

•Invited Review•

Evolutionary Genomics of Weedy Rice in the USA

Kenneth M. Olsen^{1*}, Ana L. Caicedo² and Yulin Jia³

¹Department of Biology, Washington University, St Louis, MO 63130, USA;

²Department of Biology, University of Massachusetts, Amherst, MA 01003, USA;

³USDA-ARS Dale Bumpers National Rice Research Center, Stuttgart, AR 72160, USA)

Abstract

Red rice is an interfertile, weedy form of cultivated rice (*Oryza sativa* L.) that competes aggressively with the crop in the southern US, reducing yields and contaminating harvests. No wild *Oryza* species occur in North America and the weed has been proposed to have evolved through multiple mechanisms, including “de-domestication” of US crop cultivars, accidental introduction of Asian weeds, and hybridization between US crops and Asian wild/weedy *Oryza* strains. The phenotype of US red rice ranges from “crop mimics”, which share some domestication traits with the crop, to strains closely resembling Asian wild *Oryza* species. Assessments of genetic diversity have indicated that many weed strains are closely related to Asian taxa (including *indica* and *aus* rice varieties, which have never been cultivated in the US, and the Asian crop progenitor *O. rufipogon*), whereas others show genetic similarity to the *tropical japonica* varieties cultivated in the southern US. Herein, we review what is known about the evolutionary origins and genetic diversity of US red rice and describe an ongoing research project to further characterize the evolutionary genomics of this aggressive weed.

Key words: crop-weed introgression; de-domestication; *Oryza sativa*; weed evolution; weedy rice.

Olsen KM, Caicedo AL, Jia Y (2007). Evolutionary genomics of weedy rice in the USA. *J. Integr. Plant Biol.* 49(6), 811–816.

Available online at www.blackwell-synergy.com/links/toc/jipb, www.jipb.net

Over half of the world's top 10 crop species possess con-specific weedy forms (Ellstrand et al. 1999), which compete aggressively with crops, reduce yields, and contaminate harvests. Because they are interfertile with crops, weedy relatives can also serve as an avenue of escape for engineered traits or introduce undesirable traits into crops. Despite their detrimental impact on crop yields worldwide, little is known about the population genetic dynamics by which weed-associated traits arise and persist in weedy crop relatives. Weed phenotypes are typically assumed to arise through

a combination of genomic introgression (resulting from hybridization between crops and related wild species) and selection acting to favor weed-adaptive traits.

Rice (*Oryza sativa* L.) is among the world's most important food crops; rice consumption constitutes approximately 20% of the world's caloric intake (Food and Agricultural Organization of the United Nations (FAO) 2003). Weedy rice (also commonly referred to as red rice) co-occurs with the crop wherever it is cultivated. Strains of the weed, although phenotypically highly variable, often have traits characteristic of non-domesticated *Oryza* species, including seed dispersal mechanisms (e.g. seed shattering, awns), seed dormancy, and small, dark-pigmented pericarps. However, unlike truly wild species, these weedy strains are best adapted to growing in agricultural fields, where they mimic and/or outcompete the crop.

In regions of rice cultivation where native *Oryza* species occur, such as Asia and Latin America, weedy rice may have arisen through a continual process of gene flow between the cultivated crop and neighboring wild populations. In world regions where no wild crop relatives occur, including the US and the Mediterranean, a fundamentally different set of population dynamics has likely governed the weed's emergence and

Received 9 Jan. 2007 Accepted 15 Mar. 2007

Supported by the National Science Foundation Plant Genome Research Program (DBI-0638820).

Publication of this paper is supported by the National Natural Science Foundation of China (30624808).

Tel: +1 314 935 7013;

Fax: +1 314 935 4432;

E-mail: <kolsen@wustl.edu>.

© 2007 Institute of Botany, the Chinese Academy of Sciences

doi: 10.1111/j.1672-9072.2007.00506.x

aggressive spread in rice fields. Our discussion of weedy rice in the present paper focuses on the latter situation, and specifically on the weedy rice found in US rice fields.

