Theor Appl Genet (1992) 83:565–581



# Polymorphism and phylogenetic relationships among species in the genus *Oryza* as determined by analysis of nuclear RFLPs

**Z. Y. Wang, G. Second, and S. D. Tanksley****

Department of Plant Breeding and Biometry, 252 Emerson Hall, Cornell University, Ithaca NY 14853, USA

Received April 15, 1991; Accepted July 18, 1991
Communicated by F. Salamini

**Summary.** Ninety-three accessions representing 21 species from the genus *Oryza* were examined for restriction fragment length polymorphism. The majority (78%) of the accessions, for which five individuals were tested, were found to be monomorphic. Most of the polymorphic accessions segregated for only one or two probes and appeared to be mixed pure lines. For most of the *Oryza* species tested, the majority of the genetic variation (83%) was found between accessions from different species with only 17% between accessions within species. Tetraploid species were found to have, on average, nearly 50% more alleles (unique fragments) per individual than diploid species reflecting the allopolyploid nature of their genomes.

Classification of *Oryza* species based on RFLPs matches remarkably well previous classifications based on morphology, hybridization and isozymes. In the current study, four species complexes could be identified corresponding to those proposed by Vaughan (1989): the *O. ridleyi* complex, the *O. meyeriana* complex, the *O. officinalis* complex and the *O. sativa* complex. Within the *O. sativa* complex, accessions of *O. rufipogon* from Asia (including *O. nivara*) and perennial forms of *O. rufipogon* from Australia clustered together with accessions of cultivated rice *O. sativa*. Surprisingly, indica and japonica (the two major subspecies of cultivated rice) showed closer affinity with different accessions of wild *O. rufipogon* than to each other, supporting a hypothesis of independent domestication events for these two types of rice. Australian annual wild rice *O. meridionalis* (previously classified as *O. rufipogon*) was clearly distinct from all other *O. rufipogon* accessions supporting its recent reclassification as *O. meridionalis* (Ng et al. 1981).

Using genetic relatedness as a criterion, it was possible to identify the closest living diploid relatives of the currently known tetraploid rice species. Results from these analyses suggest that BBCC tetraploids (*O. malampuzhaensis*, *O. punctata* and *O. minuta*) are either of independent origins or have experienced introgression from sympatric C-genome diploid rice species. CCDD tetraploid species from America (*O. latifolia*, *O. alta* and *O. grandiglumis*) may be of ancient origin since they show a closer affinity to each other than to any known diploid species. Their closest living diploid relatives belong to C genome (*O. eichingeri*) and E genome (*O. australiensis*) species. Comparisons among African, Australian and Asian rice species suggest that *Oryza* species in Africa and Australia are of polyphyletic origin and probably migrated to these regions at different times in the past.

Finally, on a practical note, the majority of probes used in this study detected polymorphism between cultivated rice and its wild relatives. Hence, RFLP markers and maps based on such markers are likely to be very useful in monitoring and aiding introgression of genes from wild rice into modern cultivars.

**Key words:** *Oryza* – Rice – Restriction fragment length polymorphism – Phylogeny

## Introduction

*Oryza* is an agronomically important genus containing species with highly diverse morphology (Clayton and Renvoize 1986). Included in this genus is cultivated rice (*O. sativa*), which constitutes an important part of the

* *Current address:* International Rice Research Institute, P.O. Box 933, Manila, Los Baños, The Philippines
** To whom correspondence should be addressed

diet of more than half of the world's population (Coffman and Herrera 1980). While the other species in this genus have not been domesticated (with the exception of *O. glaberrima*), they contain genes for many valuable traits including those for resistance to diseases, insects and stress tolerances (Chang et al. 1975; Sitch et al. 1989). Currently, only a small portion of these valuable genes have been introgressed into cultivated rice (Jena and Khush 1990), but they have had a major impact on rice production (Sharma 1983). The development of wide-hybridization techniques coupled with a recently developed RFLP map of rice makes it likely that there will be much more use of wild *Oryza* species in rice breeding in the future (Jena and Khush 1990; McCouch et al. 1988; Sitch et al. 1989; Tanksley et al. 1989).

The genus *Oryza* belongs to the tribe *Oryzeae* of subfamily *Oryzoideae* in the family *Poaceae* (*Gramineae*). In a recent review, Vaughan (1989) summarizes the taxonomy of genus *Oryza*. The genus can be divided into four species complexes as well as into two other discrete species. One of these latter two species, *O. schlechteri*, was known only from herbarium specimens but has been recently rediscovered (Vaughan 1990). The other, *O. brachyantha*, grows in Africa and is the only species bearing the F genome. The four species complexes are: the *O. ridleyi* complex, the *O. meyeriana* complex, the *O. officinalis* complex and the *O. sativa* complex. The *O. ridleyi* complex includes two tetraploid species, *O. ridleyi* and *O. longiglumis*, which are distributed in Southeast Asia, New Guinea, Iran and Indonesia. The *O. meyeriana* complex contains two diploid species (*O. meyeriana* and *O. granulata*) found in South and Southeast Asia including Southern China. The *O. officinalis* complex, also called the *O. latifolia* group by Second (1985b), is highly diverse genetically and geographically and contains both diploid (B, C and E genomes) and tetraploid (CD and BC genomes) species. The *O. sativa* complex or the *O. sativa* group (Second 1985a) contains all A genome species including the two cultivated species, *O. sativa* L. and *O. glaberrima*. *O. sativa* with two main subspecies (indica and japonica) is Asian in origin, but is now cultivated worldwide. *O. glaberrima*, on the other hand, is cultivated only in Africa. *O. rufipogon* (Asian form) and *O. barthii* (Africa) (also known as *O. breviligulata*) are generally accepted as the wild progenitors of the two cultigens respectively (Oka 1988).

The genus *Oryza* has been subjected to cytogenetic (Hu 1970; Katayama T 1982; Morinaga 1964; Nayar 1973), morphometric (Ghesquière 1988; Morishima and Oka 1960; Tateoka 1962), isozyme (Second 1982, 1985a, b) and chloroplast DNA restriction (Dally and Second 1990) analyses, and all have contributed in complementary ways to our current understanding of the genus (Oka 1988; Vaughan 1989). RFLP (restriction fragment length polymorphism) analysis of nuclear DNA provides an additional tool for studying genetic variation and phylogenetic relationships among populations and species (Song et al. 1988, 1990; Miller and Tanksley 1990). Since RFLP analysis is done directly at the DNA level, it reflects heritable changes in the nucleotide sequence both in coding and non-coding regions. As a result, RFLP studies are more sensitive to genetic changes than isozymes, which reflect only those changes resulting in specific amino acid substitutions. Moreover, since there are many more RFLP markers available in rice than isozyme markers it is possible to base a study on a relatively large number of loci scattered throughout the genome (McCough et al. 1988).

The objective of the study reported here was to determine the level of RFLP variation both within accessions and within species in the genus *Oryza* and to determine the phylogenetic relationships among species in this genus. Specific questions that we wished to address were the following. (1) How much genetic variation exists between versus within species? (2) What percentage of RFLP clones detect polymorphism between cultivated rice (*O. sativa*) and other races and species within the genus? (3) Was cultivated rice (*O. sativa*) the result of a single or more than one independent domestication events? (4) What is the relationship between extant diploid species and their tetraploid relatives? (5) What is the general relationship of species in the genus *Oryza*? (6) What has been the general worldwide pattern of dispersal since the early origins of the genus in Asia? Results from this study shed light on all of these topics.

## Materials and methods

### *Plant material*

In order to assess within accession variation, three to five individuals were sampled from each of 22 accessions representing a total of 19 species from the genus *Oryza* (Table 1). Each sample was tested with 15 RFLP probes using a single restriction enzyme (EcoRI). These accessions were provided by the International Rice Germplasm Center (IRGC) at the International Rice Research Institute (IRRI), Los Baños, The Philippines. For a larger survey of variation in the genus, DNA was extracted from single plants from each of the above accessions, as well as from 71 additional accessions obtained from ORSTOM previously studied for isozyme and chloroplast DNA diversity (Second 1984, 1985a; Dally and Second 1990; Table 2), and assayed with a total of 25 probes (including 9 of the same 15 probes used earlier).

### *RFLP probing*

DNA extraction, restriction endonuclease digestion, electrophoresis, Southern blotting, hybridization, genomic library construction and autoradiography have been described previously (McCouch et al. 1988). Previous experience with cultivated rice suggested that, for a single probe, polymorphism detected with different restriction enzymes is likely to be due to the same mutation (probably insertion, deletion or other DNA rearrangements) (McCouch et al. 1988; Wang and Tanksley 1989). To

**Table 1.** Accessions from the genus *Oryza* used in determining within accession polymorphism. Within accession polymorphism is expressed as the number of polymorphic clones out of the 15 clones used in this study

| IRRI accession | Number species name | Genome | Number of individuals sampled per accession | Proportion of clones detecting polymorphism within accessions |
|---|---|---|---|---|
| 101395 | *O. alta* | CCDD | 5 | 0/15 |
| 100882 | *O. australiensis* | EE | 4 | 3/15 |
| 100122 | *O. barthii* | AA | 5 | 0/15 |
| 103410 | *O. collina* | CC | 4 | 2/15 |
| 101422 | *O. eichingeri* | CC | 4 | 2/15 |
| 101425 | *O. eichingeri* | CC | 5 | 0/15 |
| 102201 | *O. glaberrima* | AA | 5 | 0/15 |
| 100968 | *O. glumaepatula* | AA | 5 | 4/15 |
| 101405 | *O. grandiglumis* | CCDD | 4 | 0/15 |
| 100914 | *O. latifolia* | CCDD | 5 | 14/15 |
| 101378 | *O. longistaminata* (obake) | AA | 5 | 0/15 |
| 100957 | *O. malampuzhaensis* | BBCC | 5 | 0/15 |
| 101147 | *O. meridionalis* | AA | 4 | 0/15 |
| 101089 | *O. minuta* | BBCC | 5 | 0/15 |
| 101125 | *O. minuta* | BBCC | 5 | 0/15 |
| 101508 | *O. rufipogon* (nivara) | AA | 5 | 0/15 |
| 100896 | *O. officinalis* | CC | 5 | 0/15 |
| 101150 | *O. officinalis* | CC | 3 | 0/15 |
| 101409 | *O. punctata* | BB, BBCC | 5 | 0/15 |
| 100821 | *O. ridleyi* | Tetraploid | 5 | 0/15 |
| 103831 | *O. rufipogon* (spontanea) | AA | 4 | 0/15 |
| 101141 | *O. minuta* | BBCC | 5 | 0/15 |

avoid analytical problems engendered by this lack of independence with multiple enzymes, we opted to use only a single restriction endonuclease, EcoRI, for digestion of all DNA samples.

