

## Evolutionary relationships in the *Sativa* group of *Oryza* based on isozyme data

Gérard SECOND \*

O.R.S.T.O.M. (Institut Français de Recherche Scientifique  
pour le Développement en Coopération)

\* Present address : Unité de Biologie des Populations et des Peuplements  
C.N.R.S., Centre « Louis Emberger », B.P. 5051  
F 34033 Montpellier Cedex

### Summary

The wild and weedy species of the *Sativa* group of *Oryza* were studied for their isozyme polymorphism at 24 loci for 181 plants, each taken in a different population, the total representing the entire area of pan-tropical distribution. One to 10 electromorphs were distinguished per locus with an average of 4.1. A multivariate analysis of the data on an individual strain basis showed several important discrepancies with taxonomy based on morphological characters. In particular, the Australian form was largely divergent from the Asian and American forms of *O. rufipogon*, which were closely related to each other and to *O. breviligulata*. A Chinese form of *O. rufipogon* could be distinguished from a South Asian form although many intermediates existed. The South Asian form of *O. rufipogon* was the most variable of all although it was not possible to distinguish the annual from the perennial life forms at the isozyme level. The electromorph frequencies and the distribution of genetic distances in 2 by 2 strain comparisons were analysed within and between the geographic groups or species. Large values of Nei's genetic distances (up to 1.26) and large proportions of loci with no allele in common (up to 9 out of 24) were found between groups. The closest relationships of the 3 basic isozyme patterns of, *indica*, *japonica* and *glaberrima* cultivated rice were found with South Asian *O. rufipogon*, Chinese *O. rufipogon* and *O. breviligulata*, respectively. It was concluded that the *Sativa* group was composed of 2 species endemic to Africa — *O. longistaminata* and *O. breviligulata* — and a complex species *O. rufipogon* with South Asian, Chinese, New Guinean, Australian and American geographic forms. How such differentiation has occurred was discussed in relation to : 1) the major tectonic and climatic events in the Old World over the past 15 million years ; 2) the independent domestication of rice in South Asia, China and Africa ; 3) the role played by man in carrying seeds across geographic barriers and in disturbing primitive habitats.

*Key words* : Poaceae (= Gramineae), *Oryza*, isozyme, evolution, domestication.

### Résumé

*Les relations évolutives entre lignées du groupe Sativa (génôme AA) d'Oryza L. étudiées à 24 locus d'isozymes*

Les espèces sauvages et adventices du groupe *Sativa* du genre *Oryza* sont étudiées pour leur polymorphisme isozymique à 24 locus sur 181 plantes, chacune prélevée dans une

population différente et représentant toute l'aire de la distribution pantropicale. De 1 à 10 électromorphes sont distingués à chaque locus avec une moyenne de 4,1. Une analyse multivariée des résultats lignée par lignée montre plusieurs anomalies importantes par rapport à la taxinomie basée sur des caractères morphologiques. En particulier, la forme australienne est largement divergente des formes asiatiques et américaines d'*O. rufipogon*, étroitement apparentées entre elles et avec *O. breviligulata*. Une forme chinoise d'*O. rufipogon* peut être distinguée d'une forme sud-asiatique, bien que de nombreux intermédiaires existent. La forme sud-asiatique d'*O. rufipogon* est la plus variable de tout le groupe sans que l'on puisse y distinguer les types biologiques annuels et pérennes au niveau des isozymes. Les fréquences d'électromorphes et la distribution des distances génétiques entre lignées 2 à 2 sont analysées parmi les groupes géographiques ou espèces retenus. On observe des distances génétiques (NEI) importantes (jusqu'à 1,26) avec une large proportion de loci sans allèles communs atteignant 9 sur 24 entre groupes. Un apparemment maximal des 3 combinaisons isozymiques de base des riz cultivés, *indica*, *japonica* et *glaberrima* est trouvé avec *O. rufipogon* d'Asie du Sud, *O. rufipogon* de Chine et *O. breviligulata*, respectivement. Il est conclu que le groupe *Sativa* est composé de 2 espèces endémiques en Afrique — *O. longistaminata* et *O. breviligulata* — et d'une espèce complexe *O. rufipogon* avec des formes géographiques d'Asie du Sud, de Chine, de Nouvelle-Guinée, d'Australie et d'Amérique. On discute comment une telle différenciation est apparue et discutée en liaison avec : 1) les principaux événements tectoniques et climatiques dans l'Ancien Monde depuis 15 millions d'années ; 2) la domestication indépendante des riz en Asie du Sud, Chine et Afrique ; 3) le rôle de l'homme dans le transport des graines à travers les barrières géographiques et dans la perturbation des habitats primitifs.

*Mots clés* : Graminées, *Oryza*, isozyme, évolution, domestication.

## I. Introduction

The heterogeneous and world wide tribe *Oryzae* is considered on comparative chromosomal and anatomical grounds to be primitive within the grass family (STEBBINS, 1956 ; TATEOKA, 1957). Within the genus *Oryza*, the genome AA, defined by MORINAGA (1964) on the basis of normal chromosome pairing at meiosis in F1 hybrids, corresponds to the *Sativa* group (MORISHIMA & OKA, 1960). It comprises the 2 cultivated species *O. glaberrima* and *O. sativa* and the 3 wild or weedy species *O. rufipogon* Griff. (= *O. perennis* Moench), *O. longistaminata* A. Chev. et Roer. (sometimes misnamed *O. barthii* A. chev. or *O. perennis* subsp. *barthii*) and *O. breviligulata* A. chev. et Roer. (= *O. barthii* A. Chev.). The importance of the *Sativa* group lies in its close relationship with cultivated rice. *O. rufipogon* is, by itself, a Complex with various Asian, American and Oceanian forms and perennial, annual and intermediate life forms (MORISHIMA, 1969).

Although numerous crossing barriers are commonly found within the *Sativa* group (CHU *et al.*, 1969) hybrids between the various species, taxa or life forms, are not generally totally sterile. Gamete sterility, which is the commonest barrier encountered, is generally more pronounced for pollen than for ovules. Introgression of genes from one species to another is possible and occurs spontaneously (OKA & CHANG, 1961 ; CHU & OKA, 1970). Also, some weedy forms seem to have evolved from hybrid swarms between independently domesticated cultivars (SECOND, 1982).

Research on rice isozyme variability began in the mid 1960's in Japan and was reviewed by ENDO & MORISHIMA (1983). It was limited to esterases, peroxidases and acid phosphatases until SECOND & TROUSLOT (1980) extended it to more than 13 different enzyme systems. The genic diversity of cultivated rice and wild and weedy *O. breviligulata* was studied at 40 isozyme loci by SECOND (1982).

The morphology and F1 pollen sterility relationships of wild rice was studied extensively in MORISHIMA (1969). However, only *O. rufipogon* and *O. longistaminata* were considered and reliable samples from South China were missing. The present report includes the most typical strains studied by MORISHIMA (1969) and, in addition, *O. breviligulata* and samples of *O. rufipogon* recently made available (including a reliable collection from South China). It gives for the first time, as far as the author knows, an overall picture of the isozyme polymorphism of the *Sativa* group including all its ecogeographic forms as available in collections.

From these data, an attempt has been made to interpret the observed genetic structure, to reconstruct the distribution of the group and its structure in the paleo-environment and to determine in which area domestication occurred.

## II. Material and methods

### A. *Oryza* strains or individuals

Material studied on 24 isozyme loci included 181 strains (in the case of annual, predominantly autogamous forms) or individual plants (in the case of perennial, eventually self incompatible forms) chosen each in a different population among a great number of wild and weedy populations collected across most of the range of the *Sativa* group distribution. They were obtained from the collections of the National Institute of Genetics (NIG) in Misima (Japan), the Institut Français de Recherche Scientifique pour le Développement en Coopération (ORSTOM) in Ivory Coast, the National Institute of Agricultural Science in Tsukuba (Japan) and the International Rice Research Institute (IRRI) in the Philippines. When available, seeds collected on the original sites were used and the strains were propagated by selfing. In some cases and particularly for the Chinese strains, multiplication of seeds might have been done previously to the accession in unknown conditions. In the case of self compatible *O. longistaminata* but also for most of the perennial forms of *O. rufipogon*, plants grown from original seeds and vegetatively propagated in green house were used.

