Theor Appl Genet (1990) 80:209–222

THEORETICAL AND APPLIED GENETICS

© Springer-Verlag 1990

Chloroplast DNA diversity in wild and cultivated species of rice (Genus *Oryza*, section Oryza). Cladistic-mutation and genetic-distance analysis

A. M. Dally* and G. Second**

French Institute of Scientific Research for Development in Cooperation (ORSTOM, Montpellier Center, BP 5045, F-34032 Montpellier Cedex, France

Received June 29, 1989; Accepted January 22, 1990
Communicated by K. Tsunewaki

Summary. Using a novel nonaqueous procedure, chloroplast DNA was isolated from 318 individual adult rice plants, representing 247 accessions and the breadth of the diversity in section Oryza of genus *Oryza*. Among them, 32 different cpDNA restriction patterns were distinguished using the restriction endonucleases EcoRI and AvaI, and they were further characterized by restriction with BamHI, HindIII, SmaI, PstI, and BstEII enzymes. The differences in the electrophoretic band patterns were parsimoniously interpreted as being the result of 110 mutations, including 47 restriction site mutations. The relationships between band patterns were studied by a cladistic analysis based on shared mutations and by the computation of genetic distances based on shared bands. The deduced relationships were compared with earlier taxonomical studies. The maternal parents for BC genome allotetraploids were deduced. Within species, cpDNA diversity was found larger in those species with an evolutionary history of recent introgression and/or allotetraploidization. Occasional paternal inheritance and recombination of cpDNA in rice was suggested.

Key words: *Oryza* – rice cpDNA diversity – CpDNA inheritance – Allotetraploid progenitors – Cladistic mutation

Introduction

Studies of chloroplast DNA (cpDNA) diversity have been used on several occasions to determine the taxonomy and to retrace the phylogeny within particular groups of plants. Recent examples of such studies are those of Doebley (1987) on corn, of Lehväslaiho et al. (1987) on Festuceae, of Palmer et al. (1987) on legumes, of Ogihara and Tsunewaki (1988) on wheat, of Ishii et al. (1988) on rice, of Hosaka and Hanneman (1988) on potatoes, and of Doyle and Doyle (1990) on wild perennial soybean.

Extensive comparative studies using RFLPs of total cpDNA are quite rare. This is mainly because of the low efficiency of most cpDNA isolation procedures for many types of plants, making it necessary either to process large amounts of material or to use Southern hybridization on total cell DNA (e.g., Palmer 1985). Also, the DNA usually has to be extracted from very young tissues so that it is not possible to study individual plants.

To overcome the problems of cpDNA isolation, we have modified a nonaqueous method to isolate rice chloroplasts so that enough cpDNA could be obtained from single plants for analysis (Dally and Second 1989).

We present here the results of an extensive comparison of cpDNA restriction fragment banding patterns visualized by ethidium bromide staining after agarose gel electrophoresis. The cpDNAs from 318 individual plants representing 247 accessions in the 13 species that comprise the Oryza section were investigated.

Oryza, section Oryza, corresponds exactly to the section Sativa of Roschevicz (1931), later named Eu-Oryza by Roshevits (1937), and also corresponds exactly to the section Sativae of Tateoka (1964). As retained here, its name follows the international code for botanical nomenclature. It includes two cultivated species (*Oryza sativa* of Asiatic origin and *O. glaberrima* of African origin) and approximately 11 wild species as listed in Table 1.

Cytogenetic observations of chromosome pairing distinguished five genomes in section Oryza (Nayar 1973; Katayama 1982). Genome A is found in all cultivated

* Present address: Institute for the Developmental and Molecular Biology of Plants, Universitätsstr. 1, D-4000 Düsseldorf, FRG
** To whom offprint request should be addressed

Table 1. List of species in *Oryza* section Oryza (revised from Second 1985b)

	Geographic repartition	Life form[a]	Rep. syst.[b]	Gen. symb.[c]
The SATIVA group				
Cultivated species:				
O. sativa L. (*japonica* & *indica* subsp. Kato)	Asian origin, worldwide	I	S&I	A
O. glaberrima Steud	Africa (& S. America)	A	S	A
Wild species:				
O. rufipogon Griff[d] (complex species)	Tropical Asia Australia & America	A, I&P	S&I	A
O. longistaminata A. Chev. & Roehr.	Africa	P	O	A
O. breviligulata A. Chev. & Roehr. (=*O. barthii* A. Chev.)	Africa	A	S	A
The LATIFOLIA group				
O. officinalis[e] Wall. ex Watt (Complex species)	Asia, New Guinea	P	S	C
O. eichingeri A. Peter	Africa	P	S	C
O. alta Swallen *O. grandiglumis* (Doll.) Prod *O. latifolia* Desv.	Tropical	P	S	CD
O. punctata Kotschy	Africa	A(2×)	S	B
(2× & 4× species)		P(4×)	S	BC
O. minuta J.S. Presl	Africa	P	S	BC
O. australiensis Domin	Tropical arid Australia	A	S	E

[a] Life forms: A – annual; P – perennial; I – intermediate

[b] Reproductive system: S – selfing largely predominant: O – self incompatible; I – intermediate

[c] Cytogenetic genome symbols (Nayar 1973), A, B, C, and E are diploid (2× =24), BC and CD are allotetraploid (2× =48)

[d] *O. glumaepatula* Steud is sometimes used for (all?) American forms. *O. meridionalis* Ng et al. was proposed for (all?) Australian forms (Chang 1988). Many other names were also given (Oka 1988)

[e] *O. collina* (Trimen) S.D. Sharma & Shastry was proposed for forms originated in Sri Lanka. A tetraploid form found in India and similar to tetraploid *O. punctata* is called *O. malampuzahensis*. Krishnaw. & Chandra. This name was not retained by Tateoka (1963)

forms and their closest wild relatives and only in diploids. Genomes B and C are found in diploid and allotetraploid BC and CD species. Genome D is not known at the diploid level. Genome E is known only at the diploid level.

Extensive isozyme surveys of the genetic diversity in rice (Second 1982, 1984, 1985a, b; Glaszmann 1987) have shown the validity of recognizing two main types in *O. sativa* (these are the *indica* and *japonica* subspecies of Kato 1930) with intermediates, and has allowed a comprehensive treatment of the taxonomy and phylogenetic relationships among the section *Oryza*. Two groups of species, both with a pan-tropical distribution, were recognized in this section, according to their relatedness as determined by isozyme analysis. The Sativa group comprises all A-genome species. The Latifolia group comprises B, C, E, BC, and CD-genome species. Table 1 summarizes the available information on the geographic origin, life forms, reproductive systems, and genome constitution of the species in section Sativa.

As most of the plants and lines analyzed for cpDNA variation had been analyzed earlier for polymorphism of isozymes, a direct comparison of the information obtained in both ways is possible. In this paper, we limit the comparison mostly to that at phenetic and taxonomic levels.

Materials and methods

Plant materials

The plants were grown from seeds in the greenhouse. The complete list of accession numbers is given by Dally (1988). Accession numbers of the original collections of the National Institute of Genetics (NIG) in Japan, the French Institute of Scientific Research for Development in Cooperation (ORSTOM), and the Institute of Agronomical and Tropical Research (IRAT) in France, and the International Rice Research Institute (IRRI) in the Philippines were maintained in this paper. Most accessions of wild species from Africa, Australia, and West India were grown from seeds collected from the natural populations by ORSTOM in the breadth of their distribution in these areas. Others were obtained from lines multiplied previously from the original collections (for the wild species, most of them were obtained from the NIG collection and were originally collected from their natural habitat by this institute).

cpDNA isolation, digestion, fragment electrophoresis

Chloroplasts were isolated from freeze-dried, pulverized leaves of individual adult or juvenile rice plants and DNA was extracted using the method described by Dally and Second (1989). The method was based on that developed for wheat seedlings by Bowman and Dyer (1982). A change from this procedure was the use of less toxic organic solvents, which allowed mechanical blending of the leaves. Other modifications included a pelleting of the contaminant nuclear DNA. The average yield of cpDNA obtained from 12 g leaves was about 10 μg.

After digestion of the DNA, electrophoresis was carried out in submarine, horizontal agarose. Details on the procedure are in Dally (1988).

Determination of the sizes of restriction fragments

In addition to making a graphic reconstruction of each electrophoretic pattern obtained, the distance migrated by each fragment was measured to within 0.2 mm, using photographs of the original gel size.

Slightly sigmoid calibration curves, relating the migrated distances of electrophoretic bands to hyperbolic transformations of the fragment sizes, were deduced from a fine set of molecular weight markers ("Raoul I", Appligene) electrophoresed alongside the samples. The applied hyperbolic functions of the type $L = a:(M+b)+c$ (where L is the length of a fragment and M its migrated distance), proposed by Rickwood

and Hames (1982), were defined by three selected bands of the marker, two of them enclosing just the range of the bands in the pattern to be analyzed and the third one close to the middle of the pattern.