Fifteen single-copy probes were used in the first set of accessions for estimating within population variations and 25 single-copy probes were used to probe the second set of 93 accessions. All the probes were from a Pst I genomic library and had previously been shown to detect polymorphism between an indica and japonica variety used to construct an RFLP linkage map of rice (McCouch et al. 1988). These clones are also known to correspond to loci mapping to different chromosomes (Table 3; McCouch et al. 1988).

*Data analysis*

Each hybridizing fragment detected by Southern analysis was treated as a unit character. All hybridizing fragments detected in a single accession, regardless of frequency, were used to make comparisons among accessions or species. Genetic distances both within and between accessions or species were quantified according to Nei (1987) (formula 5.53 to 5.55) using a Macintosh computer software, HyperRFLP®, developed in our laboratory. Dendrograms were then constructed by the same software using the unweighted pair-group method with arithmetic mean (UPGMA, Sokal and Michener 1958). PCA analysis (principal component analysis) was done using DATA DESK, a Macintosh computer program. Variation within and between species were quantified using Nei genetic distances.

## Results and discussion

*Polymorphism*

An average of 11.2 unique fragments was detected per probe (determined over all accessions) – a value much larger than that observed in a previous study which included only cultivated rice (*O. sativa*) in which an average of 3.4 unique fragments were reported per probe (Wang and Tanksley 1989). This result is not surprising considering the wide range of species sampled.

*Variation within accessions.* Five of 22 accessions, for which more than one plant was sampled, were found to be polymorphic for at least 1 probe. This percentage of polymorphic accessions (23%) is very close to the percentage of polymorphic accessions (26%) previously reported for a larger study of rice cultivars (Wang and Tanksley 1989). The number of polymorphic clones per accession is given in Table 1. Variant individuals normally displayed a unique fragment(s), in place of the common fragment found in other individuals, suggesting that they were homozygous and therefore that the accessions tested are actually mixed pure lines. Similar results, based

**Table 2.** The 93 accessions from genus *Oryza* used in this study. Number of fragments equals the total number of restriction fragments present in each accession across 25 clones

| Accession number | Accession code | Original classification | RFLP classification | Genome | Number of fragments | Origin |
|---|---|---|---|---|---|---|
| 1 | 17054 | *O. sativa* (japonica) | *O. sativa* (japonica) | AA | 28 | China-Tai |
| 2 | 51064 | *O. sativa* (indica) | *O. sativa* (indica) | AA | 32 | Sri Lanka |
| 3 | 102201 | *O. glaberrima* | *O. glaberrima* | AA | 32 | West Africa |
| 4 | 103831 | *O. rufipogon* (spontanea) | *O. rufipogon* | AA | 32 | Bangladesh |
| 5 | DS14 | *O. rufipogon* | *O. rufipogon* | AA | 35 | West India |
| 6 | DR38 | *O. rufipogon* | *O. rufipogon* | AA | 31 | West India |
| 7 | W555 | *O. rufipogon* | *O. rufipogon* | AA | 34 | Sri Lanka |
| 8 | W162 | *O. rufipogon* | *O. rufipogon* | AA | 33 | Thailand |
| 9 | W1699 | *O. rufipogon* | *O. rufipogon* | AA | 34 | Thailand |
| 10 | W133 | *O. rufipogon* | *O. rufipogon* | AA | 33 | India |
| 11 | W135 | *O. rufipogon* | *O. rufipogon* | AA | 32 | India |
| 12 | W593 | *O. rufipogon* | *O. rufipogon* | AA | 39 | Malaysia |
| 13 | 103822 | *O. rufipogon* | *O. rufipogon* | AA | 33 | China |
| 14 | Plat 88-791 | *O. rufipogon* | *O. rufipogon* | AA | 36 | Thailand |
| 15 | W1669 | *O. rufipogon* | *O. rufipogon* | AA | 34 | India |
| 16 | W1654 | *O. rufipogon* | *O. rufipogon* | AA | 33 | China |
| 17 | W1655 | *O. rufipogon* | *O. rufipogon* | AA | 34 | China |
| 18 | 100968 | *O. glumaepatula* | *O. glumaepatula* | AA | 36 | Paramaribo, Surinam |
| 19 | W1185 | *O. rufipogon* | *O. glumaepatula* | AA | 36 | America |
| 20 | OR7 | *O. rufipogon* | *O. meridionalis* | AA | 30 | Australia |
| 21 | OR10 | *O. rufipogon* | *O. rufipogon* | AA | 32 | Australia |
| 22 | OR39 | *O. rufipogon* | *O. rufipogon* | AA | 28 | Australia |
| 23 | OR54 | *O. rufipogon* | *O. meridionalis* | AA | 36 | Australia |
| 24 | W1627 | *O. rufipogon* | *O. meridionalis* | AA | 32 | Australia |
| 25 | 101147 | *O. meridionalis* | *O. meridionalis* | AA | 31 | Australia |
| 26 | 101508 | *O. nivara* | *O. rufipogon* | AA | 34 | India |
| 27 | 100122 | *O. barthii* | *O. barthii* | AA | 37 | Gambia |
| 28 | Wb01 | *O. barthii* | *O. barthii* | AA | 37 | Botswana |
| 29 | TL81 | *O. longistaminata* | *O. longistaminata* | AA | 41 | Tchad |
| 30 | 101378 | *O. longistaminata* (Obake) | *O. rufipogon* | AA | 44 | Mali |
| 31 | ILL116 | *O. longistaminata* | *O. longistaminata* | AA | 35 | Mali |
| 32 | IL52 | *O. longistaminata* | *O. longistaminata* | AA | 41 | Ivory Coast |
| 33 | WL02 | *O. longistaminata* | *O. longistaminata* | AA | 39 | Botswana |
| 34 | UL12-6 | *O. longistaminata* | *O. longistaminata* | AA | 38 | Cameroun |
| 35 | YL244 | *O. longistaminata* | *O. longistaminata* | AA | 37 | Guinea |
| 36 | ZL14 | *O. longistaminata* | *O. longistaminata* | AA | 37 | Zambia |
| 37 | CL7-2 | *O. longistaminata* (Obake) | *O. longistaminata* | AA | 39 | Senegal |
| 38 | W1590 | *O. punctata* | *O. punctata* | BB | 30 | Cameroun |
| 39 | W1515 | *O. punctata* | *O. punctata* | BB | 30 | Tanzania |
| 40 | TP43 | *O. punctata* | *O. punctata* | BB | 30 | Tchad |
| 41 | 101089 | *O. minuta* | *O. minuta* | BBCC | 52 | Philippines |
| 42 | 101125 | *O. minuta* | *O. minuta* | BBCC | 52 | Philippines |
| 43 | 101141 | *O. minuta* | *O. minuta* | BBCC | 52 | Philippines |
| 44 | 103865 | *O. minuta* | *O. minuta* | BBCC | 51 | Philippines |
| 45 | W1331 | *O. minuta* | *O. punctata* | BBCC | 49 | Philippines |
| 46 | IP27 | *O. punctata* | *O. punctata* | BBCC | 49 | Ivory Coast |
| 47 | W1408 | *O. punctata* | *O. punctata* | BBCC | 44 | Nigeria |
| 48 | 101409 | *O. punctata* | *O. punctata* | BBCC | 49 | labo labo, Ghana |
| 49 | 100957 | *O. malampuzhaensis* | *O. malampuzhaensis* | BBCC | 48 | India |
| 50 | 103410 | *O. collina* | *O. collina* | CC | 29 | Sri Lanka |
| 51 | 103421 | *O. collina* | *O. collina* | CC | 27 | Sri Lanka |
| 52 | 101422 | *O. eichingeri* | *O. eichingeri* | CC | 31 | Uganda |
| 53 | W1526 | *O. eichingeri* | *O. eichingeri* | CC | 30 | Uganda |
| 54 | 101425 | *O. eichingeri* | *O. officinalis* | CC | 32 | Uganda |
| 55 | IP7 | *O. eichingeri* | *O. eichingeri* | CC | 29 | Ivory Coast |
| 56 | 100896 | *O. officinalis* | *O. officinalis* | CC | 32 | Thailand |
| 57 | 101150 | *O. officinalis* | *O. officinalis* | CC | 32 | Sabah, East Malaysia |
| 58 | D04 | *O. officinalis* | *O. officinalis* | CC | 31 | India |
| 59 | W1278 | *O. officinalis* | *O. officinalis* | CC | 32 | Sarawak |
| 60 | 100181 | *O. officinalis* | *O. minuta* | BBCC | 52 | Burma |

**Table 2.** (Continued)

| Accession number | Accession code | Original classification | RFLP classification | Genome | Number of fragments | Origin |
|---|---|---|---|---|---|---|
| 61 | 100180 | *O. officinalis* | *O. punctata* | BBCC | 51 | Malaysia |
| 62 | 104618 | *O. officinalis* | *O. officinalis* | CC | 32 | China |
| 63 | ch 83-3 | *O. officinalis* | *O. latifolia* | CCDD | 42 | China |
| 64 | 105392 | *O. officinalis* | *O. officinalis* | CC | 30 | China |
| 65 | 105393 | *O. officinalis* | *O. officinalis* | CC | 29 | China |
| 66 | 105394 | *O. officinalis* | *O. officinalis* | CC | 29 | China |
| 67 | 105395 | *O. officinalis* | *O. officinalis* | CC | 32 | China |
| 68 | 105396 | *O. officinalis* | *O. officinalis* | CC | 32 | China |
| 69 | 101395 | *O. alta* | *O. alta* | CCDD | 47 | South America |
| 70 | 101405 | *O. grandiglumis* | *O. grandiglumis* | CCDD | 45 | Brazil |
| 71 | 100914 | *O. latifolia* | *O. latifolia* | CCDD | 55 | Campo cotaxtla, Mexico |
| 72 | W1168 | *O. latifolia* | *O. latifolia* | CCDD | 46 | Cuba |
| 73 | W1144 | *O. latifolia* | *O. alta* | CCDD | 58 | South America |
| 74 | 100963 | *O. latifolia* | *O. latifolia* | CCDD | 55 | Guatemala |
| 75 | OA4 | *O. australiensis* | *O. australiensis* | EE | 28 | Australia |
| 76 | OA27 | *O. australiensis* | *O. australiensis* | EE | 61 | Australia |
| 77 | OA36 | *O. australiensis* | *O. australiensis* | EE | 33 | Australia |
| 78 | 100882 | *O. australiensis* | *O. australiensis* | EE | 33 | Australia |
| 79 | EY25 | *O. brachyantha* | *O. brachyantha* | FF | 38 | Tanzania |
| 80 | W654 | *O. brachyantha* | *O. brachyantha* | FF | 33 | West Africa |
| 81 | W656 | *O. brachyantha* | *O. brachyantha* | FF | 32 | West Africa |
| 82 | W615 | *O. meyeriana* | *O. meyeriana* | Diploid | 33 | Burma |
| 83 | W3 | *O. meyeriana* | *O. meyeriana* | Diploid | 34 | India |
| 84 | W5 | *O. meyeriana* | *O. meyeriana* | Diploid | 31 | Sri Lanka |
| 85 | W67 | *O. meyeriana* | *O. meyeriana* | Diploid | 38 | Thailand |
| 86 | W609 | *O. meyeriana* | *O. meyeriana* | Diploid | 33 | China |
| 87 | W1348 | *O. meyeriana* | *O. meyeriana* | Diploid | 50 | Borneo |
| 88 | W1228 | *O. longiglumis* | *O. longiglumis* | Tetraploid | 41 | New Guinea |
| 89 | W1220 | *O. longiglumis* | *O. longiglumis* | Tetraploid | 43 | New Guinea |
| 90 | 100821 | *O. ridleyi* | *O. ridleyi* | Tetraploid | 42 | Thailand |
| 91 | W2033 | *O. ridleyi* | *O. ridleyi* | Tetraploid | 45 | Thailand |
| 92 | W604 | *O. ridleyi* | *O. ridleyi* | Tetraploid | 47 | Malaysia |
| 93 | W1 | *O. ridleyi* | *O. ridleyi* | Tetraploid | 47 | Thailand |

on RFLPs and isozymes, have been observed in other self-pollinated species (Miller and Tanksley 1990).

