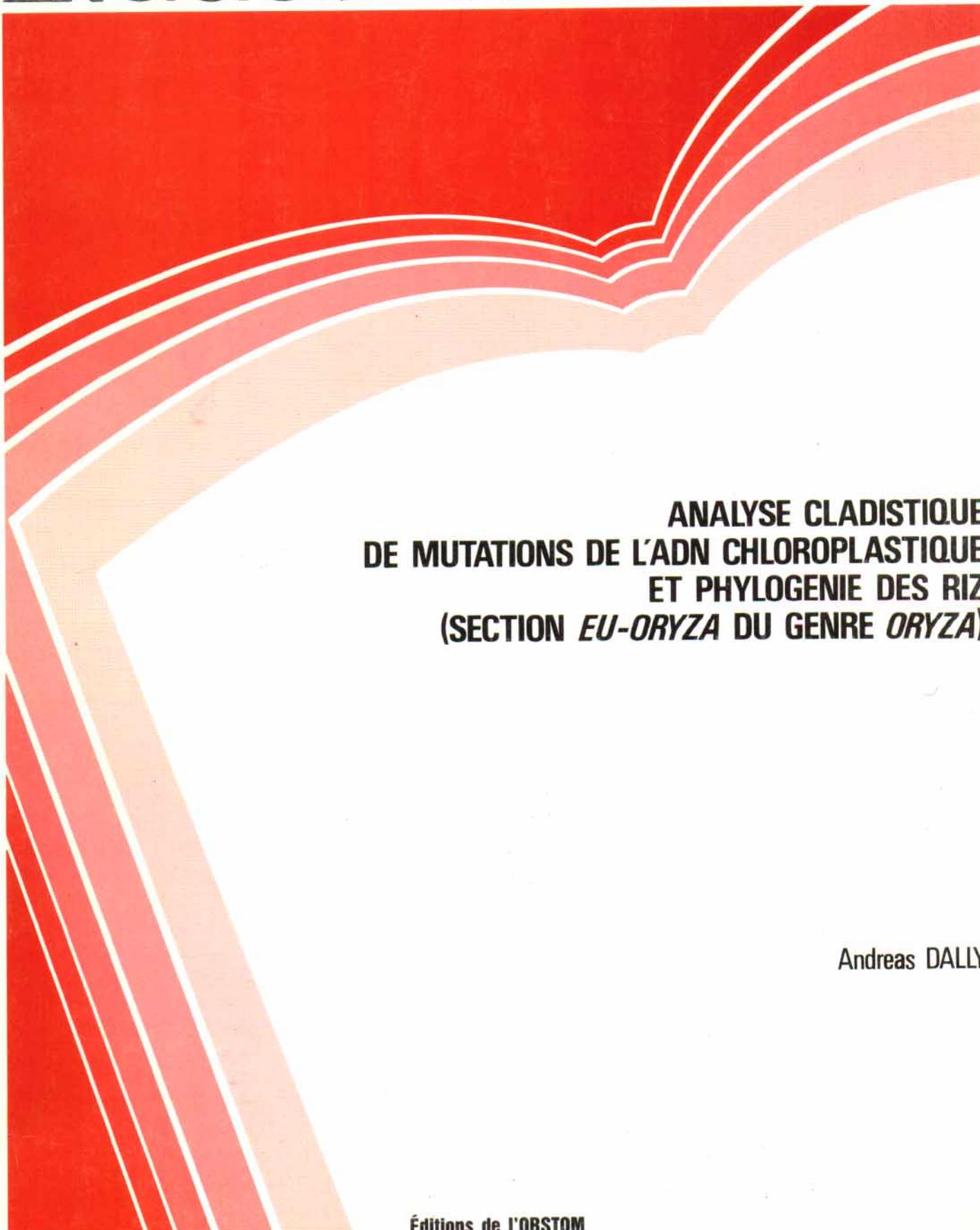


# Etudes et Thèses



**ANALYSE CLADISTIQUE  
DE MUTATIONS DE L'ADN CHLOROPLASTIQUE  
ET PHYLOGENIE DES RIZ  
(SECTION *EU-ORYZA* DU GENRE *ORYZA*)**

Andreas DALLY

Éditions de l'ORSTOM

INSTITUT FRANÇAIS DE RECHERCHE SCIENTIFIQUE POUR LE DÉVELOPPEMENT EN COOPÉRATION

Andreas DALLY

**ANALYSE CLADISTIQUE  
DE MUTATIONS DE L'ADN CHLOROPLASTIQUE  
ET PHYLOGENIE DES RIZ  
(SECTION *EU-ORYZA* DU GENRE *ORYZA*)**

---

**Éditions de l'ORSTOM**  
INSTITUT FRANCAIS DE RECHERCHE SCIENTIFIQUE POUR LE DÉVELOPPEMENT EN COOPÉRATION

Collection **ÉTUDES et THÈSES**

PARIS 1988

Cette étude a fait l'objet d'une thèse en biologie, intitulée  
"Polymorphisme des Longueurs des Fragments de Restriction de l'ADN  
chloroplastique dans la Section Eu-Oryza du genre Oryza (Riz)  
et Implications phylogénétiques",  
dirigée par M. Gérard SECOND et soutenue le 9 juin 1988  
devant l'Université des Sciences et des Techniques du Languedoc  
- Académie de Montpellier II -

< La loi du 11 mars 1957 n'autorisant, aux termes des alinéas 2 et 3 de l'article 41, d'une part, que les  
< "copies ou reproductions strictement réservées à l'usage privé du copiste et non destinées à une utilisation  
< collective" et, d'autre part, que les analyses et les courtes citations dans un but d'exemple et d'illustration,  
< "toute représentation ou reproduction intégrale, ou partielle, faite sans le consentement de l'auteur ou de  
< ses ayants droit ou ayants cause, est illicite" (alinéa 1<sup>er</sup> de l'article 40).

< Cette représentation ou reproduction, par quelque procédé que ce soit, constituerait donc une contre-  
< façon sanctionnée par les articles 425 et suivants du Code Pénal >.

---

ISSN : 0767-2888  
ISBN : 2-7099-0932-4

---

© ORSTOM 1988

# SOMMAIRE

	page
Introduction .....	7
Matériel et Méthodes .....	11
1) Les plantes .....	11
2) La préparation de l'ADN chloroplastique .....	13
a) Généralités.....	13
b) Protocole d'extraction de l'ADNcp .....	13
c) Remarques au protocole d'extraction d'ADNcp.....	17
d) Complément au protocole d'extraction: traitement des déchets .....	19
e) Une autre possibilité d'extraction non-aqueuse .....	19
3) La digestion de l'ADNcp par les enzymes de restriction .....	20
4) L'électrophorèse de l'ADNcp .....	21
5) La coloration et photographie des gels d' électrophorèse .....	22
6) La lecture des gels d'électrophorèse .....	23
7) L'analyse des mutations .....	25
a) Généralités.....	25
b) La détection des différents types de mutations .....	26
8) La construction d'un cladogramme sur la base des relations définies par mutations entre les différents types moléculaires .....	28
9) La construction des arbres phylogénétiques hypothétiques .....	29
a) Une méthode de repérage par paires non-pondérées .....	29
b) Une méthode de repérage par paires pondérées .....	30
Les Résultats .....	31
1) Résultats concernant les techniques et méthodes .....	31
a) Amélioration de la technique non-aqueuse de purification des ADNcp appliquée au riz .....	31
b) La non-spécificité au riz de la technique de préparation de l'ADNcp.....	32
c) La régularité des électrophorèses .....	34
d) La qualité des préparations d'ADNcp .....	34
2) Les patrons de restriction des ADNs chloroplastiques du riz .....	36
a) Les patrons <u>Eco R I</u> .....	37 + 39
b) Les patrons <u>Ava I</u> .....	37 + 41
c) Les patrons <u>Bam H I</u> .....	38 + 43

d) Les patrons <u>Hind III</u> .....	38 + 44
e) Les patrons <u>Sma I</u> .....	38 + 45
f) Les patrons <u>Pst I</u> .....	38 + 46
g) Les patrons <u>Bst E II</u> .....	38 + 47
h) Polymorphisme à l'intérieur d'une plante .....	48
i) Résolution de la distinction des différents plastotypes .....	48
j) La régularité entre électrophorèses .....	49
L'Analyse des Mutations .....	51
a) Premier niveau d'intégration .....	53
b) Deuxième niveau d'intégration .....	55
c) Troisième niveau d'intégration .....	57
d) Quatrième niveau d'intégration .....	58
e) Cinquième niveau d'intégration .....	60
f) Sixième niveau d'intégration .....	62
g) Dernier niveau d'intégration .....	63
h) Remarques sur les coïncidences et réversions .....	64
Remarques sur les Résultats .....	71
1) Comparaisons avec des résultats antérieurs d'analyse des ADNcp du riz .....	71
a) Détermination des tailles des fragments de restriction .....	71
b) Interprétations des PLFR	
A) Comparaison avec l'analyse d'ICHIKAWA, HIRAI et KATAYAMA .....	72
B) Comparaison avec l'analyse d'ISHII, TERACHI et TSUNEWAKI .....	73
2) La situation des variétés cultivées mâle-stériles cytoplasmiques .....	75
3) Polymorphisme intravariétal .....	76
Implications Phylogénétiques de l'Analyse des Mutations .....	77
1) Transformation du cladogramme en phylogrammes hypothétiques .....	77
2) Interprétabilité du phylogramme .....	81
a) Justification des branches de vitesse 0 .....	82
b) Fiabilité des relations évolutives entre les plastotypes .....	82
Insertion: Aspects de la génétique de transmission des ADNcp .....	82
c) Le calibrage de l'échelle de temps et le scénario d'évolution .....	85
i) L'estimation du taux de mutations attendu .....	86
ii) Calibrage du phylogramme par le taux de mutations attendu .....	88
iii) Comparaison des deux approches de calibrage .....	89
iv) Conséquences du calibrage .....	91
Conclusion .....	95

Annexe A: Liste des plantes analysées dans la section <u>Eu-Oryza</u> .....	97
Annexe B: Liste des plastotypes déterminés dans la section <u>Eu-Oryza</u> .....	107
Annexe C: Résultats d'électrophorèse d'ADNcp .....	121
Annexe D: Restitutions dessinées des patrons de restriction d'ADNcp .....	135
Annexe E: Protocole de la digestion de l'ADNcp par l'enzyme <u>Msp I</u> suivie par marquage radioactif au <sup>35</sup> S et électrophorèse en gel polyacrylamide à gradient de densité continu .....	140
Références Bibliographiques .....	143
Bibliographie .....	147
Remerciements .....	155

## INTRODUCTION

Leurs propriétés remarquables de se transmettre uniparentalement, sans recombinaison, et de ségréger végétativement, toujours en haploïdie, ont fait des génomes cytoplasmiques des marqueurs appréciés par les généticiens. Chez les plantes, ce sont notamment les marqueurs moléculaires des plastomes qui alimentent des recherches évolutives, taxonomiques et phylogénétiques importantes, qui actuellement se concentrent sur les analyses par enzymes de restriction de l'ADN chloroplastique (ADNcp) (voir PALMER 1985 et 1987 pour des revues récentes). C'est d'ailleurs sa forte conservation au cours de l'évolution et la simplicité de son organisation qui prédestine l'ADNcp, par rapport à l'ADN mitochondrial (ADNmt), pour ce genre d'études.

Malheureusement les avantages théoriques de cette approche sont contrariés par nombre de difficultés pratiques, avant tout celle de la purification rapide de l'ADNcp en quantités facilement manipulables.