The US ranks third among rice-exporting nations (Childs 2005) and US rice production is among the most efficient in the world (Moldenhauer et al. 2004). The major region of rice cultivation in the US is in the southern Mississippi floodplain, an area that includes Arkansas, Louisiana, Mississippi, southern Missouri, and eastern Texas; rice cultivated in this region is exclusively of the *tropical japonica* variety. A secondary region of rice cultivation occurs in California, where mostly *temperate japonica* varieties are grown.

The US red rice is most problematic in the Mississippi floodplain, where it reduces crop yields by up to 80% (Estorninos et al. 2005) and contaminates harvests with its small, dark-pericarp seed. An estimated 30% of rice production areas in this region are infested with red rice, leading to an annual loss exceeding US\$ 50 million (Gealy 2005). Because it is the same species as cultivated rice, red rice is difficult to control by herbicide application; moreover, crop traits such as targeted herbicide resistance can escape into weed populations by gene flow (Rajguru et al. 2005). Red rice was first recorded as a weed in the southern US in 1846, and studies conducted around 1850 reported four morphologically distinct weedy forms (Craigmiles 1978). Weed infestations of southern US rice fields became common by the 1940s, after which a shift to weed-free seed stocks and changes in cultivation practices reduced the proliferation of the weed to some extent (Craigmiles 1978). Nonetheless, red rice infestations remain a primary constraint on US rice productivity (Estorninos et al. 2005; Gealy 2005).

Phenotypic and Genetic Diversity of US Red Rice

The US red rice shows a very wide range of phenotypic variation: strains range from crop-like plants, which share many domestication traits with cultivated rice (e.g. short stature, lack of seed shattering, early and synchronous flowering, straw-colored hulls, lack of awns), to strains closely resembling Asian wild *Oryza* species (e.g. tall stature, presence of seed shattering, asynchronous flowering, dark hulls, presence of awns). These divergent forms occur in other world regions with weedy rice infestations (e.g. LingHwa and Morishima 1997; Suh et al. 1997; Federici et al. 2001; Arrieta-Espinoza et al. 2005), and they have been proposed to represent two adaptive weed strategies (Federici et al. 2001): (i) the "crop mimic" (characterized by low stature, low degree of shattering, and overall similarity to the crop), which persists via crypsis in rice fields and contamination of seed stocks; and (ii) the "progenitor-type", which aggressively overtops cultivated rice, lodges,

and scatters its seed into the soil, where the seed can remain dormant for 20 years or more.

Assessments of genetic diversity in US red rice have relied primarily on simple sequence repeat (SSR) markers (e.g. Estorninos et al. 2000; Vaughan et al. 2001; Gealy et al. 2002). These studies have consistently found that the weed is highly genetically diverse, with clear differentiation between crop germplasm and the weed, as well as among different weed strains. Many US red rice strains appear to be genetically closely related to *indica* rice based on SSR allele sharing (Vaughan et al. 2001; Gealy et al. 2004; Londo 2006; for a review, see Gealy et al. 2005). This is particularly true for strains characterized by straw-colored hulls and a lack of awns, two traits that are more typical of cultivated rice than wild *Oryza* species (the latter of which typically have black hulls and long awns).

Gealy et al. (2002) examined SSR diversity at 18 loci (distributed across all 12 chromosomes) and found clear differentiation among the following groups: (i) awnless, strawhull strains; (ii) awned, blackhull strains; (iii) known weed-crop hybrids (which fell out as intermediate between weed strains and the crop in multidimensional scaling plots); and (iv) cultivated rice. Vaughan et al. (2002) used a different panel of 18 SSR markers and also observed high genetic diversity in the weed; some strains clustered most closely to *indica* rice, others to *japonica* varieties, and still others to non-cultivated wild relatives (*O. rufipogon* and *nivara*).

