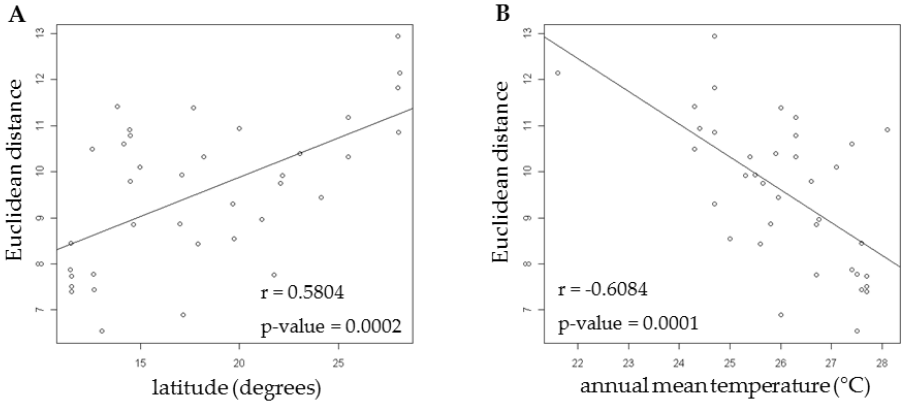


Figure 6. Plots showing the correlation of Euclidean distance with latitude (**A**) and mean annual temperature (**B)** in sympatric populations of *O. nivara* and *O. rufipogon*.

Correlations of geographic position and environmental factors with intraspecific character variations

Ten characters in *O. nivara* and 11 in *O. rufipogon* show a geographical trend (Table 4). In *O. nivara,* a particularly strong correlation is displayed by latitude with leaf width, spikelet width, sterile lemma width and anther length and also by longitude with leaf width and spikelet width. Their correlation plots are shown in Figure 7.

In *O. nivara* and *O. rufipogon*, seedling height, culm number, culm diameter and leaf length exhibit moderate correlation with geographic data and display opposing correlation directions for the two species (e.g., seedling height and longitude are correlated positively in *O. rufipogon* and negatively in *O. nivara*).

Characters that are significantly correlated (at $p < 0.05$ level) with altitude, annual mean temperature and annual precipitation are presented in Table 5. Mean annual temperature and anther length are positively correlated in *O. nivara* and negatively in *O. rufipogon* while leaf width and annual precipitation are negatively correlated in *O. nivara* and positively in *O. rufipogon.*

Table 4. Significant correlations of latitude and longitude to characters within species at $p < 0.05$ significance level. Correlation coefficient is denoted by r. Values indicating particularly strong correlation are highlighted.

Character	latitude		longitude	
	r	P-value	r	P-value
within *O. nivara*				
seedling height	0.3867	0.0087	-0.4390	0.0026
culm number	-0.3831	0.0094	0.4612	0.0014
culm diameter	0.4094	0.0052	-0.4865	0.0007
leaf length	0.4334	0.0029	-0.3527	0.0175
leaf width	0.6065	0.0000	-0.7227	0.0000
spikelet width	0.6460	0.0000	-0.6070	0.0000
awn length	-0.3549	0.0167	0.5238	0.0002
style length	0.3541	0.0170	-0.3127	0.0365
sterile lemma width	0.6443	0.0000	-0.4364	0.0027
anther length	-0.6018	0.0000		
within *O. rufipogon*				
number of days from seeding to first heading	-0.2928	0.0317		
culm diameter	-0.4038	0.0025		
spikelet length	-0.2734	0.0455		
culm number	0.3547	0.0085	-0.3649	0.0067
culm length	-0.3542	0.0086	0.4799	0.0002
leaf length	-0.4426	0.0008	0.4670	0.0004
panicle length	-0.3167	0.0197	0.3324	0.0141
distance from panicle base to lowest spikelet insertion	-0.2684	0.0498	0.3289	0.0152
stigma length	-0.3340	0.0136	0.3685	0.0061
awn length			0.3279	0.0055
seedling height			0.3205	0.0182

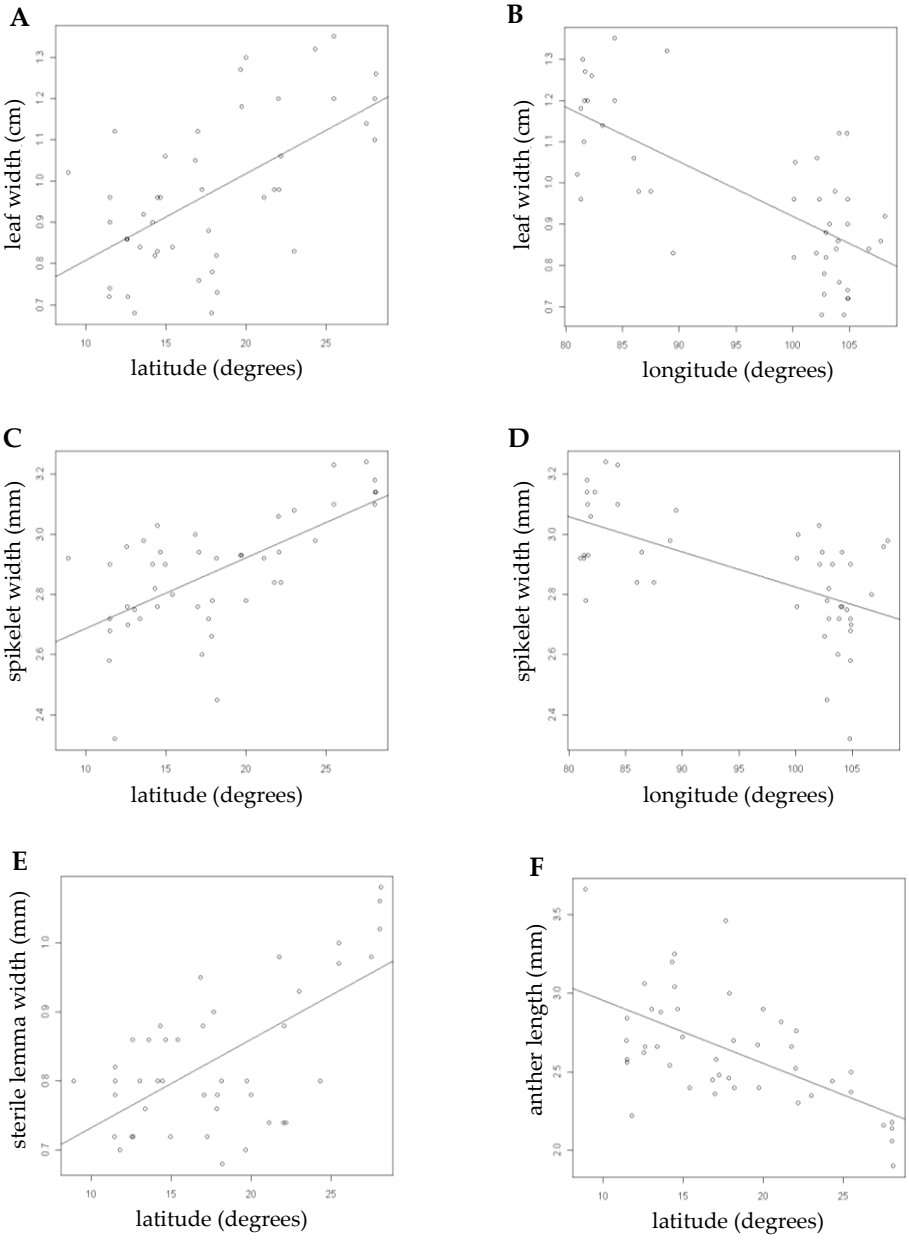


Figure 7. Correlation plots of geographic data (x axis) and selected quantitative characters (y axis) in *O. nivara*. **A.** latitude vs. leaf width**, B.** longitude vs. leaf width, **C.** latitude vs. spikelet width, **D**. longitude vs. spikelet width, **E.** latitude vs. sterile lemma width, and **F.** latitude vs. anther length.

Table 5. Correlation of altitude, annual mean temperature and annual precipitation to characters within species at $p < 0.05$ significance level. Correlation coefficient is denoted by r.

Character	altitude		annual mean temperature		annual precipitation	
	r	P-value	r	P-value	r	P-value
within *O. nivara*						
culm length	-0.3122	0.0368				
spikelet length	0.3110	0.0376				
number of days from seeding to first heading	-0.3136	0.0359				
leaf width	0.3466	0.0197	-0.3770	0.0107	-0.3414	0.0217
spikelet width			-0.4359	0.0028	-0.3316	0.0261
sterile lemma width			-0.3883	0.0084		
anther length			0.4419	0.0024		
sterile lemma length			-0.3746	0.0112		
within O. *rufipogon*						
anther length			-0.4004	0.0027		
number of days from seeding to first heading	-0.3486	0.0098	0.4821	0.0002		
culm diameter	-0.2698	0.0485			0.3335	0.0137
stigma length					0.2821	0.0388
leaf width					0.3636	0.0069
number of basal primary branches of the panicle					0.4591	0.0005

Analysis of Variance

Euclidean distances of sympatric species in geographic regions

The ANOVA reveals a significant difference (at $p < 0.05$ level) between the Euclidean distances of sympatric *O. nivara* and *O. rufipogon* populations in South Asia (mean = 10.29) and Southeast Asia (mean = 9.07).

Interspecific character differences

The combined effect of the 34 characters on the differences between species is highly significant at $p < 0.001$ level.

Differences between all three species or certain pairs of species have been observed in 30 characters. Number of panicles is significant at $p < 0.001$ level; all others are statistically highly significant at $p < 0.0001$ level (Table 6).

Twelve characters differentiated *O. meridionalis, O. nivara* and *O. rufipogon* from each other. Eleven traits separated *O. rufipogon* from the two annual species while six discriminated *O. meridionalis* from *O. nivara* and *O. rufipogon.* Only one character (spikelet width to spikelet length ratio) distinguished *O. nivara* from the other species (Table 6).

Table 6. Morphological variation within and between *O. meridionalis, O. nivara* and *O. rufipogon* accessions (N = number of samples). Mean standard error as well as *F* and *P* values of each character are presented. Means with the same superscripts do not differ significantly at $P < 0.05$ of Tukey`s HSD test.

Character (unit)	*O. meridionalis* (N=22)	*O. nivara* (N=215)	*O. rufipogon* (N=240)	ANOVA *F*	*P*
culm length (cm)	113.8 ± 6.13[a]	93.6 ± 1.21[b]	136.0 ± 1.76[c]	186.02	0.0000
culm habit	3.7 ± 0.21[a]	4.5 ± 0.07[b]	6.1 ± 0.13[c]	74.07	0.0000
leaf length (cm)	54.9 ± 2.3[a]	48.9 ± 0.75[b]	37.0 ± 0.75[c]	75.03	0.0000
flag leaf width (cm)	1.45 ± 0.038[a]	1.09 ± 0.017[b]	0.95 ± 0.013[c]	62.55	0.0000
flag leaf attitude	3.6 ± 0.4[a]	2.9 ± 0.1[b]	4.6 ± 0.07[c]	87.14	0.0000
panicle type	2.3 ± 0.36[a]	3.5 ± 0.11[b]	5.1 ± 0.06[c]	101.46	0.0000
spikelet width (mm)	2.41 ± 0.038[a]	2.86 ± 0.015[b]	2.30 ± 0.012[c]	461.93	0.0000
sterile lemma width (mm)	0.7 ± 0.014[a]	0.83 ± 0.008[b]	0.76 ± 0.007[c]	26.57	0.0000
awn length (mm)	128.5 ± 3.58[a]	88.5 ± 1.21[b]	67.0 ± 1.14[c]	172.91	0.0000
awn thickness (mm)	0.31 ± 0.01[a]	0.27 ± 0.00[b]	0.20 ± 0.00[c]	225.80	0.0000
spikelet fertility (%)	50.1 ± 4.39[a]	64.2 ± 1.08[b]	32.4 ± 1.4[c]	151.85	0.0000
anther length to spikelet length ratio	0.28 ± 0.005[a]	0.31 ± 0.003[b]	0.6 ± 0.005[c]	1653.80	0.0000
culm number	98.2 ± 17.01[a]	167.6 ± 4.18[b]	142.0 ± 3.52[b]	20.19	0.0000
culm diameter (mm)	5.6 ± 0.24[a]	4.7 ± 0.05[b]	4.6 ± 0.06[b]	12.62	0.0000
leaf width (cm)	1.31 ± 0.031[a]	0.96 ± 0.014[b]	0.9 ± 0.013[b]	43.77	0.0000
panicle number	70.0 ± 9.4[a]	115.0 ± 4.0[b]	112.0 ± 3.0[b]	7.57	0.0005
number of basal primary branches of the panicle	1.4 ± 0.12[a]	1.1 ± 0.02[b]	1.2 ± 0.02[b]	12.28	0.0000

Table 6. (Continued) Morphological variation within and between *O. meridionalis, O. nivara* and *O. rufipogon* accessions (N = number of samples). Mean standard error as well as *F* and *P* values of each character are presented. Means with the same superscripts do not differ significantly at $P < 0.05$ of Tukey`s HSD test.

Character (unit)	*O. meridionalis* (N=22)	*O. nivara* (N=215)	*O. rufipogon* (N=240)	ANOVA *F*	*P*
panicle length (cm)	24.3 ± 0.6[a]	20.8 ± 0.18[b]	20.3 ± 0.27[b]	13.36	0.0000
spikelet width to spikelet length ratio	0.28 ± 0.005[a]	0.33 ± 0.002[b]	0.27 ± 0.001[a]	434.63	0.0000
culm lodging resistance	3.5 ± 0.18[a]	3.8 ± 0.09[a]	4.6 ± 0.07[b]	26.64	0.0000
rhizome and stolon formation	1.0 ± 0.00[a]	1.1 ± 0.00[a]	1.8 ± 0.03[b]	248.05	0.0000
number of days from seeding to first heading	114.0 ± 4.90[a]	110.0 ± 1.30[a]	161.5 ± 1.97[b]	241.09	0.0000
flag leaf length (cm)	30.3 ± 0.96[a]	30.7 ± 0.48[a]	20.1 ± 0.39[b]	155.93	0.0000
ligule length (mm)	24.2 ± 1.17[a]	21.5 ± 0.39[a]	18.5 ± 0.43[b]	19.11	0.0000
panicle exsertion	7.4 ± 0.47[a]	7.3 ± 0.14[a]	8.9 ± 0.05[b]	63.20	0.0000
distance from panicle base to lowest spikelet insertion (mm)	19.5 ± 1.11[a]	18.8 ± 0.33[a]	26.1 ± 0.62[b]	54.40	0.0000
stigma length (mm)	1.4 ± 0.03[a]	1.5 ± 0.01[a]	1.8 ± 0.02[b]	76.93	0.0000
anther length (mm)	2.4 ± 0.04[a]	2.6 ± 0.03[a]	5.1 ± 0.04[b]	1401.10	0.0000
leaf senescence	3.5 ± 0.18[a]	3.2 ± 0.05[a]	4.1 ± 0.08[b]	45.58	0.0000
life cycle	1.0 ± 0.00[a]	1.0 ± 0.02[a]	3.0 ± 0.02[b]	2458.60	0.0000

Intra-specific character differences

The combined effects of the 34 characters on the differences between geographic populations within *O. nivara* and within *O. rufipogon* were also both highly significant at $p < 0.001$ level.

In *O. nivara,* 17 characters were significantly different (at $p < 0.0001$ level) between the South Asian and Southeast Asian populations (Table 7). Within *O. rufipogon* 16 characters differentiated certain pairs or groups of geographic populations (at $p < 0.001$ significance level; Table 8). Seven of these discriminated the Australasian populations from the rest of the species (Table 8).

Table 7. Characters with highly significant differences ($P < 0.001$) between South Asian and Southeast Asian populations of *O. nivara* (N = number of samples).

Character (unit)	South Asian (N=81)	South East Asian (N=134)	ANOVA	
			F	*P*
seedling height (cm)	23.4 ± 0.50	20.3 ± 0.41	23.18	0.0000
culm habit	4.0 ± 0.12	4.7 ± 0.08	24.11	0.0000
culm number	136.1 ± 5.64	186.6 ± 5.13	40.74	0.0000
culm diameter (mm)	5.1 ± 0.09	4.4 ± 0.06	35.49	0.0000
leaf length (cm)	53.3 ± 1.15	46.3 ± 0.92	22.12	0.0000
leaf width (cm)	1.13 ± 0.019	0.86 ± 0.014	131.94	0.0000
flag leaf length (cm)	33.4 ± 0.81	29.0 ± 0.55	21.94	0.0000
flag leaf width (cm)	1.28 ± 0.024	0.98 ± 0.017	106.35	0.0000
flag leaf attitude	2.3 ± 0.16	3.2 ± 0.13	21.01	0.0000
panicle exsertion	6.0 ± 0.22	8.0 ± 0.15	61.25	0.0000
panicle type	3.0 ± 0.18	3.9 ± 0.12	18.64	0.0000
spikelet width (mm)	3.01 ± 0.019	2.77 ± 0.016	94.21	0.0000
sterile lemma width (mm)	0.9 ± 0.02	0.8 ± 0.01	42.05	0.0000
awn length (mm)	78.9 ± 1.81	94.4 ± 1.38	46.84	0.0000
anther length (mm)	2.5 ± 0.05	2.7 ± 0.03	23.28	0.0000
anther length to spikelet length ratio	0.29 ± 0.005	0.32 ± 0.003	27.07	0.0000
spikelet width to spikelet length ratio	0.35 ± 0.003	0.32 ± 0.001	99.20	0.0000

Table 8. Characters with highly significant differences ($P < 0.001$) among the different geographical populations of *O. rufipogon* (N = number of samples). Means with the same superscripts do not differ significantly at $P < 0.05$ of Tukey`s HSD test.

Character (unit)	South Asia (N=62)	Continental Southeast Asia (N=111)	Insular Southeast Asia (N=42)	Australasia (N=20)	ANOVA	
					F	P
culm number	151.8 ± 7.27^{a}	154.6 ± 4.86^{a}	121.4 ± 6.36^{b}	84.7 ± 6.27^{c}	14.42	0.0000
anther length (mm)	5.2 ± 0.07^{ab}	4.9 ± 0.06^{a}	5.3 ± 0.09^{b}	5.7 ± 0.13^{c}	13.06	0.0000
leaf length (cm)	32.3 ± 1.27^{a}	37.2 ± 1.10^{ab}	38.9 ± 1.83^{b}	46.7 ± 1.67^{c}	9.29	0.0000
culm length (cm)	120.0 ± 2.53^{a}	140.7 ± 2.65^{b}	139.6 ± 3.82^{b}	151.9 ± 5.34^{b}	12.63	0.0000
distance from panicle base to lowest spike-let insertion (mm)	22.0 ± 1.22^{a}	27.2 ± 0.87^{ab}	27.3 ± 1.52^{b}	30.6 ± 1.52^{b}	6.43	0.0003

Table 8. (Continued) Characters with highly significant differences ($P < 0.001$) among the different geographical populations of *O. rufipogon* (N = number of samples). Means with the same superscripts do not differ significantly at $P < 0.05$ of Tukey`s HSD test.

Character (unit)	South Asia (N=62)	Continental Southeast Asia (N=111)	Insular Southeast Asia (N=42)	Australasia (N=20)	ANOVA F	ANOVA P
number of days from seeding to first heading	133.8 ± 2.91[a]	173.5 ± 2.05[b]	183.7 ± 3.37[b]	134.3 ± 5.71[a]	66.18	0.0000
panicle number	126.3 ± 5.42[a]	120.4 ± 4.49[a]	87.7 ± 5.17[b]	70.3 ± 5.73[b]	14.86	0.0000
culm diameter (mm)	4.4 ± 0.10[a]	4.5 ± 0.08[a]	5.3 ± 0.15[b]	4.8 ± 0.17[a]	11.19	0.0000
culm habit	6.4 ± 0.23[a]	6.5 ± 0.20[a]	5.6 ± 0.29[ab]	4.7 ± 0.22[b]	6.28	0.0004
seedling height (cm)	20.4 ± 0.56[a]	18.7 ± 0.34[a]	18.6 ± 0.69[a]	27.7 ± 0.93[b]	30.19	0.0000
ligule length (mm)	18.7 ± 0.67[a]	17.6 ± 0.61[a]	17.7 ± 1.2[a]	24.1 ± 1.1[b]	6.15	0.0005
flag leaf length (cm)	19.3 ± 0.67[a]	19.5 ± 0.53[a]	20.4 ± 1.06[a]	25.5 ± 1.35[b]	6.67	0.0002
panicle length (cm)	19.6 ± 0.45[a]	20.0 ± 0.39[a]	20.4 ± 0.68[a]	24.3 ± 0.72[b]	7.73	0.0001
spikelet length (mm)	8.5 ± 0.06[a]	8.3 ± 0.06[a]	8.6 ± 0.06[a]	9.2 ± 0.09[b]	16.59	0.0000
awn length (mm)	63.9 ± 1.8[a]	67.5 ± 1.65[a]	61.7 ± 2.37[a]	85.4 ± 4.48[b]	10.48	0.0000
stigma length (mm)	1.8 ± 0.04[a]	1.8 ± 0.03[a]	1.9 ± 0.04[a]	2.1 ± 0.05[b]	8.52	0.0000

Box and whisker plot analysis

Six characters displayed clear-cut differences between certain pairs of species or geographical populations (Figure 8). Anther length was well differentiated between the annual species and *O. rufipogon* (Figure 8A). Spikelet width also differentiated *O. nivara* from *O. rufipogon* (Figure 8B). *O. meridionalis* was discriminated from the South Asian and insular Southeast Asian *O rufipogon* by awn length (Figure 8C) and from the Australasian *O. rufipogon* by flag leaf width (Figure 8D) and stigma length (Figure 8E). The number of days from seeding to first heading was higher in both insular and continental Southeast Asian *O. rufipogon* and distinguished the former from continental Southeast Asian *O. nivara* (Figure 8F).

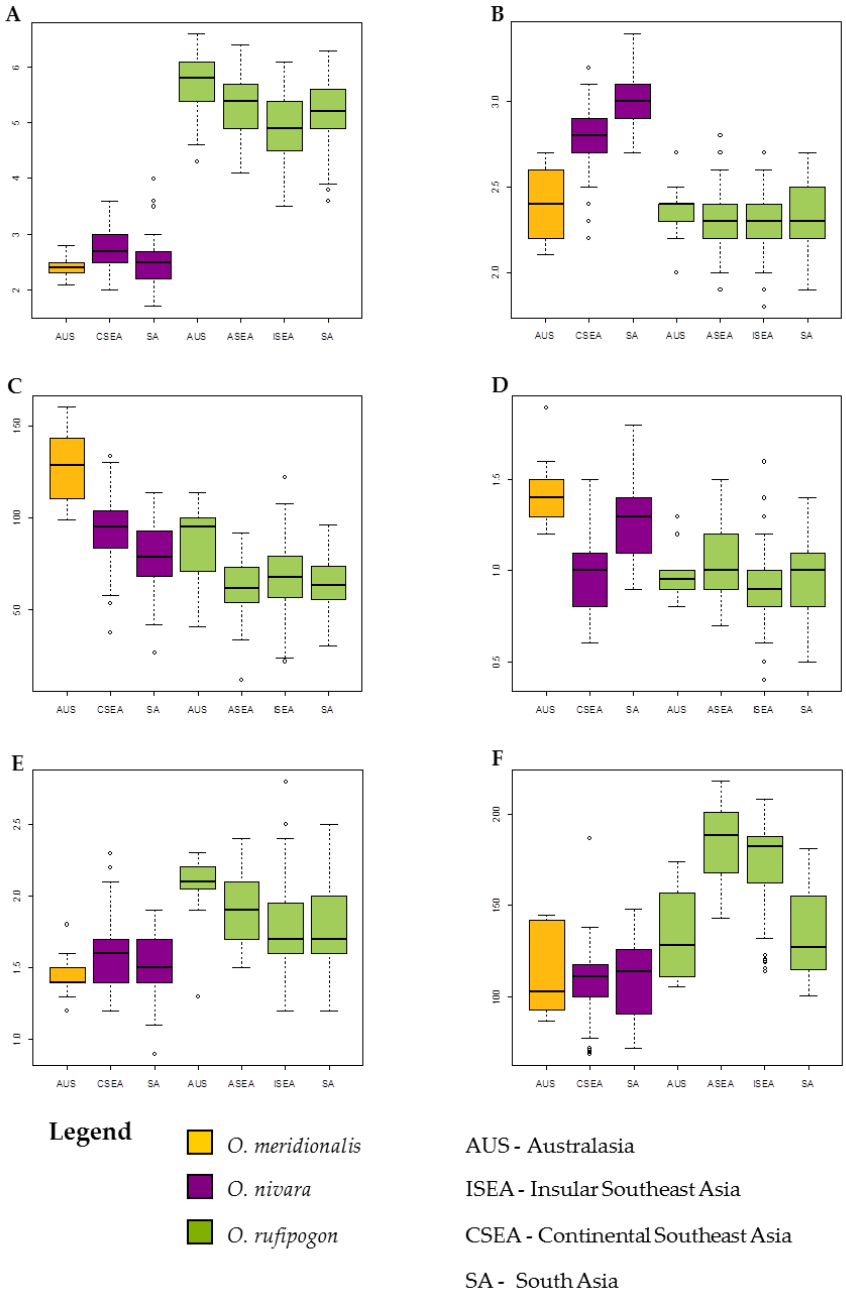


Figure 8. Box and whisker plots of characters showing differences between species and geographical populations. **A**. anther length (mm), **B**. spikelet width (mm), **C**. awn length (mm), **D**. flag leaf width (cm), **E**. stigma length(mm) and **F**. number of days from seeding to first heading.

Discussion

Two morphological species

Hierarchical clustering and various partitional (k-means) clustering methods all discriminate *O. rufipogon* from the annual species *O. nivara* and *O. meridionalis*. This is in accordance with previous numerical studies that verified the phenotypic separation of *O. nivara* and *O. rufipogon* (Ng et al. 1981a; Barbier 1989; Cai et al. 2004), and reinforces Sharma and Shastry`s (1965) decision to identify the former as a distinct morphological species.

Principal component analysis did not yield such a clear partitioning at first (Figure 3C), but several accessions that held an intermediate position were suspected to represent hybrids. Based on herbarium observations, six of these 18 accessions (N11, N31, R10, R15, R43 and R50) were tentatively identified as weedy forms or hybrid swarms between *O. sativa* and either *O. nivara* or *O. rufipogon*. Hybrid swarms are produced when neighboring populations of wild and cultivated rice interact and exchange genes (Oka and Chang 1961). In continental Asia, gene flow between populations of *O. rufipogon* (*sensu lato*) and *O. sativa* is highly probable due to their overlapping distribution in this region (Vaughan et al. 2003). The remaining 12 accessions (including subpopulations R5A and R29A) appear to be intermediate between *O. nivara* and *O. rufipogon* and so, are assumed to be hybrids or introgressed populations of the two wild species. Such hybrid or intermediate populations were reported by Morishima et al. (1980; 1984) and are most likely represented in the IRG collection. Hybrids and non-hybrid individuals can coexist in mixed accessions as exemplified by the two *O. rufipogon* accessions R5 and R29 (Appendix 1). While R5A and R29A seemed to represent intermediate forms, their corresponding co-subpopulations R5B and R29B consistently grouped with *O. rufipogon* in PCA, HCA and k-means analysis.

Although intermediate populations are present, *O. nivara* and *O. rufipogon* are morphologically distinct from each other as shown by the two distinguishable clusters in the PCA plot (Figure 3C). In Laos, natural populations of *O. nivara* and *O. rufipogon* can be easily differentiated by flag leaf length and width, panicle branching type and distance from the panicle base to the lowest spikelet insertion (Banaticla-Hilario and Almazan 2010). Ng et al. (1981a) and Barbier (1989) also found phenotypic discontinuities between the two species, disputing the Asian wild rice perennial-annual continuum postulated by Morishima et al. (1980; 1984).

Although one could argue that the existence of intermediate forms or hybrids points to the need to recognize the taxa only at an intraspecific taxonomic level, we deem the outcomes of especially the clustering and k-means analyses sufficiently strong to maintain the distinction between *O. rufipogon* and both annual species at the level of species. To support our assumption about the nature of the intermediate populations, an in-depth genetic study would be necessary.

Some of the phenotypic differences between the perennial *O. rufipogon* and the two annual species can be seen as responses to the differences in their habitat, breeding system and life cycle. The presence of stolons and long, strong and spreading culms make *O. rufipogon* suitable for permanently inundated habitats while the shorter, less decumbent culms of *O. nivara* and *O. meridionalis* are probably more suited for seasonally dry habitats. Stigmas and anthers are longer and panicles are more exserted and open in the outcrossing *O. rufipogon* than in the two annual, inbreeding species. Leaf senescence tends to be earlier, and flowering is not as photoperiod sensitive in the annual species. Adaptation to different habitat conditions (particularly to different hydrological regimes) led to differentiation of morphology, reproductive system and life cycle (Morishima et al. 1984; Vaughan et al. 2003). All this further supports our conclusion that *O. rufipogon* should be recognized as being different from both the other taxa at species level.

O. meridionalis consistently groups with *O. nivara* populations in our PCA and HCA results (Figures 3-4). Ng et al. (1981a) obtained a similar result when hierarchical cluster analysis was applied to 30 accessions of series *Sativae* species using 41 morphological characters. K-means clustering distinguishes *O. meridionalis* and the geographic groups of *O. nivara* and *O. rufipogon* at k = 5, although at k = 3, it clusters with the South Asian *O. nivara* populations. In spite of its morphological similarities to *O. nivara*, *O. meridionalis* is a clearly separated species based on molecular phylogenetic evidence (Wang et al. 1992; Bautista et al. 2001; Park et al. 2003; Ohtsubo et al. 2004; Juliano et al. 2005; Xu et al. 2005; Zhu and Ge 2005; Kwon et al. 2006; Duan et al. 2007). In addition, it is possible to distinguish *O. meridionalis* from *O. nivara* and *O. rufipogon* using the set of diagnostic characters developed here. Univariate tests recognized significant character differences that multivariate failed to detect. ANOVA identified 12 characters that can differentiate the three species from each other and another six characters that can distinguish *O. meridionalis* from the other two species (Table 6). The marked morphological overlap between *O. meridionalis* and *O. nivara* can probably be attributed to their identical ecological niche. Both species occupy open, swampy areas that are seasonally dry (Vaughan et al. 2003).

This study implies a greater influence of ecological factors (compared to geographical) on the morphology of species. *O. nivara* and *O. rufipogon* have probably undergone and are still undergoing ecological speciation. On the other hand, *O. meridionalis* and *O. nivara* are two allopatric species with resembling morphology possibly caused by ecological adaptation and/or parallel evolution.

More distinct in South Asia, more alike in Southeast Asia

Across their entire distribution, *O. nivara* and *O. rufipogon* populations tend to be more phenotypically different with increasing spatial distance (Table 3). However, when examined separately, geographic populations exhibit different variation patterns. In continental Southeast Asia, *O. nivara* and *O. rufipogon* appear to be more similar as they become more geographically close to each other, implying gene flow between adjacent populations. Yet in South Asia, they retain their morphological differences even at close spatial distances. The variation between *O. meridionalis* and *O. rufipogon* in Australasia also seemed to be not spatially influenced (Table 3). It should be noted that the correlations obtained from the distances between *O. nivara* and *O. rufipogon* were significant but weaker (across distribution r = 0.13; within continental Southeast Asia r = 0.18) than the correlations obtained from distances within species (r ranging from 0.24 to 0.36, except in Southeast Asian *O. nivara* where r = 0.12) implying that intra-specific gene flow is more spatially sensitive than inter-specific gene exchange.

The PCA also shows that *O. nivara* and *O. rufipogon* are more widely separated morphologically in South Asia than in Southeast Asia (Figure 3C). The ANOVA and correlation analysis likewise reveal that sympatric populations of South Asian *O. nivara* and *O. rufipogon* are more morphologically distinct than their Southeast Asian counterparts as Euclidean distance is correlated positively with latitude and negatively with longitude (Figure 6A). The strong negative correlation of Euclidean distance and annual mean temperature (Figure 6B) is largely influenced by the inversely proportional relationship between the latter and latitude. These results indicate that in Southeast Asia, the conditions may allow more active gene flow between species (hence more frequent hybridization events) resulting in less phenotypic differences while in South Asia, the reproductive barriers are more effective and thus maintain two morphologically more distinct entities despite geographical sympatry. Molecular analysis using simple sequence repeats (SSRs) and single nucleotide polymorphisms (SNPs) data and hybridization studies could be applied to validate this conclusion and provide more detailed information on

the migration of genes between *O. nivara* and *O. rufipogon*.

Geographical patterns of variation within species

Variations in insular populations not influenced by spatial distance
The positive mantel correlation of Euclidean and geographic distances implies the dependence of gene flow in *O. nivara* (across distribution and within geographical populations) to geographic distance. The same spatial influence on morphology was observed in *O. rufipogon* across its distribution. However, within geographic regions, spatial correlation was observed primarily in the continental (South and Southeast) Asians and not in the insular populations. In Australasia and insular Southeast Asia, *O. rufipogon* populations are generally more geographically isolated from each other apparently resulting in a more restricted gene flow. Such spatial independence of intra-specific genetic distance was observed in other insular plant species as well, for example *Jepsonia malvifolia* (Greene) Small (Helenurm 2001), *Castilleja grisea* Dunkle (Helenurm et al. 2005), *Pilosocereus euchlorus* (Weber) Taylor and *Pilosocereus machrisii* Dawson (Moraes et al. 2005).

South Asian and Southeast Asian phenotypes in O. nivara
The regional separation of *O. nivara* into South and Southeast Asian populations is consistently captured in the PCA, HCA and k-means results (Figures 3C - 5). The ANOVA (Table 7) indicates that South Asian populations tend to have taller seedlings [(15) 23 – 37 cm], fewer, more erect and broader culms [(3.3) 5.1 – 7.6 mm], larger leaves (~53 cm long, ~1.13 cm wide) and flag leaves (~33.4 cm long, ~1.28 cm wide), more erect flag leaf attitude, less exserted and more compact panicles, broader spikelets [(2.7) 3 – 3.4 mm] and sterile lemmas [(0.6) 0.9 – 1.2 mm] and shorter awns [(42) 78 – 114 cm] and anthers [1.7 – 2.5 mm] compared to the Southeast Asian populations. However, because these are all tendencies and no clear morphological distinction can be made, we do not think it is advisable to grant these regional populations a formal taxonomic rank.

Australasian populations differ from the rest of O. rufipogon
The PCA (Figure 3) and HCA (Figure 4) did not divide *O. rufipogon* into geographic groups but at k = 5 the k-means analysis separated all the Australasian and most (6 out of 9) of the insular Southeast Asian populations from the majority (70.7%) of the continental Southeast Asian ones (Figure 5A). This continental-insular variation reflects the typically high genetic differentiation observed in island populations resulting from their geographic isolation (Fedorenko et al. 2009). A similar

divergence of mainland and island populations at the phenotype and genotype levels has been reported in the tree species *Pinus merkusii* Junghuhn & de Vriese (Howcroft and Davidson, 1973) and *Pterocarpus officinalis* Jacq. (Rivera-Ocasio et al. 2006), respectively. The ANOVA (Table 8) revealed that Australasian populations are quite distinct from the rest as they possess significantly fewer culms [39 – 55 (150)], taller seedlings (22 to more than 30 cm), and longer leaves [(36) 46 – 62 cm], flag leaves [(16) 25 – 40 cm], ligules [(18) 24 – 38 mm], panicles [(19) 24 – 32 cm], spikelets [(8.3) 9.2 – 9.8 mm], awns [(40) 85 – 113 cm], anthers [(4.3) 5.7 – 6.6 mm] and stigmas [(1.3) 2.1 – 2.3 mm]. This corresponds with the correlation results (Table 4) summarized into four trends: 1) culms are narrower and the number of days from seeding to flowering is shorter at higher latitudes (continental Asia); 2) seedlings are taller and awns are longer at higher longitude (Australasia); 3) distance from panicle base to lowest spikelet insertion and length of culms, leaves, panicles, spikelets and stigmas decrease with latitude and increase with latitude; and 4) culm number increases with latitude and decrease with longitude. The distinctiveness of the Australasian populations can most likely be attributed to their geographic separation from the Asian populations but, again, we do not see the need to grant this distinction a formal taxonomic status.

Contrasting responses to geographical and environmental gradients

In continental Asia, regional differentiation is evident in *O. nivara* but lacking in *O. rufipogon.* This contrasting variation pattern agrees with the genetic diversity patterns observed in the two species (Kuroda et al. 2007; Zhou et al. 2008; Zheng and Ge 2010) in which *O. nivara* showed higher inter-population genetic differentiation than *O. rufipogon.* Again, differences in habitat and breeding system may have influenced the different behavior of the two species. Inbreeding plants like *O. nivara* generally exhibit low intra-population diversity but in situations where local adaptation is important, the different populations that are isolated from each other would maintain different alleles at certain loci, hence resulting in greater inter-population than intra-population diversity (Silvertown and Charlesworth 2001). *O. rufipogon* is found in comparatively stable habitats while *O. nivara* occupies disturbed habitats (Kariali et al. 2008). Fluctuating environmental conditions could have induced *O. nivara* to become more responsive to the regional differences in climate and other environmental factors. In fact, four characters in *O. nivara* and none in *O. rufipogon* exhibited strong correlation with latitude and annual mean temperature (Table 4).

In the two species eighteen characters exhibit weak to moderate correlations to geographic and/or climatic data (Table 4 & 5). Eight characters that are

significantly correlated to one or more of the factors are common to *O. nivara* and *O. rufipogon*. In both species, awn length tends to increase with longitude while the number of days from seeding to first heading tends to decrease with increasing altitude. The remaining common characters exhibit opposing correlation directions. As latitude increases, culm diameter and leaf length increase in *O. nivara* and decrease in *O. rufipogon* while culm number decreases in *O. nivara* and increases in *O. rufipogon.* As longitude increases, seedling height and leaf length decrease in *O. nivara* and increase in *O. rufipogon* while culm number increases in *O. nivara* and decreases in *O. rufipogon*. With increasing annual mean temperature, anthers tend to lengthen in *O. nivara* and shorten in *O. rufipogon* while with increasing annual precipitation, leaves tend to narrow in *O. nivara* and broaden in *O. rufipogon.* Such contrasting correlations suggest that the two species behave differently towards the existing geographic and climatic gradients. The opposing morphological responses probably contribute to the phenotypic dissimilarities observed between sympatric *O. nivara* and *O. rufipogon* populations.

Diagnostic characters

The ANOVA detected 12 characters that can discriminate between the three species (Table 6). Leaf length, flag leaf width, awn length and awn thickness are highest in *O. meridionalis* and lowest in *O. rufipogon. O. nivara* has the highest spikelet width, spikelet fertility (both are lowest in *O. rufipogon*) and sterile lemma width (lowest in *O. meridionalis*) while *O. rufipogon* has the highest culm length, flag leaf attitude (both are lowest in *O. nivara*), culm habit, panicle type and anther length to spikelet length ratio (all three are lowest in *O. meridionalis*) (Table 6). Box and whisker plot analysis revealed that anther length can readily distinguish *O. rufipogon* and the combination of spikelet width and awn length can separate *O. nivara* and *O. meridionalis* from each other (Figure 8). Even so, it is not advisable to rely mainly on these characters. The following descriptions provide a clear delineation of each species.

O. rufipogon is characterized by a spreading habit, stolon formation, long culms [(74) 136 to more than 170 cm] that are moderately resistant to lodging, short leaves [15 – 37 (76) cm], flag leaves [7 – 20 (40) cm] and ligules [(2.4) 18.4 – 30 (48)], well exserted panicles with spreading branches, long distances from the panicle base to lowest spikelet insertion [(7) 16 – 26 (54) mm], long anthers (>3.7mm) and stigmas [(1.2) 1.8 – 2.8 mm], high anther to spikelet length ratios (0.4 – 0.8), narrow spikelets [1.8 – 2.3 (2.8)mm], low spikelet fertilities [2 – 32 (88)%], late leaf

senescence, photoperiod sensitivity and a perennial life cycle.

The other two species share the following character traits: semi-erect to open habits, absence of stolons, short plant heights (<170 cm), weak culms, long flag leaves [(16) 30 - 49 cm) and ligules [(9.4) 21 – 44 mm], moderately well exserted panicles with compact to semi-compact branches, short distances from the panicle base to the lowest spikelet insertion [(9 – 18 (31) mm], short anthers (<3.7mm) and stigmas [(0.9) 1.4 – 1.8 (2.3) mm], low anther to spikelet length ratios (usually <0.4), early leaf senescence, lack of photoperiod sensitivity and an annual life cycle.

O. nivara can be distinguished from the other two species by broad spikelets [(2.2) 2.8 – 3.4 mm], consequently high spikelet width to length ratios (0.28 to 0.4) and broad sterile lemmas [(0.6) 0.83 – 1.2 mm]. *O. meridionalis* can be discriminated by thick culms [(4.1) 5.6 - 8.2 mm), broad leaves (1 – 1.6 cm), long panicles [(18.6) 24 – 28 cm] and narrow sterile lemmas [0.6 – 0.7 (0.8) mm]. Culm number and panicle number were detected by ANOVA as discriminating for *O. meridionalis*. However, these characters should only be considered when dealing with potted plants or plants in a contained environment. It was reported that in their natural habitat, *O. nivara* produces fewer tillers while *O. rufipogon* generates abundant tillers (Kariali et al 2008).

These descriptions delineate the three species morphologically and can be used to identify *Oryza* series *Sativae* plants (from Asia-Pacific) or reclassify those that have been misidentified at the species level. We hope this will help gene banks and herbaria in providing correct taxonomic labels for their *Oryza* accessions and specimens.

References

Aggarwal, R. K., Brar, D. S., Nandi, S., Huang, N. & Khush, G. S. (1999). Phylogenetic relationships among *Oryza* species revealed by AFLP markers. *Theoretical and Applied Genetics* **98,** 1320-1328.

Banaticla-Hilario, M. C. N. & Almazan, M. S. R. (2010). Field observation of wild rice populations in Vientiane, Laos. pp. 12. International Rice Research Institute, Los Baños, Laguna, Philippines.

Barbier, P. (1989). Genetic variation and ecotypic variation in the wild rice species *Oryza rufipogon.* I. Population differentiation in life-history traits and isozymic loci. *The Japanese Journal of Genetics* **64,** 259-271.

Bautista, N. S., Solis, R., Kamijima, O. & Ishii, T. (2001). RAPD, RFLP and SSLP analyses of phylogenetic relationships between cultivated and wild species of rice. *Genes & Genetic Systems* **76,** 71-79.

Bioversity International, International Rice Research Institute and Africa Rice Center. (2007). Descriptors for wild and cultivated rice (*Oryza* spp.), Bioversity International, Rome, Italy; International Rice Research Institute, Los Baños, Philippines; WARDA, Africa Rice Center, Cotonou, Benin.

Cai, H. W., Wang, X. K. & Morishima, H. (2004). Comparison of population genetic structures of common wild rice (*Oryza rufipogon* Griff.), as revealed by analyses of quantitative traits, allozymes, and RFLPs. *Heredity* **92,** 409-417.

Ding, C. & He, X. (2004). K-means Clustering via Principal Component Analysis. *In* "Proceedings of the 21st International Machine Learning Conference" (R. Greiner, and D. Schuurmans, , 2004, Eds.), pp. 225-232. ACM Press.

Duan, S., Lu, B., Li, Z., Tong, J., Kong, J., Yao, W., Li, S., Zhu, Y. (2007). Phylogenetic analysis of AA-genome *Oryza* species (Poaceae) based on chloroplast, mitochondrial, and nuclear DNA sequences. *Biochemical Genetics* **45,** 113-129.

Fedorenko, O. M., Gritskikh, M. V., Malysheva, I. E. & Nikolaevskaya, T. S. (2009). Genetic diversity of insular natural populations of *Festuca pratensis* Huds.: RAPD Analysis. *Russian Journal of Genetics* **45,** 1134-1138.

Furrer, R., Nychka, D. & Sain, S. (2010). fields: Tools for spatial data. R package version 6.3.

Helenurm, K. (2001). High levels of genetic polymorphism in the insular endemic herb *Jepsonia malvifolia*. *Journal of Heredity* **92,** 427-432.

Helenurm, K., West, R. & Burckhalter, S. J. (2005). Allozyme variation in the endangered insular endemic *Castilleja grisea*. *Annals of Botany* **95,** 1221-1227.

Hijmans, R. J., Guarino, L., Cruz, M. & Rojas, E. (2001). Computer tools for spatial analysis of plant genetic resources data: 1. DIVA-GIS. *Plant Genetic Resources Newsletter* **124,** 15-19.

Hiroi, K., Al Mamun, A., Wada, T. & Takeoka, Y. (1990). A study on the interspecific variation of panicle structure in the genus *Oryza*. *Japan Journal of Crop Science* **59,** 815-823.

Howcroft, N. & Davidson, J. (1973). An international provenance trial of *Pinus merkusii* in Papua New Guinea. *Tropical Forestry Research Note* **15,** 27.

Iwamoto, M., Nagashima, H., Nagamine, T., Higo, H. & Higo, K. (1999). A Tourist element in the flanking region of the catalase gene CatA reveals evolutionary relationships among *Oryza* species with various genome types. *Molecular Genetics and Genomics* **262,** 493-500.

Juliano, A. B., Naredo, M. E., Lu, B.-R. & Jackson, M. T. (2005). Genetic differentiation in *Oryza meridionalis* Ng based on molecular and crossability analyses. *Genetic Resources and Crop Evolution* **52,** 435-445.

Kariali, E., Rani Kuanar, S. & Kumar Mohapatra, P. (2008). Individual tiller dynamics of two wild *Oryza* species in contrasting habitats. *Plant Production Science* **11,** 355-360.

Kuroda, Y., Sato, Y.-I., Bounphanousay, C., Kono, Y. & Tanaka, K. (2007). Genetic structure of three *Oryza* AA genome species (*O. rufipogon, O. nivara* and *O. sativa*) as assessed by SSR analysis on the Vientiane Plain of Laos. *Conservation Genetics* **8,** 149-158.

Kwon, S.-J., Lee, J. K., Hong, S.-W., Park, Y.-J., McNally, K. L. & Kim, N.-S. (2006). Genetic diversity and phylogenetic relationship in AA *Oryza* species as revealed by Rim2/Hipa CACTA transposon display. *Genes & Genetic Systems* **81,** 93-101.

Lu, B.-R. (1999). Taxonomy of the genus *Oryza* (Poaceae): historical perspective and current status. *International Rice Research Notes* **24,** 4-8.

Lu, B.-R., Ge, S., Sang, T., Chen, J.-K. & Hong, D.-Y. (2001). The current taxonomy and perplexity of the genus *Oryza* (Poaceae). *Acta Phytotaxonomica Sinica* **39,** 373-388.

Maechler, M., Rousseeuw, P., Struyf, A. & Hubert, M. (2005). *Cluster Analysis Basics and Extensions*; unpublished.

Mendiburu de, F. (2010). *Agricolae: Statistical Procedures for Agricultural Research*. R package version 1.0-9.

Moraes, E. M., Abreu, A. G., Andrade, S. C. S., Sene, F. M. & Solferini, V. N. (2005). Population genetic structure of two columnar cacti with a patchy distribution in eastern Brazil. *Genetica* **125,** 311-323.

Morishima, H., Sano, Y. & Oka, H. I. (1980). *Observations on wild and cultivated rices and companion weeds in the hilly areas of Nepal, India and Thailand.* National Institute of Genetics, Misima, Japan.

Morishima, H., Sano, Y. & Oka, H. I. (1984). Differentiation of perennial and annual types due to habitat conditions in the wild rice *Oryza perennis*. *Plant Systematics and Evolution* **144,** 119-135.

Morishima, H., Sano, Y. & Oka, H. I. (1992). Evolutionary studies in cultivated rice and its wild relatives. *Oxford Surveys in Evolutionary Biology* **8,** 138-184.

Ng, N. Q., Chang, T. T., Williams, J. T. & Hawkes, J. G. (1981a). Morphological studies of Asian rice and its related wild species and the recognition of a new Australian taxon. *Biological Journal of the Linnean Society* **16,** 303-313.

Ng, N. Q., Hawkes, J. G., Williams, J. T. & Chang, T. T. (1981b). The recognition of a new species of rice (Oryza) from Australia. . *Botanical Journal of the*

Linnean Society **82,** 327 - 330.

Ohtsubo, H., Cheng, C., Ohsawa, I., Tsuchimoto, S. & Ohtsubo, E. (2004). Rice retroposon p-SINE1 and origin of cultivated rice. *Breeding Science* **54,** 1-11.

Oka, H. I. (1988). *Origin of cultivated rice.* Japan Science Society Press, Tokyo.

Oka, H. I. & Chang, W. T. (1961). Hybrid swarms between wild and cultivated rice species, *Oryza perennis* and *O. sativa*. *Evolution* **15,** 418-430.

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., O'Hara, R. B., Simpson, G. L., Solymos, P., Henry, M., Stevens, H. & Wagner, H. (2010). *Vegan: Community ecology package*. R package 1.17-4.

Park, K. C., Kim, N. H., Cho, Y. S., Kang, K. H., Lee, J. K. & Kim, N. S. (2003). Genetic variations of AA genome *Oryza* species measured by MITE-AFLP. *Theoretical and Applied Genetics* **107,** 203-209.

R Development Core Team. (2009). *R: A Language and Environment for Statistical Computing* (R. F. f. S. Computing, Ed.), Vienna, Austria.

Ren, F., Lu, B.-R., Li, S., Huang, J. & Zhu, Y. (2003). A comparative study of genetic relationships among the AA-genome Oryza species using RAPD and SSR markers. *Theoretical and Applied Genetics* **108,** 113-120.

Rivera-Ocasio, E., Aide, T. M. & McMillan, W. O. (2006). The influence of spatial scale on the genetic structure of a widespread tropical wetland tree *Pterocarpus officinalis* (Fabaceae). *Conservation Genetics* **7,** 251-266.

Second, G. (1985). Evolutionary relationships in the Sativa group of *Oryza* based on isozyme data. *Genetics Selection Evolution* **17,** 89-114.

Sharma, S. D. & Shastry, S. V. S. (1965). Taxonomic studies in the genus *Oryza* L. III. *O. rufipogon* Griff. sensu stricto and *O. nivara* Sharma et Shastry *nom. nov. Indian Journal of Genetics and Plant Breeding* **25,** 157-167.

Silvertown, J. & Charlesworth, D. (2001). *Introduction to plant population biology.* Blackwell Publishing, Oxford.

Tateoka, T. (1963). Taxonomic studies of *Oryza*. III. Key to the species and their enumeration. *Botanical Magazine Tokyo* **76,** 165-173.

Vaughan, D. A., Morishima, H. & Kadowaki, K. (2003). Diversity in the *Oryza* genus. *Current Opinion in Plant Biology* **6,** 139-146.

Wang, Z. Y., Second, G. & Tanksley, S. D. (1992). Polymorphism and phylogenetic relationships among species in the genus *Oryza* as determined by analysis of nuclear RFLPs. *Theoretical and Applied Genetics* **83,** 565-581.

Xu, J.-H., Kurata, N., Akimoto, M., Ohtsubo, H. & Ohtsubo, E. (2005). Identification and characterization of Australian wild rice strains of *Oryza meridionalis* and *Oryza rufipogon* by SINE insertion polymorphism. *Genes & Genetic Systems* **80,** 129-134.

Xu, X., Liu, X., Ge, S., Jensen, J. D., Hu, F., Li, X., Dong Y., Gutenkunst R. N., Fang, L., Huang, L. Li, J., He, W., Zhang, G., Zheng, X., Zhang, F., Li, Y., Yu, C., Kristiansen, K., Zhang, X., Wang, J., Wright, M., McCouch, S., Nielsen, R., Wang, J. & Wang, W. (2012). Resequencing 50 accessions of cultivated and wild rice yields markers for identifying agronomically important genes. *Nature Biotechnology* **30,** 105–111.

Zheng, X. & Ge, S. (2010). Ecological divergence in the presence of gene flow in two closely related *Oryza* species (*Oryza rufipogon* and *O. nivara*). *Molecular Ecology* **19,** 2439-2454.

Zhou, H.-F., Zheng, X.-M., Wei, R.-X., Second, G., Vaughan, D. A. & Ge, S. (2008). Contrasting population genetic structure and gene flow between *Oryza rufipogon* and *Oryza nivara*. *Theoretical and Applied Genetics* **117,** 1181-1189.

Zhu, Q. & Ge, S. (2005). Phylogenetic relationships among A-genome species of the genus *Oryza* revealed by intron sequences of four nuclear genes. *New Phytologist* **167,** 249-265.

Appendix 1. Codes, accession numbers, and geographic origin of the studied populations.

Code[a]	Species	Country of origin	IRGC number[b]	Coordinates[c]	
				Latitude	Longitude
M1	*O. meridionalis*	Australia	86539	-12.5333	131.7167
M2	*O. meridionalis*	Indonesia	93260	-8.2835	140.4592
M3	*O. meridionalis*	Australia	105282	-12.5833	131.3333
M4	*O. meridionalis*	Australia	105294	-15.6667	145.2500
M5	*O. meridionalis*	Australia	105302	-13.5000	141.7500
N1	*O. nivara*	Bangladesh	105882	23.0166	89.4666
N2	*O. nivara*	Bangladesh	103830	24.3333	88.9000
N3[e]	*O. nivara*	Cambodia	93129	10.6333	103.7833
N4	*O. nivara*	Cambodia	105719	11.4167	104.9667
N5	*O. nivara*	Cambodia	89216	11.4467	104.4889
N6	*O. nivara*	Cambodia	89187	11.4969	104.8183
N7	*O. nivara*	Cambodia	89172	11.5094	104.8189
N8	*O. nivara*	Cambodia	89185	11.5297	104.8342
N9	*O. nivara*	Cambodia	89212	11.5317	104.8283
N10	*O. nivara*	Cambodia	106320	11.8333	104.7933
N11	*O. nivara*	Cambodia	88939	12.5333	106.6166
N12	*O. nivara*	Cambodia	92886	12.6166	103.9833
N13	*O. nivara*	Cambodia	92677	12.6500	104.9166
N14	*O. nivara*	Cambodia	92823	13.0500	104.5166
N15	*O. nivara*	Cambodia	106339	13.3833	103.8333
N16	*O. nivara*	Cambodia	92943	14.1667	103.2667
N17[d]	*O. nivara*	China	103821	23.1833	114.4500
N18	*O. nivara*	India	80549	19.6667	81.6833
N19	*O. nivara*	India	80560	19.7500	81.3333
N20	*O. nivara*	India	80601	20.0000	81.5000
N21	*O. nivara*	India	80677	21.1333	81.3333
N22	*O. nivara*	India	106101	21.7500	87.5000
N23	*O. nivara*	India	80645	22.0333	81.9167
N24	*O. nivara*	India	106051	22.0833	86.4167
N25	*O. nivara*	India	106054	22.1667	86.0000
N26	*O. nivara*	India	81837	25.5000	84.3333
N27	*O. nivara*	Laos	86699	15.4100	106.7000
N28	*O. nivara*	Laos	106148	17.9167	102.7500
N29	*O. nivara*	Laos	106151	18.2000	102.7500
N30[d]	*O. nivara*	Myanmar	106358	16.6667	96.4667
N31	*O. nivara*	Myanmar	106347	17.2500	96.6667
N32	*O. nivara*	Nepal	93192	27.5119	83.2428
N33	*O. nivara*	Nepal	93185	28.0200	81.6033
N34	*O. nivara*	Nepal	93191	28.0211	81.6011
N35	*O. nivara*	Nepal	93184	28.0283	81.6322
N36	*O. nivara*	Nepal	93195	28.1008	82.2683
N37	*O. nivara*	Sri Lanka	103422	8.9167	81.0000
N38	*O. nivara*	Thailand	105803	17.0833	104.0833

Appendix 1. (Continued) Codes, accession numbers, and geographic origin of the studied populations.