When an accession was polymorphic for more than 1 probe, it was common that the same individual contained the variant fragments. An extreme example of this was found in one accession of *O. latifolia* (accession 10094) in which a single individual contained unique polymorphisms for 14 out of 15 probes tested. It seems unlikely that such extreme within-population differentiation could be maintained naturally and is more likely to represent a seed mixture. This individual was thus excluded from further study.

It is interesting to note that the only accession from a self-incompatible species (*O. longistaminata*) was monomorphic. Self-incompatibility is thought to be associated with increased within-population variation – a prediction borne out by both isozyme and RFLP studies in other plants (Miller and Tanksley 1990). Isozyme studies showed that *O. longistaminata* has a high level of variation both within and between populations even though morphologically it is not so variable (Ghesquière 1988; M. Causse personal communication). The particular accession of *O. longistaminata* (101378) used in this study has been maintained in the IRRI collection for a number of years and may have been unintentionally selected for self-compatibility. There is also some question of whether this accession represents a true *O. longistaminata* or a derivative from a *O. longistaminata* × *O. sativa*. Such introgressive derivatives are not uncommon and have been referred to as 'Obake' types (Ghesquière 1988).

*Variation within and between species.* Genetic variation within species was defined as the mean distance of all pair-wise comparisons between different accessions from the same species. Genetic variation between species was defined as the mean distances between accessions from *different* species. Based on RFLP analysis, some of the accessions received for this study were apparently mis-

**Table 3.** Cloned probes used in RFLP study

| Probes | Chromosome[b] | Number unique fragment detected |
|---|---|---|
| RG120 | ? | 32 |
| RG139 | 2 | 39 |
| RG144[a] | 2 | 34 |
| RG157 | 2 | 33 |
| RG181 | 12 | 40 |
| RG191[a] | 3 | 49 |
| RG214 | 4 | 51 |
| RG220 | 1 | 27 |
| RG222[a] | 1 | 36 |
| RG236 | 1 | 32 |
| RG246 | 1 | 36 |
| RG341[a] | 12 | 31 |
| RG386[a] | 9 | 45 |
| RG396 | 12 | 33 |
| RG400[a] | 1 | 50 |
| RG424 | 6 | 39 |
| RG437 | 2 | 53 |
| RG450 | 3 | 35 |
| RG451 | 9 | 33 |
| RG462[a] | 1 | 52 |
| RG472[a] | 1 | 45 |
| RG479[a] | ? | 47 |
| RG484 | 12 | 36 |
| RG532 | 1 | 48 |
| RG553 | 9 | 42 |

? Unmapped clone
[a] Used in within-accession variation study
[b] New rice chromosome numbering system adopted by committee chaired by Dr. Kinoshita et al. 2nd Int. Rice Genetics Symposium, Los Baños, The Philippines (1990)

classified as to species (Table 2, see next section for details). The corrected species assignment is used for discussions in this section.

The average genetic distance between species (0.077) is 6 times larger than those within species (0.013). *O. australiensis* exhibited the greatest within-species variation. One of the accessions of *O. australiensis* (accession OA 27) contained nearly twice as many fragments per individual (61) than other accessions of this species or other diploid species (see next section). Even when this accession of *O. australiensis* was excluded from the analysis, variation for this species as a whole still remained high. That *O. australiensis* was found to be so variable at the DNA level was surprising since it is a relatively uniform species morphologically (G. Second, personal communication). Also high in within-species variation was *O. latifolia* and *O. rufipogon* (which includes cultivated rice, *O. sativa*).

Recently created polyploids are likely to have limited genetic variation due to the genetic bottleneck imposed by the polyploidization event. For example, wheat is believed to be a relatively new polyploid species, and very limited variation can be detected among wheat varieties using RFLP probes (Chao et al. 1988). In rice, however, no significant difference (based on a *t*-test) was observed between diploids and tetraploids with respect to within species polymorphism. In fact, the tetraploid species, *O. latifolia*, was among the highly polymorphic species, suggesting that this species may not be of recent origin (see later section for further discussion).

*Total number of fragments per individual is higher in tetraploids than diploids.* Since allotetraploids contain two different genomes, they are expected to possess more alleles (per individual) than their diploid counterparts (Gale and Sharp 1988; W. Burnquist personal communication). RFLP data provide a unique opportunity to test this hypothesis over a relatively large number of loci. Figure 1 depicts the number of fragments per individual for all rice accessions examined (diploid and tetraploids). The mean number of fragments per invididual for diploids (32.5) and tetraploids was (47.3) different significantly based on a *t*-test ($p < 0.01$). Moreover, the distributions were almost non-overlapping.

The major exceptions were three diploid accessions that had high fragment numbers (Fig. 1). As mentioned previously, *O. australiensis* (accession 76), had nearly twice as many bands (61 versus 32) as other *O. australiensis* accessions. Moreover, this accession, while grouping with other *O. australiensis* accessions in the dendrogram, was less closely related to this species group than any of the other accessions (Fig. 3). A similar situation was seen with *O. meyeriana* accession 87: this accession also had a significantly greater number of fragments per individual than other accessions of this species and was the least closely related (Fig. 3).

We can imagine at least two possible explanations for these anomalous accessions. First, since these accessions were obtained only as leaf tissue from ORSTOM, it is possible that the samples were actually a mixture of tissue from more than a single accession or a mixture of DNA. The second possibility that explains these results equally well is that these two accessions are actually previously unidentified allotetraploids and not diploids. This would explain the increased numbers of fragments per individual and the looser affinity with other diploid accessions of the same species classification. We are currently working on importing seeds of these accessions in order to check their chromosome numbers. The third diploid accession with a higher fragment number was number 30, which is classified as *O. longistaminata*. Since this accession was previously found to be monomorphic, heterozygosity is probably not the cause of the greater number of fragments, and we currently do not have an explanation for this result.

*Percentage polymorphic loci between O. sativa (indica and japonica) and other taxa.* One of the major reasons for developing an RFLP map in rice is to use the markers to detect and monitor genes of interest in a breeding pro-

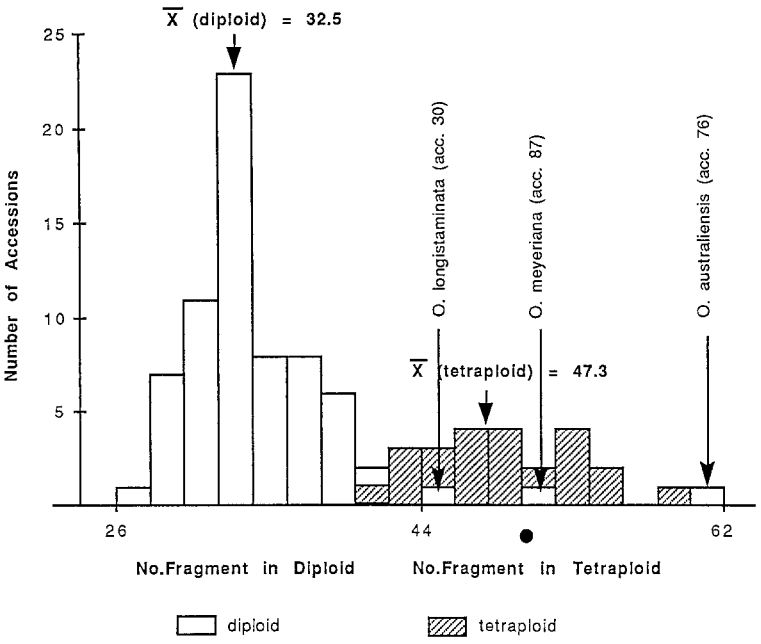



**Fig. 1.** Average number of unique restriction fragments detected per individual for each species in the genus *Oryza*

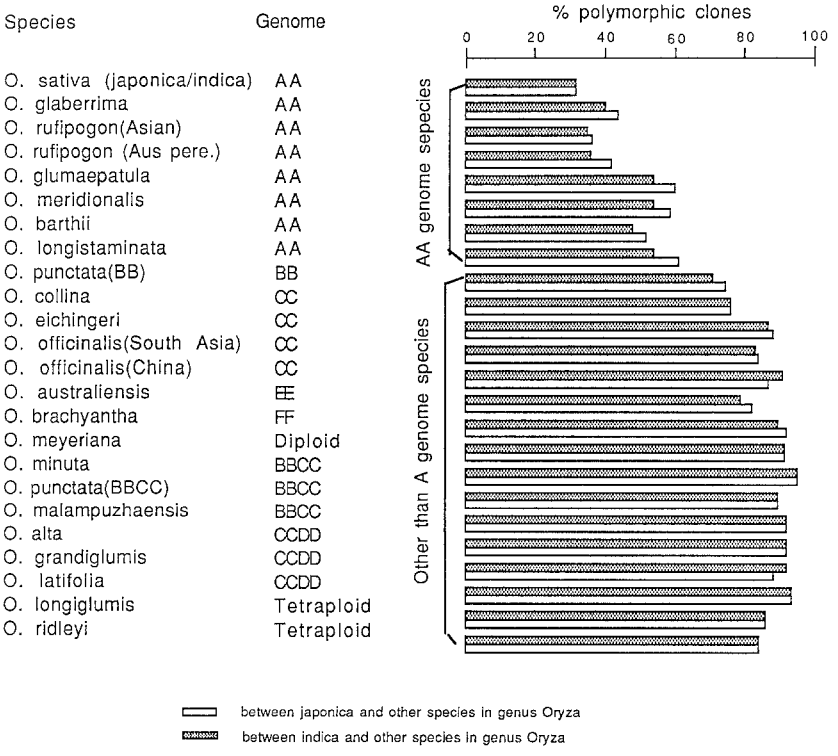



**Fig. 2.** Percentage of polymorphic clones between *O. sativa* (indica and japonica) and other taxa in genus *Oryza* for a single restriction enzyme

gram (McCouch et al. 1988). For a marker to be informative it must segregate in the progeny from a cross between two parents. It is thereafter useful to know what proportion of the markers (clones) are likely to be polymorphic in a cross between two accessions, whether they be from the same species or different species.