In view of the numerous accessions available from Asia and Africa, limited numbers were selected to represent most of the area of distribution and of the ecological origins. For Africa, 20 *O. longistaminata* plants and 20 *O. breviligulata* strains collected in populations truly isolated from cultivated fields were selected. For South Asia, only a few of the original populations were really isolated from cultivated rice fields in terms of possible pollen flow ; 361 strains or plants were analysed for some of the loci (see table 2) and showed a complex array of intermixed genotypes. 85 strains or plants were chosen among them, each from a different population, so as to represent as equitably as possible the ecological situations or life forms in the various South Asian geographical origins. They were classified as « weedy » or « wild » according to the location of collection in, or outside, a rice field ; « wild » forms were further classified into « annual », « perennial » or « intermediate annual-perennial » life forms in the sense in use at the moment in the NIG collection (see in particular MORISHIMA *et al.*, 1961). Although based on repeated observations, for most plants or strains, this classification into life forms remained partly ambiguous. All populations or strains available from China (28), New Guinea (5), Australia (10) and America (16) were

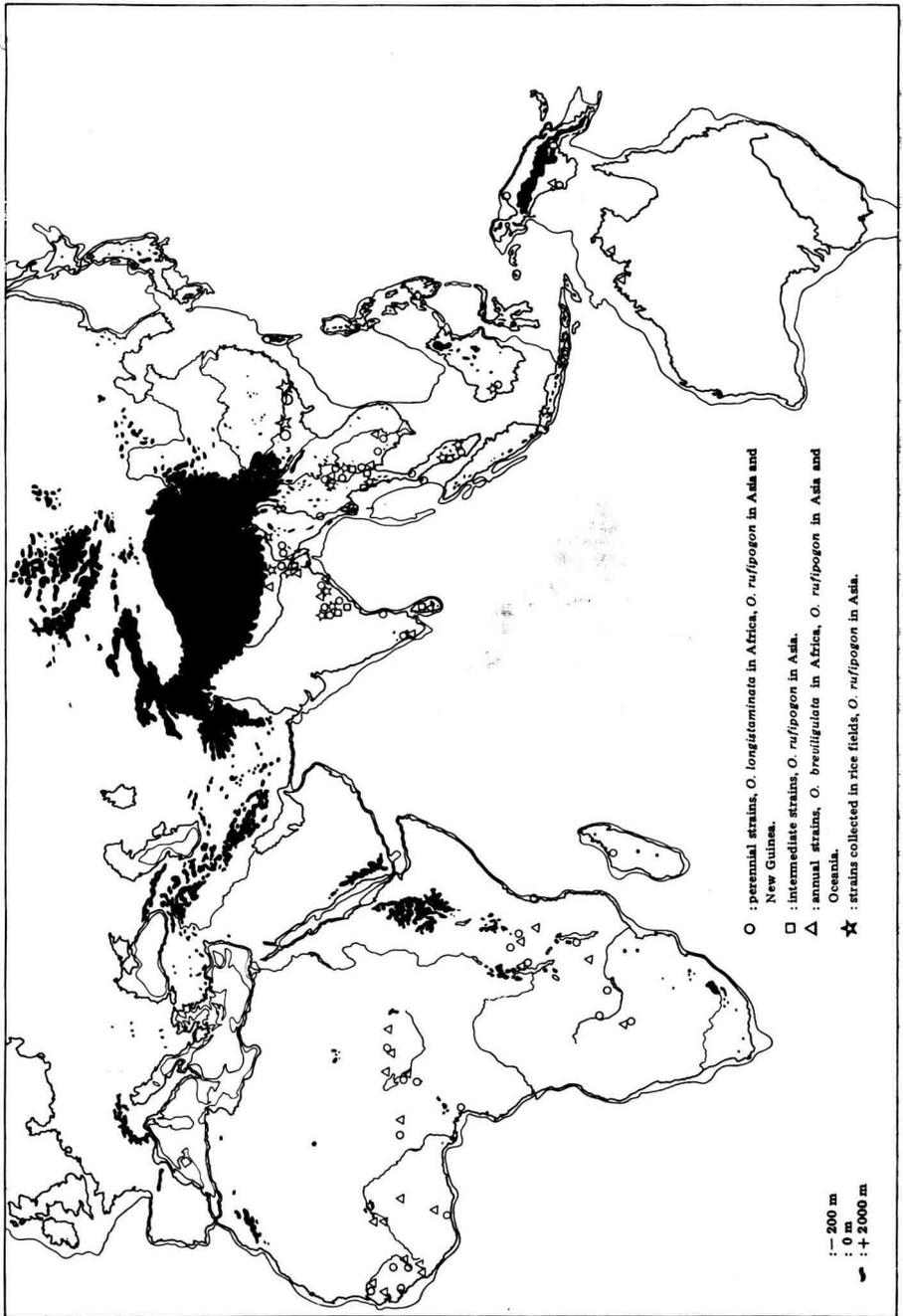


FIG. 1

*The geographical origins of the spontaneous strains of the Sativa group studied from the Old World and Australia shown by conventional symbols according to their classification. A dot may correspond to several strains extracted from different populations. The map shows in addition the - 200 m isobathic line and the areas more than 2 000 meters above sea level.*

*Les origines géographiques des formes spontanées du groupe Sativa étudiées dans l'Ancien Monde et l'Australie.*

selected and represented by one individual. Geographic origins are shown on figure 1 (except for America). The exact location of most Chinese strains was not known in detail but they came from three different origins : 13 were from Kwangsi, 10 from Boluo County in Kwangtung Province and 4 from Taiwan. They belonged to 3 different original collections. The American accessions were from various origins of the Amazon basin and Cuba.

### B. Electrophoresis and analysis of isozymes

Seeds were germinated for 5 to 7 days in distilled water on filter paper. The tip of one coleoptile was cut to perform a preliminary electrophoresis (see table 1). The rest of the same seedling and 2 other seedlings were transplanted the next day on heat-disinfected soil. Usually, the cut seedling survived and the same plant was further analysed or another plant of the same accession number was used. In the case of plants vegetatively propagated, a preliminary electrophoresis was done on the white part of a growing leaf. Another part of a growing leaf and a blade of a young but developed leaf were further sampled and immediately used for electrophoresis, according to the scoring sequence exposed in table 1. The samples were crushed in distilled water and subjected to horizontal starch gel electrophoresis with only minor modifications (indicated in a forthcoming paper) of the technique described in SECOND & TROUSLOT (1980) and SECOND (1982).

TABLE 1

Scoring sequence of the 24 loci at 3 different plant stages and 3 different buffer systems.

Séquence de lecture des 24 locus à 3 stades différents avec 3 systèmes de tampons.

Stage	Buffer systems		
	Histidine pH 8.0	Histidine pH 6.0	Borate pH 8.0
Coleoptile . . . . .	<i>Idh-A</i> <i>Pgm-A</i> <i>Cat-A</i> <i>Pgd-A</i> and B <i>Pgi-A</i> and B		
White part of a growing leaf . . . . .	( <i>Idh-A</i> ) ( <i>Pgm-A</i> ) <i>Adh-A</i> + <i>To-A</i> ( <i>Cat-A</i> ) ( <i>Pgd-A</i> and B) <i>Gdh-A</i> and B <i>Pox-B</i> <i>Est-A</i>	<i>Mdh-A, B</i> and C <i>Lap-E</i> <i>Est-B</i> and E <i>Pox-B</i>	
Developped leaf . . .	( <i>Pox-B</i> ) <i>Pgi-A</i> and B	<i>Got-A</i> and C <i>Lap-E</i> + <i>Gp-A</i> ( <i>Est-B</i> and E) <i>Pox-B</i>	<i>Px-1</i> <i>Acp-1</i>

Parentheses indicate a possibility of scoring not always used. *Pgi-A* and B, *Lap-E* and *Pox-B* loci are always scored twice at least. For a given stage and buffer system, loci are listed according to the slicing sequence of the gel plates.

In this way, the most polymorphic loci were scored twice at least and, if necessary, individuals with similar bands were placed side by side to check for minor differences in migration distances.

At the two highly variable *Pgi* loci, hidden variability was screened by heat treatment as described in SECOND (1982), using a set of slices of the same gel.

### C. Nomenclature of genetic variants

The electrophoretic or thermostability variants referred to in this paper as electromorphs or electro-thermomorphs are assumed to represent classes of alleles indistinguishable according to the method utilized. The nomenclature for most of them and the criteria used as evidence for the existence of a locus are described elsewhere (SECOND & TROUSLOT, 1980 ; SECOND, 1982).

The loci studied were the following : Alcohol dehydrogenase (*Adh-A*) ; Glutamate dehydrogenase, 2 loci (*Gdh-A*, B) ; Malate dehydrogenase, 3 loci (*Mdh-A*, B, C) ; Isocitrate dehydrogenase (*Idh-A*) ; Phosphogluconate dehydrogenase, 2 loci (*Pgd-A*, B) ; Catalase (*Cat-A*) ; Peroxidase, 2 loci (*Pox-B* and *Px-1*, PAI *et al.*, 1973) ; Tetrazolium oxidase (*To-A*) ; Aspartate aminotransferase or Glutamate oxaloacetate transferase, 2 loci (*Got-A*, C) ; Acid phosphatase (*Acp-1*, PAI *et al.*, 1975) ; Phosphoglucosmutase (*Pgm-A*) ; Esterase, 3 loci (*Est-B*, E, Ca) ; Leucine aminopeptidase (*Lap-E*) ; Glucose-phosphate isomerase, 2 loci (*Pgi-A*, B). In addition the major band of general protein was scored as it appeared in the strain for LAP or ACP with extracts of green leaves. It probably corresponds to the fraction 1 protein (Rubisco) which is known to be encoded by both nuclear and cytoplasmic factors (KUNG, 1976). As variation was observed in only two closely related strains, a single locus was assumed : *Gp*.