Au niveau de l'ADNcp, à notre connaissance, une seule analyse impliquant un grand nombre d'individus fut publiée en 1985 par BANKS et BIRKY. Ils ont étudié 134 sites de restriction (par 7 enzymes) sur cent individus, issus de 21 populations de Lupinus texensis, une Papillonacée sauvage. Deux études portaient sur un grand nombre de mutations:

- l'un par SYTSMA et GOTTLIEB (1986), qui ont décélé 119 mutations de sites parmi 9 représentants d'espèces de Clarkia observées à l'aide de 29 enzymes de restriction;

- l'autre par PERL-TREVES et GALUN (1985), qui déduisirent l'arbre phylogénique le plus parcimonieux sur les plastomes de 20 représentants d'espèces sauvages et 6 de variétés cultivées de Cucumis, à partir de 73 mutations détectées.

Les Graminées, avec leurs feuilles peu pulpeuses et souvent très fibreuses ou

cutinisées à l'état adulte, comptent en bonne partie parmi les plantes problématiques. Pour celles-ci, un certain progrès fut atteint avec la mise au point d'une méthode non-aqueuse pour la préparation des chloroplastes de blé par BOWMAN et DYER (1982), qui, outre l'augmentation des rendements augmentés en ADNcp, présente l'avantage d'être adapté au matériel lyophilisé, rendant les extractions indépendantes de la date de récolte des feuilles.

Une autre voie vers une plus grande efficacité de la détection directe de l'ADNcp, c'est à dire en évitant les hybridations avec des sondes selon SOUTHERN (1975) fut poursuivie par LEHVÄSLAIHO, SAURA et LOKKI (1987) chez des Graminées de la famille des Festucées. Ils sont parvenus à réduire la quantité d'ADNcp nécessaire pour une digestion de restriction à 40 ng (au lieu de 1 ou 2 µg), en marquant les bouts collants des fragments au <sup>35</sup>S par des nucléotides libres, à l'aide de l'enzyme de Klenow, et en exposant les gels d'électrophorèse séchés à l'autoradiographie.

Le blé fut parmi les premières plantes étudiées pour leur patron de restriction de l'ADNcp (VEDEL, QUETIER, BAYEN 1976) et il donna lieu à plusieurs études phylogénétiques (VEDEL et al. 1978; BOWMAN, BONNARD, DYER 1983; TERACHI, OGIHARA, TSUNEWAKI 1984).

Les chloroplastes d'autres céréales furent également analysés sur le plan de l'ADN, comme le le mil (CLEGG, RAWSON, THOMAS 1984), l'orge (CLEGG, BROWN, WHITFIELD 1984; HOLWERDA, JANA, CROSBY 1986) et le maïs (ZURAWSKI, CLEGG, BROWN 1984; DOEBLY 1987). L'avoine fit partie d'une comparaison avec ces trois dernières céréales par les ADNcp et mt (VEDEL, LEBACQ, QUETIER 1980).

Le riz fut inclus dans une étude sur les relations phylogénétiques parmi dix céréales par comparaison des patrons de fragments de restriction générés par les enzymes Kpn I, Pst I, et Sal I, par ENOMOTO, OGIHARA et TSUNEWAKI en 1985.

Egalement en 1985 fut publiée la première carte physique de l'ADNcp d'un riz cultivé, enregistrant à côté des sites de restriction Sal I, Pst I et Pvu II les localisations des deux gènes principaux (i.e. pour la grande sous-unité de la rubisco et pour la protéine de 32 KD du centre de réaction du photosystème II), par HIRAI, ICHIKAWA, IWATSUKI et SUGIURA (1985).

En 1986 deux travaux furent consacrés à l'ADN chloroplastique du riz:

- ISHII, TERACHI et TSUNEWAKI (1986) se concentraient sur une comparaison de certains fragments de restriction, générés par les enzymes Eco R I, Hind III et Pst I,

parmi 30 variétés de riz cultivé. Ils décrivaient deux types de patrons pour chaque enzyme, une fois distinguant Oryza sativa par rapport à Oryza glaberrima (Eco R I), et pour les deux autres enzymes O. sativa japonica avec O. glaberrima par rapport à O. sativa indica.

- ICHIKAWA, HIRAI et KATAYAMA (1986) décrivaient une observation de 9 représentants de 7 espèces du genre Oryza sur une dizaine de bandes Bam H I ensemble avec des résultats d'électrophorèse en électro-focalisation de la protéine fraction I sur 19 représentants de 13 espèces. Entre autres, ils montraient une différence intra-spécifique dans O. latifolia, une espèce sud-américaine portant un génome CCDD.

Le riz semblait cependant compter parmi les espèces difficiles à étudier avec les méthodes aqueuses, peut-être à cause de ses petites cellules et de sa haute teneur en matière sèche. La méthode aqueuse était utilisée après broyage d'un grand nombre de jeunes plantules dans l'azote liquide.

Notre objectif d'analyser un grand nombre de variétés et d'espèces nous a conduit à choisir la méthode non-aqueuse comme étant a priori la plus apte à permettre l'extraction d'ADNcp à partir de plantes uniques - donc adultes -, sans compter les possibilités de conservation et d'expédition du matériel qu'elle permet. Nous l'avons modifiée pour atteindre cet objectif et pour réduire son inconvénient principal: l'utilisation en grandes quantités de solvants dangereux.

La présente étude poursuit l'analyse des relations phylogénétiques entre ADNcp, en restant à l'intérieur de la section Eu-Oryza du genre. Elle complète les résultats obtenus sur les marqueurs moléculaires nucléaires, notamment par SECOND en électrophorèse d'isozymes (pour des synthèses voir SECOND 1985 a et b).

Nos analyses se basent principalement sur:

- les patrons Eco R I des ADN cp de 320 plantes individuelles, appartenant à 248 variétés de 15 espèces,
  - les patrons Ava I de 240 de ces plantes,
  - les patrons Sma I et Bst I d'un choix de 36 variétés, ainsi que
  - les patrons Pst I, Hind III et Bst E II d'un choix de 19 variétés,
- développés sur des gels d'agarose de 24 cm de longueur, qui seront tous quantifiés et précisément examinés pour des mutations. Les patrons Sal I de la quasi-totalité

des plantes, obtenus sur mini-gels, et les patrons Msp I, générés en gels de polyacrylamide à gradient de densité continu, ont permis quelques observations qualitatives.

Le lecteur trouvera d'abord une présentation de la technique préparative de l'ADNcp utilisée, une technique non-aqueuse modifiée, dont la mise au point fut ouverte la voie à des extractions efficaces et, par là, à l'élargissement et à un certain raffinement de l'échelle d'analyse à la fois.

Le rapport des résultats bruts sera suivi par une première étape d'interprétation, l'analyse des mutations. De celle-ci sera déduit un cladogramme, intégrant les différents plastotypes observés.

Finalement, au travers de quelques estimations quantitatives de certains paramètres concernant la génétique de transmission des plastomes dans les populations naturelles, la possibilité d'une traduction de ce phénogramme en un arbre phylogénique sera discutée.

Ce travail s'inscrit dans le cadre des recherches sur l'évolution du riz, en faveur de la connaissance et pour la conservation et l'utilisation des ressources génétiques, conduit par l'ORSTOM (Institut Français de Recherche Scientifique pour le Développement en Coopération). Il a été réalisé dans les laboratoires du CEPE / CNRS (Centre d'Etudes Phytosociologiques et Ecologiques - Centre National de la Recherche Scientifique) à Montpellier.

## MATERIEL ET METHODES

### 1) Les plantes

Toutes les plantes incluses dans la présente étude proviennent de la culture d'une collection vivante entretenue par le Docteur G. SECOND au C.N.R.S à Montpellier sous forme de graines conservées au froid et en partie en état végétatif en serre.

Il s'agit de représentants des espèces suivantes:

A) dans le **Groupe Sativa**:

Cultivés:

- Oryza sativa: 75 variétés par 92 individus, dont
  - a) O. sativa indica: 21 variétés par 27 individus;
  - b) O. sativa japonica: 42 variétés par 49 individus;
  - c) mâle-stériles: 11 variétés par 14 individus;
  - d) adventice de type Purpure: 1 variété par 2 individus;
- Oryza glaberrima: 7 variétés par 13 individus;

Sauvages ou adventices:

- Oryza rufipogon: 41 variétés par 57 individus, dont originaires
  - a) de l'Inde: 11 variétés par 16 individus;
  - b) de l'Asie de Sud-Est: 5 variétés par 8 individus;
  - c) de la Chine: 10 variétés par 11 individus;
  - d) de l'Amérique: 5 variétés par 9 individus;
  - e) de l'Australie: 10 variétés par 13 individus;
- Oryza longistaminata: 15 variétés par 18 individus;

- Oryza breviligulata: 8 variétés par 14 individus;

B) dans le Groupe Latifolia:

- Oryza punctata: 16 variétés par 27 individus, dont porteurs

a) d'un génome BB: 6 variétés par 13 individus;

b) d'un génome BBCC: 10 variétés par 14 individus;

- Oryza minuta: 6 variétés par 7 individus, dont porteurs

a) d'un génome BB: 2 variétés par 2 individus;

b) d'un génome BBCC: 4 variétés par 5 individus;

- Oryza malampuzhaensis (BBCC, Inde): 1 variété par 2 individus;

- Oryza officinalis: 49 variétés par 54 individus, dont porteurs

a) d'un génome BBCC\* : 5 variétés par 7 individus;

b) d'un génome CC: 44 variétés par 47 individus;

- Oryza eichingeri: 6 variétés par 8 individus;

- Oryza latifolia: 11 variétés par 16 individus;

- Oryza latifolia, type paraguayensis: 1 variété par 2 individus;

- Oryza alta: 2 variétés par 2 individus;

- Oryza grandiglumis: 3 variétés par 3 individus;

- Oryza australiensis: 5 variétés par 9 individus.

En général quand il y a plusieurs individus par variétés, ceux-là ne proviennent pas de la même plante mère. Mais ceci ne constitue pas un problème, puisque toute la variabilité intravariétale rencontrée se limite sur une seule variété mâle-stérile d'O. sativa\*\*.

Dans tous les autres cas, les différents exemplaires d'une même variété se comportaient comme des vraies soeurs.

Une liste détaillée de la totalité des échantillons examinés se trouve dans l'annexe A.

\* Il est possible que ces échantillons représentent des erreurs dans la collection d'origine au Japon, car O. officinalis a été décrit seulement avec le génome CC

\*\* voir page

## 2) La préparation de l'ADN chloroplastique

### a) Généralités

La méthode non-aqueuse pour l'isolation du matériel chloroplastique permettant la purification de l'ADNcp, mise au point par BOWMAN et DYER (1982), introduisit un nombre d'avantages sur les méthodes aqueuses courantes. Outre l'amélioration des rendements d'ADN, elle rend possible le stockage du matériel végétal pour des périodes de temps étendues, même aux températures ambiantes, ce qui facilite par exemple son expédition par poste. Il n'y a plus de besoin d'exécuter les extractions immédiatement pendant la bonne saison de croissance des plantes. Certaines contraintes sur la procédure d'extraction elle-même sont également supprimées grâce à l'inactivité des enzymes telles que DNAses en milieu non-aqueux.

Un inconvénient de cette technique réside dans la corrélation inverse entre les rendements en ADN cytoplasmiques nets et les contaminations par de l'ADN nucléaire (ADNn) - problème particulièrement délicat dans notre expérience sur le riz: trop souvent après l'électrophorèse une bonne partie des fragments de restriction de l'ADNcp se trouve dérobée par un voile opaque d'ADN chromosomique (cf. fig. 1 dans l'annexe C). Un autre inconvénient lié à la sécurité au laboratoire est l'utilisation massive des solvants dangereux tétrachlorure de carbone et n-hexane, tous les deux neurotoxiques et cancérogènes.