Rice breeders use not only cultivated varieties but also wild species as resources of genetic variation. In the interest of using RFLP markers in rice breeding, we have calculated the average percentage of clones showing polymorphism with a single enzyme between the two subspecies, indica and japonica, and between *O. sativa* (indica and japonica) and other taxa in the genus *Oryza* (Fig. 2). Thirty-two percent of the clones revealed the polymorphism with a single restriction enzyme between the indica and japonica varieties used in this study. A

similar value was previously obtained from a study of a larger set of 67 cultivated rice varieties (Wang and Tanksley 1989).

When comparing *O. sativa* (both indica and japonica) with other A genome species in the genus *Oryza*, the percentages of polymorphic clones varied from 35% to 61% (Fig. 2). The value for *O. sativa* versus Asian *O. rufipogon* was 36%, which is very similar to that observed between indica and japonica. This value was also the smallest for any of the comparisons between *O. sativa* and other taxa in the genus and supports the hypothesis that rice was domesticated from wild forms of *O. rufipogon* from Asia (Oka 1988; see also next section).

For comparison of *O. sativa* to species with genomes other than the A genome, the percentage polymorphism ranged from 71% to 96% (Fig. 2). These high values indicate that most RFLP clones will be informative in crosses involving cultivated and wild rice containing other genomes. Since these values were obtained with only a single enzyme, even greater variation might be expected when more than one restriction enzyme is employed.

RFLP markers and maps represent a potentially powerful tool to aid in the introgression of genes from alien species. RFLP techniques provide the means by which to select for desirable rare recombinants that will eliminate or mitigate linkage drag (Tanksley et al. 1989). The high level of polymorphism between cultivated rice and its wild ancestors indicates that RFLP will likely be very useful in this endeavor. Recent studies by Kochert et al. (personal communication) have demonstrated the use of this technique in analyzing introgression products involving crosses with *O. sativa* and *O. officinalis*.

### *Genetic relationships among accessions*

A genetic distance matrix was generated in which each accession was treated as an operational taxonomic unit (OTU) making no assumptions about species assignment. From this matrix a dendrogram was produced depicting the genetic relationships among all accessions (Fig. 3).

### *Diploid species clusters*

*AA genome species.* Accessions of Asian *O. rufipogon* (also called *O. nivara* for its annual life form), Australian perennial *O. rufipogon* and *O. sativa* grouped together to form a loosely knit cluster. Annual forms of *O. rufipogon* from Australia, however, were very distinct and showed no obvious affinity with Asian *O. rufipogon*. Recently, it has been proposed that these Australian annual forms constitute a discrete species (*O. meridionalis*) (Ng et al. 1981). The RFLP data presented here lend strong support to this proposal, and hereafter we refer to these accessions as *O. meridionalis*.

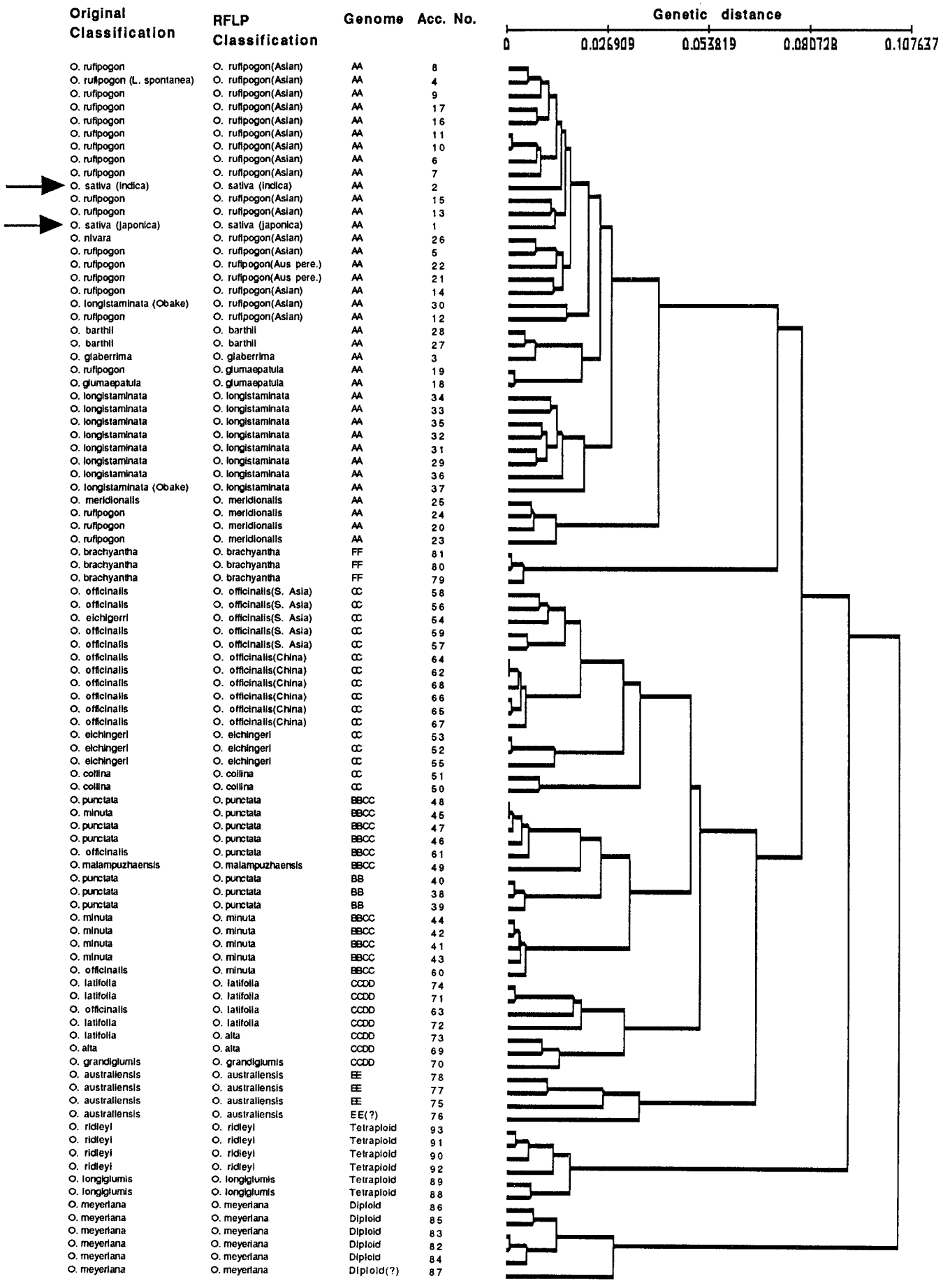
A single accession of *O. rufipogon* from South America was also included and grouped very tightly with the only accession classified as *O. glumaepatula*. This latter species, which is found in South America, has often been considered to be a subtype of *O. rufipogon*. Our results, however, show that it has slightly closer affinity with two African species, *O. barthii* and *O. glaberrima*, than with Asian *O. rufipogon* accessions as a whole and raise the question about the origin of *O. glumaepatula*.

All but one of the remaining A genome accessions from Africa could be partitioned into three species groups: *O. barthii* (also called *O. brevigulata*), *O. glaberrima* and *O. longistaminata*. One accession (accession 30 = 101378) grouped with Asian *O. rufipogon*, but was originally classified as *O. longistaminata*. This particular accession is thought to be an 'obake' – a natural hybrid derivative from a cross between *O. longistaminata* and *O. sativa* – which might explain its association with Asian *O. rufipogon* (Ghesquière 1988).

*CC genome species.* Four species groups could be detected among accessions classified as containing the C genome. Accessions classified as *O. officinalis* separated in the dendrogram to give two separate groups. One set, comprised of accessions from China, formed a cohesive group that could be differentiated from the remaining accessions coming from South Asia (Fig. 3). One accession (54 = IRRI number 101425) clustered with Chinese *O. officinalis* but was originally classified as *O. eichingeri*. Since other accessions classified as *eichingeri* formed a separate, discrete cluster, we assumed that accession 54 had been misclassified. The two remaining C genome accessions clustered together and represent the species *O. collina*.

*Other diploid species.* From the remaining diploid accessions, four species clusters could be clearly identified: *O. brachyantha* (F genome), *O. punctata* (B genome), *O. australiensis* (E genome) and *O. meyeriana* (genome not classified). Diploid *O. punctata* (B genome) was clearly differentiated from tetraploid *O. punctata* (BBCC) – a result consistent with isozyme and chloroplast DNA studies (Second 1985b; Dally and Second 1990). All of these data suggest that these two forms should really be considered as two separate species.

**Fig. 3.** A computer-generated dendrogram of all 93 accessions of genus *Oryza* based on RFLPs detected with 25 genomic clones. *Left side, first column* – original classification; *second column* – RFLP classification; *third column* – genome designation; *fourth column* – accession number (see Table 2). Genetic distance is according to Nei (Nei 1987)


