Theor Appl Genet (2002) 105:841–846
DOI 10.1007/s00122-002-1056-4

I. M. Mullins · K. W. Hilu

Sequence variation in the gene encoding the 10-kDa prolamin in *Oryza* (Poaceae). I. Phylogenetic Implications

Received: 21 August 2001 / Accepted: 18 February 2002 / Published online: 13 September 2002

Abstract *Oryza* L. (Poaceae) contains approximately 20 wild and two domesticated species and nine genomes. Major disagreements exist on its systematics and genome evolution. Sequence polymorphism in the gene that encodes the 10-kDa prolamin polypeptide (a seed storage protein) was used to determine phylogenetic relationships and evaluate current systematics for 19 *Oryza* species. This gene in *Oryza* is approximately 402-bp long, and includes a 72-bp signal peptide region. A strict consensus tree shows *Oryza brachyantha* (FF) as the most basal species, followed by a polytomy of three clades that can be delineated based on genome composition: (1) the GG clade: *Oryza granulata* and *Oryza meyeriana*, (2) the EE clade: *Oryza australiensis*, and (3) the ABCD clade: the remaining *Oryza* species. Two subclades within the ABCD clade emerge, one containing species with the AA genome, the other with components of the BC and D genomes. Members of the AA subclade form a polytomy and were delineated by a single 3-base deletion. The African species *Oryza punctata* (BB) and the South American-endemic CCDD genome species form a strong lineage, pointing to a close genetic affinity of *O. punctata* to the missing DD genome donor. The strong association between the CC and BBCC species implies convergence at the gene level. The study supports the following sectional units of *Oryza*: Section *Oryza* (Series sativae and officinaliae), Section *australiensis*, Section *Granulata*, Section *Brachyantha*.

Keywords *Oryza* · Poaceae · Prolamin · Phylogeny · Genomes

Communicated by P.L. Pfahler

I.M. Mullins · K.W. Hilu (✉)
Department of Biology,
Virginia Polytechnic Institute and State University,
Blacksburg, VA 24061, USA
e-mail: hilukw@vt.edu, Fax: +1-540-231-9307

Introduction

The genus *Oryza* (Poaceae, grasses) contains approximately 22 species and nine genomes (Aggarwal et al. 1997; reviewed in Vaughan 1994). *Oryza sativa* (Asian rice) and *Oryza glaberrima* (African rice) are cultivated species; the remaining twenty are wild. *O. sativa* is third only to wheat and maize in global economic production (World Almanac 1998). Major disagreements still exist on the systematics and genome evolution of this economically important genus. Since the first taxonomic treatment of *Oryza* by Roshevicz (1931), up to six sections (Oka 1988) and eight series (Sharma and Shastry 1965) have been recognized. The two recent taxonomic treatments of *Oryza* (Vaughan 1994; Lu 1999) differ at the sectional and intersectional levels. The composition of section *Oryza* differs in the treatment of *Oryza australiensis*. Species within the sections *Ridleyanae* and *Granulata* in Vaughan's system were merged under section *Padia* by Lu (1999). Unlike Vaughan (1994), the latter author raised *Oryza brachyantha* to a sectional level, erecting the monotypic section *Brachyantha*. *Oryza schelecteri* was also recognized as a distinct series (Schlechterianae) under *Padia*. Vaughan (1994) placed *Oryza schlechteri* and *O. brachyantha* as unclassified species in section *Ridleyanae*.

Not only is the systematics of the genus disputed, but also the evolution of its genomes is not well understood. Currently, a comprehensive hypothesis regarding the *Oryza* genome evolution is unavailable. The objectives of this study were to use sequence polymorphism in the 10-kDa prolamin gene (a seed storage protein) to determine the phylogenetic relationships within *Oryza*, provide insight into the evolution of its genomes and evaluate current taxonomic treatments.