Code[a]	Species	Country of origin	IRGC number[b]	Coordinates[c]	
				Latitude	Longitude
N39	*O. nivara*	Thailand	104724	14.4667	100.1166
N40	*O. nivara*	Thailand	105765	14.3333	102.9167
N41	*O. nivara*	Thailand	105755	14.5000	102.0833
N42	*O. nivara*	Thailand	104743	14.9666	102.1166
N43	*O. nivara*	Thailand	104756	15.1000	104.3333
N44	*O. nivara*	Thailand	105859	14.6667	102.3333
N45	*O. nivara*	Thailand	104473	16.8333	100.2500
N46	*O. nivara*	Thailand	105801	17.0000	104.0833
N47	*O. nivara*	Thailand	105809	17.2500	103.7500
N48	*O. nivara*	Thailand	105825	17.6667	102.9167
N49	*O. nivara*	Thailand	105828	17.8450	102.5841
N50	*O. nivara*	Thailand	104736	18.1500	100.1333
N51	*O. nivara*	Vietnam	86496	12.5681	107.7556
N52	*O. nivara*	Vietnam	86493	13.6178	108.1171
R1	*O. rufipogon*	Bangladesh	105881	23.0666	89.3500
R2	*O. rufipogon*	Bangladesh	103827	24.1500	89.0500
R3	*O. rufipogon*	Cambodia	93085	10.6333	103.7833
R4	*O. rufipogon*	Cambodia	105720	11.4167	104.9833
R5	*O. rufipogon*	Cambodia	89228	11.4467	104.4889
R6	*O. rufipogon*	Cambodia	89230	11.4969	104.8183
R7	*O. rufipogon*	Cambodia	89223	11.5094	104.8189
R8	*O. rufipogon*	Cambodia	89227	11.5297	104.8342
R9	*O. rufipogon*	Cambodia	89232	11.5317	104.8283
R10	*O. rufipogon*	Cambodia	106321	11.8333	104.7833
R11	*O. rufipogon*	Cambodia	89007	12.5333	106.6166
R12	*O. rufipogon*	Cambodia	110408	12.6166	103.9833
R13	*O. rufipogon*	Cambodia	99538	12.6500	104.9166
R14	*O. rufipogon*	Cambodia	93063	13.0500	104.5166
R15	*O. rufipogon*	Cambodia	106335	13.3833	103.8500
R16	*O. rufipogon*	Cambodia	110409	14.1667	103.2667
R17	*O. rufipogon*	China	103823	23.2333	114.0333
R18	*O. rufipogon*	India	80550	19.6667	81.7667
R19	*O. rufipogon*	India	80562	19.7500	81.3167
R20	*O. rufipogon*	India	80600	20.0000	81.5000
R21	*O. rufipogon*	India	80680	21.1333	81.5667
R22	*O. rufipogon*	India	82983	21.7500	87.5000
R23	*O. rufipogon*	India	80643	22.0333	82.8167
R24	*O. rufipogon*	India	82982	22.0833	86.3667
R25	*O. rufipogon*	India	106055	22.1667	85.9167
R26	*O. rufipogon*	India	81881	25.5000	84.1333
R27	*O. rufipogon*	Laos	86697	15.4100	106.7000
R28	*O. rufipogon*	Laos	106149	17.9167	102.7500

Appendix 1. (Continued) Codes, accession numbers, and geographic origin of the studied populations.

Code[a]	Species	Country of origin	IRGC number[b]	Coordinates[c]	
				Latitude	Longitude
R29	*O. rufipogon*	Laos	106152	18.2000	102.7500
R30	*O. rufipogon*	Myanmar	106357	16.6667	96.5000
R31	*O. rufipogon*	Myanmar	106346	17.2500	96.6333
R32	*O. rufipogon*	Nepal	93221	27.5119	83.2439
R33	*O. rufipogon*	Nepal	93218	28.0200	81.6033
R34	*O. rufipogon*	Nepal	93220	28.0211	81.6011
R35	*O. rufipogon*	Nepal	93210	28.0283	81.6322
R36	*O. rufipogon*	Nepal	93216	28.1008	82.2683
R37	*O. rufipogon*	Sri Lanka	103423	8.5000	81.1667
R38	*O. rufipogon*	Thailand	105804	14.4667	100.1166
R39	*O. rufipogon*	Thailand	104713	14.5000	102.0833
R40	*O. rufipogon*	Thailand	105766	14.5000	102.9167
R41	*O. rufipogon*	Thailand	105758	14.6667	102.3333
R42	*O. rufipogon*	Thailand	104742	14.9666	102.1166
R43	*O. rufipogon*	Thailand	104757	15.1000	104.3333
R44	*O. rufipogon*	Thailand	105860	16.8333	100.2500
R45	*O. rufipogon*	Thailand	104474	17.0000	104.0833
R46	*O. rufipogon*	Thailand	105800	17.0833	104.0833
R47	*O. rufipogon*	Thailand	82979	17.1667	103.7500
R48	*O. rufipogon*	Thailand	105823	17.6667	102.9167
R49	*O. rufipogon*	Thailand	105829	17.8450	102.5841
R50	*O. rufipogon*	Thailand	104737	18.1500	100.1333
R51	*O. rufipogon*	Vietnam	86512	12.5681	107.7556
R52	*O. rufipogon*	Vietnam	86506	13.8333	107.8667
R53	*O. rufipogon*	Philippines	80774	7.8803	125.0061
R54	*O. rufipogon*	Indonesia	81976	-6.5569	106.7617
R55	*O. rufipogon*	Indonesia	81977	-6.1081	106.7097
R56	*O. rufipogon*	Indonesia	81978	-6.3333	106.3000
R58	*O. rufipogon*	Indonesia	105567	-0.7500	117.2667
R59	*O. rufipogon*	Indonesia	105952	-6.4167	107.0000
R60	*O. rufipogon*	Indonesia	105958	-6.0631	106.1189
R61	*O. rufipogon*	Indonesia	106452	-3.2500	104.6667
R62	*O. rufipogon*	Indonesia	106453	-3.0833	104.5000
R63	*O. rufipogon*	Australia	86542	-12.5333	131.7167
R64	*O. rufipogon*	Indonesia	93274	-8.2930	140.4089
R65[f]	*O. rufipogon*	Australia	105283	-12.5833	131.3333
R66	*O. rufipogon*	Australia	105293	-15.6667	145.2500
R67	*O. rufipogon*	Australia	105303	-13.5000	141.7500
AR1	*O. sativa*	Iran	12880		
AU1	*O. sativa*	India	32561		
IN1	*O. sativa*	Philippines	66970		
IN2	*O. sativa*	India	108921		

Appendix 1. (Continued) Codes, accession numbers, and geographic origin of the studied populations.

Code[a]	Species	Country of origin	IRGC number[b]	Coordinates[c]	
				Latitude	Longitude
J1	*O. sativa*	Philippines	328		
J2	*O. sativa*	Japan	12731		
O1	*O. officinalis*	Philippines	80780		
O2	*O. officinalis*	India	100947		

[a] Code assigned by the author

[b] IRG accession number

[c] The author georeferenced *O. meridionalis, O. nivara* and *O. rufipogon* accessions based on available location data in the International Rice Genebank Information System (IRGCIS). Figures in decimal degrees.

[d] Seeds did not germinate

[e] Currently labeled as *O. rufipogon* in IRGCIS, tentatively re-classified by the author as *O. nivara* based on seed morphology

[f] Currently labeled as *O. meridionalis* in IRGCIS, tentatively re-classified by the author as *O. rufipogon* based on seed morphology

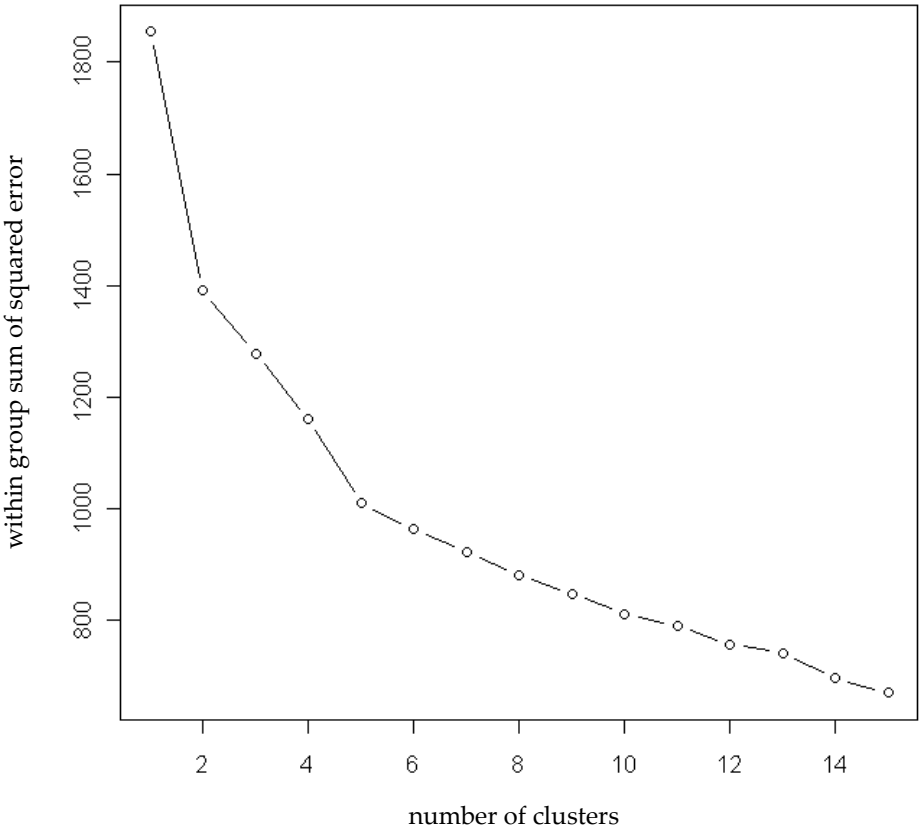
Appendix 2. Characters used in phenotyping populations of *Oryza meridionalis, O. nivara* and *O. rufipogon* (descriptions taken from Bioversity, IRRI and AfricaRice 2007).

	Character (unit)	Character code	Description
Quantitative characters			
1.	Seedling height (cm)	SDL_HT	measured from the base of the shoot to the tip of the tallest leaf blade, to the nearest centimeter (cm), at 20 days after sowing
2.	Culm number	CU_NO	total number of grain-bearing and non-bearing culms, after flowering
3.	Culm diameter (mm)	CU_DI	the outer diameter of basal portion of the main culm, at late reproductive stage
4.	Culm length (cm)	CU_LT	measured from ground level to the base of the panicle, to the nearest cm, at 7 days after anthesis
5.	Flag leaf length (cm)	FL_LT	measured from the ligule to the tip of the blade of the flag leaf, at 7 days after anthesis
6.	Flag leaf width (cm)	FL_WD	the widest portion of the flag leaf, at 7 days after anthesis
7.	Ligule length (mm)	LI_LT	measured from the base of the collar to the tip of the ligule of the penultimate leaf (i.e. highest leaf below the flag leaf), after anthesis
8.	Leaf length (cm)	LE_LT	measured from the ligule to the tip of the blade of the penultimate leaf , at 7 days after anthesis
9.	Leaf width (cm)	LE_WD	the widest portion of the penultimate leaf, at 7 days after anthesis
10.	Panicle number	PAN_NO	number of panicles per plant, at full ripening
11.	Number of basal primary branches of the panicle	NO_BPB	the number of primary panicle branches attached to the basal whorl of the panicle, after anthesis
12.	Panicle length (cm)	PAN_LT	length of main axis of panicle measured from the panicle base to the tip, at 7 days after anthesis
13.	Awn length (mm)	AW_LT	measured from the base to the tip of the awn, after anthesis
14.	Awn thickness (mm)	AW_TH	measured at 1cm from the apiculus of the spikelet, at anthesis
15.	Spikelet length (mm)	SP_LT	the distance from the base of the lowermost glume to the apiculus of the fertile lemma or palea, whichever is longer, after flowering
16.	Spikelet width (mm)	SP_WD	the distance across the fertile lemma and palea at the widest point, after flowering
17.	Sterile lemma length (mm)	STL_LT	length of longer sterile lemma, measured after flowering
18.	Sterile lemma width (mm)	STL_WD	width of longer sterile lemma measured after flowering

Appendix 2. (Continued) Characters used in phenotyping populations of *Oryza meridionalis*, *O. nivara* and *O. rufipogon* (descriptions taken from Bioversity, IRRI and AfricaRice 2007).

	Character (unit)	Character code	Description
19.	Stigma length[a] (mm)	STI_LT	measured at anthesis
20.	Style length[a] (mm)	STY_LT	measured at anthesis
21.	Anther length (mm)	AN_LT	measured at anthesis
22.	Spikelet fertility (%)	SP_FRT	percentage of filled grains over total number of spikelets
23.	Anther length: Spikelet length[a]	ANLT_SPLT	anther length to spikelet length ratio
24.	Spikelet width: Spikelet length[a]	SPWD_SPLT	spikelet width to spikelet length ratio
25.	Distance from panicle base to lowest spikelet insertion (mm)	PAN_DIS	measured at full panicle exsertion
26.	Number of days from seeding to first heading	NO_DAYS	the number of days from effective seeding date to first heading date
Qualitative characters			
27.	Flag leaf attitude	FL_AT	estimated angle of attachment between the flag leaf blade and the main panicle axis, at 7 days after anthesis
28.	Culm habit	CU_HB	estimated angle of inclination of the base of the main culm from vertical, after flowering
29.	Rhizome and stolon formation	RHST_FO	observed at flowering to maturity
30.	Attitude of panicle branches	PAN_TY	compactness of the panicle, classified according to its mode of branching, angle of primary branches, and spikelet density, at 7 days after anthesis
31.	Panicle exsertion	PAN_EX	extent to which the panicle is exserted above the flag leaf sheath, observed near maturity
32.	Culm lodging resistance	CU_LR	degree of lodging, at maturity
33.	Leaf senescence	LE_SE	observed as retention of greenness of all leaves below the flag leaf
34.	Life cycle	LI_CY	the completeness of plant growth in a growing season, observed as recovery after ratooning

[a]Additional characters, not included in the list of descriptors of Bioversity, IRRI and AfricaRice (2007)



Appendix 3. Within group sum of squared error (SSE) against the number of clusters (plot line added to help visualize trends). The decline in SSE starts to slow down markedly at k = 5.

Appendix 4. Grouping of accessions at different cluster solutions produced by the k-means clustering analysis.

Accession	Species	Geographic region	Cluster solution			
			k = 5	**k = 4**	**k = 3**	**k = 2**
M1	*O. meridionalis*	Australasia	1	1	1	1
M2	*O. meridionalis*	Australasia	1	1	1	1
M3	*O. meridionalis*	Australasia	1	1	1	1
M4	*O. meridionalis*	Australasia	1	1	1	1
M5	*O. meridionalis*	Australasia	1	1	1	1
N1	*O. nivara*	South Asia	2	1	1	1
N2	*O. nivara*	South Asia	2	1	1	1
N18	*O. nivara*	South Asia	2	1	1	1
N19	*O. nivara*	South Asia	2	1	1	1
N20	*O. nivara*	South Asia	2	1	1	1
N21	*O. nivara*	South Asia	3	2	2	1
N22	*O. nivara*	South Asia	2	1	1	1
N23	*O. nivara*	South Asia	3	2	2	1
N24	*O. nivara*	South Asia	3	2	2	1
N25	*O. nivara*	South Asia	1	2	2	1
N26A	*O. nivara*	South Asia	2	1	1	1
N26B	*O. nivara*	South Asia	2	1	1	1
N32	*O. nivara*	South Asia	2	1	1	1
N33	*O. nivara*	South Asia	2	1	1	1
N34	*O. nivara*	South Asia	2	1	1	1
N35	*O. nivara*	South Asia	2	1	1	1
N36	*O. nivara*	South Asia	2	1	1	1
N37	*O. nivara*	South Asia	2	1	1	1
N6	*O. nivara*	Continental Southeast Asia	3	2	2	1
N7	*O. nivara*	Continental Southeast Asia	3	2	2	1
N8	*O. nivara*	Continental Southeast Asia	3	2	2	1
N9	*O. nivara*	Continental Southeast Asia	3	2	2	1
N10	*O. nivara*	Continental Southeast Asia	3	2	2	1
N12	*O. nivara*	Continental Southeast Asia	3	2	2	1
N13	*O. nivara*	Continental Southeast Asia	3	2	2	1
N14	*O. nivara*	Continental Southeast Asia	3	2	2	1
N15	*O. nivara*	Continental Southeast Asia	3	2	2	1
N16	*O. nivara*	Continental Southeast Asia	3	2	2	1
N27	*O. nivara*	Continental Southeast Asia	3	2	2	1
N28	*O. nivara*	Continental Southeast Asia	3	2	2	1
N29	*O. nivara*	Continental Southeast Asia	3	2	2	1
N38	*O. nivara*	Continental Southeast Asia	3	2	2	1
N39	*O. nivara*	Continental Southeast Asia	3	2	2	1
N40	*O. nivara*	Continental Southeast Asia	3	2	2	1
N41	*O. nivara*	Continental Southeast Asia	3	2	2	1
N42	*O. nivara*	Continental Southeast Asia	3	2	2	1
N44	*O. nivara*	Continental Southeast Asia	3	2	1	1

Appendix 4. (Continued) Grouping of accessions at different cluster solutions produced by the k-means clustering analysis.

Accession	Species	Geographic region	Cluster solution			
			k = 5	**k = 4**	**k = 3**	**k = 2**
N45	*O. nivara*	Continental Southeast Asia	2	1	1	1
N46	*O. nivara*	Continental Southeast Asia	3	2	2	1
N47	*O. nivara*	Continental Southeast Asia	3	2	2	1
N48	*O. nivara*	Continental Southeast Asia	3	2	2	1
N49	*O. nivara*	Continental Southeast Asia	3	2	2	1
N50	*O. nivara*	Continental Southeast Asia	3	2	2	1
N51	*O. nivara*	Continental Southeast Asia	3	2	2	1
N52	*O. nivara*	Continental Southeast Asia	2	1	1	1
R1	*O. rufipogon*	South Asia	4	3	3	2
R2	*O. rufipogon*	South Asia	5	4	3	2
R18	*O. rufipogon*	South Asia	5	4	3	2
R19	*O. rufipogon*	South Asia	4	3	3	2
R20	*O. rufipogon*	South Asia	4	3	3	2
R21	*O. rufipogon*	South Asia	4	3	3	2
R22	*O. rufipogon*	South Asia	4	3	3	2
R24	*O. rufipogon*	South Asia	5	4	3	2
R25	*O. rufipogon*	South Asia	4	3	3	2
R26	*O. rufipogon*	South Asia	5	4	3	2
R33	*O. rufipogon*	South Asia	4	3	3	2
R34	*O. rufipogon*	South Asia	4	3	3	2
R35	*O. rufipogon*	South Asia	4	3	3	2
R36	*O. rufipogon*	South Asia	4	3	3	2
R3	*O. rufipogon*	Continental Southeast Asia	5	4	3	2
R4	*O. rufipogon*	Continental Southeast Asia	4	3	3	2
R5B	*O. rufipogon*	Continental Southeast Asia	4	3	3	2
R6	*O. rufipogon*	Continental Southeast Asia	4	3	3	2
R7	*O. rufipogon*	Continental Southeast Asia	4	3	3	2
R8	*O. rufipogon*	Continental Southeast Asia	5	4	3	2
R9	*O. rufipogon*	Continental Southeast Asia	4	3	3	2
R11	*O. rufipogon*	Continental Southeast Asia	4	3	3	2
R12	*O. rufipogon*	Continental Southeast Asia	4	3	3	2
R13	*O. rufipogon*	Continental Southeast Asia	5	4	3	2
R14	*O. rufipogon*	Continental Southeast Asia	4	3	3	2
R16	*O. rufipogon*	Continental Southeast Asia	4	3	3	2
R28	*O. rufipogon*	Continental Southeast Asia	4	3	3	2
R29B	*O. rufipogon*	Continental Southeast Asia	4	3	3	2
R30	*O. rufipogon*	Continental Southeast Asia	5	4	3	2
R31	*O. rufipogon*	Continental Southeast Asia	4	3	3	2
R38	*O. rufipogon*	Continental Southeast Asia	4	3	3	2
R39	*O. rufipogon*	Continental Southeast Asia	5	4	3	2
R40	*O. rufipogon*	Continental Southeast Asia	4	3	3	2
R41	*O. rufipogon*	Continental Southeast Asia	4	3	3	2

Appendix 4. (Continued) Grouping of accessions at different cluster solutions produced by the k-means clustering analysis.

Accession	Species	Geographic region	Cluster solution			
			k = 5	**k = 4**	**k = 3**	**k = 2**
R42	*O. rufipogon*	Continental Southeast Asia	4	3	3	2
R44	*O. rufipogon*	Continental Southeast Asia	4	3	3	2
R46	*O. rufipogon*	Continental Southeast Asia	5	4	3	2
R47	*O. rufipogon*	Continental Southeast Asia	5	4	3	2
R48	*O. rufipogon*	Continental Southeast Asia	4	3	3	2
R51	*O. rufipogon*	Continental Southeast Asia	5	4	3	2
R52	*O. rufipogon*	Continental Southeast Asia	4	3	3	2
R53	*O. rufipogon*	Insular Southeast Asia	5	4	3	2
R54	*O. rufipogon*	Insular Southeast Asia	5	4	3	2
R55	*O. rufipogon*	Insular Southeast Asia	5	4	3	2
R56	*O. rufipogon*	Insular Southeast Asia	4	3	3	2
R58	*O. rufipogon*	Insular Southeast Asia	5	4	3	2
R59	*O. rufipogon*	Insular Southeast Asia	5	4	3	2
R60	*O. rufipogon*	Insular Southeast Asia	4	3	3	2
R61	*O. rufipogon*	Insular Southeast Asia	5	4	3	2
R62	*O. rufipogon*	Insular Southeast Asia	4	3	3	2
R63	*O. rufipogon*	Australasia	5	4	3	2
R64	*O. rufipogon*	Australasia	5	4	3	2
R66	*O. rufipogon*	Australasia	5	4	3	2
R67	*O. rufipogon*	Australasia	5	4	3	2

CHAPTER 3

Local differentiation amidst extensive allele sharing in *Oryza nivara* and *O. rufipogon*

Maria Celeste N. Banaticla-Hilario[1,2], Ronald G. van den Berg[2], Nigel Ruaraidh Sackville Hamilton[1] and Kenneth L. McNally[1]

[1]*T.T. Chang Genetic Resources Center, International Rice Research Institute, Los Baños, Laguna, Philippines*
[2]*Biosystematics Group, Wageningen University and Research Center, Wageningen, The Netherlands*

Abstract

Genetic variation patterns within and between species may change along environmental gradients and at different spatial scales. This is studied in the 119 Asia Pacific *Oryza* series *Sativae* accessions genotyped with 29 SSR markers. The material includes locally sympatric populations of *Oryza nivara* and *O. rufipogon* and of *O. meridionalis* and *O. rufipogon.* Genetic similarities between *O. nivara* and *O. rufipogon* across distribution are evident in the phylogenetic and ordination results and in the large proportion of shared alleles between these taxa. At the regional scale, the two species seem more differentiated in South Asia than in Southeast Asia as revealed by a F_{ST} analysis and Mantel test. Local-level species separation is evident from Bayesian and phylogenetic analyses. The presence of strong gene flow barriers in smaller spatial units is also suggested in the AMOVA results where 64% of the genetic variation is contained among populations (as compared to 26% within populations and 10% among species).

O. nivara (H_E = 0.67) exhibits slightly lower diversity, lower correlation between genetic and geographic distances (r = 0.257, p-value = 0.000) and greater population differentiation than *O. rufipogon* (H_E = 0.70; r = 0.422, p-value = 0.000). These species also display different regional patterns of intra-specific differentiation, as isolation by distance has primarily been detected in South Asian *O. nivara* and in continental Southeast Asian *O. rufipogon*.

Bayesian inference identified four optimal population groups composed of: 1) South Asian *O. nivara*; 2) Southeast Asian *O. nivara*; 3) continental and insular Southeast Asian *O. rufipogon*; and 4) *O. meridionalis* and Australasian *O. rufipogon.* However, a more stable and phylogenetically concordant clustering solution was observed at K = 8. The eight population groups were: C1) Indian and Bangladeshi *O. nivara*; C2) Cambodian *O. nivara*; C3) Southeast Asian *O. rufipogon*; C4) *O. meridionalis*; C5) Nepalese *O. nivara*; C6) Non-Cambodian Southeast Asian *O. nivara*; C7) Australasian *O. rufipogon*; and C8) South Asian *O. rufipogon*. The Nepalese *O. nivara* appeared to be genetically isolated from all the population groups within the series while the Australasian *O. rufipogon* seemed distinct from the rest of the species. The different varietal groups of *O. sativa* cluster with different wild population groups. The aromatic and japonica cultivar groups join *O. rufipogon* from South Asia while indica and aus group with *O. nivara* from Thailand and Cambodia, respectively.

Altitude is correlated negatively with genetic diversity within population and

positively with local-scale species differentiation. Nevertheless, the modest correlations obtained suggest the possible contribution of other environmental factors in the observed variation patterns.

Introduction

The unwavering pursuit to fully understand the rice gene pool is reflected in the continuously accumulating publications on *O. rufipogon* Griff. and *O. nivara* Sharma & Shastry. These two closest relatives of Asian cultivated rice (*O. sativa* L.) are morphologically distinct (Ng. et al. 1981; Uga et al. 2003; Banaticla-Hilario et al. in Chapter 2) and isolated from each other by differences in habitat, mating system and flowering time. Their geographic distributions show overlap in tropical continental Asia with *O. rufipogon* extending southeastward to insular Southeast Asia and Australasia.

O. nivara is variously treated as a distinct species (Sharma and Shastry 1965; Ng et al. 1981; Lu 1999; Lu et al. 2001) or as an ecotype of *O. rufipogon* (Tateoka 1963; Oka 1988; Vaughan et al. 2003). Duistermaat (1987) even regards *O. nivara* as a full synonym of *O. sativa* (but this could be the result of an erroneous interpretation of the type material). This taxonomic ambivalence is also reflected in incongruences between the results of different molecular data where isozymes (Second 1985), RAPDs (Martin et al. 1997; Ren et al. 2003), allozymes and RFLPs (Cai et al. 2004), transposon display markers (Kwon et al. 2006), tourist sequences (Iwamoto et al. 1999), MITE-AFLPs (Park et al. 2003); simple sequence repeats (SSRs) (Ren et al. 2003), and various genes sequences (Zhu and Ge 2005; Zhou et al. 2008; Zheng and Ge, 2010) did not detect divergence between *O. nivara* and *O. rufipogon* while AFLPs (Aggarwal et al. 1999), SSRs (Kuroda et al. 2007), combined sequences from chloroplast, mitochondrial and nuclear DNA (Duan et al. 2007) and single nucleotide polymorphisms (SNPs) (Xu et al. 2012) did indicate a separation at species level. In this study, the annual taxon is tentatively considered as a distinct species.

Genetic differentiation between *O. nivara* and *O. rufipogon* has been examined globally using populations sampled across the species' total geographic distribution (Zheng and Ge, 2010) and at a regional scale by comparing patterns in South and Southeast Asia (Lu et al., 2008). For example, a local-scale study was conducted by Kuroda et al. (2007) using Lao populations. Yet, spatial patterns of

intra- and interspecific differentiation remain unclear for these two taxa.

The closely related taxon, *O. meridionalis* Ng, is a genetically distinct species (Xu et al. 2005; Kwon et al. 2006) that exhibits a similar life cycle, breeding habit, phenology and habitat to that of *O. nivara* but geographically overlaps only with the southern limit of *O. rufipogon*. Therefore, it could be worthwhile to compare the genetic differences between *O. meridionalis* and *O. rufipogon* with those between *O. nivara* and *O. rufipogon*.

A number of questions remain to clarify the relationship between *O. nivara* and *O. rufipogon*: 1) Are the observed genetic similarities/differences consistent along spatial gradients and across varying geographical units? 2) Are locally sympatric populations of *O. nivara* and *O. rufipogon* more differentiated than the non-sympatric ones? 3) How do geoclimatic factors (e.g., latitude, longitude, altitude, annual mean temperature and annual precipitation) influence the genetic variation within and between these *Oryza* species?

In an effort to answer these questions and uncover underlying spatial variation patterns, this study analyzes locally sympatric accession pairs (i.e., populations of different species collected from the same locality) of *O. nivara* and *O. rufipogon* from across South Asia and continental Southeast Asia and of *O. meridionalis* and *O. rufipogon* in Australasia as well as *O. rufipogon* populations from insular Southeast Asia. These three taxa, along with cultivated rice compose *Oryza* series *Sativae* in the Asia Pacific area.

For this study we use SSR markers to: 1) determine global-, regional-, and local-scale differentiation between *O. nivara* and *O. rufipogon*; 2) infer geographic population groups in Asia-Pacific *Oryza* series *Sativae*; 3) assess genetic diversity at the population-, population group- and species level; and 4) test the correlation of certain geographic and environmental factors with intra-specific (within population) genetic diversity and with species differentiation.

Materials and methods

Plant material

One hundred nineteen accessions from the International Rice Genebank (IRG) at the International Rice Research Institute (IRRI) in the Philippines were selected to

represent sympatric populations of *O. nivara* and *O. rufipogon* and of *O. meridionalis* and *O. rufipogon* across their distribution range, as well as *O. rufipogon* populations that are non-sympatric to both annual species (Appendix 1). Due to limited availability and germination issues, only one accession from China was sampled. The same plant material was used in a previous phenotyping experiment wherein some accessions were tentatively classified as intermediate forms (i.e., intermediate between two wild species or between *O. sativa* and a wild species; Appendix 1).

Six *O. sativa* accessions were also included for comparison (Appendix 1). IRGC 81837, 89228 and 106152 displayed two different plant types within the accession and were thus represented as two separate subpopulations (N26A and N26B, R5A and R5B, and R29A and R29B, respectively). The plants material was grown in the Genetic Resources Center screenhouse at IRRI. Leaf samples were harvested from five individual plants per accession.

DNA extraction

Genomic DNA was extracted from fresh leaf samples by applying the modified CTAB (cetyl trimethyl ammonium bromide) extraction protocol (Fulton et al. 1995). The DNA samples were quantified using spectrophotometry (NanoDrop™ 1000 spectrophotometer) and gel densitometry (using Lambda DNA as a standard), and then normalized to 5ng/μl concentration.

SSR genotyping

The markers used (Table 1) were from the panel of 30 standard SSR markers developed by the Generation Challenge Program for rice diversity analysis (http://gramene.org/markers/microsat/50_ssr.html). However, RM514 did not amplify well in most of the samples and was dropped from the analysis.

Polymerase chain reaction (PCR) was conducted in 20 μl reaction volume composed of 5.92 μl sterilized ultrapure water, 2 μl each of 10x $MgCl_2$ free buffer, 10mM DNTPs and 25mM $MgCl_2$ (Intron), 0.08 μl of 5U/μl *i-Taq*™ DNA polymerase (Intron), 1 μl of 1μM labelled M13 forward primer (LI-COR IRDye 700 or 800), 1 μl of 1μM M13-tailed SSR forward primer (Invitrogen), 2 μl of 1μM SSR reverse primer (Invitrogen) and 4 μl of genomic DNA. The program started with denaturation at 95°C for 2 minutes, succeeded by 32 cycles of denaturation at 95°C (30 seconds), annealing at 55°C (30 seconds) and elongation at 72°C (50 seconds), then followed by a 2 minute final extension step at 72°C. The annealing temperature was adjusted to match the optimal value for each marker as indicated in Table 1.

Table 1. Basic information and overall diversity of the 29 SSR markers used in the study.

Marker name	Chr	Taxon (germplasm)	Motif	Annealing temperature (°C)	A_N	RA_N	MA_F	PIC
RM237	1	*O. sativa*: indica (IR36)	(CT)18	55	23	17	0.1860	0.9056
RM283	1	*O. sativa*: indica (IR36)	(GA)18	61	20	13	0.2140	0.8869
RM431	1	*O. sativa*: japonica (Nipponbare)	(AG)16	55	19	11	0.1280	0.8925
RM495	1	*O. sativa*: japonica (Nipponbare)	(CTG)7	55	4	0	0.7480	0.3870
RM154	2	*O. sativa* (unknown)	(GA)21	61	24	19	0.2480	0.8416
RM452	2	*O. sativa*: japonica (Nipponbare)	(GTC)9	61	9	5	0.7500	0.3985
OSR13	3	*O. sativa* (unknown)	(GA)n	53	19	12	0.2080	0.8783
RM338	3	*O. sativa*: indica (IR36)	(CTT)6	55	4	2	0.9280	0.1293
RM124	4	*O. sativa*: japonica (Nipponbare)	(TC)10	67	9	6	0.7100	0.4477
RM161	5	*O. sativa*: japonica (Nipponbare)	(AG)20	61	17	11	0.2570	0.8442
RM413	5	*O. sativa*: japonica (Nipponbare)	(AG)11	53	22	15	0.2580	0.8745
RM507	5	*O. sativa*: japonica (Nipponbare)	(AAGA)7	55	7	4	0.7320	0.3984
RM133	6	*O. sativa*: japonica (Nipponbare)	(CT)8	63	7	3	0.4470	0.6093
RM162	6	*O. sativa* (unknown)	(AC)20	61	22	15	0.2790	0.8467
RM118	7	*O. sativa*: japonica (Nipponbare)	(GA)8	67	9	6	0.4000	0.6216
RM125	7	*O. sativa*: japonica (Nipponbare)	(GCT)8	63	13	6	0.2420	0.8362
RM455	7	*O. sativa*: japonica (Nipponbare)	(TTCT)5	57	3	1	0.8970	0.1691
RM44	8	*O. sativa*: indica (IR36)	(GA)16	53	18	7	0.1690	0.8897
RM152	8	*O. sativa* (unknown)	(GGC)10	53	9	5	0.4450	0.6453
RM408	8	*O. sativa*: japonica (Nipponbare)	(CT)13	55	11	8	0.4150	0.6948
RM447	8	*O. sativa*: japonica (Nipponbare)	(CTT)8	55	15	8	0.1610	0.8723
RM284	8	*O. sativa*: indica (IR36)	(GA)8	55	10	4	0.5990	0.5447

Table 1. (Continued) Basic information and overall diversity of the 29 SSR markers used in the study.

Marker name	Chr	Taxon (germplasm)	Motif	Annealing temperature (°C)	A_N	RA_N	MA_F	PIC
RM433	8	*O. sativa*: japonica (Nipponbare)	(AG)13	53	18	11	0.2030	0.8591
RM215	9	*O. sativa*: indica (IR36)	(CT)16	55	21	13	0.1460	0.8904
RM316	9	*O. sativa*: indica (IR36)	(GT)8-(TG)9(TTTG)4(TG)4	55	32	28	0.1900	0.9048
RM271	10	*O. sativa*: indica (IR36)	(GA)15	55	23	14	0.1590	0.9135
RM484	10	*O. sativa*: japonica (Nipponbare)	(AT)9	55	8	3	0.3080	0.7381
RM536	11	*O. sativa*: japonica (Nipponbare)	(CT)16	55	12	7	0.3540	0.7064
RM277	12	*O. sativa*: indica (IR36)	(GA)11	55	9	5	0.3510	0.6941
Mean					14	9	0.3909	0.6934

Chr – chromosome number; A_N - number of alleles; RA_N – number of rare alleles (Allele frequency ≤ 5%); MA_F - frequency of major allele; PIC - Polymorphism information content

PCR products were multiplexed by combining 2 μl each of IRDye 700- and IRDye 800-labeled samples, 5 μl sterilized nanopure water and 5 μl loading dye. Gel electrophoresis was performed on the 4300 LI-COR DNA analyzer system. The LI-COR IRDye 50–350 bp size standard ladder was used to estimate allele size. The gels were analyzed and scored with the SAGA Generation 2 software (LI-COR).

Data analyses

Genetic diversity measures were estimated for markers, populations, inferred population groups and species. The number of alleles and rare alleles (frequency ≤ 5%), frequency of major allele (allele with the highest frequency), observed heterozygosity, unbiased estimate of gene diversity (Weir 1996), polymorphism information content and inbreeding coefficient (F_{IS}) were obtained using PowerMarker 3.25 (Liu and Muse 2005). Allelic richness values were determined with FSTAT (Goudet 2001).

A phylogenetic analysis was conducted with PowerMarker 3.25 (Liu and Muse 2005). A neighbor joining (NJ) tree based on C.S. Chord 1967 distance (Cavalli-Sforza and Edwards 1967) was constructed and bootstrapping with 1000 replicates was performed. The output trees were viewed and edited with MEGA 5.05 (Tamura et al. 2011).

Principal coordinate analysis (PCoA) of C.S. Chord distance matrices between species (across distribution and within geographic regions) and within species was performed with GenAlEx 6.4 (Peakall and Smouse 2006).

Genetically distinct populations were inferred by applying the non-spatial and spatially explicit Bayesian clustering algorithms of STRUCTURE 2.2 (Pritchard et al. 2000; Falush et al. 2003) and TESS 2.3.1 (Durand et. al 2009), respectively.

A STRUCTURE model allowing for admixture and assuming correlated allele frequencies was implemented. Cluster values ranging from K = 1 to K = 15 were tested with 20 independent runs for each K with a burn in of 20000 iterations and run length of 20000 iterations per run. The 10 runs with the highest posterior probability *Ln P(D)* values were selected from each K and their average *Ln P(D)* were used in calculating the delta K (ΔK), a statistic based on the rate of changes in the likelihood distribution between successive K values (Evanno et al. 2005). The cluster value corresponding to the highest peak in the ΔK plot is considered as the appropriate cluster solution.

Since the TESS program uses spatial prior information, the geographically under-represented set of *O. sativa* accessions were excluded from the analysis. Misidentified accessions detected by NJ, PCoA and STRUCTURE analysis were also removed from the runs. Prior to analysis, the "generate spatial coordinates" option of TESS was used to create individual sample coordinates based on accession coordinates. In the TESS runs, the conditional auto-regressive (CAR) model of admixture (admixture parameter = 1.0, spatial interaction parameter = 0.6) was implemented with a linear trend surface. The maximal number of clusters was set to range from Kmax = 2 to Kmax = 10. Each Kmax was tested with 100 runs, and each run had 20000 burn in sweeps followed by another 30000 sweeps. To determine the appropriate number of TESS clusters, the deviance information criterion (DIC) should be analyzed and the stability of the bar plots should be considered (Durand et al. 2009). From each Kmax, the 10 runs with the lowest DIC were selected and their mean DIC values were plotted against Kmax. The optimum cluster solution is the Kmax value that coincides with the plateau of the DIC curve.

For each K/Kmax (from K = 2 to K = 8), the ten STRUCTURE runs with maximal Ln(P(D)) values as well as the ten TESS runs with the lowest DIC values were aligned and averaged using CLUMPP 1.1.2 (Jakobsson and Rosenberg 2007), employing the Greedy algorithm with 10000 random permutations (except for K = 8 where the LargeKGreedy algorithm with 10000 random permutations was used). The output files (bar graphs) were viewed and edited with the DISTRUCT software (Rosenberg 2004).

GenAlEx 6.4 (Peakall and Smouse 2006) was used in conducting an analysis of molecular variance (AMOVA), in estimating pairwise F_{ST} values between inferred population groups and between sympatric accessions of *O. nivara* and *O. rufipogon*, and also in testing the Mantel correlation of pairwise F_{ST} and geographic distance matrices within and between species as well as within population groups.

DIVA-GIS (Hijmans 2001) was used to obtain altitude, annual mean temperature and annual precipitation data from 5-arc minute grids provided by the WorldClim website (http://worldclim.org/). The correlation of geographic (e.g., latitude, longitude and altitude) and climatic (e.g., mean annual temperature and annual precipitation) data to allelic richness and gene diversity of accessions and to the pairwise F_{ST} values and C.S. Chord distances of sympatric *O. nivara* and *O. rufipogon* population pairs were tested using the cor.test() function of R 2.14 (R Development Core Team 2011).

Results

Overall microsatellite diversity in Asia Pacific *Oryza* series *Sativae*

Across the 119 *Oryza* series *Sativae* populations from Asia Pacific, 417 alleles (62% of which are rare) were detected at the 29 SSR loci. The number of alleles per locus varied from three (RM455) to 32 (RM316), with an average of 14. The number of rare alleles ranged from zero (RM495) to 28 (RM316), with an average of nine. The most common allele in each locus had a mean frequency of 0.39 and varied from 0.13 (RM431) to 0.93 (RM338). The PIC values differed from 0.13 (RM338) to 0.91 (RM237 and RM271) and had a mean value of 0.69 (Table 1). Based on the mentioned parameters, the most diverse loci were RM154, RM271, RM237 and RM316 while the least diverse were RM338, RM455 and RM495.

Diversity within accessions

The different genetic diversity measures computed for each accession/population are tabulated in Appendix 1. On average, *O. rufipogon* accessions are the most diverse among the species as they exhibit the highest values for all the parameters, followed by *O. nivara* accessions. *O. meridionalis* and *O. sativa* accessions display comparatively low diversity (Appendix 1).

In terms of allelic richness and gene diversity, the most diverse *O. rufipogon* accessions are R56 (Indonesia), R9 (Cambodia), R58 (Indonesia) and R3 (Cambodia) while the least diverse are R66 (Australia), R30 (Myanmar), R64 (Indonesia) and R43 (Thailand) (Appendix 1). The most diverse *O. nivara* accessions are N14 (Cambodia), N26 (India), N45 (Thailand), N21 (India), N5 (Cambodia) and N47 (Thailand) and the least diverse are N41 (Thailand), N18 (India), N2 (Bangladesh), N48 (Thailand) and N51 (Vietnam) (Appendix 1). Among the five *O. meridionalis* accessions, M2 (Irian Jaya, Indonesia) appears to be the most genetically depauperate (allelic richness = 1.0, gene diversity = 0.0) while M4 (Cooktown, Australia) seems the most diverse (Appendix 1). Among the *O. sativa* accessions, the highest gene diversity is exhibited by J1 (Philippines) and the lowest by IN2 (India) and J2 (Japan) while the highest allelic richness is displayed by AU1 (India), and the lowest by J1 (Philippines) and IN1 (Philippines).

No correlation was established between accession diversity (allelic richness and gene diversity) and location (latitude and longitude) data in both *O. rufipogon* and

O. nivara. However, allelic richness in *O. rufipogon* show weak correlation with altitude (r = -0.25, p-value < 0.05) and mean annual temperature (r = 0.25, p-value < 0.05) while in *O. nivara*, allelic richness and gene diversity both display weak correlation (p-value < 0.05) with altitude (r = -0.32 and r = -0.29, respectively).

Phylogenetic analysis

The neighbor joining tree (Figure 1) shows the general tendency of accessions from the same species to cluster together.

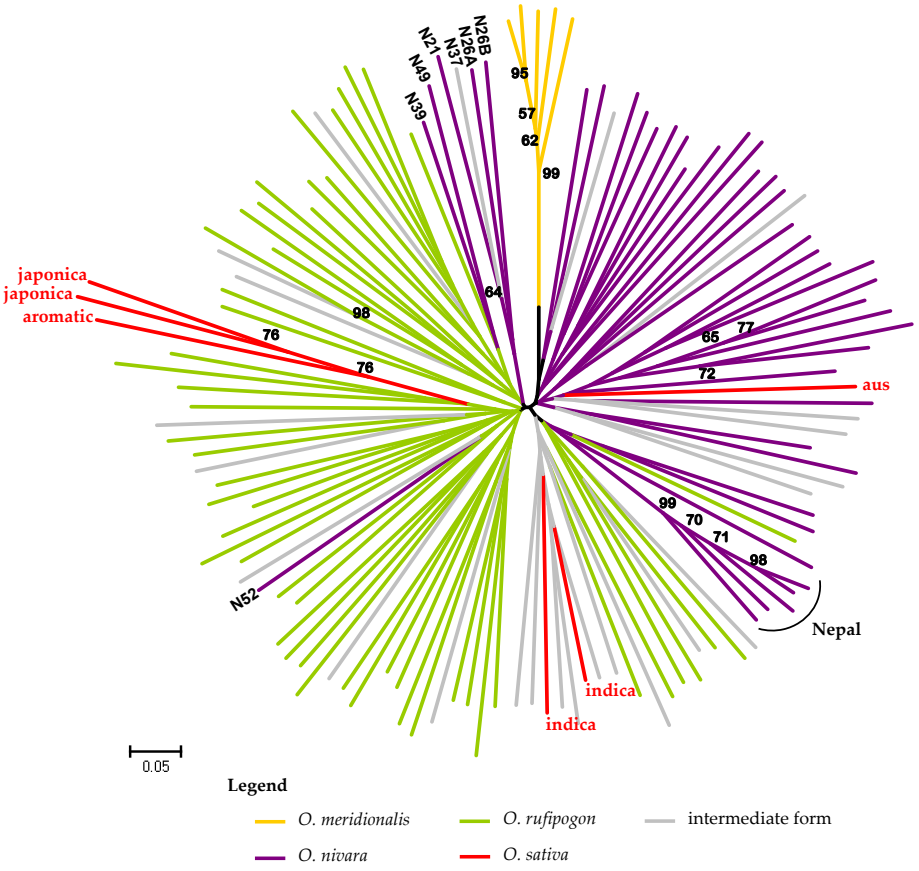


Figure 1. Neighbor joining tree of Asia Pacific *Oryza* series *Sativae* populations based on C.S. Chord 1967 distance. Bootstrap values above 50% are displayed at the root of the supported cluster. The populations are colored according to their morphological classification.

An admixed cluster composed of (mainly Nepalese) *O. nivara*, *O. rufipogon*, the two indica accessions of *O. sativa* and several phenotypically intermediate forms was also produced. However, the *O. nivara* and *O. rufipogon* clusters as well as the admixed cluster have very low bootstrap values (less than 50%) and only the *O. meridionalis* and Nepalese *O. nivara* clusters have strong bootstrap support (both with 99% bootstrap value). *O. nivara* and *O. meridionalis* join in a large cluster where the latter forms a distinct group that seems closer to the Southeast Asian populations of the former (Figure 1). The Nepalese *O. nivara* forms a separate branch within the admixed cluster. The aromatic and japonica populations of *O. sativa* group with several South Asian accessions of *O. rufipogon* while the aus population joins a cluster composed of both South and Southeast Asian *O. nivara* (Figure 1). Three *O. nivara* populations from India (N21, N26A and N26B) and an intermediate population from Sri Lanka (N37) separated from the annual species and form a cluster (Figure 1). Three *O. nivara* accessions (N39, N49 and N52) group with *O. rufipogon* while the intermediate populations are distributed among different clusters (Figure 1).

Principal Coordinate Analysis

Between the three species

The first two principal coordinate axes reflect separate but partially overlapping clusters of *O. nivara* and *O. rufipogon* (Figure 2A). *O. meridionalis* and Nepalese *O. nivara* accessions form distinct clusters isolated by axes 1 (21.83% proportion of variance) and 2 (19.10%) (Figure 2A). *O. sativa* accessions are distributed throughout the plot with aromatic and japonica populations joining *O. rufipogon*, aus and one indica accession (IN1) grouping with *O. nivara*, and the other indica accession (IN2) in the middle of the complex (Figure 2A).

The third and fourth principal coordinate axes do not separate *O. nivara* from *O. rufipogon* (Figure 2B). Axis 3 (16.43%) isolates *O. meridionalis* while axis 4 (15.52%) separates the Nepalese *O. nivara* from the rest of the taxa (Figure 2B). The succeeding principal coordinate axes displayed uninformative clustering patterns.

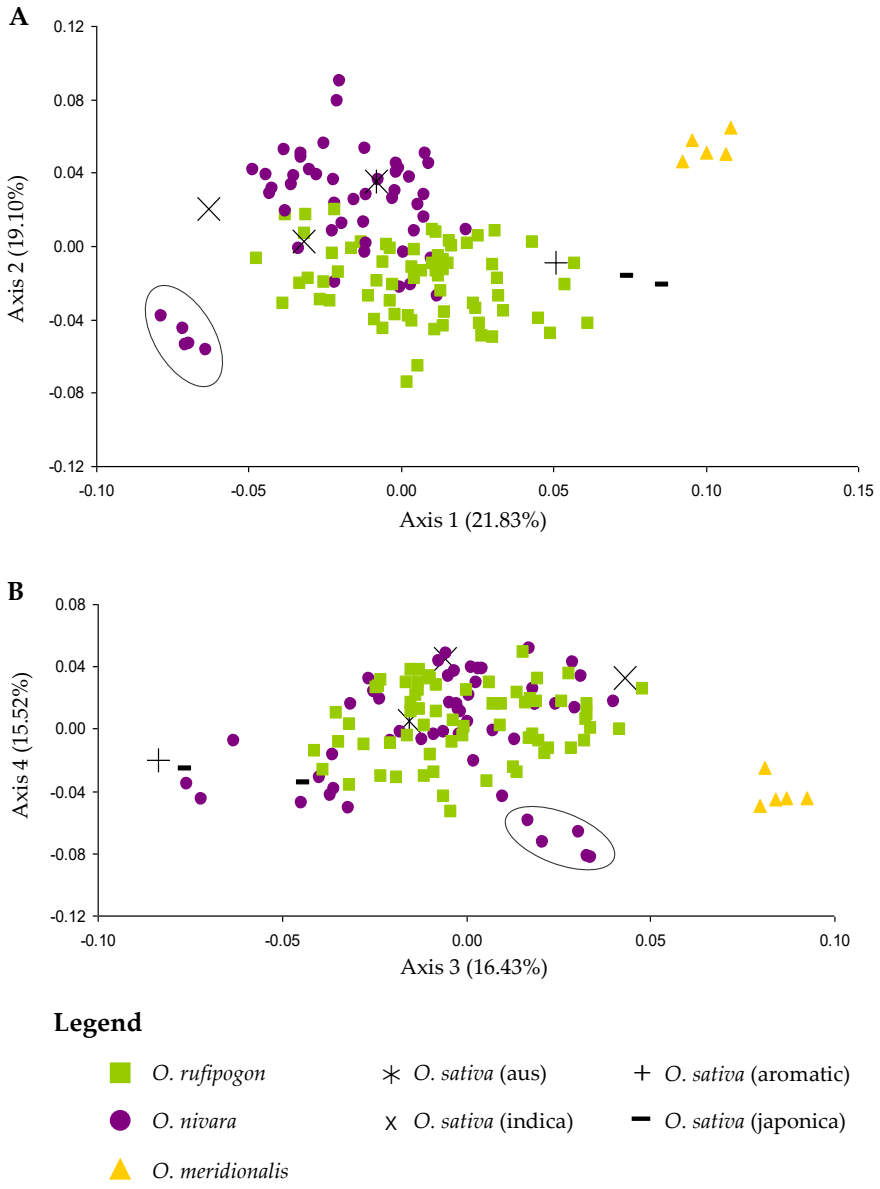


Figure 2. Principal coordinate plots revealing the genetic distinctiveness among *O. rufipogon, O. meridionalis* and *O. nivara* populations and *O. sativa* accessions. *O. nivara* accessions from Nepal are encircled. **A)** axis 1 vs. axis 2. **B)** axis 3 vs. axis 4.

Between sympatric populations from different geographic regions

The PCA plot of South Asian accessions (Figure 3A) shows three distinct clusters separated by axes 1 (27.25%) and 2 (21.64%). *O. rufipogon* is well separated from both the Nepalese and the Indian-Bangladeshi clusters of *O. nivara.* Subpopulation N26A together with populations N21 and N37 cluster with *O. rufipogon* while N26B joins the Indian-Bangladeshi *O. nivara* cluster (Figure 3A).

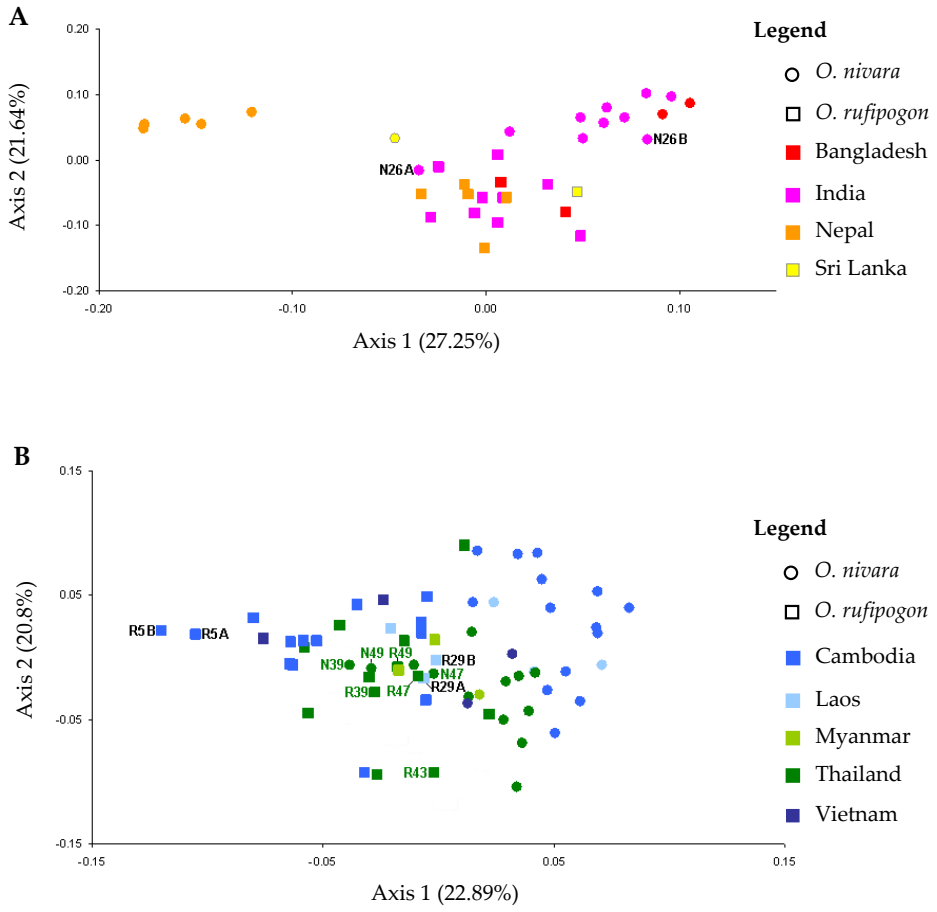


Figure 3. Principal coordinate plots depicting the clustering patterns of sympatric populations of *O. nivara* and *O. rufipogon* in South Asia (**A**) and Southeast Asia (**B**). Subpopulations within accessions as well as sympatric accession pairs exhibiting close genetic affinity are labeled.

In contrast, no distinct clusters can be observed in Southeast Asian populations (Figure 3B), although axis 1 (22.89%) displays substantial species separation while axis 2 (20.8%) shows a weak tendency of separation between the western (Thailand and Myanmar) and the eastern (Cambodia, Laos and Vietnam) populations (Figure 3B). The *O. nivara* and *O. rufipogon* components of sympatric pairs N39-R39, N47-R47 and N49-R49 display close genetic relationships (Figure 3B). *O. nivara* accessions N39 and N49 are included in the *O. rufipogon* group while *O. rufipogon* populations R40 and R50 cluster with *O. nivara* (Figure 3B). R5A and R5B cluster together as well as R29A and R29B (Figure 3B).

Within species

O. nivara exhibits geographic partitions as axis 1 (26.78%) isolates the Nepalese from the rest of the populations while axis 2 (22.07%) separates the Indian-Bangladeshi populations from the Nepalese and Southeast Asian populations (Figure 4A). Within the Southeast Asian cluster, western and eastern populations form subgroups (Figure 4A).

In *O. rufipogon,* the Australasian populations are isolated from the rest by axis 1 (22.46%) (Figure 4B). The South-Southeast Asia split is not as clear-cut as in *O. nivara,* although axis 2 (19.25%) shows a partial separation between the South and the continental Southeast Asian populations (Figure 4B).

Bayesian clustering

The clustering patterns produced by STRUCTURE and TESS from K = 2 to K = 8 are shown in Figure 5. TESS exhibits more consistent runs and produces more stable population clusters (with less fragmented members) than STRUCTURE as indicated in the average membership coefficients of each K obtained by CLUMPP (Figure 5).

Across different K values in the STRUCTURE runs, populations N39 and N49 cluster with *O. rufipogon* while R10, R43 and R50 are grouped with *O. nivara* (Figure 5). These seemingly mislabelled populations also do not cluster with their supposed species groups in the NJ (Figure 1) and PCA (Figure 3B) results and were excluded from the TESS runs.