Original Classification
RFLP Classification
Genome
Acc. No.
Genetic distance
0
0.026909
0.053819
0.080728
0.107637
O. rufipogon O. rufipogon(Asian) AA 8
O. rufipogon (L. spontanea) O. rufipogon(Asian) AA 4
O. rufipogon O. rufipogon(Asian) AA 9
O. rufipogon O. rufipogon(Asian) AA 17
O. rufipogon O. rufipogon(Asian) AA 16
O. rufipogon O. rufipogon(Asian) AA 11
O. rufipogon O. rufipogon(Asian) AA 10
O. rufipogon O. rufipogon(Asian) AA 6
O. rufipogon O. rufipogon(Asian) AA 7
O. sativa (indica) O. sativa (indica) AA 2
O. rufipogon O. rufipogon(Asian) AA 15
O. rufipogon O. rufipogon(Asian) AA 13
O. sativa (japonica) O. sativa (japonica) AA 1
O. nivara O. rufipogon(Asian) AA 26
O. rufipogon O. rufipogon(Asian) AA 5
O. rufipogon O. rufipogon(Aus pere.) AA 22
O. rufipogon O. rufipogon(Aus pere.) AA 21
O. rufipogon O. rufipogon(Asian) AA 14
O. longistaminata (Obake) O. rufipogon(Asian) AA 30
O. rufipogon O. rufipogon(Asian) AA 12
O. barthii O. barthii AA 28
O. barthii O. barthii AA 27
O. glaberrima O. glaberrima AA 3
O. rufipogon O. glumaepatula AA 19
O. glumaepatula O. glumaepatula AA 18
O. longistaminata O. longistaminata AA 34
O. longistaminata O. longistaminata AA 33
O. longistaminata O. longistaminata AA 35
O. longistaminata O. longistaminata AA 32
O. longistaminata O. longistaminata AA 31
O. longistaminata O. longistaminata AA 29
O. longistaminata O. longistaminata AA 36
O. longistaminata (Obake) O. longistaminata AA 37
O. meridionalis O. meridionalis AA 25
O. rufipogon O. meridionalis AA 24
O. rufipogon O. meridionalis AA 20
O. rufipogon O. meridionalis AA 23
O. brachyantha O. brachyantha FF 81
O. brachyantha O. brachyantha FF 80
O. brachyantha O. brachyantha FF 79
O. officinalis O. officinalis(S. Asia) CC 58
O. officinalis O. officinalis(S. Asia) CC 56
O. eichigerri O. officinalis(S. Asia) CC 54
O. officinalis O. officinalis(S. Asia) CC 59
O. officinalis O. officinalis(S. Asia) CC 57
O. officinalis O. officinalis(China) CC 64
O. officinalis O. officinalis(China) CC 62
O. officinalis O. officinalis(China) CC 68
O. officinalis O. officinalis(China) CC 66
O. officinalis O. officinalis(China) CC 65
O. officinalis O. officinalis(China) CC 67
O. eichingeri O. eichingeri CC 53
O. eichingeri O. eichingeri CC 52
O. eichingeri O. eichingeri CC 55
O. collina O. collina CC 51
O. collina O. collina CC 50
O. punctata O. punctata BBCC 48
O. minuta O. punctata BBCC 45
O. punctata O. punctata BBCC 47
O. punctata O. punctata BBCC 46
O. officinalis O. punctata BBCC 61
O. malampuzhaensis O. malampuzhaensis BBCC 49
O. punctata O. punctata BB 40
O. punctata O. punctata BB 38
O. punctata O. punctata BB 39
O. minuta O. minuta BBCC 44
O. minuta O. minuta BBCC 42
O. minuta O. minuta BBCC 41
O. minuta O. minuta BBCC 43
O. officinalis O. minuta BBCC 60
O. latifolia O. latifolia CCDD 74
O. latifolia O. latifolia CCDD 71
O. officinalis O. latifolia CCDD 63
O. latifolia O. latifolia CCDD 72
O. latifolia O. alta CCDD 73
O. alta O. alta CCDD 69
O. grandiglumis O. grandiglumis CCDD 70
O. australiensis O. australiensis EE 78
O. australiensis O. australiensis EE 77
O. australiensis O. australiensis EE 75
O. australiensis O. australiensis EE(?) 76
O. ridleyi O. ridleyi Tetraploid 93
O. ridleyi O. ridleyi Tetraploid 91
O. ridleyi O. ridleyi Tetraploid 90
O. ridleyi O. ridleyi Tetraploid 92
O. longiglumis O. longiglumis Tetraploid 89
O. longiglumis O. longiglumis Tetraploid 88
O. meyeriana O. meyeriana Diploid 86
O. meyeriana O. meyeriana Diploid 85
O. meyeriana O. meyeriana Diploid 83
O. meyeriana O. meyeriana Diploid 82
O. meyeriana O. meyeriana Diploid 84
O. meyeriana O. meyeriana Diploid(?) 87

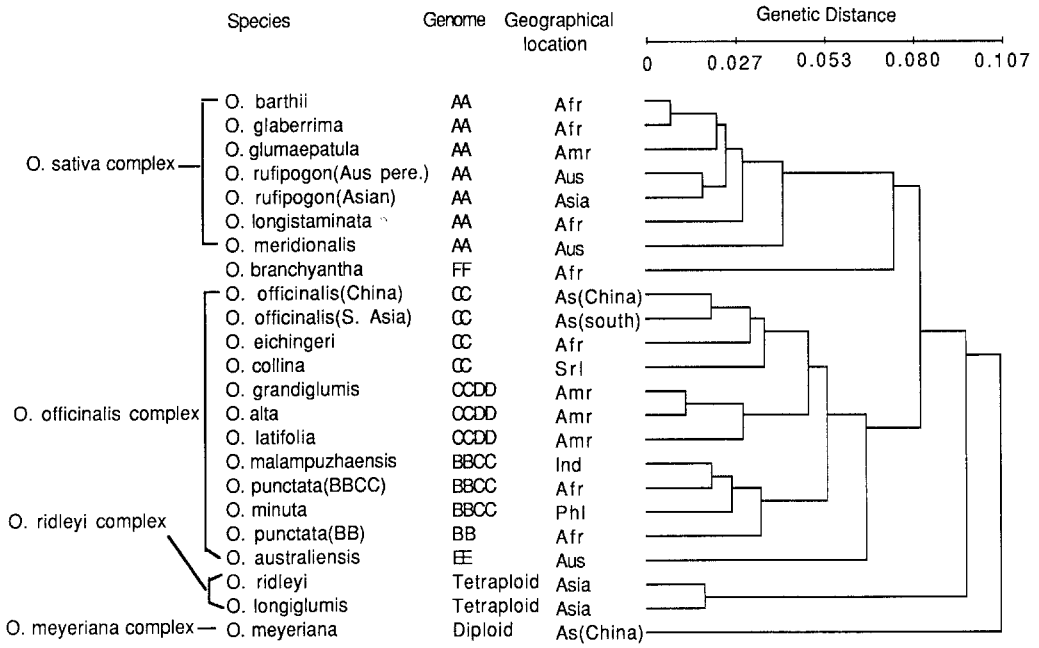



**Fig. 4.** Dendrogram of species in genus *Oryza* based on mean genetic distances between species

*Tetraploid species clusters*

Tetraploid accessions clustered into six species clusters: *O. punctata* (BBCC), *O. malampuzhaensis* and *O. minuta*, all BBCC species, formed three of the groups and showed close association with diploid *O. punctata* (BB) (see next section for further discussions about *O. punctata*). *O. latifolia* and *O. alta*/*O. grandiglumis* (CCDD) formed two more clusters, and *O. ridleyi*/*O. longiglumis* (tetraploids of unknown genome) make up the last cluster. Accession 45 was originally classified as *O. minuta*, but was clearly clustered with *O. punctata* (BBCC). Accession 60 was also apparently misclassified as *O. officinalis* since it showed close affinity with *O. minuta* (BBCC) and, as discussed in the previous section, it displayed a large number of fragments per individual, indicative of polyploidy.

*Genetic relationships among species*

To make a species dendrogram we pooled data from the accessions into species. This was done based on the RFLP classification (Fig. 3, Table 2). *O. sativa* ssp. indica and *O. sativa* ssp. japonica were inseparable from Asian *O. rufipogon*, so they were included in the Asian *O. rufipogon* subgroup for construction of the dendrogram. Distances between two species were calculated by averaging all distances of pair-wise comparisons between all accessions in one species and all accessions in another species. The resulting distances[1] were used to generate dendrograms.

[1] A table of genetic distances (minimum, average, maximum) among all species in this study is available upon request

*I. Relationships among species in genus Oryza (including tetraploids)*. The relationships among *Oryza* species are shown in Fig. 4. Four groups appear in the species dendrogram and *O. brachyantha* (FF) stands as an independent species (Fig. 4). The first group contains all A genome species, which corresponds to the *O. sativa* complex. The second group, corresponding to the *O. officinalis* complex, is composed of B, C, E, BC and CD genome species. *O. ridleyi* (tetraploid) and *O. longiglumis* (tetraploid) appeared to be closely related to one another, forming the third group corresponding to the *O. ridleyi* complex. *O. meyeriana* (diploid), representing the *O. meyeriana* complex, stands out as an independent branch. *O. meyeriana* is the only one of two species of the *O. meyeriana* complex described by Vaughan (1989), and thus this branch is considered as the complex. The first group (the *O. sativa* complex) has some affinity with *O. brachyantha* (FF), and then together they merge with the second group (the *O. officinalis* complex). The next to be merged was the third group (the *O. ridleyi* complex), and the final merger is between the fourth group (the *O. meyeriana* complex) and the rest of the groups (species), suggesting a more distant relationship of *O. meyeriana* with other species in the genus *Oryza* (Fig. 4).

Within the first group (the *O. sativa* complex), *O. barthii* and *O. glaberrima* clustered first. Together they then clustered with *O. glumaepatula*. Australian *O. rufipogon* (perennial form) and Asian *O. rufipogon* (including indica and japonica) clustered together; then all together clustered with *O. longistaminata*, an African wild species. The last merger in the complex was with *O. meridionalis*, which is certainly the most distant species in this complex (group). Within the second group,

three C genome species (*O. officinalis*, *O. eichingeri* and *O. collina*) and three CCDD species (*O. alta*, *O. grandiglumis* and *O. latifolia*) form two subgroups. *O. minuta* (BBCC), *O. punctata* (BB and BBCC) and *O. malampuzhaensis* (BBCC) form another subgroup. *O. australiensis* (EE) appears to be allied with these three subgroups and is the most distant species in this group (complex).

Among CD genome species *O. alta* and *O. grandiglumis* are more closely related to each other than to *O. latifolia*. However, one of four accessions of *O. latifolia* from South America was separated from other accessions and clustered with *O. alta*, and the rest of accessions of *O. latifolia* formed another subcluster. It is interesting to note that our results are consistent with those based on hybridization to rice repeat sequences (Wu and Wu, in preparation). For the BC genome, *O. malampuzhaensis* and *O. minuta* are most closely related, and *O. punctata* (BBCC) is the next closely related species.

In general, our results are in accordance with previous authors' classification of genus *Oryza* (Morishima and Oka 1960; Morishima et al. 1984; Second 1984, 1985a, b; Tateoka 1962; Dally and Second 1990; Vaughan 1989). The four groups appearing in the dendrogram correspond to previously proposed species complexes (Vaughan 1989). Two cultivated species, *O. sativa* and *O. glaberrima*, each clustered with their previously proposed ancestral species, Asian *O. rufipogon* and *O. barthii*, respectively. PCA analysis was also performed on the same data, and the results are, in general, in agreement with those of the dendrograms (Fig. 5). On the first axis the *O. sativa* complex was separated from the *O. ridleyi* complex and the *O. meyeriana* complex, and the second axis put the *O. sativa* complex from the *O. officinalis* complex and the *O. ridleyi* complex from the *O. meyeriana* complex. It is interesting to note that there is one discrepancy between the species dendrogram and the accession dendrogram: *O. malampuzhaensis* is clustered together with *O. minuta* in the species dendrogram, while in the accession dendrogram it is merged with *O. punctata* (BBCC)

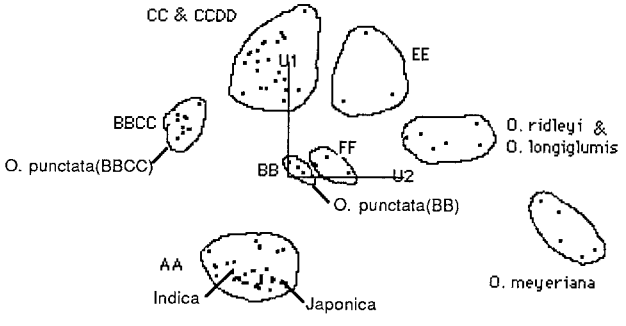



**Fig. 5.** Principal component analysis (PCA) of 93 accessions from the genus *Oryza*. The first two axes extract 73% of total variation. The first axis extracts 46% genetic variation and the second, 27%

(Figs. 3 and 4). This could be due to the fact that in the construction of species dendrogram *O. malamuzhaensis* has only one accession, and it could sometimes reshape the fine structure of the dendrogram.