A recent study by Londo (2006) has used 16 SSR loci together with DNA sequence data from the *p-VATPase* locus to explicitly examine genetic relationships between US red rice (a sample of 27 phenotypically diverse strains) and Asian wild/cultivated rice varieties. The two datasets in Londo's work present a consistent picture of the weed's origins. As has been found in previous SSR analyses, most strawhull, awnless red rice appears to be closely related to *indica* rice. In addition, Londo's data reveal a relationship between most blackhull, awned varieties and Asian cultivated rice of the *aus* variety group. The *aus* rice is cultivated in a fairly restricted geographical area in the northern Indian subcontinent (India, Bangladesh); it has recently been proposed to have been domesticated independently of *indica* and *japonica* rice varieties (Londo et al. 2006). The work by Londo (2006) further indicates that although most red rice strains examined do not show evidence of introgression from US crop germplasm, some weed strains do show evidence of such introgression, as well as past hybridization between *indica*-like and *aus*-like strains of the weed.

Several SSR loci have been identified that are diagnostic in distinguishing cultivated rice and red rice (e.g. Vaughan et al. 2001). A large-scale SSR genotyping project for red rice is currently under way at the USDA National Rice Research Center (Stuttgart, AR, USA). This project includes 30–35 SSR

markers and preliminary analyses again indicate clear differentiation among weed strains; patterns of differentiation are similar to those observed in the earlier SSR studies (D Gealy, pers. comm., 2006).

Several studies have examined genetic diversity in weedy rice occurring in regions outside the US (e.g. Cho et al. 1995; Suh et al. 1997; Federici et al. 2001; Ferrero et al. 2001). However, these studies may be less relevant for understanding US red rice genetic diversity, because the weed appears to be genetically quite distinct from one world region to another. For example, Ferrero et al. (2001) found that, in the Mediterranean region, where *japonica* rice is traditionally grown, red rice is most closely related to *japonica* rice, whereas in Brazil, where *indica* rice is typically grown, the weed is closely related to *indica* crop varieties.

Evolutionary Origins of US Red Rice

Various hypotheses have been proposed to account for the presence and diversity of US red rice, and these hypotheses lead to different predictions about the population genomic structure underlying weed phenotypes. Because no *Oryza* species are native in North America, one hypothesis is that red rice can arise through a process of “de-domestication”, in which weediness is incidentally selected for in feral crop volunteers (i.e. unharvested seed that can give rise to unmanaged plants in crop field; Bres-Patry et al. 2001; Lu and Snow 2005). Alternatively, the genetic similarity between many red rice strains and Asian taxa strongly suggests that the introduction of non-cultivated Asian germplasm has likely played a role in the weed’s origin. Following the origin of weed traits by either of these mechanisms, introgression via hybridization could shuffle weed-associated alleles into plants of any genomic background. Although both rice and red rice are predominantly selfing, out-crossing does occur at a low level; most estimates suggest crop-weed introgression occurs at rates of <1% (Gealy 2005), although in some cases levels may exceed 30% (Langevin et al. 1990; Gealy 2005) and may be significant at the phenotypic level after only a few growing seasons (Langevin et al. 1990). Thus, weed strains may well represent genomic mosaics of crop and wild germplasm. Introgression into red rice may increase the weed’s adaptability by introducing genetic diversity (Langevin et al. 1990), particularly in portions of the genome where introgressed genotypes are selectively favored in the weed.

Several characteristics of this study system make it an excellent model for examining the population genomic mechanisms underlying weed trait evolution. First, red rice is unique in that it belongs to a species with a completely sequenced and mapped genome (Goff et al. 2002; Yu et al. 2002). This makes it possible to target genomic regions interspersed across the genome for

DNA sequencing. The DNA sequence data themselves offer a key advantage over SSRs (or other polymerase chain reaction (PCR) fragment-based markers) in population genetic analyses, in that the vast majority of molecular population genetic theory was specifically developed for analyzing nucleotide variation. Population-based SNP (single nucleotide polymorphism) analysis also avoids ascertainment bias problems common to other molecular markers, such as differential targeting of non-coding genomic regions. In addition, SNPs can more easily be compared across datasets generated by different research groups than fragment-based molecular markers (where the homology of alleles and/or loci between laboratories is often less certain); this permits an incremental accumulation of data that can be pooled across datasets and continuously subjected to re-analysis.