Si aujourd'hui cette technique s'applique avec efficacité sur le riz, c'est grâce à la mise au point de certaines modifications par rapport à la méthode originale. Elles seront expliquées après la présentation du protocole d'extraction actuel.

### b) Protocole d'extraction de l'ADNcp \*

1) D'une plante adulte ou juvénile 10 à 12 grammes de feuilles sont coupés, immédiatement emballés dans un sac plastique et plongés dans l'azote liquide; une conservation prolongée est ensuite possible à -80°C ou même -20°C.

2) Après pesée, les feuilles toujours congelées sont pulvérisées dans un

---

\* Le lecteur qui n'est pas intéressé à reproduire la technique est prié d'excuser la forme détaillée de ce protocole

broyeur vibrant à billes en acier, refroidi par l'azote liquide, pendant 2,5 minutes (broyeur type Dangoumill 300 ; fréquence de va-et-vient 700 coups/min; pot acier, capacité 300 ml; jeu de billes, diamètres: 20 mm (1x), 12 mm (4x), 8 mm (9x)).

3) En évitant sa décongélation la poudre est transférée dans un cristalliseur et celui-ci couvert d'un papier poreux à l'aide d'un élastique; de nouveau une conservation longue est possible à  $-80^{\circ}\text{C}$ .

4) Le matériel est lyophilisé. Sa conservation se fait ensuite sous vide à  $-20^{\circ}\text{C}$ , sur  $\text{P}_2\text{O}_5$  dans un désiccateur - jusqu'à plus d'un an dans notre expérience, mais des conservations beaucoup plus longues sont sans doute possibles.

5) Après la décongélation sous vide, évitant toute condensation ou adsorption d'eau, la poudre sèche (environ 2 à 4 grammes) est divisée en deux ou trois portions. Une par une celles-ci sont transférées dans un tube de centrifugation de 50 ml, diamètre intérieur 24 mm, et chaque fois 10 ml d'un mélange de solvants sont ajoutés. Ce mélange consiste de 4 volumes de 1,2,4-Trichlorobenzène (purum, Fluka; densité 1,45) et d'un volume de Decahydronaphthalène (cis + trans) (purum, Fluka; densité 0,89).

6) Pour chaque portion, la suspension de la poudre et l'homogénéisation de l'ensemble sont effectués par l'action d'un homogénéiseur Polytron (équipé avec la sonde: PMA 20 TN) à la vitesse la plus faible pendant 15 à 60 secondes - selon la coriacité du tissu.

7) L'homogénéisat est versé sur un filtre nylon (maille 25  $\mu\text{m}$ ) placé dans un entonnoir. La filtration est aidée en remuant et par pression avec une spatule et un pilon. Par cette voie, les filtrats des homogénéisats des deux ou trois portions sont réunis dans un autre tube identique de centrifugation.

8) Les résidus réunis des filtrations sont retransférés dans le tube de broyage, remêlés avec 5 ml du mélange 4:1 et resoumis aux pas 6) et 7) (et 8) à volonté). Le volume final du filtrat fait environ 30 ml. Il se conserve plusieurs jours dans le réfrigérateur.

9) Le filtrat est réfrigéré pendant 5 minutes dans un bain d'eau à  $4^{\circ}\text{C}$ , ensuite prudemment superposé par 10 ml du mélange 4:1 susdit, ce qui produit un gradient de "propreté". Le gradient est fermé par une couche de 2ml d'un mélange 1:3 (v/v) des solvants susdits.

10) Une centrifugation est exécutée dans un rotateur à pendulation libre à 2.300 g pendant 17 minutes à  $4^{\circ}\text{C}$ . Cette centrifugation provoque l'accumulation de matériel chloroplastique dans une bande sur l'interface entre les deux mélanges des solvants

des densités différentes.

11) La bande vert foncé, bien définie, localisée directement au-dessous de la couche de couverture transparente, est aspirée dans une pipette pasteur recourbée et récupérée dans un tube de centrifugation de type Corex de 15 ml siliconé.

12) 2 volumes au moins de cyclohexane ("normapur", Prolabo) sont rajoutés et mélangés avec le concentrat.

13) Le matériel chloroplastique est culoté par une centrifugation pendant 10 minutes sous les mêmes conditions qu'avant (étape 10).

14) Le surnageant est décanté, le tube retourné et bien égoutté sur un chiffon. On peut essuyer les parois intérieures du tube. De préférence, le culot lui-même n'est pas complètement séché. Il se conserve pendant des semaines dans le tube fermé au réfrigérateur.

15) Sont rajoutés au culot:

- 0,5 ml d'une solution de Triton X 100 (Boehringer) de 20%,

- 2 ml de tampon de resuspension "RS" (Tris-HCl 80 mM, NaCl 100 mM, EDTA 30 mM, pH 8,0),

- une bille en acier, diamètre 12 mm,

tout cela à 0°C. Le tube est vortexé pendant quelques secondes à la vitesse minimale, jusqu'à la resuspension complète du culot. La bille s'enlève avec un aimant.

16) Le tube est fermé par un bouchon bien étanche et mis à tourner sur une roue à 0°C (ou simplement placé dans la glace, où il faut le remuer de temps en temps) pendant 2 ou 3 heures.

17) Le lysat est centrifugé à 3.300 g ou plus pendant 10 minutes ou plus, afin de culoter le matériel nucléosomique non-dissout.

18) Le surnageant est versé dans un tube à hémolyse type Pyrex de 5 ml, contenant déjà 30 µl d'une solution de Protéinase K (Boehringer) à 10 mg/ml. On y ajoute encore 170 µl d'une solution de SDS à 20%, puis on mélange et on incube à 65°C pendant une nuit (minimum 2 heures).

(L'exécution des étapes 19) à 24) à température basse est facultative:)

19) 1 volume d'un mélange 1:1 (p/p) de chloroforme et phénol (Merck), saturé en Tris 1 M, pH 8, est rajouté, le tube bouché est retourné 100 fois.

20) Centrifugation à 2.000 g pendant 25 minutes.

21) La phase aqueuse (c'est la supérieure) est prélevée avec une pipette pasteur et transvasée dans un nouveau tube, contenant 180 µl de Na Acétate 3 M, pH 6,2.

22) Répétition des étapes 19) et 20) (extraction au chloroforme / phénol).

23) On récupère à nouveau la phase aqueuse. Pour le troisième tour des extractions organiques 1 volume de chloroforme pur est rajouté. On ferme le tube et on le retourne de nouveau maintes fois.

24) Centrifugation à 2.000 g pendant 15 minutes.

25) Par aliquots d'environ 0,5 ml, la phase aqueuse est répartie dans 4 tubes Eppendorff de 1,5 ml. Dans chacun, 1 ml d'éthanol absolu est ajouté. Les tubes sont fermés, retournés à plusieurs reprises et exposés à -20°C pendant une nuit (3 heures minimum), afin de permettre à l'ADN de se précipiter complètement.

26) Les tubes sont centrifugés pendant 30 minutes à 11.000 g dans une chambre froide.

27) L'éthanol est enlevé sans reste à l'aide d'une pipette pasteur effilée et remplacé par 1ml d'éthanol à 70%.

28) Centrifugation pendant 15 minutes.

29) Après l'enlèvement de l'éthanol, les acides nucléiques sont séchés à l'air libre.

30) Si les culots séchés sont bien visibles, on permet à chacun de se redissoudre dans 25 µl de tampon TE (Tris 10 mM, EDTA 1mM); si les culots sont plus petits, on réduit le volume à rajouter en fonction. Selon le rapport de l'ADN aux ARNs dans l'échantillon, on arrive à une concentration finale d'ADNcp d'environ 0,07 à 0,2 µg/µl.

31) La titration est effectuée visuellement à partir d'une électrophorèse en mini-gel de 2,5 µl de la solution d'ADNcp digérée par Sal I, en présence d'un standard: 0,48 µg d'ADN du phage lambda donne la même intensité des bandes que 1,3 µg d'ADNcp.

### c) Remarques sur le protocole d'extraction d'ADNcp

Pas 1): Pour le riz aucune nécessité ne fut observée d'enfermer les plantes à l'obscurité (afin de faire digérer l'amidon des chloroplastes) pendant plus d'une nuit; par conséquent les feuilles peuvent être récoltées directement dans la serre, de préférence très tôt le matin.

En principe, au lieu de les congeler, on pourrait aussi sécher les feuilles à l'air chaude dans une étuve. Mais le succès dans un nombre limité d'essais, aux températures 40°C et 60°C, fut irrégulier: en général l'ADN avait subi une dégradation et les rendements étaient plus faibles que chez les témoins\*. Quelques résultats valables furent obtenus en broyant les feuilles séchées dans l'azote liquide (comme dans le pas 2), mais en supprimant la bille de 20 mm) et en poursuivant ensuite le protocole normal.

L'évaluation fiable de la possibilité d'un broyage directement à sec des feuilles séchées à la chaleur, supprimant toute utilisation d'azote liquide et de la lyophilisation, mériterait encore des essais soigneux; un seul de nos six essais produisit un résultat prometteur (c'est l'ADN de cet échantillon qui figure sur la piste 13 de la fig. IV A dans l'annexe C).

Pas 2) Pour les quantités de feuilles faibles (jusqu'à 4 grammes) le broyage au Dangoumo reste valable, mais il est préférable de réduire le poids des billes utilisées (par exemple utiliser 12 de 8 mm seulement) plutôt que le temps de vibration. Pour les quantités encore plus faibles (éprouvé jusqu'à 1 gramme), l'utilisation d'un mortier et un pilon est à recommander.

Pas 5) D'après la fiche toxicologique n° 151 de l'I.N.R.S., Paris, les principaux risques pathologiques et toxicologiques des trichlorobenzènes seraient des irritations non-allergiques de la peau, des muqueuses oculaires et de celles des voies respiratoires. Leur pénétration dans l'organisme se ferait par inhalation des vapeurs, l'absorption percutané par contre serait très lente.

Le décahydronaphthalène, un carbure d'hydrogène bicyclique saturé, n'est porteur d'aucune classification de danger par ses distributeurs.

Etant donné la très faible volatilité des deux solvants utilisés, la protection pen-

---

\* Dans les annexes A et B, les échantillons pour lesquels l'extraction après séchage au chaud a réussi sont marqués par l'indice de la température appliquée.

dant les extractions par une hotte aspirante, une blouse et des gants jetables en latex, changés après chaque contamination accidentelle, paraît suffisante. (N.B.: les solvants de remplacement coûtent moins cher que les originaux.)

Pas 6) A cause de la forte volatilité et de la faible viscosité du tetrachlorure de carbone et du n-hexane, s'ajoutant à l'inflammabilité de ce dernier, un tel broyage mécanique n'était pas possible dans la version originale de la technique.

Pas 9) Le dépôt de la couche propre, servant comme filtre liquide pour les chloroplastes lors de la centrifugation, n'est qu'une mesure de précaution et n'est pas obligatoire. En effet, on peut ici économiser du solvant et se servir de tubes de centrifugation d'une plus petite section, ce qui peut faciliter ensuite le prélèvement des chloroplastes.

Pas 12) Le cyclohexane est un solvant volatil nocif (cf. la fiche toxicologique de l'I.N.R.S. n<sup>o</sup> 17) mais recommandé comme produit de remplacement pour le n-hexane. Il est à manipuler sous la hotte aspirante, les tubes doivent rester bouchés lors de la centrifugation.