Analysis of mutations

The different electrophoretic types were grouped together in a hierarchy of similarity. In the first step, the differences between pairs of patterns that were most similar to each other were interpreted by assuming the smallest possible number of mutations to have taken place (no more than two in practice). By comparison with all other patterns, the primitive and the derived state were determined for each mutation.

In the subsequent steps, each group previously formed was identified with its primitive (hypothetical) type. The most similar types (hypothetical ones or remaining original ones) were again grouped in pairs, and the differences within these pairs were presumed to result from the smallest possible number of likely mutations. When a large number of band differences had to be interpreted simultaneously, we sought to explain the differences by the least possible number of mutations. When possible, mutations were presumed to be located on the inverted repeats (IRs) and were counted only once.

In order to define the differences between two patterns of fragments, restriction site mutations were presumed to have occurred when two bands replaced one band and the sizes of the replacements together equalled the size of the original (to within 0.01 kb), or vice versa. Small additions or deletions were presumed to have taken place when shifts of bands corresponding to size differences of 0.04 kb or less were found. Large additions or deletions were defined as those causing band displacements corresponding to fragment size differences of 0.05 kb or more and when these shifts were found with at least two enzymes. If these size differences were not found with at least two enzymes, they were interpreted as being due to site mutations with one of the fragments involved too small to be detected in the gel. An inversion was presumed to have taken place when two fragments replaced two other fragments, with the respective sums of their sizes being equal.

Genetic distances

Genetic distances were computed according to the method of Nei and Li (in Nei 1987, pp 106–107), based on the presence or absence of size-characterized fragments. When a band was assumed to comprise more than one fragment, each fragment was considered individually. Either distances were computed for each enzyme separately, using all fragments, or patterns of the five enzymes for which all plastotypes had been analyzed (EcoRI, AvaI, BamHI, PstI, SmaI) were combined in the following way: based on the mutation analysis, the bands considered to represent the same addition/deletion or inversion event were counted only once and only the largest of the two subfragments involved in a site substitution event was considered. The effect of this was to eliminate the SmaI patterns because the SmaI recognition sequence partially coincides with AvaI recognition sequence, to render negligible the additional information contained in HindIII and BstEII restriction patterns (for which some data were missing) and also the contribution of fragments smaller than 1 kb.

Results

Interpretation of electrophoregrams

On average, EcoRI generates 57 fragments longer than 400 bp, and AvaI, 60 fragments of 280 bp or more. The number of fragments produced longer than 600 bp was 34 for BamHI and HindIII, 16 for SmaI, 15 for PstI, and 14 for BstEII.

Most of the 318 individual plants were studied with both EcoRI and AvaI, and 32 chloroplast genotypes ("plastotypes") named a1–q2 were distinguished. A reduced number of plants was further studied with BamHI, HindIII, SmaI, PstI, and BstEII as well, with representatives of all plastotypes generally studied with all enzymes (exceptions included plastotypes d, e2 to e5, f2, h2, i2, m3, n3, p1, p2, q1 and q2 for HindIII and plastotypes d, e2 to e5, f2, h2, i2, m3, n3, p1, p2 and q1 for BstEII). Details of the restriction patterns of all the accessions are described in Dally (1988).

Distribution of the plastotypes among the taxa of wild and cultivated species

The distribution of the observed plastotypes among the taxa of wild and cultivated species of rice is shown in Table 2 and is further detailed below.

Cultivated rice. The two species of cultivated rice were shown to have different plastotypes. Nine different plastotypes were observed among 65 *O. sativa* cultivars but only one plastotype was found in the six *O. glaberrima* cultivars studied. The *indica* and *japonica* types of *O. sativa* had generally different plastotypes.

For further details, the classification in groups of *O. sativa* based on nuclear encoded isozyme markers (Glaszmann 1987) was adopted. Group VI, which included temperate and tropical *japonica* cultivars, usually had the e1 plastotype. Exceptions were observed in only the most cold-resistant cultivar, found in Madagascar (cv "Latsika"), and in one cultivar from Thailand (cv "Howm Om"), which had unique plastotypes e2 and e4, respectively. Group I, closest to typical *indica* cultivars, usually had an a1 plastotype. Exceptions included an upland *indica* cultivar from Thailand (cv "Dao Pao" with a plastotype c), an African cultivar (with plastotype b), and also two cultivars from Africa and Madagascar (with the plastotype e1 usually found in *japonica* cultivars). Intermediate *indica/japonica* cultivars (including groups II, V, and the somewhat divergent and rare, deep-water cultivars of groups III and IV, as well as other intermediates) usually shared the e1 plastotype of the *japonica* type. Exceptions included six group II cultivars. Two are from Iran with a unique plastotype e3, and the others are from Africa with plastotype a (generally found in *indica* cultivars).

Finally, a primitive landrace, which originated in West India and has a strong affinity with *O. glaberrima* at isozyme level but a long ligule as *O. sativa* (Lolo and Second 1988), had an e5 plastotype.

Table 2. Number of *Oryza* accessions having the different plastotypes, classified by taxa

Cultivars *O. sativa* (classified according to Glaszman 1987)	Plastotypes																
	a[a]	a1	a2	b	c	d	e1	e2	e3	e4	e5	f1	f2	g	h1	h2	i1	i2
Group I (*indica*)	3	12		1	1													
Group II	1	3					4		2									
Group III							1											
Group IV							2											
Group V							3											
Group VI (*japonica*)							22	1		1								
Intermediate I–VI							4											
Group I, CMS[b]		3	9															
Other[c]											1							
O. glaberrima																	6	
Wild species, Sativa group																		
O. rufipogon, China	1	1	1				7											
South Asia		1	5	2	1		1											
Weast India							2				2							
America						1	1								1	1		
Australia												9	1					
O. breviligulata																	8	1
O. longistaminata[d]														18				

	Plastotypes															
	k	l/m[a]	l	m1	m2	m3	m4	n1	n2	n3	o	p1	p2	p3	q1	q2
Latifolia group																
O. punctata (diploid)	6															
O. punctata (tetraploid)		1	7	1												
O. officinalis				12	9	1	2									
O. malampuzahensis	1															
O. minuta	2															
O. eichingeri								3	1	1						
O. latifolia											5	4	2			
O. alta													2			
O. grandiglumis														3		
O. australiensis															3	2

[a] The AvaI restriction pattern was not analyzed and the plastotypes a1 and a2 or l, m1, m2, m3, and m4 could not be distinguished
[b] Number of plants, not accessions. In addition, one plant showed a composite plastotype a1 and a2; see text
[c] A landrace from Gujarat, western India, closely related to a local form of *O. rufipogon* with particular affinity to *O. glaberrima* and *O. brevilligulata* at isozyme level (Lolo and Second 1988)
[d] Including five spontaneous derivatives from *O. longistaminata* × *O. sativa* hybrids "Obake" (Ghesquiere 1988)

Male sterility inducing cytoplasm. Thirteen cytoplasmic male-sterile (cms) plants representing 11 different lines of the *indica* type and at least the WA and Gambiaca sources of cms (i.e., the two main sources of cms known in *indica* rice, Virmani et al. 1986) were also analyzed. As shown in Table 2, most of them had an a2 plastotype but the WA cytoplasm was found heterogeneous with plants carrying either an a1 or a2 plastotype within a particular line. One plant even carried a mixture of these two plastotypes (see below).

The cultivar "Chinsurah Boro II", which is the source of cms in some *japonica* cultivars (Shinjyo 1975), was analyzed from two different accessions and showed an e1 plastotype, as do other isozyme group II cultivars from Asia.

Wild species of the Sativa group. Oryza longistaminata was represented by 13 plants derived from seeds collected directly in nine populations, within a large part of its distribution in intertropical Africa and Madagascar. In addition, five hybrid derivatives of spontaneous *O. longistaminata* × *O. sativa* crosses ("Obake", Ghesquiere 1988) were also included. The 18 plants shared the same plastotype, g.

O. breviligulata was represented by 15 plants issued from nine wild or weedy populations originating in intertropical Africa. One population from Botswana had the i2 plastotype; all others (including a population from Zambia) shared the i1 plastotype with *O. glaberrima.*

O. rufipogon is a complex species with annual, perennial, and intermediate life history forms and contains

more or less divergent Asian, American, and Australian geographic forms (Second 1985a). The Asian form shows a trend of differentiation at the isozyme level between strains from China and from South/Southeast Asia, which parallels the *japonica-indica* differentiation of cultivated rice. A peculiar annual form with affinity to *O. breviligulata* at the isozyme level is also found (Lolo and Second 1988). As detailed in Table 2, all the main geographic forms were represented (except, however, strains from New Guinea with affinity to the Australian form, Second 1985a). The various life forms were also represented, although the number of accessions was too limited to make any comparison at this level. Australian accessions included both annual and perennial life forms.