## III. Results

### A. Electrophoretic data

A number of zymograms were drawn or pictured in SECOND & TROUSLOT (1980). The genetic determinism was largely inferred from their polymorphism observed within and between autogamous and self incompatible species. The simplest genetic interpretation compatible with the observations was adopted. In the case of self-incompatible *O. longistaminata*, agreement was found between observed and expected frequencies of genotypes under the hypothesis of panmixis for all polymorphic loci in this species, with populations of 50 to 200 individuals (SECOND & TROUSLOT, 1980 and unpublished data). Further, Mendelian segregations were observed for the loci *Mdh-A*, *Cat-A*, *Pox-B*, *Est-E*, *Lap-E*, *Pgi-A* and B (SECOND, 1982 and unpublished data). *Px-1* and *Acp-1* loci have already been described in more detail (PAI *et al.*, 1973 and 1975). The linkage relationships between these loci are under study by several groups of workers. Clearly, the linkage groups are scattered over many chromosomes (SECOND & MORISHIMA, 1981 ; A. GHESQUIERE, J.L. LACROIX, J.L. PHAM, L. SANO, personal communications).

Only the *To-A* locus was found to be completely monomorphic among the 24 loci scored. Two loci had a null electromorph : *Est-B* and *Est-E*. A total of 99 electromorphs (4.1/locus) were observed with a maximum of 10 for the *Pgi-A* locus. Although heterozygotes were frequently observed in cross-pollinated strains, their frequencies were not considered because many lines had been artificially sexually multiplied, generally only once.

It appeared that the interrelationships among strains at the isozyme level did not always reflect their classification on the basis of conventional taxonomy, based on morphological characters, geographic origins or life forms. Prior to analysing the genic diversity within groups of strains or species, each strain was thus considered on an individual basis and a multivariate analysis of its interrelationships with all other strains was studied.

### B. Multivariate analysis of the observed genetic differentiation

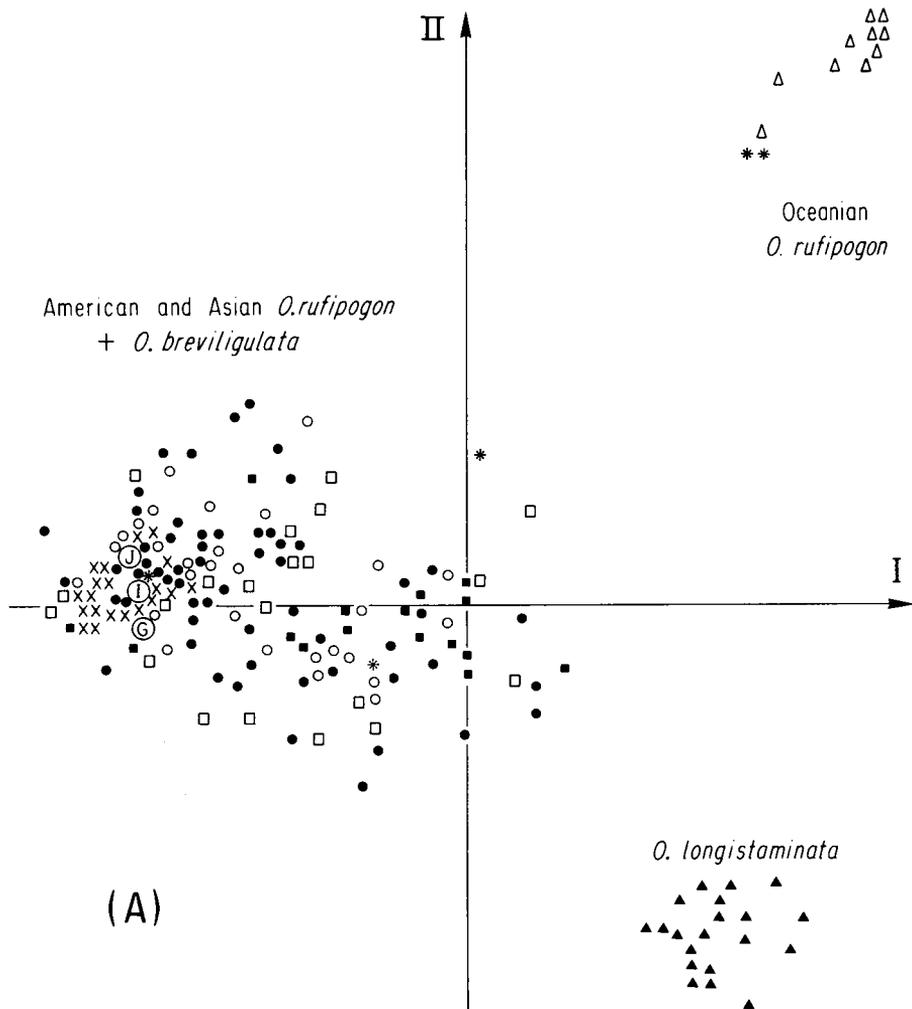
Genetic distances (NEI, 1975) were calculated (a) between strains in 2 by 2 comparisons among the 181 strains studied at 24 loci and (b) between each strain and the 3 basic isozyme patterns found in cultivated rice : *indica*, *japonica* and *glaberrima* (SECOND, 1982, tabl. 6). A multivariate analysis of the generated distances was made by the method of principal coordinate analysis of Gower (in PIELOU, 1969, p.p. 257-260).

Four series of analyses were carried out as follows :

- 1) all strains were included without any *a priori* classification ;
- 2) the oceanian forms of *O. rufipogon* and *O. longistaminata* were excluded ;
- 3) only the Asian forms of *O. rufipogon* (including those of Chinese origins) were considered, (a) alone and (b) together with the collection of 40 cultivated varieties of *O. sativa* analysed in SECOND (1980) ;
- 4) only the *O. rufipogon* strains of South and continental South-East Asian origins were considered with regards to their classification on a life form basis.

The percentage of total variation explained by the first 2 vectors varied between 27 and 31 p. 100 depending on the analysis. It appears to be rather low. However it depends on the criterion used to calculate it. A comparison of the distances on the figures 2 and 3 with the true distances given later in table 3 shows that the overall representation is essentially correct.

When all strains were considered (fig. 2), only *O. longistaminata* and the Oceanian forms of *O. rufipogon* were clearly distinguished. Other strains were in the main cluster. Among the 5 strains of New Guinean origin, 3 were found among the Asian forms of *O. rufipogon* while 2 were close to the Australian strains of *O. rufipogon* according to the classification of the same strains previously made by MORISHIMA (1969) on the basis of numerical taxonomy and analysis of F1 pollen sterility. The 2 strains classified with the Oceanian form of *O. rufipogon* are nevertheless much different from the Australian strains on morphological grounds (MORISHIMA, 1969). In figure 2 they are closer to the Asian form of *O. rufipogon* than the Australian strains according to the exact electrophoretic distances given in table 3.



Δ : Australia ; \* : New Guinea ; ▲ : *O. longistaminata* , Africa ;  
 x : *O. breviligulata* , Africa ; ● : India , Sri - Lanka , Nepal , Bengladesh ,  
 Burma , Thailand , Cambodia ; ○ : China ; □ : Malaya , Indonesia ,  
 Philippines ; ■ : South - America , Cuba .  
 ⊙ : Glaberrima ; ① : Indica ; ② : Japonica .

FIG. 2

181 strains of the spontaneous forms of the Sativa group plotted in the first plane  
 of a principal coordinate analysis of the genetic distances in strain comparison at 24 loci.  
 The geographical origins or taxonomic classifications are indicated by a conventional sign.  
 The positioning of cultivated rice is indicated by the 3 basic isozyme patterns of indica,  
 japonica and glaberrima types of cultivars.

181 lignées des formes spontanées du groupe Sativa distribuées dans le premier plan  
 d'une analyse en coordonnées principales de la diversité à 24 locus lue en électrophorèse.

When the strains of the main cluster only were considered, no group was clearly distinguished but the following origins tended to segregate from others : 1) the strains from China together with the *japonica* isozyme pattern of cultivated rice on the 1<sup>st</sup> vector, 2) the strains from America on the 2<sup>nd</sup> vector and 3) *O. breviligulata*, together with the *glaberrima* pattern of cultivated rice, on the 3<sup>rd</sup> vector.