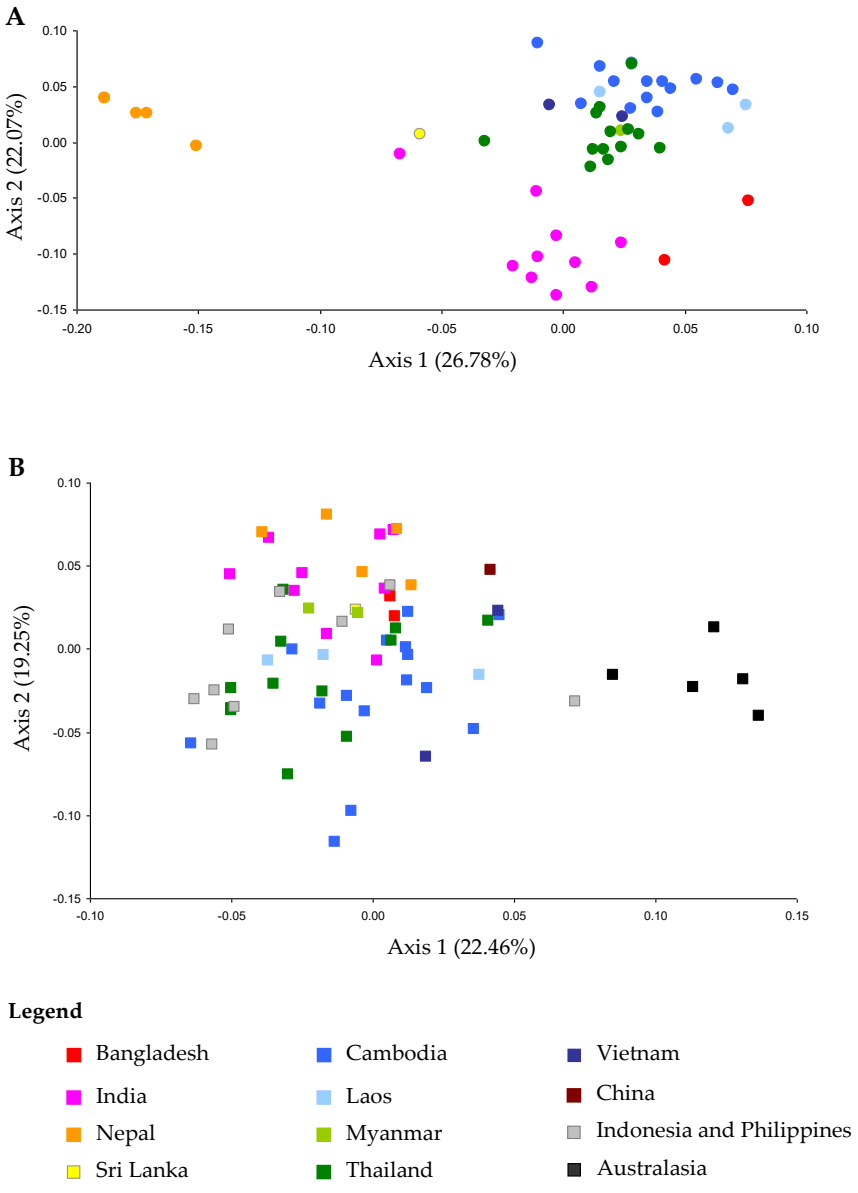


Figure 4. Principal coordinate plots showing the inter-population variation patterns in *O. nivara* (**A**) and *O. rufipogon* (**B**) across their geographic range.

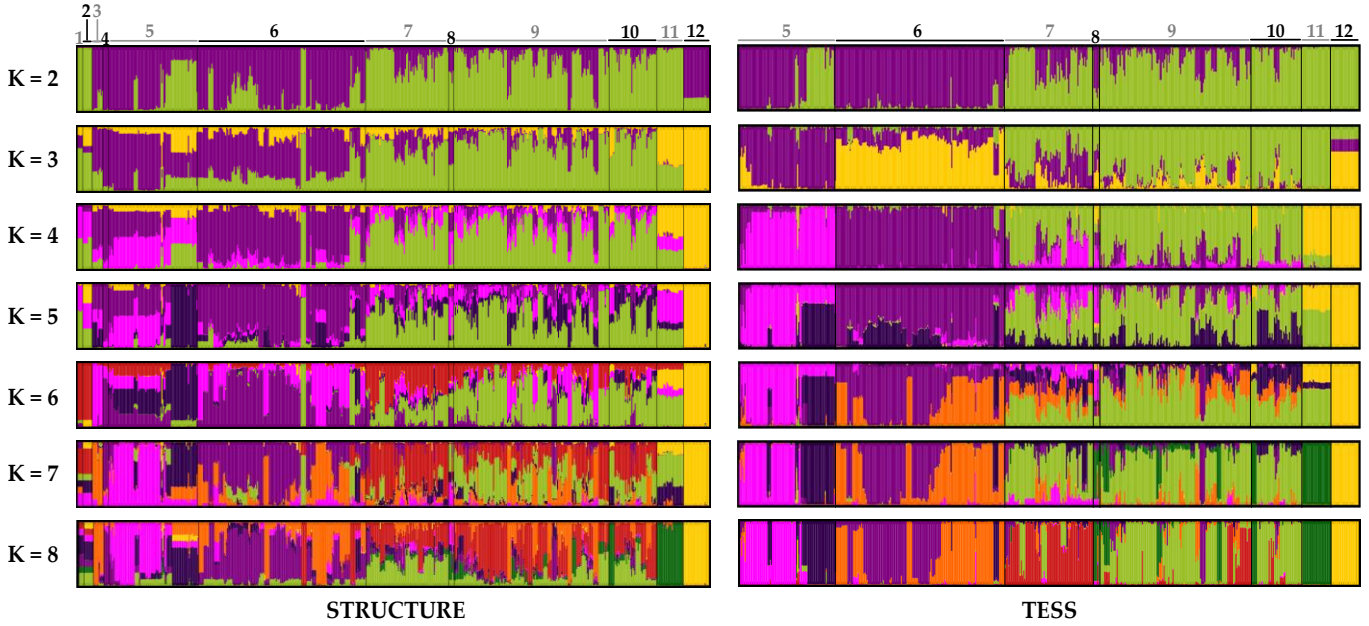


Figure 5. Cluster solutions produced by STRUCTURE and TESS from K = 2 to K = 8. The average membership coefficient of 10 runs from each K are shown. The pre-defined populations are: 1 – *O. sativa* (aromatic); 2 – *O. sativa* (japonica); 3 – *O. sativa* (indica); 4 – *O. sativa* (aus); 5 – *O. nivara* (from South Asia); 6 – *O. nivara* (from Southeast Asia); 7 – *O. rufipogon* (from South Asia); 8 – *O. rufipogon* (from China); 9 – *O. rufipogon* (from continental Southeast Asia); 10 – *O. rufipogon* (from insular Southeast Asia); 11 – *O. rufipogon* (from Australasia); 12 – *O. meridionalis* (from Australasia).

The ΔK plot of the STRUCTURE runs displays distinct peaks at K = 2 (the highest value), K = 4 and K = 6 (Figure 6A). However, K = 2 is rejected as an optimal cluster value since the cluster solution produced by STRUCTURE fails to distinguish *O. meridionalis* as a distinct population (Figure 5). The relatively stable membership coefficient plots of both STRUCTURE and TESS runs at K = 4 (Figure 5) suggest that four clusters optimally define the population structure of the dataset.

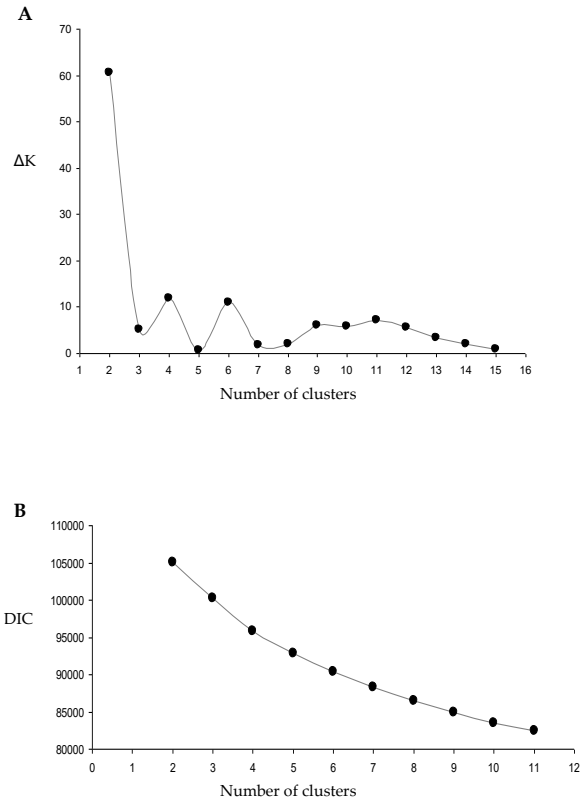


Figure 6. Criteria used in determining the appropriate cluster solution: **A**, Delta K plot of STRUCTURE runs based on *Ln P(D)* values; and **B**, Plot of the mean DIC value of each cluster solution (from TESS). Plot lines were added to help visualize trends.

The four population groups depicted similarly by STRUCTURE and TESS are: C1) South Asian *O. nivara*; C2) Southeast Asian *O. nivara*; C3) Asian *O. rufipogon* (including the continental Asians and insular Southeast Asians); and C4) *O.*

meridionalis and Australasian *O. rufipogon* (Figures 7 and 8). However, at K = 4, certain populations are occasionally swapped between different clusters across the ten STRUCTURE runs (Appendix 2). Australasian *O. rufipogon* frequently joins *O. meridionalis* but also groups with either the rest of *O. rufipogon* or the japonica group of *O. sativa* in the other runs (Appendix 2). The Nepalese and the Indian-Bangladeshi populations of *O. nivara* cluster together once, but in the rest of the runs, the former groups with either *O. rufipogon* or *O. meridionalis* while the latter joins either the Southeast Asian *O. nivara* or *O. meridionalis* (Appendix 2). The ten TESS runs show more consistent population clustering at K = 4, with 8 runs grouping the Australasian *O. rufipogon* with *O. meridionalis* and 9 runs splitting the South Asian and Southeast Asian populations of *O. nivara* (Appendix 3B).

At K = 6, the recognized groups are (Appendix 2; Figure 7): an *O. meridionalis* cluster (in 70% of the runs); two clusters in *O. rufipogon* (a South Asian group joined by the aromatic and japonica accessions of *O. sativa* and a Southeast Asian cluster in 40% of the runs); and three clusters in *O. nivara* (one cluster is predominantly Cambodian and groups with *O. sativa* aus, another cluster is mainly Nepalese and the third cluster comprises the rest of *O. nivara* and is grouped with *O. sativa* indica, in 20% of the runs). Nevertheless, the output of the six-cluster solution of STRUCTURE (and even TESS) seems unstable such that certain populations (particularly the Australasian *O. rufipogon* and the non-Nepalese South Asian *O. nivara*) appears fragmented and/or are swapped between different clusters (Figure 5; Appendices 2 and 3C).

The DIC plot of the TESS runs does not exhibit a well-defined plateau as the DIC values continuously decrease at higher Kmax (Figure 6B). The ten aligned runs in each Kmax are presented in Appendices 3A – 3E. K = 8 shows the most consistent grouping of populations across the ten runs (Figure 5; Appendix 3D). Moreover, higher Kmax values (K = 9 and K = 10), display less stable clustering and do not recognize additional distinct population clusters aside from the groups inferred at K = 8 (Appendices 3D – 3E). This indicates that the eight-cluster solution fits the lower population structure level of the dataset. K = 4 and K = 7 also produce stable bar plots (Figure 5; Appendices 3B – 3C) and will be discussed for comparison purposes. The clustering pattern of K = 4 was discussed previously with the STRUCTURE results. At K = 7, the inferred groups are: C1) Indian and Bangladeshi *O. nivara*; C2) Cambodian *O. nivara*; C3) continental and insular Asian *O. rufipogon*; C4) *O. meridionalis*; C5) Nepalese *O. nivara*; C6) non-Cambodian *O. nivara*; and C7) Australasian *O. rufipogon* (Figures 5 and 8). At K = 8, the same population groups are recognized except for C3 (Asian *O. rufipogon*) that is split into the Southeast

Asian *O. rufipogon* (C3 of K = 8) and South Asian *O. rufipogon* (C8 of K = 8) clusters (Figures 5 and 8).

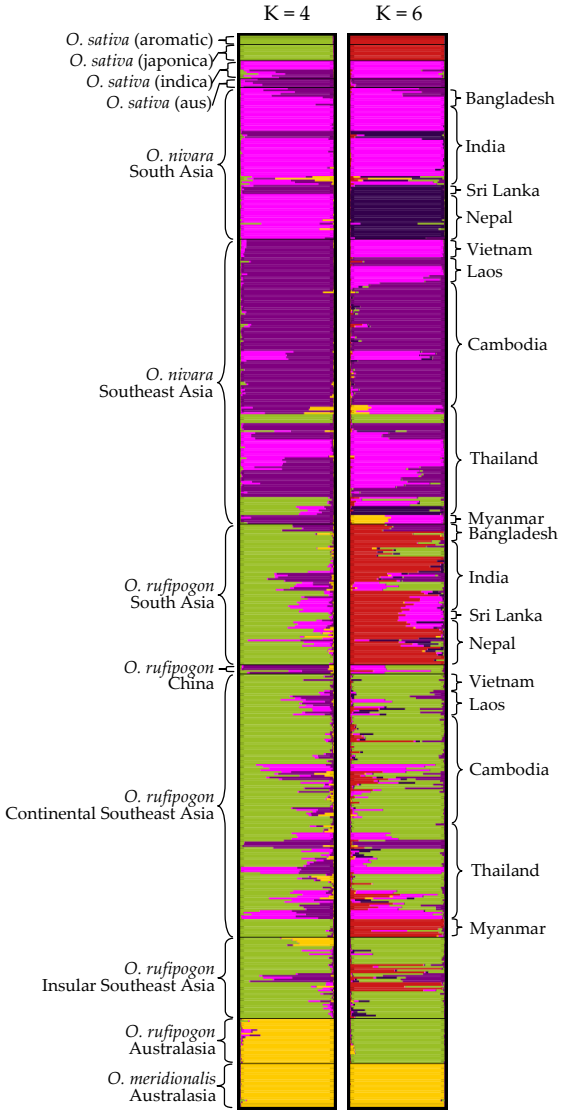


Figure 7. Population clusters of STRUCTURE at K = 4 and K =6 (based on the modal clustering pattern). The pre-defined population assignments are on the left hand side and the geographic origin of sympatric populations are on the right hand side of the figure.

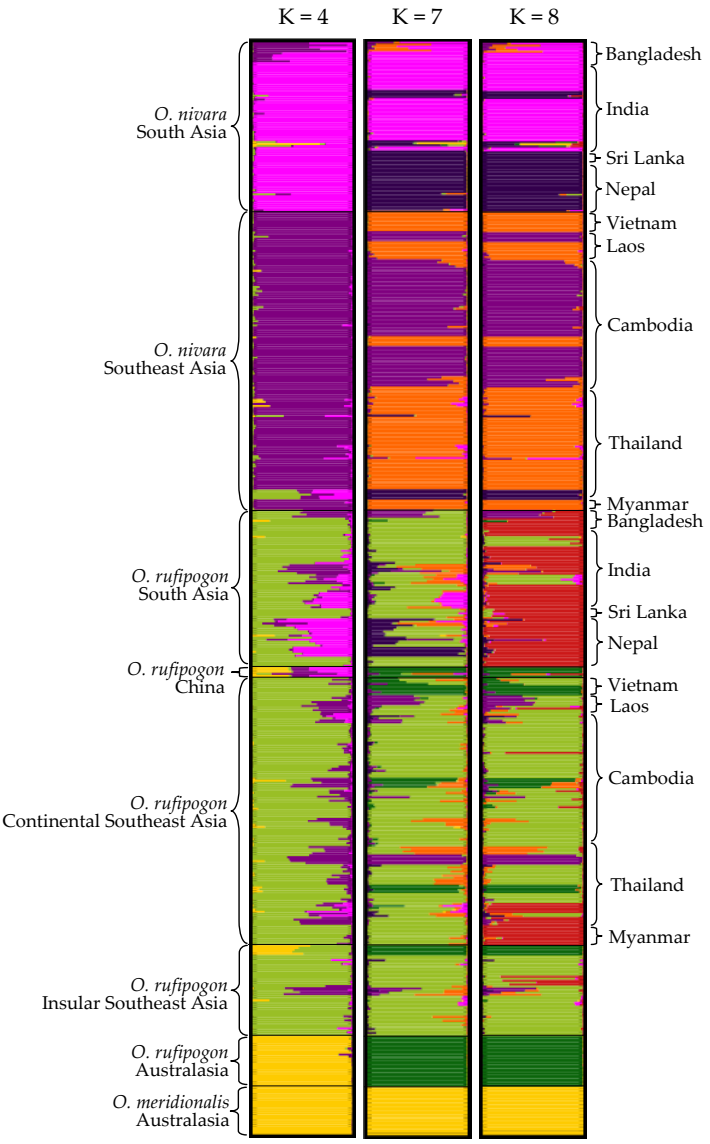


Figure 8. Population clusters of TESS at K = 4, K = 7 and K = 8 (based on the modal clustering pattern). The pre-defined population assignments are on the left hand side and the geographic origin of sympatric populations are on the right hand side of the figure.

The relationships between population clusters at K = 4, K = 7 and K = 8 are depicted by the PCA plots in Figures 9 and 10.

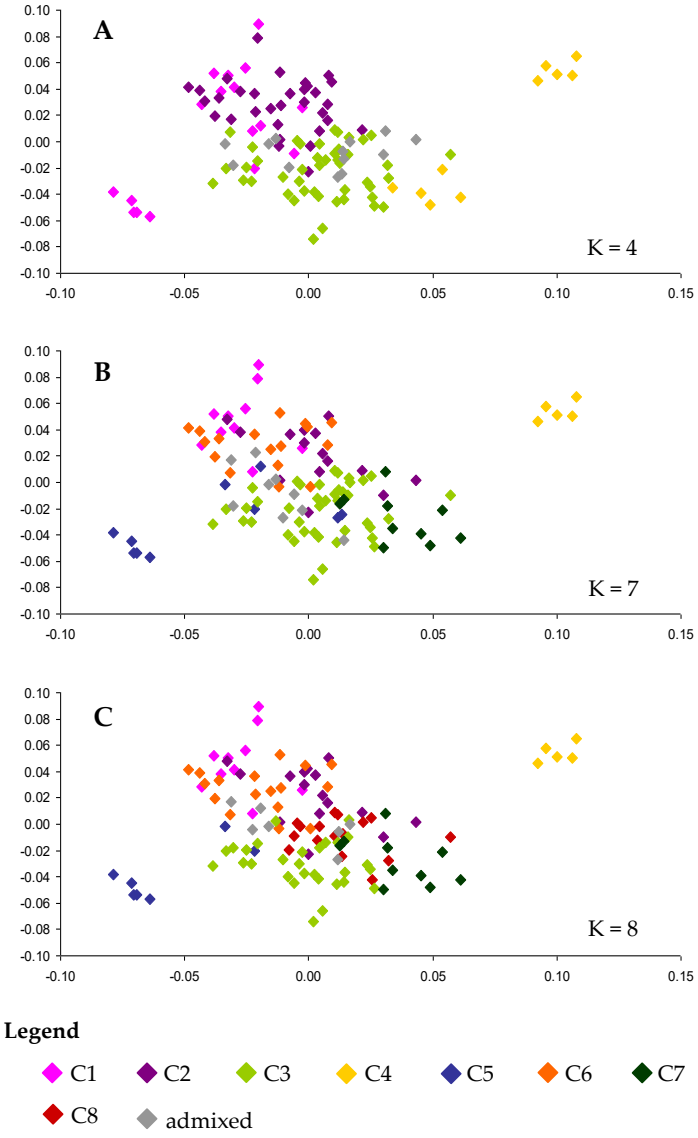


Figure 9. Placement of TESS clusters in the plot of the first two principal coordinate axes (axis 1 on x and axis 2 on y). **A)** at K = 4, C1 - South Asian *O. nivara*; C2 - Southeast Asian *O. nivara*; C3 – continental Asian and insular Southeast Asian *O. rufipogon*; and C4 - *O. meridionalis* and Australasian *O. rufipogon*. **B)** at K = 7, C1 - Indian and Bangladeshi *O. nivara*; C2 - Cambodian *O. nivara*; C3 – continental Asian and insular Southeast Asian *O. rufipogon*; C4 - *O. meridionalis*; C5 - Nepalese *O. nivara*; C6 - non-Cambodian *O. nivara*; and C7 - Australasian *O. rufipogon*. **C)** at K = 8, C1 - Indian and Bangladeshi *O. nivara*; C2 - Cambodian *O. nivara*; C3 – Southeast Asian *O. rufipogon*; C4 - *O. meridionalis*; C5 - Nepalese *O. nivara*; C6 - non-Cambodian *O. nivara*; C7 - Australasian *O. rufipogon*; and C8 – South Asian *O. rufipogon*. Admixed populations are those with less than 0.6 membership coefficient.

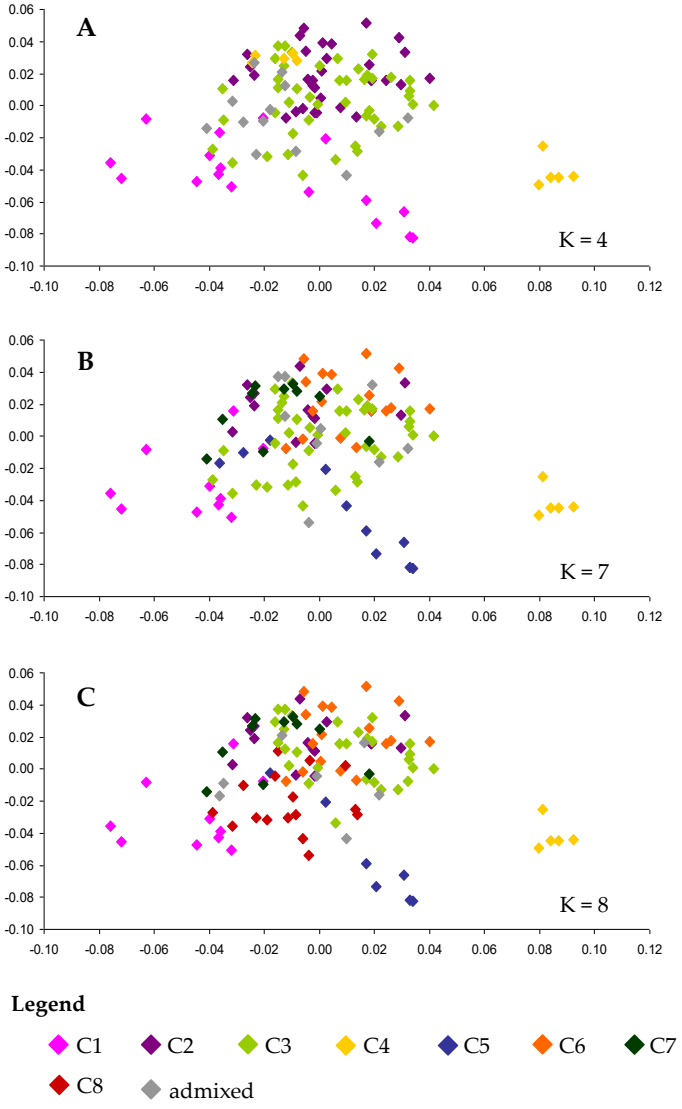


Figure 10. Placement of TESS clusters in the plot of the third and fourth principal coordinate axes (axis 3 on x and axis 4 on y). **A)** at K = 4, C1 - South Asian *O. nivara*; C2 - Southeast Asian *O. nivara*; C3 – continental Asian and insular Southeast Asian *O. rufipogon*; and C4 - *O. meridionalis* and Australasian *O. rufipogon*. **B)** at K = 7, C1 - Indian and Bangladeshi *O. nivara*; C2 - Cambodian *O. nivara*; C3 – continental Asian and insular Southeast Asian *O. rufipogon*; C4 - *O. meridionalis*; C5 - Nepalese *O. nivara*; C6 - Non-Cambodian *O. nivara*; and C7 - Australasian *O. rufipogon*. **C)** at K = 8, C1 - Indian and Bangladeshi *O. nivara*; C2 - Cambodian *O. nivara*; C3 – Southeast Asian *O. rufipogon*; C4 - *O. meridionalis*; C5 - Nepalese *O. nivara*; C6 - Non-Cambodian *O. nivara*; C7 - Australasian *O. rufipogon*; and C8 – South Asian *O. rufipogon*. Admixed populations are those with less than 0.6 membership coefficient.

The fourth inferred group (C4 – Australasian *O. rufipogon* and *O. meridionalis*) at K = 4 seems dubious since the genetic uniqueness of *O. meridionalis* detected by PCA is not recognized in the said cluster solution (Figures 9A and 10A). At K = 7 and K = 8, C1, C2 and C6 form the *O. nivara* cluster and C3, C7 and C8 (for K = 8) comprises the *O. rufipogon* group (Figures 9B – 9C) while C4 (*O. meridionalis*) and C5 (Nepalese *O. nivara*) form distinct clusters (Figures 9B – 9C and 10B – 10C). Principal coordinate axis 1 separates C7 (Australasian *O. rufipogon*) from the rest of *O. rufipogon* at K = 7 and K = 8 (Figures 9B – 9C) while axis 4 separates the South Asian clusters (C1, C5, C8) from the Southeast Asian groups (C2, C3, C6) at K = 8 (Figure 10C).

Appendix 4 shows the cluster membership of populations at K = 8. Following Garris et al. (2005) and Agrama et al. (2010), a cut-off membership coefficient of ≥0.6 was imposed to assign each population to a cluster. The cluster size and proportion of absolute (membership coefficient = 1) and slightly admixed (membership coefficient ranging from 0.6 to 0.9) members in each population cluster are displayed in Figure 11. C4 exhibits 100% absolute membership while the rest of the population groups contain slightly admixed members varying from 10% (C1) to 56% (C8). Seven populations were not assigned to any cluster. They are: N21 (intermediate between C1 and C5), N26A (admixture of C5, C8, C7 and C4), R22 (admixture of C2, C6 and C8), R29A (admixture of C1, C2, C6, C8 and C3), R49 (admixture of C6, C8 and C3), R56 (intermediate between C8 and C3) and R58 (admixture of C2, C6 and C3) (Appendix 4).

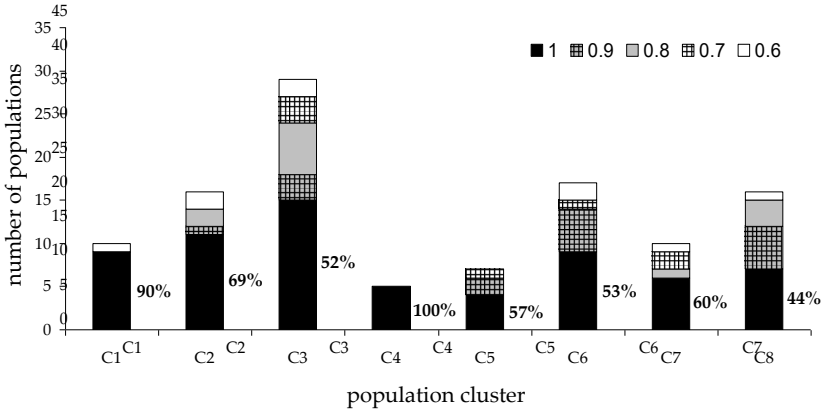


Figure 11. The number of populations and proportion of absolute (membership coefficient = 1) and slightly admixed (membership coefficient ranging from 0.6 to 0.9) members in each TESS cluster at K = 8. The percentages of absolute membership in each group are displayed.

The geographic subdivisions in *O. nivara* and *O. rufipogon* are illustrated in the distribution map of the eight population clusters (Figure 12), where the local separation of the two species across their range (except for three *O. rufipogon* accessions R1, R39 and R40 that grouped with *O. nivara* clusters) is also depicted.

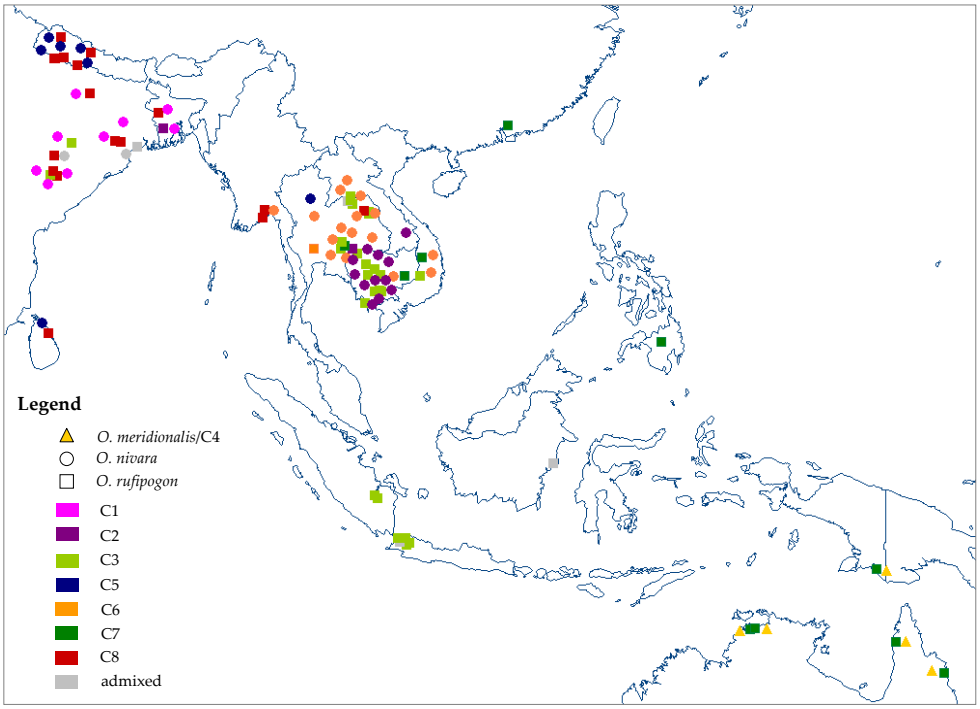


Figure 12. Distribution map of the eight population groups of Asia Pacific *Oryza* series *Sativae* inferred by TESS. The geographic coordinates were slightly adjusted to allow better graphical representation of overlapping accessions.

Genetic diversity of species and population groups

The inbreeding coefficient (F_{IS}) and genetic diversity estimates (from PowerMarker and FSTAT analyses) of species and population groups are presented in Table 2. The annual species and population groups (*O. meridionalis* and *O. nivara*) exhibit higher F_{IS} values (0.94 – 0.98) than the perennial taxa (0.85 – 0.87). Among the three species, *O. rufipogon* contains the largest genetic variation as it displays the highest

values in all diversity parameters. In contrast, *O. meridionalis* has the lowest values, rendering it the least diverse species and population group (Table 2). This is indicative of a genetic bottleneck probably caused by the geographic isolation of this Australasian species. However, a larger number of accessions and multiple individuals from the accessions should be analyzed to confirm this conclusion.

Table 2. Microsatellite diversity of species and population groups in Asia Pacific *Oryza* series *Sativae* (based on 29 SSR loci).

Taxon (TESS cluster number)	S	A_L	RA_L	R_S	H_O	H_E	PIC	F_{IS}
Species								
O. nivara	229	9.83	5.03	7.47	0.02	0.67	0.65	0.97
O. rufipogon	226	11.79	6.72	8.43	0.09	0.70	0.68	0.87
O. meridionalis (C4)	24	2.24	0.21	2.24	0.01	0.24	0.23	0.98
Population clusters								
Indian and Bangladeshi *O. nivara* (C1)	46	4.28	0.86	4.06	0.02	0.52	0.48	0.97
Nepalese *O. nivara* (C5)	35	3.59	0.76	3.47	0.02	0.39	0.37	0.94
Cambodian *O. nivara* (C2)	72	6.55	2.31	5.70	0.03	0.59	0.56	0.95
Non-Cambodian *O. nivara* (C6)	73	6.48	1.59	5.88	0.02	0.61	0.59	0.97
South Asian *O. rufipogon* (C8)	69	7.97	3.31	6.85	0.10	0.64	0.62	0.86
Southeast Asian *O. rufipogon* (C3)	122	9.48	5.07	7.36	0.10	0.66	0.64	0.85
Australasian *O. rufipogon* (C7)	44	5.38	1.41	5.04	0.07	0.59	0.56	0.88

S – number of samples; A_L – mean number of alleles per locus; RA_L – mean number of rare alleles per locus; R_S – allelic richness; H_O – observed heterozygosity; H_E - gene diversity (unbiased estimate); PIC – polymorphism information content; F_{IS} – inbreeding coefficient

Based on allelic richness and gene diversity, C3 (Southeast Asian *O. rufipogon*) is the most diverse among the population groups, followed by C8 (South Asian *O. rufipogon*) (Table 2). The genetic variation in Southeast Asian *O. nivara* clusters C2 and C6 is comparable to that of C7 (Australasian *O. rufipogon*) and greater than those of South Asian *O. nivara* clusters C1 and C5. Next to C4 (*O. meridionalis*), C5

(Nepalese *O. nivara*) shows the least diversity among the population groups. Heterozygosity is greater in the *O. rufipogon* clusters C3, C7 and C8 (0.07 – 0.1) than in the rest of the population groups (0.01 – 0.03). Clusters with the highest proportion of rare alleles are C3 (53.5%), C8 (41.6%) and C2 (35.3%) (Table 2).

Unique and shared alleles

The proportions of unique and shared alleles in each species are shown in Figure 13 while the allele frequencies in each population group are presented in Appendix 5. Forty seven (47) alleles are common to the three species. *O. meridionalis* share two alleles with *O. nivara* and five alleles with *O. rufipogon*. In stark contrast, *O. nivara* and *O. rufipogon* share 192 alleles making up more than half of the total alleles detected in the annual (68.6%) and perennial (56.1%) taxa (Figure 13). Of the 192 alleles, 14 are exclusively present in Southeast Asian populations (C2, C3 and C6), only one allele is endemic to South Asian populations (C1 and C8), while the remaining 177 alleles are not restricted to regionally sympatric populations (Appendix 5).

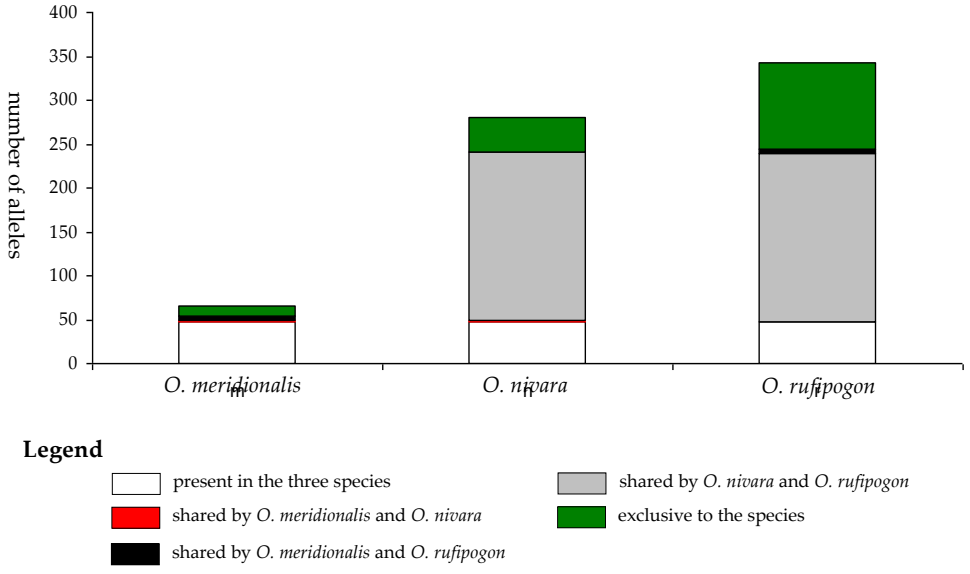


Figure 13. The proportion of private and shared alleles in *O. meridionalis*, *O. nivara* and *O. rufipogon* in Asia Pacific.

Figure 14 displays the proportion of shared alleles found in locally sympatric and non-sympatric population pairs. Sixteen out of the 52 alleles (31%) shared by *O. meridionalis* and *O. rufipogon* are detected in at least one of the five sympatric population pairs while the remaining 36 alleles (69%) are shared by non-sympatric populations (Figure 14). Among the 239 shared alleles of *O. nivara* and *O. rufipogon*, 98 (41%) are found in sympatric populations (82 and 16 of which are detected in ≤20% and >20% of the sympatric population pairs, respectively) and 141 are found in non-sympatric populations (Figure 14).

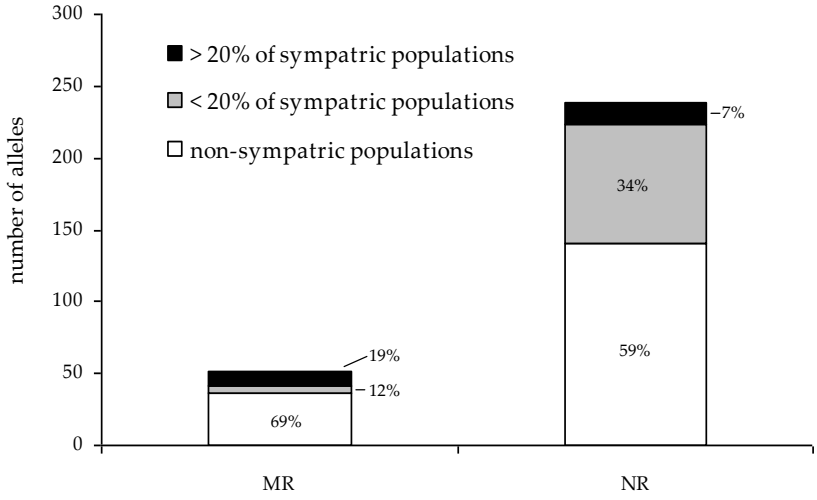


Figure 14. Proportion of shared alleles detected in sympatric and non-sympatric population pairs of *O. meridionalis* and *O. rufipogon* (MR) and of *O. nivara* and *O. rufipogon* (NR).

O. rufipogon has the largest proportion of unique alleles (98 alleles, 28.7%), followed by *O. meridionalis* (11 alleles, 16.9%), and *O. nivara* (39 alleles, 13.9%) (Appendix 5; Figure 13). Among the population groups, the highest numbers of unique alleles are detected in the *O. rufipogon* clusters C3 (29), C8 (25) and C7 (12). Southeast Asian C3 exclusively shares 11 alleles with South Asian C8 and 12 alleles with Australasian C7. Three alleles are present in both C8 and C7 while another six alleles are unique to the three *O. rufipogon* groups. The *O. nivara* clusters has fewer unique alleles (C1 - 7; C2 – 9; C5 – 3; C6 - 9). The Indian-Bangladeshi cluster C1 exclusively shares three alleles each with Southeast Asian groups C2 and C6. One allele is endemic to C2, C6 and the Nepalese cluster C5. Four alleles are unique to C2 and C6 while no alleles are exclusively shared by the South Asian

clusters C1 and C5 (Appendix 5).

The most highly discriminating markers for *O. meridionalis* are RM124, RM316 and RM413 as they distinguish all accessions of the Australasian species from the rest of Asia Pacific *Oryza* series *Sativae* populations (Appendix 5). RM44, RM431, RM118 and RM161 discriminate 12.5 – 20.8% of *O. meridionalis* populations while RM237 and RM433 distinguish less than 5% of the taxon (Appendix 5). Certain alleles of RM154, RM413, RM44, RM433 and RM495 are found exclusively in all geographic populations of *O. rufipogon* but in limited frequencies ranging from 0.007 to 0.432 (Appendix 5). RM118 differentiates 47% of Australasian *O. rufipogon* and at least one allele from each of the 26 loci (RM277, RM455 and RM536 are not included) discriminates a small proportion (allele frequencies ranging from 0.004 to 0.205) of one or two *O. rufipogon* population group/s (Appendix 5). No allele is present throughout the distribution range of *O. nivara.* The 39 unique alleles from 20 markers discriminate at least one of the four *O. nivara* population groups in frequencies ranging from 0.007 to 0.304 (Appendix 5).

Genetic differentiation

Based on the AMOVA results, genetic variation in Asia Pacific *Oryza* series *Sativae* (excluding *O. sativa*) resides mainly among accessions (explaining 64% of the total variance) and to a lower degree within accessions (26%) as well as among the three species (10%) (Table 3). Significant and moderate differentiation can be observed between accessions (Φ_{PT} = 0.74) and between species (Φ_{RT} = 0.1), respectively (both at $p < 0.001$ level).

Table 3. Analysis of molecular variance (AMOVA) among species, among populations and within populations of Asia Pacific *Oryza* series *Sativae* accessions based on 29 SSR markers. The P-values are based on 999 permutations.

Source	DF	SS	MSS	Est var	% var	P- value*
Among species	2	1449.896	724.948	4.408	10	0.001
Among populations	100	13853.252	138.533	27.102	64	0.001
Within populations	382	4205.283	11.009	11.009	26	0.001
Total	484	19508.431		42.518	100	

DF – degree of freedom; SS – sum of squares; MSS – mean sum of squares; Est var – estimated variance; % var – percentage of variance

The population clusters identified by TESS at K = 8 display different degrees of differentiation (Figure 15) with pairwise F_{ST} values ranging from 0.08 (between South (C8) and Southeast Asian (C3) *O. rufipogon*) to 0.58 (between *O. meridionalis* (C4) and Nepalese *O. nivara* (C5)). *O. meridionalis* is clearly the most distinct population group. The mean pairwise F_{ST} value between *O. nivara* clusters (0.23) is greater than between *O. rufipogon* clusters (0.13) and even between clusters of *O. nivara* and *O. rufipogon* (0.19) suggesting deep genetic divisions within the annual species (Figure 15). *O. nivara* from Nepal (C5) is clearly differentiated from the rest of the clusters. The Australasian cluster C7 seems the most distinct among the *O. rufipogon* population groups (Figure 15). The relatively low pairwise F_{ST} values between Southeast Asian clusters of *O. nivara* (C2 and C6) and *O. rufipogon* (C3) indicate less interspecific differentiation in that region compared to South Asia.

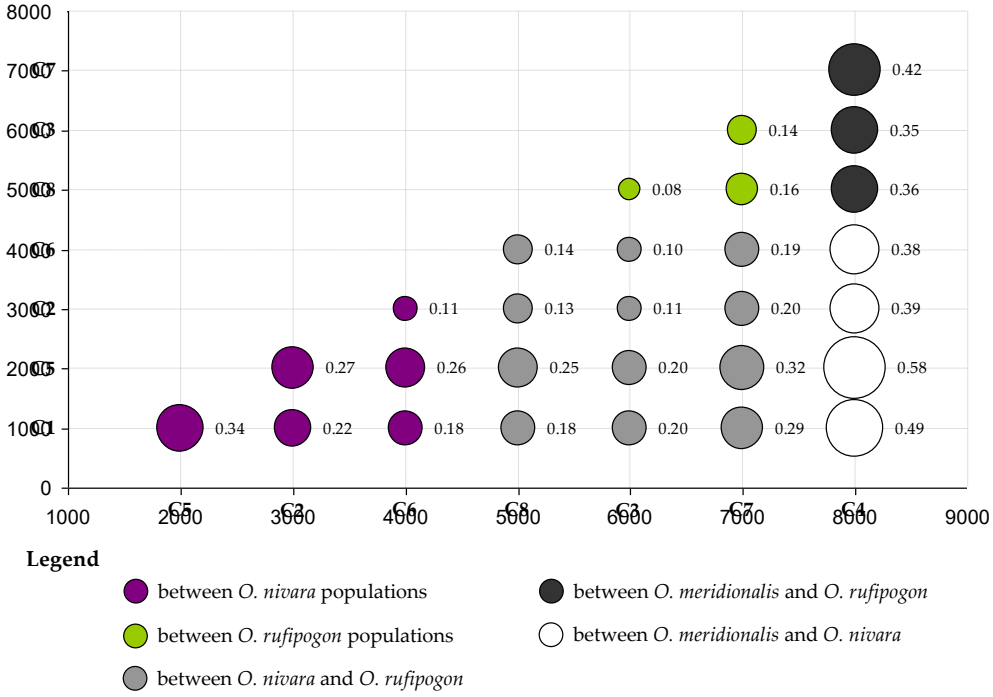


Figure 15. Pairwise F_{ST} of the eight population groups of *Oryza* series *sativae* from Asia Pacific. All values are significant at p < 0.001 level based on 999 permutations. The groups are coded according to TESS cluster numbers: C1 - Indian and Bangladeshi *O. nivara*; C5 - Nepalese *O. nivara*; C2 - Cambodian *O. nivara*; C6 - non-Cambodian Southeast Asian *O. nivara*; C8 - South Asian *O. rufipogon*; C3 - Southeast Asian *O. rufipogon*; C7 - Australasian *O. rufipogon*; and C4 - *O. meridionalis*. Circle size is proportional to its corresponding F_{ST} value (indicated at its right hand side).

Table 4 shows the Mantel correlations of geographic distances and F_{ST} values within and between species of Asia Pacific *Oryza* series *Sativae*. Significant correlations can be observed within *O. nivara* (r = 0. 257, p-value = 0.000), within *O. rufipogon* (r = 0.422, p-value = 0.000) and between the two species (r = 0. 220, p-value = 0.001) across their distribution range. However, when regional populations are tested, significant correlations are found only within South Asian *O. nivara* (r = 0.286, p-value = 0.016), within continental Southeast Asian *O. rufipogon* (r = 0.336, p-value = 0.012) and between *O. nivara* and *O. rufipogon* in Southeast Asia (r = 0.201, p-value = 0.026) (Table 4). Among the population groups, significant correlation is detected only in the Nepalese *O. nivara* cluster C5 (r = 0.689, p-value = 0.008).

Table 4. Mantel correlations of geographic distances and pairwise F_{ST} values within and between species of Asia Pacific *Oryza* series *Sativae*. Statistically significant r values are highlighted.

Taxon/taxa	r	p-value
Within species		
O. nivara (across distribution)	0.257	0.000
in South Asia	0.286	0.016
in Southeast Asia	0.195	0.093
O. rufipogon (across distribution)	0.422	0.000
in South Asia	0.074	0.351
in continental Southeast Asia	0.336	0.012
in insular Southeast Asia	0.171	0.210
in Australasia	0.283	0.111
O. meridionalis (across distribution)	-0.194	0.195
Between species		
O. meridionalis and *O. rufipogon*	-0.116	0.177
O. nivara and *O. rufipogon* (across distribution)	0.220	0.001
in South Asia	0.136	0.082
in continental Southeast Asia	0.201	0.026

The C.S. Chord distances and pairwise F_{ST} values of sympatric *O. nivara* and *O. rufipogon* population pairs are not significantly correlated with location (longitude and latitude) data. C.S. Chord distance shows positive moderate correlation with altitude (r = 0.39 at 0.01 significance level) while pairwise F_{ST} exhibits positive moderate correlation with altitude (r = 0.55 at 0.001 significance level) and negative moderate correlation with mean annual temperature (r = -0.42, at 0.01 significance level).

Discussion

Global overlapping and local differentiation

Phylogenetic (Figure 1) and ordination (Figure 2) methods reveal a lack of clear-cut genetic division between *O. nivara* and *O. rufipogon* across their distribution range, concurring with results of previous molecular studies (Second 1985; Barbier 1989; Iwamoto et al. 1999; Ren et al. 2003; Park et al. 2003; Cai et al. 2004; Zhu and Ge 2005; Zhou et al. 2008; Zheng and Ge 2010). None of the markers used in this study can discriminate the majority of either *O. nivara* or *O. rufipogon* accessions from the rest of the series (Appendix 5). The relatively large extent of allele sharing between non-sympatric populations from different geographic regions (Appendix 5; Figure 14) renders it more probable that most of the similarities can be traced to common ancestry although gene flow cannot be ruled out as an explanation for the genetic overlap of the two species (Zhou et al. 2008; Zheng and Ge 2010). However, genetic separation of *O. nivara* and *O. rufipogon* was detected by Bayesian methods at the highest population structure level (K = 2), even earlier than the recognition of *O. meridionalis* as a distinct group (at K = 3).

Based on the PCA plots (Figure 3) and pairwise F_{ST} chart (Figure 15), the genetic similarities seem to be greater in Southeast Asia. The fact that only one allele is endemic to South Asian *O. nivara* and *O. rufipogon* (compared to 14 in Southeast Asian populations) (Appendix 5) suggests lesser interspecific gene flow in the region. Furthermore, the absence of correlation between geographic distances and pairwise F_{ST} values between *O. nivara* and *O. rufipogon* in South Asia indicate that genetic differences are maintained regardless of spatial distance within the said geographic area (Table 4). This conforms to morphological variation patterns (reported in Chapter 2) but contradicts an earlier SSR experiment that reported greater species differentiation in Southeast Asia (Lu et al. 2008).

Despite the genetic overlap, species separation is apparent at a local scale. *O. nivara* and *O. rufipogon* populations from the same locality cluster apart from each other (except N43 and R43) in the neighbor joining tree while Bayesian analysis differentiated the two species in sympatric population pairs throughout their distribution range (Figures 8 and 12). Kuroda et al. (2007) have reported species separation of *O. nivara* and *O. rufipogon* populations in Vientiane, Laos. Therefore, contrary to the claim of Zheng and Ge (2010), molecular divergence is not completely absent between the two species and exists locally in sympatric populations indicative of adequately strong barriers to gene flow (e.g., differences

in phenology and mating system) operating over smaller spatial units. It is evident from the AMOVA results that gene flow is more restricted between populations rather than between species (Table 3). Similar situations of extensive allele sharing along with local genetic differentiation were observed in geographically overlapping populations of closely related bromeliad (*Pitcairnia*) species (Palma-Silva et al. 2011) and *Lupinus microcarpus* varieties (Drummond and Hamilton 2007).

Deeper divisions and less within-population diversity in *O. nivara*

The high genetic diversity (at the population and species level) (Table 2) and low intra-specific differentiation detected in *O. rufipogon* (compared to *O. nivara*; Figure 15) have been observed in similar studies (Kuroda et al. 2007; Lu et al. 2008; Zhou et al. 2008) and are typical variation patterns exhibited by outcrossing species (Hamrick and Godt 1996). As seen in *O. nivara* (Table 2; Figure 4), inbreeding species possess relatively low within-population diversity but tend to have higher inter-population diversity as different alleles are maintained by different isolated populations at specific loci (Silvertown and Charlesworth 2001). Indeed, our Mantel correlations suggest greater intra-specific gene flow in *O. rufipogon* than in *O. nivara* across distribution (Table 4). Interestingly, isolation by distance is detected mainly in South Asian *O. nivara* and mainland Southeast Asian *O. rufipogon* implying different geographic patterns of intra-specific gene flow in the two taxa.

STRUCTURE indicates an optimal four-cluster solution for Asia Pacific *Oryza* series *Sativae* (Figure 6) that recognizes the South Asian and Southeast Asian populations of *O. nivara* as two genetically distinct groups (Figures 7 - 8). Regional separation of *O. nivara* was also reflected by morphological data (Banaticla-Hilario et al. in Chapter 2).

At K = 8, spatially explicit Bayesian clustering infers four geographic populations in *O. nivara* (Figures 8 and 12) composed of two South Asian and two Southeast Asian groups. However, ordination (Figure 4A) and F_{ST} (Figure 15) analyses seem to combine the two Southeast Asian clusters, revealing greater differentiation of the India-Bangladeshi and Nepalese clusters.

O. nivara in Nepal: a discrete genetic entity

The genetic distinctiveness of *O. nivara* populations in Nepal is comparable to that of *O. meridionalis* as explicitly shown in the phylogenetic (Figure 1), ordination (Figures 2 – 3A) and F_{ST} (Figure 15) results. However, the Nepalese *O. nivara* group seems distinguishable at lower population structure levels. Bayesian methods detect this group at K = 5 following the recognition of the South-Southeast Asia split in *O. nivara* at K = 4 (Figures 5, 7 - 8). The uniqueness of *O. meridionalis* is evident also at higher population structure levels (K = 3) (Figure 5).

Low diversity and genetic isolation from the rest of the species expose the Nepalese *O. nivara* to inbreeding depression and genetic erosion. More in depth studies are needed not just to confirm the unique genetic identity of these regional populations but also to further establish variation patterns that will aid in formulating in- and ex-situ conservation strategies.

Australasian populations are distinct within *O. rufipogon*

The Australasian populations appear to be quite different from the rest of *O. rufipogon* as signified by ordination results (Figure 4B), Bayesian inference (Figures 5, 7 - 8) and relatively high F_{ST} values (Figure 15). Geographic isolation probably restrained gene exchange with other population groups. This corresponds to the morphological divergence of Australasian *O. rufipogon* reported by Banaticla-Hilario et al. (in Chapter 2). Genetic and morphological differentiation between continental and insular populations has been reported in many other plant species (for example Howcroft and Davidson 1973; Rivera-Ocasio et al. 2006; Fievet et al. 2007; Fedorenko et al. 2009).

O. rufipogon populations in Australasia have been reported to flourish vegetatively and produce less seeds in their natural habitats (Vaughan et al. 2003; 2008), which could be a reason for the low genetic diversity exhibited by this population group.

The geographic separation and fairly low diversity (compared to other populations of the species) of Australasian *O. rufipogon* predispose them to inbreeding depression and subsequent genetic deterioration as observed in other island populations (Frankham 1997 and 1998; Fedorenko et al. 2009). Therefore, this population group should be carefully examined and considered in reviewing and designing management practices for their protection and preservation.

Associations with *O. sativa*

The reported consanguinity of *O. rufipogon* with *O. sativa* var. *japonica* and of *O. nivara* with *O. sativa* var. *indica* (Cheng et al. 2003; Yamanaka et al. 2003; Ohtsubo et al. 2004; Xu et al. 2007; Xu et al. 2012) is evident from the present study (Figures 1, 2, 5 and 7). At the uppermost hierarchical level of population structure (K = 2), the japonica and aromatic varietal groups join up with *O. rufipogon* while the indica and aus groups do so with *O. nivara* (Figure 5).

Geographic clustering patterns are further displayed by the cultivated varieties at lower structural levels. Starting at K = 4, aus consistently groups with Cambodian *O. nivara* (Figure 7) while starting at K = 7, indica clusters with *O. nivara* from Thailand. This is analogous to the clustering patterns revealed by 6.5 million SNPs where indica and aus appeared similar to different populations of *O. nivara* (Xu et al. 2012). However, the limited number of cultivated populations analyzed in this study limits the validity of the clustering patterns obtained in this study. Phylogeographic results (Londo et al. 2006) agree with the genetic association of indica with wild rice in Thailand (Figure 5) but are in discordance with the observed merging of aromatic and japonica with South Asian *O. rufipogon* at K = 6 (Figure 7).

It is worth mentioning that out of the six populations morphologically classified as weedy forms (i.e., intermediate between *O. sativa* and either *O. nivara* or *O. rufipogon*) (Banaticla-Hilario et al. in Chapter 2), one was detected by STRUCTURE (at K = 6) as a genetic admixture of *O. nivara* and *O. rufipogon* while the other populations were included in the *O. nivara*-indica group. Caution should be taken when interpreting SSR diversity patterns as the presence of interaction between cultivated and wild taxa could be masked by the genetic similarities within *Oryza* series *Sativae*. Vaughan et al. (2008) warned that some genebank accessions of the Asian wild rice might have introgressed with cultivated rice, as most of these accessions were collected from disturbed habitats.

Altitudinal variation patterns

In South Asia, population groups of *O. nivara* and *O. rufipogon* exhibit lower diversity (Table 2) and higher intra- and inter-specific differentiation (Figure 15) indicating stronger gene flow barriers in the said geographic region. Nonetheless, population/accession diversity in both species across the entire distribution does

not correlate with latitude and longitude data. Species differentiation at the local level (represented by C.S. Chord distances and pairwise F_{ST} values between sympatric population pairs of *O. nivara* and *O. rufipogon*) also seems uncorrelated with geographic location.

On the other hand, altitude displays a weak negative correlation with population genetic diversity and a positive moderate correlation with local-scale species differentiation suggesting that reproductive barriers are intensified to a certain degree at high altitude conditions. Similar correlations were observed in other plant species (Chen et al. 2008; Hou and Lou 2011) as well as other altitudinal patterns where diversity increased with elevation (Shi et al. 2011) or where diversity and differentiation peaked at mid- elevations (Herrera and Bazaga 2008; Yan et al. 2009; Tanto Hadado et al. 2010). The inherent negative relationship between altitude and temperature (Korner 2007) can be accounting for the weak to moderate correlation of mean annual temperature with allelic richness in *O. rufipogon* and with F_{ST} values of sympatric accession pairs. Regional differences in the genetic variation patterns of *O. nivara* and *O. rufipogon* can be attributed to the altitudinal differences between South and Southeast Asia. The elevations of the sampled localities are much higher in South Asia (mean elevation = 253 meters above sea level) than in Southeast Asia (mean elevation = 124 m a.s.l.). However, the modest correlations, established between elevation and genetic variation, implicate other factors that could also have contributed to the observed genetic patterns. For example, several other environmental aspects such as soil, light and humidity were not incorporated in the analyses but may well be associated with species and population differentiation.

Conclusions

This research imparts a more detailed account of the genetic variation patterns in *O. nivara* and *O. rufipogon*, less so in the geographically restricted *O. meridionalis*. The recognition of local differentiation in the midst of global similarities reconciles the conflicting results of prior studies (Second 1985; Barbier 1989; Aggarwal et al. 1999; Iwamoto et al. 1999; Ren et al. 2003; Park et al. 2003; Cai et al. 2004; Zhu and Ge 2005; Duan et al. 2007; Kuroda et al. 2007; Takahashi et al. 2008; Zhou et al. 2008; Zheng and Ge 2010; Xu et al. 2012). Furthermore, regional differences in the strength of interspecific gene flow has been detected indicating that the extent of genetic differentiation between *O. nivara* and *O. rufipogon* varies at different

geographic scales.

Evolutionary divergence in these two closely related *Oryza* species is also evident from their opposing patterns of intra-specific diversity and differentiation with *O. nivara* having less diverse and more differentiated populations and *O. rufipogon*.

