

Geographic Origins, Genetic Diversity and the Molecular Clock
Hypothesis in the Oryzae

G. SECOND

Biologie des Populations et des Peuplements, Centre Louis
Emberger, C.N.R.S., B.P. 5051. F-34033 Montpellier Cedex (France).
And Institut Français de Recherche Scientifique pour le
Développement en Coopération (ORSTOM), 24 rue Bayard. 75008 Paris.

ABSTRACT

The two species groups Sativa and Latifolia of the genus Oryza seem to be genetically independent in their variation. They have both a pan-tropical distribution. A study using isozyme electrophoresis of strains representing most of the area of distribution at 16 to 40 loci, shows a genetic structure in disagreement with that previously established on a morphological basis (Morishima 1969 and others). The isozyme structure shows, in both groups, a maximum divergence between Australian and other taxa, while American taxa are closely related to their Asian counterparts.

The application of the calibration of the electrophoretic clock by Sarich (1977) to the distances found points to a time of divergence for the Australian taxa compatible with geological data of a collision between Australasia and South-East Asia some 15 millions years ago. The same calibration gives estimates compatible with our knowledge on the palaeoenvironment in particular with regards to :

- 1)- The possibilities of migration between Eurasia and Africa (interrupted earlier for species of rice adapted to forest or inundated plain than for those species adapted to temporary pools in arid savannas)
- 2)- The emergence of the Himalayan barrier between China and South Asia.

The evolutionary picture which emerges is that of an Eurasian

Fonds Documentaire ORSTOM

Cote : Bx 15193 Ex : 1

Fonds Documentaire ORSTOM



010015193

NATO ASI Series, Vol. G5
Genetic Differentiation and Dispersal in Plants
Edited by P. Jacquard et al.
© Springer-Verlag Berlin Heidelberg 1985

origin of the genus with migration to Australia and to Africa during the Tertiary period and to America probably during the historical epoch. The morphological and adaptive divergence of these species in America would be a case of rapid adaptation to new ecological niches. Allotetraploid species of rice would be recently evolved species induced by the man's activities.

A way to further test the validity of the molecular clock hypothesis among the Oryzeae tribe is suggested. An explanation of why the opposite view of a direct adaptive significance of isozyme polymorphism has so often been envisaged in plants is put forward.

INTRODUCTION

The use of the diversity of genomes to determine a time scale according to the so-called molecular evolutionary clock hypothesis is an attractive possibility to explore the "fourth dimension".

An example may illustrate in particular the importance of a time scale in the interpretation of a genetic structure: the case of the subspecies differentiation of the common cultivated rice. It is presented in a simplistic frame in Fig. 1 which reflects two main ways of looking at the evolutionary paths leading to an observed differentiation at the subspecies level.

The existence of two subspecies among the common cultivated rice was evidenced in Japan in the 1920's. It appeared later that an intergradation of forms existed between them and their origin was considered in terms of disruptive selection in the course of domestication (Oka, 1974, 1982).

However, it was recently discovered that a large isozyme electrophoretic distance (D of Nei of the order of 0.3 over 40 loci) was also related to this subspecies differentiation. According to the widely accepted view of protein polymorphism evolution and also because no new isozyme was found in the African cultivated rice species compared to its immediate wild ancestor,

it was inferred that such an enormous amount of molecular differentiation should need more than a few thousands generations to accumulate.