Figure 4 provides information on the pattern of worldwide dispersal of the *Oryza* genus. By far the greatest diversity of species from the genus resides in Asia and is consistent with an Asian origin of the genus (Oka 1988; Second 1985b). That African species were dispersed in six different places in the dendrogram instead of clustering into a single cluster suggests that rice in Africa was a result of multiple introduction events. In other words, rice species from Asia probably migrated to Africa several times in the past. Similarly, rice in Australia has apparently migrated from Asia more than once since species in Australia were distributed into different clusters rather than in a single cluster.

*II. Relationships among diploid species in the genus Oryza.* Figure 6 is a dendrogram depicting relationships among diploid accessions only. Each genome type (A, B, C, etc.) forms a discrete cluster with branch points of differentiation between genome types occurring at genetic distances between 0.07 and 0.11 compared with the largest distances (0.04) between accessions of the same genome. These results indicate that genome classification in rice, based on meiotic chromosome pairing, reflects genetic distance and that species classified as containing the same genome are probably of monophyletic origin. From figure 6, it is also possible to make inferences about relationships among genomes. For example, the A genome appears to be most closely related to the B genome and the C, E and F genomes appear to be more closely related to each other than to the A and B genomes. The most divergent genome is that found in the diploid species *O. meyeriana*.

III. Relationship between wild species and cultivated rice

*Asian rice.* As indicated earlier, the two subspecies of *O. sativa* (indica and japonica) are closely associated, in the accession dendrogram, with accessions of *O. rufipogon* from Asia (Fig. 3). It is interesting to note that the *O. sativa* ssp. japonica and *O. sativa* spp. indica accessions used in this study show a closer affinity to specific accessions of *O. rufipogon* than they do to each other (Fig. 3). The japonica variety showed the closest relationship to *O. rufipogon* accessions 13 and 15 (India), whereas the indica variety was most closely allied with accessions from India, Sri Lanka, Thailand and China. These results suggest two things. First, *O. rufipogon* from Asia probably represents the genetic stock from which both forms of cultivated Asian rice (*O. sativa* spp indica and japonica) derived. Second, since indica and japonica shown closer affinities to specific accessions of *O. ru-*

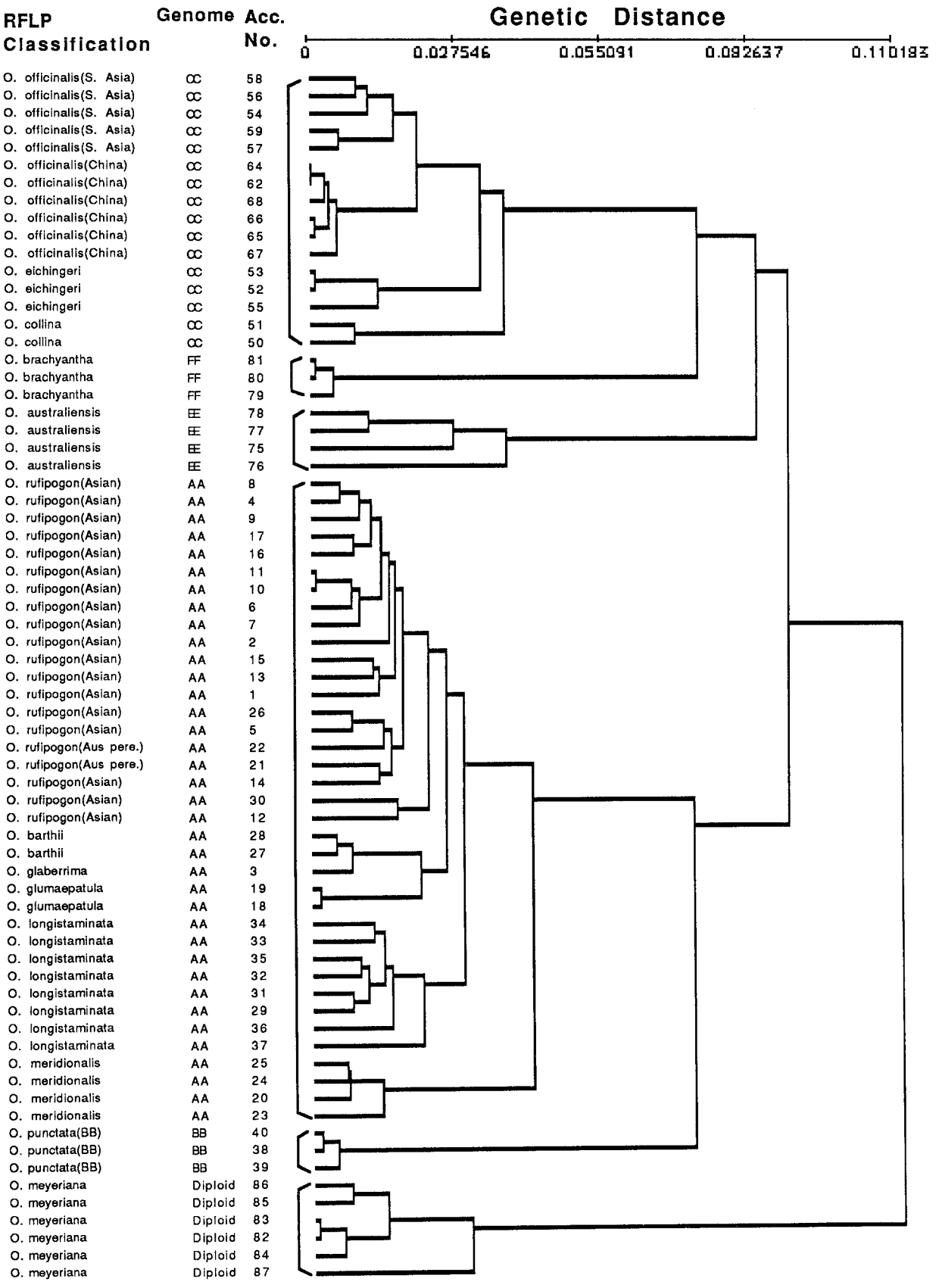

RFLP Classification
Genome
Acc. No.
Genetic Distance
0
0.027546
0.055091
0.082637
0.110183
O. officinalis(S. Asia) CC 58
O. officinalis(S. Asia) CC 56
O. officinalis(S. Asia) CC 54
O. officinalis(S. Asia) CC 59
O. officinalis(S. Asia) CC 57
O. officinalis(China) CC 64
O. officinalis(China) CC 62
O. officinalis(China) CC 68
O. officinalis(China) CC 66
O. officinalis(China) CC 65
O. officinalis(China) CC 67
O. eichingeri CC 53
O. eichingeri CC 52
O. eichingeri CC 55
O. collina CC 51
O. collina CC 50
O. brachyantha FF 81
O. brachyantha FF 80
O. brachyantha FF 79
O. australiensis EE 78
O. australiensis EE 77
O. australiensis EE 75
O. australiensis EE 76
O. rufipogon(Asian) AA 8
O. rufipogon(Asian) AA 4
O. rufipogon(Asian) AA 9
O. rufipogon(Asian) AA 17
O. rufipogon(Asian) AA 16
O. rufipogon(Asian) AA 11
O. rufipogon(Asian) AA 10
O. rufipogon(Asian) AA 6
O. rufipogon(Asian) AA 7
O. rufipogon(Asian) AA 2
O. rufipogon(Asian) AA 15
O. rufipogon(Asian) AA 13
O. rufipogon(Asian) AA 1
O. rufipogon(Asian) AA 26
O. rufipogon(Asian) AA 5
O. rufipogon(Aus pere.) AA 22
O. rufipogon(Aus pere.) AA 21
O. rufipogon(Asian) AA 14
O. rufipogon(Asian) AA 30
O. rufipogon(Asian) AA 12
O. barthii AA 28
O. barthii AA 27
O. glaberrima AA 3
O. glumaepatula AA 19
O. glumaepatula AA 18
O. longistaminata AA 34
O. longistaminata AA 33
O. longistaminata AA 35
O. longistaminata AA 32
O. longistaminata AA 31
O. longistaminata AA 29
O. longistaminata AA 36
O. longistaminata AA 37
O. meridionalis AA 25
O. meridionalis AA 24
O. meridionalis AA 20
O. meridionalis AA 23
O. punctata(BB) BB 40
O. punctata(BB) BB 38
O. punctata(BB) BB 39
O. meyeriana Diploid 86
O. meyeriana Diploid 85
O. meyeriana Diploid 83
O. meyeriana Diploid 82
O. meyeriana Diploid 84
O. meyeriana Diploid 87

**Table 4.** Genetic distances between diploid and tetraploid species in genus *Oryza*. The underlined values are the smallest distances between diploids and tetraploids

| | | *O. minuta* BBCC | *O. punctata* BBCC | *O. malampuzhaensis* BBCC | *O. alta* CCDD | *O. grandiglumis* CCDD | *O. latifolia* CCDD |
|---|---|---|---|---|---|---|---|
| *O. rufipogon* (Asian)[a] | AA | 0.054794 | 0.067143 | 0.060544 | 0.087956 | 0.0854 | 0.088424 |
| *O. glaberrima* | AA | 0.062916 | 0.064512 | 0.057783 | 0.087918 | 0.096414 | 0.083373 |
| *O. glumaepatula* | AA | 0.062699 | 0.071763 | 0.078172 | 0.085057 | 0.081776 | 0.083641 |
| *O. meridionalis* | AA | 0.056558 | 0.047426 | 0.063553 | 0.078777 | 0.078627 | 0.088989 |
| *O. rufipogon* (Aust. Pere.) | AA | 0.063212 | 0.069694 | 0.058672 | 0.090404 | 0.103615 | 0.102714 |
| *O. barthii* | AA | 0.053716 | 0.071009 | 0.068479 | 0.084229 | 0.082585 | 0.079868 |
| *O. longistaminata* | AA | 0.069158 | 0.073425 | 0.064642 | 0.070636 | 0.083727 | 0.082095 |
| *O. punctata* (BB) | BB | 0.045714 | 0.023136 | 0.03434 | 0.082105 | 0.092548 | 0.066501 |
| *O. collina* | CC | 0.04416 | 0.03897 | 0.033774 | 0.046817 | 0.044285 | 0.0486157 |
| *O. euchigerri* | CC | 0.039454 | 0.030168 | 0.034387 | 0.045318 | 0.043636 | 0.0463252 |
| *O. officinalis* (South Asia) | CC | 0.033944 | 0.038529 | 0.035021 | 0.053033 | 0.049856 | 0.0508813 |
| *O. officinalis* (China) | CC | 0.040816 | 0.044944 | 0.03807 | 0.052745 | 0.057021 | 0.0550513 |
| *O. australiensis* | EE | 0.065574 | 0.055528 | 0.061003 | 0.056735 | 0.057619 | 0.0573231 |
| *O. branchyantha* | FF | 0.077451 | 0.053168 | 0.059482 | 0.093666 | 0.113246 | 0.078555 |
| *O. meyeriana* | Diploid | 0.088765 | 0.098097 | 0.094913 | 0.093225 | 0.091514 | 0.11185 |

[a] *O. sativa* subspecies indica/japonica included

*fipogon* than to each other, their domestication from *O. rufipogon* was probably not a monophyletic event, rather the two types were separately domesticated from different gene pools of *O. rufipogon*. This conclusion is also supported by isozyme data (Second 1982).