Six plastotypes were found among 25 accessions of the Asian form. Most South Asian strains (five out of ten) had the plastotype a2, as found in the *indica* cms cultivars. Most Chinese strains (seven out of nine, including those strains collected in areas isolated from cultivated rice) had the e1 plastotype of most *japonica* cultivars. All other plastotypes found in Asian *O. rufipogon* were also observed in cultivated rice. On the other hand, rare plastotypes e2, e3, and e4 found in *O. sativa* were not observed in our samples of *O. rufipogon*.

The American form of *O. rufipogon* was most variable, with the four accessions studied each having a different plastotype, among which three were peculiar to them. These were d, h1, and h2 and the fourth one was e1.

The Australian form of *O. rufipogon* had two unique plastotypes: f1 and f2. The most common one was shared by the annual and the perennial life forms.

Wild species of the Latifolia group. The complex species *O. officinalis* (genome C) was represented by 24 accessions covering most of its distribution area (including China but not New Guinea or Sri Lanka). Four specific different plastotypes, m1, m2, m3, and m4, were distinguished. Plastotypes m1 and m2 were observed most frequently, m1 in accessions originating from most parts of the distribution, including one accession from China, m2 only in accessions from Borneo.

The diploid *O. eichingeri* (genome C) a rare African species, was represented by three accessions from Uganda and two accessions from the Ivory Coast. Three plastotypes, n1, n2, and n3, were distinguished, one in Uganda and two in the Ivory Coast accessions. The largest deletion of 400 bp detected in this study characterized plastotype n3.

The diploid form of *O. punctata* (genome B) was represented by 12 plants obtained from seeds directly collected in six populations from Nigeria, Cameroon, Chad, and Tanzania. They all shared a unique plastotype, k.

The tetraploid form of *O. punctata* (genome BC) was represented by nine accessions originally collected in West Africa. Seven of them showed the plastotype l. One of them (from the Ivory Coast) showed a plastotype m1, and another one was not analyzed with AvaI to determine which of I or m it was.

Other allotetraploids with genome BC include *O. malampuzahensis* from India (acc. W1159) and *O. minuta* from the Philippines. Both shared plastotype k observed in diploid *O. punctata*. (Some other lines of these two species have been reported earlier to be indistinguishable from tetraploid *O. punctata* for isozyme, (Second 1984, 1985b) and cpDNA diversity (Dally 1988). Latter information indicates, however, that these lines correspond to some mislabellings in the original collection).

The genome BC thus comprises forms with plastotypes indistinguishable or very similar to either of the two parental B and C genomes.

The tetraploid American species of the Latifolia group are considered to form the *O. latifolia* complex (genome CD), although three or more species are often distinguished (Tateoka 1963). Among those available, 16 accessions were studied. Four plastotypes, o, p1, p2, and p3, were distinguished. Plastotype p3 was specific to the species *O. grandiglumis* but the three others were found in *O. latifolia* (sensu stricto). *Oryza alta* shared plastotype p2 with *O. latifolia*.

Finally, the diploid Australian species *O. australiensis* (genome EE) was represented by five accessions sampled in the range of its distribution in northern Australia. Two plastotypes, q1 and q2, were found with no apparent geographic segregation.

Mutation analysis

Table 3 characterizes the presumed mutations, following a hierarchy of similarity of the plastotypes differentiated by those mutations. The mutations are numbered from 1 to 110. Between the groups of plastotypes a to i and k to q, the primitive state of the mutation is arbitrarily attributed to the Latifolia group of species, because of a missing reference taxa out of the Sativa section.

Indication that the root of this cladistic system is situated between the Sativa and Latifolia species groups came from a preliminary cpDNA RFLP study at the level of the Oryzeae tribe with the enzyme BspXI. Two bands were specific to the representatives of the Latifolia group (including *O. australiensis*, type q). Another result from this study, shown in Table 3, is mutation 57: one band is found only in plastotypes f, g, h, and i.

There are altogether 47 presumed restriction site mutations (among which, both subfragments involved in the mutation were observed for 35 of these), 62 addition/deletion (46 "small" and 16 "large"), and one presumed inversion.

Table 3. Presumed mutations observed in cpDNA of *Oryza*

No.	Enzyme	Fragments lost	Fragments gained	Mutation type[a]	Carrier plastotype	No.	Enzyme	Fragments lost	Fragments gained	Mutation type[a]	Carrier plastotype
1	EcoRI	3.74	3.79	sa	b	48	EcoRI	2.9	2.61	ld	l–n
	AvaI	0.86	0.89				AvaI	5.1	4.82		
2	EcoRI	3.74	3.70	sd	c		BamHI	6.94	6.62		
3	AvaI	2×2.44	2×2.42	sd	c		HindIII	2.44	2.14		
4	EcoRI	6.12	6.27	la	e2		PstI	4.66	4.34		
	AvaI	3.07	3.23				BstEII	11.9	11.6		
	BamHI	4.64	4.8			49	EcoRI	(0.05+)6.34	6.39	sl	l–n
5	EcoRI	1.64	1.09+0.55	sg	e3	50	EcoRI	2.29	2.25	sd	l–n
6	EcoRI	3.70	3.66	sd	e4	51	EcoRI	0.71	0.75	sa	l–n
7	EcoRI	6.12	3.92+2.20	sg	e5	52	EcoRI	0.69	0.67	sd	l–n
8	EcoRI	2×1.71	2×1.68	sd	f2		BamHI	2.55	2.52		
9	EcoRI	3.53	3.49	sd	h2	53	EcoRI	2.9	1.76+1.18	sg	o/p
10	AvaI:	0.82+0.86	0.81+0.80	2sd	h2	54	EcoRI	0.84	0.83	sd	o/p
11	EcoRI	3.82	3.90	la	i2	55	EcoRI	0.80	0.81	sa	o/p
	AvaI	2.70	2.80			56	EcoRI	2×3.70	7.4	sl	a–e
12	AvaI	2.42	2.44	sa	m3	57	BspXI	....	1.10+....	sg	f–i
13	EcoRI	6.39+3.70	9.7	ld	n3	58	EcoRI	0.65+0.84	1.48	sl	k
	AvaI	3.33+1.55	4.49			59	AvaI	2.65	1.59+1.06	sg	k
	BamHI	5.17	4.75			60	AvaI	4.62	3.27+1.36	sg	k
	SmaI	9.6	9.2			59/60	SmaI	19.2	12.3+6.91		
	PstI	16.3	15.9			61	BamHI	5.30	3.74+1.56	sg	k
14	EcoRI	3.79	3.90	la	p3	62	BamHI	6.94	4.55+2.40	sg	k
	AvaI	2.71	2.80			63	EcoRI	0.89	0.91	sa	k
	HindIII	6.94	7.04			64	EcoRI	2.90	2.94	sa	k
	SmaI	4.75	4.86			65/66	AvaI	3×0.62	3×0.61	2sd	k
15	EcoRI	2.73	2.69	sd	q2	67	EcoRI	1.00+0.80	1.17+0.64	inv	k
	AvaI	2.38	2.35				BamHI	5.17+0.76	5.23+0.70		
	BamHI	1.35	1.33			68/69	AvaI	(0.19+)5.47	3.20+2.46	sg+sl	k
16	BamHI	1.39	1.42	sa	q2	70	AvaI	3.33	3.20	ld(a)	k
17	AvaI	2.15	1.78+0.37	sg	a2/b/c	71	AvaI	2.53	2.46	ld(b)	k
	SmaI	14.6	11.4+3.17			70/71	PstI	15.0	14.8	ld(a+b)	k
18	AvaI	1.55	1.54	sd	g	72	AvaI	11.1+0.82	11.6	ld(c)	k
19	AvaI	1.55	1.56	sa	h		BstEII	6.45	6.15		
20	AvaI	2.42	2.46	sa	l	70–72	BamHI	8.95	8.5	ld	k
21	AvaI	(0.05+)3.33	3.38	sl	m4					(a+b+c)	
22	AvaI	6.85+0.62	7.5	sl	m2/3		HindIII	7.75	7.30		
23	EcoRI	0.75+0.4	1.15	sl	n1		SmaI	8.8	8.4		
24	EcoRI	1.00	1.02	sa	n2/3	73/74	EcoRI	3.79+2.90	3.85+1.55 +0.98	2sg+ld(c)	k
25	EcoRI	1.76	1.73	sd	o						
26	AvaI	2.15+0.82	3.00	sl	o	75	BamHI	1.39	1.42	sa	l–p
27	BamHI	2.55	2.58	sa	p1	76	EcoRI	2×1.45	2.9	sl	k–p
28	EcoRI	0.89	0.83	ld	a/b/c	77	BamHI	3.10	3.12	sa	k–p
	AvaI	4.49	4.43			78/79	EcoRI	2.9	1.16+0.96 +0.77	2sg	q
	HindIII	3.41	3.36								
	PstI	3.87	3.80			80	EcoRI	1.83+0.89	2.73	sl	q
	BstEII	2.22	2.16			81	EcoRI	1.38+0.48	1.86	sl	q
29	EcoRI	0.84	0.86	sa	a/b/c	82	EcoRI	1.38	0.86+0.52	sg	q
30	EcoRI	(0.06+)2.35	2.41	sl	d	83	BamHI	7.9	7.15+0.76	sg	q
31	EcoRI	3.70	3.05+0.65	sg	g/h	84	BamHI	(0.13+)3.10	3.23	sl	q
32	HindIII	2.67	1.46+1.22	sg	g/h	85	BamHI	2.00+1.35	3.36	sl	q
33	EcoRI	6.12+3.70	9.8	sl	i	86	SmaI	(0.3+)9.3	9.6	sl	q
34	EcoRI	1.00+0.80	1.79	sl	i	87	BstEII	(0.18+)4.07	4.25	sl	q
35	AvaI	2×3.07	2×3.00 (+2×0.07)	sg	i	88	EcoRI	1.00	0.90	ld	q
							AvaI	5.47	5.37		
36	AvaI	2×2.42	2×2.40	ld	i		SmaI	8.8	8.7		
37	AvaI	0.86	0.87	sa	p2/3	89	EcoRI	3.53	3.60	la	q
38	AvaI	2.42	2.44	sa	p2/3		BamHI	4.86	4.92		
39	BamHI	8.95	0.84+8.1	sg	o/p1		SmaI	9.3	9.4		
40	HindIII	7.75	6.6+1.17	sg	p2/3		PstI	5.03	5.10		
41	EcoRI	3.70	3.74	sa	a–d	90	EcoRI	2×1.45	2×1.42	sd	q
42	AvaI	2×2.42	2×2.44	sa	a–d	91	EcoRI	1.45	1.42	sd	q
43	EcoRI	1.33	0.77+0.56	sg	f		AvaI	3.33	3.29		
44	AvaI	1.55	1.54	sd	f	92	EcoRI	2.9	2.86	sd	q
45	AvaI	4.49	4.46	sd	f		AvaI	2.05	2.00		
46	HindIII	3.41	3.44	sa	f	93	EcoRI	0.95	0.96	sa	q
47	EcoRI	3.79	3.82	sa	g–i		AvaI	0.88	0.89		
	AvaI	2.67	2.70								