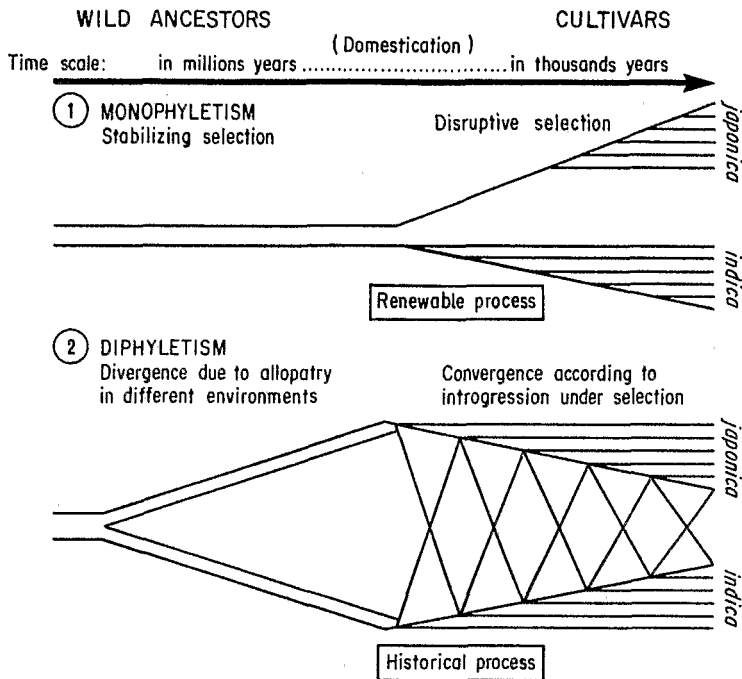


Fig. 1 - Two main hypothesis for the evolution of a subspecific differentiation with intergradation of forms. The case of the *indica* and *japonica* subspecies of cultivated rice is presented as an example. Hypothesis 1 is the author's interpretation of the ideas of Oka (1974, 1982). Hypothesis 2 is according to Second (1982).

The latter finding thus cast doubt on the former interpretation and pointed to a differentiation in the wild ancestors, previous to domestication (Second, 1982). This hypothesis was supported by the finding that Chinese wild rice are genetically differentiated from their South Asian counterpart, an information that was not previously available (Second, 1984a).

It is intended to show here that a judicious comparison of the genetically diversified geographical origins in the section *Fu-Oryza* of genus *Oryza* allows to obtain estimates of time of divergence which can be put together in a reasonable way with past

tectonic and climatic events in the last 20 millions years. Then it is suggested that the tribe *Oryzeae* is suitable to further test the molecular clock hypothesis in plants against vertebrates possibly in the last 100 millions years. Finally, a way to explain why isozymes so often seem to be associated with environmental factors, particularly in plants, is pointed out.

GEOGRAPHIC ORIGINS: THE TRIBE ORYZEAE AND THE SECTION EU-ORYZA

The tribe *Oryzeae* is a small tribe considered as primitive among the grasses (Stebbins 1956, Tateoka 1957, Clayton 1975). Its distribution, with generally different genera on different continents, suggests its existence at the time of the fragmentation of Gondwanaland (when grasses might have been quite unobtrusive in the plant cover and thus do not appear in the fossil record until the Eocene epoch).

Among the genus *Oryza*, two species groups have a pan-tropical distribution and we may thus ask whether the same geographic barriers have acted upon their isozyme differentiation in the same way, independently of their life history characteristics. Their species composition is outlined in Table 1. One group, the *Sativa* group comprises, as is suggested by its name, all cultivated forms and their closest wild and weedy relatives. They are all diploid and share only one genome (A) as reported in the literature on the basis of chromosome pairing at meiosis in *F1* hybrids.

The other, the *Latifolia* group, includes no domesticated form, but four genomes are recognized: B, C, D and E. The B, C and E genomes are known at the diploid level with their maximum distribution in Africa, South Asia and Australia, respectively. The D genome is not known at the diploid level. It was presumed to correspond to *O. officinalis* in China (Second, 1984b) although this remains to be proved as its cytogenetic relationship with other genomes has not yet been studied. Two allotetraploid genomes are known in nature: BC in Africa and Asia and CD, mainly in America but probably in Asia also.