The prolamin is a class of alcohol-soluble seed storage proteins unique to the grass family (Shewry et al. 1995). Molecular weights of prolamin range from 10 to about 100 kDa (Hilu 2000). *Oryza* contains low-molecular-weight prolamins, which are organized into three size classes, 10, 13 and 16 kDa (Barbier and Ishihama 1990), that are encoded by three nuclear multi-gene families (Kim and Okita 1988). The gene that encodes the 10-kDa prolamin polypeptide is 402 base pairs (bp) in length (Masumura et al. 1989), has no intron (Barbier and Ishihama 1990), and contains a 72-bp signal-peptide region (Masumura et al. 1989). The gene exists in multiple copies, perhaps as many as 80–100 copies per haploid genome (Kim and Okita 1988), which codes for a polypeptide consisting of 134 amino acids that lacks major repetitive sequences (Masumura et al. 1989). Barbier and Ishihama (1990) observed little nucleotide variation among *Oryza rufipogon* strains or between *O. rufipogon* and *Oryza longistaminata*, indicating homogeneity at the locus level. Examination of nucleotide sequences of the 10-kDa prolamin gene (Hilu and Sharova 1998, 2002) in two *Oryza* species and in *Phyllostachys aurea* (Bambusoideae) demonstrated systematic utility at and above the species level.

Materials and methods

Seed acquisition and plant growth

Nineteen *Oryza* species, represented by 26 accessions, were examined (Table 1). Grains were germinated on moist filter paper in Petri plates and seedlings were transferred to individual pots containing a clay/organic soil mixture in a greenhouse. Voucher specimens are located at International Rice Research Institute (IRRI, The Philippines) and the Virginia Tech Massey Herbarium (VT). All IRRI seed lots used in this project have been positively identified by Dr. B.R. Lu, an IRRI germplasm specialist.

Genomic DNA extraction

DNA was isolated from leaf material with the CTAB extraction method following Hilu (1994). Genomic DNA was amplified by PCR using *Taq* polymerase (Promega). Two sets of primers were used to amplify the 10-kDa prolamin gene. In most cases, the primers PR10.1F2 (5′ ACG TGA ATT CCA CCA TCT GGA ATC TGG 3′) and PR10.3RV (5′ ACG TTC TAG AG TGT TTG CAC ACG ATA GTA 3′) (Fig. 1) were employed. When these two primers did not amplify, PR10.1E (5′ ACG TGA ATT CAT GGC AGC ATA CAC CAG CAA G 3′) and PR10.2RB (5′ ACG TGG ATC CAA CCA CAG GAA GAG AGT TGG 3′) primers were used (Hilu and Sharova 2002). Primers were designed based on the Masumura et al. (1989) nucleotide sequence. The forward primer PR10.1E and reverse primer PR10.2RB are located within the conserved coding region (Fig. 1). PCR amplification conditions used for this set of primers were 40 cycles of: 0.75 min, 94 °C (denaturation); 1 min, 40–50 °C (annealing); and 1 min, 72 °C (extension). The forward primer PR10.1F2 and the reverse primer PR10.3RV are located outside the coding region of the 10-kDa prolamin gene (Fig. 1), and therefore provide complete sequence information for the coding region and the signal peptide. The amplification conditions in this case were 40 cycles of: 0.5 min, 94 °C (denaturation);1.5 min, 48–50 °C (annealing); and 1 min, 72 °C (extension). PCR reactions were electrophoresed in a 1.5% agarose gel to isolate the individual bands. The appropriate bands were cut out and the DNA cleaned using the Quiagen Gel Extraction Kit (Quiagen Inc., Valencia, Calif.). Sequencing reactions were carried out directly on the clean PCR products using two different ABI PrismTM Dye Terminator Cycle Sequencing Kits (Perkin Elmer, Norwalk, Conn.). Samples were electrophoresed in an ABI 373A automated sequencer or in an ABI 310 Genetic Analyzer (Applied Biosystems, Inc., Foster City, Calif.).