Another advantage of this system is that cultivated rice in the southern US represents a very narrow, genetically well-characterized germplasm pool, with nearly all varieties originating from approximately 50 *tropical japonica* introductions made during the 20th century (Mackill and McKenzie 2003; Lu et al. 2005). Although *indica* varieties have served as gene donors for specific traits that now characterize southern US rice (e.g. semi-dwarfism, blast resistance), *indicas* have never been widely cultivated in the US (only one cultivar, Jasmine 85, has been grown in the US; McClung et al. 1998). Detailed pedigree information exists for nearly all US varieties (Mackill and McKenzie 2003), and genome-wide assessments of genetic diversity have been made for large portions of the US rice germplasm. This genetically narrow, well-characterized crop germplasm provides a well-defined genomic context for detecting introgression from US rice into the weed.

Population Genomic Characterization of US Red Rice

A complete understanding of the evolution of weedy rice and of the emergence of weedy invasive species in general requires an understanding of the origin of the weedy germplasm and how introgression and selection contribute to shaping the weed’s genomic and phenotypic diversity. The wealth of genetic and genomic information for cultivated rice provides an unparalleled opportunity to address the evolutionary genomics of its weedy relative. As part of a recently funded US National Science Foundation Plant Genome research project on red rice, we are currently conducting research to elucidate the evolutionary origins and molecular population genetic dynamics of weedy trait evolution in US red rice, as described below.

To determine the population genomic structure underlying US red rice, it is necessary to assess the patterns of genetic polymorphism across the weed genome and compare these with the *Oryza* species/varieties that are potential contributors

to the genome. To accomplish this goal, we are characterizing SNPs in the red rice genome as well as in representative samples of Asian cultivated rice varieties, US cultivated rice varieties, the wild ancestor of rice (*O. rufipogon*), and a sample of more distantly related *Oryza* species. The data are derived from 48 approximate 500-bp fragments, known as sequence tagged sites (STS). These 48 markers are part of a larger set of 111 STS fragments that has been used to successfully characterize populations of domesticated Asian rice and the wild progenitor *O. rufipogon* (AL Caicedo, 2007, unpubl. data). The original set of STS loci was chosen randomly from a set of 6 591 expressed sequence tags (ESTs; Wu et al. 2002). All fragments correspond to single-copy expressed genes and most contain both exon and intron sequence. Bayesian analysis of population structure (Pritchard et al. 2000) with the 48 fragments compared with the original set of 111 indicates that these 48 loci can adequately recapture the population structure of Asian rice and *O. rufipogon* (Figure 1).

Although our own analyses are focused specifically on US red rice, it is our hope that the genome-wide set of 48 diagnostic loci will become a shared tool in the rice community for comparative assessments of genomic SNP diversity. The data from the STS set provide an unbiased sample of sequence diversity across the genomes of the various *Oryza* varieties sampled, and they offer the benefits of SNPs over SSRs described above. Levels of genomic nucleotide diversity have been estimated previously for *O. sativa* Asian landraces and *O. rufipogon*, and were found to vary between 0.05 and 0.4% per site (AL Caicedo, 2007, unpubl. data).

Data from the set of 48 STS markers enable the following analyses of US red rice diversity: (i) quantification of nucleotide variation in US red rice and comparison of these data to previously characterized levels in cultivated and wild rice; (ii) determination of the population structure of US red rice, especially whether accessions form well-defined populations or represent admixture from different sources; and (iii) determination of which *Oryza* species/varieties have contributed most

to the genomes of US red rice and the role of de-domestication in the evolution of weed strains.