Bayesian analyses infers four, and at a finer structural level, eight genetic groups that correspond to geographic populations of the three *Oryza* series *Sativae* species in Asia Pacific. The revealed geographic partitions within species as well as the inferred population groupings within the series can be considered in assessing the genetic representativeness of genebank collections and in selecting plant materials for in- and ex-situ conservation and research purposes. Especially the uniqueness and vulnerability to genetic degradation of *O. nivara* in Nepal and of *O. rufipogon* in Australasia calls for immediate conservation measures. Furthermore, the vast amount of genetic variation detected among populations justifies the maintenance of a large collection of Asian wild rice germplasm.

Locally sympatric populations of *O. nivara* and *O. rufipogon* tend to become more differentiated and within-population diversity of both species tends to decrease with increasing elevation. This is one of the first reports (if not the first) showing altitudinal patterns of intra-specific genetic diversity and species differentiation in the two taxa. This study demonstrates the dynamic interplay among geography and other environmental factors that shaped the current variation within and between *O. nivara* and *O. rufipogon*.

References

Aggarwal, R. K., Brar, D. S., Nandi, S., Huang, N. & Khush, G. S. (1999). Phylogenetic relationships among *Oryza* species revealed by AFLP markers. *Theoretical and Applied Genetics* **98,** 1320-1328.

Barbier, P. (1989). Genetic variation and ecotypic variation in the wild rice species *Oryza rufipogon.* I. Population differentiation in life-history traits and isozymic loci. *The Japanese Journal of Genetics* **64,** 259-271.

Cai, H. W., Wang, X. K. & Morishima, H. (2004). Comparison of population genetic structures of common wild rice (*Oryza rufipogon* Griff.), as revealed by analyses of quantitative traits, allozymes, and RFLPs. *Heredity* **92,** 409-417.

Cavalli-Sforza, L. L. & Edwards, A. W. F. (1967). Phylogenetic analysis models and estimation procedures. *American Journal of Human Genetics* **19,** 233-257.

Chen, G., Wang, Y., Zhao, C., Korpelainen, H. & Li, C. (2008). Genetic diversity of *Hippophae rhamnoides* populations at varying altitudes in the Wolong Natural Reserve of China as revealed by ISSR markers. *Silvae Genetica* **57,** 29-36.

Cheng, C., Motohashi, R., Tsuchimoto, S., Fukuta, Y., Ohtsubo, H. & Ohtsubo, E. (2003). Polyphyletic origin of cultivated rice: based on the interspersion pattern of SINEs. *Molecular Biology and Evolution* **20,** 67-75.

Drummond, C. S. & Hamilton, M. B. (2007). Hierarchical components of genetic variation at a species boundary: population structure in two sympatric varieties of Lupinus microcarpus (Leguminosae). *Molecular Ecology* **16,** 753-769.

Duan, S., Lu, B., Li, Z., Tong, J., Kong, J., Yao, W., Li, S., Zhu, Y. (2007). Phylogenetic analysis of AA-genome *Oryza* species (Poaceae) based on chloroplast, mitochondrial, and nuclear DNA sequences. *Biochemical Genetics* **45,** 113-129.

Duistermaat, H. (1987). A revision of *Oryza* (Gramineae) in Malesia and Australia. *Blumea* **32,** 157-193.

Durand, E., Chen, C. & Francois, O. (2009). TESS version 2.3 - Reference Manual [Documentation file, online]. Document last modified in August, 2009 (accessed on September 28, 2011). Available at http://membres-timc.imag.fr/Olivier.Francois/manual.pdf

Evanno, G., Regnaut, S. &Goudet, J. (2005). Detecting the name of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14,** 2611 - 2620.

Falush, D., Stephens, M. & Pritchard, J. K. (2003). Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* **164,** 1567-1587.

Fedorenko, O. M., Gritskikh, M. V., Malysheva, I. E. & Nikolaevskaya, T. S. (2009). Genetic diversity of insular natural populations of *Festuca pratensis* Huds.: RAPD Analysis. *Russian Journal of Genetics* **45,** 1134-1138.

Fievet, V., Touzet, P., Arnaud, J.-F. & Cuguen, J. (2007). Spatial analysis of nuclear and cytoplasmic DNA diversity in wild sea beet (*Beta vulgaris* ssp. *maritima*) populations: do marine currents shape the genetic structure? *Molecular Ecology* **16,** 1847-1864.

Frankham, R. (1997). Do island populations have less genetic variation than mainland populations? *Heredity* **78,** 311-327.

Frankham, R. (1998). Inbreeding and Extinction: Island Populations. *Conservation Biology* **12,** 665-675.

Fulton, T., Chunwongse, J. & Tanksley, S. (1995). Microprep protocol for extraction of DNA from tomato and other herbaceous plants. *Plant Molecular Biology Reporter* **13,** 207-209.

Goudet, J. (2001). FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). At http://www2.unil.ch/popgen/softwares/fstat.htm.

Hamrick, J. L. & Godt, M. J. W. (1996). Effects of Life History Traits on Genetic Diversity in Plant Species. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences* **351,** 1291-1298.

Herrera, C. M. & Bazaga, P. (2007). Adding a third dimension to the edge of a species' range: altitude and genetic structuring in mountainous landscapes. *Heredity* **100,** 275-285.

Hijmans, R. J., Guarino, L., Cruz, M. & Rojas, E. (2001). Computer tools for spatial analysis of plant genetic resources data: 1. DIVA-GIS. *Plant Genetic Resources Newsletter* **124,** 15-19.

Hou, Y. & Lou, A. (2011). Population Genetic Diversity and Structure of a Naturally Isolated Plant Species, *Rhodiola dumulosa* (Crassulaceae). *PLoS ONE* **6(9),** e24497. doi:10.1371/journal.pone.0024497

Howcroft, N. & Davidson, J. (1973). An international provenance trial of *Pinus merkusii* in Papua New Guinea. *Tropical Forestry Research Note* **15,** 27.

Iwamoto, M., Nagashima, H., Nagamine, T., Higo, H. & Higo, K. (1999). A Tourist element in the flanking region of the catalase gene CatA reveals evolutionary relationships among *Oryza* species with various genome types. *Molecular Genetics and Genomics* **262,** 493-500.

Jakobsson, M. & Rosenberg, N. A. (2007). CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **23,** 1801-1806.

Körner, C. (2007). The use of altitude in ecological research. *Trends in Ecology & Evolution (Personal edition)* **22,** 569-574.

Kuroda, Y., Sato, Y.-I., Bounphanousay, C., Kono, Y. & Tanaka, K. (2007). Genetic structure of three *Oryza* AA genome species (*O. rufipogon, O. nivara* and *O. sativa*) as assessed by SSR analysis on the Vientiane Plain of Laos. *Conservation Genetics* **8,** 149-158.

Kwon, S.-J., Lee, J. K., Hong, S.-W., Park, Y.-J., McNally, K. L. & Kim, N.-S. (2006). Genetic diversity and phylogenetic relationship in AA *Oryza* species as revealed by Rim2/Hipa CACTA transposon display. *Genes & Genetic Systems* **81,** 93-101.

Liu, K. & Muse, S. V. (2005). PowerMarker: Integrated analysis environment for genetic marker data. *Bioinformatics* **21,** 2128-2129.

Londo, J. P., Chiang, Y. C., Hung, K. H., Chiang, T. Y. & Schaal, B. A. (2006). Phylogeography of Asian wild rice, *Oryza rufipogon*, reveals multiple independent domestications of cultivated rice, *Oryza sativa*. *Proceedings of the National Academy of Sciences of the United States of America* **103,** 9578-9583.

Lu, B.-R. (1999). Taxonomy of the genus *Oryza* (Poaceae): historical perspective and current status. *International Rice Research Notes* **24,** 4-8.

Lu, B.-R., Ge, S., Sang, T., Chen, J.-K. & Hong, D.-Y. (2001). The current taxonomy and perplexity of the genus *Oryza* (Poaceae). *Acta Phytotaxonomica Sinica* **39,** 373-388.

Lu, J., Zhang, X., Wang, H., Yuan, X., Xu, Q., Wang, Y., Yu, H., Tang, S. & Wei, X. (2008). SSR Analysis on Diversity of AA Genome *Oryza* Species in the Southeast and South Asia. *Rice Science* **15,** 289-294.

Martin, C., Juliano, A., Newbury, H. J., Lu, B. R., Jackson, M. T. & Ford-Lloyd, B. V. (1997). The use of RAPD markers to facilitate the identification of *Oryza* species within a germplasm collection. *Genetic Resources and Crop Evolution* **44,** 175-183.

Ng, N. Q., Chang, T. T., Williams, J. T. & Hawkes, J. G. (1981). Morphological studies of Asian rice and its related wild species and the recognition of a new Australian taxon. *Biological Journal of the Linnean Society* **16,** 303-313.

Ohtsubo, H., Cheng, C., Ohsawa, I., Tsuchimoto, S. & Ohtsubo, E. (2004). Rice retroposon p-SINE1 and origin of cultivated rice. *Breeding Science* **54,** 1-11.

Oka, H. I. (1988). *Origin of cultivated rice.* Japan Science Society Press, Tokyo.

Palma-Silva, C., Wendt, T., Pinheiro, F., Barbara, T., Fay, M. F., Cozzolino, S. & Lexer, C. (2011). Sympatric bromeliad species (*Pitcairnia* spp.) facilitate tests of mechanisms involved in species cohesion and reproductive isolation in Neotropical inselbergs. *Molecular Ecology* **20,** 3185-3201.

Park, K. C., Kim, N. H., Cho, Y. S., Kang, K. H., Lee, J. K. & Kim, N. S. (2003). Genetic variations of AA genome *Oryza* species measured by MITE-AFLP. *TAG Theoretical and Applied Genetics* **107,** 203-209.

Peakall, R. O. D. & Smouse, P. E. (2006). GenAlEx 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* **6,** 288-295.

Pritchard, J. K., Stephens, M. & Donelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* **155,** 945-959.

R Development Core Team (2011). R: A Language and Environment for Statistical Computing (R. F. f. S. Computing, Ed.). Vienna, Austria.

Ren, F., Lu, B.-R., Li, S., Huang, J. & Zhu, Y. (2003). A comparative study of genetic relationships among the AA-genome Oryza species using RAPD and SSR markers. *Theoretical and Applied Genetics* **108,** 113-120.

Rivera-Ocasio, E., Aide, T. M. & McMillan, W. O. (2006). The influence of spatial scale on the genetic structure of a widespread tropical wetland tree *Pterocarpus officinalis* (Fabaceae). *Conservation Genetics* **7,** 251-266.

Rosenberg, N. A. (2004). *Distruct*: a program for the graphical display of population structure. *Molecular Ecology Notes* **4,** 137-138.

Second, G. (1985). Evolutionary relationships in the Sativa group of *Oryza* based on isozyme data. *Genetics Selection Evolution* **17,** 89-114.

Sharma, S. D. & Shastry, S. V. S. (1965). Taxonomic studies in the genus *Oryza* L. III. *O. rufipogon* Griff. sensu stricto and *O. nivara* Sharma et Shastry *nom. nov. Indian Journal of Genetics and Plant Breeding* **25,** 157-167.

Shi, M.-M., Michalski, S. G., Chen, X.-Y. & Durka, W. (2011). Isolation by elevation: genetic structure at neutral and putatively non-neutral loci in a dominant tree of subtropical forests, *Castanopsis eyrei*. *PLoS ONE* **6(6),** e21302. doi:10.1371/journal.pone.0021302

Silvertown, J. & Charlesworth, D. (2001). *Introduction to plant population biology*. Blackwell Publishing, Oxford.

Takahashi, H., Sato, Y.-I. & Nakamura, I. (2008). Evolutionary analysis of two plastid DNA sequences in cultivated and wild species of *Oryza*. *Breeding Science* **58,** 225-233.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. & Kumar, S. (2011). MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Molecular Biology and Evolution* **28,** 2731-2739.

Tanto Hadado, T., Rau, D., Bitocchi, E. & Papa, R. (2010). Adaptation and diversity along an altitudinal gradient in Ethiopian barley (*Hordeum vulgare* L.) landraces revealed by molecular analysis. *BMC Plant Biology* **10,** 121.

Tateoka, T. (1963). Taxonomic studies of *Oryza*. III. Key to the species and their enumeration. *Botanical Magazine Tokyo* **76,** 165-173.

Uga, Y., Fukuta, Y., Ohsawa, R. & Fujimura, T. (2003). Variations of floral traits in Asian cultivated rice (*Oryza sativa* L.) and its wild relatives (*O. rufipogon* Griff.). *Breeding Science* **53,** 345-352.

Vaughan, D. A., Lu, B.-R. & Tomooka, N. (2008). The evolving story of rice evolution. *Plant Science* **174,** 394-408.

Vaughan, D. A., Morishima, H. & Kadowaki, K. (2003). Diversity in the *Oryza* genus. *Current Opinion in Plant Biology* **6,** 139-146.

Weir, B. S. (1996). *Genetic data analysis II*. Sinauer Associates, Sunderland, Massachusetts.

Xu, J.-H., Cheng, C., Tsuchimoto, S., Ohtsubo, H. & Ohtsubo, E. (2007). Phylogenetic analysis of *Oryza rufipogon* strains and their relations to *Oryza sativa* strains by insertion polymorphism of rice SINEs. *Genes & Genetic Systems* **82,** 217-229.

Xu, J.-H., Kurata, N., Akimoto, M., Ohtsubo, H. & Ohtsubo, E. (2005). Identification and characterization of Australian wild rice strains of *Oryza meridionalis* and *Oryza rufipogon* by SINE insertion polymorphism. *Genes & Genetic Systems* **80,** 129-134.

Xu, X., Liu, X., Ge, S., Jensen, J. D., Hu, F., Li, X., Dong Y., Gutenkunst R. N., Fang, L., Huang, L. Li, J., He, W., Zhang, G., Zheng, X., Zhang, F., Li, Y., Yu, C., Kristiansen, K., Zhang, X., Wang, J., Wright, M., McCouch, S., Nielsen, R., Wang, J. & Wang, W. (2012). Resequencing 50 accessions of cultivated and

wild rice yields markers for identifying agronomically important genes. *Nature Biotechnology* **30,** 105–111.

Yamanaka, S., Nakamura, I., Nakai, H. & Sato, Y.-I. (2003). Dual origin of the cultivated rice based on molecular markers of newly collected annual and perennial strains of wild rice species, *Oryza nivara* and *O. rufipogon*. *Genetic Resources and Crop Evolution* **50,** 529-538.

Yan, X.-B., Guo, Y.-X., Zhao, C., Liu, F.-Y. & Lu, B.-R. (2009). Intra-population genetic diversity of two wheatgrass species along altitude gradients on the Qinghai-Tibetan Plateau: its implication for conservation and utilization. *Conservation Genetics* **10,** 359-367.

Zheng, X. & Ge, S. (2010). Ecological divergence in the presence of gene flow in two closely related *Oryza* species (*Oryza rufipogon* and *O. nivara*). *Molecular Ecology* **19,** 2439-2454.

Zhou, H.-F., Zheng, X.-M., Wei, R.-X., Second, G., Vaughan, D. A. & Ge, S. (2008). Contrasting population genetic structure and gene flow between *Oryza rufipogon* and *Oryza nivara*. *Theoretical and Applied Genetics* **117,** 1181-1189.

Zhu, Q. & Ge, S. (2005). Phylogenetic relationships among A-genome species of the genus *Oryza* revealed by intron sequences of four nuclear genes. *New Phytologist* **167,** 249-265.

Appendix 1. Genetic diversity exhibited by the 128 *Oryza* series *Sativae* populations across 29 microsatellite loci.

Code	IRGC number	Species	Geographic origin	S	A	Rs	H	Gene diversity
N1	105882	*O. nivara*	Bangladesh	5	1.55	1.22	0.02	0.16
N2	103830	*O. nivara*	Bangladesh	5	1.07	1.03	0.00	0.02
N18	80549	*O. nivara*	India	5	1.07	1.03	0.01	0.02
N19	80560	*O. nivara*	India	4	1.24	1.09	0.04	0.07
N20	80601	*O. nivara*	India	5	1.21	1.08	0.02	0.06
N21	80677	*O. nivara*	India	5	1.69	1.26	0.03	0.19
N22	106101	*O. nivara*	India	5	1.34	1.15	0.03	0.11
N23	80645	*O. nivara*	India	5	1.34	1.13	0.01	0.09
N24	106051	*O. nivara*	India	5	1.21	1.08	0.01	0.06
N25	106054	*O. nivara*	India	5	1.28	1.10	0.01	0.07
N26[a]	81837	*O. nivara*	India	5	2.10	1.20	0.03	0.31
N26A	81837	*O. nivara*	India	3	1.52	1.28	0.02	0.15
N26B	81837	*O. nivara*	India	2	1.21	1.13	0.03	0.05
N32	93192	*O. nivara*	Nepal	5	1.31	1.12	0.03	0.09
N33	93185	*O. nivara*	Nepal	5	1.24	1.09	0.04	0.07
N34	93191	*O. nivara*	Nepal	5	1.38	1.13	0.02	0.10
N35	93184	*O. nivara*	Nepal	5	1.62	1.22	0.03	0.16
N36	93195	*O. nivara*	Nepal	5	1.45	1.18	0.03	0.13
N37	103422	*O. nivara*	Sri Lanka	5	1.21	1.08	0.00	0.06
N3[bc]	93129	*O. nivara*	Cambodia	3	1.34	1.17	0.05	0.10
N4[c]	105719	*O. nivara*	Cambodia	5	1.31	1.09	0.01	0.07
N5[c]	89216	*O. nivara*	Cambodia	5	1.59	1.25	0.02	0.18
N6	89187	*O. nivara*	Cambodia	5	1.59	1.21	0.04	0.15
N7	89172	*O. nivara*	Cambodia	5	1.24	1.08	0.01	0.06
N8	89185	*O. nivara*	Cambodia	5	1.55	1.24	0.01	0.17
N9	89212	*O. nivara*	Cambodia	5	1.38	1.14	0.02	0.10
N10	106320	*O. nivara*	Cambodia	5	1.52	1.19	0.01	0.14
N11[c]	88939	*O. nivara*	Cambodia	5	1.38	1.15	0.04	0.11
N12	92886	*O. nivara*	Cambodia	5	1.17	1.07	0.02	0.05
N13	92677	*O. nivara*	Cambodia	5	1.52	1.21	0.03	0.16
N14	92823	*O. nivara*	Cambodia	5	1.83	1.35	0.03	0.25
N15	106339	*O. nivara*	Cambodia	5	1.48	1.18	0.02	0.13
N16	92943	*O. nivara*	Cambodia	5	1.34	1.14	0.02	0.10
N27	86699	*O. nivara*	Laos	5	1.31	1.12	0.01	0.08
N28	106148	*O. nivara*	Laos	5	1.24	1.10	0.01	0.08
N29	106151	*O. nivara*	Laos	4	1.24	1.10	0.02	0.07
N31[c]	106347	*O. nivara*	Myanmar	5	1.45	1.17	0.01	0.13
N38	105803	*O. nivara*	Thailand	5	1.48	1.21	0.01	0.15
N39	104724	*O. nivara*	Thailand	5	1.21	1.07	0.01	0.05
N40	105765	*O. nivara*	Thailand	5	1.59	1.20	0.02	0.14
N41	105755	*O. nivara*	Thailand	4	1.03	1.02	0.00	0.01

Appendix 1. (Continued) Genetic diversity exhibited by the 128 *Oryza* series *Sativae* populations across 29 microsatellite loci.

Code	IRGC number	Species	Geographic origin	S	A	Rs	H	Gene diversity
N42	104743	*O. nivara*	Thailand	5	1.17	1.07	0.00	0.05
N43[c]	104756	*O. nivara*	Thailand	5	1.14	1.06	0.01	0.04
N44	105859	*O. nivara*	Thailand	5	1.45	1.17	0.02	0.13
N45	104473	*O. nivara*	Thailand	2	1.41	1.26	0.05	0.11
N46	105801	*O. nivara*	Thailand	5	1.34	1.10	0.02	0.08
N47	105809	*O. nivara*	Thailand	5	1.69	1.25	0.03	0.18
N48	105825	*O. nivara*	Thailand	5	1.07	1.03	0.00	0.02
N49	105828	*O. nivara*	Thailand	5	1.59	1.21	0.03	0.15
N50	104736	*O. nivara*	Thailand	5	1.21	1.07	0.02	0.05
N51	86496	*O. nivara*	Vietnam	5	1.10	1.04	0.03	0.03
N52	86493	*O. nivara*	Vietnam	5	1.21	1.07	0.02	0.05
R1	105881	*O. rufipogon*	Bangladesh	4	1.45	1.19	0.08	0.13
R2	103827	*O. rufipogon*	Bangladesh	5	1.83	1.25	0.12	0.19
R18	80550	*O. rufipogon*	India	4	1.55	1.22	0.13	0.16
R19	80562	*O. rufipogon*	India	5	1.59	1.22	0.07	0.16
R20	80600	*O. rufipogon*	India	4	1.38	1.14	0.06	0.10
R21	80680	*O. rufipogon*	India	5	1.48	1.18	0.07	0.13
R22	82983	*O. rufipogon*	India	5	1.79	1.28	0.11	0.21
R23[c]	80643	*O. rufipogon*	India	5	1.76	1.28	0.23	0.22
R24	82982	*O. rufipogon*	India	4	1.69	1.26	0.12	0.18
R25	106055	*O. rufipogon*	India	5	1.97	1.32	0.12	0.24
R26	81881	*O. rufipogon*	India	3	1.62	1.28	0.07	0.17
R32[c]	93221	*O. rufipogon*	Nepal	4	1.97	1.37	0.14	0.26
R33	93218	*O. rufipogon*	Nepal	5	1.59	1.24	0.03	0.18
R34	93220	*O. rufipogon*	Nepal	5	1.76	1.31	0.08	0.23
R35	93210	*O. rufipogon*	Nepal	5	1.48	1.17	0.10	0.13
R36	93216	*O. rufipogon*	Nepal	5	1.45	1.18	0.11	0.14
R37[c]	103423	*O. rufipogon*	Sri Lanka	5	1.34	1.15	0.03	0.11
R17[c]	103823	*O. rufipogon*	China	5	1.72	1.24	0.02	0.17
R3	93085	*O. rufipogon*	Cambodia	5	2.14	1.39	0.10	0.29
R4	105720	*O. rufipogon*	Cambodia	5	1.59	1.23	0.03	0.17
R5[a]	89228	*O. rufipogon*	Cambodia	5	2.03	1.30	0.16	0.27
R5A[c]	89228	*O. rufipogon*	Cambodia	3	1.62	1.28	0.15	0.17
R5B	89228	*O. rufipogon*	Cambodia	2	1.59	1.32	0.17	0.15
R6	89230	*O. rufipogon*	Cambodia	5	1.83	1.31	0.09	0.23
R7	89223	*O. rufipogon*	Cambodia	5	1.62	1.26	0.06	0.19
R8	89227	*O. rufipogon*	Cambodia	2	1.69	1.37	0.17	0.17
R9	89232	*O. rufipogon*	Cambodia	5	2.34	1.40	0.16	0.30
R10[c]	106321	*O. rufipogon*	Cambodia	4	1.31	1.12	0.08	0.09
R11	89007	*O. rufipogon*	Cambodia	5	1.69	1.25	0.08	0.19
R12	110408	*O. rufipogon*	Cambodia	5	1.97	1.35	0.08	0.26

Appendix 1. (Continued) Genetic diversity exhibited by the 128 *Oryza* series *Sativae* populations across 29 microsatellite loci.

Code	IRGC number	Species	Geographic origin	S	A	Rs	H	Gene diversity
R13	99538	*O. rufipogon*	Cambodia	5	2.03	1.33	0.16	0.25
R14	93063	*O. rufipogon*	Cambodia	5	2.03	1.33	0.12	0.25
R15[c]	106335	*O. rufipogon*	Cambodia	5	1.55	1.24	0.03	0.18
R16	110409	*O. rufipogon*	Cambodia	5	1.48	1.21	0.06	0.15
R28	106149	*O. rufipogon*	Laos	5	1.79	1.28	0.08	0.21
R29[a]	106152	*O. rufipogon*	Laos	4	1.93	1.22	0.09	0.24
R29A[c]	106152	*O. rufipogon*	Laos	2	1.45	1.27	0.12	0.11
R29B	106152	*O. rufipogon*	Laos	2	1.31	1.18	0.07	0.08
R30	106357	*O. rufipogon*	Myanmar	5	1.10	1.04	0.03	0.03
R31	106346	*O. rufipogon*	Myanmar	5	1.72	1.29	0.08	0.21
R38	105804	*O. rufipogon*	Thailand	4	1.28	1.12	0.04	0.08
R39	104713	*O. rufipogon*	Thailand	4	1.69	1.27	0.05	0.18
R40	105766	*O. rufipogon*	Thailand	5	1.55	1.17	0.08	0.13
R41	105758	*O. rufipogon*	Thailand	5	1.69	1.25	0.10	0.19
R42	104742	*O. rufipogon*	Thailand	5	1.69	1.25	0.11	0.18
R43[c]	104757	*O. rufipogon*	Thailand	4	1.17	1.08	0.00	0.05
R44	105860	*O. rufipogon*	Thailand	4	1.38	1.17	0.03	0.11
R46	105800	*O. rufipogon*	Thailand	5	1.72	1.27	0.06	0.20
R47	82979	*O. rufipogon*	Thailand	5	1.97	1.30	0.17	0.23
R48	105823	*O. rufipogon*	Thailand	2	1.24	1.16	0.05	0.07
R49[c]	105829	*O. rufipogon*	Thailand	4	1.41	1.17	0.09	0.12
R50[c]	104737	*O. rufipogon*	Thailand	5	1.28	1.11	0.01	0.08
R51	86512	*O. rufipogon*	Vietnam	4	1.79	1.29	0.18	0.21
R52	86506	*O. rufipogon*	Vietnam	5	1.72	1.26	0.06	0.19
R53	80774	*O. rufipogon*	Philippines	5	1.48	1.18	0.07	0.14
R54	81976	*O. rufipogon*	Indonesia	5	1.59	1.24	0.06	0.17
R55	81977	*O. rufipogon*	Indonesia	5	1.55	1.20	0.11	0.15
R56	81978	*O. rufipogon*	Indonesia	5	2.21	1.42	0.18	0.31
R58	105567	*O. rufipogon*	Indonesia	5	2.00	1.39	0.16	0.29
R59	105952	*O. rufipogon*	Indonesia	5	1.66	1.17	0.10	0.13
R60	105958	*O. rufipogon*	Indonesia	5	1.90	1.37	0.18	0.28
R61	106452	*O. rufipogon*	Indonesia	5	1.59	1.24	0.14	0.19
R62	106453	*O. rufipogon*	Indonesia	5	1.41	1.14	0.02	0.10
R63	86542	*O. rufipogon*	Australia	5	1.55	1.22	0.17	0.17
R64	93274	*O. rufipogon*	Indonesia	5	1.21	1.06	0.01	0.04
R65[cd]	105283	*O. rufipogon*	Australia	5	1.38	1.17	0.06	0.13
R66	105293	*O. rufipogon*	Australia	5	1.10	1.04	0.02	0.03
R67	105303	*O. rufipogon*	Australia	5	1.62	1.24	0.14	0.19
M1	86539	*O. meridionalis*	Australia	5	1.10	1.04	0.00	0.03
M2	93260	*O. meridionalis*	Indonesia	5	1.00	1.00	0.00	0.00
M3	105282	*O. meridionalis*	Australia	4	1.10	1.04	0.01	0.03

Appendix 1. (Continued) Genetic diversity exhibited by the 128 *Oryza* series *Sativae* populations across 29 microsatellite loci.

Code	IRGC number	Species	Geographic origin	S	A	Rs	H	Gene diversity
M4	105294	*O. meridionalis*	Australia	5	1.41	1.21	0.00	0.15
M5	105302	*O. meridionalis*	Australia	5	1.34	1.11	0.02	0.08
AR1	12880	*O. sativa*	Iran	5	1.21	1.08	0.00	0.06
J1	328	*O. sativa*	Philippines	5	1.38	1.03	0.01	0.02
J2	12731	*O. sativa*	Japan	5	1.07	1.12	0.00	0.09
IN1	66970	*O. sativa*	Philippines	5	1.24	1.03	0.03	0.02
IN2	108921	*O. sativa*	India	5	1.07	1.09	0.00	0.07
AU1	32561	*O. sativa*	India	4	1.28	1.14	0.05	0.10

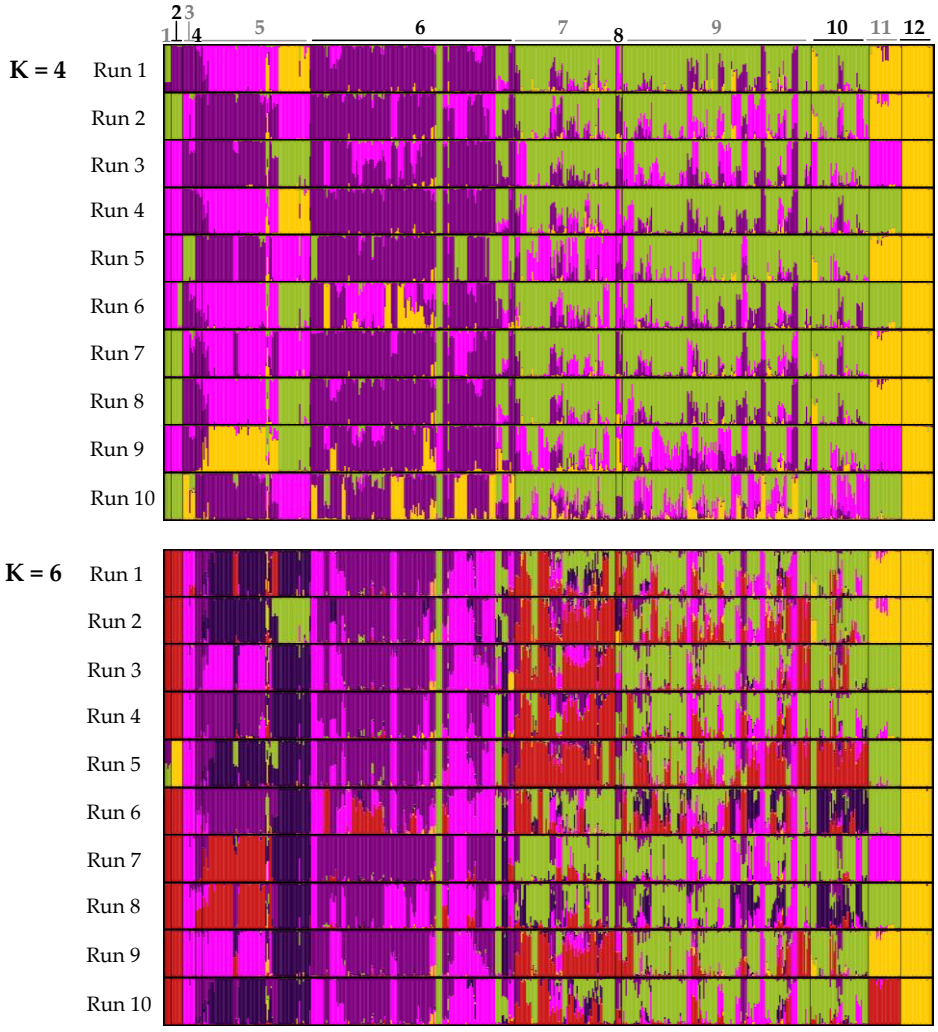
S –sample size; A –number of alleles; Rs – allelic richness; H – observed heterozygosity

[a] Composed of two plant types; partitioned into two subpopulations in the succeeding analyses

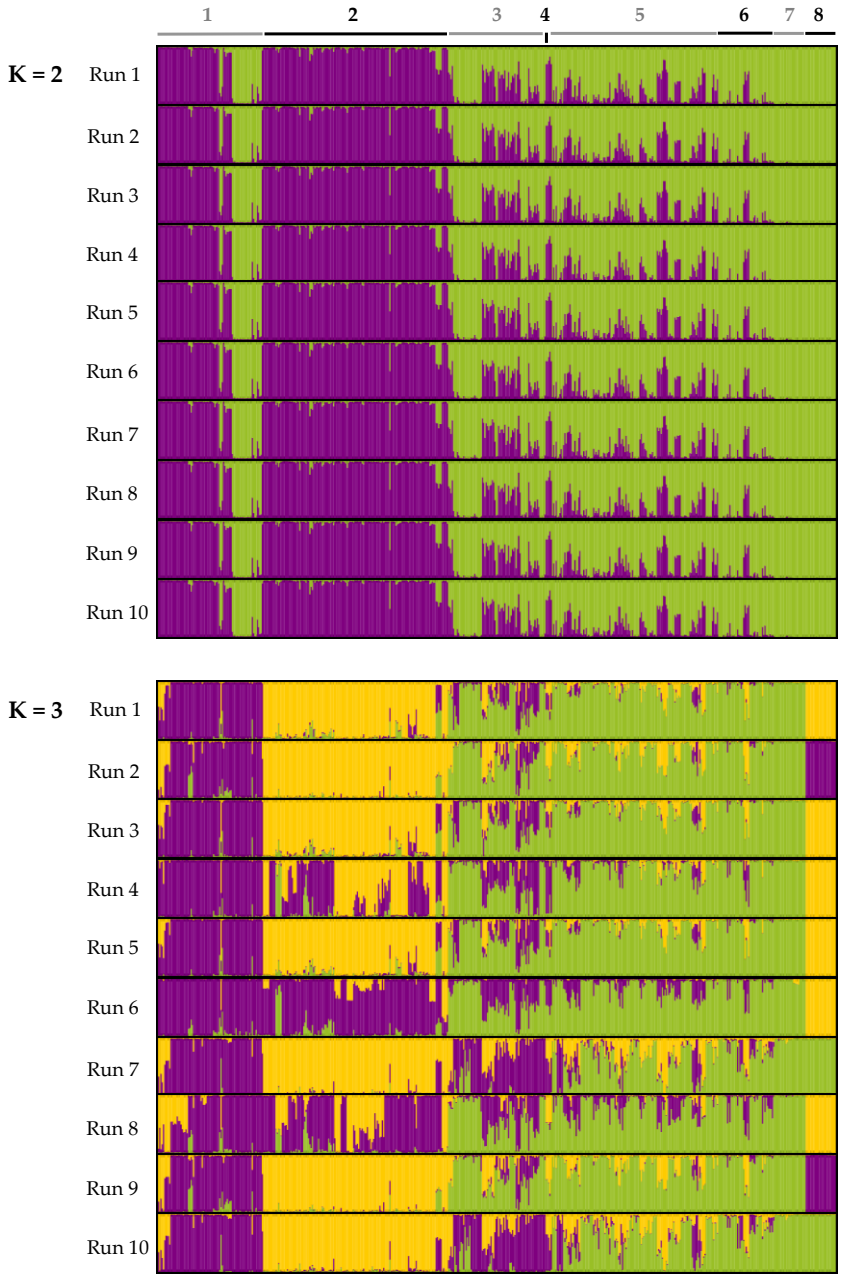
[b] Currently labeled as *O. rufipogon* in the International Rice Genebank Collection Information System (IRGCIS), tentatively re-classified by the author as *O. nivara* based on seed morphology

[c] Classified as intermediate form in previous morphological analysis

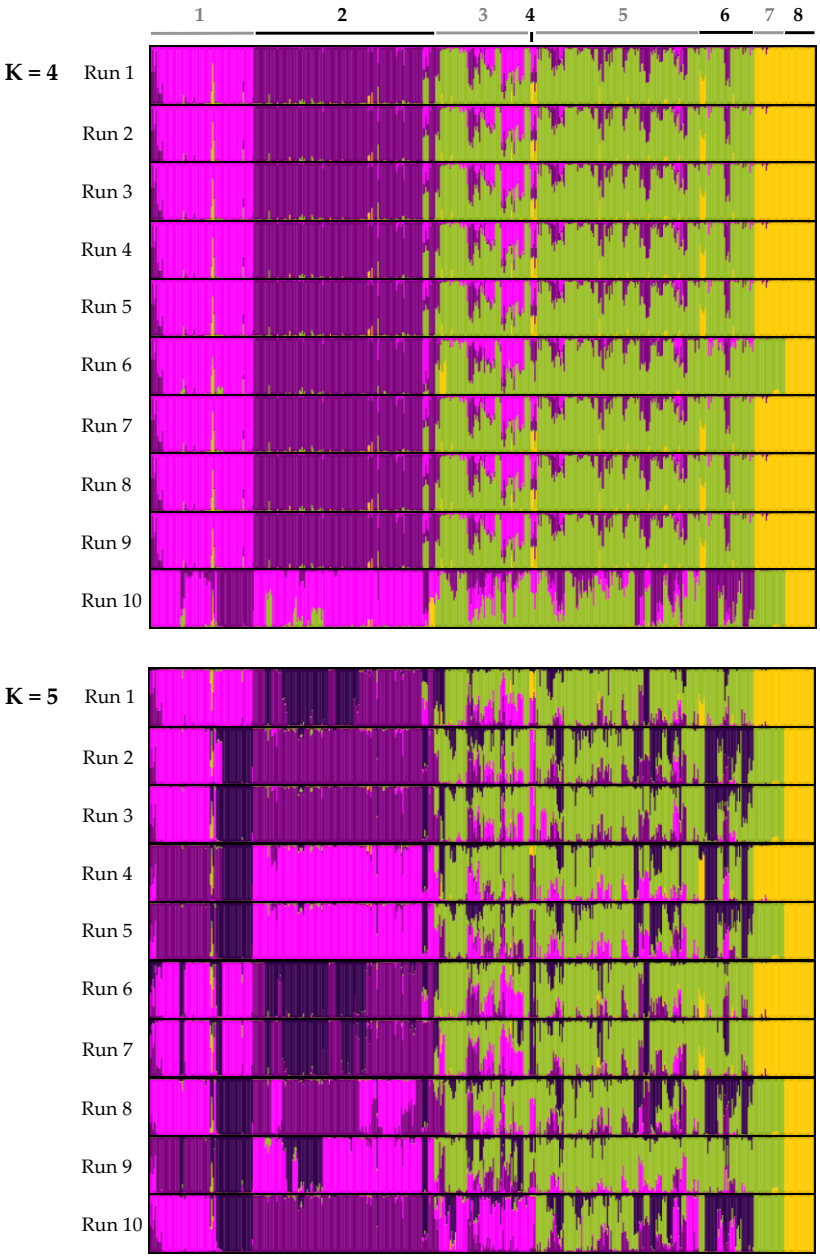
[d] Currently labeled as *O. meridionalis* in IRGCIS, tentatively re-classified by the author as *O. rufipogon* based on seed morphology



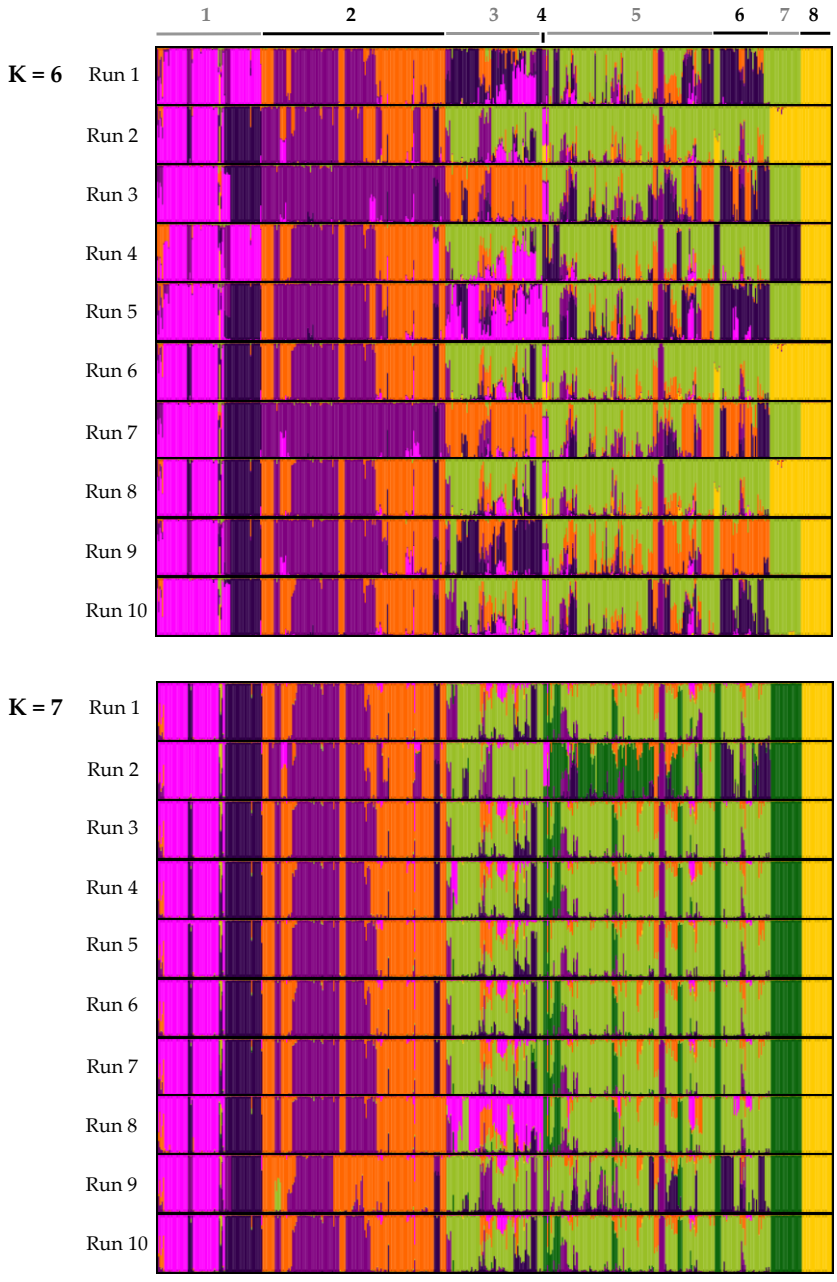
Appendix 2. Membership coefficients of 10 aligned STRUCTURE runs at K = 4 and K = 6. The pre-defined populations are: 1 – *O. sativa* (aromatic); 2 – *O. sativa* (japonica); 3 – *O. sativa* (indica); 4 – *O. sativa* (aus); 5 – *O. nivara* (from South Asia); 6 – *O. nivara* (from Southeast Asia); 7 – *O. rufipogon* (from South Asia); 8 – *O. rufipogon* (from China); 9 – *O. rufipogon* (from continental Southeast Asia); 10 – *O. rufipogon* (from insular Southeast Asia); 11 – *O. rufipogon* (from Australasia); 12 – *O. meridionalis* (from Australasia).



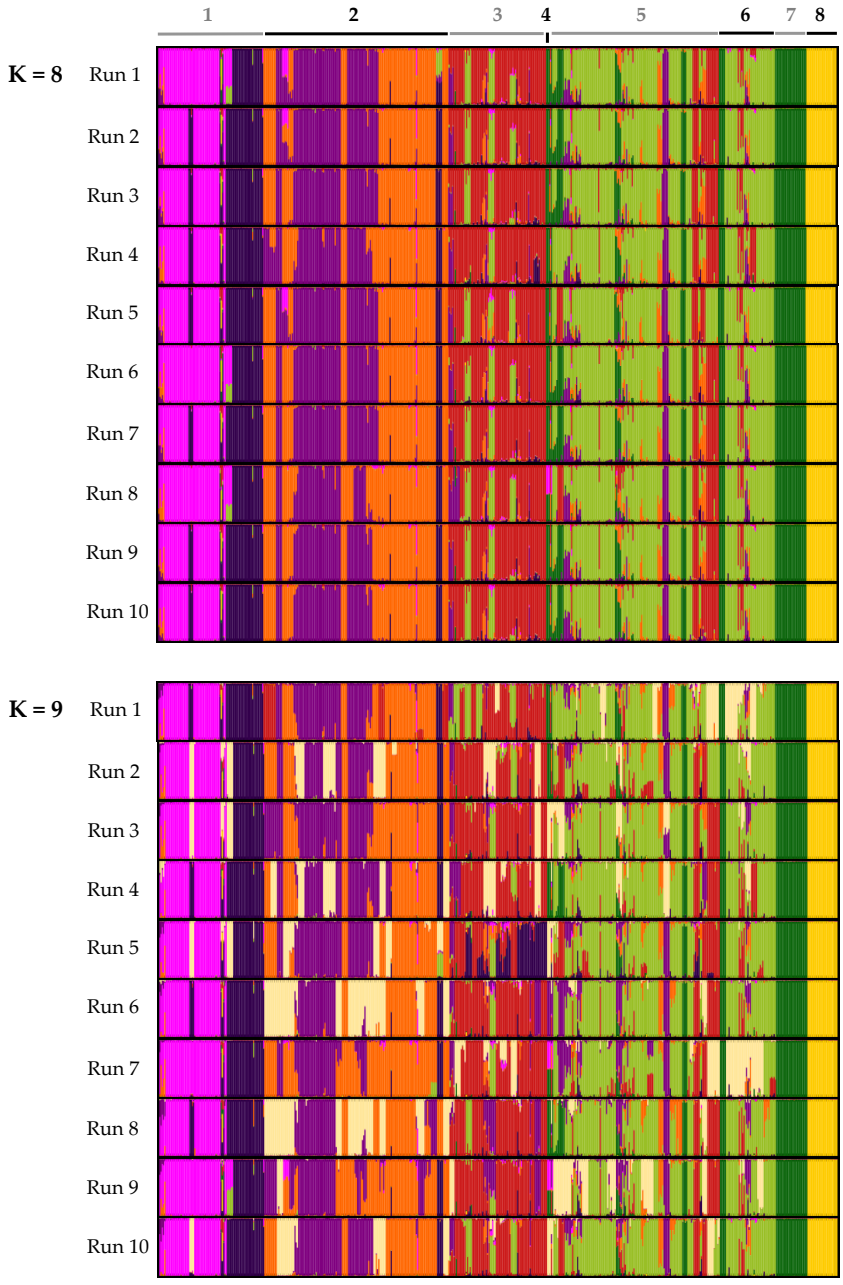
Appendix 3A. Membership coefficients of the 10 aligned TESS runs at K = 2 and K = 3. The predefined populations are: 1 – South Asian *O. nivara*; 2 – Southeast Asian *O. nivara*; 3 – South Asian *O. rufipogon*; 4 – Chinese *O. rufipogon*; 5 – continental Southeast Asian *O. rufipogon*; 6 – insular Southeast Asian *O. rufipogon*; 7 – Australasian *O. rufipogon*; and 8 – *O. meridionalis*.



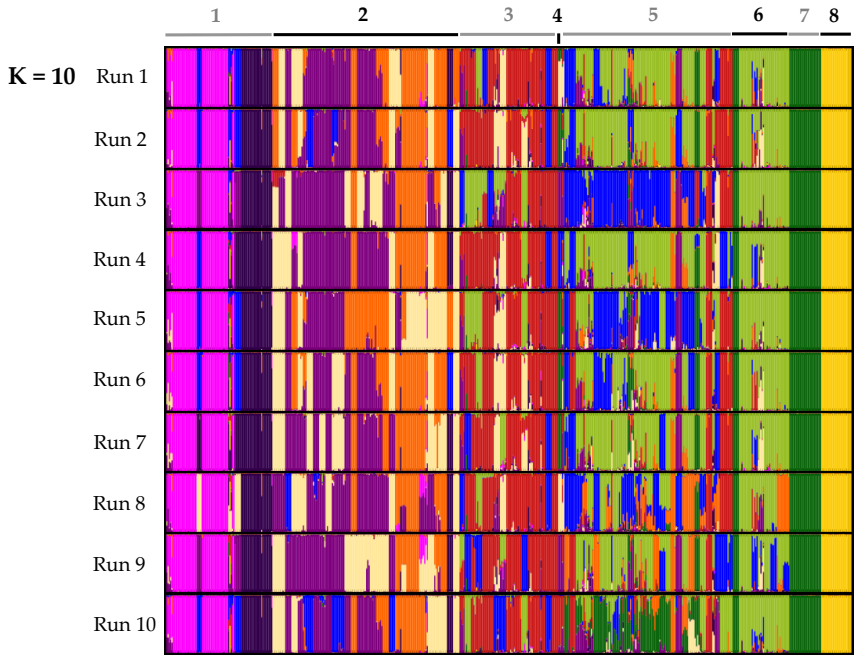
Appendix 3B. Membership coefficients of the 10 aligned TESS runs at K = 4 and K = 5. The predefined populations are: 1 – South Asian *O. nivara*; 2 – Southeast Asian *O. nivara*; 3 – South Asian *O. rufipogon*; 4 – Chinese *O. rufipogon*; 5 – continental Southeast Asian *O. rufipogon*; 6 – insular Southeast Asian *O. rufipogon*; 7 – Australasian *O. rufipogon*; and 8 – *O. meridionalis*.



Appendix 3C. Membership coefficients of the 10 aligned TESS runs at K = 6 and K = 7. The predefined populations are: 1 – South Asian *O. nivara*; 2 – Southeast Asian *O. nivara*; 3 – South Asian *O. rufipogon*; 4 – Chinese *O. rufipogon*; 5 – continental Southeast Asian *O. rufipogon*; 6 – insular Southeast Asian *O. rufipogon*; 7 – Australasian *O. rufipogon*; and 8 – *O. meridionalis*.



Appendix 3D. Membership coefficients of the 10 aligned TESS runs at K = 8 and K = 9. The predefined populations are: 1 – South Asian *O. nivara*; 2 – Southeast Asian *O. nivara*; 3 – South Asian *O. rufipogon*; 4 – Chinese *O. rufipogon*; 5 – continental Southeast Asian *O. rufipogon*; 6 – insular Southeast Asian *O. rufipogon*; 7 – Australasian *O. rufipogon*; and 8 – *O. meridionalis*.



Appendix 3E. Membership coefficients of the 10 aligned TESS runs at K = 10. The predefined populations are: 1 – South Asian *O. nivara*; 2 – Southeast Asian *O. nivara*; 3 – South Asian *O. rufipogon*; 4 – Chinese *O. rufipogon*; 5 – continental Southeast Asian *O. rufipogon*; 6 – insular Southeast Asian *O. rufipogon*; 7 – Australasian *O. rufipogon*; and 8 – *O. meridionalis*.

Appendix 4. TESS cluster membership of each population group at K = 8 (averaged over 10 runs).

	Partial membership to each cluster							
Population code	C1	C2	C3	C4	C5	C6	C7	C8
N1	0.5888	0.1442	0.0033	0.0004	0.0005	0.2592	0.0004	0.0032
N2	0.9905	0.0044	0.0004	0.0002	0.0003	0.0027	0.0001	0.0014
N18	0.9967	0.0003	0.0009	0.0000	0.0005	0.0005	0.0001	0.0011
N19	0.9910	0.0005	0.0028	0.0001	0.0018	0.0017	0.0001	0.0021
N20	0.9965	0.0005	0.0007	0.0000	0.0006	0.0004	0.0001	0.0012
N21	0.4343	0.0017	0.0029	0.0001	0.5394	0.0020	0.0008	0.0188
N22	0.9939	0.0008	0.0010	0.0001	0.0010	0.0012	0.0001	0.0021
N23	0.9902	0.0011	0.0016	0.0002	0.0011	0.0015	0.0001	0.0041
N24	0.9941	0.0005	0.0007	0.0001	0.0010	0.0008	0.0001	0.0027
N25	0.9834	0.0008	0.0061	0.0001	0.0035	0.0015	0.0010	0.0036
N26A	0.0111	0.0038	0.0188	0.1729	0.4794	0.0027	0.1624	0.1489
N26B	0.9534	0.0169	0.0053	0.0002	0.0025	0.0051	0.0083	0.0085
N37	0.2048	0.0016	0.0936	0.0002	0.6968	0.0005	0.0001	0.0024
N32	0.0004	0.0003	0.0004	0.0000	0.9974	0.0004	0.0001	0.0009
N33	0.0003	0.0004	0.0004	0.0000	0.9974	0.0004	0.0001	0.0010
N34	0.0003	0.0003	0.0004	0.0001	0.9973	0.0004	0.0001	0.0009
N35	0.0022	0.0024	0.0383	0.0001	0.9453	0.0082	0.0015	0.0020
N36	0.0342	0.0013	0.0010	0.0000	0.9565	0.0055	0.0001	0.0013
N51	0.0004	0.0998	0.0020	0.0003	0.0002	0.8947	0.0015	0.0011
N52	0.0007	0.0666	0.0020	0.0002	0.0005	0.9284	0.0003	0.0014
N27	0.0026	0.9843	0.0034	0.0002	0.0003	0.0041	0.0017	0.0034
N28	0.1292	0.1456	0.0023	0.0003	0.0003	0.7206	0.0005	0.0012
N29	0.0027	0.0747	0.0013	0.0002	0.0010	0.9134	0.0010	0.0058
N3	0.0042	0.8200	0.0119	0.0005	0.0053	0.1548	0.0008	0.0025
N4	0.0005	0.9794	0.0024	0.0013	0.0007	0.0142	0.0007	0.0007
N5	0.0018	0.9933	0.0016	0.0001	0.0007	0.0013	0.0004	0.0008
N6	0.0005	0.9730	0.0202	0.0002	0.0023	0.0018	0.0008	0.0011
N7	0.0005	0.9958	0.0013	0.0001	0.0006	0.0009	0.0003	0.0005
N8	0.0009	0.9536	0.0038	0.0002	0.0027	0.0250	0.0008	0.0130
N9	0.0020	0.9881	0.0023	0.0002	0.0035	0.0014	0.0010	0.0013
N10	0.0011	0.9662	0.0013	0.0002	0.0002	0.0277	0.0007	0.0026
N11	0.0010	0.0047	0.0022	0.0051	0.0023	0.9814	0.0005	0.0028
N12	0.0004	0.8945	0.0017	0.0003	0.0024	0.0997	0.0006	0.0004
N13	0.0016	0.9848	0.0033	0.0009	0.0004	0.0070	0.0004	0.0017
N14	0.0007	0.9904	0.0011	0.0006	0.0011	0.0049	0.0004	0.0007
N15	0.0018	0.7955	0.0031	0.0007	0.0016	0.1962	0.0004	0.0008
N16	0.0075	0.5837	0.0069	0.0013	0.0007	0.3963	0.0013	0.0025
N38	0.0284	0.0162	0.0007	0.0032	0.0007	0.9463	0.0008	0.0036
N40	0.0023	0.0072	0.0028	0.0006	0.0939	0.8922	0.0004	0.0006
N41	0.0014	0.0033	0.0020	0.0002	0.0008	0.9900	0.0012	0.0012
N42	0.0007	0.0009	0.0014	0.0001	0.0007	0.9944	0.0009	0.0009

Appendix 4. (Continued) TESS cluster membership of each population group at K = 8 (averaged over 10 runs).

	Partial membership to each cluster							
Population code	C1	C2	C3	C4	C5	C6	C7	C8
N43	0.0004	0.0009	0.0013	0.0007	0.0005	0.9951	0.0005	0.0005
N44	0.0256	0.0027	0.0026	0.0001	0.0011	0.9654	0.0010	0.0015
N45	0.2654	0.1550	0.0129	0.0004	0.0011	0.5516	0.0014	0.0125
N46	0.0012	0.0039	0.0024	0.0002	0.0004	0.9898	0.0012	0.0010
N47	0.0009	0.0068	0.0057	0.0001	0.0011	0.9838	0.0010	0.0007
N48	0.0001	0.0021	0.0020	0.0004	0.0005	0.9940	0.0004	0.0004
N50	0.0039	0.0195	0.0475	0.0001	0.9202	0.0039	0.0016	0.0033
N31	0.0030	0.0020	0.0013	0.0009	0.0005	0.9886	0.0005	0.0033
R1	0.0017	0.6273	0.0214	0.0002	0.0006	0.0016	0.0008	0.3466
R2	0.0015	0.0889	0.0023	0.0004	0.0013	0.0014	0.0529	0.8514
R18	0.0037	0.0054	0.0281	0.0002	0.0015	0.0015	0.0004	0.9593
R19	0.0016	0.0018	0.7642	0.0071	0.0208	0.0010	0.0078	0.1957
R20	0.0026	0.0010	0.0044	0.0004	0.0063	0.0007	0.0003	0.9843
R21	0.0073	0.0006	0.0042	0.0002	0.0375	0.0006	0.0012	0.9484
R22	0.0017	0.1465	0.0055	0.0005	0.0312	0.3004	0.0005	0.5137
R23	0.0937	0.0189	0.8025	0.0006	0.0032	0.0490	0.0047	0.0274
R24	0.0033	0.0072	0.0035	0.0003	0.0078	0.0092	0.0003	0.9685
R25	0.0199	0.0118	0.0047	0.0001	0.0059	0.0103	0.0019	0.9453
R26	0.0054	0.0012	0.0043	0.0001	0.0164	0.0180	0.0005	0.9541
R37	0.0422	0.0011	0.3714	0.0002	0.0016	0.0228	0.0002	0.5605
R32	0.0169	0.0039	0.0027	0.0004	0.1174	0.0957	0.0002	0.7629
R33	0.0016	0.0021	0.0019	0.0007	0.0038	0.0108	0.0014	0.9777
R34	0.0117	0.0928	0.0055	0.0001	0.0456	0.0029	0.0008	0.8407
R35	0.0062	0.0028	0.0071	0.0003	0.0787	0.0038	0.0027	0.8982
R36	0.0005	0.0005	0.0029	0.0001	0.0033	0.0007	0.0003	0.9918
R17	0.0520	0.0170	0.0009	0.0003	0.0007	0.0867	0.7809	0.0614
R51	0.0012	0.0200	0.7370	0.0002	0.0027	0.0348	0.2030	0.0013
R52	0.0006	0.0253	0.1024	0.0002	0.0011	0.0039	0.6678	0.1986
R28	0.0037	0.3831	0.5851	0.0001	0.0039	0.0035	0.0162	0.0043
R29A	0.1048	0.1169	0.4336	0.0004	0.0028	0.0794	0.0301	0.2323
R29B	0.0010	0.0525	0.8293	0.0005	0.0028	0.0972	0.0032	0.0138
R3	0.0127	0.0838	0.7085	0.0005	0.0881	0.0911	0.0021	0.0132
R4	0.0009	0.0056	0.9884	0.0001	0.0005	0.0016	0.0015	0.0013
R5A	0.0006	0.0081	0.9780	0.0004	0.0032	0.0011	0.0024	0.0063
R5B	0.0006	0.0018	0.9866	0.0003	0.0048	0.0011	0.0036	0.0014
R6	0.0039	0.0129	0.8636	0.0005	0.0031	0.0061	0.0012	0.1088
R7	0.0006	0.0032	0.9881	0.0002	0.0007	0.0016	0.0035	0.0020
R8	0.0009	0.0093	0.9605	0.0003	0.0073	0.0022	0.0057	0.0141
R9	0.0014	0.0149	0.9534	0.0003	0.0091	0.0067	0.0035	0.0107
R11	0.0087	0.0788	0.0093	0.0019	0.0051	0.2011	0.6413	0.0540
R12	0.0049	0.0998	0.6863	0.0003	0.0240	0.0734	0.0202	0.0911

Appendix 4. (Continued) TESS cluster membership of each population group at K = 8 (averaged over 10 runs).