Sequence data analysis

DNA sequences were manually edited and aligned using Sequence Navigator 1.0 (PE Biosystems Inc., Foster City, Calif.). Both for-

Table 1 Description of *Oryza* accessions examined. IRRI = International Rice Research Institute; PI = Plant Introduction Number of U. S. Department of Agriculture; KH = Khidir Hilu; W = DNA obtained from Dr. Y. Sano, Hokkaido University, Japan; B = Plant material obtained from the University of Bonn, Germany

Taxon	Genome	KH no.	Accession no.	Geographic origin	Prime pair used for sequencing
O. alta	CCDD	KH 7016	IRRI 101395	Unknown	PR10.1F2/PR10.3RV
O. australiensis	EE	–	W0008	N. Australia	PR10.1E/PR10.3RV
O. australinesis	EE	KH 7043	IRRI 101144	Australia	PR10.1E/PR10 2RB
O. barthii	AA	KH 7055	IRRI 100933	Sudan	PR10.1F2/PR10.3RV
O. brachyantha	FF	KH 7024	IRRI 105171	Cameroon	PR10.1E/PR10.2RB
O. brachyantha	FF	KH 7025	IRRI 105172	Cameroon	PR10.1F2/PR10.3RV
O. eichingeri	CC	–	W1522	Uganda	PR10.1F2/PR10.3RV
O. glaberrima	AA	KH 7004	PI 450198	Nigeria	PR10.1F2/PR10.3RV
O. grandiglumis	CCDD	KH 7017	IRRI 101405	Brazil	PR10.1F2/PR10.3RV
O. granulata	GG	KH 7029	IRRI106468	Laos	PR10.1E/PR10.2RB
O. latifolia	CCDD	KH 7040	IRRI 100165	Guatemala	PR10.1E/PR10.2RB
O. meridionalis	AA	KH 7027	IRRI 105279	Australia	PR10.1F2/PR10.3RV
O. meridionalis	AA	KH 7015	IRRI 101145	Australia	PR10.1F2/PR10.3RV
O. meyeriana	GG	KH 7030	IRRI 106474	Philippines	PR10.1F2/PR10.3RV
O. minuta	BBCC	KH 7014	IRRI 101128	Philippines	PR10.1F2/PR10.3RV
O. minuta	BBCC	KH 7013	IRRI 101082	Philippines	PR10.1F2/PR10.3RV
O. nivara	AA	KH 7035	IRRI 101524	India	PR10.1F2/PR10.3RV
O. officinalis	CC	–	W0002	Thailand	PR10.1F2/PR10.3RV
O. punctata	BB	–	W1514	Kenya	PR10.1F2/P10.3RV
O. rhizomatis	CC	KH 7020	IRRI 103421	Sri Lanka	PR10.1F2/PR10.3RV
O. rhizomatis	CC	KH 7019	IRRI 103410	Sri Lanka	PR10.1E/PR10.2RB
O. rufipogon	AA	KH 7031	IRRI 100907	Taiwan	PR10.1F2/PR10.3RV
O. rufipogon	AA	KH 7032	PI 590418	Myanmar	PR10.1F2/PR10.3RV
O. rufipogon	AA	KH 7037	PI 239671	India	PR10.1F2/PR10.3RV
O. sativa	AA	KH 7000	–	United States	PR10.1F2/PR10.3RV

Fig. 1 An illustration of the gene encoding the 10-kDa prolamin seed storage protein. The position of primers, the signal peptide, and the coding region are indicated

ward and reverse sequences, which usually show 100% overlap, were examined for the presence of polymorphic signals. *Phyllostachys aurea* (Hilu and Sharova 1998) was used as an outgroup. Candidate outgroup taxa originally included *Zizania aquatica*, *Leersia virginica* and *HygrOryza aristata*; however direct sequencing of PCR products for these species yielded multiple genetic types in each taxon. PCR products for the species were cloned but none of the clones yielded sequences alignable with the *Oryza* sequences.