Table 1. The species of section Eu-Oryza of genus Oryza

	Area of distribution	Biological type *	Genome **
SATIVA GROUP			
Cultivated species :			
O.sativa (two sub-species indica and japonica)	Asian origin	I	A
O.glaberrima	African origin	A	A
Wild species :			
O.rufipogon (complex species)			
- Oceanian form	Australasia	A or I	A
- Asian form	China, South Asia	A+I+P	A
- American form	Tropical America	I	A
O.longistaminata	Africa	P	A
O.breviligulata	Africa	A	A
LATIFOLIA GROUP (all wild)			
O.punctata(diploid)	Africa	A	B
O.punctata (tetraploid)	Africa	P	BC
O.minuta	South-East Asia	P	BC
O.malampuzahensis	Asia	P	BC
O.officinalis	Asia	P	C+D(?)
O.eichingeri	Africa	P	C
O.latifolia			
O.alta	Tropical	P	CD
O.grandiglumis	America		
O.australiensis	Australia	I	E

* A : annual ; I : intermediate ; P : perennial

** Genomes according to the observation of pairing of chromosomes at meiosis, as reported in the literature. For the full nomenclature of species, readers may consult Tateoka (1963).

GENETIC DIVERSITY: ISOZYME ELECTROPHORETIC STUDIES

Representatives of worldwide collections (except O.officinalis from China) were studied at a few discriminant isozyme loci, and a condensed collection was constituted on the basis of the taxonomical classification, the geographical or ecological origins and the preliminary isozyme survey. The condensed assembly of the strains was further studied by isozyme electrophoresis.

For the O.sativa group, 181 strains were analyzed at 24 isozyme

loci, including 15 enzyme systems commonly stained in plant isozyme studies (Second, 1984 b). Only one locus was monomorphic and up to 10 alleles were distinguished at a single locus with a mean of 4.1 allele per locus.

In the Latifolia group, 155 strains were screened at 8 loci and a condensed assembly of 25 strains was studied at 17 loci (Second 1984b). One locus (a different one than in the Sativa group) was also monomorphic. A mean of 3.3 alleles per locus were distinguished.

Nei's standard genetic distances (D) were computed from strain to strain in all possible combinations within groups in order not to make any a priori classification. While the maximum distances found were $D = 1.3$ in the two groups, the median values were much greater in the comparisons within the Latifolia group than within the Sativa group in accordance with the values of the average calculated heterozygosity ($1 - \sum x_i^2$, x_i = frequency of allele i). The table 2 summarizes these various data.

The data were further analysed through several methods of numerical taxonomy, using the presence or absence of alleles in particular strains or the matrix of distances between strains. They gave congruent results. Fig. 2 and 3 show the clusters of strains assembled on the plane defined by the two first vectors extracted in a Principal Coordinate Analysis of the matrix of Nei's distances calculated for pairs of strains in the Sativa and Latifolia groups, respectively.

In Fig. 2A, which includes the world accessions in the Sativa group, three clusters are separated. One includes the strains from Australia and some of the strains from New Guinea (O.rufipogon, Oceanian form); it is the most wide apart. Another one includes only the strains of the African perennial life form (O.longistaminata). The third cluster is composite, comprising all Asian and American strains and the other New Guinean strains (O.rufipogon, Asian and American forms), the African annual life form (O.breviligulata) and all cultivated forms (which all

clustered together with O.breviligulata).

TABLE 2. Various data concerning the isozyme electrophoretic analysis of section Eu-Oryza.

	Sativa group		Latifolia group	
	1 *	2	1	2
Number of strains or vegetatively propagated plants studied	181	25	25	25
Number of loci considered	24	15	17	15
Number of alleles per locus	4.1	3.6	3.3	3.3
Average calculated heterozygosity	0.32	0.32	0.47	0.45
Median value of Nei's distance between lines	0.33	0.39	0.49	0.51**
Maximum value of Nei's distance between lines	1.26	1.30	1.45	1.32

* In view of the fact that different numbers of loci and lines were studied in each group, two sets of data are presented for each group : 1) those concerning the entire assembly of individuals and loci studied, 2) a restricted assembly more liable for a comparison between groups with only, for both groups, 25 individuals chosen at random in the various species or geographic forms and the 15 homologous loci studied in common.