*African rice. O. glaberrima*, the cultivated species of rice from Africa, clustered with African wild species *O. barthii*, suggesting that *O. barthii* is probably the ancestor of African cultivated rice *O. glaberrima*. These results support the hypothesis that the domestication of African rice was independent from domestication of Asian rice (Morishima et al. 1963; Oka 1974, 1988).

*IV. Origin of tetraploid rice species*

While most *Oryza* species are diploid ($2n = 2x = 24$), six of the species in genus *Oryza* are considered to be allotetraploids. They are: *O. minuta* (BC genome), *O. punctata* (BC genome), *O. malampuzhaensis* (BC genome), *O. alta* (CD genome), *O. grandiglumis* (CD genome) and *O. latifolia* (CD genome). On the basis of chromosome pairing in meiosis, researchers have attempted to deduce the genome composition of each of these polyploids (Kihara 1963; Nayar 1973; Katayama 1982). Analysis of RFLPs permits an independent method for deducing the genetic composition of polyploids. In order to deduce which diploid species are most likely to have contributed to the respective tetraploids, a distance matrix of all diploid and tetraploid species was generated (Table 4).

It is obvious from Table 4 that species of the BC genome have smaller distances with B and C genome species (0.034 and 0.038) than with other diploid species (0.066) and that CD genome species have smaller distances with C and E genome species (0.049 and 0.057) than with the other diploid species (0.088). Therefore, species of B and C genomes are the possible relatives of parents of tetraploid species (BBCC), and C genome of all three CCDD tetraploids is possibly from C genome species. The D genome is unknown in *Oryza* at diploid level, and the next most closely related diploid species to the CD genome species is *O. australiensis* (EE) with a distance of 0.057 (Table 4). This suggests that the E genome is related to the D genome and that an E genome ancestor may have played a role in the formation of CCDD tetraploids.

Further, from these putative ancestral species (based on Table 4) the most closely related accessions to the tetraploid species were identified based on genetic distances. To test the fit of complementary diploid genomes to their tetraploid counterparts, 'synthetic' tetraploid accessions were created by pooling fragment data from pairs of putative ancestral diploid accessions. The genetic distances between these synthetic tetraploids and the actual tetraploids was then computed and compared with the genetic distances between the actual tetraploids and the corresponding diploid accessions (Table 5). If the diploids comprising the synthetic tetraploid are indeed the most closely related to the two genomes of the actual tetraploid, the distance between the synthetic tetraploid and an actual tetraploid should be less than the distance

**Fig. 6.** A computer-generated dendrogram of only diploid species in genus *Oryza* based on RFLPs detected with 25 genomic clones. *Left side, first column* – RFLP classification; *second column* – genome designation; *third column* – accession number (see Table 2). Genetic distance is according to Nei (Nei 1987)

**Table 5.** The smallest distances of actual tetraploid accessions with synthetic tetraploid accessions (diploid + diploid) and the distances with the corresponding diploid accessions

| Tetraploid species | Accessions | | | |
|---|---|---|---|---|
| | | BB<br>*O. punctata*<br>(BB)<br>W1515 | CC<br>*O. officinalis*<br>100896 | BB+CC |
| *O. minuta* (BBCC) | 101089 | 0.041318 | 0.039801 | 0.015921 |
| *O. minuta* (BBCC) | 101125 | 0.044487 | 0.039801 | 0.018863 |
| *O. minuta* (BBCC) | 101141 | 0.047851 | 0.036956 | 0.017369 |
| *O. minuta* (BBCC) | 103865 | 0.047085 | 0.036227 | 0.018322 |
| | | X=0.044 | X=0.038 | X=0.0177 |
| | | BB<br>*O. punctata*<br>(BB)<br>W1590 | CC<br>*O. eichingeri*<br>IP7 | BB+CC |
| *O. punctata* (BBCC) | IP27 | 0.02329 | 0.022545 | 0.006821 |
| *O. punctata* (BBCC) | W1408 | 0.017903 | 0.02422 | 0.005534 |
| *O. punctata* (BBCC) | 101409 | 0.018979 | 0.022545 | 0.004414 |
| | | X=0.0200 | X=0.0231 | X=0.0056 |
| | | BB<br>*O. punctata*<br>(BB)<br>W1590 | CC<br>*O. collina*<br>103410 | BB+CC |
| *O. malampuzhaensis* (BBCC) | 100957 | 0.032465 | 0.031689 | 0.016509 |
| | | CC<br>*O. eichingeri*<br>101422 | EE<br>*O. australiensis*<br>OA4 | CC+EE |
| *O. alta* (CCDD) | 101395 | 0.041397 | 0.047085 | 0.032006 |
| *O. alta* (CCDD) | W1144 | 0.046389 | 0.055255 | 0.036436 |
| *O. grandiglumis* (CCDD) | 101405 | 0.039801 | 0.049357 | 0.032998 |
| *O. latifolia* (CCDD) | 100914 | 0.04426 | 0.042182 | 0.028575 |
| *O. latifolia* (CCDD) | W1168 | 0.047491 | 0.054591 | 0.038511 |
| *O. latifolia* (CCDD) | 100963 | 0.039801 | 0.039801 | 0.025528 |
| | | X=0.043 | X=0.047 | X=0.0323 |

between the actual tetraploid and each of the diploid counterparts. Table 5 lists the smallest distance between the synthetic tetraploids and actual tetraploids as well as the distances of the actual tetraploid with each of the diploid counterparts. The extant ancestral diploid species most closely related to the progenitors of tetraploid species deduced from distances of actual diploids and tetraploids were confirmed from distances of synthetic and actual tetraploids.

*O. minuta/O. punctata/O. malampuzhaensis (BBCC species).* *O. minuta* is endemic to The Phillipines and, based on chromosome pairing in crosses to diploids, contains both the B and C genomes (Chang 1964; Katayama 1967; Nayar 1973). We found that the closest related diploid accessions are *O. officinalis* (100896) (C genome) from Thailand and three B genome accessions of *O. punctata* from Africa (Table 5). The mean genetic distance between tetraploid *O. minuta* accessions and *O. punctata* was 0.044, whereas with *O. officinalis* the average distance was 0.038. With the synthetic tetraploid (comprised of the *O. punctata* and *O. officinalis* accessions), the genetic distance to *O. minuta* dropped to 0.017. This latter distance is comparable to the average genetic distance observed between diploid accessions of the same species. This small value suggests two things: (1) that the identified diploid accessions are likely to be close relatives of the accessions that created the actual tetraploid and (2) that the tetraploid *O. minuta* has not diverged much since the original polyploidization event.

The *O. officinalis* accession identified as the most closely related to *O. minuta* is from Thailand, which is consistent with the likely origin of *O. minuta* in Asia. However, the other diploid accession, *O. punctata*, is found only in Africa. A possible explanation of this result is that African *O. punctata* is a close relative of a diploid B genome species now extinct in Asia.

Similarly, the closest diploid accessions to *O. malampuzhaensis* are *O. punctata* (B genome) (accession W1590) from Africa and *O. collina* (CC) (103410) from Sri Lanka (Table 5). The distance of *O. malampuzhaensis* with the synthetic tetraploid (0.0165) was much less than with either of the closest diploids (0.03245 and 0.0317). Thus, these two diploid accessions are likely to be the close relatives of the actual accessions that created *O. malampuzhaensis*.

Unlike the case with *O. minuta* and *O. malampuzhaensis*, the two closest diploid accessions to *O. punctata* (BBCC) are both from Africa: *O. punctata* (B genome) (W1590) and *O. eichingeri* (C genome) (IP 7). The average distance of *O. punctata* with the synthetic tetraploid (0.0059) was much less than with either of the corresponding diploids (0.0201 and 0.02310).

In summary, while the closest B genome diploid relative (*O. punctata*) to three BBCC tetraploid species is found in Africa, the closest C genome diploid accessions are found in different locations and close to the locations of the corresponding BBCC species. This suggests that there is a possibility of independent origins of C genome of BBCC species or that introgression has been occurring

between BBCC tetraploids and their sympatric C genome relatives.

*O. latifolia*/*O. grandiglumis*/*O. alta (CCDD species).* CCDD species are found only in America. Chromosome pairing has identified the C genome as one of the likely components of this species. However, the other genomic component is unknown and has been assigned the tentative 'D' genome symbol (Chang 1964; Katayama 1982; Nayar 1973).

Based on genetic distances, two diploid accessions were identified as being most closely related to all three CCDD species (Table 5). *O. eichingeri* (C genome diploid from Africa) and *O. australiensis* (E genome diploid from Australia) gave average genetic distances with the tetraploids of 0.043 and 0.047, respectively. The lower value obtained with *O. eichingeri* probably reflects a closer affinity with the C genome than with the E genome, a result consistent with chromosome pairing studies (Chang 1964; Nayar 1973). Synthetic tetraploids, constructed from these two diploid accessions, gave genetic distances (with CCDD species) only *slightly* smaller (0.032) than those obtained with diploid/tetraploid comparisons (0.043 and 0.047) (Table 5). Since decreases in distances of CCDD tetraploids with synthetic tetraploids, compared with the corresponding diploids, were not as dramatic as those observed with other tetraploids (e.g., *O. minuta*, see previous section), we conclude that while *O. eichingeri* (C genome) and *O. australiensis* (E genome) are the closest living relatives (reported thus far) of the CCDD tetraploids, neither is very closely related to the ancestral diploids that generated this species.