Edited sequences were phylogenetically analyzed using Fitch parsimony in PAUP* 4.0b3a (Swofford 1998). Heuristic searches were performed with 1,000 random stepwise-addition replicates, "MulTrees" on, and the TBR (tree-bisection-reconnection) branch-swapping algorithm for two data sets: one that included the signal peptide and the mature peptide regions of the gene, and another that included the mature peptide region only. Gaps were treated as missing data except for one unique ATG indel that was manually coded as a binary character. A strict consensus tree was constructed from the equally parsimonious trees. G_1 values were created for 1,000 random trees using the "evaluate random trees" option in PAUP* 4.0b3a, and compared to the values given for 25 taxa with 250 characters (variable positions) with a P value of 0.01 (Hillis and Huelsenbeck 1992) to test for the presence of non-random structure in the data set. Decay indices (Bremmer 1988; Donaghue et al. 1992) and bootstrap values (Felsenstein 1985) for 100 replicates were calculated to measure support for individual clades. AutoDecay (Eriksson 1998) was used to perform the decay analysis. Deduced amino acids were determined using Lasergene Navigator (DNASTAR Inc., Madison, Wis.). The open reading frame for each of the sequences was determined to exclude the signal peptide region and then translated using the standard genetic code.

Results

The 10-kDa prolamin gene in *Oryza* consists of approximately 402 bases, which includes a 72-bp signal peptide region (Fig. 1). The open reading frame (ORF) for all species that contain the AA genome is 330 bases. All remaining *Oryza* species contain an ORF of 333 bases. The ORF terminated with TGA for all species except for *Oryza latifolia* (GGA), *O. australiensis* 7043 (GGA) and *O. meyeriana* (GGA). The sequence for *O. granulata* terminated approximately 24 bases prior to the 3′ end of the remaining *Oryza* sequences. A three base-pair deletion was synapomorphic to all AA genome sequences (Fig. 2). Two 9-bp gaps were inserted into the *Oryza* sequences to align them to *P. aurea* (Bambusoideae). A 12-bp gap was detected in *O. meyeriana*, and a 5-bp indel ten bases from the 5′ end of the gene for *O. latifolia* and *O. brachyantha* (KH7024) was present.

Phylogenetic analysis based on the data set that excluded the signal peptide region showed that the first, second and third codon positions provided 13 (26%), 13 (26%), and 24 (48%) of the 50 parsimony informative characters, respectively. These data indicate that 26 of the 50, or 52%, of the parsimony informative characters are candidate positions for amino-acid substitutions.