**0.95 when only the diploid strains were considered.

It was a surprise to note that the morphologically distinct American strains (Morishima 1969) shared all their frequent alleles with their Asian relatives. On the contrary, O.breviligulata, along with the African cultivated rice O.glaberrima, could be clearly distinguished from others on the basis of one specific allele Pgi-A3 as well as of their short ligule. Also, when only Asian strains were considered, as shown in Fig. 2B, it appeared that some Chinese accessions were clearly distinguishable from their South Asian relatives .

The differentiation between Chinese and South Asian strains was, for a large part, parallel to the differentiation of cultivars in indica and japonica type but the amount of variation was much greater in wild strains than in cultivated ones.

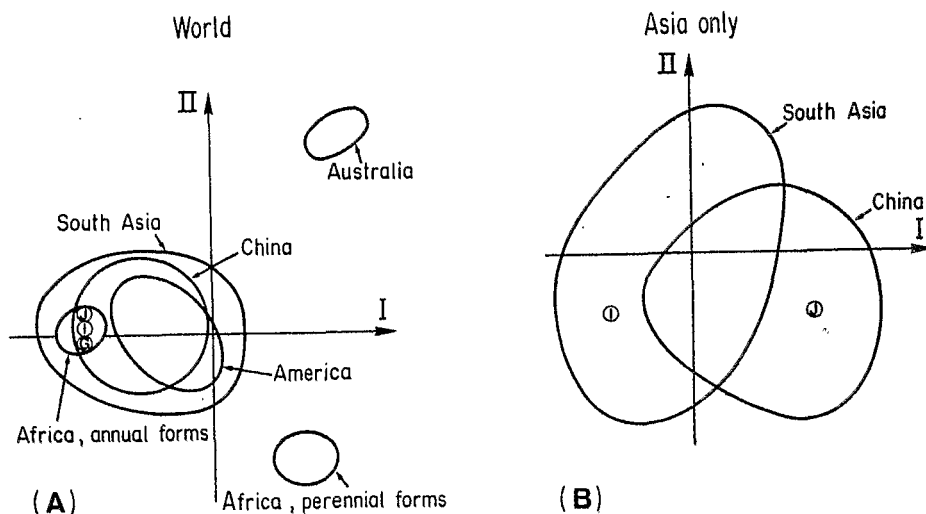


Fig. 2 . The geographic distribution of the spontaneous forms of the Sativa group as they cluster in the first plane of a principal coordinate analysis of the genetic diversity scored at 24 isozyme loci.

A)- 181 strains representative of the whole species group are considered.

B)- 110 strains originating in Asia are only considered.

The three basic isozyme patterns of the indica (I), japonica (J), and glaberrima (G) types of cultivated rice are plotted as summarary variables. The taxonomical classification is given in table 1.

Up to 43 isozyme loci involving 15 different enzyme systems were further stained in cultivated rice and in some wild representatives of the Sativa group (Second, 1982 and unpublished data). The genetic structure described above was only reinforced when more loci were considered. In short, four main areas of geographic differentiation appear in the Old World: Africa, South Asia, China and Australasia. The American taxa were distinguishable from the Asian ones on the basis of their genotypic combinations but not of any specific allele. Only within African accessions could the annual and perennial life forms be unambiguously distinguished on an isozyme basis.

In Fig. 3, which shows the distribution of strains of the

Latifolia group, diploid strains originating in Africa, South Asia and Australia were found at the extremes in the distribution while allotetraploid strains, but also some diploid strains belonging to the genome C, were found to be intermediate. With exceptions at only two loci (in the species O. minuta) the intermediate strains with BC and C genomes showed no allele of their own but alleles of both the B and C genomes. A putative D genome isozyme pattern was thus tentatively determined by comparing the isozymes found in strains with CD and C genomes. It showed specific alleles. Assuming, as stated above, the origin of D genome in China, the four areas of geographic differentiation noticed in the Sativa group, are thus remarkably found also in the Latifolia group.