Table 6 showed the smallest genetic distances of tetraploid species to each other and to the synthetic tetraploids. It is obvious that CCDD tetraploid species are more closely related to each other than to any diploid or synthetic tetraploid (Table 5 and 6). This raises the possibility that all CCDD species were derived from a single polyploidization event. If this is correct, the event likely occurred long ago since there is now considerable geographic divergence in the CCDD tetraploids. Accessions from North America (Mexico and Cuba) show significant genetic differentiation from those in South America (Fig. 3). Sufficient time may have elapsed since original polyploidization events that the real diploid progenitors are either extinct or have diverged considerably.

## Conclusions

Using 25 single-copy RFLP probes we have examined genetic variation and relationships in the genus *Oryza*. From this study we derive several conclusions. First, most accessions of wild species are pure lines or mixed pure lines, and heterozygosity within accessions is rare. Tetraploid species (such as *O. latifolia* and *O. minuta*) harbor 1.5 times more alleles per plant than their diploid counterparts, a result consistent with the allopolyploid nature of these species. Using genetic relatedness as criterion, it was possible to identify the closest living diploid

**Table 6.** Most closely related extant tetraploid species (center) and synthetic tetraploid (diploid + diploid) species (right) to selected tetraploid species (left). Affinities are based on genetic distances (see Table 5)

| BBCC | Closest BBCC species | Closest BB + CC species |
|---|---|---|
| *O. minuta*<br>(Philippines) | *O. malampuzhaensis*<br>(India)<br>0.0214 | *O. punctata* (W1515) + *O. officinalis* (100896)<br>(Tanzania) (Thailand)<br>0.0142 |
| *O. punctata*<br>(Africa) | *O. malampuzhaensis*<br>(India)<br>0.0177 | *O. punctata* (W1590) + *O. eichingeri* (IP7)<br>(Cameroun) (Ivory Coast)<br>0.0044 |
| *O. malampuzhaensis*<br>(South India) | *O. punctata* (W1408)<br>(Nigeria)<br>0.0177 | *O. punctata* (W1590) + *O. collina* (103410)<br>(Cameroun) (Sri Lanka)<br>0.0165 |
| **CCDD** | **Closest CCDD species** | **Closest CC + EE species** |
| *O. alta*<br>(America) | *O. grandiglumis*<br>(Brazil)<br>0.0118 | *O. eichingeri* (101422) + *O. australiensis* (OA4)<br>(Uganda) (Australia)<br>0.032 |
| *O. grandiglumis*<br>(Brazil) | *O. alta*<br>(America)<br>0.0118 | *O. eichingeri* (101422) + *O. australiensis* (OA4)<br>(Uganda) (Australia)<br>0.033 |
| *O. latifolia*<br>(America) | *O. grandiglumis*<br>(Brazil)<br>0.017 | *O. eichingeri* (101422) + *O. australiensis* (OA4)<br>(Uganda) (Australia)<br>0.0255 |

relatives of the currently known tetraploid rice species. We conclude from these analyses that BBCC tetraploids (*O. malampuzhaensis*, *O. punctata* and *O. minuta*) are likely to be of independent origins or, if they originated from a single ancestral tetraploid species, introgression has been occurring between these tetraploids and sympatric C genome relatives. CCDD tetraploids (*O. latifolia*, *O. alta* and *O. grandiglumis*) species have a closer affinity to each other than to any known diploid species suggesting the possibility of an ancient origin for these American rice species. The closest living diploid relatives (to CCDD tetraploids) were identified as those containing the C genome and the E genome (*O. australiensis*).

Dendrograms constructed for the entire genus reveal four species complexes which correspond to those proposed by Vaughan (1989): the *O. sativa* complex, the *O. officinalis* complex, the *O. ridleyi* complex and the *O. meyeriana* complex. Within the *O. sativa* complex, *O. sativa* and *O. rufipogon* (including *O. nivara*) accessions form an intertwined group. *O. sativa* spp. japonica and *O. sativa* spp. indica showed closer affinity with different accessions of *O. rufipogon* than with each other, suggesting the possibility that these subspecies reflect independent domestication events.

Finally, the RFLP markers used in this study represent a subset of those included on a complete rice RFLP map (McCouch et al. 1988; Causse et al. in preparation). The majority of these RFLP probes reveal polymorphism between cultivated rice and wild rice accessions, indicating that they will be useful for monitoring and facilitating introgression of new genes from wild species into modern cultivars.

*Acknowledgements.* We would like to thank Dr. Bonierbale M. W. for reviewing this paper. The authors are grateful to International Rice Research Institute and ORSTOM for the plant material they provided and to The Rockefeller Foundation for the financial support of this project.

## References

Chang TT (1964) Present knowledge of rice genetics and cytogenetics. Tech Bull 1 Int Rice Res Inst

Chang TT, Qu SH, Pathak MD, Ling KC, Kauffman HE (1975) The search for disease and insect resistance in rice germplasm. In: Frankel OH, Hawkes JG (eds) Crop genetic resources for today and tomorrow. Cambridge University Press, Cambridge, pp 183–200

Chao S, Sharp PJ, Gale MD (1988) A linkage map of wheat homoeologous group 7 chromosomes using RFLP markers. In: Miller TE, Koebner RMD (eds) Wheat Genet Symp, vol 1. Cambridge, England, pp 493–498

Clayton WD, Renvoize SA (1986) Genera *graminum*. Grasses of the world. Kew Bulletin Series XIII. London, Her Majesty's Stationary Office

Coffman WR, Herrera RM (1980) Rice. In: Hybridization of crop plants. American Society of Agronomy-Crop Science of America, 28 Wisc., pp 101–131

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

Gale MD, Sharp PJ (1988) Genetic markers in wheat – developments and prospects. In: Miller TE, Koebner RMD (eds) Proc 7th Int Wheat Genet Symp, vol 1. Cambridge, England, pp 469–475

Ghesquière A (1988) Diversité génétique de l'espèce sauvage de riz, *Oryza longistaminata* A. Chev. & Roehr et dynamique des flux géniques au sein du groupe Sativa. en afrique. Thèse Drès-Sciences, Université, Paris XI, Orsay

Hu Chao-Hwa (1970) Cytogenetic studies of *Oryza officinalis* complex. III. The genomic constitution of *O. punctata* and *O. eichingeri*. Cytologia 35:304–318

Jena KK, Khush GS (1990) Introgression of genes from *Oryza officinalis* Well ex WAtt to cultivated rice, *O. sativa* L. Theor Appl Genet 80:737–745

Katayama T (1967) Cytogenetical studies on *Oryza* $F_1$ hybrids of the crosses BB × CC, BBCC × a diploid strain of *O. punctata* and CC × a diploid strain of *O. punctata*. Proc Jpn Acad 43:327–331

Katayama T (1982) Cytogenetical studies on the genus *Oryza*. XIII. Relationship between the genomes E and D. Jpn J Genet 57:613–621

Kihara H (1963) Genome analysis in *Oryza*. In: Symp on rice problems. Tenth Pacific Sci Congr Pacific Sci Assoc, University of Hawaii, Honolulu, pp 57–66

McCouch SR, Kochert G, Yu ZH, Wang ZY, Khush GS, Coffman WR, Tanksley SD (1988) Molecular mapping of rice chromosomes. Theor Appl Genet 76:815–829

Miller JC, Tanksley SD (1990) RFLP analysis of phylogenic relationships and genetic variation in the genus *Lycopersicon*. Theor Appl Genet 80:437–448

Morinaga T (1964) Cytogenetical investigation on *Oryza* species. In: IRRI (ed) Rice genetics and cytogenetics. Elsevier, Amsterdam, pp 91–103

Morishima H, Oka H (1960) The pattern of interspecific variation in the genus *Oryza*: its quantitative representation by statistical methods. Evolution 14:153–165

Morishima H, Hinata K, Oka HI (1963) Comparison of modes of evolution of cultivated forms from two wild species, *Oryza breviligulata* and *O. perennis*. Evolution 17:170–181

Morishima H, Sano Y, Oka HI (1984) Differentiation of perennial and annual types due to habitat conditions in the wild rice *Oryza perennis*. Plant Syst Evol 144:119–135

Nayar NM (1973) Origin and cytogenetics of rice. Advances in genetics, vol 17. Academic Press, New York London, pp 153–292

Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York, pp 106–107

Ng NQ, Hawkes JG, Williams JT, Chang TT (1981) The recognition of a new species of rice (*Oryza*) from Australia. Bot J Linn Soc 82:327–330

Oka HI (1974) Experimental studies on the origin of cultivated rice. Genetics 78:475–486

Oka HI (1988) Origin of cultivated rice. Jpn Sci Soc Press, Amsterdam New York

Second G (1982) Origin of the genic diversity of cultivated rice (*Oryza* spp.): study of the polymorphism scored at 40 isozyme loci. Jpn J Genet 57:25–57

Second G (1984) Relations évolutives chez le genre *Oryza* et processus de domestication des riz. These Doctorat d'Etat, Université Paris-Sud, Orsay

Second G (1985a) Evolutionary relationships in the *Sativa* group of *Oryza* based on isozyme data. Genet Sel Evol 17:89–114

Second G (1985 b) A new insight into the genome differentiation in *Oryza* L. through isozymic studies. In: Sharma AR, Sharma A (eds) Advances in chromosome and cell genetics. Oxford & IBH Pub, New Delhi Bombay Calcutta, pp 75

Sharma SD (1983) Conservation of wild species of *Oryza*. In: Rice germplasm conservation workshop. The International Rice Research Institute and the International Board for Plant Genetic Resources, 21–25

Sitch LA, Dalmacio RD, Romero GO (1989) Crossability of wild *Oryza* species and their potential use for improvement of cultivated rice. Rice Genet Lett 6:58–60

Sokal RR, Michener CD (1958) A statistic method for evaluating systematic relationships. Univ Kansas Sci Bull 28:1409–1438

Song KM, Osborn TC, William PH (1988) Brassica taxonomy based on nuclear restriction fragment length polymorphisms (RFLPs) Theor Appl Genet 75:784–794

Song KM, Osborn TC, William PH (1990) Brassica taxonomy based on nuclear restriction fragment length polymorphisms (RFLPs) 3. Genome relationships in Brassica and related genera and the origin of *B. oleracea* and *B. rapa* (syn. *campestris*). Theor Appl Genet 79:497–506

Tanksley SD, Young ND, Paterson AH, Bonierbale MW (1989) RFLP mapping in plant breeding – new tools for an old science. Biotechnology Bio/technology 7:257–264

Tateoka T (1962) Taxonomic studies of *Oryza* II. Several species complex. Bot Mag Tokyo 75:455–461

Vaughan DA (1989) The genus *Oryza* L. current status of taxonomy. IRRI res paper ser no 138

Vaughan DA (1990) The relatives of rice in Papua New Guinea. Report of collaborative germplasm collecting in Papua New Guinea between the Department of Primary Industry and IRRI. IRRI, Manila

Wang ZY, Tanksley SD (1989) Restriction fragment length polymorphism in *Oryza sativa* L. Genome 32:1113–1118