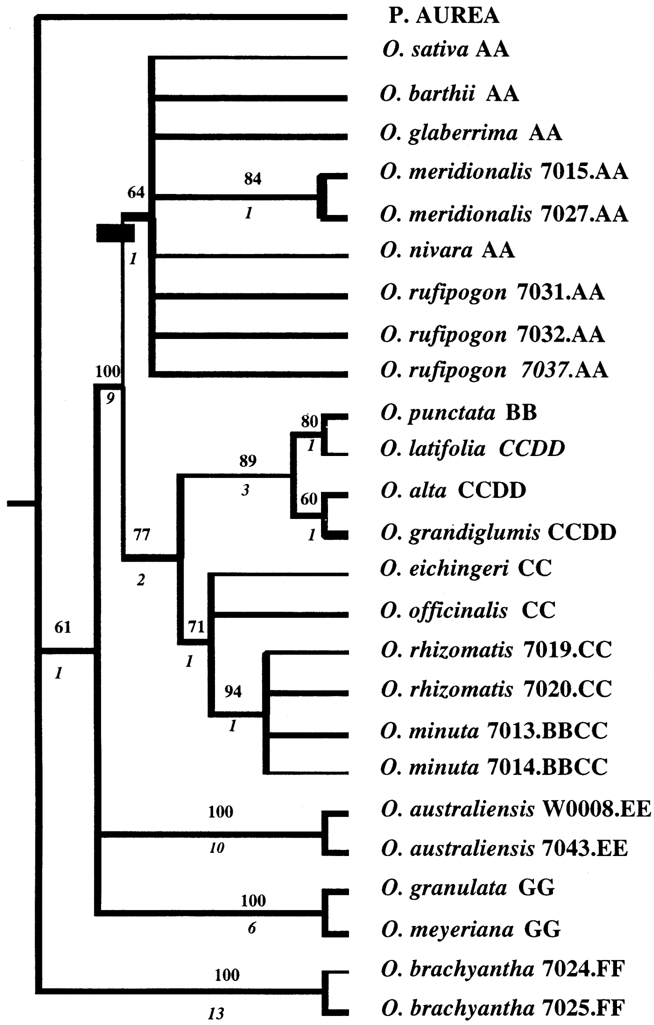

Fig. 2 Strict consensus tree of *Oryza* based on sequences from the signal peptide and coding region of the gene encoding the 10-kDa prolamin. Bootstrap values are indicated above branches and decay values are below branches. The *vertical bar* denotes a single indel event. CI = 0.864, RI = 0.867, g_1 = –1.123, tree length = 173 steps

The open reading frame of the gene was translated into its deduced amino acids. Residues with the greatest variability among species include: glutamine, leucine, methionine, threonine, asparagine, cysteine and serine. The *Oryza* species that exhibit the greatest variability in amino-acid composition relative to *O. sativa* (cultivated rice) are: *Oryza punctata* [7 substitutions (s)], *Oryza minuta* (7s), *Oryza rhizomatis* (7s), *Oryza alta* (6s), *O. latifolia* (8s), and *Oryza grandiglumis* (5s). Those taxa are members of the South American-endemic CCDD/BB and the BBCC/CC subclades.

The 26 aligned sequences for the mature peptide and the signal peptide contained 430 characters, of which 130 (30%) were variable. Of the 130 variable characters, 69 (53%) were parsimony informative. The –1.123 g_1 value of the *Oryza* data set provides significant ($P < 0.01$) evidence of non-random structure. The cladistic analysis yielded 24 most-parsimonious trees that were 173 steps in length. The Consistency Index (CI) and Retention Index (RI) were 0.864 and 0.867 respectively. The predicted CI based on the polynomial regression analysis of Sanderson and Donoghue (1989) is 0.3279, which indicates that there are low levels of homoplasy in this data set.

The exclusion of the signal peptide region resulted in increased polytomies at the base and internal nodes of the phylogeny. Otherwise, the two phylogenies are congruent in topology. Consequently, the rest of the paper will focus on the data set that included the signal peptide region (Fig. 2). *O. brachyantha* (FF) is sister to all *Oryza* species examined. The remaining species formed a lineage with relatively low support (bootstrap 61%, decay 1). In this lineage, a polytomy is formed representing three clades that can be delineated based on the genome composition of their respective species: (1) the GG clade encompassing *O. granulata* and *O. meyeriana* (bootstrap 100%, decay 6), (2) the EE clade of *O. australiensis* (bootstrap 100%, decay 10), and (3) the ABCD clade (bootstrap 100%, decay 9). Two subclades emerge within the ABCD clade representing the AA and BCD genome species. Members of the AA subclade formed a polytomy. The BCD clade consisted of two lineages. One strongly supported lineage (bootstrap 89%, decay 3) contains the African *O. punctata* (BB) and the South American-endemic *O. latifolia*, *Oryza alta*, and *Oryza grandiglumis* (all CCDD). The other encompasses members of the CC genome (*Oryza eichingeri*, *Oryza officinalis* and *Oryza rhizomatis*) and the BBCC genome (*Oryza minuta*) groups.

Discussion

The gene encoding the 10-kDa prolamin appears homogeneous at the locus level in the *Oryza* species examined, as indicated by the lack of multiple genetic species in the pooled PCR products used for direct sequencing. This result provides yet another support for concerted evolution operating on tandem repeat units. The length of the gene and the signal peptide region are in agreement with previous reports (Masumura et al. 1989). The two 9-bp deletions required to align *Oryza* with *P. aurea* may represent a molecular marker for the genus. However, the gene has to be sequenced from other members of the Ehrhartoideae to confirm this possibility.