Table 3. (continued)

No.	Enzyme	Fragments lost	Fragments gained	Mutation type[a]	Carrier plastotype	No.	Enzyme	Fragments lost	Fragments gained	Mutation type[a]	Carrier plastotype
94	EcoRI	0.94	0.90	sd	q	104	PstI	7.87	7.73 (+0.14)	sg	a–i
	AvaI	2.42	2.38			105	EcoRI	1.14	1.10	sd	a–i
95	EcoRI	0.65	0.63	sd	q		HindIII	3.44	3.41		
	AvaI	1.59	1.57			106	EcoRI	1.02	1.00	sd	a–i
96	EcoRI	0.56	0.59	sa	q		HindIII	2.44	2.42		
	AvaI	1.55	1.57			107	EcoRI	0.69	0.67	sd	a–i
97	BstEII	2×2.41	2×2.45	sa	q	108	EcoRI	3.79	3.70	ld(a)	a–i
98	EcoRI	6.34	6.12 (+0.22)	sg	a–i		SmaI	7.5	7.41		
99	AvaI	5.47	4.49+1.04	sg+la	a–i	109	EcoRI	0.48	0.4	ld(b)	a–i
+100	SmaI	8.8	7.8+1.03				AvaI	2.53+0.62	3.07		
101	AvaI	2.71	2.42+0.28	sg	a–i		SmaI	9.4	9.3		
	SmaI	9.6	9.3 (+0.3)			108/109	PstI	16.3	16.1	ld(a+b)	a–i
102	BamHI	5.17	4.64 (+0.53)	sg	a–i	110	AvaI	3.33	2.67+0.58	sg+ld(a)	a–i
103	BamHI	2.84	1.72+1.12	sg	a–i						

[a] sg – site gain, sl – site loss, sa – small addition, sd – small deletion, la – large addition, ld – large deletion, inv – inversion

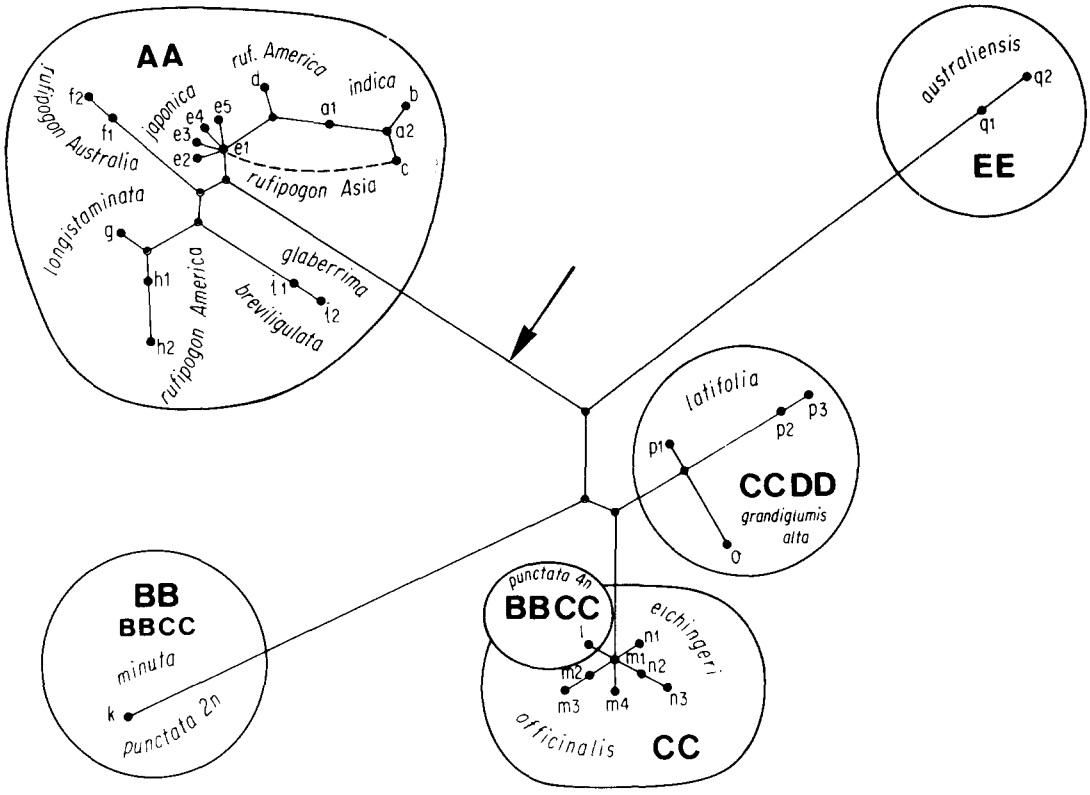

Fig. 1. Cladogram showing relationships among 32 plastotypes in *Oryza* and their correspondence with taxa and earlier definition of genomes. Branch lengths are proportional to the number of branch-specific mutations. Plastotypes are symbolized in small letters and genomes in capital letters. The *broken line* indicates a possible hybrid origin of plastotype c. The *arrow* indicates the position of the root

Possible coincidences, reversions, and recombinations include the following. In mutations numbered 46 and 105, we assumed that the similarities between one HindIII band (3.44 kb) of type f and type k to q result from a coincidence due to a reversion in type f. Other coincidences involved the bands AvaI-32 (2.44 kb) and AvaI-33 (2.42 kb). These are mutations numbered 12, 38, and 42. The genetic separation of the plastotypes concerned (m3 – p2/p3 – a/b/c/d) is confirmed by many other mutations. In contrast, among other possibilities, we suggest that mutation 3, also involving bands AvaI-32 and -33, could result from recombination between an *indica* type (like a2) and a *japonica* type (like e1); as for mutations 2, plastotype c regained the primitive phenotype (see discussion).