Pas 16) C'est l'utilisation d'un détergent non-ionique comme le Triton qui permet la libération préférentielle de l'ADN cytoplasmique par rapport à l'ADNn, car, à basse température, son activité lysante sur les nucléosomes, des complexes d'ADN avec des histones (des protéines de nature basique), est moins puissante que celle sur les complexes des fragments de membranes de chloroplastes avec l'ADNcp; celui-ci n'est pas autant emballé dans des protéines (HERRMANN, POSSINGHAM 1980).

Si LEHVÄSLAIHO (1986) utilise le Triton dans sa technique de préparation d'ADNcp, c'est également pour casser les membranes des organites, mais sans se servir de son activité différentielle, qui dépend d'ailleurs crucialement de la baisse de la température \*.

Pas 18) Pour le riz, les remplacements du Sarcosyl par le SDS et de la Pronase par la Protéinase K, associés à l'augmentation de la température et de la durée d'incubation, de 37°C à 60°C et de 1 heure à plusieurs heures, respectivement, ont améliorés la digérabilité de l'ADN isolé par les enzymes de restriction.

Les étapes 19) et suivants correspondent à un procédé classique de purification d'ADN, documenté par MANIATIS, FRITSCH et SAMBROCK (1982).

Pas 30) Les culots et les solutions finales présentent parfois un aspect brun,

---

\* résultat d'une série d'essais

orange ou jaune, témoignant la contamination des préparations, soit par des pigments (chlorophylle, carotinoïdes ou xanthophylles) ou leurs dérivés, soit par de la plastochinone ou des hydroquinones (cf. JACOBI 1974). Il faut soupçonner ces dernières d'un considérable effet inhibiteur sur les enzymes de restriction.

#### d) Complément au protocole d'extraction: le traitement des déchets

Tous les déchets organiques liquides sont réunis dans une bouteille conservée sous la hotte, et de là périodiquement emmenés pour traitement par une entreprise spécialisée. Les déchets solides (résidus de filtration, chiffons trempés, gants souillés) sont scellés dans un sac plastique et jetés. Les filtres nylon sont essuyés, laissés sécher et lavés à plusieurs reprises sous la hotte dans une lessive chaude (decon 90), qui après est versée dans l'égout de la hotte. Puis ils sont séchés et réutilisés.

#### e) Une autre possibilité d'extraction non-aqueuse

Le culot du matériel chloroplastique, obtenu par l'une ou l'autre version de la méthode non-aqueuse (p.e. après le pas 14) ci-dessus), même s'il est coriace (par exemple à cause d'un séchage complet au vide), se disperse facilement par agitation légère dans un petit volume de 1,1,1-trichloroéthane. Cette suspension peut être extraite avec du tampon RS, contenant 3 ou 4 % de Triton X 100, en émulsionnant l'ensemble pendant 3 heures à 0°C sur une roue tournante. L'agressivité modérée du solvant libère l'ADN des organites, ce qui permet son absorption dans la phase aqueuse. La durée de l'extraction est cependant critique, car sa réduction entraîne des pertes de rendement, son extension entraîne par contre une contamination nucléaire.

### 3) La digestion de l'ADNcp par les enzymes de restriction

Toutes les digestions furent exécutées selon les recommandations des fournisseurs des enzymes (Appligène, Boehringer). Notamment, leurs précisions sur les tampons d'incubation spécifiques aux enzymes furent respectés.

La quantité d'ADN utilisée fut fixée en fonction de l'enzyme: environ 2 à 2,5 µg pour les enzymes coupants fréquemment (Eco R I, Ava I), 1 à 1,5 µg pour les enzymes générant moins de fragments (Sal I, Pst I, Bam H I voir, Bst I, Hind III, Sma I, Bst E II).

Les temps d'incubation s'élevaient à 60 minutes pour les enzymes Bst I et Bst E II, qui fonctionnent à chaud, à 15 heures pour l'enzyme Sma I, qui digère à température ambiante, et à 3 heures pour les autres, qui demandent l'incubation à 37°C.

Par essai, 10 unités d'enzyme de restriction (4 unités pour Ava I) étaient mises en oeuvre, complémentées par 2µg de RNase (Merck, 20 U/mg, bouillie pendant 10 min lors de sa préparation).

Voici la composition d'un essai de digestion de restriction typique:

solution ADN + tampon TE	:	17,0 µl
tampon d'incubation 10x	:	2,1 µl
enzyme de restriction (10 U/µl )	:	1,0 µl
solution de RNase (2 mg/ml)	:	0,9 µl

- pour une série de 20 ou 40 digestions, les trois composants à rajouter à l'ADN (le tampon et les deux enzymes concentrés) étaient mélangés sur la glace, distribués par aliquots sur les échantillons d'ADN refroidis et remélangés avec ceux-ci par tapotement des tubes. Les mélanges étaient rassemblés par un coup de centrifugeuse et ensuite incubés dans un bain-marie.

Les réactions étaient terminées en remplaçant les tubes dans la glace. Chaque échantillon était mêlé avec 2,3 µl d'un tampon de charge (10 mM EDTA, 0,25 % bleu de bromophénol, 90 % glycérol), chauffé à 65°C pendant 5 minutes afin de bien isoler

les fragments d'ADN, et de nouveau refroidi.

#### 4) L'électrophorèse de l'ADNcp

La séparation des fragments de restriction fut effectuée par électrophorèse "sous-marine" dans des gels d'agarose ("ultrapure" Biorad) d'une épaisseur de 4 mm (200 ml). Comme tampons, furent utilisés le TAE (Tris, acétate, EDTA) et le TBE (Tris, borate, EDTA; réutilisé deux fois) selon MANIATIS (1982), et ceci en fonction des enzymes de restriction utilisées, comme précisé dans le tableau 1.

tab. 1: conditions d'électrophorèse

enzyme	concentration d'agarose	tampon	voltage	durée approximative
<u>Eco R I</u>	1,30 %	TAE	50 V	20 h
<u>Ava I</u>	1,75 %	TBE	110 V	7 h
<u>Bam H I</u>	0,85 %	TAE	70 V	10 h
<u>Sal I</u>	0,70 %	TAE	40 V	7 h *
autres	0,75 %	T(A + B)E **	45 V	20 h

\* conditions de mini-essai: gel de 24 cm avec 111 puits en trois lignes, taille du puits 3 mm x 1mm, volume final de l'échantillon 7 µl;

\*\* T(A + B)E : 33% TAE + 67% TBE (réutilisé deux fois);

Pour assurer une bonne définition des bandes, toutes les électrophorèses sont initiées par l'application d'un voltage de 25 V pendant 45 minutes. Pendant l'électrophorèse le tampon est recirculé du bac de l'anode vers le bac de la cathode par une pompe péristaltique (débit 5 ml/min).

Dans chaque électrophorèse est inclu, soit un marqueur de poids moléculaires ("Raoul I", Appligène, mêlé avec le tampon d'incubation de l'enzyme d'analyse et un aliquot de RNase), soit un marqueur interne, c'est à dire donnant un patron déjà connu.

On trouvera dans l'annexe E le protocole pour une digestion de restriction à petite échelle (100 µg d'ADN), le marquage radioactif au  $^{35}\text{S}$ -dCTP en 5' des fragments, et l'électrophorèse en gel polyacrylamide à gradient de concentration, tel qu'appliqué pour les enzymes Msp I et Hpa II (se distinguant par un déterminisme de reconnaissance de sites à 4 bases), générateurs de plus de 200 fragments chez l'ADNcp du riz.

### 5) La coloration et la photographie des gels d'électrophorèse

Les gels sont colorés directement dans leurs moules rescellés aux deux bouts avec du ruban adhésif, en leur superposant un volume égal (200 ml) d'une solution de bromure d'éthidium à 1µg/ml pour 30 minutes. Le fond est ensuite décoloré dans un bain d'eau déminéralisée pendant 20 minutes.

Transilluminés par des UV à 365 nm, les gels sont photographiés sur négatifs de 8.5 cm x 10.8 cm (Polaroid 665, ISO 80/20°) à travers (du bas en haut)

- un filtre ultraviolet,
- une gélatine filtre jaune faible,
- un filtre rouge foncé (Cokin),

par une exposition de 18 minutes avec le diaphragme ouvert.

## 6) La lecture des gels d'électrophorèse

Chaque bande observée est caractérisée par sa distance de migration absolue, au quart de millimètre près, par son bord antérieur. L'agrandissement des photos à la taille originale des gels permet la précision de la lecture.

L'estimation des poids moléculaires correspondant aux différentes bandes est effectuée à l'aide d'une courbe de calibrage presque-linéarisée, déduite du patron du marqueur. Voici comment:

a) parmi les bandes du marqueur moléculaire, trois sont choisies pour la détermination d'une hyperbole (suivant une proposition de RICKWOOD et HAMES, 1982) du type

$$\underline{h}: L = \frac{a}{c + M} + b ,$$

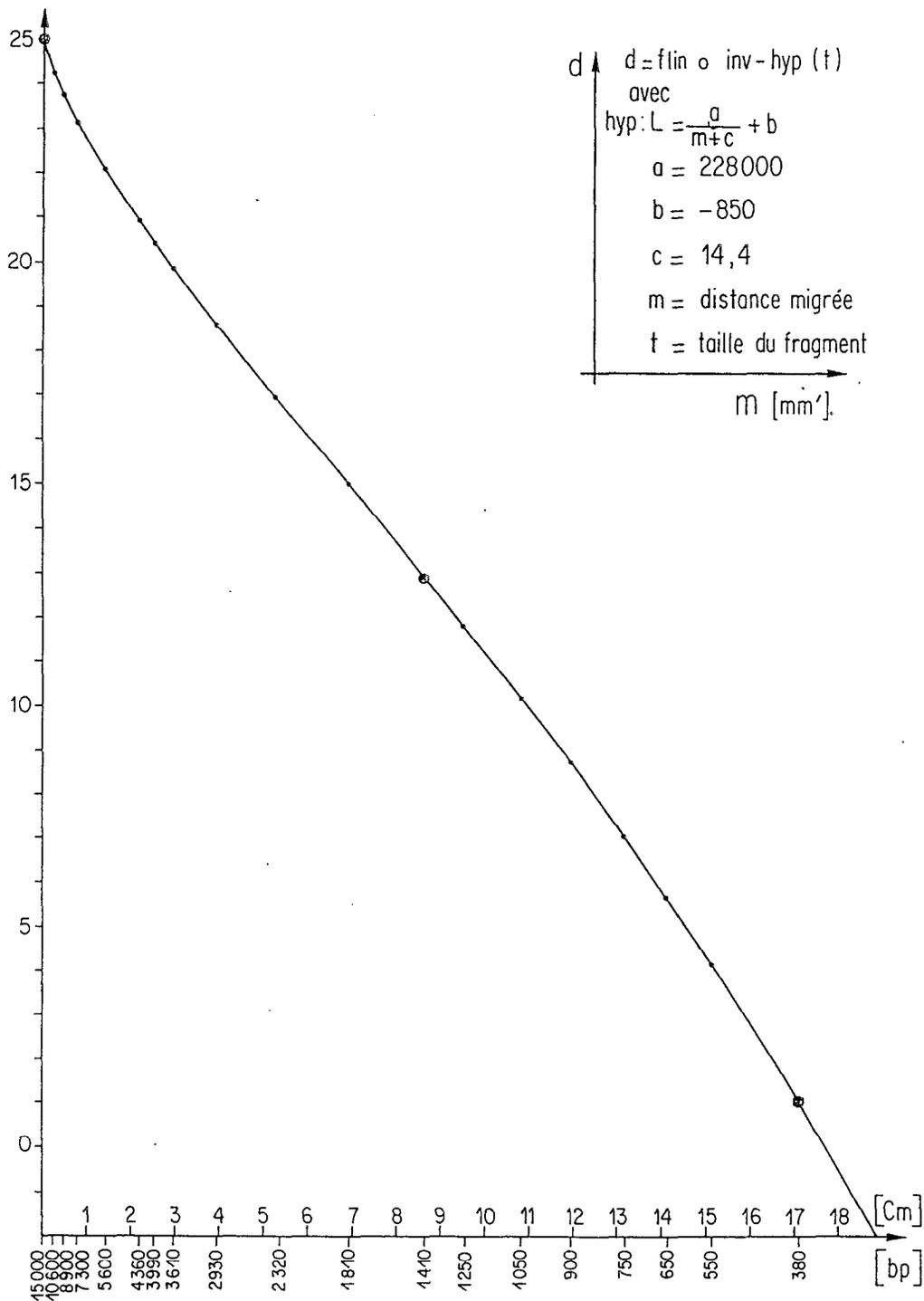
où M symbolise la distance migrée et L la longueur en paires de bases (ou le poids moléculaire) d'un fragment d'ADN. Le critère de choix pour ces trois premières bandes de référence est qu'ils encadrent au mieux les patrons de l'ADNcp: une à chaque extrémité et une au milieu.