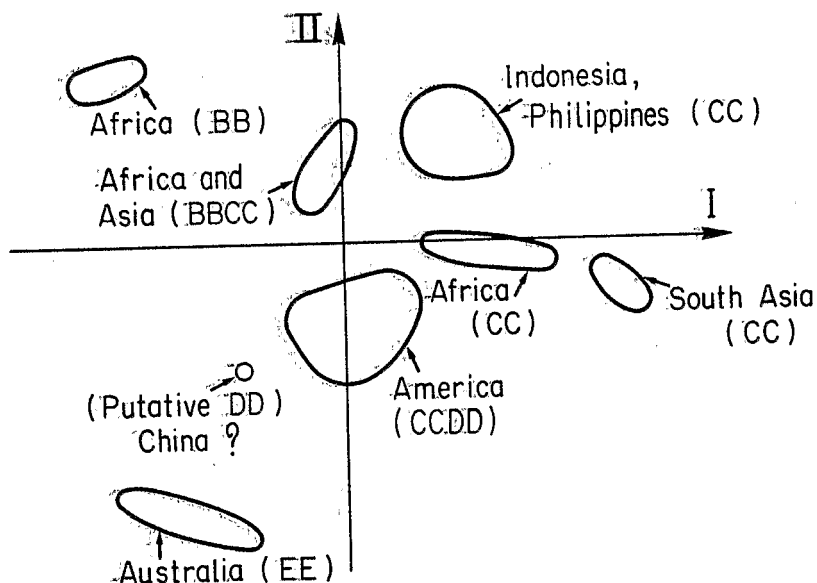


Fig. 3 - A similar study as in Fig. 2 of 26 strains of the Latifolia species group scored at 17 isozyme loci. The taxonomical classification according to the genomes or subgenomes is given in table 1.

THE ELECTROPHORETIC CLOCK AND A SCENARIO OF EVOLUTION

Turning to the interpretation, it appeared that, although apparent convergence in isozyme patterns between geographically isolated strains sometimes could be distinguished (as for example between the annual african O.breviligulata and all cultivated forms), no clear overall association between types of habitat (or life forms) and isozyme patterns could be seen at a continental scale. On the contrary, it appeared that a simple interpretation could be reached in the light of the mutation-random drift hypothesis of molecular evolution (Kimura, 1983); in other words, considering the divergence in isozyme patterns as reflecting a long time of genetic isolation (after the gene flow between the two groups of strains compared was suppressed). The intermediate patterns would then represent recently intermixed genotypes.

Sarich (1977) gave a good understanding of the molecular clock based on electrophoretic data and attempted its calibration in vertebrates on the basis of immunological distances related to paleontological data. His calibration (see the medium curve in Fig. 4 which corresponds to our mixed assembly of loci, Second 1984a and c) points, in our case to the Miocene epoch as the time of divergence to account for the maximum distances found in both the Sativa and Latifolia groups between Australian and non Oceanian strains.

This dating can be put together with the well documented collision of Australasia with South-East Asia in the course of its Northward rafting, some 15 millions years ago. The isolation of some populations of rice in the moving topography of the Malesian archipelago could have taken place at this epoch. These populations could have subsequently migrated up to Australia and remained isolated according to the so-called Wallace line. It seems thus reasonable to calibrate our electrophoretic clock upon this geological event which allows the following further determinations of times of divergence :

- 15 million years also between African and Asian representatives

of the Latifolia group.

- 7 million years between the perennial African (O. longistaminata) and Asian O. rufipogon.

- 2 or 3 million years between the ancestors of the three basic types of cultivated rice (indica, japonica and glaberrima), that is between african O. breviligulata and its South Asian and Chinese relatives.