The cladogram

Apart from the probable recombination just mentioned and assuming a few coincidences, the cladistic analysis presented above does not result in any contradiction or intrinsic ambiguity. It is presented in Fig. 1 as a cladogram drawn by hand, in which the length of the branches is proportional to the number of presumed mutations. Figure 2 represents a computer-generated cladogram based on the mutations listed in Table 3 and shows their distribution. The site mutations are labelled. If only site mutations were considered, 17 plastotypes would be distinguished, i.e., about half the total number, but the general topology of the cladogram would not be appreciably modified. Note, however, that plastotypes such as a1 and e1 characterizing, respectively, the *indica* and *japonica* subspecies of cultivated rice would be no longer distinct.

Addition/deletion modifications make up a higher proportion of the diversity than restriction site modifications within species than between species, as can be verified by studying Fig. 2. For example, taking the following plastotypes or groups of plastotypes as corresponding to different species, a to e, f, g, h, i, k, l to m4, n, o to p3, and q, there are 24 additions/deletions and 10 sites gain/loss within species, while there are, respectively, 39 and 37 of these mutation types between species.

It appears that some of the branches of the cladogram are longer than others in terms of accumulated mutations. In particular, the branches leading to plastotypes k and q are relatively longer. Also, some plastotypes appear to be particularly conserved compared to others in the same cluster (see, in particular, plastotypes e1 and c, i2 or h2). Finally, with the root located between the Sativa and the Latifolia groups of species, the Latifolia group appears to be much more diverse than the Sativa group. However, as shown in Fig. 1, there is a remarkable congruence between the clusters of plastotypes and the genomes as defined on the basis of pairing of chromosomes.

Genetic distance analysis

In both comparisons, restriction patterns from each enzyme considered independently or combining those of EcoRI, AvaI, BamHI, and PstI (see "Materials and methods"), plastotypes k and q appeared the most different from one another and from all others. Plastotypes a–i were further distinguished from the group of l–p plastotypes as shown in Fig. 3 for the combined data. The only exception to this observation is shown in Fig. 4. It concerns the plastotypes of the *O. latifolia* complex, o–p. In the EcoRI patterns, plastotypes o–p were found more similar to k than to i–m plastotypes. The same trend was also observed in HindIII and PstI patterns. However, if plastotypes o–p are discarded from the comparison, the overall relationships obtained by analysis of

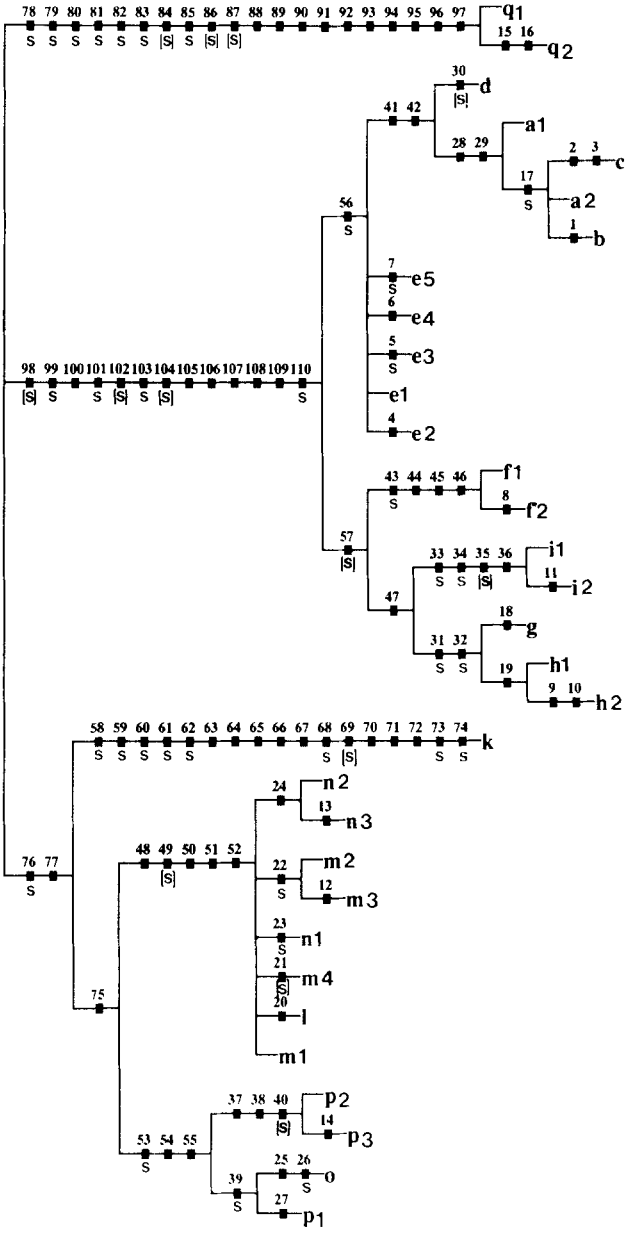

Fig. 2. A cladogram generated using a prototype of the package CLADOS of K. Nixon (Bailey Hortorium, Cornell), showing same relationships among plastotypes as in Fig. 1 and the distribution of mutations listed in Table 3. A "s" indicates a site mutation; "(s)" means that a site mutation was presumed but the smallest fragment involved was not seen in the gel. Symbols of plastotypes are as in the text; mutations are numbered as in Table 3

EcoRI digests are comparable to that obtained with other enzymes, with only expected minor differences. In the matrices of distances (not shown, but available on request), it was also noted that, from the EcoRI patterns (involving, on average, 57 fragments), plastotypes o–p are closer to a–i and q plastotypes, in terms of genetic distances, than to the l–m plastotypes with which they

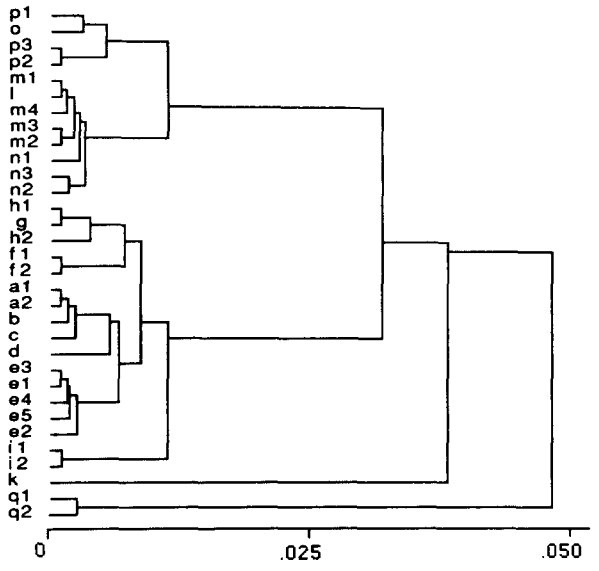

Fig. 3. An average linkage dendrogram based on the genetic distances between plastotypes, obtained from the combined data of all patterns. Symbols as in text

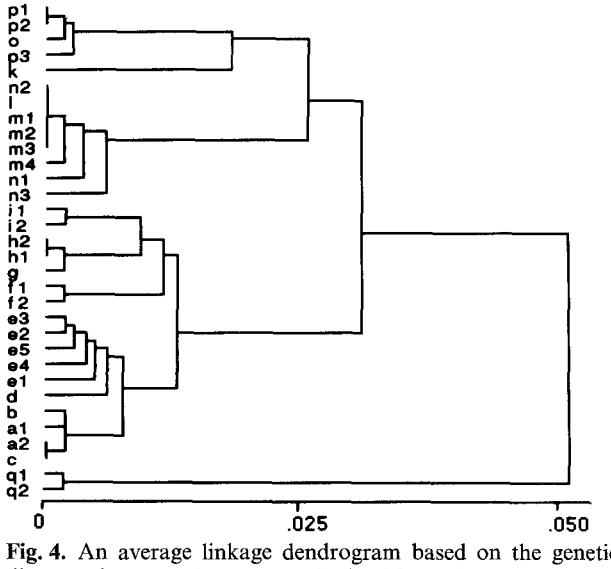

Fig. 4. An average linkage dendrogram based on the genetic distances between plastotypes, obtained from the EcoRI restriction patterns only. Compared with Fig. 3, it shows the closer relationships of o–p and k plastotypes as revealed by this enzyme. Symbols as in text

are most related in the combined data. Apart from this and the fact that the distance "tree" of Fig. 3 is unrooted, the distance analysis showed a close congruence with that of the cladistic analysis described above.

Specific observations

Inheritance of plastotypes. Inheritance of cpDNA was studied by surveying two artificially produced interspecific hybrids (*O. glaberrima* × *O. rufipogon* and *O. rufipogon* American form × *O. rufipogon* Asian form) and four hybrid derivatives ("Obake", Ghesquiere 1988) of *O. longistaminata* × *O. sativa*. Strict maternal inheritance of cpDNA restriction pattern was observed. This inheritance analysis does not, however, rule out occasional leakage of the paternal plastome as suggested by the following observation.