	Partial membership to each cluster							
Population code	C1	C2	C3	C4	C5	C6	C7	C8
R14	0.0008	0.0024	0.9857	0.0012	0.0006	0.0015	0.0037	0.0041
R15	0.0007	0.0058	0.8163	0.0067	0.0013	0.1549	0.0104	0.0040
R16	0.0012	0.0084	0.9720	0.0006	0.0005	0.0133	0.0020	0.0020
R38	0.0003	0.0040	0.9871	0.0004	0.0016	0.0024	0.0037	0.0007
R39	0.0015	0.0009	0.4102	0.0002	0.0020	0.5576	0.0003	0.0273
R40	0.0003	0.9656	0.0237	0.0023	0.0003	0.0013	0.0027	0.0039
R41	0.0006	0.0044	0.8863	0.0017	0.0384	0.0642	0.0006	0.0039
R42	0.0029	0.0049	0.9541	0.0008	0.0016	0.0329	0.0010	0.0016
R44	0.0009	0.0033	0.2137	0.0009	0.0112	0.0158	0.7483	0.0059
R46	0.0006	0.0244	0.9576	0.0018	0.0004	0.0061	0.0022	0.0069
R47	0.0031	0.0095	0.0074	0.0013	0.0046	0.0179	0.0039	0.9523
R48	0.0114	0.0087	0.5963	0.0006	0.1612	0.2118	0.0021	0.0082
R49	0.0025	0.0449	0.2255	0.0005	0.0019	0.3871	0.0031	0.3347
R30	0.0004	0.0013	0.1005	0.0002	0.0003	0.0011	0.0004	0.8958
R31	0.0027	0.0044	0.1270	0.0003	0.0003	0.0197	0.0013	0.8443
R53	0.0001	0.0011	0.0015	0.0002	0.0003	0.0013	0.9953	0.0002
R54	0.0203	0.0012	0.9467	0.0005	0.0270	0.0020	0.0005	0.0017
R55	0.0011	0.0022	0.9925	0.0011	0.0007	0.0005	0.0004	0.0014
R56	0.0110	0.0060	0.4576	0.0009	0.0029	0.0026	0.0050	0.5141
R58	0.0010	0.2264	0.5374	0.0014	0.0344	0.1807	0.0050	0.0138
R59	0.0385	0.0014	0.8204	0.0064	0.0006	0.0006	0.0139	0.1180
R60	0.0009	0.0019	0.9915	0.0004	0.0010	0.0005	0.0009	0.0029
R61	0.0011	0.0013	0.9760	0.0009	0.0100	0.0084	0.0006	0.0018
R62	0.0025	0.0008	0.9775	0.0004	0.0099	0.0011	0.0015	0.0062
R64	0.0002	0.0006	0.0012	0.0006	0.0001	0.0008	0.9964	0.0001
R63	0.0003	0.0012	0.0035	0.0008	0.0004	0.0016	0.9920	0.0002
R65	0.0002	0.0018	0.0028	0.0013	0.0002	0.0017	0.9914	0.0004
R66	0.0000	0.0005	0.0007	0.0022	0.0001	0.0003	0.9961	0.0001
R67	0.0000	0.0006	0.0008	0.0006	0.0001	0.0004	0.9972	0.0001
M2	0.0000	0.0005	0.0002	0.9982	0.0000	0.0003	0.0007	0.0000
M1	0.0001	0.0002	0.0004	0.9987	0.0000	0.0001	0.0003	0.0001
M3	0.0000	0.0002	0.0007	0.9983	0.0001	0.0002	0.0005	0.0001
M4	0.0000	0.0004	0.0004	0.9979	0.0001	0.0003	0.0008	0.0000
M5	0.0000	0.0005	0.0009	0.9973	0.0000	0.0003	0.0008	0.0000

Appendix 5. Allele sizes and frequencies of 29 SSR markers in Asia Pacific *Oryza* series *Sativae*. Alleles that can distinguish all or certain population groups within a species are highlighted. Highly discriminating alleles are boldfaced.

Marker/ Allele size (bp)	Allele frequencies							
	O. nivara				*O. rufipogon*			*O. meridionalis*
	C1	C5	C2	C6	C8	C3	C7	C4
OSR13								
106	0.000	0.071	0.000	0.007	0.080	0.000	0.000	0.000
108	0.000	0.000	0.000	0.000	0.029	0.000	0.034	0.000
110	0.000	0.000	0.000	0.000	0.051	0.000	0.114	0.000
112	0.000	0.000	0.000	0.185	0.304	0.125	0.045	0.000
114	0.152	0.086	0.125	0.089	0.159	0.146	0.341	0.000
116	0.109	0.000	0.194	0.116	0.051	0.217	0.386	1.000
118	0.000	0.000	0.181	0.000	0.000	0.192	0.000	0.000
120	0.207	0.214	0.028	0.014	0.051	0.000	0.000	0.000
122	0.000	0.486	0.000	0.247	0.007	0.133	0.080	0.000
124	0.000	0.143	0.000	0.000	0.000	0.000	0.000	0.000
126	0.163	0.000	0.000	0.068	0.000	0.000	0.000	0.000
128	0.000	0.000	0.042	0.137	0.022	0.000	0.000	0.000
130	0.152	0.000	0.000	0.082	0.014	0.042	0.000	0.000
132	0.109	0.000	0.118	0.000	0.000	0.050	0.000	0.000
134	0.109	0.000	0.118	0.034	0.232	0.092	0.000	0.000
136	0.000	0.000	0.104	0.000	0.000	0.000	0.000	0.000
138	0.000	0.000	0.000	0.000	0.000	0.004	0.000	0.000
140	0.000	0.000	0.090	0.007	0.000	0.000	0.000	0.000
142	0.000	0.000	0.000	0.014	0.000	0.000	0.000	0.000
RM44								
111	0.000	0.000	0.000	0.000	0.029	0.213	0.102	0.000
113	0.000	0.029	0.000	0.000	0.036	0.317	0.000	0.000
117	0.000	0.000	0.000	0.000	0.094	0.008	0.330	0.000
119	0.000	0.143	0.097	0.000	0.080	0.025	0.045	0.083
121	0.370	0.000	0.042	0.233	0.254	0.058	0.000	0.208
123	0.000	0.057	0.028	0.068	0.196	0.050	0.295	0.167
125	0.065	0.143	0.014	0.151	0.036	0.104	0.091	0.000
127	0.000	0.486	0.222	0.068	0.029	0.075	0.000	0.125
129	0.554	0.143	0.250	0.308	0.029	0.088	0.068	0.000
131	0.000	0.000	0.250	0.000	0.159	0.000	0.000	0.000
133	0.000	0.000	0.097	0.110	0.058	0.000	0.068	0.000
135	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.208
137	0.000	0.000	0.000	0.000	0.000	0.017	0.000	0.000
139	0.000	0.000	0.000	0.055	0.000	0.046	0.000	0.000
141	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.208
143	0.011	0.000	0.000	0.000	0.000	0.000	0.000	0.000
149	0.000	0.000	0.000	0.007	0.000	0.000	0.000	0.000

Appendix 5. (Continued) Allele sizes and frequencies of 29 SSR markers in Asia Pacific *Oryza* series *Sativae*. Alleles that can distinguish all or certain population groups within a species are highlighted. Highly discriminating alleles are boldfaced.

Marker/ Allele size (bp)	Allele frequencies							
	O. nivara				*O. rufipogon*			*O. meridionalis*
	C1	C5	C2	C6	C8	C3	C7	C4
RM118								
153	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.125
171	0.000	0.000	0.014	0.000	0.000	0.000	0.000	0.875
173	0.000	0.000	0.000	0.000	0.000	0.000	0.466	0.000
175	0.422	0.743	0.569	0.178	0.565	0.447	0.227	0.000
177	0.000	0.143	0.000	0.014	0.101	0.148	0.000	0.000
179	0.556	0.114	0.417	0.630	0.246	0.340	0.307	0.000
181	0.000	0.000	0.000	0.041	0.087	0.066	0.000	0.000
185	0.022	0.000	0.000	0.082	0.000	0.000	0.000	0.000
195	0.000	0.000	0.000	0.055	0.000	0.000	0.000	0.000
RM124								
280	0.043	0.000	0.000	0.000	0.000	0.000	0.000	0.000
282	0.217	0.000	0.000	0.192	0.000	0.033	0.000	0.000
284	0.739	0.943	0.806	0.753	0.812	0.852	0.227	0.000
286	0.000	0.057	0.194	0.055	0.188	0.025	0.682	0.000
288	0.000	0.000	0.000	0.000	0.000	0.090	0.045	0.000
290	0.000	0.000	0.000	0.000	0.000	0.000	0.023	0.000
296	0.000	0.000	0.000	0.000	0.000	0.000	0.023	0.000
298	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000
RM125								
129	0.000	0.000	0.000	0.068	0.043	0.000	0.000	0.000
132	0.000	0.029	0.000	0.000	0.022	0.116	0.000	0.000
135	0.000	0.000	0.000	0.000	0.000	0.198	0.091	1.000
138	0.109	0.529	0.028	0.000	0.630	0.012	0.000	0.000
141	0.217	0.000	0.313	0.740	0.000	0.273	0.000	0.000
144	0.196	0.000	0.097	0.000	0.254	0.219	0.773	0.000
147	0.000	0.088	0.125	0.007	0.000	0.000	0.045	0.000
150	0.000	0.294	0.000	0.000	0.036	0.116	0.091	0.000
153	0.370	0.059	0.181	0.068	0.000	0.041	0.000	0.000
156	0.109	0.000	0.257	0.116	0.014	0.017	0.000	0.000
162	0.000	0.000	0.000	0.000	0.000	0.008	0.000	0.000
RM133								
243	0.413	0.971	0.882	0.514	0.174	0.246	0.205	0.208
245	0.152	0.000	0.056	0.375	0.652	0.324	0.534	0.792
247	0.435	0.000	0.014	0.014	0.174	0.225	0.023	0.000
249	0.000	0.014	0.049	0.097	0.000	0.164	0.193	0.000

Appendix 5. (Continued) Allele sizes and frequencies of 29 SSR markers in Asia Pacific *Oryza* series *Sativae*. Alleles that can distinguish all or certain population groups within a species are highlighted. Highly discriminating alleles are boldfaced.

Marker/ Allele size (bp)	Allele frequencies							
	O. nivara				*O. rufipogon*			*O. meridionalis*
	C1	C5	C2	C6	C8	C3	C7	C4
251	0.000	0.000	0.000	0.000	0.000	0.008	0.045	0.000
255	0.000	0.014	0.000	0.000	0.000	0.033	0.000	0.000
RM152								
149	0.380	0.000	0.076	0.055	0.087	0.008	0.000	0.000
152	0.000	0.000	0.125	0.205	0.181	0.016	0.000	0.000
155	0.087	0.000	0.000	0.014	0.000	0.000	0.011	1.000
158	0.174	0.857	0.104	0.000	0.391	0.541	0.284	0.000
161	0.359	0.143	0.611	0.658	0.283	0.373	0.659	0.000
164	0.000	0.000	0.000	0.000	0.029	0.000	0.000	0.000
167	0.000	0.000	0.076	0.027	0.014	0.053	0.000	0.000
170	0.000	0.000	0.007	0.041	0.014	0.000	0.045	0.000
173	0.000	0.000	0.000	0.000	0.000	0.008	0.000	0.000
RM154								
178	0.000	0.000	0.000	0.000	0.015	0.020	0.023	0.000
180	0.000	0.000	0.000	0.000	0.000	0.012	0.091	0.000
182	0.000	0.000	0.000	0.096	0.067	0.012	0.170	0.000
184	0.087	0.000	0.583	0.260	0.082	0.389	0.114	0.208
186	0.239	0.171	0.236	0.068	0.164	0.160	0.000	0.771
188	0.000	0.000	0.069	0.164	0.030	0.020	0.011	0.000
189	0.000	0.000	0.000	0.000	0.007	0.000	0.000	0.000
190	0.652	0.829	0.000	0.260	0.112	0.082	0.170	0.000
192	0.022	0.000	0.069	0.000	0.045	0.000	0.023	0.000
194	0.000	0.000	0.014	0.000	0.052	0.020	0.193	0.000
196	0.000	0.000	0.000	0.000	0.000	0.012	0.205	0.000
198	0.000	0.000	0.000	0.000	0.097	0.004	0.000	0.000
200	0.000	0.000	0.028	0.068	0.037	0.094	0.000	0.000
202	0.000	0.000	0.000	0.000	0.022	0.090	0.000	0.021
204	0.000	0.000	0.000	0.000	0.090	0.000	0.000	0.000
206	0.000	0.000	0.000	0.068	0.015	0.029	0.000	0.000
208	0.000	0.000	0.000	0.000	0.075	0.000	0.000	0.000
210	0.000	0.000	0.000	0.014	0.015	0.000	0.000	0.000
212	0.000	0.000	0.000	0.000	0.000	0.020	0.000	0.000
216	0.000	0.000	0.000	0.000	0.022	0.000	0.000	0.000
218	0.000	0.000	0.000	0.000	0.045	0.000	0.000	0.000
224	0.000	0.000	0.000	0.000	0.000	0.033	0.000	0.000
228	0.000	0.000	0.000	0.000	0.007	0.000	0.000	0.000

Appendix 5. (Continued) Allele sizes and frequencies of 29 SSR markers in Asia Pacific *Oryza* series *Sativae*. Alleles that can distinguish all or certain population groups within a species are highlighted. Highly discriminating alleles are boldfaced.

Marker/ Allele size (bp)	Allele frequencies							
	O. nivara				*O. rufipogon*			*O. meridionalis*
	C1	C5	C2	C6	C8	C3	C7	C4
RM161								
171	0.000	0.000	0.000	0.000	0.000	0.025	0.000	0.000
173	0.000	0.000	0.056	0.000	0.457	0.434	0.000	0.000
175	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.125
177	0.000	0.000	0.000	0.000	0.000	0.033	0.443	0.583
179	0.000	0.000	0.063	0.000	0.014	0.017	0.114	0.292
181	0.587	0.286	0.148	0.123	0.000	0.008	0.000	0.000
183	0.000	0.714	0.169	0.534	0.130	0.099	0.330	0.000
185	0.217	0.000	0.148	0.137	0.188	0.140	0.114	0.000
187	0.000	0.000	0.000	0.137	0.138	0.025	0.000	0.000
189	0.000	0.000	0.028	0.000	0.058	0.033	0.000	0.000
191	0.000	0.000	0.387	0.000	0.000	0.116	0.000	0.000
193	0.000	0.000	0.000	0.000	0.000	0.033	0.000	0.000
197	0.196	0.000	0.000	0.000	0.000	0.000	0.000	0.000
199	0.000	0.000	0.000	0.068	0.014	0.037	0.000	0.000
RM162								
220	0.065	0.000	0.000	0.000	0.000	0.000	0.000	0.000
222	0.000	0.000	0.000	0.000	0.043	0.000	0.000	0.000
224	0.043	0.000	0.000	0.068	0.007	0.008	0.000	0.000
226	0.000	0.000	0.014	0.000	0.051	0.033	0.000	0.000
228	0.000	0.000	0.056	0.014	0.000	0.045	0.000	0.958
230	0.000	0.686	0.299	0.000	0.181	0.120	0.182	0.000
232	0.098	0.000	0.139	0.123	0.232	0.186	0.000	0.000
234	0.413	0.000	0.326	0.425	0.116	0.244	0.455	0.000
236	0.033	0.029	0.000	0.068	0.036	0.136	0.341	0.000
238	0.043	0.143	0.042	0.219	0.080	0.025	0.000	0.000
240	0.000	0.143	0.118	0.000	0.007	0.116	0.000	0.000
242	0.000	0.000	0.000	0.000	0.101	0.021	0.000	0.000
244	0.000	0.000	0.000	0.000	0.101	0.000	0.000	0.000
246	0.000	0.000	0.000	0.000	0.000	0.008	0.000	0.000
250	0.000	0.000	0.007	0.000	0.000	0.041	0.000	0.000
256	0.000	0.000	0.000	0.000	0.014	0.008	0.000	0.000
260	0.000	0.000	0.000	0.000	0.000	0.000	0.023	0.000
262	0.304	0.000	0.000	0.027	0.000	0.000	0.000	0.000
264	0.000	0.000	0.000	0.000	0.014	0.000	0.000	0.000
266	0.000	0.000	0.000	0.055	0.000	0.000	0.000	0.042
268	0.000	0.000	0.000	0.000	0.014	0.008	0.000	0.000

Appendix 5. (Continued) Allele sizes and frequencies of 29 SSR markers in Asia Pacific *Oryza* series *Sativae*. Alleles that can distinguish all or certain population groups within a species are highlighted. Highly discriminating alleles are boldfaced.

Marker/ Allele size (bp)	Allele frequencies							
	O. nivara				*O. rufipogon*			*O. meridionalis*
	C1	C5	C2	C6	C8	C3	C7	C4
RM215								
146	0.000	0.029	0.000	0.000	0.000	0.000	0.000	0.000
148	0.000	0.000	0.014	0.000	0.022	0.020	0.000	0.917
150	0.000	0.000	0.000	0.000	0.036	0.000	0.023	0.000
152	0.000	0.486	0.000	0.034	0.051	0.127	0.000	0.042
154	0.022	0.086	0.097	0.082	0.145	0.205	0.034	0.000
156	0.000	0.200	0.000	0.014	0.174	0.184	0.659	0.000
158	0.000	0.029	0.333	0.034	0.174	0.107	0.057	0.000
160	0.207	0.000	0.042	0.055	0.130	0.160	0.000	0.042
162	0.326	0.000	0.250	0.212	0.065	0.045	0.068	0.000
164	0.185	0.000	0.042	0.185	0.094	0.090	0.034	0.000
166	0.109	0.029	0.069	0.288	0.051	0.041	0.080	0.000
168	0.130	0.000	0.042	0.082	0.014	0.008	0.000	0.000
170	0.022	0.143	0.014	0.000	0.007	0.004	0.045	0.000
172	0.000	0.000	0.000	0.014	0.014	0.008	0.000	0.000
176	0.000	0.000	0.083	0.000	0.000	0.000	0.000	0.000
178	0.000	0.000	0.000	0.000	0.022	0.000	0.000	0.000
184	0.000	0.000	0.014	0.000	0.000	0.000	0.000	0.000
RM237								
120	0.000	0.000	0.000	0.000	0.029	0.000	0.000	0.000
126	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.065
128	0.000	0.057	0.007	0.014	0.000	0.000	0.000	0.000
130	0.000	0.029	0.083	0.000	0.138	0.020	0.034	0.000
132	0.000	0.000	0.021	0.021	0.217	0.000	0.136	0.000
134	0.022	0.029	0.042	0.027	0.094	0.057	0.000	0.000
136	0.000	0.000	0.014	0.000	0.000	0.049	0.250	0.000
138	0.000	0.000	0.000	0.000	0.014	0.160	0.068	0.022
140	0.783	0.114	0.139	0.000	0.203	0.061	0.000	0.000
142	0.109	0.000	0.056	0.082	0.043	0.090	0.102	0.065
144	0.065	0.000	0.292	0.164	0.072	0.012	0.000	0.000
146	0.000	0.000	0.028	0.082	0.000	0.053	0.034	0.000
148	0.022	0.057	0.014	0.233	0.051	0.000	0.000	0.000
150	0.000	0.143	0.118	0.349	0.036	0.172	0.136	0.543
152	0.000	0.000	0.000	0.000	0.014	0.090	0.000	0.000
154	0.000	0.486	0.097	0.000	0.022	0.041	0.125	0.304
156	0.000	0.000	0.014	0.027	0.000	0.070	0.000	0.000
158	0.000	0.086	0.028	0.000	0.029	0.041	0.091	0.000
160	0.000	0.000	0.000	0.000	0.029	0.016	0.000	0.000

Appendix 5. (Continued) Allele sizes and frequencies of 29 SSR markers in Asia Pacific *Oryza* series *Sativae*. Alleles that can distinguish all or certain population groups within a species are highlighted. Highly discriminating alleles are boldfaced.

Marker/ Allele size (bp)	Allele frequencies							
	O. nivara				*O. rufipogon*			*O. meridionalis*
	C1	C5	C2	C6	C8	C3	C7	C4
166	0.000	0.000	0.000	0.000	0.000	0.004	0.000	0.000
168	0.000	0.000	0.000	0.000	0.000	0.029	0.000	0.000
170	0.000	0.000	0.049	0.000	0.007	0.000	0.023	0.000
176	0.000	0.000	0.000	0.000	0.000	0.033	0.000	0.000
RM271								
104	0.000	0.114	0.160	0.055	0.000	0.144	0.000	0.208
106	0.000	0.000	0.000	0.041	0.125	0.021	0.000	0.333
108	0.000	0.000	0.000	0.000	0.199	0.030	0.000	0.000
110	0.435	0.014	0.174	0.048	0.162	0.047	0.000	0.000
112	0.174	0.229	0.201	0.137	0.000	0.042	0.068	0.000
114	0.011	0.000	0.201	0.123	0.059	0.038	0.182	0.000
116	0.000	0.029	0.125	0.000	0.000	0.008	0.443	0.083
118	0.054	0.014	0.007	0.315	0.191	0.322	0.023	0.208
120	0.000	0.000	0.000	0.000	0.000	0.123	0.000	0.000
122	0.207	0.000	0.000	0.000	0.029	0.034	0.000	0.167
124	0.120	0.000	0.000	0.253	0.029	0.004	0.159	0.000
126	0.000	0.600	0.021	0.000	0.000	0.008	0.000	0.000
128	0.000	0.000	0.049	0.000	0.015	0.017	0.000	0.000
130	0.000	0.000	0.007	0.000	0.088	0.068	0.000	0.000
132	0.000	0.000	0.000	0.000	0.000	0.008	0.011	0.000
134	0.000	0.000	0.000	0.000	0.000	0.030	0.000	0.000
136	0.000	0.000	0.000	0.000	0.000	0.034	0.114	0.000
138	0.000	0.000	0.000	0.000	0.074	0.000	0.000	0.000
140	0.000	0.000	0.056	0.000	0.000	0.000	0.000	0.000
142	0.000	0.000	0.000	0.027	0.015	0.000	0.000	0.000
144	0.000	0.000	0.000	0.000	0.015	0.000	0.000	0.000
146	0.000	0.000	0.000	0.000	0.000	0.013	0.000	0.000
156	0.000	0.000	0.000	0.000	0.000	0.008	0.000	0.000
RM277								
128	0.000	0.000	0.014	0.000	0.000	0.000	0.000	0.000
130	0.000	0.057	0.000	0.000	0.000	0.000	0.000	0.000
132	0.000	0.000	0.000	0.027	0.000	0.020	0.000	0.000
134	0.000	0.000	0.458	0.082	0.587	0.176	0.193	1.000
136	0.000	0.000	0.250	0.315	0.000	0.402	0.000	0.000
138	0.848	0.800	0.222	0.466	0.341	0.250	0.239	0.000
140	0.152	0.143	0.042	0.110	0.072	0.152	0.455	0.000

Appendix 5. (Continued) Allele sizes and frequencies of 29 SSR markers in Asia Pacific *Oryza* series *Sativae*. Alleles that can distinguish all or certain population groups within a species are highlighted. Highly discriminating alleles are boldfaced.

Marker/ Allele size (bp)	Allele frequencies							
	O. nivara				*O. rufipogon*			*O. meridionalis*
	C1	C5	C2	C6	C8	C3	C7	C4
142	0.000	0.000	0.014	0.000	0.000	0.000	0.114	0.000
RM283								
151	0.000	0.029	0.000	0.000	0.015	0.004	0.000	0.000
153	0.000	0.000	0.000	0.000	0.000	0.008	0.114	0.000
155	0.000	0.000	0.000	0.021	0.044	0.000	0.000	0.000
157	0.000	0.400	0.014	0.000	0.000	0.109	0.000	0.000
159	0.000	0.114	0.000	0.000	0.029	0.050	0.000	0.000
161	0.000	0.000	0.000	0.127	0.015	0.084	0.307	0.000
163	0.023	0.000	0.153	0.070	0.015	0.034	0.193	0.000
165	0.000	0.029	0.000	0.014	0.081	0.042	0.000	0.000
167	0.114	0.000	0.028	0.000	0.000	0.000	0.000	0.000
169	0.114	0.171	0.000	0.085	0.140	0.239	0.000	0.167
171	0.205	0.143	0.486	0.056	0.316	0.134	0.239	0.000
173	0.000	0.114	0.069	0.049	0.000	0.076	0.000	0.000
175	0.205	0.000	0.028	0.155	0.250	0.021	0.000	0.625
177	0.182	0.000	0.000	0.127	0.059	0.084	0.034	0.208
179	0.068	0.000	0.181	0.028	0.000	0.008	0.000	0.000
181	0.000	0.000	0.042	0.014	0.000	0.042	0.114	0.000
183	0.091	0.000	0.000	0.000	0.000	0.046	0.000	0.000
185	0.000	0.000	0.000	0.183	0.000	0.008	0.000	0.000
187	0.000	0.000	0.000	0.070	0.037	0.000	0.000	0.000
197	0.000	0.000	0.000	0.000	0.000	0.008	0.000	0.000
RM284								
159	0.000	0.186	0.000	0.000	0.000	0.107	0.000	0.000
161	0.957	0.000	0.542	0.795	0.841	0.459	0.182	1.000
163	0.022	0.143	0.319	0.068	0.072	0.266	0.580	0.000
165	0.022	0.671	0.042	0.068	0.029	0.127	0.159	0.000
167	0.000	0.000	0.000	0.068	0.051	0.029	0.080	0.000
169	0.000	0.000	0.097	0.000	0.000	0.008	0.000	0.000
173	0.000	0.000	0.000	0.000	0.007	0.000	0.000	0.000
177	0.000	0.000	0.000	0.000	0.000	0.004	0.000	0.000
RM316								
134	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000
180	0.000	0.000	0.000	0.000	0.000	0.025	0.000	0.000
186	0.000	0.000	0.000	0.000	0.000	0.008	0.023	0.000
187	0.000	0.000	0.000	0.000	0.000	0.041	0.000	0.000

Appendix 5. (Continued) Allele sizes and frequencies of 29 SSR markers in Asia Pacific *Oryza* series *Sativae*. Alleles that can distinguish all or certain population groups within a species are highlighted. Highly discriminating alleles are boldfaced.

Marker/ Allele size (bp)	Allele frequencies							
	O. nivara				*O. rufipogon*			*O. meridionalis*
	C1	C5	C2	C6	C8	C3	C7	C4
215	0.000	0.000	0.000	0.000	0.051	0.000	0.000	0.000
216	0.011	0.000	0.000	0.000	0.000	0.000	0.000	0.000
217	0.000	0.000	0.000	0.000	0.000	0.029	0.023	0.000
218	0.000	0.114	0.083	0.397	0.181	0.182	0.291	0.000
219	0.000	0.000	0.000	0.205	0.000	0.000	0.000	0.000
220	0.000	0.000	0.000	0.068	0.000	0.008	0.105	0.000
221	0.000	0.000	0.000	0.000	0.000	0.000	0.070	0.000
222	0.000	0.000	0.472	0.082	0.022	0.091	0.023	0.000
223	0.000	0.000	0.000	0.000	0.000	0.017	0.000	0.000
224	0.989	0.029	0.069	0.123	0.000	0.037	0.000	0.000
225	0.000	0.000	0.014	0.000	0.043	0.004	0.000	0.000
226	0.000	0.143	0.000	0.000	0.130	0.012	0.047	0.000
227	0.000	0.543	0.000	0.000	0.116	0.074	0.000	0.000
228	0.000	0.000	0.007	0.000	0.000	0.025	0.174	0.000
229	0.000	0.143	0.000	0.000	0.000	0.004	0.000	0.000
230	0.000	0.000	0.139	0.000	0.225	0.182	0.163	0.000
231	0.000	0.000	0.000	0.000	0.000	0.017	0.000	0.000
233	0.000	0.029	0.000	0.123	0.014	0.025	0.000	0.000
235	0.000	0.000	0.146	0.000	0.007	0.099	0.000	0.000
236	0.000	0.000	0.000	0.000	0.036	0.000	0.000	0.000
237	0.000	0.000	0.014	0.000	0.014	0.008	0.000	0.000
238	0.000	0.000	0.056	0.000	0.000	0.045	0.000	0.000
239	0.000	0.000	0.000	0.000	0.058	0.000	0.000	0.000
240	0.000	0.000	0.000	0.000	0.080	0.062	0.000	0.000
242	0.000	0.000	0.000	0.000	0.000	0.004	0.012	0.000
245	0.000	0.000	0.000	0.000	0.000	0.000	0.070	0.000
274	0.000	0.000	0.000	0.000	0.022	0.000	0.000	0.000
RM338								
193	0.000	0.000	0.000	0.068	0.000	0.000	0.000	0.000
199	0.000	0.000	0.194	0.000	0.000	0.057	0.114	0.000
202	1.000	1.000	0.806	0.932	1.000	0.926	0.886	1.000
205	0.000	0.000	0.000	0.000	0.000	0.016	0.000	0.000
RM408								
130	0.022	0.000	0.000	0.000	0.000	0.000	0.000	0.000
134	0.000	0.000	0.014	0.137	0.000	0.061	0.000	0.375
136	0.196	0.829	0.000	0.164	0.109	0.246	0.080	0.000
138	0.130	0.000	0.806	0.336	0.094	0.066	0.091	0.625

Appendix 5. (Continued) Allele sizes and frequencies of 29 SSR markers in Asia Pacific *Oryza* series *Sativae*. Alleles that can distinguish all or certain population groups within a species are highlighted. Highly discriminating alleles are boldfaced.

Marker/ Allele size (bp)	Allele frequencies							
	O. nivara				*O. rufipogon*			*O. meridionalis*
	C1	C5	C2	C6	C8	C3	C7	C4
140	0.489	0.171	0.056	0.205	0.493	0.508	0.830	0.000
142	0.000	0.000	0.097	0.041	0.116	0.045	0.000	0.000
144	0.098	0.000	0.000	0.068	0.101	0.070	0.000	0.000
146	0.065	0.000	0.028	0.041	0.051	0.000	0.000	0.000
148	0.000	0.000	0.000	0.007	0.000	0.000	0.000	0.000
150	0.000	0.000	0.000	0.000	0.000	0.004	0.000	0.000
154	0.000	0.000	0.000	0.000	0.036	0.000	0.000	0.000
RM413								
81	0.000	0.000	0.000	0.000	0.000	0.004	0.000	0.000
83	0.109	0.000	0.000	0.000	0.000	0.008	0.114	0.000
85	0.522	0.829	0.194	0.171	0.391	0.260	0.091	0.000
87	0.087	0.000	0.118	0.000	0.217	0.091	0.114	0.000
89	0.000	0.000	0.000	0.000	0.065	0.008	0.000	0.000
91	0.065	0.143	0.194	0.219	0.000	0.099	0.000	0.000
93	0.000	0.000	0.021	0.137	0.043	0.033	0.000	0.000
95	0.000	0.000	0.111	0.075	0.080	0.099	0.114	0.000
97	0.000	0.029	0.194	0.075	0.043	0.198	0.000	0.000
99	0.217	0.000	0.007	0.000	0.000	0.054	0.000	0.000
101	0.000	0.000	0.000	0.000	0.072	0.012	0.000	0.000
103	0.000	0.000	0.083	0.068	0.000	0.074	0.227	0.000
105	0.000	0.000	0.000	0.000	0.007	0.025	0.227	0.000
107	0.000	0.000	0.000	0.000	0.000	0.012	0.114	0.000
109	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.708
111	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.292
113	0.000	0.000	0.000	0.000	0.036	0.008	0.000	0.000
115	0.000	0.000	0.056	0.000	0.000	0.000	0.000	0.000
117	0.000	0.000	0.021	0.000	0.000	0.008	0.000	0.000
119	0.000	0.000	0.000	0.199	0.000	0.004	0.000	0.000
121	0.000	0.000	0.000	0.000	0.043	0.000	0.000	0.000
123	0.000	0.000	0.000	0.055	0.000	0.000	0.000	0.000
RM431								
248	0.000	0.000	0.000	0.055	0.000	0.000	0.000	0.000
250	0.000	0.000	0.000	0.000	0.000	0.008	0.000	0.000
252	0.109	0.714	0.014	0.000	0.058	0.008	0.000	0.167
254	0.109	0.000	0.007	0.000	0.043	0.102	0.000	0.000
256	0.022	0.029	0.000	0.000	0.130	0.090	0.364	0.000
258	0.011	0.000	0.056	0.068	0.319	0.238	0.114	0.000

Appendix 5. (Continued) Allele sizes and frequencies of 29 SSR markers in Asia Pacific *Oryza* series *Sativae*. Alleles that can distinguish all or certain population groups within a species are highlighted. Highly discriminating alleles are boldfaced.

Marker/ Allele size (bp)	Allele frequencies							
	O. nivara				*O. rufipogon*			*O. meridionalis*
	C1	C5	C2	C6	C8	C3	C7	C4
260	0.000	0.029	0.000	0.041	0.116	0.283	0.341	0.000
262	0.652	0.000	0.215	0.068	0.181	0.004	0.000	0.208
264	0.022	0.000	0.153	0.151	0.094	0.078	0.045	0.208
266	0.000	0.229	0.194	0.260	0.014	0.090	0.000	0.083
268	0.054	0.000	0.194	0.226	0.007	0.082	0.102	0.000
270	0.022	0.000	0.069	0.055	0.022	0.016	0.000	0.000
272	0.000	0.000	0.000	0.068	0.014	0.000	0.000	0.125
276	0.000	0.000	0.083	0.000	0.000	0.000	0.023	0.000
278	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.208
280	0.000	0.000	0.014	0.007	0.000	0.000	0.000	0.000
282	0.000	0.000	0.000	0.000	0.000	0.000	0.011	0.000
RM433								
226	0.000	0.000	0.000	0.000	0.000	0.021	0.000	0.000
232	0.000	0.000	0.000	0.000	0.044	0.000	0.000	0.000
234	0.000	0.000	0.000	0.000	0.140	0.045	0.080	0.000
236	0.000	0.400	0.222	0.342	0.125	0.087	0.432	0.688
238	0.000	0.114	0.194	0.055	0.199	0.062	0.068	0.000
240	0.000	0.029	0.132	0.068	0.169	0.273	0.273	0.188
242	0.174	0.200	0.028	0.027	0.037	0.260	0.000	0.083
244	0.022	0.000	0.000	0.055	0.051	0.033	0.000	0.000
246	0.304	0.257	0.056	0.151	0.088	0.120	0.011	0.000
248	0.391	0.000	0.000	0.000	0.096	0.050	0.114	0.000
250	0.109	0.000	0.181	0.116	0.015	0.029	0.000	0.000
252	0.000	0.000	0.111	0.068	0.000	0.021	0.011	0.000
254	0.000	0.000	0.069	0.089	0.000	0.000	0.000	0.000
256	0.000	0.000	0.000	0.000	0.037	0.000	0.000	0.000
278	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.042
280	0.000	0.000	0.000	0.027	0.000	0.000	0.000	0.000
314	0.000	0.000	0.000	0.000	0.000	0.000	0.011	0.000
326	0.000	0.000	0.007	0.000	0.000	0.000	0.000	0.000
RM447								
122	0.000	0.057	0.000	0.014	0.000	0.017	0.000	1.000
125	0.000	0.000	0.000	0.014	0.022	0.000	0.000	0.000
128	0.000	0.629	0.183	0.090	0.179	0.054	0.114	0.000
131	0.000	0.000	0.000	0.042	0.007	0.004	0.159	0.000
134	0.283	0.000	0.042	0.063	0.366	0.017	0.193	0.000
137	0.000	0.000	0.014	0.167	0.142	0.308	0.000	0.000

Appendix 5. (Continued) Allele sizes and frequencies of 29 SSR markers in Asia Pacific *Oryza* series *Sativae*. Alleles that can distinguish all or certain population groups within a species are highlighted. Highly discriminating alleles are boldfaced.

Marker/ Allele size (bp)	Allele frequencies							
	O. nivara				*O. rufipogon*			*O. meridionalis*
	C1	C5	C2	C6	C8	C3	C7	C4
140	0.000	0.000	0.056	0.181	0.000	0.250	0.250	0.000
143	0.000	0.143	0.528	0.250	0.037	0.125	0.193	0.000
146	0.000	0.171	0.000	0.000	0.187	0.133	0.091	0.000
149	0.522	0.000	0.014	0.125	0.000	0.063	0.000	0.000
152	0.152	0.000	0.070	0.000	0.000	0.004	0.000	0.000
155	0.000	0.000	0.000	0.000	0.022	0.000	0.000	0.000
158	0.000	0.000	0.092	0.056	0.000	0.000	0.000	0.000
161	0.000	0.000	0.000	0.000	0.000	0.025	0.000	0.000
164	0.043	0.000	0.000	0.000	0.037	0.000	0.000	0.000
RM452								
209	0.000	0.000	0.000	0.000	0.029	0.000	0.000	0.000
212	0.000	0.000	0.014	0.000	0.000	0.000	0.000	0.000
215	0.000	0.000	0.000	0.000	0.000	0.000	0.057	0.000
218	0.511	0.000	0.000	0.000	0.094	0.025	0.000	0.000
221	0.000	0.000	0.000	0.000	0.000	0.000	0.068	0.000
224	0.315	0.857	0.903	0.699	0.725	0.934	0.875	0.792
227	0.109	0.000	0.069	0.164	0.145	0.000	0.000	0.208
230	0.065	0.000	0.014	0.137	0.000	0.033	0.000	0.000
233	0.000	0.143	0.000	0.000	0.007	0.008	0.000	0.000
RM455								
145	0.891	1.000	0.986	0.986	0.746	0.775	0.989	1.000
149	0.109	0.000	0.014	0.000	0.254	0.217	0.011	0.000
153	0.000	0.000	0.000	0.014	0.000	0.008	0.000	0.000
RM484								
302	0.000	0.000	0.069	0.000	0.000	0.045	0.000	0.000
304	0.000	0.143	0.194	0.404	0.210	0.049	0.000	0.000
306	0.000	0.829	0.153	0.014	0.203	0.447	0.568	0.000
308	0.261	0.000	0.431	0.137	0.529	0.352	0.318	0.000
310	0.739	0.029	0.153	0.377	0.000	0.078	0.114	0.000
312	0.000	0.000	0.000	0.068	0.058	0.000	0.000	1.000
316	0.000	0.000	0.000	0.000	0.000	0.029	0.000	0.000
RM495								
165	0.467	0.500	0.972	0.932	0.594	0.803	0.523	1.000
171	0.000	0.000	0.000	0.000	0.080	0.102	0.432	0.000
174	0.533	0.500	0.007	0.062	0.232	0.020	0.011	0.000

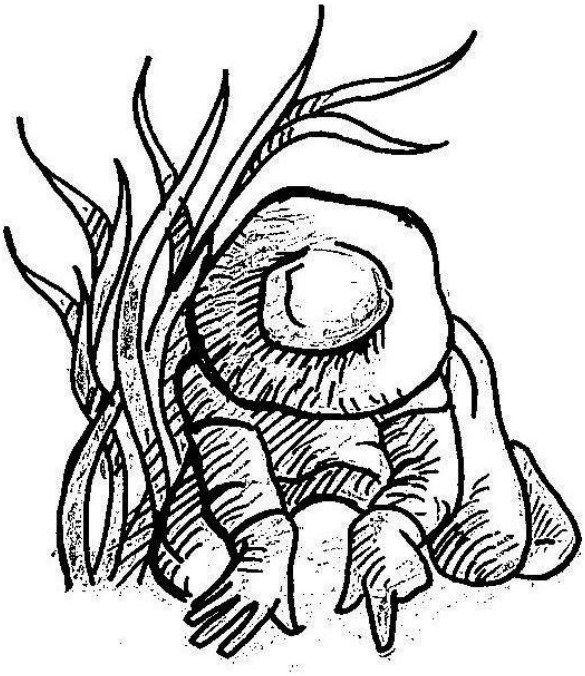
Appendix 5. (Continued) Allele sizes and frequencies of 29 SSR markers in Asia Pacific *Oryza* series *Sativae*. Alleles that can distinguish all or certain population groups within a species are highlighted. Highly discriminating alleles are boldfaced.

Marker/ Allele size (bp)	Allele frequencies							
	O. nivara				*O. rufipogon*			*O. meridionalis*
	C1	C5	C2	C6	C8	C3	C7	C4
177	0.000	0.000	0.021	0.007	0.094	0.074	0.034	0.000
RM507								
250	0.000	0.000	0.000	0.000	0.000	0.000	0.114	1.000
258	0.000	0.000	0.000	0.068	0.007	0.020	0.443	0.000
266	0.000	0.000	0.000	0.000	0.000	0.000	0.057	0.000
270	0.674	1.000	0.778	0.842	0.746	0.898	0.239	0.000
274	0.326	0.000	0.222	0.089	0.246	0.078	0.114	0.000
278	0.000	0.000	0.000	0.000	0.000	0.000	0.011	0.000
282	0.000	0.000	0.000	0.000	0.000	0.004	0.023	0.000
RM536								
235	0.109	0.000	0.028	0.000	0.000	0.000	0.000	0.000
237	0.435	0.000	0.000	0.260	0.072	0.037	0.000	0.000
239	0.109	0.000	0.042	0.000	0.000	0.000	0.000	0.000
241	0.000	0.000	0.070	0.000	0.145	0.123	0.364	0.000
243	0.109	0.000	0.493	0.164	0.667	0.328	0.500	0.750
245	0.022	0.886	0.225	0.507	0.080	0.504	0.136	0.250
247	0.000	0.000	0.056	0.000	0.000	0.000	0.000	0.000
249	0.000	0.000	0.014	0.068	0.007	0.008	0.000	0.000
251	0.174	0.114	0.070	0.000	0.029	0.000	0.000	0.000
253	0.043	0.000	0.000	0.000	0.000	0.000	0.000	0.000

CHAPTER 4
Crossability patterns in Asia Pacific *Oryza* series *Sativae*

Maria Celeste N. Banaticla-Hilario[1,2], Kenneth L. McNally[1], Ronald G. van den Berg[2] and Nigel Ruaraidh Sackville Hamilton[1]

[1]*T.T. Chang Genetic Resources Center, International Rice Research Institute, Los Baños, Laguna, Philippines*
[2]*Biosystematics Group, Wageningen University and Research Center, Wageningen, The Netherlands*



Abstract

Artificial crossing experiments using 15 parental accessions of *O. meridionalis*, *O. nivara*, and *O. rufipogon* were conducted to assess the extent of post-pollination reproductive isolation within and among these Asia Pacific *Oryza* series *Sativae* species.

Reproductive incompatibility was observed within and between *O. meridionalis* and *O. nivara*. Intra- and interspecific crosses of these selfing species had very low seed set and produced inviable hybrid seeds indicative of strong pre- and post-zygotic barriers. Contrastingly, the outcrossing *O. rufipogon* exhibited high intraspecific crossability and modest compatibility with *O. nivara* and *O. meridionalis* in terms of seed set suggesting substantial pre-zygotic reproductive isolation of the species. Among the intraspecific combinations of *O. rufipogon*, only those with parental distances that ranged from 1062 to 3813 kilometers produced viable and non-sterile F1 hybrids while those involving maritime Southeast Asian and/or Australasian accessions were the most reproductively successful.

Crossability between *O. meridionalis* and *O. rufipogon* was asymmetric and favoured combinations with maternal *O. meridionalis* rather than crosses with maternal *O. rufipogon*. Concurrently, *O. nivara* and *O. rufipogon* showed symmetric compatibility where reciprocal crosses displayed similar crossability estimates. *O. nivara* and *O. meridionalis* manifested comparable degrees of isolation from *O. rufipogon* despite differences in strength of several post-zygotic barriers. Crosses between *O. nivara* and *O. rufipogon* produced few hybrids (14 hybrids from 42 panicles) with annual habit and relatively high fertility (36.9% mean F1 fertility) while crosses between *O. meridionalis* and *O. rufipogon* produced more hybrids (39 hybrids from 30 panicles) with perennial habit and reduced fertility (17.2% mean F1 fertility).

Mating compatibility within and between the Asia Pacific species of *Oryza* series *Sativae* is not strongly spatially influenced, but some resistance to gene flow under sympatric conditions was observed.

Although more vegetatively robust and more late-flowering than their parent accessions, intraspecific *O. rufipogon* hybrids conformed to the morphology and life cycle of their parental species. Intra- and interspecific hybrids of Australasian *O. rufipogon* differed phenotypically from crosses with non-Australasian populations. Interspecific hybrids displayed both intermediate and parental character traits. *O.*

nivara and *O. rufipogon* generated early-flowering hybrids that are more similar to the former. *O. meridionalis* and *O. rufipogon* produced hybrids that varied in phenology and morphology. Crosses with maternal *O. meridionalis* yielded early flowering hybrids while crosses with maternal *O. rufipogon* resulted in late flowering F1 plants. Moreover, hybrids with Australasian *O. rufipogon* as paternal parent appeared similar to *O. meridionalis* while those with non-Australasian *O. rufipogon* as either maternal or paternal parent generally appeared similar to *O. rufipogon*.

Introduction

Natural hybridization and its contributions to speciation have been commonly recognized in higher plants (Stebbins 1959; Ellstrand et al. 1996; Arnold 1997). On the basis of the assumption that mating compatibility weakens with increasing genetic distance (Rieseberg and Carney 1998; Mallet 2005; Mallet 2008), data from crossing experiments have been and are still used to evaluate intra- and interspecific genetic relationships and test the accuracy of taxonomic schemes in many different plant groups (for example Andersson 1993; Raimondi et al. 2003; Favero et al. 2006; Pellegrino et al. 2008; Marcussen and Borgen 2011). Novel phylogenetic approaches are being developed to be able accommodate reticulate evolutionary patterns resulting from hybridization (Vriesendorp and Bakker 2005).

In the genus *Oryza*, one of the first accounts of natural crossing between wild and cultivated taxa was reported by Roschevics in 1931. In 1960, Nezu et al. analyzed the crossability among 17 *Oryza* species and detected interfertility among the Asian members as well as among the African species of series *Sativae* although some taxonomically dubious species names (e.g., *O. perennis* Moench, *O. stapfii* Roshev.) were used in their study. Much later, more extensive hybridization experiments on *Oryza* series *Sativae* were conducted within the T.T. Chang - Genetic Resources Center (TTC-GRC) of the International Rice Research Institute (IRRI), the Philippines, where Naredo et al. (1997, 1998) assessed the crossability among the Asian (*O. nivara* Sharma & Shastry and *O. rufipogon* Griff.), Australasian (*O. meridionalis* Ng) and tropical American (*O. glumaepatula* Steud.) species and Juliano et al. (2005) examined the reproductive compatibility between geographic populations of *O. meridionalis*.

The current research focuses on the three wild species of *Oryza* series *Sativae* in Asia Pacific. *O. rufipogon* is perennial, predominantly outcrossing, photoperiod sensitive and thrives in permanently submerged areas of tropical Asia and northern Australia. *O. nivara* and *O. meridionalis* are both annual, inbreeding, photoperiod insensitive and can be found in seasonally wet habitats. *O. rufipogon* overlaps with the former in continental Asia and with the latter in Australasia. *O. meridionalis* is widely accepted as a taxonomic species (Ng et al. 1981; Duistermaat 1987; Juliano et al. 2005) while *O. nivara* (considered in this study as a species) is sometimes treated as an ecotype of *O. rufipogon* (Tateoka 1963; Oka 1988; Vaughan et al. 2003). Recent studies revealed that these two Asian taxa differ morphologically (Banaticla-Hilario et al. in Chapter 2) and have genetically distinct populations that overlap across their distribution but are differentiated at the local scale (Banaticla-Hilario et al. in Chapter 3). Despite previous experiments (Naredo et al. 1997, 1998), the full extent of reproductive isolation between *O. nivara* and *O. rufipogon* remains to be established.

This study examines the strength of post-pollination isolation within and between the three species in relation to other factors that were not accounted for in the earlier crossing studies such as the spatial distance between the parents' geographic origins, and sympatricity vs. non-sympatricity of parents. The specific objectives are to: a) compare the crossability exhibited by different intra and interspecific combinations of Asia Pacific *Oryza* series *Sativae* species; b) determine the correlation between crossability and spatial distance between parental origins; c) test whether reproductive success differs between sympatric and non-sympatric crosses; and d) evaluate the morphology of the obtained F1 hybrids.

Materials and methods

Crossing experiments

A set of 15 parental accessions from the International Rice Genebank (IRG) at the International Rice Research Institute (IRRI) in the Philippines was selected to represent the three species across their geographic range and include sympatric population pairs of *O. nivara* and *O. rufipogon* and of *O. meridionalis* and *O. rufipogon* (Figure 1, Table 1). This material has been previously phenotyped (Banaticla-Hilario et al. in Chapter 2) and genotyped with simple sequence repeat (SSR) markers (Banaticla-Hilario et al. in Chapter 3).

Figure 1. Geographic distribution of the 15 parental accessions used in the crossing experiments.

Table 1. Parental accessions and their mean seed set obtained from 25 self-pollinated panicles (five panicles from each of the five plants per accession).

Code	IRGC No.	Species	Geographic origin	Seed set (Mean ± Standard deviation)
M1	86539	*O. meridionalis*	Australia	43.7 ± 31.7
M2	93260	*O. meridionalis*	Indonesia	48.4 ± 18.9
M4	105294	*O. meridionalis*	Australia	34.5 ± 24.6[a]
N2	103830	*O. nivara*	Bangladesh	35.8 ± 19.7
N18	80549	*O. nivara*	India	55.5 ± 22.7[a]
N51	86496	*O. nivara*	Vietnam	71.2 ± 11.0
R2	103827	*O. rufipogon*	Bangladesh	24.4 ± 20.6[b]
R18	80550	*O. rufipogon*	India	52.5 ± 16.0[c]
R30	106357	*O. rufipogon*	Myanmar	49.6 ± 32.3
R51	86512	*O. rufipogon*	Vietnam	37.2 ± 23.2
R53	80774	*O. rufipogon*	Philippines	55.0 ± 15.3
R59	105952	*O. rufipogon*	Indonesia	40.3 ± 18.2
R63	86542	*O. rufipogon*	Australia	47.6 ± 31.4
R64	93274	*O. rufipogon*	Indonesia	32.1 ± 11.4
R66	105293	*O. rufipogon*	Australia	52.4 ± 21.1

[a] mean of 15 panicles; [b] mean of 20 panicles; [c] mean of 24 panicles

Five plants from each accession were grown at the TTC-GRC Screenhouse at IRRI where all artificial pollination experiments and succeeding phenotyping of hybrids were conducted. To achieve self-pollination, five panicles from each plant were bagged prior to anthesis and harvested once the grains ripened. The seed set of each selfed plant was then determined by obtaining the average percentage of filled spikelets from the five bagged panicles (Appendix 1). The seed set of each accession was also derived from the mean seed set of the 25 bagged panicles (Table 1).

A total of 25,692 spikelets from 803 panicles were artificially pollinated in 114 combinations: 8 within accessions, 37 between accessions of the same species, and 69 between species (Appendix 2). Emasculation of female parents, artificial pollination, and harvesting of hybrid seeds was performed as described by Naredo et al. (1997) while hybrid seed germination and hybrid plant establishment adhered to the methods of Naredo et al. (1998).

Evaluation of hybrids

The 145 F1 plants obtained from the artificial crosses were evaluated using ten characters that can discriminate between all or certain pairs of the three parental species (based on Table 6 and Figure 8 in Banaticla-Hilario et al. in Chapter 2). These were culm length (cm), leaf length (cm), flag leaf width (cm), spikelet length (mm) and width (mm), awn length (mm), anther length (mm), anther length to spikelet length ratio, spikelet fertility (i.e., seed set of selfed panicles) and number of days from seeding to first heading. Measurements were obtained and fresh leaf samples (for DNA extraction) were collected from each hybrid plant. Herbarium voucher specimens of each hybrid were collected and deposited at the IRG Herbarium.

DNA samples were extracted and genotyped as described in Fulton et al. (1995) and Banaticla-Hilario et al. (in Chapter 3), respectively. Based on the SSR profile of the parents, certain combinations of 26 microsatellite markers were used to validate the true F1 hybrids of different crosses (Appendix 3).

Data analyses

All crossing and phenotype data were analysed using R 2.14 (R Development Core Team 2011). Five crossability estimates were examined: 1) seed set (the number of filled spikelets as a percentage of the total number of pollinated spikelets); 2) germinability (the number of germinated seeds as a percentage of the total number

of filled spikelets; 3) seedling survival (the number of established F1 plants as a percentage of the total number of germinated seeds); 4) cross fertility (the number of non-sterile hybrids as a percentage of the total number of established F1 plants); and 5) F1 fertility (the mean seed set of the selfed (bagged) panicles of the resulting mature fertile F1 hybrids).

Box and whisker plots of these estimates were constructed using the boxplot() function. To determine whether crossability across all and within certain species combinations varies with different factors (e.g., maternal/paternal species, sympatricity of parents), a one-way analysis of variance (ANOVA) was conducted with the car package (Fox and Weisberg 2011). The correlation of crossability estimates with spatial distance between the geographic origin of parents was also tested using the cor.test() function of the base package (R Development Core Team 2011).

Parental and hybrid phenotypes were compared by conducting a principal components analysis (PCA) among distinguishing traits using the prcomp() function. Anther length to spikelet length ratio was excluded in the PCA due to its high correlation with anther length.

Anther length, anther length to spikelet length ratio, awn length and spikelet width can readily distinguish the three parental species from each other (Banaticla-Hilario et al. in Chapter 2). Scatter plots of these highly discriminating characters were constructed to determine whether the interspecific hybrids would exhibit intermediate morphology. Measurements from ten genetically confirmed hybrid accessions (identified by Banaticla-Hilario et al. in Chapter 3 as intermediate forms between *O. nivara* and *O. rufipogon*) were incorporated in the plots for comparison and validation.

Results

Intraspecific crossability

Self-compatibility of parental accessions

Fertility of self-pollinated panicles does not differ significantly among the three species, among accessions within *O. meridionalis* and among accessions within *O. rufipogon*. However, accessions of *O. nivara* differ significantly at the $p < 0.05$ level.

All species show high variability in seed set among plants of the same accession and in a few cases, even among panicles from the same plant (Appendix 1, Table 1).

Artificial crosses within and between accessions

The crossability results of each cross combination are tabulated in Appendix 2 while the ANOVA results are listed in Table 2. Comparisons of crossability among the different intra- and interspecific crosses are presented in Figure 2. Variations in the mean seed set of artificially crossed panicles and in the mean seed set of self-pollinated F1 hybrids among specific parental combinations are displayed in Figures 3 and 4, respectively.

Three out of five crosses within *O. meridionalis* and four out of six within *O. nivara* had zero seed set (Appendix 2, Figure 3). In *O. meridionalis,* the combinations M1 x M2 and M1 x M1 only produced seeds that however, failed to germinate (Appendix 2). In *O. nivara*, the crosses between N18 and N51 produced filled spikelets that germinated but did not survive the seedling stage (Appendix 2). Intraspecific combinations (artificial crosses within and between accessions) produced fewer seeds than the selfed (i.e., bagged) panicles in both annual species (Figure 3).