Figure 3 shows the results on the first plane of an analysis including the Asian strains of *O. rufipogon* and the collection of 40 varieties of *O. sativa*. The variability of wild strains largely exceeded but also included that of cultivars for the most

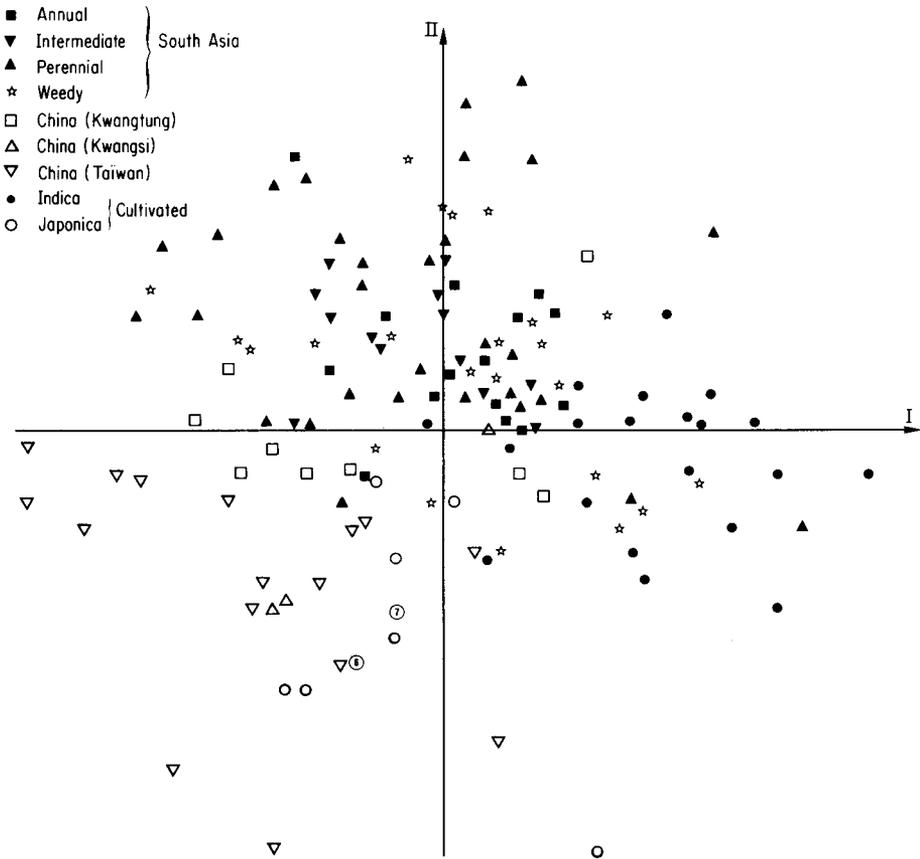


FIG. 3

110 strains of Asian *O. rufipogon* and 40 varieties of *O. sativa* plotted in the first plane of a principal coordinate analysis of the genetic distances in strain comparison at 24 loci. Are indicated by a conventional symbol, the life forms for South Asian *O. rufipogon*, the 3 different collections for Chinese *O. rufipogon* and the classification in indica and japonica types of *O. sativa* by the phenol reaction of glumes.

110 lignées des formes spontanées d'*O. rufipogon* asiatique et une collection de 40 variétés d'*O. sativa* distribuées dans le premier plan d'une analyse en coordonnées principales de la diversité à 24 locus lue en électrophorèse.

part. Both axis allowed for the distinction of the positive and negative phenol reaction type of cultivars which is clearly in relation with their differentiation in *indica* and *japonica* types (OKA, 1958). Interestingly, the pattern of variation in the wild strains was similar to the main pattern of differentiation in the cultivars. Only wild strains originating in China were found to be closely related to the *japonica* type of cultivars. There was however a sharp differentiation among the 3 different collections originating in 3 different provinces of China :

— Most strains from Kwangsi province were close to the *japonica* type although their morphological appearance was typically that of wild rices. They were generally homozygous, sometimes for rare electromorphs at *Idh* and *Pgm* loci, unique to them (see later, fig. 4).

— The strains from Kwangtung province were intermediate between the *indica* and *japonica* types (Six out of 9 were also heterozygous at such locus as *Cat-A* related to the *indica-japonica* differentiation).

— The strains from Taiwan were also close to the *japonica* type.

On the other hand, only few strains of *O. rufipogon*, most of them weedy, were found together with the *indica* type of cultivars. Most *O. rufipogon* strains from South Asia were intermediate between the *indica* and *japonica* types on the 1<sup>st</sup> axis but segregated from cultivated rice on the 2<sup>nd</sup> axis due to their heterozygosity and larger electromorph diversity.

Also apparent from figure 3, is that no strong association between the genetic diversity and the life form types, as classified here, could be detected. This point was confirmed in round 4 of the multivariate analysis when only the continental South-Asian strains were considered.

It should be noted that a similar structure was revealed in principal component and factorial analysis of correspondence on the electromorph frequencies (0, 0.5 or 1 in a given strain). Using the standard genetic distance of NEI, however, less relative weight was given to the rare alleles at the detriment of the overall relationships.

Taking into account, both the above pattern of interrelationships and the phenetic study at the morphological and F1 pollen sterility level of MORISHIMA (1969), the following subclassification of the *Sativa* group is suggested on a geographical basis :

1) *O. rufipogon*, South Asian (SA) : This group comprises all South and South-East Asian strains, irrespective of life forms. It includes the 3 strains originating in New Guinea with South Asian affinity.

2) *O. rufipogon*, Chinese (CA) : From China, including Taiwan. It partly overlaps the South Asian group and deserves further studies for its characterization.

3) *O. rufipogon*, New Guinea (ON) : This group includes the two strains originating in New Guinea but with an Australian affinity.

4) *O. rufipogon*, Australia (AO).

5) *O. rufipogon*, America (AM).

6) *O. breviligulata* in Africa (BV).

7) *O. longistaminata* in Africa (AL).

8) Cultivated rices with three basic isozyme patterns : *japonica* (JP), *indica* (ID) and *glaberrima* (GL) (SECOND, 1982).

These groups are used for presenting allozyme frequencies data.

TABLE 2

*Electromorph frequencies and mean gene diversity index at 24 loci in the main geographic wild or weedy forms or species of the Sativa group.*

*Les fréquences d'électromorphes et l'index moyen de diversité génique à 24 locus parmi les principales formes géographiques et espèces sauvages ou adventices du groupe Sativa.*

Electromorphs	Groups and strains or individuals numbers							
	SA 82	SA' 173 to 361	CA 28	ON 2	OA 10	AM 16	BV 20	AL 20
<i>Adh-A</i>								
1 .....	.99	.98	1	1	.9	1	1	.98
2 .....								.02
3 .....	.01	.02						
4 .....					.1			
<i>Gdh-A</i>								
1 .....	.99		.96	1	1	1	1	.95
2 .....								.05
3 .....	.01		.04					
<i>Gdh-B</i>								
1 .....	1		1	1	1	1	1	
2 .....								1
<i>Mdh-A</i>								
1 .....	.98	.94	1	1	1	1	.90	.95
2 .....	.01	ε					.05	.03
3 .....		.01					.05	
4 .....								.03
5 .....	.01	.05						
<i>Mdh-B</i>								
1 .....	1	1	1	1	.3	1	1	.92
2 .....								.05
3 .....					.7			
4 .....								.03
<i>Mdh-C</i>								
1 .....	1	1	1	1	1	1	.90	.80
2 .....								.17
3 .....								.03
4 .....							.10	
<i>Idh-A</i>								
1 .....	1	.99	.86	1	1	1	1	1
3 .....		.01	.14					
<i>Pgd-A</i>								
1 .....	.67	.73	.84			.88	.90	1
2 .....	.10	.07				0.06	.05	
3 .....	.19	.18	.16			0.06	.05	
4 .....				1	1			
5 .....	.04	.02						

Electromorphs	Groups and strains or individuals numbers							
	SA 82	SA' 173 to 361	CA 28	ON 2	OA 10	AM 16	BV 20	AL 20
<i>Pgd-B</i>								
1 .....	.79	.80	.93			.94	1	.95
2 .....	.20	.19	.07			.06		
3 .....	.01	.01						.05
5 .....				1	1			
<i>Cat-A</i>								
1 .....	.95	.97	.30	1	.9	1	1	1
2 .....	.05	.03	.70					
3 .....					.1			
<i>Pox-B</i>								
a .....	.12		.07		.1	.56	.35	.43
b .....	.64		.93		.1	.37	.65	.57
c .....				1	.7			
d .....	.23					.06		
e .....	.01				.1			
<i>Px-1</i>								
0 .....	.08		.04		1			
2 .....	.64		.57	1		.19	1	
4 .....	.28		.40			.81		1
<i>Got-A</i>								
1 .....	.94	.97	1			.31	1	.87
2 .....								.08
3 .....	.06	.03				.69		.05
4 .....				1	1			
<i>Got-C</i>								
1 .....	.98		1	1	.7	1	1	.95
2 .....	.01				.3			.05
4 .....	.01							
<i>Pgm-A</i>								
1 .....	1	1	.91	1	.9	.81	.90	.90
2 .....						.19	.10	.05
3 .....					.1			.05
4 .....			.09					
<i>Est-B</i>								
0 .....	.13		.04					
a .....	.75		.96		.1	.94	1	.73
b .....	.06			1	.1			
c .....					.8			.15
d .....	.06					.06		.13
<i>Est-E</i>								
0 .....	.22	.18	.72	1	1			.52
1 .....	.64	.67	.14			.88	1	
2 .....	.10	.11	.14			.12		
3 .....								.44
4 .....								.04
5 .....	.04	.04						