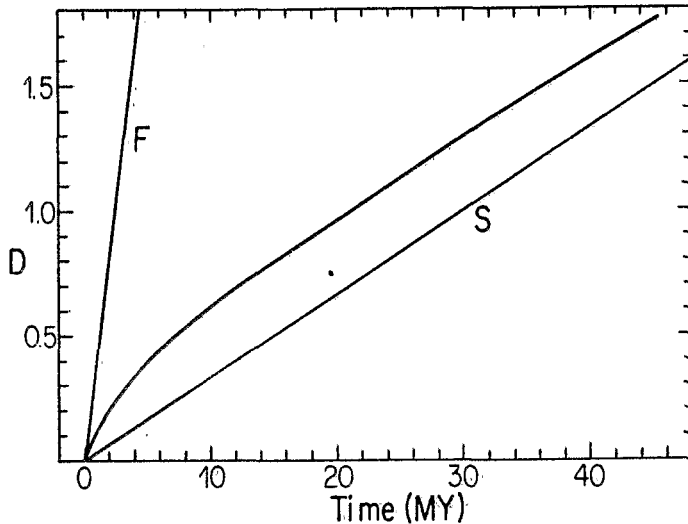


Fig. 4 - The accumulation of electrophoretic genetic distance (D) as a function of time for the rapidly evolving proteins (F), slowly evolving proteins (S), and a mixed system involving 3/4 of slowly evolving, and 1/4 of rapidly evolving proteins (unlabelled line) calculated according to Sarich (1977).

The migration was probably from Asia to Africa because the forest adapted species are found in Asia. With regards to the ecological requirements of the species concerned, these estimates are in agreement with the probable sequence for the establishment of a climatic barrier between Africa and Eurasia: tropical forest disappeared first, then humid savanna, and arid savanna before the establishment of desertic or temperate areas.

Further, the 2 or 3 million years of divergence found between the ancestors of the indica and japonica types of cultivated rice can

be put together with the geologically recent rise of the Himalayas which may not have become a barrier for the migration of the Hipparion (three-toed horse) fauna (hence of annual wild rice) until the Pliocene epoch.

On the other hand, although they appear as thoroughly naturalized, the American forms or species may represent very recent introductions. Arguments coming from different fields further support this finding (Second and Ghesquiere, in preparation).

The above stated apparent convergence of all cultivated rice with O. breviligulata could be due to their domestication from an annual life form which may have crossed the African climatic barrier and the Himalayas long after the migration of species adapted to forests or inundated plains was interrupted.

It is the reasonableness of the estimated dates that gives confidence in the use of the electrophoretic clock although the hypothesis of a universal molecular clock valid for plants as well as for animals and irrespective of generation lengths is rather controversial (see however Prager et al., 1976 and Hori and Osawa, 1979 for findings related to the question).

It seems that the validity of such a hypothesis could be further tested in the Oryzaeae tribe with regards 1) to geological events; 2) to other organisms whose distribution may have been affected in a parallel way by the same tectonic events.

To illustrate the approach proposed, one may consider 1) the possible phylogenetic tree for the genus Oryza (Fig. 5), starting some 100 million years ago, that can be taken as a working hypothesis related to geological events (Second 1984c), 2) the example of such groups as frogs which show large molecular diversity in spite of a low morphological and chromosomal polymorphism. In one example, the immunological distances found between representatives of mammals and frogs in Australia and America were comparable and allowed the determination of a time of divergence of 60 - 70 million years, consistent with the study of