In contrast, out of 34 crosses within *O. rufipogon,* 11 generated non-sterile hybrids and only four had zero seed set (Appendix 2; Figures 3 - 4). Intraspecific *O. rufipogon* combinations exhibited significantly higher seed set than all the other intraspecific and interspecific combinations (at $p < 0.0001$ level) and were also more successful than crosses between and within the two annual species in terms of survival, cross fertility and F1 fertility (Table 2; Figure 2).

In *O. rufipogon*, crosses within the same accession displayed lower seed set ($p < 0.05$) but higher germinability ($p < 0.01$) than crosses between different accessions (Table 2). The germinated seeds of these intrapopulation crosses did not survive the seedling stage (Table 2). Artificially pollinated panicles of both intra- and interpopulation crosses exhibited significantly lower seed set ($p < 0.0001$) than the selfed panicles (43.81%) (Figure 3). The seed set of combinations R30 x R2 (32.3%), R51 x R64 (46.4%) and R53 x R64 (51.4%) exceeded the seed set (of self-pollinated panicles) of one or both of their corresponding parental accessions (Figure 3). The most successful combinations in terms of F1 fertility were R59 x R30 (69.3%), R63 x R66 (60.1%) and R53 x R59 (55.1%) (Figure 4).

Table 2. Variations in intra- and interspecific crossability of Asia Pacific *Oryza* series *Sativae*. Mean ± standard error and ANOVA *F* and *P* values of each parameter are presented (N = number of replicates in each cross).

Type of cross	Seed set		Germinability		Survival		Cross fertility		F1 fertility	
	N	mean ± SE	N	mean ± SE	N	mean ± SE	N	mean ± SE	N	mean ± SE
Intraspecific										
M x M	25	2.66 ± 1.04	6	0.00 ± 0.00	-	-	-	-	-	-
N x N	28	2.58 ± 1.67	5	100.00 ± 0.00	5	0.00 ± 0.00	-	-	-	-
R x R	266	14.01 ± 1.16	163	55.69 ± 3.18	125	13.63 ± 2.19	40	85.36 ± 4.13	39	35.73 ± 3.49
Interspecific										
M x N	24	0.00 ± 0.00	-	-	-	-		-	-	-
N x M	16	0.48 ± 0.48	1	100.00	1	0.00	-	-	-	-
M x R	87	6.37 ± 1.57	21	43.47 ± 7.67	17	42.23 ± 9.65	11	100.00 ± 0.00	11	18.67 ± 2.45
R x M	76	2.38 ± 0.49	27	38.89 ± 9.29	12	16.67 ± 11.24	2	50.00 ± 50.00	1	0.77
N x R	153	6.26 ± 1.20	54	58.57 ± 5.74	41	4.29 ± 2.08	5	100.00 ± 0.00	5	31.79 ± 3.99
R x N	128	5.46 ± 0.89	57	74.03 ± 5.90	38	5.26 ± 3.00	4	100.00 ± 0.00	4	43.33 ± 10.42
ANOVA										
F		10.532		4.943		5.997		2.700		2.696
P		0.000		0.000		0.000		0.039		0.040
Within *O. rufipogon*										
Intrapopulation	18	3.59 ± 1.43	7	100.00 ± 0.00	7	0.00 ± 0.00	-	-	-	-
Interpopulation	248	14.77 ± 1.23	156	53.70 ± 3.24	118	14.44 ± 2.30	40	85.36 ± 4.13	39	35.73 ± 3.49
ANOVA										
F		5.937		9.131		2.327		1.923		2.332
P		0.015		0.003		0.130		0.167		0.128

*M = *O. meridionalis;* N = *O. nivara;* R = *O. rufipogon*

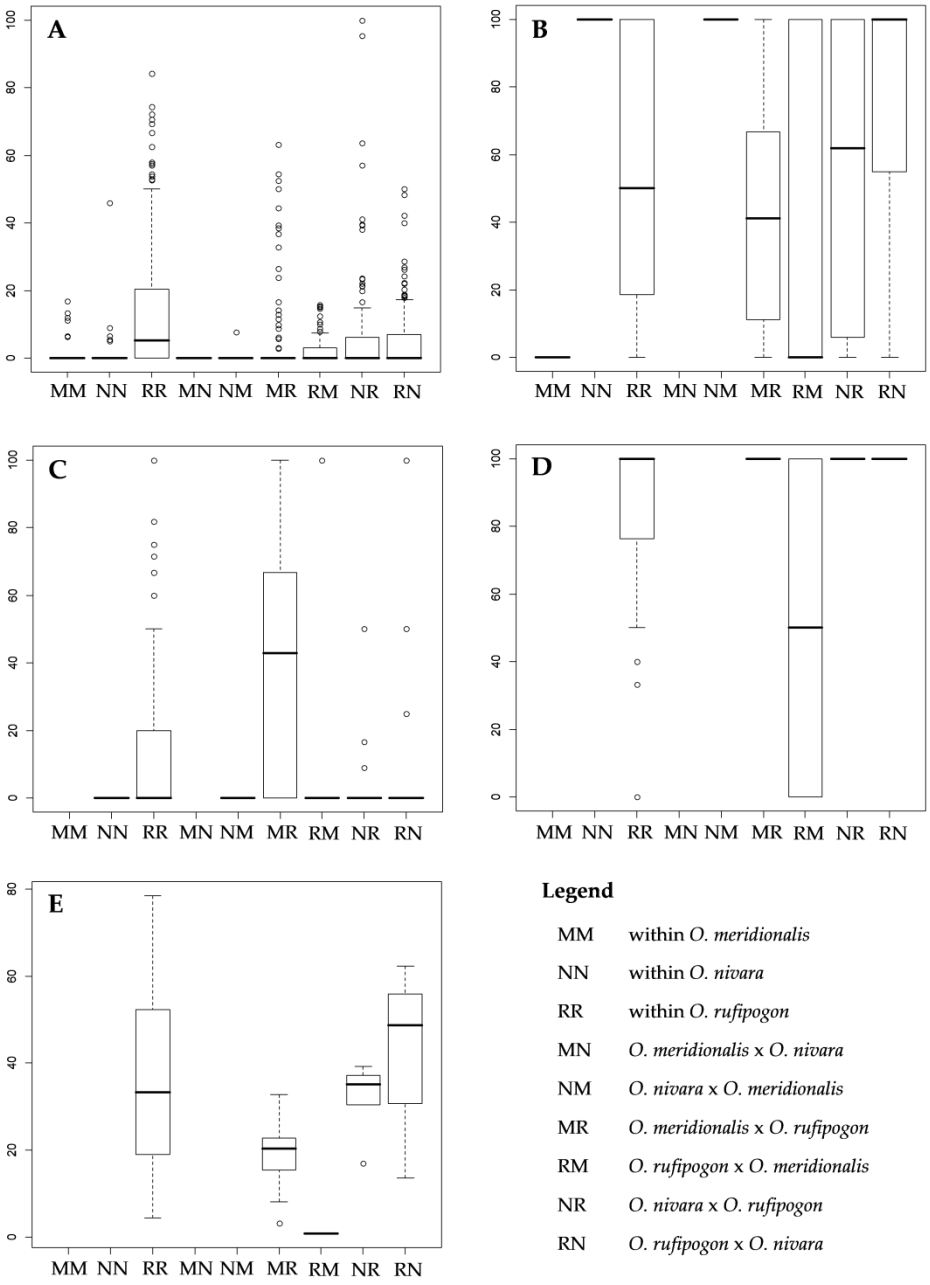


Figure 2. Crossability within and between *Oryza* series *Sativae* species in Asia Pacific: **A**) seed set; **B**) germinability; **C**) seedling survival; **D**) cross fertility; and **E**) F1 fertility.

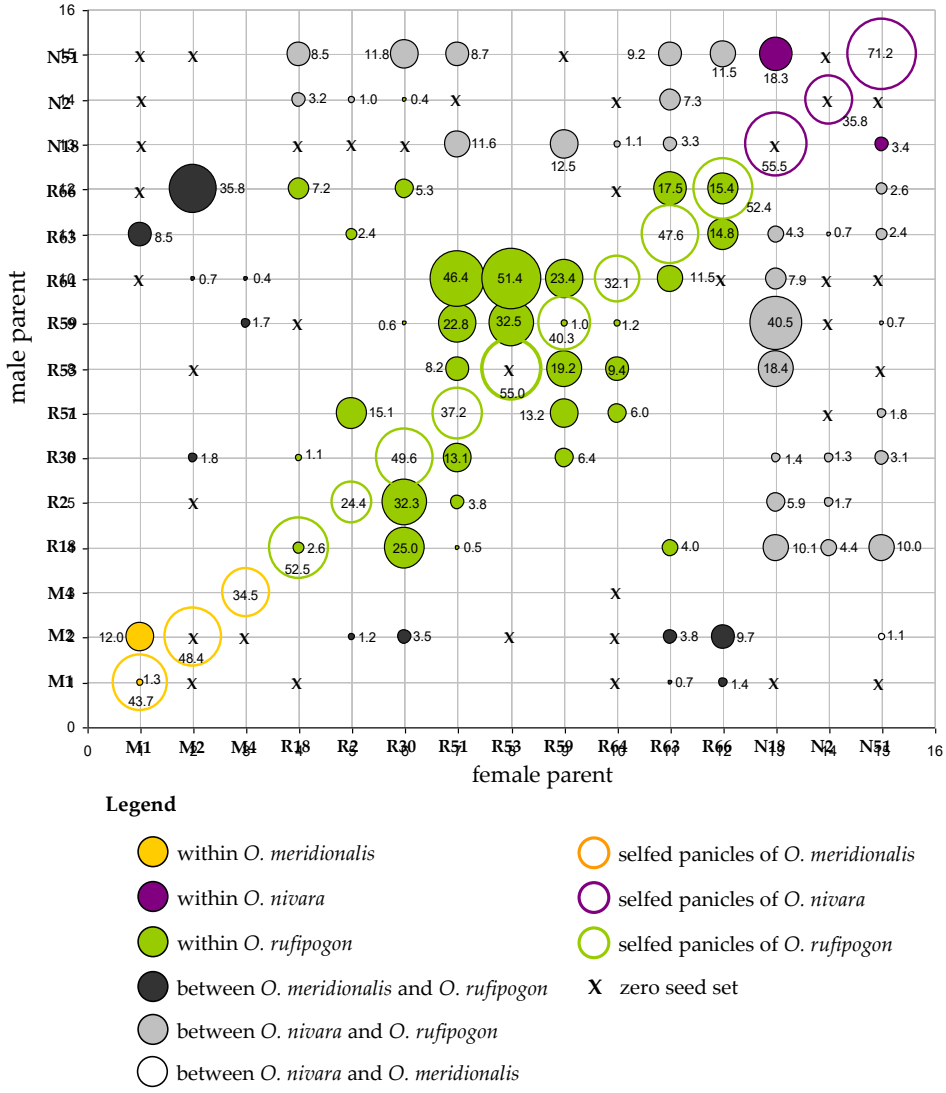


Figure 3. Mean seed set of selfed panicles and of crosses between and within the Asia Pacific species of *Oryza* series *Sativae*.

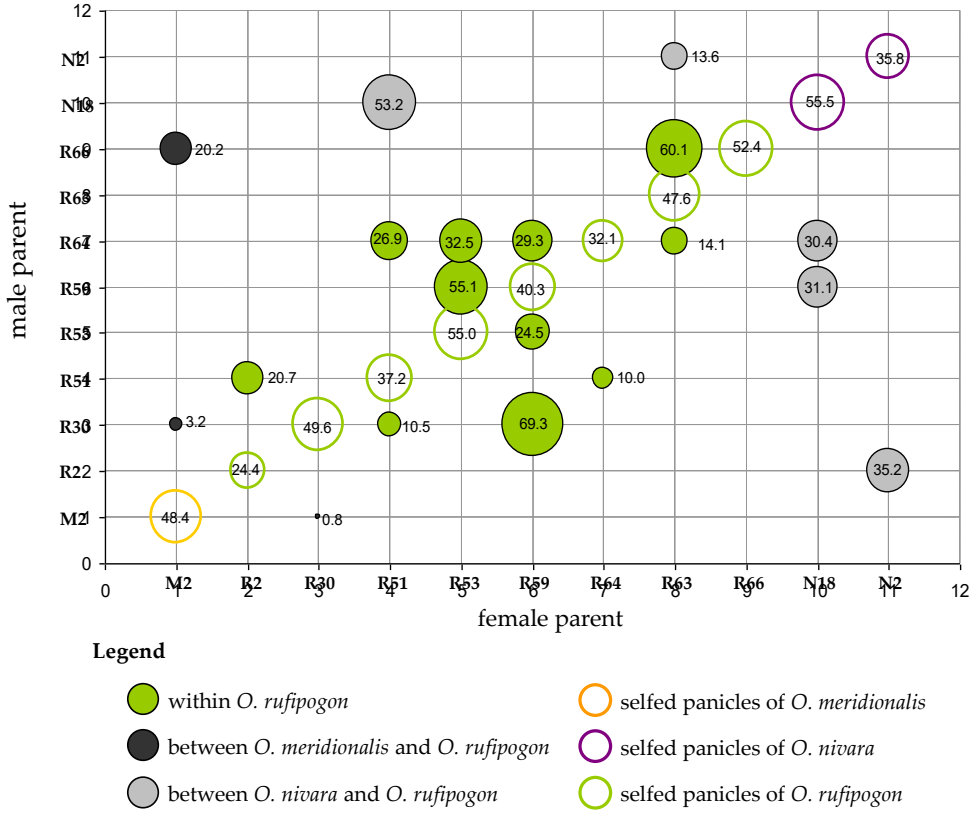


Figure 4. Mean seed set of selfed F1 hybrids from inter and intra-specific crosses of Asia Pacific *Oryza* series *Sativae* compared to the mean seed set of selfed parental accessions.

Geographic distance between parents was not significantly correlated with the different estimates of crossability in *O. rufipogon*. The spatial distance between parents of combinations that produced non-sterile hybrids ranged from 1062 – 3813 kilometers (Figure 5). Crosses with distance below and beyond the said limit exhibited zero seedling survival (Figure 5).

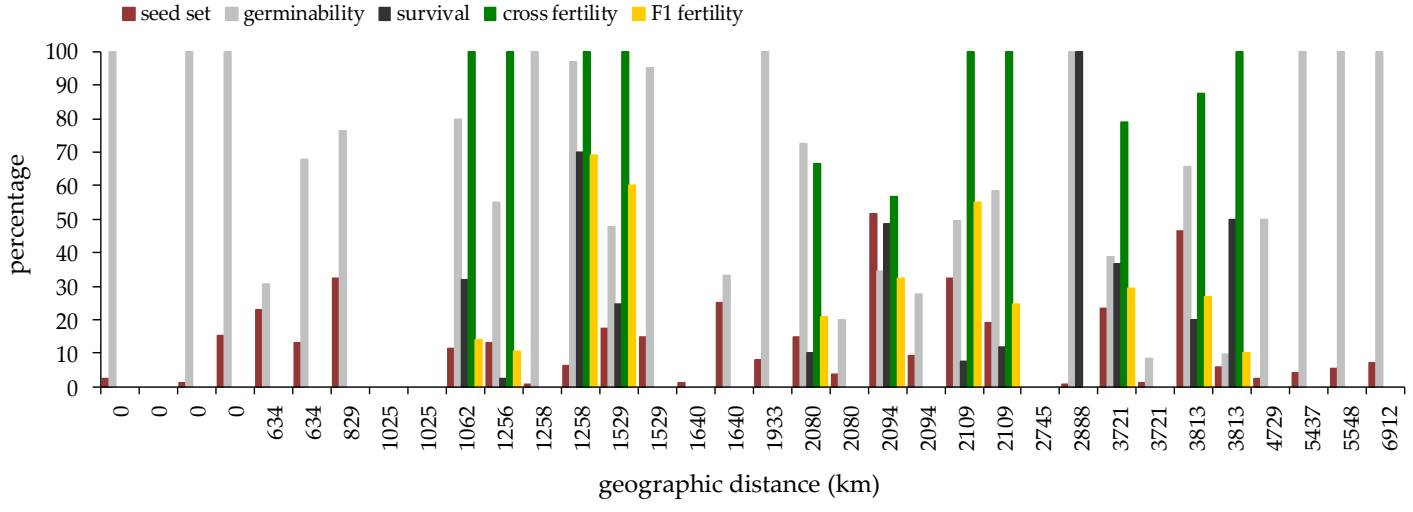


Figure 5. Crossability of intraspecific *O. rufipogon* crosses within (distance = 0) and between accessions, graphed according to increasing geographic distance between parental origins. Estimates in reciprocal crosses are presented separately.

Interspecific crossability

Crosses between *O. meridionalis* and *O. nivara* were as unproductive as their corresponding intraspecific crosses (Table 2; Figure 2). Six out of seven combinations had zero seed set (Appendix 2; Figure 3). Only N51 x M2 produced a single seed that did not survive the seedling stage (Appendix 2; Figure 3).

In contrast, 12 out of 21 crosses between *O. meridionalis* and *O. rufipogon* had seed sets that ranged from 0.37% to 35.75% (Figure 3). Yet, the 38 fertile hybrids obtained were mainly from three combinations: M2 x R66 (20.22% F1 fertility), M2 x R30 (3.21%) and R30 x M2 (0.77%) (Figure 4). Crosses with maternal *O. meridionalis* had significantly higher seed set ($p < 0.0001$), survival ($p < 0.0001$), proportion of fertile hybrids ($p < 0.05$) and F1 fertility ($p < 0.05$) than crosses with maternal *O. rufipogon* (Figure 2).

Among the 41 crosses between *O. nivara* and *O. rufipogon*, 30 combinations had seed sets that ranged from 0.4% to 40.5% (Appendix 2; Figure 3). Nonetheless, only five combinations generated 14 fertile hybrids: R51 x N18 (53.2% mean F1 fertility); N18 x R59 (31.1%); N2 x R2 (35.2%); N18 x R64 (30.4%); and R63 x N2 (13.6%) (Figure 4). All crossability estimates were comparable between crosses with maternal *O. nivara* and crosses with maternal *O. rufipogon* (Table 2 and Figure 2). The F1 hybrids of *O. nivara* and *O. rufipogon* exhibited lower survival ($p< 0.0001$) but higher fertility ($p< 0.05$) than the offspring of *O. meridionalis* and *O. rufipogon.*

In interspecific combinations involving *O. rufipogon*, all crossability measures were not significantly linearly correlated with geographic distance between parental origins and did not differ significantly between crosses with sympatric and non-sympatric parents (except for germinability being significantly higher in sympatric than in non-sympatric crosses of *O. nivara* and *O. rufipogon* at $p < 0.05$ level). However, it is worth mentioning that non-sympatric combinations exhibited higher seed set, seedling survival and F1 fertility in crosses between *O. rufipogon* and *O. meridionalis* and between *O. rufipogon* and *O. nivara*. No F1 hybrids of sympatric combinations of *O. meridionalis* and *O. rufipogon* survived the seedling stage.

F1 hybrids

SSR markers for F1 validation

RM44, RM316 and RM154 discriminate among most of the intra- and interspecific combinations of Asia Pacific *Oryza* series *Sativae*. All F1 offspring of *O. nivara* and *O. rufipogon* is being distinguished by RM152 while all intraspecific *O. rufipogon* hybrids (except the offspring of R2 x R51 and R51 x R64) are being discriminated by RM237 (Appendix 3). RM44, RM316, RM154, RM271, RM484 and RM124 differentiate the hybrids of *O. meridionalis* and *O. rufipogon*.

F1 life cycle and phenology

Hybrids resulting from crosses within *O. rufipogon* and between *O. meridionalis* and *O. rufipogon* exhibit a perennial life cycle while all F1s between *O. nivara* and *O. rufipogon* display an annual life cycle (except for the single offspring of R63 x N2).

Intraspecific *O. rufipogon* hybrids flower significantly later (mean of 271 days from seeding to first heading) than the parental accessions (mean number of days: *O. rufipogon* – 128; *O. nivara* – 104; *O. meridionalis* – 98) and other interspecific hybrids (mean number of days: F1 of *O. meridionalis* and *O. rufipogon* – 121; F1 of *O. nivara* and *O. rufipogon* – 80) at $p < 0.0001$ level. Unlike most intraspecific *O. rufipogon* combinations, R63 x R66 and R2 x R51 produced offspring that flowers relatively early (number of days from seeding to first heading ranged from 109 to 122).

Among the F1 offspring of *O. meridionalis* and *O. rufipogon*, hybrids with maternal *O. rufipogon* take longer to flower (247 days) than hybrids with maternal *O. meridionalis* (107 days). On the other hand, in crosses between *O. nivara* and *O. rufipogon*, the average number of days from seeding to first heading of hybrids with maternal *O. nivara* (85 days) is higher but does not differ significantly from that of hybrids with maternal *O. rufipogon* (70 days).

F1 morphology

The PCA results are shown in Table 3 and Figure 6. The first two principal components account for 57.6% of the total variance. Hybrids from the same cross combination generally cluster together and F1 plants from interspecific crosses seem more morphologically similar to their annual parental species (Figure 6). The first principal component (34.74%) separated the intraspecific hybrids of non-Australasian *O. rufipogon* (long-culmed, long-anthered, and late flowering) from two groups of short-culmed, short-anthered, and early flowering interspecific hybrids (*O. meridionalis x* Australasian *O. rufipogon* and *O. nivara* x *O. rufipogon*)

(Figure 6). The parental *O. rufipogon* accessions together with the F1 offspring of *O. meridionalis* and non-Australasian *O. rufipogon* and most of the intraspecific Australasian *O. rufipogon* hybrids are positioned in between these groups (Figure 6).

Table 3. Factor loadings and importance of the first three principal components of the PCA of the nine morphological characters of parental accessions and F1 hybrids. Characters with the highest factor loading are highlighted.

Character	Principal component		
	1	2	3
Culm length	0.5007	-0.1380	-0.1509
Leaf length	0.0158	-0.2633	-0.6231
Flag leaf width	0.1352	0.2561	-0.5914
Spikelet length	-0.2333	-0.4760	0.1659
Spikelet width	-0.3515	0.3598	-0.2417
Anther length	0.4837	-0.0737	0.2430
Awn length	-0.1219	-0.5839	-0.0642
Spikelet fertility	0.1904	0.3658	0.2777
Number of days from seeding to first heading	0.5174	-0.0988	-0.1143
Standard deviation	1.7683	1.4338	1.2704
Proportion of Variance	0.3474	0.2284	0.1793
Cumulative Proportion	0.3474	0.5758	0.7552

The second principal component (22.84%) separates the interspecific hybrids of *O. nivara* and *O. rufipogon* from the F1 offspring of *O. meridionalis* and *O. rufipogon* (Figure 6, Table 3). The former has shorter and broader spikelets, shorter awns and more fertile panicles than the latter.

The third principal component explains 17.93% of the total variance and does not show any distinct clustering pattern. Leaf length and flag leaf width has the highest factor loading along this principal component (Table 3). The rest of the principal components also do not display any clustering patterns.

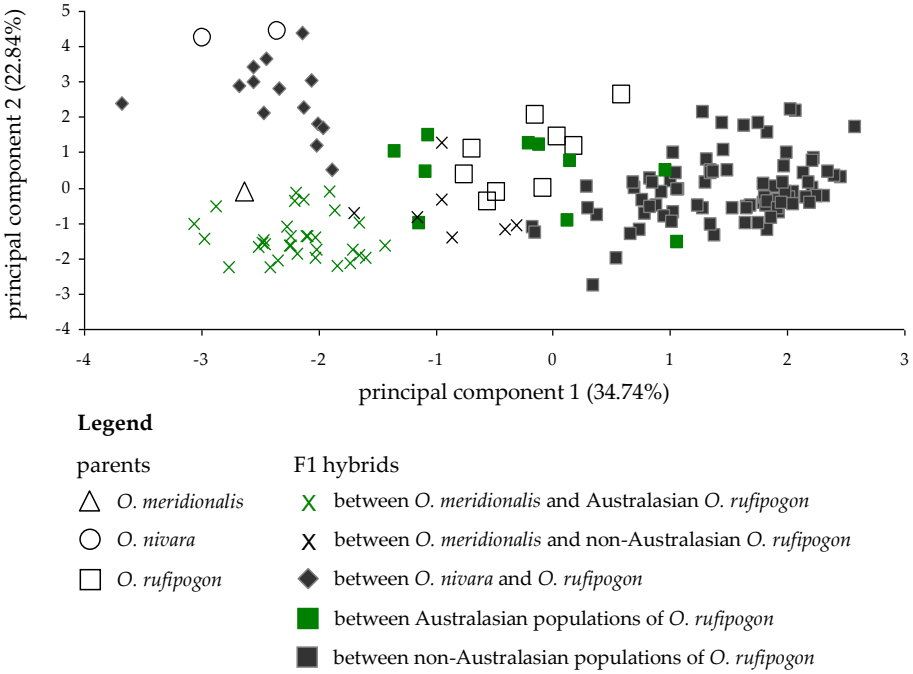


Figure 6. PCA plot showing the clustering patterns of F1 hybrids and parental accessions based on nine quantitative morphological characters.

Figure 7 shows that anther, awn and spikelet measurements of intraspecific *O. rufipogon* hybrids fall within the range of their parental species. An offspring of M2 x R30 also appears to be more similar to *O. rufipogon* based on the four traits.

Anther length and anther length to spikelet length ratio of interspecific hybrids are intermediate between *O. rufipogon* and the annual species while awn length and spikelet width tend to be more similar to the annual parent. Most hybrids between *O. nivara* and *O. rufipogon* have shorter awns and broader spikelets than hybrids between *O. meridionalis* and *O. rufipogon*. Accessions genetically identified as hybrids between *O. nivara* and *O. rufipogon* also exhibit intermediate anther length and anther length to spikelet length ratio but some populations had slightly longer awns and narrower spikelets than the F1 hybrids of *O. nivara* and *O. rufipogon* obtained in this experiment (Figure 7).

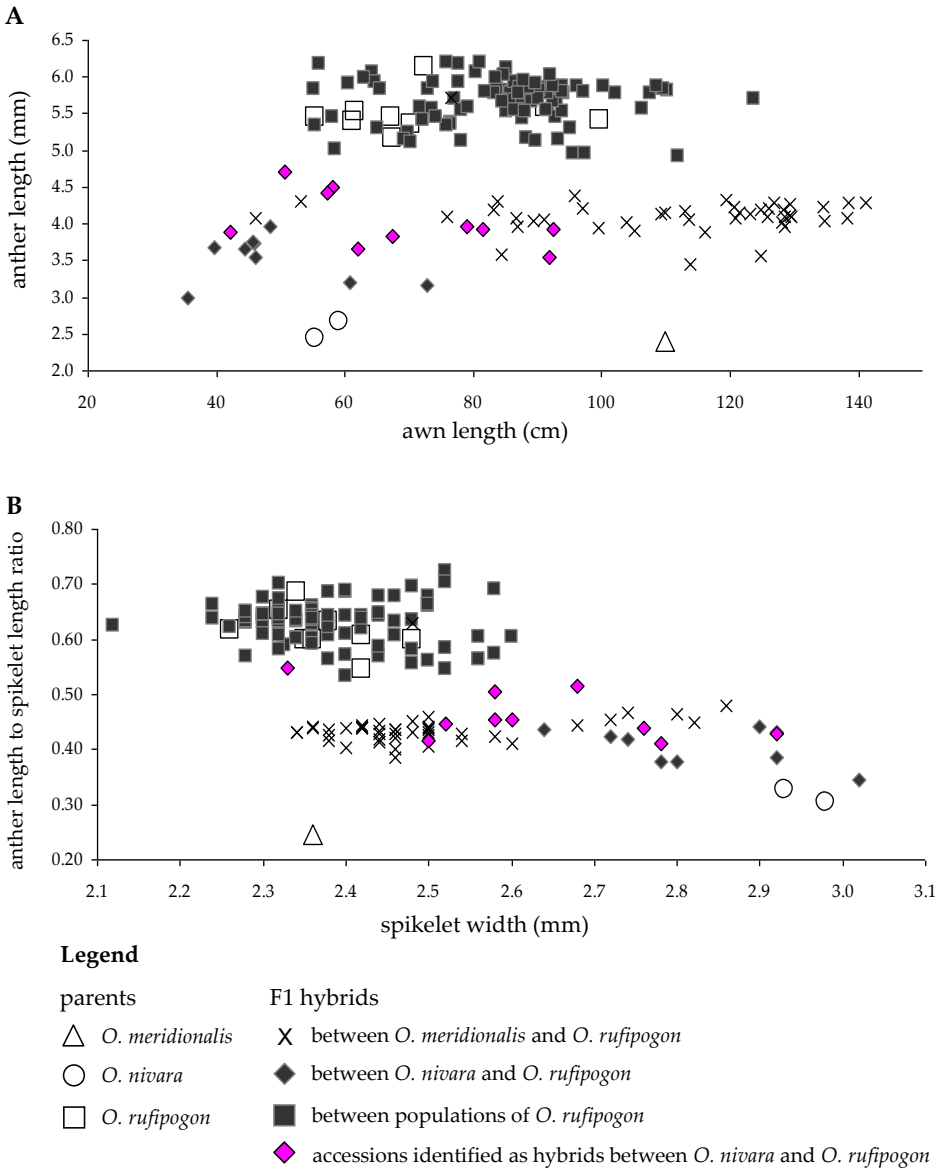


Figure 7. Scatter plot of the four quantitative characters that can readily discriminate the species and hybrids of Asia Pacific *Oryza* series *Sativae*: **A**) awn length vs. anther length; **B**) spikelet width vs. anther length to spikelet length ratio.

Discussion

Varying fertilities of self-pollinated populations within species

The reported differences between the seed productivity of *O. nivara* (high) and *O. rufipogon* (low) (Vaughan and Morishima 2003) is not evident from our results. This may be due to the limited number of accessions used in this study. Banaticla-Hilario et al. (in Chapter 2) examined the morphology of 124 populations (including the parental accessions in this experiment) and found significant differences in the spikelet fertility among the three Asia Pacific *Oryza* series *Sativae* species where *O. rufipogon* displayed the lowest and *O. nivara* the highest value.

The high variability of seed set observed among accessions of *O. nivara* (Table 1) and among plants of the same accession (Appendix 1) reflects the high genetic diversity between and within wild rice populations. The plants in each accession might be segregating both for self- and cross-compatibility. Wei et al. (2010) detected segregation at the S_5 locus (one of the major genes that control female sterility) in *O. rufipogon* accessions.

Incompatibility within and between the selfing species

The high seed production of bagged parental panicles and generally poor seed set of artificially crossed panicles (Figure 3) demonstrates the highly selfing nature of *O. meridionalis* and *O. nivara*.

The absence of seed set in most of the crosses within and between these two annual species suggests a strong pre-zygotic barrier. Concurrently, the inviability of the few generated seeds implies post-zygotic isolation. Although Naredo et al. (1998) obtained several F1 plants from crosses within and between *O. nivara* and *O. meridionalis*, the hybrids exhibited very low fertility. Substantial post-zygotic barriers (i.e., low fitness and reduced fertility of intraspecific hybrids) were also observed in other selfing species (Grundt et al. 2006; Bomblies et al. 2007; Widmer et al. 2009).

Nevertheless, Juliano et al. (2005) reported high F1 fertility (>50% spikelet fertility) in crosses between *O. meridionalis* populations from the same geographic province indicating that reproductive compatibility is higher within the regions of Australasia (e.g. Irian Jaya, Northern Territory and Queensland). This cannot be

confirmed in the current study as only three *O. meridionalis* accessions from different regions were used allowing only for intra-population and inter-regional comparisons. It would be interesting to know whether interfertility at the regional scale also exists in *O. nivara*.

The observed incompatibilities within and between *O. meridionalis* and *O. nivara* probably reflect the genetic dissimilarities among the geographically isolated parental accessions used in this experiment. Pollen-pistil interactions should be investigated to confirm and provide insights into the pre-zygotic mechanisms involved in post-pollination isolation of these selfing species.

Reproductive coherence in *O. rufipogon*

As expected, *O. rufipogon* shows greater intraspecific compatibility than the two inbreeding species. Spontaneous hybridization strongly favours outcrossing perennials that can propagate vegetatively like *O. rufipogon* (Ellstrand et al. 1996). However, it should be noted that this wild rice species is also self-compatible as exemplified by the abundant seed set of selfed panicles (Appendix 1, Table 1, Figure 3).

Intraspecific *O. rufipogon* crosses also had significantly higher seed set than all other interspecific combinations (Table 2, Figures 2 - 3) signifying the presence of a pre-zygotic isolation mechanism.

Reciprocal crosses with *O. rufipogon*

O. meridionalis and *O. rufipogon* exhibit asymmetrical compatibility where combinations with maternal *O. meridionalis* have more reproductive success than crosses with maternal *O. rufipogon* (Table 2, Figure 2). Reciprocal crosses of *O. nivara* and *O. rufipogon* display comparable crossability estimates. Similar compatibility patterns were detected by Naredo et al. (1997).

Asymmetrical reproductive isolation appears to be widespread in plants (Rieseberg and Carney 1998; Tiffin et al. 2001; Yasumoto and Yahara 2006) and may result from parental differences in style/stigma length, breeding system and fruit abortion and also from nuclear-cytoplasmic incompatibilities (Tiffin et al. 2001). The stigma length factor seems plausible in the case of the three *Oryza*

species since this trait separates *O. meridionalis* from *O. rufipogon* (specifically the Australasian populations) but does not differ between *O. nivara* and the perennial species (Banaticla-Hilario et al. in Chapter 2). Molecular studies identified *O. meridionalis* as the first lineage to diverge from the rest of the series (Zhu and Ge 2005; Kwon et al. 2006; Duan et al. 2007). It was even reported to be genetically more affiliated to the African species than the Asian series *Sativae* members (Duan et al. 2007), hence some incongruencies in nuclear-cytoplasmic interactions were expected in interspecific crosses. Conversely, the strong genetic affinity of *O. nivara* with *O. rufipogon* (Zhu and Ge 2005; Zhou et al. 2008; Zheng and Ge 2010) has probably masked the effect of differences in breeding system.

It is worth mentioning that asymmetric compatibilities have also been detected in several intraspecific crosses of *O. rufipogon*, in line with the large diversity within this outcrossing species.

Crossability with *O. rufipogon*

In agreement with Mallet (2008), the reproductive failure of crosses within species and the relative prolificness of interspecific crosses precluded the use of crossability estimates in confirming the species status of the two annual *Oryza* taxa. It is well known that *O. meridionalis* is a genetically distinct species (Xu et al. 2005; Kwon et al. 2006; Banaticla-Hilario et al. in Chapter 3) and that *O. nivara* and *O. rufipogon* are closely related taxa (Vaughan et al. 2003; Kwon et al. 2006; Zhu and Ge 2005; Duan et al. 2007; Zheng and Ge 2010). Hence, the reproductive compatibilities of *O. meridionalis* and *O. nivara* with *O. rufipogon* were discussed in the context of their genetic relationships, not their taxonomic statuses.

Theoretically, *O. nivara* and *O. meridionalis* would be reproductively isolated from *O. rufipogon* in the wild by similar pre-pollination barriers (e.g., differences in habitat, phenology and breeding system). Such pre-mating isolation mechanisms are potent yet leaky reproductive barriers between species and post-zygotic mechanisms are needed to counter residual gene flow that penetrates these pre-mating barriers (Widmer et al. 2009). In this experiment, post-pollination pre-zygotic isolation (i.e., seed set) was quite comparable in interspecific combinations of *O. meridionalis* and *O. rufipogon* and of *O. nivara* and *O. rufipogon* (Table 2). However, since symmetrically compatible crosses are likely to have more chances of successful mating than asymmetrical ones, *O. nivara* could have a reproductive advantage over *O. meridionalis* under natural conditions.

Disparities in the strength of several post-zygotic barriers were detected when studying the results of crosses with different annual parental species. As predicted, reduced F1 fertility is more pronounced in crosses between the genetically distant *O. meridionalis* and *O. rufipogon* (17.2% mean F1 fertility) when compared to those between O. *nivara* and *O. rufipogon* (36.9% mean F1 fertility). Yet, the reduced F1 fertility is counterbalanced by the greater abundance of generated established F1 plants (38 hybrids, mean of 12.7 hybrids per cross vs. 14 hybrids, 2.8 hybrids per cross in *O. nivara* and *O. rufipogon* combinations). Furthermore, unlike the predominantly annual F1 offspring of *O. nivara* and *O. rufipogon*, hybrids between *O. meridionalis* and *O. rufipogon* are sustained for longer periods by their perennial habit. Thus, in crosses involving *O. meridionalis*, high production and (to a certain degree) perenniality of hybrids are compensating the low fertility of F1 offspring while high F1 fertility in crosses involving *O. nivara* atoned for the low production and short life span of F1 hybrids.

Despite differences in the intensity of several post-zygotic mechanisms, the extent of post-mating isolation of both *O. meridionalis* and *O. nivara* against *O. rufipogon* seems comparable. This proves that the magnitude of reproductive isolation may not always correspond to genetic relationships and that gene flow barriers may remain permeable even between distantly related taxa. Mallet (2005, 2008) claimed that reproductive barriers could remain porous for millions of years after species divergence. Furthermore, Widmer et al. (2009) reported the ability of certain recently diverged species to rapidly develop strong post-zygotic barriers. It can be said that within Asia Pacific *Oryza* series *Sativae* reproductive isolation is incomplete and a certain amount of gene flow transgresses species boundaries. However, to confirm this conclusion, hybrid breakdown or its absence must be confirmed and the viability and fertility of succeeding hybrid generations should be monitored to ascertain the ultimate reproductive success of interspecific crosses.

Reproductive isolation in plants is a product of the complex synergy between numerous pre- and post-zygotic barriers (Rieseberg and Willis 2007; Widmer et al. 2009). This study was limited to several easily measurable isolating mechanisms. The contributions of gametic, habitat, temporal and other barriers in the maintenance of reproductive isolation within the series remain unaccounted for and await exploration.

Spatial crossability patterns

The lack of significant correlation between distance of parental origins and crossability suggest a weak (if any) spatial influence on the mating compatibility within and between the species of *Oryza* series *Sativae* in Asia Pacific. This complements the weak mantel correlation of genetic distance with geographic distance detected within and between the species of the series (Banaticla-Hilario et al. in Chapter 3).

In intraspecific *O. rufipogon* crosses among the selected accession, production of viable and non-sterile F1 hybrids seemed feasible only within a limited range of parental spatial distance (approximately 1100 – 3900 kilometers). Reproductive failure was expected in crosses with spatially remote and consequently genetically diversified parents. Surprisingly though, post-zygotic barriers also seemed to favour combinations with spatially proximate parents. This suggests that even geographically close *O. rufipogon* populations may be genetically differentiated and reflects the complex diversity patterns within this perennial wild rice species.

Within *O. rufipogon*, crossability estimates vary considerably in crosses between those from the same region (e.g., South Asia, continental Southeast Asia, maritime Southeast Asia and Australasia) as well as in inter-regional combinations (Appendix 1). Still, it can be discerned that crosses within and between maritime Southeast Asian and Australasian populations (especially those involving R53, R59 and R64) are generally the most successful in terms of seed set and F1 fertility (Figures 3 & 4).

Geographical patterns are more obscure in crosses between *O. meridionalis* and *O. rufipogon* as well as between *O. nivara* and *O. rufipogon*. Crossability estimates are highly variable and both intra- and inter-regional combinations produce viable and non-sterile hybrids (Appendix 2).

Crossability differences between sympatric and non-sympatric combinations are insignificant but hint at slightly stronger reproductive isolation under sympatric conditions. Banaticla- Hilario et al. (in Chapter 3) observed similar genetic patterns where *O. nivara* and *O. rufipogon* exhibited local differentiation (i.e., genetically distinct sympatric populations) amidst global similarities.

Variations in F1 hybrids

F1 life cycle and phenology also differs between interspecific combinations with the two annual parental species. Crosses involving *O. meridionalis* produced perennial hybrids while crosses involving *O. nivara* generated mostly annual hybrids.

The asymmetry observed in the reproductive compatibility of *O. meridionalis* and *O. rufipogon* extends to the phenology of their hybrids as F1 plants seem to inherit the photoperiod sensitivity or insensitivity of their maternal species. Hybrids between *O. nivara* and *O. rufipogon* flower much earlier than their parental accessions regardless of which species is the maternal or paternal parent. On the other hand, intraspecific *O. rufipogon* hybrids exhibit prolonged periods of vegetative growth and flower much later than their parents.

The number of days from seeding to flowering, culm length and anther length are greater in the majority of the *O. rufipogon* intraspecific hybrids compared to the parental accessions and other hybrids (Table 3; Figure 6), conforming to the general assumption that F1 hybrids between geographic races tend to be vegetatively superior to their parents (Rieseberg and Carney 1998). However, crosses between Australasian *O. rufipogon* populations produced hybrids that were morphologically comparable to the parental *O. rufipogon* accessions (Figure 6). The slightly different morphology displayed by the hybrids of Australasian populations is unsurprising as this geographic population group is indeed genetically distinct (Banaticla-Hilario et al. in Chapter 3) and can even be morphologically differentiated from the rest of *O. rufipogon* by the length of ligules, flag leaves, panicles, spikelets, awns and stigmas (Banaticla-Hilario et al. in Chapter 2).

In interspecific combinations with *O. meridionalis,* Australasian and non-Australasian populations of *O. rufipogon* produced fractionally different offspring. Hybrids with Australasian *O. rufipogon* as parents appeared more similar to *O. meridionalis* while hybrids with non-Australasian *O. rufipogon* as parents appeared more similar to the perennial species (Figure 6). Crosses between *O. nivara* and *O. rufipogon* resulted in hybrids that were more similar to the former, regardless of the geographic origin of the latter implying that a considerable portion of phenotype-determining genes are inherited predominantly from *O. nivara* by the F1 hybrids. The variations in F1 morphology of different interspecific crosses can most likely be attributed to maternal or cytoplasmic effects, as seen in the hybrids between the indica and japonica groups of *O. sativa* (Li et al. 1997). Certain cytoplasmic genes

that control hybrid sterility seem to influence the agronomic traits of F1 japonica hybrid rice (Wang et al. 1998).

Species coherence in terms of morphology is evident from the scatter plots of highly discriminating characters where all parental accessions and intraspecific hybrids of *O. rufipogon* form a distinct cluster separate from the other parental species and interspecific hybrids (Figure 7). Conversely, similar to the observations of Rieseberg and Carney (1998), our interspecific hybrids display some character traits that are similar to their annual parental species and some that are intermediate between the two parents (Figure 7). The poor morphological resolution of interspecific hybrids causes difficulty in distinguishing the hybrid- from the non-hybrid plants of Asia Pacific *Oryza* series *Sativae*. Therefore, the use of more reliable molecular techniques such as SSR and single nucleotide polymorphism (SNP) genotyping in evaluating these hybrids is highly recommended.

Conclusions

A clearer perception of how post-pollination barriers operate within and between the Asia Pacific species of *Oryza* series *Sativae* has been provided.

In the case of the studied taxa, post-zygotic barriers do not necessarily intensify with increasing genetic distance. Hybridization data is not a reliable basis in gauging taxonomic relationships among these wild rice species since gene flow patterns could either transcend or fall short of species limits.

To prevent unwanted hybridization, the crossability patterns recognized in this study should be considered in layout planning for genebank regeneration activities.

Since *O. nivara* and *O. rufipogon* as well as *O. meridionalis* and *O. rufipogon* seem more reproductively isolated when they are geographically proximal, caution should be exercised when re-introducing populations in the wild, especially in areas of sympatry.

References

Andersson, S. (1993). Morphometric differentiation, patterns of interfertility, and the genetic-basis of character evolution in *Crepis tectorum* (Asteraceae). *Plant Systematics and Evolution* **184,** 27-40.

Arnold, M. L. **(**1997). Natural Hybridization and Evolution. Oxford Univ. Press, Oxford.

Bomblies, K., Lempe, J., Epple, P., Warthmann, N., Lanz, C., Dangl, J. L. & Weigel, D. (2007). Autoimmune response as a mechanism for a Dobzhansky-Muller-type incompatibility syndrome in plants. *PLoS Biology* **5(9),** e236. doi:10.1371/journal.pbio.0050236

Duan, S., Lu, B., Li, Z., Tong, J., Kong, J., Yao, W., Li, S., Zhu, Y. (2007). Phylogenetic analysis of AA-genome *Oryza* species (Poaceae) based on chloroplast, mitochondrial, and nuclear DNA sequences. *Biochemical Genetics* **45,** 113-129.

Duistermaat, H. (1987). A revision of *Oryza* (Gramineae) in Malesia and Australia. *Blumea* **32,** 157-193.

Ellstrand, N. C., Whitkus, R. & Rieseberg, L. H. (1996). Distribution of spontaneous plant hybrids. *Proceedings of the National Academy of Sciences of the United States of America* **93,** 5090-5093.

Favero, A. P., Simpson, C. E., Valls, J. F. M. & Vello, N. A. (2006). Study of the evolution of cultivated peanut through crossability studies among *Arachis ipaensis*, *A. duranensis*, and *A. hypogaea*. *Crop Science* **46,** 1546-1552.

Fox, J. & Weisberg, S. (2011). *An R Companion to Applied Regression*. Sage, Thousand Oaks, California.

Fulton, T., Chunwongse, J. & Tanksley, S. (1995). Microprep protocol for extraction of DNA from tomato and other herbaceous plants. *Plant Molecular Biology Reporter* **13,** 207-209.

Grundt, H. H., Kjølner, S., Borgen, L., Rieseberg, L. H. & Brochmann, C. (2006). High biological species diversity in the arctic flora. *Proceedings of the*

National Academy of Sciences of the United States of America **103,** 972-975.

Juliano, A. B., Naredo, M. E., Lu, B.-R. & Jackson, M. T. (2005). Genetic differentiation in *Oryza meridionalis* Ng based on molecular and crossability analyses. *Genetic Resources and Crop Evolution* **52,** 435-445.

Kwon, S.-J., Lee, J. K., Hong, S.-W., Park, Y.-J., McNally, K. L. & Kim, N.-S. (2006). Genetic diversity and phylogenetic relationship in AA *Oryza* species as revealed by Rim2/Hipa CACTA transposon display. *Genes & Genetic Systems* **81,** 93-101.

Li, Z., Pinson, S., Paterson, A. H., Park, W. D. & Stansel, J. W. (1997). Genetics of hybrid sterility and hybrid breakdown in an intersubspecific rice (*Oryza sativa* L.) population. *Genetics* **145,** 1139-1148.

Mallet, J. (2005). Hybridization as an invasion of the genome. *Trends in Ecology & Evolution* **10,** 229-237.

Mallet, J. (2008). Hybridization, ecological races and the nature of species: empirical evidence for the ease of speciation. *Philosophical Transactions of the Royal Society B: Biological Sciences* **363,** 2971-2986.

Marcussen, T. & Borgen, L. (2011). Species delimitation in the Ponto-Caucasian *Viola sieheana* complex, based on evidence from allozymes, morphology, ploidy levels, and crossing experiments. *Plant Systematics and Evolution* **291,** 183-196.

Naredo, M. E. B., Juliano, A. B., Lu, B. R. & Jackson, M. T. (1997). Hybridization of AA genome rice species from Asia and Australia .1. Crosses and development of hybrids. *Genetic Resources and Crop Evolution* **44,** 17-23.

Naredo, M. E. B., Juliano, A. B., Lu, B. R. & Jackson, M. T. (1998). Taxonomic status of *Oryza glumaepatula* Steud. II. Hybridization between New World diploids and AA genome species from Asia and Australia. *Genetic Resources and Crop Evolution* **45,** 205-214.

Nezu, M., Katayama, T. C. & Kihara, H. (1960). Genetic study of the genus *Oryza*. *Seiken Ziho* **11,** 1-11.

Ng, N. Q., Hawkes, J. G., Williams, J. T. & Chang, T. T. (1981). The recognition of a new species of rice (Oryza) from Australia. *Botanical Journal of the Linnean Society* **82,** 327-330.

Oka, H. I. (1988). *Origin of cultivated rice.* . Japan Science Society Press, Tokyo.

Pellegrino, G., Bellusci, F. & Musacchio, A. (2009). Genetic integrity of sympatric hybridising plant species: the case of *Orchis italica* and *O. anthropophora*. *Plant Biology* **11,** 434-441.

R Development Core Team. (2011). R: A Language and Environment for Statistical Computing (R. F. f. S. Computing, Ed.), Vienna, Austria.

Raimondi, J., Sala, R. & Camadro, E. (2003). Crossability relationships among the wild diploid potato species *Solanum kurtzianum, S. chacoense* and *S. ruiz-lealii* from Argentina. *Euphytica* **132,** 287-295.

Rieseberg, L. H. & Carney, S. E. (1998). Plant hybridization. *New Phytologist* **140,** 599-624.

Rieseberg, L. H. & Willis, J. H. (2007). Plant Speciation. *Science* **317,** 910-914.

Roschevics, R. (1931). A contribution to the study of rice. *Trudy po Prikladnoi Botanike, Genetike i Selektsii* **27,** 3 - 133.

Stebbins, G. L. (1959). The role of hybridization in evolution. *Proceedings of the American Philosophical Society* **103,** 231-251.

Tateoka, T. (1963). Taxonomic studies of *Oryza*. III. Key to the species and their enumeration. *Botanical Magazine Tokyo* **76,** 165-173.

Tiffin, P., Olson, S. & Moyle, L. C. (2001). Asymmetrical crossing barriers in angiosperms. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **268,** 861-867.

Vaughan, D. A. & Morishima, H. (2003). Biosystematics of the genus *Oryza.* In Wayne Smith, Dilday, R. H. (eds) *Rice: Origin, History, Technology and Production*, pp. 27 - 65. John Wiley and Sons, New Jersey.

Vaughan, D. A., Morishima, H. & Kadowaki, K. (2003). Diversity in the *Oryza* genus. *Current Opinion in Plant Biology* **6,** 139-146.

Vriesendorp, B. & Bakker, F. T. (2005). Reconstructing patterns of reticulate evolution in angiosperms: what can we do? *Taxon* **54**, 593-604.

Wang, C., Tang, S. & Tang, Y. (1998). Effects of male sterile cytoplasm on yield and agronomic characters in Japonica hybrid rice, *Oryza sativa* L. *Breeding Science* **48,** 263-271.

Wei, C., Wang, L., Yang, Y., Chen, Z., Shahid, M., Li, J., Liu, X. & Lu, Y. (2010). Identification of an S_5^n allele in *Oryza rufipogon* Griff. and its effect on embryo sac fertility. *Chinese Science Bulletin* **55,** 1255-1262.

Widmer, A., Lexer, C. & Cozzolino, S. (2009). Evolution of reproductive isolation in plants. *Heredity* **102,** 31-38.

Xu, J.-H., Kurata, N., Akimoto, M., Ohtsubo, H. & Ohtsubo, E. (2005). Identification and characterization of Australian wild rice strains of *Oryza meridionalis* and *Oryza rufipogon* by SINE insertion polymorphism. *Genes & Genetic Systems* **80,** 129-134.

Yasumoto, A. & Yahara, T. (2006). Post-pollination reproductive isolation between diurnally and nocturnally flowering daylilies, *Hemerocallis fulva* and *Hemerocallis citrina*. *Journal of Plant Research* **119,** 617-623.

Zheng, X. & Ge, S. (2010). Ecological divergence in the presence of gene flow in two closely related *Oryza* species (*Oryza rufipogon* and *O. nivara*). *Molecular Ecology* **19,** 2439-2454.

Zhou, H.-F., Zheng, X.-M., Wei, R.-X., Second, G., Vaughan, D. A. & Ge, S. (2008). Contrasting population genetic structure and gene flow between *Oryza rufipogon* and *Oryza nivara*. *Theoretical and Applied Genetics* **117,** 1181-1189.

Zhu, Q. & Ge, S. (2005). Phylogenetic relationships among A-genome species of the genus *Oryza* revealed by intron sequences of four nuclear genes. *New Phytologist* **167,** 249-265.

Appendix 1. Seed set of self-pollinated plants obtained from five bagged panicles.

Accession code	Plant code	Seed set		
		Mean	Minimum	Maximum
M1	5	12.2	3.0	18.0
M1	274	26.0	0.0	57.3
M1	285	43.8	34.8	51.6
M1	503	85.7	46.8	98.2
M1	674	59.7	31.2	77.6
M2	137	35.0	0.0	49.2
M2	149	58.7	18.4	73.6
M2	391	63.1	52.4	78.3
M2	425	41.7	21.5	57.5
M2	647	43.2	30.9	53.0
M4	291	9.5	0.0	24.0
M4	544	48.5	13.2	69.4
M4	638	45.3	36.2	56.9
N2	89	57.1	51.3	72.7
N2	250	43.4	33.6	52.0
N2	326	37.5	31.1	49.3
N2	550	38.3	26.9	46.5
N2	597	2.2	0.0	9.6
N18	205	62.9	44.4	84.8
N18	386	63.7	31.1	86.4
N18	542	40.0	7.5	60.0
N51	62	61.5	51.0	65.1
N51	173	78.5	69.2	85.0
N51	297	78.5	73.8	83.1
N51	551	60.8	44.9	73.6
N51	656	76.4	60.9	85.7
R2	271	5.9	5.4	6.6
R2	302	42.2	30.7	57.6
R2	527	43.7	25.6	50.9
R2	630	5.3	1.4	13.6
R18	33	67.7	53.7	81.5
R18	265	50.1	34.0	73.9
R18	314	36.8	20.9	53.5
R18	487	47.1	22.8	60.0
R18	579	57.7	36.7	69.4
R30	36	72.2	43.8	89.2
R30	222	78.4	35.3	100.0
R30	384	8.1	0.0	16.1
R30	528	63.2	41.5	90.9
R30	637	25.8	10.5	37.7
R51	91	19.6	0.0	44.4
R51	164	18.7	6.3	27.4

Appendix 1. (Continued) Seed set of self-pollinated plants obtained from five bagged panicles.

Accession code	Plant code	Seed set		
		Mean	Minimum	Maximum
R51	387	27.8	5.2	50.0
R51	489	58.8	46.3	76.9
R51	595	61.7	46.1	74.7
R53	66	48.1	36.1	60.0
R53	264	64.9	46.2	78.3
R53	394	50.0	26.7	75.6
R53	508	53.0	47.5	60.6
R53	682	58.9	41.2	81.8
R59	7	45.2	21.1	68.0
R59	174	44.3	24.4	75.0
R59	313	20.2	14.3	24.5
R59	430	34.2	26.7	43.2
R59	610	57.4	30.8	65.8
R63	100	14.2	7.0	27.0
R63	176	87.8	79.0	97.0
R63	296	75.4	64.0	83.0
R63	471	17.6	6.0	28.0
R63	585	43.0	28.0	52.0
R64	12	29.8	22.4	39.4
R64	159	40.9	30.2	55.3
R64	338	25.1	6.4	37.5
R64	459	26.8	22.5	32.7
R64	578	37.9	25.5	49.3
R66	59	57.6	46.2	77.4
R66	206	25.0	15.4	37.9
R66	329	47.6	0.0	70.3
R66	515	66.1	54.0	75.0
R66	568	66.2	47.5	81.1

Appendix 2. Crossability within and between the species of Asia Pacific *Oryza* series *Sativae*. The crossability variables are presented as the mean ± standard deviation of the panicles in each crossing combination.

Crosses	Dist (km)	Npan	Seed set	Germinability	Survival	Cross fertility	F1 fertility
Intra-specific crosses							
O. meridionalis							
M1 x M2	1030.78	5	12.03 ± 3.78	0.00 ± 0.00	-	-	-
M2 x M1	1030.78	9	0.00 ± 0.00	-	-	-	-
M4 x M2	844.26	2	0.00 ± 0.00	-	-	-	-
M1 x M1*	0.00	5	1.29 ± 2.89	0.00	-	-	-
M2 x M2*	0.00	4	0.00 ± 0.00	-	-	-	-
O. nivara							
N51 x N18	2897.56	5	3.44 ± 3.20	100.00 ± 0.00	0.00 ± 0.00	-	-
N18 x N51	2897.56	3	18.31 ± 24.27	100.00 ± 0.00	0.00 ± 0.00	-	-
N51 x N2	2096.65	5	0.00 ± 0.00	-	-	-	-
N2 x N51	2096.65	5	0.00 ± 0.00	-	-	-	-
N2 x N2*	0.00	5	0.00 ± 0.00	-	-	-	-
N18 x N18*	0.00	5	0.00 ± 0.00	-	-	-	-
O. rufipogon							
R18 x R30	1639.61	10	1.06 ± 2.24	0.00 ± 0.00	-	-	-
R30 x R18	1639.61	10	25.04 ± 29.78	33.33 ± 51.64	0.00 ± 0.00	-	-
R18 x R59	2744.58	5	0.00 ± 0.00	-	-	-	-
R18 x R66	6911.75	8	7.24 ± 6.75	100.00 ± 0.00	0.00 ± 0.00	-	-
R2 x R51	2080.33	10	15.05 ± 11.02	72.59 ± 31.29	10.00 ± 31.62	66.67	20.65
R51 x R2	2080.33	8	3.81 ± 4.01	20.00 ± 44.72	0.00	-	-
R2 x R63	4729.22	5	2.35 ± 5.26	50.00	0.00	-	-

Appendix 2. (Continued) Crossability within and between the species of Asia Pacific *Oryza* series *Sativae*. The crossability variables are presented as the mean ± standard deviation of the panicles in each crossing combination.