Electromorphs	Groups and strains or individuals numbers							
	SA 82	SA' 173 to 361	CA 28	ON 2	OA 10	AM 16	BV 20	AL 20
<i>Est-Ca</i>					—			
1 .....	.83		.64	1	1	.88	1	1
2 .....	.13		.36			.12		
3 .....	.04							
<i>Acp-1</i>								
1 .....	.37		.86				.90	
2 .....	.40		.14			.19		
3 .....	.20					.06	.10	.95
4 .....	.02			1	1	.69		.05
5 .....	.01							
6 .....	.01							
7 .....						.06		
<i>Lap-E</i>								
1 .....	.61		.43			.34	.85	
2 .....	.17		.04			.25	.10	
3 .....	.10		.16			.19	.05	
4 .....	.10		.26	1	.9	.16		.73
5 .....	.02		.11			.06		.22
6 .....								.05
<i>Pgi-A</i>								
1 .....	.18	.18	.23			.03	.05	
2 .....	.63	.60	.77			.97	.35	.12
3 .....							.60	
4 .....	.03	.05						.13
5 .....								
6 .....				1	.1			
7 .....					.6			
8 .....	.15	.15			.3			.50
9 .....		.02						.12
10 .....								.13
<i>Pgi-B</i>								
1 .....	.69	.73	.87	1	1	.88	.70	.95
2 .....	.17	.21	.13			.06		
3 .....	.04	.01				.06		.05
4 .....	.10	.04					.30	
<i>Gp</i>								
1 .....	1		1	1	1	.88	1	1
2 .....						.12		
<i>To-A</i>								
1 .....	1		1	1	1	1	1	1
H .....	.24		.19	0	.12	.18	.11	.17

SA : South Asian *O. rufipogon* (SA' stands for a larger collection); CA : Chinese *O. rufipogon*; ON : New Guinean *O. rufipogon* with Australian affinities; OA : Australian *O. rufipogon*; AM : American *O. rufipogon*; BV : *O. breviligulata*; AL : *O. longistaminata*; H is the genic diversity averaged over 24 loci.

### C. Gene diversity within and between geographic groups or species

The frequencies of electromorphs at each of the 24 loci are presented in table 2 for the 8 geographical groups, forms or species defined above. When frequencies were averaged between groups of strains, six loci only (25 p. 100) were found to be monomorphic at the 0.95 level. Usually, only one electromorph dominated each group. For the South Asian group, in addition to the above, a number of lines (173 to 361 in total, subgroup SA') were studied at 13 loci (tabl. 2); the electromorph frequencies remained fairly similar to those obtained with the 85 strains only.

All electromorphs of the American form are shared with the Asian form with the exception of the rare *Gp-2* electromorph found only in 2 strains that originated in Surinam. However, the electromorph *Acp-1.4*, that was frequent among American strains was only found in Asia among the strains that originated in Malaysia (see fig. 4). On the contrary, one frequent electromorph *Pgi-A3* was peculiar to *O. breviligulata*.

The electromorph frequencies (for the most polymorphic loci), obtained in different ecogeographical origins of Asian *O. rufipogon* (which is believed to be the direct ancestor of cultivated *O. sativa*) are diagrammatically presented in figure 4 along with the basic isozyme patterns of *indica* and *japonica* types of cultivated rice. For the South Asian strains, four geographic origins, the life forms and the weedy or non weedy origins were distinguished. Among the Chinese strains, the 3 different geographic origins were distinguished. Geographic clines related to the differentiation in cultivated rice appeared for the electromorph frequencies at a few loci. *Cat-A2* was frequent in China only while *Est-E0* and *Acp-1.1* were frequent in China and Continental South East Asia only.

The electromorphs of the *japonica* type were found with a high frequency only in the Chinese strains with the exception of *Est-Ca2*. There was more similarities of the *indica* type with South Asian strains but no particular life form nor geographical origin appeared to be definitively more closely related to it. Electromorphs *Pgi-A1* and *Pgi-B2* of the *indica* type were only rarely found anywhere in spontaneous plants.

As a measure of the genetic variability, Nei's gene diversity  $H = 1 - \sum P_i^2$ ,  $P_i$  = frequency of electromorph *i*, averaged over the 24 loci was indicated in table 2. The highest mean gene diversity,  $H = 0.25$ , was found in South Asian *O. rufipogon*. With  $H = 0.19$  and  $0.18$ , the Chinese and the American *O. rufipogon* also had a relatively large gene diversity. On the other hand, the Australian *O. rufipogon* and *O. breviligulata* had a lower value :  $H = 0.12$  and  $0.11$ , respectively. With  $H = 0.17$ , the perennial and allogamous *O. longistaminata* had an intermediate value. The 2 « Oceanian » strains from New Guinea were found to have the same isozyme pattern.

The high value of the gene diversity  $H$  for the « South Asian » group did not result from its heterogeneity in terms of geographical origins and life forms, since it was also high and varied from  $0.14$  to  $0.25$  for single ecogeographic subgroups as indicated in fig. 4.

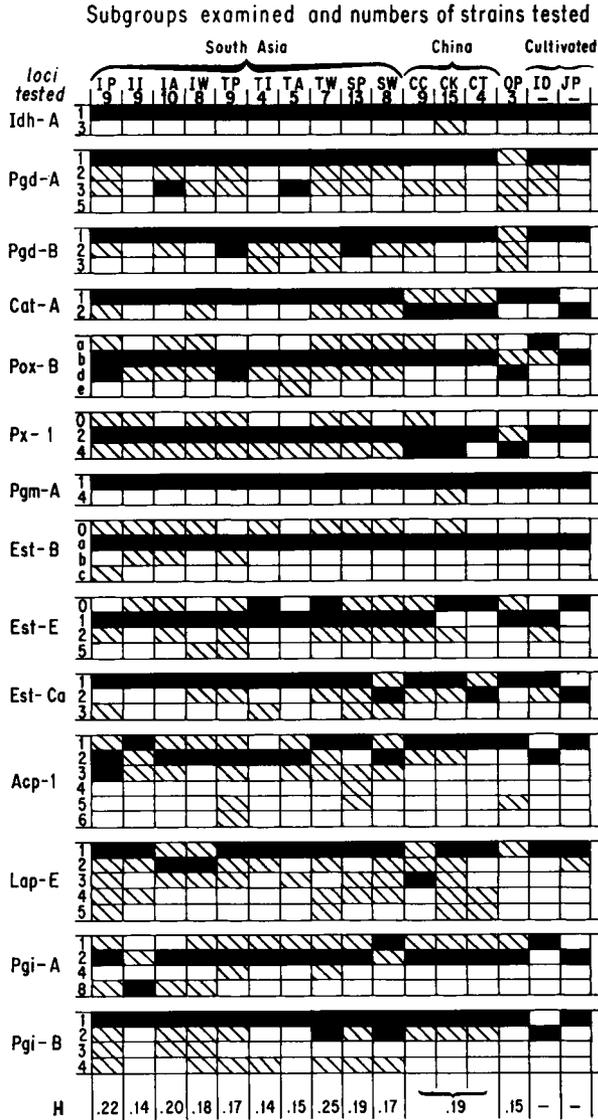


FIG. 4

Diagrammatic representation of electromorph frequencies found at the 14 most polymorphic loci in groups of wild or weedy strains of Asian *O. rufipogon* along the 2 representative isozyme patterns of cultivated *O. sativa*. For South Asian strains, the 1st letter represents the geographic origins: I for Indian subcontinent, T for continental South-East Asia and S for insular South-East Asia; the 2nd letter stands for Perennial, Intermediate, Annual or Weedy. For China, CC, CK and CT stand for strains originating in Kwangtung, Kwangsi and Taiwan respectively. OP stands for Asian *O. rufipogon* strains from New Guinea, ID and JP for indica and japonica representative patterns of cultivated rice. In each group, the predominant alleles are represented by a black box and rarer alleles (in terms of frequency within each subgroup) are hatched.

H is the gene diversity averaged over 24 loci.

Représentation graphique des électromorphes trouvés à 14 locus les plus polymorphes parmi les groupes écogéographiques de la forme asiatique d'*O. rufipogon* ainsi que dans les 2 formules isozymiques de base des riz cultivés.

H : diversité génique de Nei, moyenne sur 24 locus.

#### D. Group analysis of genetic distances between strains

A rationale for the calculation of genetic distances between individuals (each representing a population) rather than between groups lies in the genetic structure of the species considered. Various reproductive barriers are commonly encountered including autogamy, death or sterility of the F1 plants, ecological segregation and geographical disjunction. In addition, domestication has certainly disturbed the natural genetic structure. Any group constituted is genetically heterogeneous.

Furthermore, the large proportion of loci with no common allele and the largely variable gene diversity in different ecogeographical groups do not allow a reliable comparison in terms of Nei's genetic distances between such groups. In place of that, the table 3 presents various aspects of the genetic divergence among the 8 groups of strains defined in section B :

- 1) the distribution (median and range of variation) of Nei's genetic distances ( $D$ ) in strain comparison within and between groups. As the distribution may not be normal, the median value is a better indication than the mean ;
- 2) the number of loci out of 24 with no electromorph in common between groups ;
- 3) the number of loci with no predominant electromorph in common between groups.

In agreement with the multivariate positioning in figure 2, the highest values of  $D$  were found between the Australian and the non-Oceanian strains, with a median value ( $\widehat{D}$ ) of 0.60 to 0.76, while  $\widehat{D}$  between *O. longistaminata* and Asian and American strains was 0.45 only.  $\widehat{D}$  between the Chinese and South Asian groups is slightly greater than  $\widehat{D}$  within these groups (0.29 vs 0.25 and 0.20). The interval of variation of  $D$  (min-max) was considerable among South Asian strains (0.62), which appeared to be composed of a complex intergrade of genotypes.