Bayesian approaches based on the optimization of Hardy-Weinberg equilibrium and linkage equilibrium within populations, such as *structure* (Pritchard et al. 2000), have proved useful for estimating numbers of populations and degrees of admixture in diverse sample sets (e.g. Rosenberg et al. 2002; Lu et al. 2005). This approach is particularly informative for addressing whether US red rice has had a single or multiple origins and which *Oryza* taxa have contributed to the genomes of red rice. Traditional clustering methods and *Fst* (standardized variance in allele frequency) analyses can also be used to assess structure and relationships among red rice and other *Oryza* accessions. In addition, recent theoretical advances, which take into account the contribution of inbreeding to deviations from Hardy-Weinberg equilibrium within populations and that allow for a joint estimation of out-crossing rate and population structure (Gao and Bustamante 2007), can be applied to estimate the degree of inbreeding in US red rice.

Genomic characterizations of diversity are also informative about evolutionary processes that affect the entire genome, including population bottlenecks, founder events, and other demographic fluctuations. Many of these processes can be identified by their effect on the polymorphism frequency spectrum (e.g. Tajima 1989a, 1989b; Fay and Wu 2000). Thus, it is possible to use polymorphism data from the STS loci to search for the signature of demographic processes shaping genome-wide diversity. Moreover, loci that deviate from genome-wide patterns in multilocus tests of neutrality (e.g. Hudson et al. 1987; Bustamante et al. 2002) may serve as indicators of particular genomic regions that have been targets of selection, either during the domestication process or subsequent evolution of weedy forms.

Conclusions

Current assessments of phenotypic and genetic diversity in US

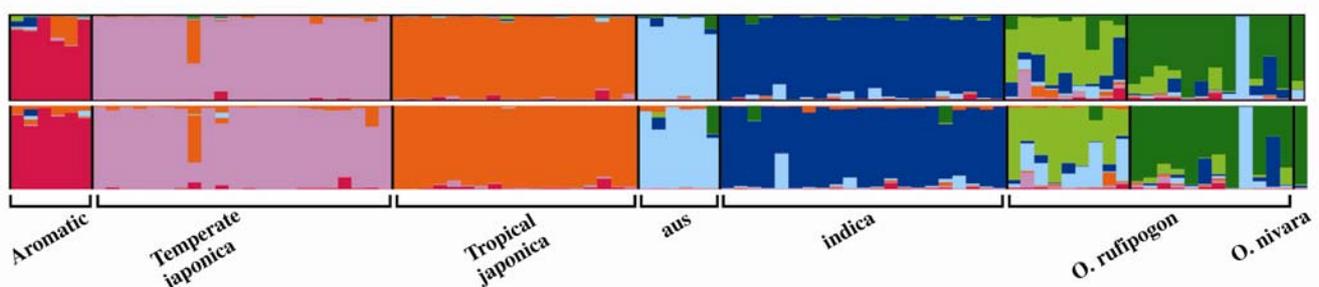


Figure 1. Estimated population structure for 97 accessions of *Oryza sativa* and *O. rufipogon* from 111 sequence tagged sites (STS) loci (top panel) and a subset consisting of 48 STS loci (bottom panel).

red rice present a picture of a genetically heterogeneous, evolutionarily dynamic system. The population genomic structure of this weed has likely been shaped by hybridization (involving both US crops and introduced Asian wild/weedy *Oryza* strains), selection for weed-adaptive traits (e.g. seed dormancy, dispersal mechanisms, competitiveness in agricultural fields), and demographic fluctuations. Our ongoing research on the genomic diversity of this weed will provide a fine-scale assessment of how these population genetic forces have interacted to generate the weed-adaptive phenotypes that have enabled red rice to compete so aggressively in US agricultural fields. The absence of native *Oryza* species in North America and the narrow germplasm pool of US rice varieties makes US red rice a particularly tractable system for examining these processes. The findings from this work will also offer a useful starting point for examining weedy rice evolution in the genetically more complex systems found in Asia.