Fragment polymorphism within a plant. One plant (accession IRRI 103919, line V41A), which was a cytoplasmic male-sterile *O. sativa*, could not be classified unambiguously as containing just one plastotype. It seemed to have an approximately 1 : 1 mixture between plastotypes a1 and a2 (with the two bands differentiating plastotypes a1 and a2 in AvaI patterns showing the relative reduction of fluorescence expected in an equimolar mixture). This plant belonged to a "WA" cytoplasm cms line, which was described above to be heterogeneous, with plastotypes a1 and a2 observed in different plants. Other plants of the same line were later shown to have also a mixed plastotype and cross hybridization with a labelled probe confirmed a mixed plastotype a1 + a2 (Second et al. 1989).

As the maintainer lines providing pollen in the reproduction of these cms plants carry an a1 plastotype while the cms plants more often have an a2 plastotype, this observation suggests occasional paternal transmission of cpDNA.

Plastotype diversity within accessions. In 48 accessions out of the 238 studied (excluding the male-sterile lines), more than one plant per accession was studied (2.5 on average). In all but one instance (acc. W135), we found no variation within an accession.

Plastotype diversity within species. Two groups of species with contrasted diversity at cpDNA level could be made. (1) The species *O. glaberrima*, *O. breviligulata*, *O. longistaminata*, *O. rufipogon* (Australian), diploid *O. punctata*, tetraploid *O. punctata*, *O. minuta*, and *O. australiensis* with one or two plastotypes each. (2) The species *O. sativa*, *O. rufipogon* (Asian), *O. rufipogon* (American), *O. officinalis*, *O. eichingeri*, and *O. latifolia* complex with three to nine plastotypes each. In view of the respective ecological distributions of these species and the evolutionary scenario proposed for the section Oryza (Second 1985b), it appears a clear trend for these two groups to correspond to species with different ecological/evolutionary characteristics, although the difference was not clearcut. The second group is constituted exclusively of species thought to have undergone extensive introgressive hybridization and/or allotetraploidization. They occupy human disturbed environments or were recently (a few centuries ago) introduced on continents new to them. The first group is composite but is mostly characterized by species such as *O. breviligulata*, *O. longistaminata*,

diploid *O. punctata*, and *O. australiensis*, which were sampled in the whole range of their distribution, yet have a limited cpDNA diversity. They are species generally found in undisturbed habitats and thought to have occupied these habitats for millions of years. A similar trend appeared in the single species, *O. officinalis* (data not shown, see Dally 1988), with the eight accessions collected in their original forest habitat having the single m1 plastotype, while four plastotypes were observed in all in that species.

Discussion

On the technique

Use of a novel cpDNA extraction technique made it possible for us to isolate sufficient amounts of this DNA from adult plants for direct visualization by ethidium bromide staining, following gel electrophoresis of restriction digests. As a result, we were able to study a large number of individual plants while keeping a high resolution of the banding patterns.

Ten plastotypes among cultivated rice and 32 plastotypes in the Oryza section of genus *Oryza* could be reliably distinguished. Use of accurate molecular size markers allowed a consistent interpretation of the variation in terms of mutations. About half of the plastotypes are distinguished only on the basis of small additions/deletions that are unlikely to be detected with a hybridization technique after a short migration.

Although there were more than twice the rate of additions/deletions compared to the site gain/loss in the mutations found within species than between species, the same overall topology of relationships between species was obtained by considering separately site gain/loss or addition/deletion.

Our data does not as yet allow the variation in the cpDNA molecule to be mapped. However, this should be possible with the sequence data of total rice cpDNA now available (Hiratsuka et al. 1989).

On the taxonomical information

Most of the relationships that were demonstrated on the basis of differences between cpDNA molecules of section Oryza are in agreement with what is known of these relationships based on other parameters, although there are noticeable differences which we will now review.

The earlier definition of genomes based on the observation of chromosome pairing in F_1 hybrids is clearly reflected in the similarity of the corresponding plastotypes, as in isozyme distances. The root of the cladogram was placed between the two species groups Sativa and Latifolia, although in terms of restriction band differences and number of inferred mutations, the A-genome plastotypes are closest to the C- (as well as to the CD- and most of the BC-) genomes plastotypes than to the B- and E-genomes plastotypes. Differences in the rate of accumulation of mutations could account for this observation.

Confirmation of the position of the root of the cladogram came from a study of cpDNA restriction patterns in the Oryzeae tribe (Zhang and Second 1989 and unpublished data), from which it appears that mutation 102 (in Table 3) is specific (as site loss) to the Latifolia group of species, while mutations 103 and 104 are specific to the Sativa group of species. This placement of the root of the cladogram supports the recognition (Second 1984) of these two groups of species as natural groups among the section Oryza.

Although kept as a separate species by Tateoka (1962) because it represents a clearly typifiable species with endemic distribution, completely isolated from its counterparts, *O. australiensis* was included in the Latifolia group of species on the basis of the geographic structure of isozyme diversity in the Oryza section (Second 1984). This treatment also agrees with the morphological affinities. In cpDNA variation, both *O. australiensis* and the diploid form of *O. punctata* appear widely divergent in the section Oryza. Diploid *O. punctata* also has an endemic distribution (in Africa) and a distinct habitat from its tetraploid counterpart (Sano 1980, and unpublished data). However, unlike *O. australiensis*, diploid *O. punctata* is not clearly distinguishable morphologically from tetraploid *O. punctata* or from the *O. officinalis* complex.

The great similarity of cpDNA of other diploids in the Latifolia group, *O. officinalis* and *O. eichingeri*, only confirms what is known at other levels. The heterogeneity of *O. officinalis* with four different plastotypes is in line with the sterility often found in its hybrids between different forms, the variety of habitat in which it is found, its wide distribution on isolated subcontinents, including continental Asia and insular Southeast Asia, and the introgression of the B genome that it has undertaken (Second 1984). At least one other plastotype exists in the *O. officinalis* complex that is worth mentioning. It was observed (Zhang and Second 1989 and unpublished data) in the accession W1805 from Sri Lanka, classified as *O. collina*. It differs in particular from all other plastotypes observed in the Sativa section by one additional restriction site (that it shares with *O. brachyantha*, *O. meyeriana*, and several other members of the Oryzeae tribe) and also shows other differences.

On the other hand, the heterogeneity of *O. eichingeri* at the cpDNA level came as a surprise, considering that only five accessions were analyzed and three plastotypes were found. It may be speculated whether a large deletion of 400 bp such as found in its plastotype n3 is deleterious or not.

As expected, the cpDNA diversity brought new information on the relationship of allotetraploid species. Allotetraploid *O. punctata* had a l plastotype differing by a single small addition from the most conserved plastotype (ml) in *O. officinalis,* indicating the likely female parental species. *O. minuta* and *O. malampuzahensis* had a plastotype k instead which is characteristic of the other parental species, diploid *O. punctata.* This indicates that the allotetraploidization of BC-genome species involved at least the two reciprocal crosses and confirms that these tetraploids have evolved recently. A similar case is reported in *Aegilops triuncialis* (Murai and Tsunewaki 1986).

The heterogeneity of the CD-genome species is also surprising, although all plastotypes o/p are similar to each other. The heterogeneity of cpDNA revealed in BamHI restriction patterns in the *O. latifolia* complex was reported earlier (Ichikawa et al. 1986) and we obtained additional intraspecific variation. These authors further suggested a greater similarity of the large subunit of the Fraction I protein of the *O. latifolia* complex to that of *O. sativa* than to any other species. We clearly established a relationship of the *O. latifolia* complex with the *O. officinalis* complex. However, reminiscent of their observation, we mentioned particularities in the *O. latifolia* cpDNA restriction patterns that place them closer to all other patterns [they would be more "primitives" in the cladistic sense; this appears partly in the cladogram but better in the distance analysis matrix (not shown), which involves all fragments, not only "derived" ones]. This fact is particularly puzzling, given the obscure origin of the D-genome in particular, and also the frequent gigantic morphology and high growth rate of the *O. latifolia* complex even compared to other tetraploids in the Sativa section. As isozyme data, the present data do not indicate any direct relationship of CD-genome species with any particular diploid species, but support the notion that the *O. officinalis* complex contains their closest relatives, although the contribution of other genomes in the constitution of the D-genome appears likely. Rapid appearance of new plastotypes has been observed in some cases, in sympatric zones of pines (Govindaraju et al. 1989), and following interspecific artificial substitution of cytoplasm through backcrossing (Frankel et al. 1979). The present data thus do not allow the hypothesis that CD-genome species evolved recently from their Old World relatives to be refuted, but further investigation is necessary.