More loci share electromorphs between the allopatric *O. longistaminata* and Asian *O. rufipogon* than between the parapatric species *O. longistaminata* and *O. breviligulata*. The distances of Australian *O. rufipogon* (annual) from *O. longistaminata* (perennial) and from *O. breviligulata* (annual) were similar. The Australian *O. rufipogon* exhibited however some rare electromorphs typical of Asian *O. rufipogon*. This explains why it has smaller number of loci with no common electromorph than its New Guinean counterpart which was fixed for a unique electromorph at every locus. All loci shared electromorphs between American and Asian *O. rufipogon* and *O. breviligulata*, although predominant electromorphs differed between them.

#### E. Hidden variability at the two Pgi loci

Hidden variability was screened by heat treatment at the 2 Pgi loci. The number of strains screened within each geographical group and the « electro-thermomorphs » found are given in table 4. The data for *O. breviligulata* were taken as recorded by SECOND (1982). Both *O. breviligulata* and Australian *O. rufipogon* seemed to have greater hidden variability than other *O. rufipogon* groups. However some perennial strains of *O. rufipogon* could have heterozygous loci, rendering the method less discriminant for them. Within *O. longistaminata* hidden variability seemed also to exist

TABLE 3

The distribution of genetic distances among individuals within or between groups and the number of loci with no electromorph, or no predominant electromorph in common between the main forms of the Sativa group.

La distribution des distances génétiques entre individus parmi les groupes et le nombre de locus sans allèle (ou sans allèle prédominant) en commun entre les principales formes du groupe Sativa.

	Groups and strains or individuals numbers									
	SA 82	CA 28	ON 2	OA 10	AM 16	BV 20	AL 20	ID 1	JP 1	GL 1
SA South-Asia	min 0	0.02	0.34	0.31	0.09	0.04	0.18	0.06	0.09	0.04
	med 0.25	0.29	0.47	0.61	0.30	0.23	0.45	0.23	0.28	0.22
	max 0.62	0.62	0.69	1.12	0.56	0.61	0.81	0.44	0.48	0.47
CA China	(a) 0	0	0.36	0.30	0.12	0.06	0.24	0.09	0	0.07
	(b) (3)	0.20 0.56	0.49 0.69	0.64 0.98	0.34 0.64	0.25 0.54	0.46 0.94	0.31 0.48	0.14 0.37	0.23 0.42
ON New-Guinea	5	7	—	0.09	0.40	0.47	0.51	0.52	0.53	0.47
	(9)	(9)	0 —	0.16 0.40	0.53 0.61	0.47 0.69	0.60 0.81	— 0.52	— 0.53	— 0.47
OA Australia	4	5	2	0	0.40	0.40	0.47	0.55	0.53	0.47
	(11)	(11)	(4)	0.09 0.34	0.60 1.10	0.61 1.10	0.76 1.26	0.68 0.98	0.68 0.86	0.61 1.10
AM America	0	0	7	6	0	0.13	0.24	0.15	0.21	0.14
	(3)	(5)	(9)	(10)	0.16 0.56	0.29 0.60	0.44 0.76	0.32 0.41	0.38 0.50	0.30 0.45
BV <i>O. breviligulata</i>	0	0	8	7	0	0	0.30	0.16	0.14	0
	(2)	(3)	(9)	(11)	(4)	0.09 0.29	0.45 0.80	0.18 0.28	0.24 0.35	0.09 0.23
AL <i>O. longistaminata</i>	1	2	9	9	2	4	0.03	0.42	0.42	0.30
	(6)	(7)	(10)	(11)	(6)	(6)	0.19 0.39	0.50 0.63	0.52 0.73	0.45 0.60
ID <i>indica</i>	0	0	9	7	0	2	6	—	—	—
	(3)	(6)	(10)	(12)	(4)	(4)	(8)	— —	0.30* —	0.22* —
IP <i>japonica</i>	0	0	9	7	3	2	4	5	—	—
	(4)	(1)	(10)	(11)	(6)	(4)	(8)	(7)	— —	0.20* —
GL <i>O. glaberrima</i>	0	0	8	6	1	0	4	3	3	—
	(2)	(3)	(9)	(11)	(4)	(0)	(6)	(4)	(4)	—

\* These value were 0.51, 0.35 and 0.34 when computed from 40 loci (SECOND, 1982).

SA : South Asian *O. rufipogon* ; CA : Chinese *O. rufipogon* ; ON : New Guinean *O. rufipogon* with Australian affinities ; OA : Australian *O. rufipogon* ; AM : American *O. rufipogon* ; BV : *O. breviligulata* ; AL : *O. longistaminata* ; ID, JP, GL : the *indica*, *japonica*, *glaberrima* types of cultivated rice presented by their basic isozyme patterns respectively.

Above and on the diagonal : the minimum, median and maximum values of Nei's genetic distances calculated on 24 loci in all combinations of pairs of strains. Below the diagonal a) the number of loci with no electromorph in common b) in parentheses, the number of loci differing in predominant allele.

TABLE 4  
*Distribution of 4 thermomorphs at each of the 2 Pgi loci, among 4 geographic groups or taxa of spontaneous strains of the Sativa group.*  
*Distribution, en nombre de lignées testées, de 4 thermomorphes à chacun des 2 locus Pgi parmi 4 sous-groupes géographiques ou espèces du groupe Sativa.*

Loci	Thermomorphs	Groups and strains or individuals numbers					
		SA (90)	CA (14)	OA (6)	AM (5)	BV (68)	
Pgi-A	59 °C . . . . .	86 (A <sup>2</sup> , A <sup>1</sup> , A <sup>8</sup> , A <sup>4</sup> , A <sup>9</sup> )	14 (A <sup>2</sup> , A <sup>1</sup> )	4 (A <sup>6</sup> , A <sup>7</sup> )	5 (A <sup>2</sup> , A <sup>1</sup> )	25 (A <sup>3</sup> )	
	56 °C . . . . .	4 (A <sup>2</sup> )		1 (A <sup>6</sup> )		38 (A <sup>2</sup> , A <sup>4</sup> , A <sup>1</sup> )	
	53 °C . . . . .			1 (A <sup>5</sup> )		4 (A <sup>2</sup> , A <sup>1</sup> )	
	47 °C . . . . .					1 (A <sup>1</sup> )	
Pgi-B	59 °C . . . . .						
	56 °C . . . . .	90 (B <sup>1</sup> , B <sup>2</sup> , B <sup>4</sup> , B <sup>8</sup> )	10 (B <sup>1</sup> , B <sup>2</sup> )	5 (B <sup>1</sup> )	5 (B <sup>1</sup> )	67 (B <sup>1</sup> , B <sup>4</sup> )	
	53 °C . . . . .		4 (B <sup>1</sup> )	1 (B <sup>1</sup> )		1 (B <sup>4</sup> )	
	50 °C . . . . .						

Number of strains are given. In parenthesis, the electromorphs corresponding to the thermomorphs within each group. See table 2 for explanation of group symbols.

for the same loci. Frequencies of thermomorphs were not determined in this species because of the high heterozygosity expected among its strains. Only a few electromorphs were common between *O. breviligulata* and other groups at the *Pgi-A* locus and between Australian *O. rufipogon* and other groups at the *Pgi-B* locus. Thus, the hidden variability study only reinforced the trend of variation found in simple electrophoretic survey. In particular, it reinforced the differentiation of *O. breviligulata* and the Oceanian form of *O. rufipogon* on the one hand but the lack of differentiation of the American form of *O. rufipogon* on the other hand.

#### IV. Discussion

A high degree of isozyme polymorphism was revealed within the *O. sativa* species group. We may discuss successively the parental relationships between its various forms or species and how such differentiation has arisen.

##### A. *The pattern of interrelationships among the Sativa group*

The *Sativa* group was defined on the basis of meiotic chromosome pairing (MORINAGA, 1964) as well as by statistical taxonomical methods based on morphological and physiological characters (MORISHIMA & OKA, 1960). In the subsequent work of the 2 last authors, the *Sativa* group was considered as composed of 2 species on the basis of the length of the ligule only : *O. breviligulata* and *O. perennis*.

MORISHIMA (1969) studied the phenetic similarities among strains of *O. perennis* on the basis of morphological and F1 sterility relationships. No reliable sample originating in China was however available at that time. Based on morphological characters only, the overall greatest dissimilarities were found between the annual and perennial life forms, irrespective of their geographic origins. Next, strains from each continent, Asia, Africa and America, tended to form a different cluster. The pattern of F1 sterility relationships coincided largely with the geographical groups but not with the classification in life forms. Even within the Asian group, a very wide range of F1 pollen sterility relationships was found within and between the annual, intermediate and perennial life forms. In New Guinea, strains related to both the Oceanian and the Asian groups, were found.