b) Les distances de migration de chaque bande  $m_j$  du marqueur sont reportés sur l'abscisse, et, sur l'ordonnée, les valeurs de référence  $\underline{d}$ , obtenues par l'application sur les logueurs des éléments du marqueur de l'inverse  $\text{inv}\underline{h}$  de la fonction hyperbolique susdite  $\underline{h}$ .

La suite des points ainsi construits décrit une courbe simple, légèrement sigmoïde: la courbe de calibrage. Comme exemple, la courbe établie pour l'évaluation des données Ava I, adapté à la taille d'une feuille millimétrée par une transformation linéaire, est montrée dans la figure 1.

Pour le calcul de la longueur d'un fragment correspondant à une bande quelconque, on monte de la distance migrée mesurée pour cette bande comme valeur d'abscisse sur la courbe (dans le cas idéal la droite liant les deux points de référence voisins, correspondant aux deux bandes du marqueur les plus proches de celle en examen) et lit la valeur  $\underline{d}$ . A partir de celle-ci la longueur nette est trouvée par l'application de la fonction  $\underline{h}$ . (Le cas de la figure 1 demande en plus l'application préala-

fig 1: courbe de calibrage des tailles des fragments Ava I  
 (⊗ désigne les points de définition de l'hyperbole h)



ble de l'inverse de la transformation linéaire impliquée.)

## 7) L'analyse des mutations

### a) Généralités

Afin d'expliquer en termes de mutations les différences entre deux patrons de restriction, produits sur des ADN homologues par une même endonucléase, on ne considère pas seulement qualitativement les présences et absences des bandes dans les patrons. En se servant des valeurs quantifiées des tailles des fragments correspondant aux différentes bandes, on recherche les relations possibles entre les éléments non-identiques des deux patrons.

A ce sujet, la notion de mutation se rapporte à celle définie par la génétique moléculaire, et par conséquent elle distingue les mêmes types que celle-ci. Aux substitutions de bases ou mutations ponctuelles correspondent les mutations de sites de restriction simples, qui se traduisent par les apparitions et disparitions de sites de coupure de l'endonucléase sur la molécule. On parle d'une addition ou insertion, si un bout d'ADN est inséré dans la molécule, d'une délétion, si un bout en est enlevé. Le terme de transposition désigne le déplacement d'une partie de la molécule par rapport au reste, celui d'inversion le changement de l'orientation relative d'un tel fragment. A noter qu'une mutation de site est (en général) spécifique de l'enzyme de restriction, alors que les mutations des autres types peuvent apparaître dans les patrons de digestion de plusieurs enzymes à la fois.

La possibilité d'identification directe n'est cependant pas parfaite. Puisque les sites de restriction sur l'ADN ne sont pas vraiment de simples points, mais de courtes séquences (en général 4 ou 6 paires de bases en ordre palindromique), il peut théoriquement arriver que les additions, délétions, transpositions et inversions à leurs extrémités créent ou détruisent des sites de coupure. Sur les patrons de fragments ceci pourrait produire des effets particuliers.

## b) La détection des différents types de mutations

### - Les mutations de sites

sont reconnues par le remplacement de deux petits fragments par un plus grand (disparition du site) ou d'un grand fragment par deux plus petits (apparition du site). En tout cas la somme des tailles des deux petits fragments est égale à la taille du grand. En réalité les deux petits fragments ne sont pas toujours détectables: c'est qu'au moins l'un d'eux est trop petit pour être détecté sur l'électrophorégramme.

### - Les additions ou insertions

en général donnent lieu au remplacement d'un fragment par

- un fragment plus grand, dans le cas normal,
- deux fragments dont la somme est plus grande, si le segment inséré dans la molécule porte un site de restriction.

Dans notre cas, pour la plupart, les différences de taille ne dépassent pas quelques dizaines de paires de bases.

La seule possibilité de distinguer à coup sûr une addition d'une perte d'un site est de la retrouver systématiquement avec différentes enzymes, ce qui est pratiquement impossible pour les plus petites d'entre elles.

### - Les délétions

se détectent comme les additions, la distinction entre les deux demandant de toute façon une référence extérieure à la paire de patrons en comparaison.

### - Les inversions,

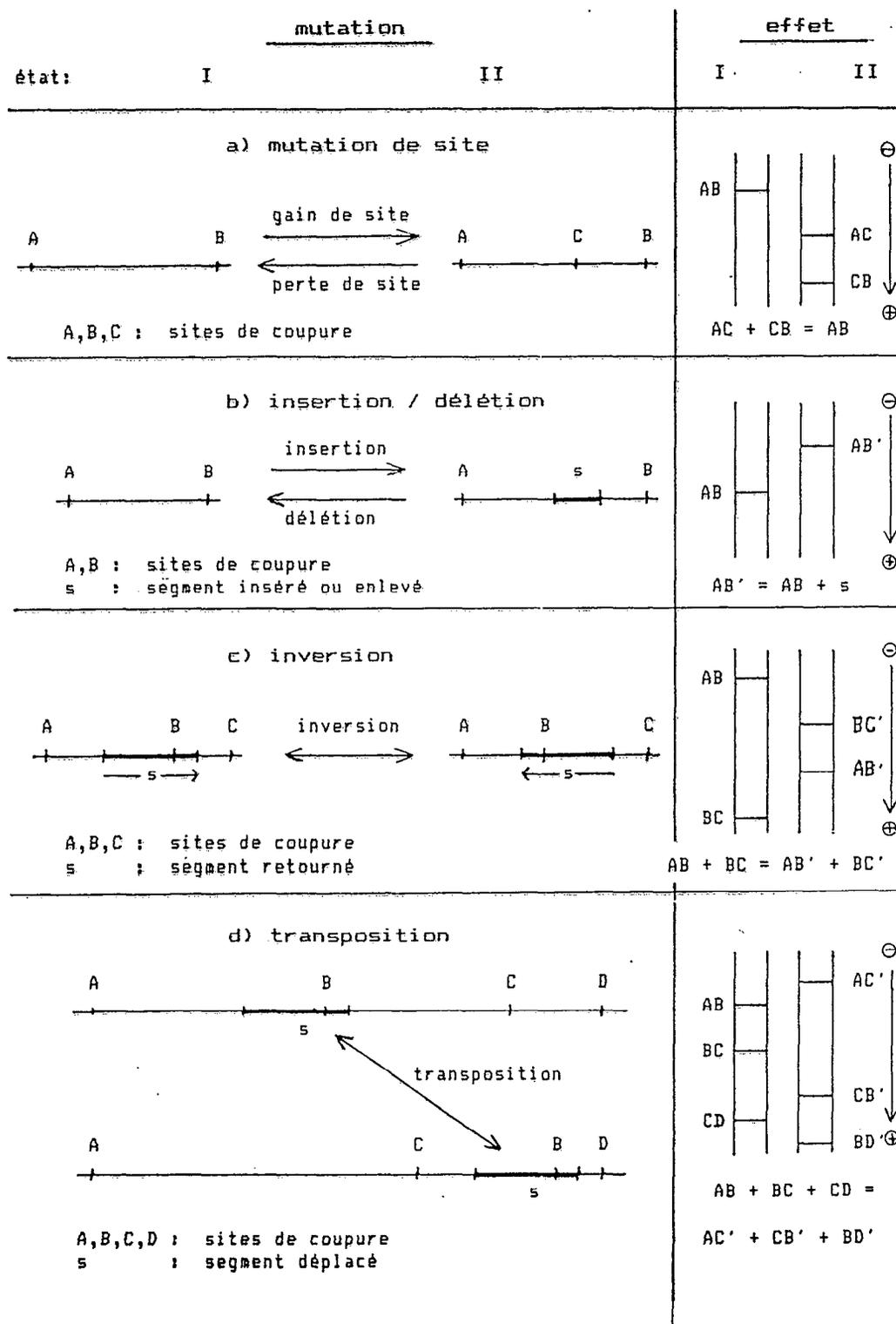
dans leur version typique, entraînent la recombinaison des deux fragments couvrants les bouts du segment inversé. Deux fragments sont ainsi remplacés par deux autres, leurs deux sommes respectives étant identiques.

### - Les transpositions

ne peuvent également être reconnues que si elles transportent au moins un site de restriction. Dans ce cas elles font typiquement disparaître trois fragments et apparaître trois autres à leur place, la somme des fragments nouveaux étant égale à celle des fragments anciens.

Dans le cas où le segment entre le point de délétion et le point d'insertion de la

fig 2: effets des mutations sur les patrons de restriction



séquence déplacée ne contient pas de site de restriction, les fragments remplacés et les fragments remplaçants sont au nombre de deux, et ainsi la distinction d'une inversion ne pourrait se faire qu'à l'aide de résultats obtenus avec plusieurs enzymes.

La figure 2 illustre les conséquences de quelques mutations typiques sur des patrons de restriction.

Comme les mutations de sites sont un type de mutation fréquent et bien caractérisable, elles furent les premières à être recherchées d'une façon systématique. Pour chaque analyse d'une paire de patrons, toutes les combinaisons simples (type  $a + b = c$ ) possibles furent vérifiées. Des additions ou délétions supérieures à 50 paires de bases ne furent supposées que quand ils étaient retrouvés indépendamment avec plusieurs enzymes.

#### **8) La construction d'un cladogramme sur la base des relations définies par mutations entre les différents types moléculaires**

Le cladogramme est construit comme un arbre hypothétique en attribuant un bout de branche à chaque type cytoplasmique individualisé, en minimisant la somme totale des mutations spécifiques aux différents branchements, en n'acceptant aucune coïncidence ou réversion parmi les mutations de site, transitions, inversions et grandes insertions ou délétions, mais en acceptant le nombre minimal nécessaire de coïncidences ou réversions sur les petites additions et délétions (à comparer avec la méthode de DeBRY et SLADE, 1985, développée pour l'analyse des cartes de restriction).