Acknowledgements

The authors thank David R. Gealy of the Dale Bumpers National Rice Research Center for his insightful comments on US red rice natural history and genetic diversity.

References

- Arrieta-Espinoza G, Sanchez E, Vargas S, Lobo J, Quesada T, Espinoza AM (2005). The weedy rice complex in Costa Rica. I. Morphological study of relationships between commercial rice varieties, wild *Oryza* relatives and weedy types. *Genet. Res. Crop Evol.* **52**, 575–587.
- Bres-Patry C, Bangratz M, Ghesquiere A (2001). Genetic diversity and population dynamics of weedy rice in Camargue area. *Genet. Selection Evol.* **33**(Suppl. 1), S425–S440.
- Bustamante CD, Nielsen R, Sawyer SA, Olsen KM, Purugganan MD, Hartl DL (2002). The cost of inbreeding in *Arabidopsis*. *Nature* **416**, 531–534.
- Childs N (2005). Recap of the 2004/2005 U.S. Rice Market. In: Evans M, ed. *Rice Situation and Outlook Yearbook*. Economic Research Service, US Department of Agriculture, Published electronically (www.ers.usda.gov). pp. 43–73.
- Cho YC, Chung TY, Suh, HS (1995). Genetic characteristics of Korean weedy rice (*Oryza sativa* L.) by RFLP analysis. *Euphytica* **86**, 103–110.
- Craigmiles JP (1978). Introduction. In: Eastin EF, ed. *Rice Research and Control*. Texas Agricultural Experiment Station Bulletin. State of Texas Publications Press, College Station, TX. pp. 5–6.
- Ellstrand NC, Prentice HC, Hancock JF (1999). Gene flow and introgression from domesticated plants into their wild relatives. *Annu. Rev. Ecol. Syst.* **30**, 539–563.
- Estorninos LE, Gealy DR, Gbur EE, Talbert RE, McClelland MR (2005). Rice and red rice interference. II. Rice response to population densities of three red rice (*Oryza sativa*) ecotypes. *Weed Sci.* **53**, 683–689.
- Food and Agricultural Organization of the United Nations (2003). FAO statistical data (on-line). <http://faostat.fao.org>
- Fay JC, Wu CI (2000). Hitchhiking under positive Darwinian selection. *Genetics* **155**, 1405–1413.
- Federici MT, Vaughan D, Tomooka N, Kaga A, Wang XW, Doi K et al. (2001). Analysis of Uruguayan weedy rice genetic diversity using AFLP molecular markers. *Electronic J. Biotechnol.* **4**, 131–145.
- Ferrero A (2001). Biology and control of red rice (*Oryza sativa* L. var. *sylvatica*) infesting European fields. In: Chataigner J, ed. *Medoyszae-Interregional Cooperative Research Network on Rice in the Mediterranean Climate*, Issue 9. FAO-CIHEAM-Institut Agronomique Méditerranéen, Montpellier, France. pp. 2–4.
- Gao H, Bustamante CD (2007). Instruct: A Bayesian clustering approach for joint inference of population structure and selfing rates from multi-locus genotype data. http://www.intl-pag.org/15/abstracts/PAG15_W02_11.html.
- Gealy DR (2005). Gene movement between rice (*Oryza sativa*) and weedy rice (*Oryza sativa*: A US Temperate Rice Perspective). In: Gressel J, ed. *Crop Fertility and Volunteerism*. CRC Press, Boca Raton. pp. 323–354.
- Gealy DR, Tai TH, Sneller CH (2002). Identification of red rice, rice, and hybrid populations using microsatellite markers. *Weed Sci.* **50**, 333–339.
- Gealy DR, Yan W, Eizanga G, Moldenhauer K, Redus M (2004). *Insights into the Parentage of Rice/Red Rice Crosses Using SSR Analysis of US Rice Cultivars and Red Rice Populations*. Rice Technology Working Group 30. US Department of Agriculture, Beaumont, TX. pp. 179.
- Goff SA, Ricke D, Lan TH, Presting G, Wang RL, Dunn M et al. (2002). A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). *Science* **296**, 92–100.
- Hudson RR, Kreitman M, Aguade M (1987). A test of neutral molecular evolution based on nucleotide data. *Genetics* **116**, 153–159.
- Langevin, SA, Clay K, Grace JB (1990). The incidence and effects of hybridization between cultivated rice and its related weed red rice (*Oryza sativa* L.). *Evolution* **44**, 1000–1008.
- LingHwa T, Morishima H (1997). Genetic characterization of weedy rices and the inference on their origins. *Breeding Sci.* **47**, 153–160.
- Londo JP (2006). *Genetic and Physiological Variation in Cultivated and Wild Rice: Domestication, Phylogeography, and Aluminum Stress Resistance*. PhD thesis, Washington University, St Louis.
- Londo JP, Chiang YC, Hung KH, Chiang TY, Schaal BA (2006). Phylogeography of Asian wild rice, *Oryza rufipogon*, reveals