It appears that the pattern of interrelationships in the present study is in global disagreement with the morphological similarities. On the other hand, it is largely in agreement with the F1 pollen relationships :

- The Oceanian form is the most distantly related to others.
- The range of F1 pollen fertilities between the American and the Asian forms fall into the range found within the Asian form itself.
- The pattern of F1 pollen relationships within the Asian form is not directly related with its subdivision into life forms.

Our work included the whole *Sativa* group. We were led to distinguish 7 groups of wild strains. In addition to the 4 geographical groups of *O. perennis* defined by MORISHIMA (1969), they corresponded (1) to *O. breviligulata* (whose distinction as a

parallel evolutive series to *O. perennis* (OKA, 1974) appears to be meaningless and based on the length of the ligule only) (2) a Chinese form and (3) a New Guinean form of *O. rufipogon*. The proposed classification of the *Sativa* group is given in table 5 in comparison with the nomenclature used in the papers by MORISHIMA & OKA cited above.

TABLE 5

*Subdivision of the Sativa group.*

*Classification du groupe Sativa.*

As per MORISHIMA & OKA (1960) and MORISHIMA (1969)	Present paper
<i>O. perennis</i> Moench. Asian form	<i>O. rufipogon</i> Griff., South Asian form
—	<i>O. rufipogon</i> , Chinese form
<i>O. perennis</i> , Oceanian form	<i>O. rufipogon</i> , Australian form
	<i>O. rufipogon</i> , Oceanian - New Guinean form
<i>O. perennis</i> , American form	<i>O. rufipogon</i> , American form
<i>O. perennis</i> , African form	<i>O. longistaminata</i> , A. Chev. & Roehr.
<i>O. breviligulata</i> , A. Chev. & Roehr.	<i>O. breviligulata</i> , A. Chev. & Roehr.

The Australian form of *O. rufipogon* was recently named *O. meridionalis* (Ng *et al.*, 1981). We can only agree with the recognition of an Australian form distinct from the others. However, as in the case of *O. longistaminata*, no simple key of determination is available besides biochemical markers. Still *O. longistaminata* is isolated from its relatives by reproductive barriers particularly developed (CHU *et al.*, 1969) while introgression of genes between Asian and Oceanian *O. rufipogon* seems to be easier and to have had a large impact on the variability of Oceanian *O. rufipogon* as is discussed in the following section. Furthermore, Oceanian *O. rufipogon* appears to be composed of 2 distinct forms at least (Australian and New Guinean). That is why we retained them as part of *O. rufipogon* but recognise *O. longistaminata* as a distinct species.

*O. breviligulata* and *O. longistaminata* are 2 species endemic to Africa. On the contrary, *O. rufipogon* is a complex species. While the divergence of an Oceanian form, first suggested by MORISHIMA (1969), was fully corroborated here, it came as a surprise that the American form was very closely related to the Asian form. All frequent electromorphs in the American form were found in Asia, particularly in some strains from Indonesia, that were unfortunately poorly represented in our collection.

Also remarkable was the distinction of some Chinese strains related to the *japonica* type of cultivars. In China, wild rice is distributed not only in Southern Pro-

vines but also in Kiangsi and Hunan Provinces up to 28°N, that is in the Yangtse river basin, in habitats isolated from rice fields (published in *Renmin Erbao*, Aug. 6, 1983, Pr WANG XIANG MING, personal communication). Only *indica* type cultivars (« Hsien ») are cultivated in lowlands in these areas. It appeared from various informations and observations that the collection studied from Kwangsi Province is very likely to be representative of the wild rice found in isolated areas in South-China. In contrast, the collection from Kwangtung Province had been harvested and multiplied in contact with the local cultivars (hence probably *indica* type). This is in accordance with the morphological appearance and the observed heterozygosity in the 2 collections : from Kwangsi, most strains looked typically wild and had a low electromorph heterozygosity. From Kwangtung, all strains looked like morphological hybrids with cultivars and displayed a high electromorph heterozygosity : they may contain genes from *indica* type cultivars.

As to the collection that originated in Taiwan, it could have introgressed genes from *japonica* varieties introduced in the lowlands of Taiwan at the time of their collection (see OKA & CHANG, 1961 ; KIANG *et al.*, 1979). However, the present data suggest that it could also be of an original constitution close to the collection studied from Kwangsi Province.

The description of a Chinese form of *O. rufipogon* seems appropriate, in spite of the relatively limited and poorly documented collection that was studied, because it showed original features compared to the South and South-East Asian forms well represented in our collection.

All 3 geographic forms of *O. rufipogon* which are most closely related to the cultivated rice (South Asian, Chinese and American forms) were the most polymorphic of the whole *Sativa* group.

### B. An evolutionary interpretation

Now the question is how such differentiation arose. It seems reasonable to assume that divergence in isozymes is correlated with time of separation among isolated populations. The validity of using the molecular clock (« electrophoretic clock » in our case) in the study of evolution of *Oryza* was discussed earlier (SECOND, 1984 c, 1984 d). Reasonable estimate of dates (of divergence) which are consistent with paleogeography are obtained for both the *Sativa* and the *Latifolia* group of species.

As pointed out by SARICH (1977), attempts to use electrophoretic data as a clock should take into consideration that the rate of fixation of mutations varies greatly among loci. Accordingly, the accumulation of electrophoretic distance is not a linear function of time. In the present case, the loci *Pox-B*, *Est-B* and *E*, *Acp-1*, *Lap-E* and *Pgi-A* and *B* may be considered as rapidly evolving loci in view of their great polymorphism among various representatives of the genus *Oryza* (unpublished data). Other loci would be more slowly evolving ones although no clear limit can be fixed between the two classes. Accordingly, the medium curve of SARICH (1977, fig. 2) was tentatively used as a calibration of the clock. It gives a time around 15 millions years for the divergence of the Australian form of *O. rufipogon* and non Oceanian forms. This dating corresponds to the Miocene collision of Laurasia with the Australia/New Guinea shard of Gondwanaland (see WHITMORE, 1981), an event which is

believed to have created Malesia and allowed considerable land migration from Asia to Australasia (see KEAST, 1971 ; WHITMORE, 1981). It is thus very much plausible.

Assuming the reasonableness of the above calibration of the electrophoretic clock, we obtained the following estimated dates based on the median values of strain to strain distances between groups (tabl. 3) :

— 7 m.y. for the divergence between *O. longistaminata* and both *O. breviligulata* and Asian *O. rufipogon* ;

— 2 or 3 m.y. years for the divergence between the ancestors of *indica*, *japonica* and *glaberrima* type of cultivated rice.

These estimates are fully compatible with what is known about the paleoenvironment, information about which was used previously to develop an evolutionary scenario for the genus *Oryza* (SECOND, 1983, 1984 a, 1984 b). Only arguments but no proof can be put forward, so it is tentative only. It may be summarized as follows.

The above times of divergence indicate the existence of the *Sativa* group during the Miocene epoch, that is at a time when tropical and sub-tropical climates were distributed further away from the Equator than at present. This is consistent with the observation made in the introduction of a primitiveness of the tribe *Oryzeae* and with the adaptation to humid areas of the tropics of genus *Oryza*.

Because *Oryza* species adapted to forest environments are primarily found in Asia, or seem to have migrated very recently from there, an Eurasiatic distribution of the genus during the Tertiary period was assumed. The genus was introduced to Australasia as stated above but remained isolated from Asia according to the so-called « Wallace line ». Its introduction into Africa seems to have occurred through a relatively arid environment. Either a few founders migrated across an inhospitable environment or a continuous distribution from Asia to Africa was interrupted due to cooling of the climate and/or increasing aridity. In accordance with their ecological requirements, the possibilities of migration were interrupted earlier (7 m.y.) for the perennial *O. longistaminata* than for the annual *O. breviligulata* (2-3 m.y.). The particularly strong reproductive barrier found between the 2 African species could be due to its reinforcement after their distributions overlapped.

Migration of Hipparian fauna was possible across the Himalayan range until 2 or 3 millions years ago. Wild rice, at least annual forms, migrated along with the fauna. The climate was suitable for rice at that time in Central Asia. Wild rice probably existed in North China and certainly in Central China until historical time.

Migration of rice would have been possible between South Asia and the South-East Asian archipelago at the time of glaciations as the sea level was lowered (see fig. 1). However, a strong geographic barrier due to thickly forested mountains remained between South Asia and China until historical time.

The naturally big and abundant seeds of rice may have evolved as an adaptation to the habitat disturbance by the feet of large animals. The seeds buried deep in a muddy substratum which later dries up are protected from predators and parasites but need more energy to germinate. Man was attracted by the seeds and domesticated rice in Neolithic times. Domestication proceeded probably independently in

China, South Asia and Africa, as suggested earlier (SECOND, 1982) and by the present finding of a trend of differentiation, related to *japonica* type cultivars, of *O. rufipogon* in China. However, Man still harvests wild rice at present in Africa and in India. He also uses wild rice as fodder for animals and as a medicine.