## 9) La construction des arbres phylogéniques hypothétiques

Les différents arbres phylogénétiques hypothétiques représentent des transformations du cladogramme primitif, maintenant sa topologie intérieure, mais tirant la totalité des extrémités des branches extérieures sur un même niveau (symbolisant le présent) et transformant l'ordre purement hiérarchique des branchements chez le premier en un ordre, à la fois hiérarchique et temporel relatif, chez le second. Deux méthodes furent utilisés pour calculer les niveaux temporels relatifs des bifurcations:

### **a) une méthode de repérage par paires non-pondérées (UPGMA; SOKAL, SNEATH 1963):**

Le niveau temporel relatif attribué à une bifurcation est proportionnel au nombre moyen des mutations nécessaires pour transformer un type porté par le bras gauche en un type du bras droit du branchement. Pour les "polyfurcations", c'est à dire s'il y a plusieurs branches, le niveau du noeud est proportionnel au nombre moyen des mutations nécessaires pour la transition d'un type porté par l'un des bras à un autre porté par un autre des bras.

### **b) une méthode de repérage par paires pondérés (WPGMA):**

La seule différence entre cette méthode et la précédente est que le traitement des deux branches d'une bifurcation, lors de la détermination du niveau relatif de cette dernière, n'est pas automatiquement égal, mais varie en fonction du nombre de bouts terminaux auxquels chaque branche descend: le poids relatif attribué à la branche étant égal à son nombre de terminaisons. Cette méthode à comme effet de réduire dans les estimations l'influence des branches minoritaires montrant des taux de mutations extrêmes.

Si, pour une bifurcation donnée, le nombre de terminaisons est égal pour ses deux branches, les deux méthodes a) et b) lui attribuent la même valeur de niveau temporel relatif.

## LES RESULTATS

### 1) Résultats concernant les techniques et méthodes

a) **Amélioration de la technique non-aqueuse de purification des ADNcp appliquée au riz.**

Les modifications apportées par notre protocole d'extraction à la méthode originale représentent des améliorations à plus d'un titre. Les voici en résumé:

- La mécanisation des deux broyages, dont la deuxième doit sa possibilité au remplacement des solvants mis en oeuvre, ainsi que la réduction des volumes à partir de la resuspension du culot chloroplastique (celle-ci s'avérant facilitée grâce à la faible volatilité des nouveaux solvants) signifient une rationalisation des extractions en termes de temps nécessaire par échantillon.

- La mécanisation est également essentielle pour l'efficacité des extractions à partir de plantes adultes.

- L'étape nouvellement introduite de solubilisation différentielle (pas 15 à 17 dans le protocole), postérieur à l'enrichissement du matériel chloroplastique, a augmenté la qualité des purifications en réduisant à un minimum la contamination de l'ADNcp par de l'ADNn (la figure 1 dans l'annexe C montre un résultat typique obtenu avec la méthode originale sur le riz - à comparer avec le reste des figures, toutes montrant des résultats obtenus avec la nouvelle méthode).

- Par conséquent, une contrainte est enlevée de l'étape d'isolation des chloroplastes, car la nécessité d'un compromis entre le rapport en chloroplastes et le degré de contamination par les ADNn se voit suspendue: l'augmentation des rendements par l'augmentation de la densité du médium de centrifugation ne s'accompagne plus par un risque élevé de contamination. Dans nos expériences sur le riz, un essai sur 12 g de feuilles adultes fraîches apporte maintenant entre 4 et 50 µg d'ADNcp, selon l'espèce et l'état de la plante, - en moyenne près de 10 µg.

- Tous les solvants de remplacement réunissent l'avantage de coûter moins cher que les produits remplacés et d'offrir une sécurité de manipulation améliorée.

#### **b) La non-spécificité au riz de la technique de préparation de l'ADNcp**

La fiabilité de la nouvelle méthode non-aqueuse fut testée dans un essai direct et unique d'application sans modifications sur 12 espèces de plantes supérieures. La qualité des ADNs isolés fut vérifiée par une digestion Bam H I, dont les résultats sont montrés dans la figure II dans l'annexe C et décrits dans le tableau 2.

tab.2: résultats d'une préparation des ADNcp de 12 espèces de plantes à fleurs avec la méthode adaptée au riz

espèce*	rendement**	cont:nucl.	état***	digérabilité
<u>Arum maculatum</u>	10 µg	--	+++	+
<u>Asphodelus fistulosus</u>	8 µg	--	+++	+
<u>Avena sterilis</u>	7 µg	(--)	+	+
<u>Centaurea collina</u>	3 µg	(--)	+	+
<u>Gossypium australe</u>	11 µg	--	+++	+
<u>Eichhornia crassipes</u>	12 µg	(--)	+	+
<u>Hibiscus sp.</u>	10 µg	--	+++	+
<u>Iris sp.</u>	5 µg	?	+++	--
<u>Plantago lanceolatum</u>	2 µg	?	--	+
<u>Quercus ilex</u>	8 µg	?	+	--
<u>Scilla Lilio-Hyazinthus</u>	6 µg	--	+++	+
<u>Thymus vulgaris</u>	1 µg	?	+	--

Conclusion: Les modifications de la technique ne doivent pas être considérées comme des adaptations spécifiques au riz.

\* récemment, dans le même laboratoire, S.H. ZHANG a obtenu des résultats positifs sur Triticum aestivum, Zea mays, Avena sativa et Dactylis glomerata, N. HOCHLANDER sur Plantago lanceolatum;

\*\* rendement d'un essai simple sur 12 g de feuilles fraîches;

\*\*\* ++ sans dégradation; + faiblement dégradé; -- dégradé;

### c) La régularité des électrophorèses

Pour une électrophorèse donnée, c'est la régularité de la courbe de calibrage, déduite d'un marqueur fin de poids moléculaire étalé lors de cette électrophorèse, qui peut être regardé comme indicatrice de la régularité de celle-ci.

Comme le montre l'exemple de la figure 1 cette régularité est remarquable, dans la mesure où la partie quasi-linéaire de la courbe est large et que les sécantes définies par deux points voisins de définition de la courbe restent toujours proches de celle-ci (il n'y a pas de zig-zag).

Si le marqueur, qui consiste en un ADN particulièrement pure produit par clonage, migre d'une façon bien régulière, ceci ne permet pas cependant de conclure avec certitude que ce soit aussi vrai pour l'ADNcp.

Au contraire on a à signaler à ce sujet, que pour les ADNs cp du riz les patrons Eco R I, mais eux seulement, se présentent différemment, par rapport à un même marqueur externe, après électrophorèse dans les deux tampons différents, TAE et TBE, en gardant les autres paramètres invariants (comp. fig. 3 et les fig. III A/B et III C dans l'annexe C). Le fait que ces différences se limitent aux placements de quelques bandes de taille moyenne fait penser qu'il pourrait s'agir d'effets de structure secondaire chez certains fragments, dépendant dans leur formation de la force ionique du milieu. Et finalement ce sont les cohérences lors de l'analyse des mutations de sites qui suggèrent que la régularité de l'électrophorèse est plutôt propre au système utilisant le TAE, donc que c'est le TBE qui semble responsable de l'artefact décrit.

### d) La qualité des préparations d'ADNcp

Quant à la pureté chimique des ADNs isolés, il peut être signalé qu'elle se montre suffisante pour assurer le fonctionnement des enzymes de restriction et pour permettre la nick-translation ainsi que le marquage en 5' des bouts de fragments de restriction par le fragment de Klenow de la DNA-polymerase I. L'enzyme Kpn I fait cependant exception, en donnant souvent des digestions incomplètes, ce qui pourrait être dû à son inhibition partielle par des substances contaminantes.

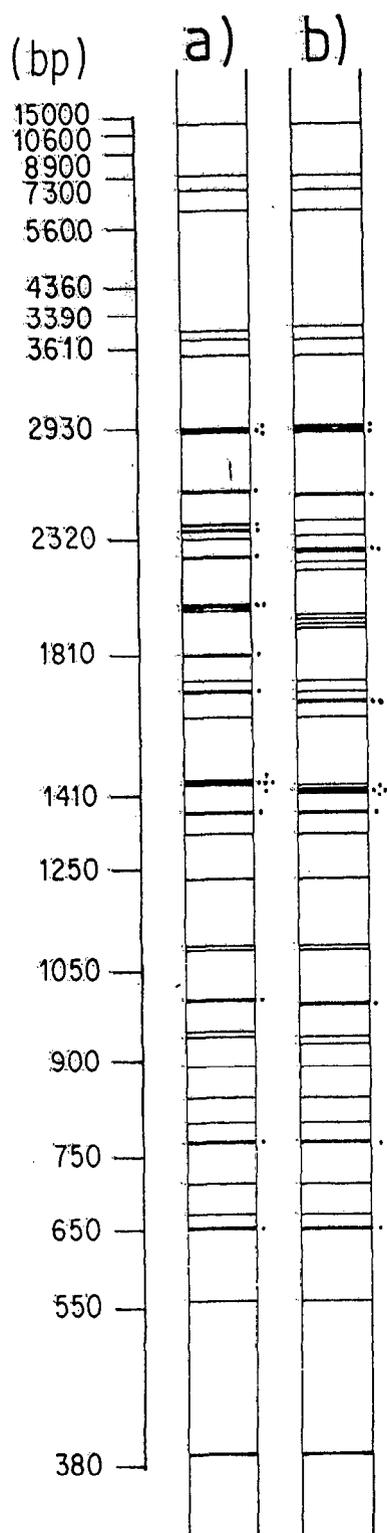


fig. 3:  
 le patron Eco R I de l'ADN cp  
 d'Oryza sativa japonica  
 par électrophorèse  
 a) en TAE , b) en TBE

En ce qui concerne les contaminations par de l'ADN nucléaire, on peut constater qu'elle est négligeable, dans la mesure où, en général, la possibilité de lecture complète des patrons spécifiques à l'ADNcp ne se voit pas affectée. Dans des cas rares il est possible de reconnaître dans des patrons Eco R I quelques bandes contaminantes faibles, qui selon leur poids moléculaire correspondent à des fragments d'ADN ribosomique nucléaire.

Les électrophorèses d'échantillons non-digérés montrent qu'en moyenne les molécules sont d'un haut poids moléculaire. Néanmoins une certaine dégradation se révèle lors de la restriction avec des enzymes qui donnent peu de fragments: les bandes correspondant aux plus grands fragments n'apparaissent pas dans l'intensité que l'on devrait attendre d'après celles des plus petits fragments. Mais cela augmente plutôt la lisibilité des gels qu'il ne la gêne; sauf peut-être dans le cas de l'enzyme Bst E II, où un fragment d'environ 60 kb, qu'elle semble produire, ne forme pas de bande nette dans le gel, mais une sorte de voile diffus (comp. fig. V dans l'annexe C).

Les contaminations par l'ADN mitochondrial sont également faibles. En molarité elles ne dépassent probablement pas l'ordre d'un pour cent, ce qui découle d'un essai d'hybridation selon SOUTHERN (1975) de l'ADNcp digéré par Ava I avec une sonde spécifique au gène de la cytochrome oxidase. Elle révèle une bande dimorphe, qui n'avait pas été visualisée par la coloration au bromure d'éthidium du gel original (comp. fig IV D dans l'annexe C; l'essai d'hybridation avec la sonde froide fut effectué par S.H. ZHANG).

## 2) Les patrons de restriction des ADNs chloroplastiques du riz

Des photos représentatives des résultats d'électrophorèse figurent dans l'annexe C, et dans l'annexe D les reconstitutions dessinées de la totalité des patrons analysés.

Pour chaque bande observée, une valeur de longueur en kilobases est déterminée, correspondant à celle des fragments de l'ADN qui constituent cette bande. Ainsi, chaque bande d'électrophorèse est identifiée avec une classe de taille des frag-

ments de restriction.