Exclusion of the signal peptide regions changed the percent substitution only slightly (27.2%), indicating similar rates of substitutions in both regions, and that the signal peptide domain contains phylogenetically useful characters. The relatively high percentages of substitutions in the first and second codon positions indicate high variability at the amino-acid level because these codon positions typically lead to nonsynonomous substitutions. Theoretically, the first, second and third codon positions are translated to 96, 100 and 31% nonsynonomous substitutions, respectively (Li and Graur 1991).

Species and genome evolution in *Oryza*

The vast majority of the studies that focused on the interspecific relationships within *Oryza* has used either phenetic approaches or was limited in species representation. Ge et al. (1999) presented the most comprehensive molecular phylogenetic study. However, discordance among their single gene phylogenies prompted them to underscore the need for the use of other genes. This study sheds light on two important issues related to *Oryza*: (1) species phylogeny and genomic relationship, and (2) taxonomy of the genus. The parsimony analysis depicts *O. brachyantha* (FF) as the most basal species and sister to a weakly supported clade (bootstrap 61%, decay 1)that contain the rest of *Oryza* (Fig. 2). This finding has important genomic and geographic implications, pointing to the FF genome as the ancestral type and suggesting an African origin for *Oryza* as opposed to the proposed Eurasian origin (Second 1985). It has to be noted that outgroup choice may influence tree topology, and that although *Phyllostachys* and *Oryza* belong to phylogenetically related subfamilies (Hilu et al. 1999), the former is not as closely related to *Oryza* as other oryzoid genera. Ge et al. (1999) *matK*- and *Adh*-based phylogenies depicted *O. meyeriana* and *O. granulata* (GG) as the most basal taxa; however, statistical support was very weak (boostrap <50 in some cases). The position of *O. brachyantha* was inconsistent among their three phylogenies. The appearance in this study of *O. brachyantha* as a distinct clade supports a sectional level treatment as proposed by Lu (1999). Vaughan's (1994) proposed relationship of *O. brachyantha* to *O. schlechteri* could not be examined because of low homology of the *O. schlechteri* sequence to other *Oryza* species.

Members of the GG genome species *O. granulata* and *O. meyeriana* and the EE genome species *O. australiensis* form two distinct and strongly supported clades (Fig. 2). Although their affinities to the ABCD clade are not resolved because of the polytomy, neither one of these two clades can be considered as a potential member of the ABCD clade because of the strong statistical support of the latter clade (bootstrap 100%, decay 9).Therefore, this phylogeny supports Vaughan's treatment of the GG genome species in the section *Granulata*. The possibility of *O. australiensis* being a sister to the ABCD lineage should not be excluded. Other studies have shown that the association between *O. australiensis* and members of the section *Oryza* is usually very weak (Dally and Second 1990; McIntyre et al. 1992; Wang et al. 1992).

The ABCD clade, except for the exclusion of *O. australiensis*, represents what has been designated as the section *Oryza* by Vaughan (1994) and Lu (1999). The

AA genome species correspond to the *O. sativa* complex of Vaughan (1994) and the Series Sativae of Lu (1999). It should be noted, however, that the AA genome clade is supported by only a synapomorphic ATG deletion. Lack of nucleotide substitutions in this prolamin gene for the AA species is in agreement with the results of Barbier and Ishihama (1990), and could imply a recent origin and a subsequent rapid radiation of those species. This phenomenon is also reflected at the morphological level because delimiting boundaries among the AA genome species is a prominent problem (Lu 1999). Vaughan (1994) contended that hybrids can occur between the AA-genome species resulting in a continuum of morphological types. The subclade containing *O. alta*, *O. eichingeri*, *O. grandiglumis*, *O. latifolia*, *O. minuta*, *O. officinalis*, *Oryza punctata* and *O. rhizomatis* corresponds to the *O. officinalis* complex of Vaughan (1994) and the Series Latifoliae of Lu (1999). The emergence of *O. alta*, *O. grandiglumis* and *O. latifolia* in a well-supported clade, is in agreement with the RFLP results of Jena and Kochert (1991).