- multiple independent domestications of cultivated rice, *Oryza sativa*. *Proc. Natl. Acad. Sci. USA* **103**, 9578–9583.
- Lu BR, Snow AA** (2005). Gene flow from genetically modified rice and its environmental consequences. *Bioscience* **55**, 669–678.
- Lu H, Redus MA, Coburn JR, Rutger NA, McCouch SR, Tai TH** (2005). Population structure and breeding patterns of 145 US rice cultivars based on SSR marker analysis. *Crop Sci.* **45**, 66–76.
- Mackill DJ, McKenzie KS** (2003). Origin and characteristics of US rice cultivars. In: Smith CW, ed. *Rice: Origin, History, Technology, and Production*. John Wiley & Sons, Hoboken, NJ. pp. 87–100.
- McClung AM, Marchetti MA, Webb BD, Bollich CN** (1998). Registration of Jasmine 85 rice. *Crop Sci.* **38**, 896.
- Moldenhauer KAK, Gibbons JH, McKenzie KS** (2004). Rice varieties. In: Champagne ET, ed. *Rice: Chemistry and Technology*. American Association of Cereal Chemists, Inc., St Paul, MN. pp. 49–75.
- Pritchard JK, Stephens M, Donnelly P** (2000). Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–959.
- Rajguru SN, Burgos NR, Shivrain VK, Stewart JM** (2005). Mutations in the red rice ALS gene associated with resistance to imazethapyr. *Weed Sci.* **53**, 567–577.
- Rosenberg NA, Pritchard JK, Weber JL, Cann HM, Kidd KK, Zhivotovsky LA et al.** (2002). Genetic structure of human populations. *Science* **298**, 2381–2385.
- Suh HS, Sato YI, Morishima H** (1997). Genetic characterization of weedy rice (*Oryza sativa* L.) based on morpho-physiology, isozymes and RAPD markers. *Theor. Appl. Genet.* **94**, 316–321.
- Tajima F** (1989a). Statistical method for testing the neutral mutation hypothesis by DNA. *Genetics* **123**, 585–595.
- Tajima F** (1989b). The effect of change in population size on DNA polymorphism. *Genetics* **123**, 597–601.
- Vaughan LK, Ottis BV, Prazak-Havey AM, Bormans CA, Sneller C, Chandler JM et al.** (2001). Is all red rice found in commercial rice really *Oryza sativa*? *Weed Sci.* **49**, 468–476.
- Wu J, Maehara T, Shimokawa T, Yamamoto S, Harada C, Takazaki Y et al.** (2002). A comprehensive rice transcript map containing 6591 expressed sequence tagged sites. *Plant Cell* **14**, 525–535.
- Yu J, Hu SN, Wang J, Wong GKS, Li SG, Liu B et al.** (2002). A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). *Science* **296**, 79–92.

(Handling editor: Bin Han)