Crosses	Dist (km)	Npan	Seed set	Germinability	Survival	Cross fertility	F1 fertility
R30 x R59	1258.28	10	0.63 ± 1.98	100.00	0.00	-	-
R59 x R30	1258.28	10	6.42 ± 13.58	96.97 ± 5.25	70.00 ± 26.46	100.00 ± 0.00	69.33 ± 12.24
R30 x R2	828.58	10	32.29 ± 32.65	76.56 ± 40.33	0.00 ± 0.00	-	-
R30 x R66	5548.38	4	5.25 ± 8.14	100.00 ± 0.00	0.00 ± 0.00	-	-
R51 x R64	3813.34	10	46.35 ± 13.51	65.85 ± 37.96	20.00 ± 23.82	87.50 ± 20.92	26.89 ± 11.75
R64 x R51	3813.34	10	5.95 ± 4.88	9.82 ± 19.07	50.00± 70.71	100.00	10.04
R51 x R59	634.06	10	22.84 ± 16.18	30.62 ± 39.28	0.00 ± 0.00	-	-
R59 x R51	634.06	10	13.21 ± 12.4	67.86 ± 47.25	0.00 ± 0.00	-	-
R51 x R53	1932.94	4	8.21 ± 7.94	100.00 ± 0.00	0.00 ± 0.00	-	-
R51 x R30	1255.90	10	13.10 ± 7.43	54.83 ± 45.05	2.38 ± 6.03	100.00	10.52
R51 x R18	2888.31	10	0.53 ± 1.66	100.00	0.00	-	-
R53 x R59	2109.06	10	32.50 ± 11.57	49.31 ± 13.49	7.54 ± 15.84	100.00 ± 0.00	55.06 ± 11.85
R59 x R53	2109.06	10	19.22 ± 10.21	58.24 ± 33.14	11.67 ± 17.21	100.00 ± 0.00	24.51 ± 11.32
R53 x R64	2094.12	10	51.41 ± 13.08	34.54 ± 12.47	48.82 ± 26.12	56.60 ± 33.84	32.46 ± 16.16
R64 x R53	2094.12	9	9.42 ± 13.87	27.78 ± 43.74	0.00 ± 0.00	-	-
R59 x R64	3720.61	10	23.4 ± 31.18	38.79 ± 8.51	36.85 ± 11.24	78.83 ± 31.45	29.27 ± 14.53
R64 x R59	3720.61	9	1.15 ± 1.63	8.33 ± 16.67	0.00 ± 0.00	-	-
R63 x R18	5437.49	5	4.03 ± 5.32	100.00 ± 0.00	0.00 ± 0.00	-	-
R63 x R66	1528.70	10	17.54 ± 9.76	47.88 ± 36.93	24.76 ± 19.01	100.00 ± 0.00	60.11 ± 23.16
R66 x R63	1528.70	4	14.82 ± 13.27	95.24 ± 8.25	0.00 ± 0.00	-	-
R63 x R64	1062.11	6	11.45 ± 15.66	80.00 ± 20.00	31.94 ± 18.79	100.00 ± 0.00	14.14 ± 8.43
R64 x R66	1025.20	6	0.00 ± 0.00	-	-	-	-
R66 x R64	1025.20	5	0.00 ± 0.00	-	-	-	-
R18 x R18*	0.00	5	2.64 ± 2.51	100.00 ± 0.00	0.00 ± 0.00	-	-
R53 x R53*	0.00	5	0.00 ± 0.00	-	-	-	-

Appendix 2. (Continued) Crossability within and between the species of Asia Pacific *Oryza* series *Sativae*. The crossability variables are presented as the mean ± standard deviation of the panicles in each crossing combination.

Crosses	Dist (km)	Npan	Seed set	Germinability	Survival	Cross fertility	F1 fertility
R59 x R59*	0.00	5	1.00 ± 2.24	100.00	0.00	-	-
R66 x R66*	0.00	3	15.44 ± 5.09	100.00 ± 0.00	0.00 ± 0.00	-	-
Intra-specific crosses							
O. meridionalis x O. nivara							
M2 x N51	3818.68	7	0.00 ± 0.00	-	-	-	-
M1 x N51	2954.06	5	0.00 ± 0.00	-	-	-	-
M1 x N18	5445.68	6	0.00 ± 0.00	-	-	-	-
M1 x N2	4742.50	6	0.00 ± 0.00	-	-	-	-
O. nivara x O. meridionalis							
N51 x M2	3818.68	7	1.10 ± 2.91	0.00	-	-	-
N51 x M1	2954.06	5	0.00 ± 0.00	-	-	-	-
N18 x M1	5445.68	4	0.00 ± 0.00	-	-	-	-
O. meridionalis x O. rufipogon							
M1 x R66	1528.70	10	0.00 ± 0.00	-	-	-	-
M1 x R63	0.00	10	8.51 ± 14.96	39.88 ± 46.06	-	-	-
M2 x R30	4968.12	8	1.82 ± 3.47	66.67 ± 47.14	50.00 ± 70.71	100.00	3.21
M2 x R66	844.26	12	35.75 ± 17.74	37.78 ± 28.59	56.17 ± 33.04	100.00 ± 0.00	20.22 ± 6.63
M4 x R64	846.99	8	0.37 ± 1.04	0.00	-	-	-
M2 x R2	5706.78	10	0.00 ± 0.00	-	-	-	-
M2 x R64	5.66	9	0.69 ± 2.08	100.00	0.00	-	-

Appendix 2. (Continued) Crossability within and between the species of Asia Pacific *Oryza* series *Sativae*. The crossability variables are presented as the mean ± standard deviation of the panicles in each crossing combination.

Crosses	Dist (km)	Npan	Seed set	Germinability	Survival	Cross fertility	F1 fertility
M1 x R64	1025.20	5	0.00 ± 0.00	-	-	-	-
M2 x R53	2098.59	5	0.00 ± 0.00	-	-	-	-
M4 x R59	4290.22	10	1.67 ± 5.27	33.33 ± 47.14	0.00	-	-
O. rufipogon x O. meridionalis							
R66 x M1	1528.70	10	1.40 ± 2.51	0.00 ± 0.00	-	-	-
R63 x M1	0.00	7	0.72 ± 1.34	50.00 ± 70.71	0.00	-	-
R30 x M2	4968.12	10	3.52 ± 5.04	86.67 ± 29.81	40.00 ± 54.77	100.00	0.77
R66 x M2	844.26	10	9.65 ± 5.39	24.07 ± 43.39	0.00 ± 0.00	-	-
R64 x M4	846.99	5	0.00 ± 0.00	-	-	-	-
R2 x M2	5706.78	10	1.16 ± 1.35	0.00 ± 0.00	-	-	-
R64 x M2	5.66	5	0.00 ± 0.00	-	-	-	-
R64 x M1	1025.20	4	0.00 ± 0.00	-	-	-	-
R53 x M2	2098.59	5	0.00 ± 0.00	-	-	-	-
R63 x M2	1030.78	5	3.77 ± 3.86	100.00 ± 0.00	0.00 ± 0.00	-	-
R18 x M1	5437.49	5	0.00 ± 0.00	-	-	-	-
O. nivara x O. rufipogon							
N18 x R64	6440.01	6	7.91 ± 9.70	62.96 ± 39.02	16.67 ± 28.87	100.00	30.38
N18 x R59	2753.41	10	40.46 ± 36.06	45.08 ± 32.77	10.82 ± 18.44	100.00 ± 0.00	31.13 ± 12.33
N2 x R63	4742.50	10	0.68 ± 2.17	0.00	-	-	-
N2 x R2	16.70	9	1.65 ± 4.94	100.00	50.00	100.00	35.20
N2 x R51	2096.65	3	0.00 ± 0.00	-	-	-	-
N2 x R30	845.12	9	1.32 ± 1.72	0.00 ± 0.00	-	-	-
N2 x R64	5715.12	8	0.00 ± 0.00	-	-	-	-
N51 x R18	2888.31	9	10.02 ± 7.66	14.29 ± 34.99	0.00 ± 0.00	-	-

Appendix 2. (Continued) Crossability within and between the species of Asia Pacific *Oryza* series *Sativae*. The crossability variables are presented as the mean ± standard deviation of the panicles in each crossing combination.

Crosses	Dist (km)	Npan	Seed set	Germinability	Survival	Cross fertility	F1 fertility
N51 x R30	1255.90	10	3.12 ± 6.75	100.00 ± 0.00	0.00 ± 0.00	-	-
N2 x R18	794.57	6	4.41 ± 8.50	78.57 ± 30.30	0.00 ± 0.00	-	-
N51 x R63	2954.06	6	2.38 ± 3.72	100.00 ± 0.00	0.00 ± 0.00	-	-
N51 x R66	4477.16	8	2.61 ± 4.88	100.00 ± 0.00	0.00 ± 0.00	-	-
N51 x R51	0.00	10	1.77 ± 4.19	100.00 ± 0.00	0.00 ± 0.00	-	-
N18 x R2	820.45	4	5.87 ± 6.13	100.00 ± 0.00	0.00 ± 0.00	-	-
N18 x R18	9.28	3	10.05 ± 5.58	95.24 ± 8.25	0.00 ± 0.00	-	-
N51 x R59	634.06	9	0.69 ± 2.08	100.00	0.00	-	-
N18 x R30	1648.89	5	1.38 ± 1.89	100.00 ± 0.00	0.00 ± 0.00	-	-
N18 x R63	5445.68	5	4.29 ± 9.58	50.00	0.00	-	-
N18 x R53	4806.58	10	18.39 ± 18.52	51.55 ± 31.50	0.00 ± 0.00	-	-
N2 x R59	1998.70	5	0.00 ± 0.00	-	-	-	-
N51 x R53	1932.94	5	0.00 ± 0.00	-	-	-	-
N51 x R64	3813.34	3	0.00 ± 0.00	-	-	-	-
O. rufipogon x O. nivara							
R64 x N18	6440.01	4	1.06 ± 2.13	0.00	-	-	-
R59 x N18	2753.41	10	12.54 ± 17.92	55.67 ± 18.62	0.00 ± 0.00	-	-
R63 x N2	4742.50	10	7.3 ± 8.77	100.00 ± 0.00	16.67 ± 40.82	100.00	13.63
R2 x N2	16.70	10	1.03 ± 3.27	0.00	-	-	-
R51 x N2	2096.65	5	0.00 ± 0.00	-	-	-	-
R30 x N2	845.12	10	0.37 ± 1.17	100.00	0.00	-	-
R64 x N2	5715.12	5	0.00 ± 0.00	-	-	-	-
R18 x N51	2888.31	10	8.54 ± 10.7	40.00 ± 54.77	0.00 ± 0.00	-	-
R30 x N51	1255.90	10	11.76 ± 17.09	80.00 ± 44.72	0.00 ± 0.00	-	-
R18 x N2	794.57	9	3.16 ± 5.56	33.33 ± 57.74	0.00	-	-
R63 x N51	2954.06	5	9.22 ± 8.03	100.00 ± 0.00	0.00 ± 0.00	-	-
R66 x N51	4477.16	6	11.49 ± 8.37	100.00 ± 0.00	0.00 ± 0.00	-	-

Appendix 2. (Continued) Crossability within and between the species of Asia Pacific *Oryza* series *Sativae*. The crossability variables are presented as the mean ± standard deviation of the panicles in each crossing combination.

Crosses	Dist (km)	Npan	Seed set	Germinability	Survival	Cross fertility	F1 fertility
R51 x N51	0.00	5	8.72 ± 5.53	100.00 ± 0.00	0.00 ± 0.00	-	-
R2 x N18	820.45	5	0.00 ± 0.00	-	-	-	-
R18 x N18	9.28	5	0.00 ± 0.00	-	-	-	-
R59 x N51	634.06	5	0.00 ± 0.00	-	-	-	-
R30 x N18	1648.89	4	0.00 ± 0.00	-	-	-	-
R63 x N18	5445.68	3	3.33 ± 5.77	100.00	0.00	-	-
R51 x N18	2897.56	7	11.64 ± 15.61	50.32 ± 44.63	33.33 ± 14.43	100.00 ± 0.00	53.23 ± 7.93

*Cross between different plants from the same accessions
Dist – geographic distance between the origin of parents
Npan – total number of pollinated panicles
Seed set = number of filled spikelets / number of pollinated spikelets *100
Germinability = number of germinated seeds / number of filled spikelets *100
Survival = number of established F1 plants / number of germinated seeds *100
Cross fertility = number of F1 plants with one or more filled spikelets / number of established F1 plants *100
F1 fertility – seed set of selfed F1 plants obtained from bagged panicles

Appendix 3. The 26 SSR markers used in validating the F1 hybrids of 19 different cross combinations of *O. meridionalis*, *O. nivara* and *O. rufipogon*. The total number of markers used in each cross combination is indicated at the bottom of the table.

SSR marker	Chr*	Motif	Cross combinations																
			R2 x R51	R51 x R64 R64 x R51	R51 x R30	R53 x R59 R59 x R53	R53 x R64	R59 x R30	R59 x R64	R63 x R66	R63 x R64	M2 x R30 R30 x M2	M2 x R66	N18 x R64	N18 x R59	N2 x R2	R51 x N18	R63 x N2
RM237	1	(CT)18			x	x	x	x	x	x	x	x		x				
RM431	1	(AG)16													x			
RM495	1	(CTG)7	x	x		x			x	x	x			x	x		x	
RM154	2	(GA)21	x	x	x			x	x			x	x	x		x		x
RM452	2	(GTC)9												x			x	x
OSR13	3	(GA)n		x											x		x	x
RM338	3	(CTT)6						x										
RM124	4	(TC)10										x	x					
RM161	5	(AG)20												x	x		x	
RM413	5	(AG)11		x		x			x			x		x				x
RM507	5	(AAGA)7		x			x		x	x	x	x		x				x
RM162	6	(AC)20	x		x			x				x		x			x	x
RM118	7	(GA)8					x		x			x						
RM125	7	(GCT)8				x			x				x	x	x			
RM433	8	(AG)13		x		x	x		x					x				x
RM215	9	(CT)16											x	x	x	x		x
RM316	9	(GT)8-(TG)9(TTTG)4(TG)4	x		x		x		x	x	x	x	x	x		x	x	
RM133	6	(CT)8	x															
RM284	8	(GA)8				x												

Appendix 3. (Continued) The 26 SSR markers used in validating the F1 hybrids of 19 different cross combinations of *O. meridionalis*, *O. nivara* and *O. rufipogon*. The total number of markers used in each cross combination is indicated at the bottom of the table.

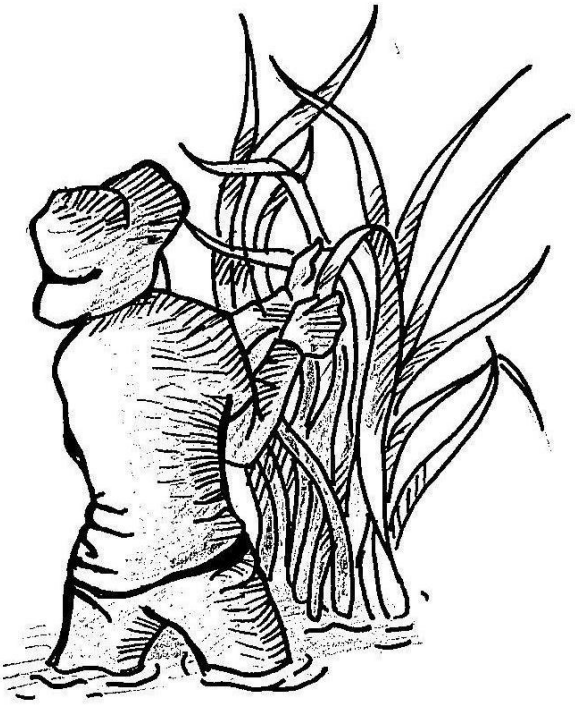
SSR marker	Chr*	Motif	Cross combinations																
			R2 x R51	R51 x R64 R64 x R51	R51 x R30	R53 x R59 R59 x R53	R53 x R64	R59 x R30	R59 x R64	R63 x R66	R63 x R64	M2 x R30 R30 x M2	M2 x R66	N18 x R64	N18 x R59	N2 x R2	R51 x N18	R63 x N2
RM408	8	(CT)13																x
RM447	8	(CTT)8							x									
RM44	8	(GA)16	x	x			x	x	x			x	x	x	x		x	
RM152	8	(GGC)10			x								x	x	x	x	x	x
RM271	10	(GA)15						x				x	x	x	x			
RM484	10	(AT)9		x				x				x	x				x	
RM277	12	(GA)11				x			x									
		total	6	8	5	7	6	7	12	4	4	11	9	15	9	4	9	10

*Chromosome number of marker

CHAPTER 5
Resolving the taxonomic ambiguity of *Oryza nivara*

Maria Celeste N. Banaticla-Hilario[1,2], Kenneth L. McNally[1], Nigel Ruaraidh Sackville Hamilton[1], Marc S.M. Sosef[2,3] and Ronald G. van den Berg[2]

[1]*T.T. Chang Genetic Resources Center, International Rice Research Institute, Los Baños, Laguna, Philippines*
[2]*Biosystematics Group, Wageningen University and Research Center, Wageningen, The Netherlands*
[3]*Naturalis Biodiversity Center, section NHN, Wageningen University, Wageningen, The Netherlands.*



Abstract

The lumping and splitting dilemma is not rare in taxonomy. A classic example of taxa that have been frequently combined and separated are *Oryza rufipogon* and *O. nivara*. New evidence from phenotyping, genetic and hybridization experiments are reviewed and considered in ascertaining the taxonomic status of *O. nivara*. The treatment of *O. nivara* and *O. rufipogon* as separate species is recommended based on the following arguments: 1) ecological distinction; 2) substantial prezygotic isolation; 3) differences in gene flow- and genetic variation patterns; 4) local-scale genetic divergence and 5) reinforced gene flow barriers under sympatric conditions. The role of ecogeography in the speciation process of *Oryza* series *Sativae* in Asia and the Pacific is discussed. The most likely scenario is that ecological speciation gave rise to the phenotypically distinct and reproductively isolated *O. nivara* and *O. rufipogon*. Within both species geographic races can be recognized.

Introduction

Asian wild rice has gone through various nomenclatural treatments (Table 1). One taxonomic turning point is the recognition of *O. nivara* Sharma & Shastry as a distinct species, restricting *O. rufipogon* Griff. to the perennial AA genome wild rice in tropical Asia and Australia (Sharma and Shastry 1965). *O. nivara* and *O. rufipogon* dwell in different habitats (the former in seasonally dry areas and the latter in swamps and other permanently submerged areas) and occur sympatrically in tropical continental Asia with the latter ranging further to maritime Southeast Asia and Australasia (Vaughan and Morishima 2003; Vaughan et al. 2008). These two taxa exhibit contrasting life cycles (annual vs. perennial), breeding systems (inbreeding vs. outcrossing) and phenologies (photoperiod insensitive vs. photoperiod sensitive) (Vaughan and Morishima 2003; Vaughan et al. 2008).

In 1981, Ng et al. recognized the annual form of Australian *O. rufipogon* as a separate species and named it *O. meridionalis* Ng. Chang (1985) eventually acknowledged three wild species of *Oryza* series *Sativae* in Asia and the Pacific : *O. nivara*, *O. rufipogon* and *O. meridionalis* (Table 1), apart from the cultigenic species *O. sativa* L. *O. meridionalis* is sympatric with *O. rufipogon* in Australasia and just like *O. nivara*, occupies seasonally dry habitats and is annual, inbreeding and photoperiod insensitive.

Table 1. Taxonomic history of Asian wild rice.

Year	Author/s	Treatment
1794	Moench	Referred to perennial wild rice in Asia, Africa and America as *O. perennis*
1851	Griffith	Published the name *O. rufipogon,* referring to wild rice found in Asia
1931	Roschevics	Considered Asian wild rice as a variation of Asian cultivated rice and named it *O. sativa* forma *spontanea*
1963	Tateoka	Defined *O. rufipogon* as all AA genome wild rice strains in tropical Asia, Australia and America
1965	Sharma and Shastry	Published the name *O. nivara* referring to the annual AA genome Asian wild rice and confined *O. rufipogon* to the perennial Asia Pacific wild rice
1972	Sharma and Shastry	Recognized the American form of *O. rufipogon* as *O. glumaepatula*
1981	de Wet	Placed *O. rufipogon* under the synonymy of *O. sativa* and published the name *Oryza sativa* subsp. *rufipogon*
1981	Ng et al.	Recognized the annual, Australian form of *O rufipogon* as a distinct species and named it *O. meridionalis*
1985	Chang	Within series *Sativae,* delimited *O. rufipogon* as the perennial Asia Pacific wild rice and recognized *O. glumaepatula, O. meridionalis* and *O. nivara* as distinct species.
1987	Duistermaat	Placed *O. nivara* under the synonymy of *O. sativa*
2003	Vaughan et al.	Placed *O. nivara* under the synonymy *O. rufipogon,* recognizing the former as the annual ecotype of the latter
2003	Vaughan and Morishima	Treated *O. nivara* as a subspecies of *O. rufipogon* suggesting the names: *O. rufipogon* subsp. *rufipogon* and *O. rufipogon* subsp. *nivara*

However, the most recent schemes relegated *O. nivara* to an ecotype (Vaughan et al. 2003) or a subspecies (Vaughan and Morishima 2003) of *O. rufipogon* (*sensu lato*). Rice scientists remain polarized on the taxonomic circumscription of *O. nivara*. Several molecular studies did not detect genetic divergence between the two taxa, thereby accepting the ecotype status of *O. nivara* (Zhu and Ge 2005; Zhou et al. 2008; Zheng and Ge 2010). Yet in GRIN taxonomy (an online system that serves as the standard reference for the taxonomic classification of crops and their wild relatives), *O. nivara* and *O. rufipogon* are considered as separate species (USDA, ARS, National Genetic Resources Program, 2012). Other taxonomic treatments synonymized *O. nivara* with *O. sativa* (Duistermaat 1987), or combined *O. nivara* and *O. rufipogon* as a subspecies of *O. sativa* (with the name *O. sativa* subsp. *rufipogon*) owing to their shared primary gene pool (De Wet 1981).

So the questions remain - Is *O. nivara* a distinct species? If *O. nivara* and *O. rufipogon* belong to the same species as suggested by Tateoka (1963), Vaughan et al. (2003) and Vaughan and Morishima (2003), how do we account for the differences in their morphology, life cycle, breeding system and phenology? *O. nivara* and *O. rufipogon* overlap genetically if we consider their entire distribution range. In contrast, *O. meridionalis* and *O. rufipogon* seem to behave as separate species despite their sympatric distribution. How do we explain the variation pattern of the three species? This chapter provides answers based on the results of the previous morphological (Chapter 2), genetic (Chapter 3) and hybridization (Chapter 4) experiments in this thesis. Sympatric population pairs of *O. nivara* and *O. rufipogon* and of *O. meridionalis* and *O. rufipogon* were analyzed to give a clearer picture of differentiation between the three species at different spatial scales.

Morphological divergence

The phenotypic separation of *O. nivara* and *O. rufipogon* has been repeatedly reported (Barbier 1989; Uga et al. 2003; Cai et al. 2004; Banaticla-Hilario et al. in Chapter 2) and interpreted as the consequence of ecological adaptation to different water regimes (Morishima et al. 1984; Vaughan et al. 2003) leading to a certain degree of genetic isolation. However, coexistence of annuals and their putative ancestral perennial forms in different ecological habitats within the same geographic area has been reported in other grass genera such as *Stipagrostis*, *Aristida, Loudetia* and *Heteropogon* (Whyte 1977). The strong influence of ecology on the morphology of Asia Pacific *Oryza* series *Sativae* species was also evident in the

phenotypic similarities between *O. meridionalis* and *O. nivara*. Thus, morphological characters may be a weak basis for separating *O. nivara* and *O. rufipogon.* Nevertheless, the three taxa have been successfully delineated phenotypically (Banaticla-Hilario et al. in Chapter 2).

Moreover, certain characters display opposing patterns of intraspecific variation for *O. nivara* and *O. rufipogon* (Tables 4 and 5 in Chapter 2). For example, anther length tends to increase in *O. nivara* and to decrease in *O. rufipogon* with increasing mean annual temperature. Similarly, culm diameter tends to increase in *O. nivara* and to decrease in *O. rufipogon* with increasing latitude. Such contrasting responses to geographic and climatic gradients indicate phenotypic differentiation beyond the scope of adaptive divergence.

The results in Chapter 2 also contradict De Wet's (1981) taxonomic approach of placing *O. nivara* and *O. rufipogon* within the circumscription of *O. sativa* and Duistermaat`s (1987) treatment of *O. nivara* as a synonym of *O. sativa.* The *O. rufipogon* cluster in the UPGMA tree (Figure 4 in Chapter 2) is clearly separated from the annual species cluster (including *O. sativa* accessions). The separation of *O. sativa* accessions from *O. nivara* populations is also clear in the tree (Figure 4 in Chapter 2) but relatively few *O. sativa* accessions were included in this study.

The treatments of De Wet (1981) and Duistermaat (1987) consider the two wild *Oryza* species as part of the variation within the cultigen *O. sativa.* Assigning wild material to a cultivated taxon is rather counterintuitive and could even lead to the interpretation of the wild taxa as escapes from cultivation. Genetic studies identified *O. rufipogon* and *O. nivara* as the wild progenitors of the japonica and indica cultivar groups of *O. sativa*, respectively (Cheng et al. 2003; Yamanaka et al. 2003; Ohtsubo et al. 2004; Xu et al. 2007). The above shows once again that the classification and nomenclature of cultivated plants should be treated separately of and differently from those of non-cultivated plants (Hetterscheid and Brandenburg 1995). The nomenclature of cultigens is regulated by the International Code of Nomenclature for Cultivated Plants (ICNCP) while the naming of plant taxa that evolve under natural conditions follows the International Code of Botanical Nomenclature (ICBN).

Molecular evidence

Studies based on isozymes (Second 1985), RAPDs (Martin et al. 1997; Ren et al. 2003), allozymes and RFLPs (Cai et al. 2004), transposon display markers (Kwon et al. 2006), tourist sequences (Iwamoto et al. 1999), MITE-AFLPs (Park et al. 2003); simple sequence repeats (SSRs) (Ren et al. 2003), and various genes sequences (Zhu and Ge 2005; Zhou et al. 2008; Zheng and Ge, 2010) did not produce a clear-cut genetic division between *O. nivara* and *O. rufipogon*. Analysis of *p-SINE1* members divided *O. rufipogon sensu lato* into four groups where one group is composed mainly of the annual strains implying that *O. nivara* is part of the intra-specific variation of *O. rufipogon* (Cheng et al. 2003). On the other hand, the existence of separate gene pools and hence the presence of distinct species was indicated by several molecular analyses using AFLPs (Aggarwal et al. 1999), SSRs (Kuroda et al. 2007), combined sequences from chloroplast, mitochondrial and nuclear DNA (Duan et al. 2007) and single nucleotide polymorphisms (SNPs) (Xu et al. 2012).

This discordance of genetic data could be reconciled when one considers the variation patterns established in Chapter 3, where the two species displayed differentiation at the local level but similarities across their distribution.

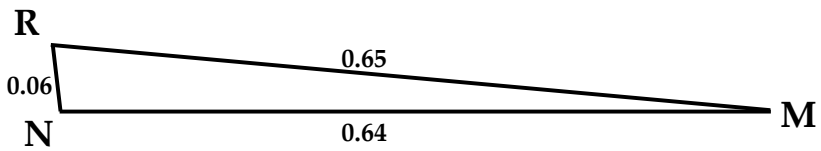


Figure 1. C.S. Chord distances between the three Asia Pacific *Oryza* series *Sativae* species (*O. rufipogon* – R; *O. nivara* – N; and *O. meridionalis* - M)

Figure 1 depicts the C.S. Chord distances of the genetic relationships among the three Asia Pacific *Oryza* series *Sativae* species based on our SSR data presented in Chapter 3 (obtained using PowerMarker 3.25 (Liu and Muse 2005)). As expected, *O. meridionalis* appears as being genetically distant from the two other species since it is the first lineage that branched out from the series (Second 1985; Wang et al. 1992; Park et al. 2003; Zhu and Ge 2005; Kwon et al. 2006; Duan et al. 2007) about 2 million years ago (mya) (Zhu and Ge 2005). Contrastingly, *O. nivara* and *O.*

rufipogon diverged much more recently at approximately 0.16 mya (Zheng and Ge 2010) as reflected in their very short genetic distance (Figure 1). More than half (59%) of the alleles shared by *O. nivara* and *O. rufipogon* reside in non-sympatric populations (Figure 14 in Chapter 3) suggesting that common ancestry primarily accounts for the extensive similarities although admixture between *O. nivara* and *O. rufipogon* may be due to (be it limited) gene flow (Zheng and Ge 2010). Bayesian clustering methods, however, detected enough differentiation to treat *O. nivara* and *O. rufipogon* as separate genetic entities even before the recognition of *O. meridionalis* as a distinct population (Figure 5 in Chapter 3).

O. rufipogon has generally less differentiated populations that are highly diverse while the populations of *O. nivara* are more differentiated and less diverse. These conflicting variation patterns have been attributed to differences in breeding systems (Kuroda et al. 2007; Lu et al. 2008; Zhou et al. 2008; Banaticla-Hilario et al. in Chapter 3). Mantel correlations (Table 4 in Chapter 3) also revealed different geographic patterns of intraspecific gene flow where isolation by distance was only observed in the South Asian populations of *O. nivara* and continental Southeast Asian populations of *O. rufipogon.* Such distinct genetic patterns clearly depict deviating evolutionary tendencies and directions for *O. nivara* and *O. rufipogon*.

The failure of molecular data to clearly separate the two taxa led some scientists to treat *O. nivara* as an ecotype of *O. rufipogon* (Ren et al. 2003; Zhu and Ge, 2005; Zheng and Ge 2010). However, we might have a different situation at hand. Recently, some authors postulated the acceptance of the "genic view" of speciation (Lexer and Widmer 2008; Wu 2011). In this view, species reproductive barriers are somewhat permeable to gene flow, and speciation can be triggered by relatively few genes that affect differential adaptation and reproductive isolation. These "speciation genes" remain diverged while neutral loci are more freely exchanged between species (Wu 2001; Feder and Nosil 2010; Rieseberg and Blackman 2010; Nosil and Schluter 2011; Southcott and Ostevik 2011). Similarly to the case of *O. nivara* and *O. rufipogon,* adaptive divergence in the face of massive allele sharing but resulting in reproductive isolation has been observed in closely related, recently diverged and geographically overlapping species of *Howea* (Savolainen et al. 2006), *Silene* (Bratteler et al. 2006), *Lupinus* (Drummond and Hamilton 2007), *Helianthus* (Yatabe et al. 2007) and *Pitcairnia* (Palma-Silva et al. 2011) where species discrimination appears to involve only a few loci/genes. Whether *O. nivara* and *O. rufipogon* are "genic" species remains to be seen as their speciation genes still await ascertainment. A good starting point is the work of Grillo et al. (2009) where quantitative trait loci (QTLs) with moderate to large effect on flowering time as

well as QTLs with small to moderate effect on floral and panicle traits associated with the mating system of *O. nivara* were identified. The same authors also implicated the role of directional selection in the fixation of majority of the QTL alleles of *O. nivara.* The loci/genes used in earlier studies that confirmed species separation could also hold clues to the identity of their supposed speciation genes. Kuroda et al. (2007) differentiated *O. nivara* and *O. rufipogon* populations in Vientiane, Laos using seven SSR markers (RM1, RM11, RM60, RM201, RM213, RM215 and RM224). Likewise, species divergence was evident in the results of Duan et al. (2007) based on sequences of the chloroplast *trn*L intron and *trn*L-*trn*F spacer, the mitochondrial *nad1* intron 2, and the nuclear internal transcribed spacer and in those of Xu et al. (2012) based on 6.5 million SNPs.

Ecological speciation

Ecological speciation occurs when reproductive isolation barriers develop as a result of divergent ecological adaptation (Rundle and Nosil 2005; Funk 2012). Environmental differences are one of the major ecological causes of divergent selection (Rundle and Nosil 2005). Hendry et al. (2007) implied that ecological speciation in plants occurs relatively easily as demonstrated in several cases where habitat disparateness appeared to have driven species and/or population divergence.

O. nivara and *O. rufipogon* differ in habitat preference, phenology as well as breeding system and these act as premating isolating mechanisms, although the level of isolation these provide still has to be assessed. It is widely accepted that ecological adaptations lead to the evolution of these reproductive barriers (Morishima et al. 1984; Vaughan et al. 2003).

Confirming the observations of Vaughan et al. (2008), Banaticla and Almazan (2010) reported that *O. nivara* and *O. rufipogon* populations in Vientiane, Laos are not spatially adjacent. The geographically nearest populations were as far as seven kilometers apart indicating mild to moderate geographical isolation. Additionally, while several intermediate forms of *O. sativa* and *O. rufipogon* and of *O. sativa* and *O. nivara* were encountered (specifically near rice fields), no intermediate forms or hybrid populations of *O. nivara* and *O. rufipogon* were found in the area (Banaticla and Almazan 2010). This implies that geographic distance effectively inhibits interbreeding between the two species at least in the mentioned locality.

Ecogeographical isolation (i.e., isolation arising from the combination of ecological barriers and geographic distance) could be an important mechanism that steers the evolution of irreversible postzygotic barriers (Lowry 2012). More surveys in other localities should be conducted to determine the actual spatial proximity of *O. nivara* and *O. rufipogon* populations and to confirm the absence/occurrence of viable hybrid populations.

Several post pollination barriers in Asia Pacific *Oryza* series *Sativae* were examined in Chapter 4. Strong prezygotic and weak postzygotic isolation was observed between *O. rufipogon* and the two annual species. Prezygotic barriers usually develop earlier and seem to predominate over postzygotic isolation in several examples of plant and animal speciation (The Marie Curie SPECIATION Network, 2012). Despite the fact that *O. nivara* and *O. meridionalis* differ markedly in their genetic similarities with *O. rufipogon*, these two annual species exhibited comparable crossabilities with the perennial one.

It was also observed that sympatric populations of *O. nivara* and *O. rufipogon* and of *O. meridionalis* and *O. rufipogon* tend to have stronger pre- and postzygotic barriers than non-sympatric populations indicating that sympatry generates a selection pressure against hybrids, favouring the evolution of stronger genetic barriers. This agrees with the genetic differentiation detected in local populations and supports the separation of *O. nivara* from *O. rufipogon*. Enhanced genetic barriers under sympatric conditions are a good indication that closely related taxa have evolved into separate species (Lowry 2012).

Geographic races

Vicariance is evident in the morphologically and genetically distinct population groups detected within the Asia Pacific species of *Oryza* series *Sativae*. These geographic races are discussed below, but are not given formal taxonomic names. More detailed morphological descriptions are provided in Chapter 2.

O. nivara can be partitioned into South Asian and Southeast Asian populations that differ in culm, leaf, panicle and spikelet characters (Table 7 in Chapter 2). Bayesian inference recognizes the genetic separation of these regional groups at higher levels of populations structure (Figures 7 and 8 in Chapter 3). *O. nivara* and *O. rufipogon* seem more differentiated in South Asia than in Southeast Asia suggesting

stronger reproductive isolation in South Asia. *O. nivara* is confined to areas with a pronounced dry season and its occurrence has not been reported in the more humid, western part of Myanmar (Vaughan et al. 2008). This area defines the regional boundary of tropical continental Asia. This geoclimatic factor probably restricts gene flow between the South and Southeast Asian populations of *O. nivara*.

Divergence of the Australasian populations of *O. rufipogon* from the rest of the species is well supported by phenotypic, SSR and hybridization data (Chapters 2, 3 and 4). This geographic race is characterized by longer leaves, ligules, panicles, spikelets, awns, anthers and stigmas (Table 8 in Chapter 2). Dissemblance of continental vs. insular populations was also observed in species of *Castilleja* (Helernum et al. 2005), *Pterocarpus* (Rivera-Ocasio et al. 2006) and *Festuca* (Fedorenko et al. 2009). Such differentiation has been attributed to geographic isolation.

The genetic uniqueness of the Nepalese populations of *O. nivara* is also worth mentioning. This geographic group does not exhibit a distinct phenotype. Yet, non-Nepalese *O. nivara* seems genetically more similar to *O. rufipogon* than to Nepalese populations of the same species (Figures 1 and 2 in Chapter 3).

Conclusion

Because *O. nivara* and *O. rufipogon* are ecologically and phenotypically distinct entities with strong pre-zygotic isolation they merit separate species status. Clear genetic separation has not yet been achieved since divergence occurred fairly recently. In spite of this, *O. nivara* and *O. rufipogon* display contrasting patterns of genetic variation, gene flow and differentiation among certain morphological traits. This is a clear indication that these two taxa have been evolving in separate directions and hence have their own evolutionary fate, which is in line with Wiley's evolutionary species concept (Wiley 1977).

Moreover, genetic divergence has been detected in local populations implying enhanced gene flow barriers on smaller spatial scales. Indeed sympatric population pairs of *O. nivara* and *O. rufipogon* and of *O. meridionalis* and *O. rufipogon* are less reproductively compatible than non-sympatric pairs. This tendency to evolve barriers to gene flow in sympatry is characteristic of good closely related species

(Wu 2001) and implies species cohesiveness in the three *Oryza* taxa, another prerequisite in several species concepts (Mishler and Donoghue 1982).

Ecology and geography are pivotal driving forces in the evolution of *Oryza* series *Sativae* in Asia and the Pacific. Ecology seems to dictate species morphology. The genetic distinctiveness of *O. meridionalis* can be explained by its geographic isolation and early divergence from the series. On the other hand, *O. nivara* and *O. rufipogon* are products of ecological speciation in a sympatric setting, where adaptation to different habitats has resulted in phenotypic differences and consequently, emergence of reproductive barriers. Allopatric divergence is apparent in the observed geographic races within the species.

References

Aggarwal, R. K., Brar, D. S., Nandi, S., Huang, N. & Khush, G. S. (1999). Phylogenetic relationships among *Oryza* species revealed by AFLP markers. *Theoretical and Applied Genetics* **98,** 1320-1328.

Banaticla-Hilario, M. C. N. & Almazan, M. S. R. (2010). Field observation of wild rice populations in Vientiane, Laos. pp. 12. International Rice Research Institute, Los Baños, Laguna, Philippines.

Barbier, P. (1989). Genetic variation and ecotypic variation in the wild rice species *Oryza rufipogon.* I. Population differentiation in life-history traits and isozymic loci. *The Japanese Journal of Genetics* **64,** 259-271.

Bratteler, M., Baltisberger, M. & Widmer, A. (2006). QTL Analysis of Intraspecific Differences between Two *Silene vulgaris* Ecotypes. *Annals of Botany* **98,** 411-419.

Cai, H. W., Wang, X. K. & Morishima, H. (2004). Comparison of population genetic structures of common wild rice (*Oryza rufipogon* Griff.), as revealed by analyses of quantitative traits, allozymes, and RFLPs. *Heredity* **92,** 409-417.

Chang, T. T. (1985). Crop history and genetic conservation: rice - a case study. *Iowa State Journal of Research* **59,** 425-455.

Cheng, C., Motohashi, R., Tsuchimoto, S., Fukuta, Y., Ohtsubo, H. & Ohtsubo, E. (2003). Polyphyletic origin of cultivated rice: based on the interspersion pattern of SINEs. *Molecular Biology and Evolution* **20,** 67-75.

De Wet, J. (1981). Species concepts and systematics of domesticated cereals. *Genetic Resources and Crop Evolution* **29,** 177-198.

Drummond, C. S. & Hamilton, M. B. (2007). Hierarchical components of genetic variation at a species boundary: population structure in two sympatric varieties of Lupinus microcarpus (Leguminosae). *Molecular Ecology* **16,** 753-769.

Duan, S., Lu, B., Li, Z., Tong, J., Kong, J., Yao, W., Li, S. & Zhu, Y. (2007). Phylogenetic analysis of AA-genome *Oryza* species (Poaceae) based on

chloroplast, mitochondrial, and nuclear DNA sequences. *Biochemical Genetics* **45,** 113-129.

Duistermaat, H. (1987). A revision of *Oryza* (Gramineae) in Malesia and Australia. *Blumea* **32,** 157-193.

Feder, J. L. & Nosil, P. (2010). The efficacy of divergence hitchhiking in generating genomic islands during ecological speciation. *Evolution* **64,** 1729-1747.

Fedorenko, O. M., Gritskikh, M. V., Malysheva, I. E. & Nikolaevskaya, T. S. (2009). Genetic diversity of insular natural populations of *Festuca pratensis* Huds.: RAPD Analysis. *Russian Journal of Genetics* **45,** 1134-1138.

Funk, D. J. (2012). Of host forms and host races: terminological issues in ecological speciation *International Journal of Ecology* **2012**. DOI 10.1155/2012/506957

Griffith, W. (1851). *Notulae ad Plantas Asiaticas*. Charles Serrao, Calcutta.

Grillo, M. A., Li, C., Fowlkes, A. M., Briggeman, T. M., Zhou, A., Schemske, D. W., Sang, T. & Juenger, T.. (2009). Genetic architecture for the adaptive origin of annual wild rice, *Oryza nivara*. *Evolution* **63,** 870-883.

Helenurm, K., West, R. & Burckhalter, S. J. (2005). Allozyme variation in the endangered insular endemic *Castilleja grisea*. *Annals of Botany* **95,** 1221-1227.

Hendry, A. P., Nosil, P. & Rieseberg, L. H. (2007). The speed of ecological speciation. *Functional Ecology* **21,** 455-464.

Hetterscheid, W. L. A. & Brandenburg, W. A. (1995). Culton versus taxon: conceptual issues in cultivated plant systematics. *Taxon* **44,** 161-175.

Iwamoto, M., Nagashima, H., Nagamine, T., Higo, H. & Higo, K. (1999). A Tourist element in the flanking region of the catalase gene CatA reveals evolutionary relationships among *Oryza* species with various genome types. *Molecular Genetics and Genomics* **262,** 493-500.

Kuroda, Y., Sato, Y.-I., Bounphanousay, C., Kono, Y. & Tanaka, K. (2007). Genetic structure of three *Oryza* AA genome species (*O. rufipogon, O. nivara* and *O. sativa*) as assessed by SSR analysis on the Vientiane Plain of Laos.

Conservation Genetics **8,** 149-158.

Kwon, S.-J., Lee, J. K., Hong, S.-W., Park, Y.-J., McNally, K. L. & Kim, N.-S. (2006). Genetic diversity and phylogenetic relationship in AA *Oryza* species as revealed by Rim2/Hipa CACTA transposon display. *Genes & Genetic Systems* **81,** 93-101.

Lexer, C. & Widmer, A. (2008). The genic view of plant speciation: recent progress and emerging questions. *Philosophical Transactions of the Royal Society B: Biological Sciences* **363,** 3023-3036.

Liu, K. & Muse, S. V. (2005). PowerMarker: Integrated analysis environment for genetic marker data. *Bioinformatics* **21,** 2128-2129.

Lowry, D. B. (2012). Ecotypes and the controversy over stages in the formation of new species. *Biological Journal of the Linnean Society* **106,** 241-257.

Lu, J., Zhang, X., Wang, H., Yuan, X., Xu, Q., Wang, Y., Yu, H., Tang, S. & Wei, X. (2008). SSR Analysis on Diversity of AA Genome *Oryza* Species in the Southeast and South Asia. *Rice Science* **15,** 289-294.

Mallet, J. (2008). Hybridization, ecological races and the nature of species: empirical evidence for the ease of speciation. *Philosophical Transactions of the Royal Society B: Biological Sciences* **363,** 2971-2986.

Martin, C., Juliano, A., Newbury, H. J., Lu, B. R., Jackson, M. T. & Ford-Lloyd, B. V. (1997). The use of RAPD markers to facilitate the identification of *Oryza* species within a germplasm collection. *Genetic Resources and Crop Evolution* **44,** 175-183.

Mishler, B. D. & Donoghue, M. J. (1982). Species concepts: a case for pluralism. *Systematic Zoology* **31(4)**, 491-503.

Moench, C. (1794). Methodus Plantas Horti Botanici et Agri Marburgensis. Marburgi cattorum.

Morishima, H., Sano, Y. & Oka, H. I. (1984). Differentiation of perennial and annual types due to habitat conditions in the wild rice *Oryza perennis*. *Plant Systematics and Evolution* **144,** 119-135.

Ng, N. Q., Hawkes, J. G., Williams, J. T. & Chang, T. T. (1981). The recognition of a new species of rice (Oryza) from Australia. *Botanical Journal of the Linnean Society* **82,** 327-330.

Nosil, P. & Schluter, D. (2011). The genes underlying the process of speciation. *Trends in Ecology & Evolution (Personal edition)* **26,** 160-167.

Ohtsubo, H., Cheng, C., Ohsawa, I., Tsuchimoto, S. & Ohtsubo, E. (2004). Rice retroposon p-SINE1 and origin of cultivated rice. *Breeding Science* **54,** 1-11.

Palma-Silva, C., Wendt, T., Pinheiro, F., Barbara, T., Fay, M. F., Cozzolino, S. & Lexer, C. (2011). Sympatric bromeliad species (*Pitcairnia* spp.) facilitate tests of mechanisms involved in species cohesion and reproductive isolation in Neotropical inselbergs. *Molecular Ecology* **20,** 3185-3201.

Park, K. C., Kim, N. H., Cho, Y. S., Kang, K. H., Lee, J. K. & Kim, N. S. (2003). Genetic variations of AA genome *Oryza* species measured by MITE-AFLP. *Theoretical and Applied Genetics* **107,** 203-209.

Ren, F., Lu, B.-R., Li, S., Huang, J. & Zhu, Y. (2003). A comparative study of genetic relationships among the AA-genome Oryza species using RAPD and SSR markers. *Theoretical and Applied Genetics* **108,** 113-120.

Rieseberg, L. H. & Blackman, B. K. (2010). Speciation genes in plants. *Annals of Botany* **106,** 439-455.

Rivera-Ocasio, E., Aide, T. M. & McMillan, W. O. (2006). The influence of spatial scale on the genetic structure of a widespread tropical wetland tree *Pterocarpus officinalis* (Fabaceae). *Conservation Genetics* **7,** 251-266.

Roschevics, R. (1931). A contribution to the study of rice. *Trudy po Prikladnoj Botanike Genetike i Selekcii.* **27,** 3-133.

Rundle, H. D. & Nosil, P. (2005). Ecological speciation. *Ecology Letters* **8,** 336-352.

Savolainen, V., Anstett, M.-C., Lexer, C., Hutton, I., Clarkson, J. J., Norup, M. V., Powell, M.P., Springate, D., Salamin, N. & Baker, W. J. (2006). Sympatric speciation in palms on an oceanic island. *Nature* **441,** 210-213.

Second, G. (1985). Evolutionary relationships in the Sativa group of *Oryza* based on isozyme data. *Genetics Selection Evolution* **17,** 89-114.

Sharma, S. D. & Shastry, S. V. S. (1965). Taxonomic studies in the genus *Oryza* L. III. *O. rufipogon* Griff. sensu stricto and *O. nivara* Sharma et Shastry *nom. nov. Indian Journal of Genetics and Plant Breeding* **25,** 157-167.

Sharma, S. D. & Shastry, S. V. S. (1972). Evolution in genus *Oryza,* in *Advancing Frontiers in Cytogenetics – Proceedings of the National Seminar,* 1972 New Delhi, India, p. 5-20.

Southcott, L. & Ostevik, K. L. (2011). Bromeliad population genetics reveals species cohesion against the odds. *Molecular Ecology* **20,** 3081-3083.

Tateoka, T. (1963). Taxonomic studies of *Oryza*. III. Key to the species and their enumeration. *Botanical Magazine Tokyo* **76,** 165-173.

The Marie Curie SPECIATION Network (2012). What do we need to know about speciation? *Trends in Ecology and Evolution* **27,** 27-39.

Uga, Y., Fukuta, Y., Ohsawa, R. &Fujimura, T. (2003). Variations of floral traits in Asian cultivated rice (*Oryza sativa* L.) and its wild relatives (*O. rufipogon* Griff.). *Breeding Science* **53,** 345-352.

USDA, ARS, National Genetic Resources Program (2012). Germplasm Resources Information Network - (GRIN) [Online Database]. National Germplasm Resources Laboratory, Beltsville, Maryland. URL: http://www.ars-grin.gov/cgi-bin/npgs/html/index.pl (April 12, 2012)

Vaughan, D. A., Lu, B.-R. & Tomooka, N. (2008). The evolving story of rice evolution. *Plant Science* **174,** 394-408.

Vaughan, D. A. & Morishima, H. (2003). Biosystematics of the genus *Oryza.* In Wayne Smith, Dilday, R. H. (eds) *Rice: Origin, History, Technology and Production*, pp. 27 - 65. John Wiley and Sons, New Jersey.

Vaughan, D. A., Morishima, H. & Kadowaki, K. (2003). Diversity in the *Oryza* genus. *Current Opinion in Plant Biology* **6,** 139-146.

Wang, Z. Y., Second, G. & Tanksley, S. D. (1992). Polymorphism and phylogenetic relationships among species in the genus *Oryza* as determined by analysis of nuclear RFLPs. *Theoretical and Applied Genetics* **83,** 565-581.

Whyte, R. O. (1977). The botanical neolithic revolution. *Human Ecology* **5,** 209-222.

Wiley, E. O. (1977). The evolutionary species concept reconsidered. *Systematic Biology* **27**, 17-26.

Wu, C.-I. (2001). The genic view of the process of speciation. *Journal of Evolutionary Biology* **14,** 851-865.

Xu, J.-H., Cheng, C., Tsuchimoto, S., Ohtsubo, H. & Ohtsubo, E. (2007). Phylogenetic analysis of *Oryza rufipogon* strains and their relations to *Oryza sativa* strains by insertion polymorphism of rice SINEs. *Genes & Genetic Systems* **82,** 217-229.

Xu, X., Liu, X., Ge, S., Jensen, J. D., Hu, F., Li, X., Dong Y., Gutenkunst R. N., Fang, L., Huang, L. Li, J., He, W., Zhang, G., Zheng, X., Zhang, F., Li, Y., Yu, C., Kristiansen, K., Zhang, X., Wang, J., Wright, M., McCouch, S., Nielsen, R., Wang J. & Wang, W. (2012). Resequencing 50 accessions of cultivated and wild rice yields markers for identifying agronomically important genes. *Nature Biotechnology* **30,** 105–111.

Yamanaka, S., Nakamura, I., Nakai, H. & Sato, Y.-I. (2003). Dual origin of the cultivated rice based on molecular markers of newly collected annual and perennial strains of wild rice species, *Oryza nivara* and *O. rufipogon*. *Genetic Resources and Crop Evolution* **50,** 529-538.

Yatabe, Y., Kane, N. C., Scotti-Saintagne, C. & Rieseberg, L. H. (2007). Rampant gene exchange across a strong reproductive barrier between the annual sunflowers, *Helianthus annuus* and *H. petiolaris*. *Genetics* **175,** 1883-1893.

Zheng, X. & Ge, S. (2010). Ecological divergence in the presence of gene flow in two closely related *Oryza* species (*Oryza rufipogon* and *O. nivara*). *Molecular Ecology* **19,** 2439-2454.

Zhou, H.-F., Zheng, X.-M., Wei, R.-X., Second, G., Vaughan, D. A. & Ge, S. (2008). Contrasting population genetic structure and gene flow between *Oryza*

rufipogon and *Oryza nivara*. *Theoretical and Applied Genetics* **117,** 1181-1189.

Zhu, Q. & Ge, S. (2005). Phylogenetic relationships among A-genome species of the genus *Oryza* revealed by intron sequences of four nuclear genes. *New Phytologist* **167,** 249-265.

CHAPTER 6
General Discussion

This thesis explored the diversity of *Oryza* series *Sativae* in Asia and the Pacific and substantiated the strong impact of ecogeography on the morphological and genetic aspects of inter- and intraspecific variations and on the speciation processes. The use of locally sympatric populations (sampled throughout the species' distribution range) of different species provided a window to look at species diversity and differentiation along geographic and climatic gradients.

Supporting evidence and new contributions to the existing knowledge of wild rice diversity

Some of the findings in this study coincided with the results of previous analyses. For example, the phenotypic separation of *O. nivara* and *O. rufipogon* observed in Chapter 2 had also been detected by Barbier (1989), Uga et al. (2003) and Cai et al. (2004). The earlier reported contrasting intraspecific diversity patterns (Oka, 1988; Lu et al. 2008; Zhou et al. 2008.) and genetic overlapping (Zhu and Ge 2005; Zhou et al. 2008; Zheng and Ge 2010) of these two species were confirmed in Chapter 3. Hybridization experiments (Chapter 4) revealed symmetric compatibility between *O. nivara* and *O. rufipogon* and asymmetric compatibility between *O. meridionalis* and *O. rufipogon*, agreeing with the results of Naredo et al. (1997).

Nonetheless, previously unrecognized variation patterns within and between *O. nivara* and *O. rufipogon* were established:

1) morphological and genetic species differentiation seemed greater in South Asia suggesting stronger gene flow barriers in that region compared to Southeast Asia;

2) sympatric populations of *O. nivara* and *O. rufipogon* are genetically distinct, indicating species differentiation at a local scale;

3) reproductive isolation between species seemed to intensify in sympatry;

4) isolation by distance was exhibited by *O. nivara* mainly in South Asia and by *O. rufipogon* mainly in continental Southeast Asia, implying different geographic patterns of intraspecific gene flow;

5) certain morphological characters (such as seedling height, culm number

and diameter, leaf length and width, and anther length) displayed opposing responses to geographic and climatic gradients for the two species;

6) altitude showed weak negative correlation with genetic diversity within populations and moderate positive correlation with species differentiation;

7) *O. nivara* seemed morphologically and genetically divided into South Asian and Southeast Asian population groups;

8) Australasian populations of *O. rufipogon* appeared to be phenotypically and genetically distinct from the rest of the species.

9) Nepalese populations of *O. nivara* are genetically diverged from the non-Nepalese populations of the same species as well as from *O. rufipogon.*

The synthesis of results from Chapters 2 to 4 ascertained that *O. nivara* and *O. rufipogon* are ecologically, morphologically and (to a limited extent) genetically distinct entities, reinforcing the treatment of these wild rice taxa as separate species.

Implications for wild rice conservation

The need for wild rice conservation and the potential benefits that it can reap from diversity studies were discussed briefly in Chapter 1. The current chapter provides detailed recommendations for effective conservation practices based on the established variation patterns.

In situ strategies

The speciation dynamics of *Oryza* series *Sativae* in Asia and the Pacific is ecogeographically driven hence habitat preservation or *in situ* conservation is imperative in ensuring the continuity and species integrity of these wild rices. If possible, national parks and protected areas should extend to local areas of sympatry. Carefulness should be applied in re-introducing populations in the wild so that locally existing gene flow barriers are not disrupted.

Oryza species dwell in generally open and accessible areas that are susceptible to man-made disturbances and habitat fragmentation can mitigate reproductive barriers. Repeated bouts of interbreeding could endanger the genetic integrity and cohesiveness of both *O. nivara* and *O. rufipogon*. Natural populations should be further examined to confirm whether ecological imbalances brought by human activities are promoting interspecific gene flow in Southeast Asia.

Natural populations of *O. meridionalis,* Australasian *O. rufipogon*, and Nepalese *O. nivara* should be located and prioritized for *in situ* conservation. The genetic uniqueness and relatively low diversity of these taxa makes them vulnerable to inbreeding depression and consequent genetic erosion.

Genebank management

As mentioned in Chapter 2, *O. nivara* and *O. rufipogon* accessions represent more than half of the wild *Oryza* collection of the International Rice Genebank in the International Rice Research Institute (IRG-IRRI) and several problems have been encountered in classifying them:

1) Misidentification is common because the species in Series *Sativae* have similar characteristics. Botanical descriptions in the literature are usually incomplete and species delineation in terms of morphology is unclear due to contradicting taxonomic views.

2) Although morphologically distinct populations of *O. nivara* and *O. rufipogon* are in the collection, a large number of accessions appear to be intermediate forms of the two species.

As stated repeatedly in this thesis, the three Asia Pacific *Oryza* series *Sativae* species can be delineated phenotypically. The species descriptions developed in Chapter 2 can be used in determining the correct taxonomic identity of misidentified or unclassified *Oryza* accessions.