The diversity found in the Sativa group of species was less than could have been expected. At an isozyme level, the main divergence is between *O. longistaminata* and the Australian form of *O. rufipogon,* while the Asian and American forms of *O. rufipogon* cluster in a third divergent group with *O. breviligulata* and cultivated rice (Second 1985a). The maximum divergence is nearly as large in the Sativa group as in the Latifolia group, although on average the Latifolia group is more variable (Second 1984, 1985b). In contrast, at the cpDNA level, *O. longistaminata* was no more divergent than *O. breviligulata.* The annual Australian *O. rufipogon* shared the same plastotypes as its perennial form, which is less divergent at the isozyme level (unpublished data). The diversity of the American form of *O. rufipogon* was remarkable, as the four accessions analyzed each showed a different plastotype, two of them related to the Asian form and the two others to *O. longistaminata.*

Ishii et al. (1988) also studied the cpDNA diversity in A-genome species of rice using a somewhat less resolving system of analysis than ours and only the three enzymes EcoRI, HindIII, and PstI for wild species (but a total of 11 restriction enzymes for cultivated rice). They found three plastotypes among *O. sativa.* Two correspond to our a1 and e1 plastotypes, respectively, and the third one, observed in a single *indica* cultivar, was not found in the present study. *Oryza longistaminata, O. breviligulata,* and the Australian *O. rufipogon* each had a single unique plastotype corresponding, respectively, to our g, i1, and f1 plastotypes. The Asian *O. rufipogon* had three plastotypes corresponding to our a and e1 plastotypes and a third, observed in only one accession, had the EcoRI pattern of one type and the PstI and HindIII patterns of the other. This particular combination was not observed in the present study but is reminiscent of our c plastotype with "recombinant" characteristics. Ishii et al. (1988) also found that American *O. rufipogon* was relatively quite variable, with four plastotypes for five accessions. Three of them are related to our g–h plastotypes but one has a low-molecular-weight PstI band that we did not observe. The fourth has a plastotype close to an a that we did not observe in this taxon. If we combine the results of their study with ours, we find six different plastotypes for (at most) nine different accessions. Three of them are related to the diversity found in the Asian *O. rufipogon* and the other three to *O. longistaminata.* On the other hand, all these American accessions were closely related to Asian *O. rufipogon* at the isozyme level (Second 1985a).

The diversity of American *O. rufipogon* at the cpDNA level can be further correlated with some other observations concerning organelle DNA as follows. (1) In an extensive study of the isozyme diversity in A-genome species (Second 1985a), it was found that two accessions of American *O. rufipogon* had a unique electromorph for the total Fraction I protein, while all other accessions in the A-genome shared the same electromorph. This character was shown to be maternally transmitted, and Pental and Barnes (1985) further demonstrated that it was due to a difference in the large subunit of Rubisco. One of the two accessions was included in the present study and had an h1 plastotype. (2) Another

accession of American *O. rufipogon* showed the only difference observed in a preliminary comparison with a cytochrome oxidase probe of the mitochondrial DNA in a set of species including A- and C-genome species (Zhang et al. 1988).

Based on the above observations, it is obvious that there is a uniquely large amount of variation in the organellar DNA of American *O. rufipogon*. This taxon was probably introduced to America by man within the past four centuries, based on its close relationship with Asian *O. rufipogon* at the isozyme level and the impossibility for seeds to travel naturally from Asia to America (Second 1985a). From its similarity with Asian *O. rufipogon* at the isozyme level, but a similarity (of some of the accessions) with *O. longistaminata* at the cpDNA level, the simplest explanation is that American *O. rufipogon* arose by hybridization of the two Old World species. The unique divergence seen at the cpDNA level in Fraction I protein and mitochondrial DNA would have arisen in the process.

The West Indian form of *O. rufipogon* is another example of differential similarity between taxa when recorded at cpDNA and isozyme levels. It is closely related to *O. glaberrima* and *O. breviligulata* at the isozyme level, but close to *O. sativa* at the cpDNA level.

Also, some *indica* cultivars have a *japonica* type of cpDNA. However, these cases appear to be exceptions to our more general observation that, in a taxon, a particular cpDNA-characterized cytoplasm is associated with a particular isozyme-characterized nucleus. Whether this is to mean primarily coadaptation of chloroplast and nuclear DNA, or simultaneous divergence with time in reproductive isolation of the taxon, or both, cannot be answered at present. However, when an odd association between plastotype and genotype markers is observed, a cytoplasmic substitution, following natural hybridization and backcrossing (wide cross hybrids in rice are often male sterile), may be suspected.

In relation to the cytoplasmic male sterility

The cytoplasmic male sterility (cms) was found to be associated with cytoplasms closely related to the normal one. Cytoplasmic male sterility in *indica* cultivars is generally associated with an a2 plastotype, most common in the direct ancestor of *indica* cultivars (South Asian *O. rufipogon*) and close to the a1 plastotype of the normal cytoplasm. Cytoplasmic male sterility in *japonica* cultivars is associated with a cytoplasm of a cultivar "Chinsurah Boro II" intermediate between *indica* and *japonica* (Second 1982, group II in Glaszmann 1987 classification) that was indistinguishable from normal *japonica* cytoplasm in our study. These pollen-sterile combinations appear to follow the intergradation of forms found at the isozyme level between *japonica* cultivars and South Asian *O. rufipogon*, which could possibly point to their evolutionary origin.

Shinjyo (1975, and also in Oka 1988) studied extensively the distribution of cytoplasm inducing male sterility in wild species and cultivars by backcrossing them with a *japonica* cultivar ("Taichung 65"). He found frequently the male sterility-inducing cytoplasm in the South/Southeast Asian *O. rufipogon*, also in the American *O. rufipogon* (once out of 25), but not in Australian *O. rufipogon* and *O. longistaminata*. This means that association of a *japonica* nucleus with f and g plastotypes (and probably others) leads to a fertile plant, but that the combination of the same nucleus with an a plastotype (and possibly others) can lead to a sterile plant. In the case of a combination of a *japonica* or an *indica* cytoplasm with the nucleus of *O. glaberrima* (i plastotype), sterility due to the non-dehiscence of anthers was observed. The pollen grains were normal. The reciprocal combination was found to be completely fertile (Yabuno 1977; Sano 1986).

In relation to the origin of O. sativa

The three main types of cultivated rice distinguished at isozyme level, *japonica* and *indica* types of *O. sativa* and *O. glaberrima*, are associated with three main different plastotypes, e1, a1, and i1, respectively. At isozyme level, these three types of cultivated rice are closely related to their presumed respective ancestors (Second 1985a): *O. rufipogon* (Chinese), *O. rufipogon* (South Asian), and *O. breviligulata*. The same trend appears at the cpDNA level, supporting the hypothesis of an independent domestication of *O. glaberrima* and the origin of the *japonica-indica* differentiation of *O. sativa* in a divergence within *O. rufipogon*, prior to domestication. Intermediate (at the isozyme level) *indica/japonica* cultivars generally have the *japonica* plastotype as characterized in the present study. This new information supports the hypothesis of an origin of the genetic diversity of *O. sativa* in the hybridization of the *indica* and *japonica* types (Second 1982).

It should be observed, moreover, that the two plastotypes a1 and e1, most common in cultivated rice, are somewhat intermediate and are closer in comparison to one another when compared to the diversity present in *O. sativa* and Asian *O. rufipogon*. The most diverse plastotypes would be represented by e3 or e5 on the one hand and by a2 or b on the other, although this needs to be confirmed with more restriction site analyses or with sequence data.

Are there occasional paternal inheritance and recombination of cpDNA in rice?

Occasional paternal inheritance of plastids has been reported in plants whose plastome inheritance is predomi-

nantly maternal (Schmitz and Kowallik 1986; Lee et al. 1988). We present here the preliminary observation that it may also occur in rice. Evidence for cpDNA recombination in nature has also been produced (Govindaraju et al. 1989). The origin of some of our plastotypes, such as c, is best explained by recombination. Others, such as h and o–p, may be presumed to have accumulated differences since they originated after hybridization events of the carrier taxa, probably less than four centuries ago. The intergradation of plastotypes found in cultivated *O. sativa* also suggests that evolution of cpDNA in rice is not purely clonal, as presumed from the mode of maternal inheritance.

If recombination between genomes from two different evolutionary lineages occurs at the cpDNA level, one may further speculate that this process could lead, through selection, to the repair of slightly deleterious mutations accumulated independently, and thus move closer to the "primitive" state of the molecule. This could explain differences in the rate of accumulation of mutations in different lineages and in the singularity of the o–p plastotypes. It will be further studied.