The introduction of wild rice into America is certainly recent as judged from the similarities in isozyme patterns with their Asian relatives, particularly those of South East Asia. As wild rice seems to have been present in South America at the end of the 18<sup>th</sup> century (GUMILLA, 1791) it may have been introduced by the Spaniards during their journeys from South East Asia to South America. This is in contradiction with the conclusion of MORISHIMA (1969) who assumed independent evolution of the 4 geographic forms of the *Sativa* group isolated by Oceans. However, isozymes may be taken as better markers of phylogenetic relationships than morphological similarities. Study of the distribution of grass genera throughout the world has led to the conclusion that « trans-oceanic dispersal may occur over moderate distances, but that it does not provide a satisfactory general explanation for global distributions » (CLAYTON, 1975). This postulate leaves only 2 alternatives for the existence of wild rice in America : many millions of years of independent evolution or an introduction by Man a few hundred years ago. Clearly, our results support the second alternative. Hence, evolution of wild rice in America is a case of rapid morphological and ecological adaptation.

Extensive migration through the activity of Man is likely to have occurred both for wild and cultivated species within the Old World as well. An example of new forms arising from introgressions between 2 geographical groups may be in the New Guinean strains with overall affinity with Australian *O. rufipogon*, but also with electromorphs typical of Asian *O. rufipogon*. The figure 5 summarizes the proposed phylogenetic relationships for the *Sativa* group, based on the estimated divergence time and the recent interchanges between geographical isolates.

Asia appears to be a central area for the evolution of the *Sativa* group. It is 1) the relict area of a probably much more wide-spread distribution in Eurasia before climatic cooling ; 2) the area where domestication was most active and successful ; 3) the area where the disturbance of primitive habitats and primitive genetic and geographic structure due to the activity of Man has been most pronounced. The cumulative effects of these factors explain the complex intergrade of various genotypes found within Asian *O. rufipogon*.

Several authors have proposed that, in the past, distinction of an annual and a perennial life form in Asian *O. rufipogon* was more pronounced than at present. If such was the case, a recent break down of reproductive barriers or isolation due to transportation of seeds by Man, concomitant with the opening up of new ecological niches would explain the complex intergrade of genotypes found at present in Asian *O. rufipogon*. There is also a greater diversity within the Asian *O. rufipogon* than in the cultivated types, including *indica*, *japonica* and *glaberrima*. This can be explained if we assume that cultivation of rice started with the annual life form which, as discussed above may have been much more distinct than at present and may have migrated between Africa, South Asia and China long after the migration of the perennial form was interrupted. In other words, the cultivated forms may represent a relatively narrow genetic base, established early in the history of domestication of rice and maintained because of selfing and selection.



### Acknowledgements

The electrophoresis work was done during my stay in National Institute of Genetics, Japan, financially supported by O.R.S.T.O.M. and N.I.G. It could not have been achieved without the warm cooperation of Drs T. ENDO, S. IYAMA, H.I. OKA, H. MORISHIMA and Y. SANO. I thank Drs. J. DEJARDIN and L. MARCHAIS for their help and advices in the statistical analyses. I appreciate greatly the critical reading of the manuscript by Drs. J. ANTONOVICS, J. DAVID, H.I. OKA, G. VALDEYRON and one anonymous referee.

### References

- CHU Y.E., OKA H.I., 1970. Introgression across isolating barriers in wild and cultivated *Oryza* species. *Evolution*, **24**, 344-355.
- CHU Y.E., MORISHIMA H., OKA H.I., 1969. Reproductive barriers distributed in cultivated rice species and their wild relatives. *Jpn. J. Genet.*, **44**, 207-233.
- CLAYTON W.D., 1975. Chorology of the genera of *Gramineae*. *Kew Bulletin*, **30**, 111-132.
- ENDO T., MORISHIMA H., 1983. Current status of isozyme research in individual plant species. 6. - Rice. In : TANKSLEY S.D., ORTON J.J. (ed.), *Isozymes in Plant Genetics and breeding*, 318 pp., Elsevier Isozyme Monograph, Amsterdam.
- GUMILLA J., 1791. Historia natural, civil y geographica de las naciones situadas en las riberas del Rio Orinoco. 2 vol., Barcelona (citations in 38th Annual Report of Smithsonian Institution, Bureau of American Ethnology, 1916-1917, p. 219).
- KEAST A. (ed), 1981. *Ecological biogeography of Australia*. 2 vol., W. Junk bv publishers, The Hague.
- KIANG Y.T., ANTONOVICS J., WU L., 1979. The extinction of wild rice (*Oryza perennis formosana*) in Taiwan. *J. Asian Ecology*, **1**, 1-9.
- KUNG S., 1976. Tobacco fraction 1 protein : a unique genetic marker. *Science*, **191**, 429-434.
- MORINAGA T., 1964. Cytogenetical investigations on *Oryza* species. *Symposium sponsored by the International Rice Research Institute, Los Banos, Philippines, February 4-8, 1963. Rice genetics and cytogenetics*, 91-103, Elsevier, Amsterdam.
- MORISHIMA H., 1969. Phenetic similarity and phylogenetic relationships among strains of *Oryza perennis*, estimated by methods of numerical taxonomy. *Evolution*, **23**, 429-443.
- MORISHIMA H., OKA H.I., 1960. The pattern of interspecific variation in the genus *Oryza* : its quantitative representation by statistical methods. *Evolution*, **14**, 153-165.
- MORISHIMA H., OKA H.I., CHANG W.T., 1961. Directions of differentiation in populations of wild rice, *Oryza perennis* and *O. sativa f. spontanea*. *Evolution*, **15**, 326-339.
- NAYAR N.M., 1973. Origin and Cytogenetics of rice. *Adv. Genet.*, **17**, 153-292.
- NEI M., 1975. *Molecular population genetics and evolution*, 278 pp., North Holland, Amsterdam.
- NG N.Q., CHANG T.T., WILLIAMS J.T., HAWKES J.G., 1981. Morphological studies of Asian rice and its related wild species and the recognition of a new Australian taxon. *Biol. J. Linn. Soc.*, **16**, 303-313.
- OKA H.I., 1958. Intervarietal variation and classification of cultivated rice. *Ind. J. Genet. Pl. Breed.*, **18**, 79-89.
- OKA H.I., 1974. Experimental studies on the origin of cultivated rice. *Genetics*, **78**, 475-486.
- OKA H.I., CHANG W.T., 1961. Hybrid swarms between wild and cultivated species, *Oryza perennis* and *O. sativa*. *Evolution*, **15**, 418-430.
- PAI C., ENDO T., OKA H.I., 1973. Genic analysis for peroxydase isozymes and their organ specificity in *Oryza perennis* and *O. sativa*. *Can. J. Genet. Cytol.*, **15**, 845-853.

- PAI C., ENDO T., OKA H.I., 1975. Genic analysis for acid phosphatase isozymes in *Oryza perennis* and *O. sativa*. *Can. J. Genet. Cytol.*, **14**, 637-650.
- PIELOU E.C., 1969. *An introduction to mathematical ecology*. Wiley, Interscience, New York.
- SARICH V.M., 1977. Rates, sample sizes, and the neutrality hypothesis for electrophoresis in evolutionary studies. *Nature*, **265**, 24-28.
- 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., 1983. The study of isozymes in relation to the distribution of the genus *Oryza* in the Palaeoenvironment and the subsequent origin of cultivated rice. *Conference on the Palaeoenvironment of East Asia from the Mid-Tertiary. Hong-Kong, January 7-12, 1983. Evolution of East Asia Environment*, vol. II, 664-682, Centre of Asian studies, Hong-Kong.
- SECOND G., 1984 a. A new insight into the genome differentiation in *Oryza* L. through isozymic studies. In : SHARMA A.K., SHARMA ARCHANA (ed.), *Advances in Chromosomes and Cell Genetics*, New Delhi Oxford and IBH (in press).
- SECOND G., 1984 b. A different rates of genome divergence presumed between two species groups in the genus *Oryza* L. *The Nucleus*, **27**, 44-48.
- SECOND G., 1984 c. Geographical origins, genetic diversity and the molecular clock hypothesis in the *Oryzaeae*. In : JACQUARD P., HEIM G., ANTONOVICS J. (eds), *Genetic differentiation and dispersal in plants*, NATO Workshop Proceedings series, Springer Verlag (in Press).
- SECOND G., 1984 d. *Relations évolutives chez le genre Oryza et processus de domestication des riz*. 171 pp., Thèse Doctorat d'Etat. Université Paris-Sud, Orsay.
- SECOND G., MORISHIMA H., 1981. Mendelian segregation analysis at 3 isozyme loci in rice cultivars. *Natl. Inst. Genet. Mishima., 1980, Annual report*, p. 118.
- SECOND G., TROUSLOT P., 1980. Electrophorèse d'enzymes de riz (*Oryza* L.). *Travaux et Documents*, **120**, O.R.S.T.O.M., Paris.
- STEBBINS G.L., Jr., 1956. Cytogenetics and evolution of the grass family. *Am. Bot.*, **43**, 890-905.
- TATEOKA T., 1957. Miscellaneous papers on the phylogeny of *Poaceae* (10). Proposition of a new phylogenetic system of *Poaceae*. *J. Jpn. Bot.*, **32**, 275-287.
- WITHMORE T.C. (ed.), 1981. *Wallace's line and plate tectonics*, 90 pp., Clarendon Press, Oxford.