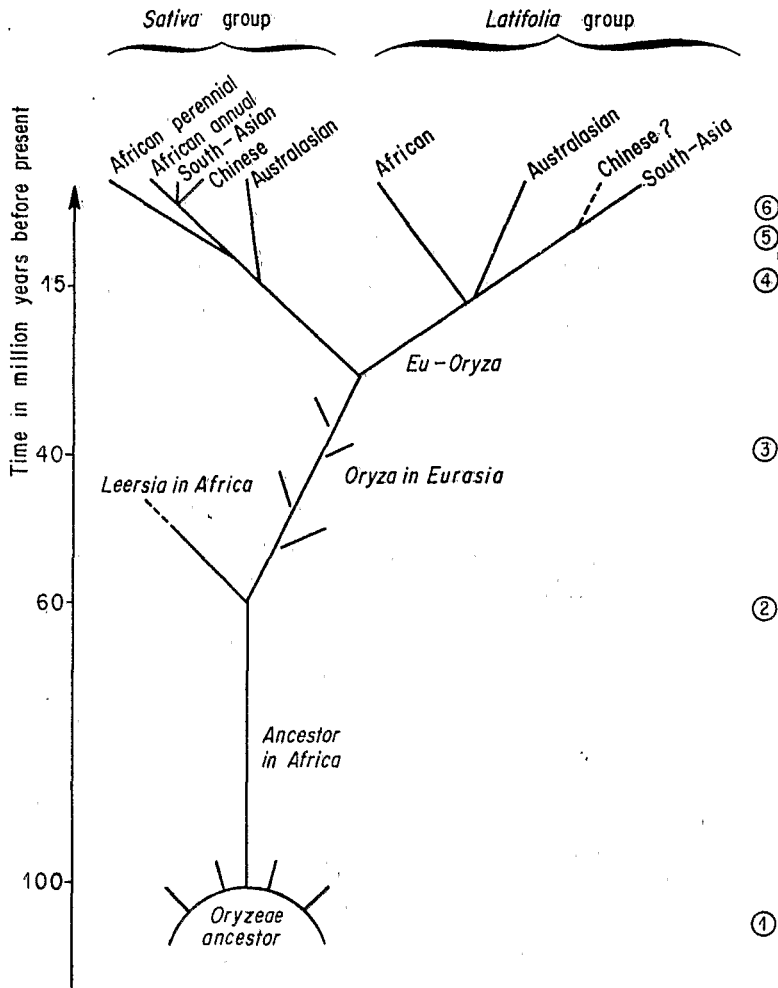


Fig. 5 - A proposed phylogenetic tree of the *Oryza* genus related to some important geological events in the Old World (Second 1984c): 1)- Progressive drift of Gondwana fragments with associated *Oryzae* ancestors. 2)- Temporary connection between Africa and Eurasia during Paleocene. Migration of the *Leersia* - *Oryza* ancestor into Eurasia. 3)- Land connection between India and Eurasia is established. 4)- Collision between Australasia and South-East Asia : formation of the Malesian archipelago. 5)- Progressive appearance of a climatic and sea barrier between Africa and Eurasia. 6)- Progressive emergence of the Himalayas as a barrier to plant migration.

plate tectonics (Wilson et al. 1977). A DNA or immunological approach should allow a comparison with representatives of the Oryzeae in Australia, Africa and South America.

In the present state of data, the picture which emerges for the Oryzeae is one of allopatric speciation. Introgression and allotetraploidization between species or semi-species would occur when genetic isolation has not been fully accomplished in spite of an independent evolution during geological ages and intermixing of diverging populations would happen through tectonic or climatic events and the intervention of animals or man. Obviously, this last factor has been of overwhelming importance for the microevolution and the recent dispersion of rice but also for many other wild or domesticated organisms.

In this scheme, the isozymes which are recognized neutral -or nearly neutral- to direct selection appear to be associated with adaptive traits because independent evolution occurs in different environments. It is thus not surprising that in phylogenetically heterogeneous species or populations, with introgressive hybridization following a long time of divergence, isozymes appear to behave as if selected by environmental factors.

As judged from paleobotanical data, the evolution of plant species has been a slow process at the organismic level since, say, the Miocene epoch, compared to what is observed in many Mammals. There are examples in the literature of plants whose ability to hybridize after a very long time of isolation was renewed when their distribution overlapped or their habitat was disturbed (Anderson and Stebbins, 1954).