The evolutionary scenario for the section Oryza, which was deduced from an evolutionary/biogeographical interpretation of the isozyme data (Second 1984, 1985a, b), was used to critically examine some of the features of the variation observed here. In particular, the higher intraspecific variation found in some species was traced to those species that have undergone introgressive hybridization or allotetraploidization through indirect human intervention. Presumed rapid divergence of some plastotypes was pointed out. Reciprocally, the results of the present study should be interpreted together with those that previously were used to establish the evolutionary relationships in the section Oryza. This will be done in a future paper.

Acknowledgements. We wish to express our appreciation to Dr. T. A. Dyer, who gave us valuable advice during the course of this study. He and Dr. K. Tsunewaki helped us considerably with preparation of this publication. Dr. J. Doyle kindly ran for us a prototype of the computer package CLADOS. Z. Y. Wang provided us with his computer program for genetic distance analysis. Many of the seeds were generously provided by Dr. Y. Sano, NIG, and Dr. T. T. Chang, IRRI. The present data were the subject of a PhD Thesis (Dally 1988) supported by a grant from the French Government and from the DAAD (FRG). The work was done in a laboratory of the Emberger Center, Montpellier, under an agreement between ORSTOM and CNRS. The paper was in part written during a stay of the second author at the Department of Plant Breeding and Biometry, Cornell University, USA.

References

Bowman CM, Dyer TA (1982) Purification and analysis of DNA from wheat chloroplasts isolated in nonaqueous media. Anal Biochem 122:108–118

Chang TT (1988) Taxonomic key for identifying the 22 species in the genus *Oryza*. Int Rice Res Newslett 13:4–5

Dally AM (1988) Analyse cladistique de mutations de l'ADN chloroplastique et phylogénie des riz (section *Eu-Oryza* du genre *Oryza*). Editions ORSTOM, Paris

Dally AM, Second G (1989) Chloroplast DNA isolation from higher plants: an improved nonaqueous method. Plant Mol Biol Rep 7:135–143

Doebley J (1987) Evolution of the *Zea* chloroplast genome. Evol Trends Plants 1:3–6

Doyle JJ, Doyle JL (1990) A chloroplast DNA phylogeny of the wild perennial relatives of soybean (*Glycine* subgenus *Glycine*): congruence with morphological and crossing groups. Evolution 44:371–389

Frankel R, Scowcroft WR, Whitfeld R (1979) Chloroplast DNA variation in isonuclear male-sterile lines of *Nicotiana*. Mol Gen Genet 169:129–135

Ghesquiere A (1988) Diversité de l'espèce sauvage de riz, *Oryza longistaminata* A. Chev. & Roehr et dynamique des flux géniques au sein du groupe *Sativa*. PhD thesis, University of Paris-Sud, Orsay

Glaszmann JC (1987) Isozymes and classification of Asian rice varieties. Theor Appl Genet 74:21–30

Govindaraju DR, Dancik BP, Wagner DB (1989) Novel chloroplast DNA polymorphism in a sympatric region of two pines. J Evol Biol 2:49–59

Hiratsuka J, Shimada H, Whittier R, Ishibashi T, Sakamoto M, Mori M, Kondo C, Honji Y, Sun CR, Meng BY, Li YQ, Kanno A, Nishizawa Y, Hirai A, Shinozaki K, Sugiura M (1989) The complete sequence of the rice (*Oryza sativa*) chloroplast genome: intermolecular recombination between distinct tRNA genes accounts for a major plastid DNA inversion during the evolution of the cereals. Mol Gen Genet 217:185–194

Hosaka K, Hanneman RE Jr (1988) Origin of chloroplast DNA diversity in the Andean potatoes. Theor Appl Genet 76:333–340

Ichikawa H, Hirai A, Katayama T (1986) Genetic analyses of *Oryza* species by molecular markers for chloroplast genomes. Theor Appl Genet 72:353–358

Ishii T, Terachi T, Tsunewaki K (1988) Restriction endonuclease analysis of chloroplast DNA from A-genome diploid species of rice. Jpn J Genet 63:523–536

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

Kato S (1930) On the affinity of the cultivated varieties of rice plants, *Oryza sativa* L. J Dep Agric Kyushu Imp Univ 2:241–276

Lee DJ, Blake TK, Smith SE (1988) Biparental inheritance of chloroplast DNA and the existence of heteroplasmic cells in alfalfa. Theor Appl Genet 76:545–549

Lehväslaiho H, Saura A, Lokki J (1987) Chloroplast DNA variation in the grass tribe Festuceae. Theor Appl Genet 74:298–302

Lolo OM, Second G (1988) Peculiar genetic characteristics of *O. rufipogon* from Western India. Rice Genet Newslett 5:67–70

Murai K, Tsunewaki K (1986) Molecular basis of genetic diversity among cytoplasms of *Triticum* and *Aegilops* species IV, CpDNA variation in *Ae. triuncialis*. Heredity 57:335–339

Nayar NM (1973) Origin and cytogenetics of rice. Adv Genet 17:153–292

Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York

Ogihara Y, Tsunewaki K (1988) Diversity and evolution of chloroplast DNA in *Triticum* and *Aegilops* as revealed by restriction fragment analysis. Theor Appl Genet 76:321–332

Oka HI (1988) Origin of cultivated rice. Japan Scientific Societies Press, Tokyo; Elsevier, Amsterdam Oxford New York Tokyo

Palmer JD (1985) Isolation and structural analysis of chloroplast DNA. Methods Enzymol 118:167–186

Palmer JD, Osorio B, Aldrich J, Thompson WF (1987) Chloroplast DNA evolution among legumes: loss of a large inverted repeat occurred prior to other sequence rearrangements. Curr Genet 11:275–286

Pental D, Barnes SR (1985) Interrelationship of cultivated rices *Oryza sativa* and *O. glaberrima* with wild *O. perennis* complex. Analysis of fraction 1 protein and some repeated DNA sequences. Theor Appl Genet 70:185–191

Rickwood D, Hames BD (ed) (1982) Gel electrophoresis of nucleic acids – a practical approach. IRL Press, Oxford

Roschevicz RJ (1931) A contribution to the knowledge of rice. Bull Appl Bot Genet Plant Breed 27:3–133 (in Russian with detailed English summary)

Roshevits RY (1937) Grasses. Engl. transl. 1980, Smithsonian Inst & NSF, Washington DC; Indian Nat Sci Doc Center, New Delhi

Sano Y (1980) Adaptive strategies compared between diploid and tetraploid forms of *Oryza punctata*. Bot Mag Tokyo 93:171–180

Sano Y (1986) Interspecific cytoplasm substitutions of an Indica strain of *Oryza sativa* L. and *O. glaberrima* Steud. Euphytica 34:587–592

Schmitz UK, Kowallik KV (1986) Plastid inheritance in *Epilobium*. Curr Genet 11:1–5

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

Second G (1984) A new insight into the genome differentiation in *Oryza* L. through isozymic studies. In: Sharma AK, Sharma A (eds) Advances in chromosomes and cell genetics. Oxford & IBH, New Delhi

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

Second G (1985b) Relations évolutives chez le genre *Oryza* et processus de domestication des riz. Coll Etudes et Thèses. ORSTOM, Paris

Second G, Dally AM, Zhang SH (1989) Occasional biparental inheritance of chloroplast DNA in rice. Rice Genet Newslett 6:150–153

Shinjyo C (1975) Genetical studies of cytoplasmic male sterility and fertility restoration in rice, *Oryza sativa* L. Sci Bull Coll Agric Univ Ryukus 22:1–51

Tateoka T (1962) Taxonomic studies of *Oryza*. II. Several species complex. Bot Mag Tokyo 75:165–173

Tateoka T (1963) Taxonomic studies of *Oryza*. III. Key to the species and their enumeration. Bot Mag Tokyo 76:165–173

Tateoka T (1964) Notes on some grasses. XVI. Embryo structure of the genus *Oryza* in relation to the systematics. Am J Bot 51:539–543

Virmani SS, Govinda Raj K, Casal C, Dalmacio RD, Aurin PA (1986) Current knowledge of and outlook on cytoplasmic-genetic male sterility and fertility restoration in rice. In: IRRI (ed) Rice genetics. IRRI, Manila, pp 633–647

Yabuno T (1977) Genetic studies on the interspecific cytoplasm substitution lines of *japonica* varieties of *O. sativa* L. and *O. glaberrima* Steud. Euphytica 26:451–463

Zhang SH, Second G (1989) Phylogenetic analysis of the tribe Oryzeae. Total chloroplast DNA restriction fragment analysis. A preliminary report. Rice Genet Newslett 6:76–80

Zhang SH, Dally AM, Second G (1988) Nonaqueous method for purification of cytoplasmic DNA from individual rice plant for RFLP analysis. Rice Genet Newslett 5:154–156