Le présent chapitre donne sous forme de tableaux les distributions des fragments sur les différentes bandes. Les valeurs de taille attribuées aux bandes sont déterminées à l'aide d'un marqueur spécial étalé lors de l'électrophorèse, comme décrit ci-dessus. Elles représentent cependant d'abord des étiquettes pour les ensembles de fragments réels réunis dans ces bandes, et ensuite des estimations approximatives des tailles réelles de ces fragments\*.

Les chiffres figurant dans les colonnes représentent alors le nombre de fragments présents dans une bande donnée (2 pour une bande double, 3 pour une bande triple etc.). Les bandes sont listées par ordre décroissant de leurs valeurs de taille respectives.

La liste des plantes de l'annexe A indique déjà le classement des échantillons sous les 32 types chloroplastiques distingués. L'annexe B donne la répartition des plantes analysées sous ces différents types.

#### **a) Les patrons Eco R I**

L'endonucléase Eco R I reconnaît et coupe la séquence GAA TTC . En moyenne elle génère 59 fragments de taille supérieure à 400 pb. Les patrons sont présentés dans le tableau 3.

#### **b) Les patrons Ava I**

Ava I reconnaît et coupe les deux séquences CCC GGG et CTC GAG . Elle produit autour de 62 fragments de taille supérieure ou égale à 280 pb. La totalité des patrons observés est porté sur le tableau 4.

---

\* une certaine réserve concerne surtout l'utilisation de la deuxième décimale, qui se propose surtout par la densité des bandes dans certains patrons de restriction

**c) Les patrons Bst I et Bam H I**

Sur l'ADNcp du riz, les deux enzymes Bst I et Bam H I se révèlent comme isoschizomères strictes. Elles créent 34 fragments supérieurs à 0,6 kb en restreignant la séquence GGA TCC . Le tableau 5 donne les patrons.

**d) Les patrons Hind III**

Séquence de reconnaissance: AAG CTT ; nombre de fragments générés: 35 (supérieurs à 0,6 kb). La liste des patrons constitue le tableau 6.

**e) Les patrons Sma I**

Séquence de reconnaissance: CCC GGG\* ; nombre de fragments générés: 20 (sup. à 0,6 kb). Liste des patrons: tableau 7.

**f) Les patrons Pst I**

Séquence de reconnaissance: CTG CAG ; nombre de fragments générés: 15 (sup. à 0,7 kb). Liste des patrons: tableau 8.

**g) Les patrons Bst E II**

Séquence de reconnaissance: GGT N ACC\*\* ; nombre de fragments générés: 14 (sup. à 0,7 kb). Liste des patrons: tableau 9.

---

\* coincidence partielle avec Ava I

\*\* N désigne un nucléotide quelconque









tab. 5: la distribution des fragments Bam H I sur les bandes électrophorétiques

bande no.	taille (kb)	types cytoplasmiques									
		<a-i>	e2	k	l/m/n	n3	o	p1	p2/3	q1	q2
1)	19.7	1	1	1	1	1	1	1	1	1	1
2)	12.8	1	1	1	1	1	1	1	1	1	1
3)	8.95	1	1	-	1	1	-	-	1	1	1
4)	8.5	-	-	1	-	-	-	-	-	-	-
5)	8.1	-	-	-	-	-	1	1	-	-	-
6)	7.9	1	1	1	1	1	1	1	1	-	-
7)	7.15	-	-	-	-	-	-	-	-	1	1
8)	6.94	1	1	1	-	-	1	1	1	1	1
9)	6.62	-	-	-	1	1	-	-	-	-	-
10)	5.30	1	1	-	1	1	1	1	1	1	1
11)	5.23	-	-	1	-	-	-	-	-	-	-
12)	5.17	-	-	-	1	-	1	1	1	1	1
13)	4.92	-	-	-	-	-	-	-	-	1	1
14)	4.86	1	1	1	1	1	1	1	1	-	-
15)	4.8-	-	1	-	-	-	-	-	-	-	-
16)	4.73	1	1	1	1	2	1	1	1	1	1
17)	4.64	1	-	-	-	-	-	-	-	-	-
18)	4.57	-	-	1	-	-	-	-	-	-	-
19)	4.10	4	4	4	4	4	4	4	4	4	4
20)	3.74	-	-	1	-	-	-	-	-	-	-
21)	3.36	1	1	1	1	1	1	1	1	2	2
22)	3.23	-	-	-	-	-	-	-	-	1	1
23)	3.12	2	2	3	3	3	3	3	3	2	2
24)	3.10	1	1	-	-	-	-	-	-	-	-
25)	2.84	-	-	1	1	1	1	1	1	1	1
26)	2.58	-	-	-	-	-	-	1	-	-	-
27)	2.55	1	1	1	-	-	1	-	1	1	1
28)	2.52	-	-	-	1	1	-	-	-	-	-
29)	2.47	1	1	1	1	1	1	1	1	1	1
30)	2.40	-	-	1	-	-	-	-	-	-	-
31)	2.00	2	2	2	2	2	2	2	2	1	1
32)	1.72	2	2	1	1	1	1	1	1	1	1
33)	1.61	1	1	1	1	1	1	1	1	1	1
34)	1.56	-	-	1	-	-	-	-	-	-	-
35)	1.52	1	1	1	1	1	1	1	1	1	1
36)	1.47	2	2	2	2	2	2	2	2	2	2
37)	1.42	-	-	-	1	1	1	1	1	-	1
38)	1.39	1	1	1	-	-	-	-	-	1	-
39)	1.35	-	-	-	-	-	-	-	-	1	-
40)	1.33	-	-	-	-	-	-	-	-	-	1
41)	1.12	3	3	2	2	2	2	2	2	2	2
42)	1.06	1	1	1	1	1	1	1	1	1	1
43)	0.84	-	-	-	-	-	1	1	-	-	-
44)	0.76	1	1	-	1	1	1	1	1	1	1
45)	0.70	1	1	2	1	1	1	1	1	1	1
	0.6	( limite de détection )									

La colonne <a-i> comprend tous les types de a1 à i2 sauf e2, et la colonne l/m/n tous les types de l à n2.

tab. 6: la distribution des fragments Hind III sur les bandes électrophorétiques

bande no.	taille (kb)	types cytoplasmiques							
		a/b/c	e+i	f	g	k	l/m/n	o	p3
1)	12	1	1	1	1	1	1	1	1
2)	9.0	1	1	1	1	1	1	1	1
3)	8.3	2	2	2	2	2	2	2	2
4)	7.75	1	1	1	1	-	1	1	-
5)	7.30	-	-	-	-	1	-	-	-
6)	7.25	1	1	1	1	1	1	1	1
7)	7.04	-	-	-	-	-	-	-	1
8)	6.94	1	1	1	1	1	1	1	-
9)	6.6	-	-	-	-	-	-	-	1
10)	6.60	1	1	1	1	1	1	1	1
11)	6.50	1	1	1	1	1	1	1	1
12)	5.42	1	1	1	1	1	1	1	1
13)	4.66	1	1	1	1	1	1	1	1
14)	3.57	1	1	1	1	1	1	1	1
15)	3.44	-	-	1	-	1	1	1	1
16)	3.41	-	1	-	1	-	-	-	-
17)	3.36	1	-	-	-	-	-	-	-
18)	3.07	1	1	1	1	1	1	1	1
19)	2.69	1	1	1	1	1	1	1	1
20)	2.67	1	1	1	-	1	1	1	1
21)	2.57	3	3	3	3	3	3	3	3
22)	2.52	2	2	2	2	2	2	2	2
23)	2.44	-	-	-	-	1	-	1	1
24)	2.42	1	1	1	1	-	-	-	-
25)	2.23	1	1	1	1	1	1	1	1
26)	2.14	-	-	-	-	-	1	-	-
27)	1.94	1	1	1	1	1	1	1	1
28)	1.46	-	-	-	1	-	-	-	-
29)	1.26	2	2	2	2	2	2	2	2
30)	1.24	1	1	1	1	1	1	1	1
31)	1.22	-	-	-	1	-	-	-	-
32)	1.17	2	2	2	2	2	2	2	3
33)	1.12	1	1	1	1	1	1	1	1
34)	1.03	1	1	1	1	1	1	1	1
35)	0.82	1	1	1	1	1	1	1	1
36)	0.73	2	2	2	2	2	2	2	2
37)	0.67	1	1	1	1	1	1	1	1
	0.6								

( limite de détection )

Les types d, e2 à e5, f2, h2, i2, m3, n3, p1, p2, q1 et q2 n'étaient pas inclus dans l'analyse.

tab. 7: la distribution des fragments Sma I sur les bandes électrophorétiques

bande no.	taille (kb)	types cytoplasmiques						
		<a-i>	a2/b/c	k	<l-p>	n3	p3	q
1)	19.2	1	1	-	1	1	1	1
2)	14.6	1	-	1	1	1	1	1
3)	12.3	1	1	2	1	1	1	1
4)	11.4	-	1	-	-	-	-	-
5)	9.6	-	-	1	1	-	1	1
6)	9.4	-	-	1	1	1	1	3
7)	9.3	4	4	2	2	2	2	-
8)	9.2	-	-	-	-	1	-	-
9)	8.8	-	-	-	1	1	1	-
10)	8.7	-	-	-	-	-	-	1
11)	8.4	-	-	1	-	-	-	-
12)	7.8	1	1	-	-	-	-	-
13)	7.5	-	-	1	1	1	1	1
14)	7.4	1	1	-	-	-	-	-
15)	6.91	-	-	1	-	-	-	-
16)	6.60	1	1	1	1	1	1	1
17)	4.86	-	-	-	-	-	1	-
18)	4.75	1	1	1	1	1	-	1
19)	4.08	3	3	3	3	3	3	3
20)	3.17	-	1	-	-	-	-	-
21)	1.76	2	2	2	2	2	2	2
22)	1.56	1	1	1	1	1	1	1
23)	1.10	2	2	2	2	2	2	2
24)	1.03	1	1	-	-	-	-	-
	0.6	( limite de détection )						

La colonne <a-i> comprend tous les types de a1 à i2 sauf a2, b et c.

tab. 8: la distribution des fragments Pst I sur les bandes électrophorétiques

bande no.	taille (kb)	types cytoplasmiques						
		a/b/c	<e-i>	k	l/m/n	n3	o/p	q
1)	20	1	1	1	1	1	1	1
2)	16.3	1	1	1	1	-	1	-
3)	16.1	1	1	-	-	-	-	-
4)	15.9	-	-	-	-	1	-	-
5)	15.0	1	1	-	1	1	1	2
6)	14.8	-	-	1	-	-	-	-
7)	14.6	1	1	1	1	1	1	1
8)	12.1	-	-	-	-	-	-	1
9)	10.7	1	1	1	1	1	1	1
10)	9.9	1	1	1	1	1	1	-
12)	8.3	2	2	2	2	2	2	2
12)	7.87	-	-	1	1	1	1	1
13)	7.73	1	1	-	-	-	-	-
14)	5.52	1	1	1	1	1	1	1
15)	5.10	-	-	-	-	-	-	1
16)	5.03	1	1	1	1	1	1	-
17)	4.66	1	1	1	-	-	1	1
18)	4.34	-	-	-	1	1	-	-
19)	3.87	-	1	1	1	1	1	1
20)	3.80	1	-	-	-	-	-	-
21)	2.13	1	1	1	1	1	1	-
22)	1.85	1	1	1	1	1	1	1
23)	1.26	-	-	-	-	-	-	1
	0.7	( limite de détection )						

Tous les types sauf f2 étaient inclus dans l'analyse. Donc a comprend a1 et a2, etc..

tab. 9: la distribution des fragments Bst E II sur les bandes électrophorétiques

bande no.	taille (kb)	types cytoplasmiques					
		a/b/c	<e-i>	k	l/m/n	o/p	q
1)	60 (?)	1	1	1	1	1	1
2)	14	1	1	1	1	1	1
3)	11.9	2	2	2	1	2	2
4)	11.6	-	-	-	1	-	-
5)	6.45	1	1	-	1	1	1
6)	6.15	-	-	1	-	-	-
7)	5.53	1	1	1	1	1	1
8)	4.25	-	-	-	-	-	1
9)	4.07	1	1	1	1	1	-
10)	3.59	1	1	1	1	1	1
11)	2.50	1	1	1	1	1	1
12)	2.45	-	-	-	-	-	2
13)	2.41	2	2	2	2	2	-
14)	2.22	-	1	1	1	1	1
15)	2.16	1	-	-	-	-	-
16)	1.06	2	2	2	2	2	2
	0.7	( limite de détection )					

Les types d, e2 à e5, f2, h2, i2, m3, n3, p1, p2 et q1 n'étaient pas inclus dans l'analyse.