The composition of the two lineages comprising the BCD subclade is particularly important (Fig. 2). The association of the BB genome African *O. punctata* and the CCDD genome South American species *O. latifolia*, *O. alta* and *O. grandiglumis* points to a close genetic affinity between these two geographically distinct groups. These data indicate that *O. punctata* could be related to the yet-unidentified DD genome donor of the CCDD genome species. Dally and Second (1990) stated that chloroplast restriction analysis and isozyme data do not indicate a direct relationship between the CD genome species and other diploid species. The subclade comprising *O. minuta* (BBCC) and the CC genome species *O. eichingeri*, *O. officinalis* and *O. rhizomatis* has strong support (bootstrap 94%, decay 3). This association may imply convergence at the gene level.

In conclusion, this study, along with that of Ge et al. (2000), clearly underscores the genetic affinities between the AA, BB, CC and DD genomes, with the AA genome species being most recently derived. Our study points to the BB genome species *O. punctata* as closely allied to the missing DD genome donor. The10-kDa prolamin gene phylogeny depicts FF as a likely ancestral genome.

It is evident from our study that the section *Oryza* is well supported. Within this section, the recognition of two series, Sativa (AA) and Latifoliae (BCD), is substantiated. This treatment is in line with those of Sharma and Shastry (1972), Vaughan (1994) and Lu (1999), and is supported by the studies of Morishima and Oka (1960), McIntyre et al. (1992), Wang et al. (1992) and Ge et al. (2000). Strong support for raising*O. brachyantha* (FF) to the sectional level as proposed by Lu (1999) is also demonstrated here. This treatment gains support from phylogenies based on the *Adh1* and *matK* data (Ge et al. 2000). The taxonomic position of *O. australiensis* (EE) remains disputable, but this study argues for its exclusion from the section *Oryza* (Vaughan 1994; Lu 1999) and treatment at the sectional rank.

Acknowledgements Seed materials were kindly provided by IRRI and the U.S. Department of Agriculture, National Plant Germplasm System. Thanks to Drs. Y. Sano, Hokaido University, Japan, and C. Neihous, University of Bonn, for supplying plant or DNA samples, and to D. Wiley for maintaining the *Oryza* greenhouse collections. This work was supported by grants from the Virginia Academy of Science, Sigma *Xi*, the Graduate Research and Development Project of Virginia Tech, and the Jeffress Foundation. The experiments comply with current laws of the U.S.A.

References

Aggarwal K, Brar D, Khush G (1997) Two new genomes in the *Oryza* complex identified on the basis of molecular divergence analysis using genomic DNA hybridization. Mol Gen Genet 254:1–12

Barbier P, Ishihama A (1990) Variation in the nucleotide sequence of a prolamin gene family in wild rice. Plant Mol Biol 15: 191–195

Bremmer K (1988) The limits of amino-acid sequence data in angiosperm phylogenetic reconstruction. Evolution 42:795–803

Dally AM, Second G (1990) Chloroplast DNA diversity in wild and cultivated species of rice (genus *Oryza*, section *Oryza*). Cladistic mutation and genetic distance analysis. Theor Appl Genet 80:209–222

Donoghue MJ, Olmstead RG, Smith JF, Palmer JD (1992) Phylogenetic relationships of Dipsacales based on *rbcL* sequences. Ann Missouri Bot Gard 79:333–345

Eriksson T (1998) AutoDecay ver. 4.0 (program distributed by the author). Department of Botany, Stockholm University, Stockholm

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the boostrap. Evolution 39:783–791