Many studies which have showed clear environmental associations of isozyme diversity and environment type were conducted in Triticeae, Hordeae and Aveneae originated in the Near East or recently introduced in California from the Near East. In that connection, it should be kept in mind that the establishment of a Mediterranean climate in the Near East may be as recent as 11000

years ago (Wright, 1976). Mediterranean grasses previously isolated in North Africa, Southern Europe and Central Asia could have been brought together at that time, creating genetically heterogeneous populations.

It could be rightly argued that many cycles of hybridization-isolation may have occurred so that the differentiation we observe now has no direct relation with any time of divergence. In that respect, the Oryzæae, with evident strong geographic barriers to migration which might have been broken down for some of the species only since the cultural evolution of Man, may be a good material to investigate divergence and convergence in the differentiation of plant population as well as of multigenic families within individuals.

REFERENCES

- ANDERSON E, STEBBINS GL (1954) Hybridization as an evolutionary stimulus. *Evolution* 8: 378-388.
- CLAYTON WD (1975) Chorology of the genera of Gramineae. *Kew bulletin* 30, 1: 111-132.
- HORI H, OSAWA S (1979) Evolutionary change in 5S RNA secondary structure and a phylogenetic tree of 54 5S RNA species. *Proc Natl Acad Sci USA* 76: 381-5.
- KIMURA M (1983) The neutral theory of molecular evolution. Cambridge University Press, p 367.
- MORISHIMA H (1969) Phenetic similarity and phylogenetic relationships among strains of Oryza perennis, estimated by methods of numerical taxonomy. *Evolution* 23: 429-443.
- OKA HI (1974) Experimental studies on the origin of cultivated rice. *Genetics* 78: 475-486.
- OKA HI (1982) Phylogenetic differentiation of cultivated rice, XXIII. Potentiality of wild progenitors to evolve the indica and japonica types of rice cultivars. *Euphytica* 31: 41-50.
- PRAGER EM, FOWLER DP, WILSON AC (1976) Rates of evolution in conifers (Pinaceae). *Evolution* 30: 637-649.

- 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: 25-27.
- SECOND G (1984a) (to be published) Evolutionary relationships in the Sativa group of Oryza based on isozyme data. Genet. Sel. Evol. 17-1.
- SECOND G (1984b) (to be published) A new insight into the genome differentiation in Oryza L. through isozymic studies. In Sharma AK, Sharma Archana (eds) Advances in Chromosomes and Cell Genetics. New Delhi Oxford and IBH.
- SECOND G (1984c) Relations évolutives chez le genre Oryza et processus de domestication des riz. Thèse Doctorat Etat. Univ. Paris XI Orsay.
- STEBBINS GL Jr (1956) Cytogenetics and evolution of the grass family. Amer. Jour Bot 43: 890-895.
- TATEOKA T (1957) Miscellaneous papers on the phylogeny of Poaceae (10). Proposition of a new phylogenetic system of Poaceae. J. Japan Bot 32,9: 275-287.
- TATEOKA T (1963) Taxonomic studies of Oryza. III. Key to the species and their enumeration. Bot Mag Tokyo 76: 165-173.
- WILSON AC, CARLSON SS, WHITE TJ (1977) - Biochemical evolution. Ann Rev Biochem 46: 573-639.
- WRIGHT HE (1976) The environmental setting for plant domestication in the Near East. Science 194: 385-389.

En application de la Loi du 11 mars 1957 et notamment son article 41, la présente reproduction est effectuée en lieu et place d'une transcription manuelle, elle est strictement destinée à des fins de recherche et est réservée à l'usage privé du demandeur.

Le Centre National de la Recherche Scientifique n'est pas responsable de l'usage qui pourrait être fait de la copie demandée, en contravention de la législation sur la propriété littéraire et artistique (Loi n° 57.298 du 11 mars 1957).