The intermediate populations are probably of hybrid origin and require further examination. There are numerous accounts of hybridization between cultivated and wild rice (Oka, 1988; Vaughan et al. 2008) but hybrid populations between *O. nivara* and *O. rufipogon* are rarely reported. Therefore, it is highly possible that some of these intermediate accessions are actually hybrids between *O. sativa* and *O. rufipogon* or *O. nivara* rather than between the two wild species. Molecular techniques such as simple sequence repeat (SSR) and single nucleotide

polymorphism (SNP) genotyping can determine the genetic identity of hybrid populations. However, SSR data from Chapter 3 revealed similarities between the aromatic and japonica varietal groups of *O. sativa* and *O. rufipogon* as well as between the indica and aus groups and *O. nivara*. Hence, caution should be exercised when drawing conclusions from SSR results since genetic similarities could conceal gene flow between cultivated and wild rice. Also, to ensure genetic purity, germplasm collections should target wild populations located away from rice fields. Moreover, farmers and gene banks should be vigilant in screening their materials for planting, distribution and other purposes to prevent the unwanted spread of these hybrid plants.

The suggested treatment of *O. nivara* as a distinct species already concurs with the standard taxonomic classification for gene banks, provided by GRIN Taxonomy (USDA, ARS, National Genetic Resources Program, 2012). However, its implications on *ex situ* conservation go beyond the correct labelling of accessions. The identified genetically distinct population groups should be evaluated further and considered in identifying gaps in the collection. Regional representation (including South Asia, Southeast Asia and Australasia) might suffice for *O. rufipogon* while *O. nivara* might require a more specific geographic representation (covering Nepal, non-Nepalese South Asia, Cambodia, and non-Cambodian Southeast Asia). For *O. rufipogon*, it might be more efficient to collect germplasm from numerous individuals of a few selected populations and plant as many as possible when regenerating gene bank accessions. On the other hand, fewer samples from many geographic populations and less plants for accession regeneration are required to capture the diversity of *O. nivara*. In growing plants for regeneration purposes, *O. rufipogon* should have wider spaces between accessions to prevent hybridization and maintain the genetic integrity of populations. Such wide spacing might not be necessary for accessions of the inbreeding species *O. nivara* and *O. meridionalis*.

Implications for rice domestication

The results in this thesis lead to the conclusion that *O. nivara* and *O. rufipogon* should be considered as separate species. How will this taxonomic treatment affect the preconceived models of Asian rice domestication? As mentioned in Chapter 1, there are two schools of thought on the origin of the japonica and indica cultivar groups of *O. sativa*. These were presented by Sang and Ge (2007) as the snow-ball

model postulating a single ancestor for the cultivated taxa and the combination model proposing independent domestications for indica and japonica. How does *O. nivara* fit in the modern schemes of the mentioned hypotheses?

Apparently, many scientists considered *O. nivara* as a variation within *O. rufipogon*, consequently dismissing the potential contribution of this annual taxon in the domestication process of *O. sativa*. *O. nivara* was poorly represented and superficially discussed (if discussed at all) in several phylogenetic studies (Bautista et al. 2001; Londo et al. 2006; Rakshit et al. 2007; Molina et al. 2011) and reviews (Kovach et al. 2007; Sweeney and McCouch 2007).

Howbeit, analyses of *p-SINE1* members (Cheng et al. 2003; Yamanaka et al. 2003; Ohstubo et al. 2004; Xu et al. 2007), SNPs (Xu et al. 2012), and the MITE system mPing (Hu et al. 2006) detected genetic similarities between *O. nivara* and the indica group as well as between *O. rufipogon* and japonica, supporting the dual origin hypothesis. A similar pattern was obtained from SSR data in this thesis (Chapter 3). Figure 1A depicts the independent domestication of indica from *O. nivara* and of japonica from *O. rufipogon* followed by hybridization/introgression events that enabled the two cultivar groups to share the same domestication alleles and the two wild species to exhibit extensive genetic overlapping. Phylogeographic evidence identified China as the ultimate geographical origin of japonica and tropical Asia (India, Myanmar and Thailand) of indica (Londo et al. 2006).

It should be noted that phylogenetic trees obtained from SNPs (Xu et al. 2012) and *p-SINE1* members (Cheng et al. 2003) nested *O. nivara* within the indica clade, strongly suggesting shared ancestry. However, recent molecular clock estimates indicate that the divergence time of *O. nivara* and *O. rufipogon* was much earlier (0.16 million years ago according to Zheng and Ge 2010) than that of indica and japonica (~8200 years ago according to Molina et al. 2011) implying a recent common ancestor for the two cultivated groups.

Indica and japonica seemed to share genes that are responsible for key domestication traits such as *sh4* for panicle shattering (Li et al. 2006), *rc* for white pericarp color (Sweeney et al. 2007) and *prog1* for erect habit (Tan et al. 2008). Based on this, Vaughan et al. (2008) inferred that the current variation in *O. sativa* can be traced to a single domestication bottleneck followed by cycles of introgression with sympatric wild and cultivated populations and farmers' selection. Studies based on SSRs (Gao and Innan 2008), SNPs (Molina et al. 2011) and other gene sequences (Duan et al. 2007) also supported the monophyletic

origin of cultivated rice. Ikehashi et al. (2009) emphasized the possible role of anthropogenic dispersal in rice domestication and proposed that japonica cultivars from China were probably introduced to India where they introgressed with local wild rice populations and developed into a genetically differentiated indica cultivar. Molina et al. (2011) adapted this perspective and suggested that the transported japonica populations hybridized with local "proto-indica" cultivars and consequently evolved into indica. Molina et al. (2011) also provided divergence time estimates that concurred to archeological evidences of major domestication sites that emerged in China around 8000-9000 years ago (corresponding to the divergence time of japonica and *O. rufipogon*) and in India nearly 4000 years ago (corresponding to the divergence time of indica and japonica about 3900 years ago).

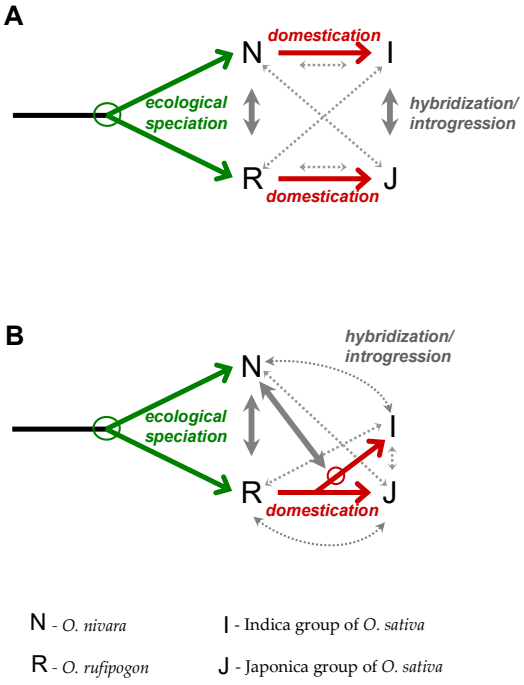


Figure 1. Simplified schemes showing the possible role of *O. nivara* in Asian rice domestication. **A**, The dual origin model suggests that indica and japonica were independently domesticated from *O. nivara* and *O. rufipogon*, respectively. **B**, The single origin model signifies that the two cultivar groups were both derived from *O. rufipogon*. Gray solid lines depict hybridization/introgression events with significant impact on the current genetic variation of Asian *Oryza* species while gray dotted lines represent less consequential gene flow. The putative ancestor of *O. nivara* and *O. rufipogon* is encircled in green while the putative recent ancestor of indica is encircled in red.

The remarkable genetic similarities between *O. nivara* and the indica cultivar group clearly validates their close interaction be it shared ancestry or gene flow. Therefore, under the single domestication setting, it is highly likely that the migrated japonica cultivars hybridized/introgressed with local *O. nivara* populations in India and eventually gave rise to the early forms of indica (as depicted in Figure 1B). Both models recognize the possibility of modest gene flow within and between cultivated and wild populations.

It is worth mentioning that population structure analyses based on SSRs (Chapter 3 in this thesis) and SNPs (Xu et al. 2012) clustered indica and aus (a minor cultivar group of *O. sativa*) to separate populations of *O. nivara.* However, these results are inconclusive since very few accessions of aus were analyzed in both studies. Still, the results do suggest that *O. nivara* has been an active player in the rice domestication process. Despite all the work done to date, the origin of cultivated rice remains debatable and continues to be a subject of scientific inquest.

References

Barbier, P. (1989). Genetic variation and ecotypic variation in the wild rice species *Oryza rufipogon.* I. Population differentiation in life-history traits and isozymic loci. *The Japanese Journal of Genetics* **64,** 259-271.

Bautista, N. S., Solis, R., Kamijima, O. & Ishii, T. (2001). RAPD, RFLP and SSLP analyses of phylogenetic relationships between cultivated and wild species of rice. *Genes & Genetic Systems* **76,** 71-79.

Cai, H. W., Wang, X. K. & Morishima, H. (2004). Comparison of population genetic structures of common wild rice (*Oryza rufipogon* Griff.), as revealed by analyses of quantitative traits, allozymes, and RFLPs. *Heredity* **92,** 409-417.

Cheng, C., Motohashi, R., Tsuchimoto, S., Fukuta, Y., Ohtsubo, H. & Ohtsubo, E. (2003). Polyphyletic origin of cultivated rice: based on the interspersion pattern of SINEs. *Molecular Biology and Evolution* **20,** 67-75.

Duan, S., Lu, B., Li, Z., Tong, J., Kong, J., Yao, W., Li, S. & Zhu, Y. (2007). Phylogenetic analysis of AA-genome *Oryza* species (Poaceae) based on chloroplast, mitochondrial, and nuclear DNA sequences. *Biochemical Genetics* **45,** 113-129.

Gao, L. & Innan, H. (2008). Nonindependent domestication of the two rice subspecies, *Oryza sativa* ssp. *indica* and ssp. *japonica*, demonstrated by multilocus microsatellites. *Genetics* **179,** 965-976.

Hu, H., Mu, J., Zhang, H.-J., Tao, Y.-Z. & Han, B. (2006). Differentiation of a miniature inverted transposable element (MITE) system in Asian rice cultivars and its inference for a diphyletic origin of two subspecies of Asian cultivated rice. *Journal of Integrative Plant Biology* **48,** 260-267.

Ikehashi, H. (2009). Why are there indica type and japonica type in rice? - History of the studies and a view for origin of two types. *Rice Science* **16,** 1-13.

Kovach, M. J., Sweeney, M. T. & McCouch, S. R. (2007). New insights into the history of rice domestication. *Trends in Genetics* **23,** 578-587.

Li, C., Zhou, A. & Sang, T. (2006). Rice domestication by reducing shattering.

Science **311,** 1936-1939.

Londo, J. P., Chiang, Y. C., Hung, K. H., Chiang, T. Y. & Schaal, B. A. (2006). Phylogeography of Asian wild rice, *Oryza rufipogon*, reveals multiple independent domestications of cultivated rice, *Oryza sativa*. *Proceedings of the National Academy of Sciences of the United States of America* **103,** 9578-9583.

Lu, J., Zhang, X., Wang, H., Yuan, X., Xu, Q., Wang, Y., Yu, H., Tang, S. & Wei, X. (2008). SSR Analysis on Diversity of AA Genome *Oryza* Species in the Southeast and South Asia. *Rice Science* **15,** 289-294.

Molina, J., Sikora, M., Garud, N., Flowers, J. M., Rubinstein, S., Reynolds, A., Huang, P., Jackson, S., Schaal, B. A., Bustamante, C. D., Boyko, A. R. & Purugganan, M. D. (2011). Molecular evidence for a single evolutionary origin of domesticated rice. *Proceedings of the National Academy of Sciences of the United States of America* **108,** 8351-8356.

Naredo, M. E. B., Juliano, A. B., Lu, B. R. & Jackson, M. T. (1997). Hybridization of AA genome rice species from Asia and Australia .1. Crosses and development of hybrids. *Genetic Resources and Crop Evolution* **44,** 17-23.

Ohtsubo, H., Cheng, C., Ohsawa, I., Tsuchimoto, S. & Ohtsubo, E. (2004). Rice retroposon p-SINE1 and origin of cultivated rice. *Breeding Science* **54,** 1-11.

Oka, H. I. (1988). *Origin of cultivated rice.* Japan Science Society Press, Tokyo.

Rakshit, S., Rakshit, A., Matsumura, H., Takahashi, Y., Hasegawa, Y., Ito, A., Ishii, T., Miyashita, N. & Terauchi, R. (2007). Large-scale DNA polymorphism study of *Oryza sativa* and *O. rufipogon* reveals the origin and divergence of Asian rice. *Theoretical and Applied Genetics* **114,** 731-743.

Sang, T. & Ge, S. (2007). The puzzle of rice domestication. *Journal of Integrative Plant Biology* **49,** 760-768.

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

Sweeney, M. T., Thomson, M. J., Cho, Y. G., Park, Y. J., Williamson, S. H.,

Bustamante, C. D. & McCouch, S. (2007). Global Dissemination of a Single Mutation Conferring White Pericarp in Rice. *PLoS Genet* **3,** e133.

Tan, L., Li, X., Liu, F., Sun, X., Li, C., Zhu, Z., Fu, Y., Cai, H., Wang, X., Xie, D. & Sun, C. (2008). Control of a key transition from prostrate to erect growth in rice domestication. *Nature Genetics* **40,** 1360-1364.

Uga, Y., Fukuta, Y., Ohsawa, R. & Fujimura, T. (2003). Variations of floral traits in Asian cultivated rice (*Oryza sativa* L.) and its wild relatives (*O. rufipogon* Griff.). *Breeding Science* **53,** 345-352.

USDA, ARS, National Genetic Resources Program (2012). Germplasm Resources Information Network - (GRIN) [Online Database]. National Germplasm Resources Laboratory, Beltsville, Maryland. URL: http://www.ars-grin.gov/cgi-bin/npgs/html/index.pl (12 April 2012).

Vaughan, D. A., Lu, B.-R. & Tomooka, N. (2008). The evolving story of rice evolution. *Plant Science* **174,** 394-408.

Xu, J.-H., Cheng, C., Tsuchimoto, S., Ohtsubo, H. & Ohtsubo, E. (2007). Phylogenetic analysis of *Oryza rufipogon* strains and their relations to *Oryza sativa* strains by insertion polymorphism of rice SINEs. *Genes & Genetic Systems* **82,** 217-229.

Xu, X., Liu, X., Ge, S., Jensen, J. D., Hu, F., Li, X., Dong Y., Gutenkunst R. N., Fang, L., Huang, L. Li, J., He, W., Zhang, G., Zheng, X., Zhang, F., Li, Y., Yu, C., Kristiansen, K., Zhang, X., Wang, J., Wright, M., McCouch, S., Nielsen, R., Wang, J. & Wang, W. (2012). Resequencing 50 accessions of cultivated and wild rice yields markers for identifying agronomically important genes. *Nature Biotechnology* **30,** 105-111.

Yamanaka, S., Nakamura, I., Nakai, H. & Sato, Y.-I. (2003). Dual origin of the cultivated rice based on molecular markers of newly collected annual and perennial strains of wild rice species, *Oryza nivara* and *O. rufipogon*. *Genetic Resources and Crop Evolution* **50,** 529-538.

Zheng, X. & Ge, S. (2010). Ecological divergence in the presence of gene flow in two closely related *Oryza* species (*Oryza rufipogon* and *O. nivara*). *Molecular Ecology* **19,** 2439-2454.

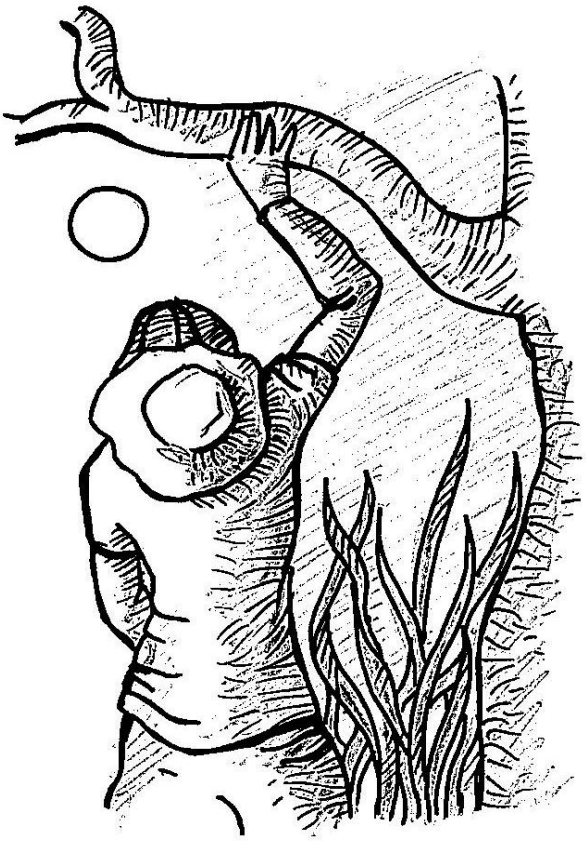
Zhou, H.-F., Zheng, X.-M., Wei, R.-X., Second, G., Vaughan, D. A. & Ge, S. (2008). Contrasting population genetic structure and gene flow between *Oryza rufipogon* and *Oryza nivara*. *Theoretical and Applied Genetics* **117,** 1181-1189.

Zhu, Q. & Ge, S. (2005). Phylogenetic relationships among A-genome species of the genus *Oryza* revealed by intron sequences of four nuclear genes. *New Phytologist* **167,** 249-265.

SUMMARY

SAMENVATTING

KABUURAN



Summary

The non-cultivated species of the genus *Oryza* can provide a genetic arsenal of useful traits for improving the widely cultivated and consumed Asian rice (*O. sativa*). The diversity of these valuable plant resources must be well understood to ensure their effective *in-* and *ex-situ* conservation. In this thesis, we examined the ecogeographic variations within and between the three species of *Oryza* series *Sativae* in Asia and the Pacific. We looked at species differentiation from different spatial scales by analysing sympatric accession pairs of *O. meridionalis* and *O. rufipogon* and of *O. nivara* and *O. rufipogon.*

We conducted phenotypic analyses in Chapter 2. The strong influence of ecology on species morphology was demonstrated in the ordination and cluster analyses results where *O. meridionalis* and *O. nivara* grouped together and were separated from *O. rufipogon*. We detected greater differentiation of *O. nivara* and *O. rufipogon* in South Asia and positive correlations between spatial and intraspecific (interpopulation) morphological distances in continental Asia. We found significant correlations between geoclimatic factors and certain character measurements within species and observed that seedling height, culm number and diameter, leaf size, and anther length exhibit contrasting responses for *O. nivara* and *O. rufipogon*. We confirmed significant morphological differences between the three species, between the South and Southeast Asian populations of *O. nivara*, and between the Australasian and the non-Australasian populations of *O. rufipogon* and provided botanical descriptions to delineate *O. meridionalis, O. nivara* and *O. rufipogon* morphologically.

In Chapter 3, we genotyped the same set of accessions with 29 SSR markers and applied a variety of methods for genetic diversity analysis. Based on ordination and phylogenetic results, we verified that *O. meridionalis* is a genetically distinct species and that *O. nivara* and *O. rufipogon* overlap genetically across their geographic distribution. However, Bayesian clustering analysis recognized local-scale species separation of *O. nivara* and *O. rufipogon* implying stronger interspecific gene flow barriers in smaller spatial units. Concurrently, AMOVA indicated that the bulk (64%) of genetic variation in Asia Pacific series *Sativae* can be found among accessions and the lesser portions within accessions (26%) and among species (10%). We captured contrasting intraspecific variation patterns for *O. nivara* and *O. rufipogon* where the former exhibited low diversity, high population differentiation and isolation by distance mainly in South Asia while the latter displayed high diversity, low population differentiation and isolation by

distance primarily in continental Southeast Asia. We established that altitude is correlated negatively to accession diversity and positively to local-scale species differentiation. Using Bayesian inference, we identified eight genetically distinct population groups: C1) Indian and Bangladeshi *O. nivara*; C2) Cambodian *O. nivara*; C3) Southeast Asian *O. rufipogon*; C4) *O. meridionalis*; C5) Nepalese *O. nivara*; C6) non-Cambodian Southeast Asian *O. nivara*; C7) Australasian *O. rufipogon*; and C8) South Asian *O. rufipogon*. Cluster analysis grouped the aromatic and japonica cultivar groups of *O. sativa* with *O. rufipogon* in South Asia and the indica and aus groups with *O. nivara* from Thailand and Cambodia, respectively. *O. nivara* from Nepal seemed genetically isolated from the other population groups. We also detected variation patterns that agreed with the results in Chapter 2 such as the South and Southeast Asian divisions of *O. nivara*, the divergence of Australasian populations from the rest of *O. rufipogon* and the greater differentiation of *O. nivara* and *O. rufipogon* in South Asia.

In Chapter 4, we conducted artificial crossing experiments to 15 selected parental accessions of *O. meridionalis*, *O. nivara*, and *O. rufipogon* and assessed the extent of several post-pollination isolating mechanisms in *Oryza* series *Sativae*. We observed reproductive incompatibility within and between the inbreeding species *O. meridionalis* and *O. nivara* and high intraspecific crossability of the outcrossing *O. rufipogon* where viable and non-sterile F1 hybrids were produced only by combinations with a parental distance that ranged from 1062 to 3813 kilometers. Insular Southeast Asian and/or Australasian accessions of *O. rufipogon* were the most reproductively successful parents. *O. rufipogon* exhibited significant pre-zygotic species isolation (in terms of seed set) and reduced post-zygotic isolation, and seemed symmetrically compatible with *O. nivara* and asymmetrically compatible with *O. meridionalis.* We obtained few annual hybrids with relatively high fertilities from crosses between *O. rufipogon* and *O. nivara* and numerous perennial hybrids with low fertilities from crosses between *O. rufipogon* and *O. meridionalis.* Crossability estimates did not show significant correlations with geographic distance between parents. However, we discerned reduced seed set and F1 fertility in interspecific combinations with sympatric parents compared to crosses with non-sympatric parents, indicative of reinforced species isolation in sympatry. We evaluated the F1 offspring of different cross combinations and found a mixture of intermediate and parental character traits in interspecific hybrids. The offspring of *O. meridionalis* and *O. rufipogon* conformed to the phenology of the maternal species while hybrids of *O. nivara* and *O. rufipogon* followed the morphology and phenology of the former regardless of parental combination. We also observed that Australasian and non-Australasian accessions of *O. rufipogon*

seemed to produce morphologically different intra- and interspecific offspring. Hybrids of Australasian *O. rufipogon* and *O. meridionalis* conformed to the morphology of the maternal species while hybrids of non-Australasian *O. rufipogon* and *O. meridionalis* appeared similar to *O. rufipogon* regardless of parental combination.

We discussed the taxonomic implications of the research results in Chapter 5 where we specifically dealt with the opposing views of lumping or splitting of *O. nivara* and *O. rufipogon*. We concluded that these two taxa deserve to be treated as separate species based on the following biosystematic evidence obtained from the thesis: 1) ecological distinction; 2) considerable prezygotic barriers; 3) opposing patterns of gene flow and genetic variation; 4) local-scale genetic divergence and 5) enhanced reproductive barriers under sympatric conditions. We identified ecogeography as a major driving force in the diversification of *Oryza* series *Sativae* in Asia and the Pacific and suggested that ecological speciation gave rise to *O. nivara* and *O. rufipogon.* We also presented recognizable geographic races within species.

Ultimately in Chapter 6, we emphasized the importance of our study in several aspects of rice science and identified results that agreed with prior *Oryza* diversity studies. At the same time, we presented previously unreported morphological and genetic variation patterns that were established in this thesis. We discussed the possible applications of the research results to wild rice conservation, covering *in situ* strategies as well as gene bank practices. We also highlighted the potential role of *O. nivara* in Asian rice domestication where it could have either directly given rise to the indica cultivar group or hybridized/introgressed with migrated japonica cultivars in India, eventually leading to the development of indica.

Samenvatting

De niet-gecultiveerde soorten van het geslacht *Oryza* kunnen een arsenaal aan nuttige eigenschappen leveren om de veel geteelde en geconsumeerde Aziatische rijst (*O. sativa*) te verbeteren. De diversiteit van deze waardevolle plantaardige hulpbronnen dient goed begrepen te worden om hun effectieve *in-* en *ex-situ* conservering te garanderen. In dit proefschrift onderzochten wij de ecogeografische variatie binnen en tussen de drie soorten van *Oryza* series *Sativae* in Azië en het pacifische gebied. We bekeken de soortsdifferentiatie op verschillende ruimtelijke schalen door sympatrische paren accessies van *O. meridionalis* en *O. rufipogon* en van *O. nivara* en *O. rufipogon* te analyseren.

Wij voerden fenotypische analyses uit in hoofdstuk 2. De sterke invloed van ecologie op de morfologie van soorten werd duidelijk in de resultaten van de ordinatie en cluster analyses waar *O. meridionalis* en *O. nivara* een groep vormden, apart van *O. rufipogon*. Wij stelden een grotere differentiatie van *O. nivara* en *O. rufipogon* vast in Zuid Azië en positieve correlaties tussen ruimtelijke en morfologische afstanden binnen de soorten in continentaal Azië. We vonden significante correlaties tussen klimatologische factoren en bepaalde gemeten kenmerken binnen soorten en namen waar dat zaailinghoogte, aantal en doorsnede van de halmen, afmeting van de bladeren en lengte van de helmhokken verschillende reacties vertoonden voor *O. nivara* and *O. rufipogon*. Wij bevestigden dat er significante morfologische verschillen aanwezig zijn tussen de drie soorten, tussen de populaties van *O. nivara* in Zuid en Zuidoost Azië, en tussen de populaties van *O. rufipogon* in Australazië en daarbuiten. Wij stelden botanische beschrijvingen op om *O. meridionalis, O. nivara* en *O. rufipogon* morfologisch af te bakenen.

In hoofdstuk 3 genotypeerden wij dezelfde set accessies met 29 SSR merkers en pasten een scala aan methoden toe om de genetische diversiteit te analyseren. Gebaseerd op de resultaten van ordinatie en fylogenetische benaderingen stelden we vast dat *O. meridionalis* genetisch een duidelijke soort is en dat *O. nivara* en *O. rufipogon* genetische overlap vertonen in hun geografisch verspreidingsgebied. Echter, Bayesiaanse clustering toonde aan dat *O. nivara* en *O. rufipogon* op locale schaal van elkaar onderscheiden konden worden, hetgeen impliceerde dat er sterkere barrieres tegen genenoverdracht aanwezig waren in kleinere ruimtelijke eenheden. Ook de AMOVA gaf aan dat het grootste deel (64%) van de genetische variatie in the Aziatische Pacifische series *Sativae* aangetroffen wordt tussen de

accessies en een kleiner deel (26%) binnen de accessies en tussen de soorten (10%). We namen verschillende infraspecifieke variatie patronen waar voor *O. nivara* en *O. rufipogon*, waarbij eerstgenoemde een lage diversiteit, een hoge differentiatie tussen populaties en een hoge isolatie over afstand vertoonde, voornamelijk in Zuid Azië, terwijl laatstgenoemde juist gekarakteriseerd werd door een hoge diversiteit, en een lage differentiatie tussen populaties en een lage isolatie over afstand, voornamelijk in continentaal Zuidoost Azië. Hoogte boven zeeniveau bleek negatief gecorreleerd met de diversiteit binnen accessies en positief met locale soortsdifferentiatie. Met Bayesiaanse methoden identificeerden wij acht genetische verschillende groepen populaties: C1) *O. nivara* in India and Bangladesh; C2) *O. nivara* in Cambodia; C3) *O. rufipogon* in Zuidoost Azië; C4) *O. meridionalis*; C5) *O. nivara* in Nepal; C6) *O. nivara* in Zuidoost Azië buiten Cambodja; C7) *O. rufipogon* in Australazië; en C8) *O. rufipogon* in Zuid Azië. Cluster analyse groepeerde de aromatische en japonica cultivar groepen van *O. sativa* met *O. rufipogon* in Zuid Azië en de indica en aus groepen met *O. nivara* van Thailand en Cambodja. *O. nivara* in Nepal leek genetisch geisoleerd van de andere groepen populaties. Wij ontdekten ook variatiepatronen die overeen kwamen met de resultaten van hoofdstuk 1, zoals de onderverdeling in Zuid tegen Zuidoost Azië binnen *O. nivara*, de divergentie van de Australaziatische populaties van de rest van *O. rufipogon* en de grotere differentiatie van *O. nivara* en *O. rufipogon* in Zuid Azië.

In hoofdstuk 4, voerden wij kruisingsexperimenten uit met 15 geselecteerde ouder accessies van *O. meridionalis*, *O. nivara*, en *O. rufipogon* en onderzochten de aanwezigheid van verscheidene isolatie mechanismen na de bestuiving in *Oryza* series *Sativae*. Wij vonden reproductieve incompatibiliteit binnen en tussen de zelfbestuivende soorten *O. meridionalis* en *O. nivara* en een hoge mate van infraspecifieke kruisbaarheid bij de kruisbestuivende *O. rufipogon*, waar levensvatbare en niet-steriele F1 hybriden slechts werden geproduceerd in combinaties van ouders die tussen de 1062 to 3813 kilometers van elkaar verwijderd waren. Accessies van *O. rufipogon* waren het meest succesvol als ouders als ze afkomstig waren van de Zuidoost Aziatische eilanden en/of Australazië. *O. rufipogon* vertoonde significante pre-zygotische isolatie (in termen van de zaad productie) en gereduceerde post-zygotische isolatie, en leek symmetrische compatibel met *O. nivara* en asymmetrisch met *O. meridionalis.* Wij verkregen weinig eenjarige hybriden, maar met relatief hoge fertiliteit, uit kruisingen tussen *O. rufipogon* en *O. nivara* en talrijke meerjarige hybriden met lage fertiliteit uit kruisingen tussen *O. rufipogon* en *O. meridionalis.* De kruisbaarheid vertoonde geen significante correlaties met de geografische afstanden tussen de ouders. Echter, de

combinaties tussen de soorten met sympatrische ouders vertoonden gereduceerde zaadzetting en F1 fertiliteit in vergelijking met niet-sympatrische ouders. Dit wijst op versterkte sympatrische soortsisolatie. In de F1 nakomelingen van de verschillende kruisingscombinaties vonden wij een mix van intermediaire en ouderlijke kenmerken in de soortshybriden. De nakomelingen van *O. meridionalis* en *O. rufipogon* kwamen overeen met de fenologie van de vrouwelijke ouder, terwijl de hybriden van *O. nivara* en *O. rufipogon* de morfologie en fenologie van *O. rufipogon* volgden, onafhankelijk van welke soort als moeder optrad. De Australaziatische en niet-Australaziatische accessies van *O. rufipogon* leken morfologisch verschillende binnen- en tussen-soortelijke nakomelingen te produceren. Hybriden van Australaziatische *O. rufipogon* en *O. meridionalis* leken morfologisch op de vrouwelijke oudersoort, terwijl de hybriden van niet-Australaziatische *O. rufipogon* en *O. meridionalis* leken op *O. rufipogon* ongeacht de combinatie van ouders.

De taxonomische implicaties van het onderzoek werden in hoofdstuk 5 besproken, waarbij we vooral ingingen op de tegenstrijdige opvattingen over het 'lumpen' en 'splitten' van *O. nivara* en *O. rufipogon*. Wij concludeerden dat deze twee taxa als aparte soorten beschouwd zouden moeten worden op basis van de volgende biosystematische argumenten: 1) ecologisch onderscheid; 2) aanzienlijke prezygotische barrieres; 3) tegengestelde patronen van gene flow and genetische variatie; 4) genetische divergentie op locale schaal en 5) verhoogde reproductieve barrieres onder sympatrische omstandigheden. Wij identificeerden de ecogeografie als de voornaamste kracht achter het ontstaan van de verscheidenheid in *Oryza* series *Sativae* in Azië en het pacifisch gebied, en suggereerden dat *O. nivara* and *O. rufipogon* door ecologische soortsvorming zijn ontstaan. Wij wezen ook op het bestaan van herkenbare geografische populatie groepen binnen de soorten.

Tenslotte benadrukten wij in hoofdstuk 6 het belang van deze studie voor verschillende aspecten van de studie van rijst en identificeerden de resultaten die overeen kwamen met eerdere studies van de diveristeit in *Oryza*. Tegelijkertijd presenteerden wij nog onbekende morfologische en genetische variatiepatronen die in dit proefschrift werden vastgesteld. Wij bespraken de mogelijke toepassingen van het onderzoek bij het conserveren van wilde rijst, zowel bij *in situ* strategien als bij de praktijk van genenbanken. Wij belichtten de mogelijke rol die *O. nivara* in de domesticatie van Aziatische rijst heeft gespeeld, waarbij de soort of direct tot de indica cultivar groep heeft geleid, of na hybridisatie/introgressie met gemigreerde japonica culitvars in India, uiteindelijk resulterend in de ontwikkeling van de indica groep.

Kabuuran

Ang mga uri o *species* ng palay na hindi itinatanim at kabilang sa pangkat ng *Oryza* ay maaaring pagkunan ng mga kapaki-pakinabang na katangian para sa ibayong pagpapabuti ng palay sa Asya (*O. sativa*) na laganap na itinatanim at kinakain ng mga tao. Ang mga uri ng ligaw na palay na ito ay mahahalagang halamang pinagkukunan at ang kanilang pagkakaiba-iba ay dapat maintindihang mabuti upang matiyak na sila ay mapangalagaan nang maayos saan man sila tumutubo – *in situ* (orihinal na lugar) o *ex situ* (pinaglipatang lugar). Sa tesis na ito, aming sinuri ang mga pagkakaibang nakapaloob at namamagitan sa tatlong uri ng *Oryza* series *Sativae* sa Asya at sa Pasipiko base sa ekolohiya at heograpiya. Tiningnan namin ang antas ng pagkakaiba ng mga uri sa iba't ibang lawak ng heyograpikong espasyo sa pamamagitan ng pagsisiyasat ng mga pares ng *sympatric* (kinolekta sa iisang lokalidad) na *accession* ng *O. meridionalis* at *O. rufipogon*, gayundin ng *O. nivara* at *O. rufipogon*.

Sa ikalawang kabanata, gumawa kami ng pagsusuri ng *phenotype* (mga katangiang bunga ng hene at kapaligiran). Ang malakas na impluwensiya ng ekolohiya sa morpolohiya o kaanyuan ay naipakita sa mga resulta ng *ordination* at *cluster analysis* kung saan ang *O. meridionalis* at *O. nivara* ay nagsama sa isang grupo at napahiwalay sa *O. rufipogon.* May napansin kaming malaking kaibahan ng *O. nivara* at *O. rufipogon* sa Timog Asya at positibong korelasyon ng distansya ng pinagmulang lugar sa pagkakaiba ng mga populasyon na kabilang sa iisang uri sa kontinenteng Asya. May nakita kaming makahulugang korelasyon sa pagitan ng mga salik ng heyograpiya at klima at sukat ng ilang katangiang nakapaloob sa iisang uri at aming napuna na magkaiba ang tugon ng *O. nivara* at *O. rufipogon* kung pag-uusapan ang taas ng punla, bilang at diyametro ng *culm* (pinakapuno ng uhay), laki ng dahon, at haba ng anter. Aming napatunayan na may mga mahahalagang pagkakaiba sa kaanyuan ng tatlong uri na aming pinag-aralan; may pagkakaiba din ang populasyon ng *O. nivara* sa Timog Asya at sa Timog-silangang Asya; hindi rin magkapareho ang populasyon ng *O. rufipogon* sa Australasia at ang mga populasyon na hindi tubo sa lugar na ito. Nagbigay kami ng paglalarawan upang makita ang kaibhan sa anyo ng *O. meridionalis*, *O. nivara*, at *O. rufipogon*.

Sa Kabanata 3, inalam namin ang *genotype* (heneng taglay) ng katulad na pangkat ng mga *accession* gamit ang 29 *SSR marker* at may mga pamamaraan kaming ginamit upang lalong matukoy ang pagkakaiba-ibang henetiko ng mga ito. Batay sa resulta ng *ordination* at *phylogenetic analysis*, nalaman namin na: a) may katangitanging pagkakaiba ang *O. meridionalis*, at b) nagkakasanib-sanib ang mga

katangiang henetiko ng *O. nivara* at *O. rufipogon* kung titingnan ang kabuuang saklaw ng kanilang distribusyon. Gayunman, ang pagsusuri na gamit ang *Bayesian clustering* ay nagpahiwatig ng paghihiwalay ng *O. nivara* at *O. rufipogon* sa lokal na antas. Ibig sabihin, mas malakas ang mga sagabal sa pagdaloy ng hene sa pagitan ng magkaibang uri kapag maiksi ang distansya ng lugar na pinagmulan ng mga populasyon. Kasabay nito, ang pagsusuring gamit ang *AMOVA* ay nagsabing ang pinakamalaking bahagdan ng henetikong pagkakaiba-iba sa series *Sativae* sa Asya-Pasipiko ay namamagitan sa mga *accession* (64%) at ang mas mas maliit na bahagi ay nakapaloob sa mga *accession* at namamagitan sa mga uri (10%). Nalaman namin na ang pagkakaibang nakapaloob sa *O. nivara* ay di tulad ng pagkakaibang nakapaloob sa *O. rufipogon.* Kakaunti ang pagkakaiba-ibang nakapaloob sa mga populasyon ng *O. nivara*, malaki ang pagkakalayo (*population differentiation*) ng mga ito at higit na nagpapakita ng pagkakabukod sa distansiya (*isolation by distance*) sa Timog Asya. Malaki naman ang pagkakaibang nakapaloob sa mga populasyon ng *O. rufipogon*, kakaunti ang pagkakalayo ng mga ito at higit na nagpapakita ng pagkakabukod sa distansya sa kontinente ng Timog-silangang Asya. Napagtibay namin na ang taas ng lugar ay may negatibong korelasyon sa pagkakaibang nakapaloob sa mga *accession* at positibong korelasyon sa pagkakalayo ng mga uri sa lokal na antas (*local-scale species differentiation*). Gamit ang *Bayesian inference*, kinilala namin ang walong magkakaibang grupo ng populasyon: C1) *O. nivara* ng India at Bangladesh; C2) *O. nivara* ng Cambodia; C3) *O. rufipogon* ng Timog-silangang Asya; C4) *O. meridionalis*; C5) *O. nivara* ng Nepal; C6) *O. nivara* ng Timog-silangang Asya na di galing sa Cambodia; C7) *O. rufipogon* ng Australasia; at C8) *O. rufipogon* ng Timog Asya. Sa pagsagawa ng *cluster analysis,* napabilang ang mga kultibar ng *O. sativa* na *aromatic* at *japonica* sa grupo ng *O. rufipogon* ng Timog Asya, habang ang *indica* ay napasama sa *O. nivara* ng Thailand at ang *aus* sa *O. nivara* ng Cambodia. Katangitangi at tila nakahiwalay ang *O. nivara* ng Nepal kumpara sa ibang grupo ng populasyon. Nalaman namin na ang ilan sa pagkakaiba-ibang henetiko na natuklasan ay sumasang-ayon sa resultang nakasaad sa ikalawang kabanata: ang pagkakahahati ng *O. nivara* sa pangkat ng Timog-Asya at Timog-silangang Asya; ang paghiwalay o paglihis ng populasyon ng *O. rufipogon* sa Australasia mula sa ibang populasyon ng naturang uri; at ang mas malaking pagkakaiba ng *O. nivara* at *O. rufipogon* sa Timog Asya.

Sa Kabanata 4, gumawa kami ng mga pagsubok sa artipisyal na paglalahi gamit ang napiling 15 *accession* ng *O. meridionalis*, *O. nivara*, at *O. rufipogon* at aming tinantiya ang saklaw ng ilang mekanismong sumasagabal sa matagumpay na reproduksyon, matapos ang polinasyon sa *Oryza* series *Sativae.* Napansin namin ang di magkatugmang pagtatambal sa pagitan ng mga uring *inbreeding* na *O.*

meridionalis at *O. nivara*, maging sa pagitan ng mga *accession* na kabilang sa isa sa mga ito, pati na ang matagumpay na pagtatambal ng mga *accession* ng uring *outcrossing* na *O. rufipogon* kung saan ang mga buhay at hindi baog na F1 haybrid ay naani mula sa mga kumbinasyon na ang agwat ng pinagmulang lugar ng mga magulang ay mula 1,062 hanggang 3,813 kilometro. Ang mga *accession* ng *O. rufipogon* na galing sa kapuluan ng Timog-silangang Asya at/o Australasia ang pinakamainam na magulang para sa matagumpay na paglalahi. Ang *O. rufipogon* ay nagpakita ng malakas na *pre-zygotic isolation* kung pag-uusapan ang *seed set*, mas mahinang *post-zygotic isolation*, tila pantay na pagkakatugma sa *O. nivara*, at di pantay na pagkakatugma sa *O. meridionalis*. Nakakuha kami ng ilang haybrid na taunan at medyo mataas ang kakayahang mamunga mula sa paglalahi ng *O. rufipogon* at *O. nivara* at maraming haybrid na pangmatagalan at mahina ang kakayahang mamunga mula sa paglalahi ng *O. rufipogon* at *O. meridionalis*. Ang mga sukatan ng kakayahang magpalahi ay walang makabuluhang korelasyon sa distansiya o layo ng pinagmulang lugar ng mga magulang. Gayunpaman, napuna namin ang mababang *seed set* at kakayahang mamunga ng mga F1 haybrid ng dalawang uri mula sa paglalahi na ginamitan *sympatric* na magulang kumpara sa paglalahi na ginamitan ng mga magulang na hindi *sympatric*; ito ay nagpapahayag na ang mga balakid sa paglalahi ng dalawang uri ay mas malakas sa pagitan ng mga populasyon na matatagpuan sa iisang lokalidad. Pinag-aralan namin ang mga supling (*F1 offspring*) ng iba't-ibang kumbinasyon ng paglalahi at nakita naming ang mga haybrid ng magkaibang uri ay nagtataglay ng magkahalong katangiang pumapagitna (*intermediate*) at kahalintulad sa mga magulang. Ang mga supling ng *O. meridionalis* at *O. rufipogon* ay gumaya sa penolohiya ng inang uri samantalang ang mga haybrid mula sa *O. nivara* at *O. rufipogon* ay sumunod sa morpolohiya at penolohiya ng *O. nivara*, anumang kumbinasyon ng magulang ang ginamit. Naobserbahan din namin na tila magkaiba ang anyo ng mga supling ng mga *accession* ng *O. rufipogon* mula sa loob at labas ng Australasia. Ang mga haybrid na galing sa *O. meridionalis* at *O. rufipogon* ng Australasia ay sumunod sa kaanyuan ng inang uri samantalang ang mga haybrid ng *O. meridionalis* at *O. rufipogon* mula sa labas ng Australasia ay tumulad sa anyo ng *O. rufipogon* anumang kumbinasyon ng magulang ang ginamit.

Aming tinalakay ang mga implikasyon sa taksonomiya ng mga resulta ng pananaliksik sa Kabanata 5. Dito ay ipinakita namin ang magkasalungat na pananaw tungkol sa pagbubuklod o paghihiwalay ng *O. nivara* at *O. rufipogon*. Ipinalagay namin na ang dalawang *taxa* na ito ay karapatdapat na ituring na magkahiwalay na uri batay sa mga ebidensyang nakalap mula sa tesis: 1) pagkakaiba ayon sa ekolohiya; 2) mga hindi mumunting hadlang sa pagbuo ang

zygote; 3) kaibahan sa pagdaloy ng hene sa mga populasyon at magkasalungat na pagkakaibang henetikong nakapaloob sa bawat uri; 4) pagkakalayong henetiko sa lokal na antas; at 5) pinagtibay na sagabal sa paglalahi sa ilalim ng *sympatric* na kundisyon. Kinilala namin ang ekoheyographiya bilang pangunahing puwersa na nagdulot ng pagkakaiba-iba sa grupong *Sativae* ng *Oryza* sa Asya at Pasipiko. Aming iminungkahi na ang *ecological speciation* ang nagbunsod sa pag-usbong ng *O. nivara* at *O. rufipogon*. Amin ding ipinakita na may mga kakila-kilalang *geographic race* na nakapaloob sa mga siniyasat na uri.

Bilang panghuli, sa Kabanata 6, binigyang diin namin ang kahalagahan ng aming pag-aaral sa ilang aspeto ng siyensiya ng palay at tinukoy ang mga resulta ng aming pananaliksik na sumang-ayon sa mga naunang ulat hinggil sa pagkakaiba-iba ng grupong *Oryza*. Kasabay nito, aming inilahad ang mga hindi pa naiulat na pagkaka-iba-iba ng kaanyuan at ng aspetong henetiko na natuklasan sa tesis na ito. Tinalakay namin ang mga posibleng aplikasyon ng aming resulta sa pangangalaga ng ligaw na palay, na sumasaklaw sa mga estratehiyang *in situ* at sa mga pamamaraan ng pagpapatakbo ng *genebank*. Gayundin, ipinakita namin ang potensiyal na papel ng *O. nivara* sa domestikasyon ng palay sa Asya: maaaring ito ang diretsong pinanggalingan ng kultibar na *indica* o ang paglahi nito sa mga kultibar na *japonica* sa India ay humantong sa pag-usbong ng *indica*.

Acknowledgements

My sincerest gratitude goes to everyone who contributed to the completion of this work.

I am deeply grateful to my thesis committee. Prof. Marc Sosef provided guidance and made the necessary arrangements for the completion of my PhD program. Dr. Ronald van den Berg served as my academic mentor and at times quasi-parent (during my short stays in Wageningen) and did it in the coolest way possible. He also translated my thesis summary in Dutch. My home institute co-supervisor Dr. Kenneth McNally introduced a lot of improvements in the research and provided the needed push to beat the deadline for manuscript submission.

I am indebted to the T.T. Chang Genetic Resources Center (TTC-GRC) in IRRI for supporting my thesis project financially and otherwise. Dr. Ruaraidh Sackville Hamilton presented to me the possibility of pursuing PhD studies in Wageningen University. He eventually imparted valuable intellectual contributions in the thesis. Ma. Socorro Almazan accompanied me in the Laos field work, arranged for the processing of my plant materials and together with the wild rice nursery support staff [Remegio Aguilar, Vicente Arcillas (former staff), Edwin Jarabejo, Nora Kuroda, Emmanuel Manaig, Yolanda Malatag, Ryan Mercado (former staff), Mae Merluza and Liza Yonzon] took care of the plants in the screenhouse and facilitated the phenotyping and crossing activities. The Genomic Diversity Laboratory team [Belinda Caspillo, Michael Gamalinda (former staff), Jocelyn Torres and Janice Victoria] with Nora Kuroda, Mae Merluza and Liza Yonzon provided laboratory support, specifically in DNA extraction and acrylamide gel preparation. Sheila Mae Mercado generously taught me about SSR genotyping and data analysis. Ma. Elizabeth Naredo gave helpful advices on dealing with ssr and crossing data. She and I wrote our manuscripts at around the same time and shared the agonies and little successes of thesis writing. The administrative support team (Teresita Santos, Anthony Telosa, Gladys Ebron and the late Digna Salisi) also provided office-related assistance.

I greatly appreciate the insights and recommendations given by: Dr. Robert Hijmans (on georeferencing); Dr. Helena Duistermaat (on *Oryza* taxonomy); Dr. Rene Smulders (on SSR data analysis); Dr. Duncan Vaughan (on *Oryza* taxonomy); the PBGB-CRIL team (Violeta Bartolome, Leilani Nora and Rose Imee Zhella Morantte) and Nadine Singson (on statistical methods).

I would like to thank: Dr. Theo van Hintum for being my external supervisor; WUR alumni Dr. Alice Laborte and Dr. Annie Boling for the survival tips; Dr. Chay Bounphanousay, Souvanh Thadavong and Tuy for organizing the field work in Vientiane, Laos; Jedi Ferrater for sharing her special recipe for the perfect acrylamide gel; Raymond Panaligan and Chris Quintana for the photography services; Jimboy Bactong for patiently designing the thesis cover and Teresita Rola for accepting the gruelling task of translating the thesis summary to Filipino.

I would also like to express my gratitude to the Filipino community in Wageningen including the batches of WUR Pinoy students from 2007 to 2012 and my fellow PhD students (Tanawat, Alvaro Monteros, Ximena Cadima Fuentes, Pulcherie Bissiengou and Mariafe Calingacion) for the camaraderie. I owe many thanks to Fe and Pulcherie for being my trustworthy paranymphs.

Finally, I would like to dedicate this book to God, and to my family and friends in the Philippines. I am blessed with a strong and dependable social support system [Ate Reggie and Jon (my pc problem fixer), Boy, Carrie and Rain, Tatay Frank and the rest of Hilario family, the Hemedez clan]. I am forever grateful to my parents, Rennie and Pida, as well as to Nanay Ampy and Yayay Maring Oclarino. They selflessly devoted extra time and energy to look after my daughter when I had to be away from home. I would not have survived my PhD program without the love, support and inspiration from my husband Paul Benjamin (who created the artwork in this book) and my daughter Yja Celestine. My gratitude for them goes beyond words.

Curriculum Vitae

Maria Celeste N. Banaticla-Hilario was born on May 4, 1977 in Calamba, Laguna, Philippines where she was raised and had her primary and secondary education. In 1994, she entered the University of the Philippines in Los Baños (UPLB) and pursued a degree of Bachelor of Science in Biology. She majored in systematics and completed her undergraduate thesis on the seed morphology of Philippine pioneer tree species while working as a student assistant in the Ethnomorphosystematics Laboratory of the Institute of Biological Sciences in the same university. After graduating in 1998, she had a one-year stint as an instructor in Central Luzon State University where she taught basic biology, botany and taxonomy courses. From 1999 – 2001, as a research specialist for a project funded by the National Research Council of the Philippines (NRCP), she participated in fieldwork surveys and co-wrote the "Identification Handbook of Philippine Commercial and Potentially Commercial Forest Vines". At around the same time, she was enrolled in the Master of Science in Botany program of the graduate school of UPLB which she finished in 2003. Her master's thesis titled "Altitudinal Gradient Distribution of Pteridophytes on Mt. Banahaw de Lucban, Luzon Island, Philippines" received multiple grants (from UPLB, NRCP and the Southeast Asian Ministers of Education Organization - Regional Center for Graduate Study in Agriculture) and spawned several publications. She was employed by the International Rice Research Institute in 2003 to serve as the resident wild rice taxonomist of the International Rice Genebank. In 2007, Wageningen University awarded her a sandwich PhD fellowship that culminated in this thesis.

List of publications

Bioversity International, IRRI and WARDA. 2007. Descriptors for wild and cultivated rice (*Oryza* spp.). Bioversity International, Rome, Italy; International Rice Research Institute, Los Baños, Philippines; WARDA, Africa Rice Center, Cotonou, Benin (contributor)

Banaticla, M. C. N. and I. E. Buot Jr. 2006. Endemic ferns on the Northeastern Slopes of Mt. Banahaw de Lucban, Sierra Madre Mountain Range, Southern Luzon, Philippines: potentials for domestication and landscaping. Journal of Nature Studies 5(1): 13 – 22

Banaticla, M. C. N. and I. E. Buot Jr. 2005. Altitudinal Zonation of Pteridophytes on Mt. Banahaw de Lucban, Luzon Island, Philippines. Plant Ecology 180:135-151

Almazan, M.S.R., M. E. B. Naredo, M. C. N. Banaticla and N.R. Sackville-Hamilton. 2005. Management of wild rice collections in the International Rice Genebank at IRRI, Philippines. Journal of Nature Studies 4(1):29-35

Banaticla, M. C. N. and I. E. Buot Jr. 2004. Leaf Architecture of Ten Philippine *Psychotria* species (Rubiaceae). The Philippine Scientist 41: 74 – 90

Banaticla, M.C.N. and M.S. Almazan. 2004. Vanishing populations of *Oryza minuta* Presl in Pangil, Laguna, Philippines. International Rice Research Notes 29 (2): 88 – 89

Banaticla, M.C.N. and IE Buot Jr. 2004. Fern Patch Structure and Species Diversity along the Altitudinal Gradient of Mt. Banahaw de Lucban, Luzon Island, Philippines. Philippine Agricultural Scientist 87(1): 49 – 60

Banaticla, M.C.N. and IE Buot Jr. 2003. Dominant Pteridophyte Species of Mt. Banahaw de Lucban, Luzon Island, Philippines. Flora Malesiana Bulletin 13 (3): 281 – 283

Banaticla, M.C.N. Cyathea contaminans (Wall. ex Hook.) Copel. and its Occurrence in Mt. Banahaw de Lucban, Quezon. In Buot, I. E. and R.T. Bagarinao eds. 2003. Biodiversity Conservation and Ecotourism: Subjects, Theories and External Pressures. Philippine Society for the Study of Nature: Los Baños, Laguna. pp. 35 – 38

Escobin, R. P. and M. C. N. Banaticla. 2000. The Identification and Utilization of Eight Philippine Major Commercial Forest Vines for the Handicraft Industry. FPRDI Journal 26 (1 & 2): 121 – 135

BOOK

Escobin, R. P. E. and M. C. N. Banaticla. 2005. Identification Handbook of Philippine Commercial and Potentially Commercial Forest Vines. College, Laguna: Forest Products Research and Development Institute. 199 p.

Education Statement of the Graduate School Experimental Plant Sciences

Issued to: **Maria Celeste N. Banaticla-Hilario**
Date: **24 October 2012**
Group: **Biosystematics, Wageningen University & Research Centre**

1) Start-up phase	*date*
► **First presentation of your project**	
Ecogeographic analysis of Oryza series Sativae from the Asia Pacific Region	Oct 09, 2007
► **Writing or rewriting a project proposal**	
Ecogeographic analysis of Oryza series Sativae from the Asia Pacific Region	Oct-Nov 2007
► **Writing a review or book chapter**	
► **MSc courses**	
Introduction to Geo-information Science (GRS-10306)	Sep-Oct 2007
► **Laboratory use of isotopes**	
Subtotal Start-up Phase	*13.5 credits**

2) Scientific Exposure	*date*
► **EPS PhD student days**	
EPS PhD Students Day 2007	Sep 13, 2007
► **EPS theme symposia**	
► **NWO Lunteren days and other National Platforms**	
► **Seminars (series), workshops and symposia**	
Biosystematics group seminar series (2 seminars)	Aug-Oct 2007
Biosystematics group seminar series (3 seminars)	Jul-Oct 2010
WEES seminar: The value of biodiversity (prof. Bas Haring)	Sep 16, 2010
Mini-symposium: 'How to write a world-class article'	Oct 26, 2010
► **Seminar plus**	
► **International symposia and congresses**	
8th Annual Scientific Convention of the Philippine Society for the Study of Nature, Ilocos Norte, Philippines	May 05-10, 2008
6th Rice Genetics Symposium, Manilla, Philippines	Nov 16-19, 2009
International Rice Congress, Hanoi, Vietnam	Nov 08-12, 2010
► **Presentations**	
8th Annual Scientific Convention of the Philippine Society for the Study of Nature (poster)	May 05-10, 2008
6th Rice Genetics Symposium (poster)	Nov 16-19, 2009
International Rice Congress(pster)	Nov 08-12, 2010
12th Annual Scientific Convention of the Philippine Society for the Study of Nature (oral)	May 22-27, 2012
► **IAB interview**	
► **Excursions**	
Subtotal Scientific Exposure	*9.9 credits**

3) In-Depth Studies	*date*
► **EPS courses or other PhD courses**	
Molecular Phylogenies: reconstruction and interpretation	Oct 15-19, 2007
PhD Summer School 'Natural variation of plants'	Aug 21-24, 2012
► **Journal club**	
Biosystematics Group	Jul-Oct 2010
► **Individual research training**	
Training in SSR techniques (Molecular Marker Lab, GRC-IRRI)	2008
Subtotal In-Depth Studies	*6.0 credits**

4) Personal development	*date*
► **Skill training courses**	
Basic Experimental Design and Data Analysis (at IRRI)	Feb 18 -22, 2008
Bioinformatics Workshop for Crop Researchers (at IRRI)	Mar 24-28, 2008
Research Data Management Course (at IRRI)	Dec 02-04, 2008
Introduction to R Course (at IRRI)	May 26-28, 2009
Introduction to R Course: statistical analysis (at IRRI)	Oct 26-30, 2009
SNP Data Analysis (at IRRI)	Mar 08-11, 2011
► **Organisation of PhD students day, course or conference**	
► **Membership of Board, Committee or PhD council**	
Subtotal Personal Development	*7.5 credits**

TOTAL NUMBER OF CREDIT POINTS*	**36.9**

Herewith the Graduate School declares that the PhD candidate has complied with the educational requirements set by the Educational Committee of EPS which comprises of a minimum total of 30 ECTS credits

** A credit represents a normative study load of 28 hours of study.*

This thesis project is a collaboration between the T.T. Chang Genetic Resources Center in the International Rice Research Institute and the Biosystematics Group of Wageningen University.

Cover art (The Germplasm Collector) and illustrations by Paul Hilario (www.facebook/Paul.Hilario)

Cover design by Jimboy Bactong

Thesis lay out by the author

Printed byWöhrmann Print Service, Zutphen, the Netherlands