Ge S, Sang T, Lu B-R, Hong D-Y (1999) Phylogeny of rice genomes with emphasis on origins of allotetraploid species. Proc Natl Acad Sci USA 96:14,400–14,405

Hillis DM, Huelsenbeck JP (1992) Signal, noise, and reliability in molecular phylogenetic analyses. J Hered 83:189–195

Hilu KW (1994) Evidence from RAPD markers in the evolution of *Echinochloa* millets (Poaceae). Plant Syst Evol 189:247–257

Hilu KW (2000) Contributions of prolamin size diversity and structure to the sytematics of the Poaceae. In: Jacobs SWL, Everett J (eds) Grass systematics and evolution 2. CSIRO Press, Melbourne, pp 241–247

Hilu KW, Sharova LV (1998) Characterization of 10-kDa prolamin genes in *Phyllostachys aurea* (Bambusoideae, Poaceae). Am J Bot 85:1033–1037

Hilu KW, Sharova LV (2002) Evolutionary implications of substitution patterns in prolamin genes of *Oryza glaberrima* (African rice, Poaceae) and related species. Am J Bot (in press)

Hilu KW, Alice LA, Liang H (1999) Phylogeny of Poaceae inferred from *matK* sequences. Ann Missouri Bot Gard 86: 835–851

Jena KK, Kochert G (1991) Restriction length polymorphism analysis of CCDD genome species of the genus *Oryza* L. Plant Mol Biol 5:109–118

Kim WT, Okita TW (1988) Structure, expression, and heterogeneity of the rice seed prolamines. Plant Physiol 88:649–655

Li WH, Graur D (1992) Fundamentals of molecular evolution. Sinauer, Sunderland, Massachusetts

Lu BS (1999) Taxonomy of the genus *Oryza* (Poaceae): historical perspective and current status. Int Rice Res Notes (IRRI) 24, 2:4–8

Masumura T, Shibata D, Hibino T, Kato T, Kawabe K, Takeba G, Tabaka K, Fuji S (1989) cDNA cloning of an mRNA encoding a sulfur-rich 10-kDa prolamin polypeptide in rice seeds. Plant Mol Biol 12:123–130

McIntyre CL, Winberg B, Houchins K, Appels R, Baum BR (1992) Relationships between *Oryza* species (Poaceae) based on 5S DNA sequences. Plant Syst Evol 183:249–264

Morishima H, Oka HI (1960) The pattern of interspecific variation in the genus *Oryza*: its quantitative representation by statistical methods. Evolution 14:13–165
Oka HI (1988) Origin of culitvated rice. Japanese Science Society Press, Tokyo
Roschevicz RJ (1931) A contribution to the knowledge of rice. Bull Bot Genet Plant Breed 27:3–133
Sanderson MJ, Donoghue MJ (1989) Patterns of variation in levels of homoplasy. Evolution 8:1781–1795
Second G (1985) Evolutionary relationships in the Sativa group of *Oryza* based on isozyme data. Génét Sél. Evol 17:89–114
Sharma SD, Shastry SVS (1965) Taxonomic studies in genus *Oryza* L. III. *O. rufipogon* Giff. *sensu stricto* and *O. nivara* Sharma et Shastry *nom. Nov*. Ind J Genet Plant Breed 25:157–167
Shewry PR, Napier JA, Tatham AS (1995) Seed storage proteins: Structures and biosynthesis. Plant Cell 7:945–956
Swofford DL (1998) PAUP: Phylogenetic Analysis Using Parsimony. Version 4.0b2a Sinaur Assoc. Sunderland, Massachusetts
Vaughan D (1994) The wild relatives of rice: a genetic resources handbook. International Rice Research Institute, Los Baños, The Philippines, pp 1–101
Wang Y, Second G, Tanksley SD (1992) Polymorphism and phylogenetic relationships among species of the genus *Oryza* as determined by analysis of nuclear RFLPs. Theor Appl Genet 83:565–581
World Almanac and Book of Facts (1998) Current population and projections for all countries: 1997, 2020 and 2050