## h) Polymorphisme à l'intérieur d'une plante

Une seule plante parmi la totalité des échantillons utilisés ne se classait pas sous un type cytoplasmique net.

Il s'agit de la plante n<sup>o</sup> 84), un O. sativa N<sup>o</sup> 103 919 (IRRI) mâle-stérile cytoplasmique, qui se montrait chimérique entre les types a1 et a2: ses patrons de restriction correspondaient à un mélange à parts égales de ces deux types cytoplasmiques (comp. fig. IV C, piste 4, et fig. VIII A, piste 2, dans l'annexe C).

## i) Résolution de la distinction des différents plastotypes

La distinction des 32 plastotypes (désignés a 1 à q 2) repose exclusivement sur les données Eco R I et Ava I, la première de ces deux enzymes ayant été appliquée sur la totalité du matériel, la deuxième sur la plus grande partie. Les autres cinq enzymes ne furent appliquées que sur des représentants des types définis par les deux premières.

Le pouvoir résolutif du couple Eco R I / Ava I s'avère cependant comparable à celui de Msp I, une endonucléase spécifique d'une séquence de 4 bases: CC GG. Celle-ci coupe l'ADNcp du riz en plus de 200 fragments (voir les figures XII A et B dans l'annexe C):

Dans une série rapide d'essais sur 210 de nos échantillons, Msp I ne révélait à première vue que 2 subdivisions sur les 32 types cytoplasmiques:

- parmi les O. rufipogon de l'Australie, W 1627 se montre différent du reste de la classe f1; et

- dans la classe b la variété cultivée S 24 (O. sativa indica) semble se distinguer de ses proches sauvages, les variétés W 1692 et W 1699 d'O. rufipogon.

Dans l'ensemble, la complexité prononcée des patrons ainsi que quelques difficultés avec la reproductibilité des patrons électrophorétiques, liées à la procédure délicate de la préparation des gels à gradient de concentration (cf. annexe E), n'ont pas permis une analyse plus approfondie des résultats Msp I. Mais en gros on peut dire que ceux-ci confirment les différences entre les types définis par le couple Eco R I / Ava I, sans aller beaucoup plus loin dans la différenciation.

### j) La régularité entre électrophorèses

La constance des patrons observés, en général peu sensible aux changements de concentration des gels, aux tampons utilisés (comp. paragraphe 1**)b**) ci-dessus) et aux voltages appliqués, manifeste une bonne régularité entre les électrophorèses.

Une indication quantitative de cette régularité peut être tirée de la comparaison de bandes communes aux enzymes Ava I et Sma I, à l'existence desquelles on doit s'attendre, puisque les sites Sma I font également partie des sites Ava I. Ainsi nous retrouvons la bande double n<sup>o</sup> 54) d'Ava I, estimée à 1.11 kb, correspondant à la bande double n<sup>o</sup> 23) chez Sma I, estimée à 1.10 kb. La bande Ava I - 56), spécifique des patrons du Groupe Sativa (types a à i) et estimée à 1.04 kb, correspond bien à la bande Sma I - 24), estimée à 1.03 kb. En outre, il y a chez Sma I une bande spécifique aux types I à p, c'est le n<sup>o</sup> 9), estimé à 8.8 kb, à laquelle peut correspondre une paire de deux éléments chez Ava I, constituant les bandes 8) et 18), estimées à 5.47 kb et à 3.33 kb, respectivement. D'autres comparaisons sont probablement possibles.

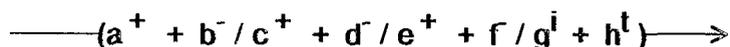
## L'ANALYSE DES MUTATIONS

Le système des mutations déchiffrées à partir du polymorphisme des longueurs des fragments de restriction dont la présentation suit reflète le résultat de leur analyse cladistique. A part quelques coïncidences peu importantes sur de petites additions ou délétions, qui seront discutées séparément, chaque mutation ne figure qu'une fois, lors de la désignation de la branche de l'arbre cladistique (fig. 4, à la fin du chapitre) dont elle est spécifique.

La présentation suit le chemin de la construction de ce cladogramme, en remontant successivement des niveaux inférieurs d'intégration par sous-ensembles génotypiques vers les niveaux supérieurs. Chaque fois, c'est l'état dit dérivé en commun sur un nombre donné de sites mutés qui sert de critère d'intégration d'un nombre (toujours le nombre minimal possible) de sous-ensembles dans le sous-ensemble plus grand précédent (à quoi correspond dans la représentation dessinée la liaison de deux branches par une bifurcation).

La notation est effectuée sous forme d'équations de transformation. Les fragments primitifs, correspondant aux états antérieurs aux mutations, figurent sur la gauche, les fragments dérivés, correspondant aux états postérieurs aux mutations figurent sur le côté droit des équations. Les types de mutation sont marqués sur les flèches symbolisant la transformation selon le code suivant:

une marque du type



symbolise une transformation qui s'explique par

a	apparitions de sites	plus
b	disparitions de sites	plus
c	petites insertions	plus
d	petites délétions	plus
e	grandes insertions	plus
f	grandes délétions	plus
g	inversions	plus
h	transpositions.	

En ce qui concerne les insertions et délétions, les petites d'entre elles sont supposées expliquer des différences de taille de fragments ne dépassant pas les 40 pb. Les "grandes" d'entre elles ne sont supposés que si elles se manifestent dans les patrons de plusieurs enzymes.

Puisque les inversions, transitions et grandes insertions ou délétions sont des mutations rares par rapport aux mutations de sites et aux petites additions ou délétions (cf. PALMER, 1985a), les indications se limitent en général sur ces dernières, donc sur les deux premières positions des indicateurs: par exemple  $\text{---}(1^+ / -)\text{---}\rightarrow$  signifie simplement l'apparition d'un site,  $\text{---}(- / 1^-)\text{---}\rightarrow$  une petite délétion.

La déduction immanente, sans repère extérieur, de la racine d'un cladogramme représente une impossibilité logique. C'est pourquoi, pour notre branche centrale, qui sépare le Groupe Sativa du Groupe Latifolia, l'orientation des mutations est fixé arbitrairement: dans le sens du dernier vers le premier. Pour les autres comparaisons entre sous-ensembles ce qui est externe aux sous-ensembles en question sert de référence pour l'identification des types primitifs et dérivés.

Les mutations sont numérotées de façon continue et classées sous les branchements auxquels elles sont spécifiques, qui sont désignés par des lettres capitales.

### a) Premier niveau d'intégration

L'intégration des différents types cytoplasmiques dans des premiers sous-ensembles représente les groupements avec les diversités internes les plus faibles possibles, généralement une seule mutation.

#### \*A1) a2 → b :

Les différences entre a2 et b, révélés par Eco R I et par Ava I, peuvent s'expliquer par une seule petite addition :

$$1) \left. \begin{array}{l} \text{Eco R I:} \quad 3.74 \\ \text{Ava I:} \quad 0.86 \end{array} \right\} \xrightarrow{-( / 1^+ )} \left\{ \begin{array}{l} 3.79 , \\ 0.89 ; \end{array} \right.$$

#### A2) a2 → c :

$$2) \text{Eco R I:} \quad 3.74 \quad \xrightarrow{-( / 1^- )} \quad 3.70 ;$$

$$3) + 3a) \text{Ava I:} \quad 2 \times 2.44 \quad \xrightarrow{-( / 2^- )} \quad 2 \times 2.42 ;$$

Comme il pourrait s'agir de deux délétions de 20 pb localisées à proximité suffisante l'une de l'autre pour tomber dans un même fragment (de 3.74 kb) chez Eco R I, mais suffisamment éloignées l'une de l'autre pour tomber sur deux fragments différents (par hasard de taille identique) chez Ava I, le nombre de mutations à retenir est 2 selon l'esprit de la parcimonie.

Nous reviendrons sur la nature et l'origine du type c.

#### B1) e1 → e2 :

$$4) \left. \begin{array}{l} \text{Eco R I:} \quad 6.12 \\ \text{Ava I:} \quad 3.07 \\ \text{Bst I:} \quad 4.64 \\ \text{Sma I:} \quad \text{néant;} \end{array} \right\} \xrightarrow{-( / - / 1^+ )} \left\{ \begin{array}{l} 6.27 , \\ 3.23 , \\ 4.8 , \end{array} \right.$$

Bien qu'elle ne soit pas observée dans le patron Sma I, où elle est peut-être masquée dans le grand fragment d'environ 19,2 kb, une grande addition d'à peu près 150 pb semble bien établi pour le cytoplasme e2.

\* rappel: A désigne une polyfurcation, donc A1 la première branche latérale de celle-ci; a2 représente le type primitif, b le type dérivé;

B2) e1 → e3 :

$$5) \quad \underline{\text{Eco R I}}: \quad 1.64 \quad \text{---} (1^+ / -) \text{---} \rightarrow \quad 1.09 + 0.55 ;$$

B3) e1 → e4 :

$$6) \quad \underline{\text{Eco R I}}: \quad 3.70 \quad \text{---} (- / 1^-) \text{---} \rightarrow \quad 3.66 ;$$

B4) e1 → e5 :

$$7) \quad \underline{\text{Eco R I}}: \quad 6.12 \quad \text{---} (1^+ / -) \text{---} \rightarrow \quad 3.92 + 2.20 ;$$

\*C) f1 → f2 :

$$8) \quad \underline{\text{Eco R I}}: \quad 2 \times 1.71 \quad \text{---} (- / 1^-) \text{---} \rightarrow \quad 2 \times 1.68 ;$$

Bien qu'une bande double soit déplacée, il suffit de supposer 1 seule mutation, située sur les deux grandes unités répétées (IR) du plastome.

D) h1 → h2 :

$$9) \quad \underline{\text{Eco R I}}: \quad 3.53 \quad \text{---} (- / 1^-) \text{---} \rightarrow \quad 3.49 ;$$

$$10) \quad \underline{\text{Ava I}}: \quad 0.82 + 0.86 \quad \text{---} (- / 2^-) \text{---} \rightarrow \quad 0.81 + 0.80 ;$$

Les deux mutations Ava I pourraient se trouver sur le même fragment Eco R I.

E) i1 → i2 :

$$11) \quad \left. \begin{array}{l} \underline{\text{Eco R I}}: \quad 3.82 \\ \underline{\text{Ava I}}: \quad 2.70 \end{array} \right\} \quad \text{---} (- / - / 1^+) \text{---} \rightarrow \quad \left\{ \begin{array}{l} 3.90 , \\ 2.80 ; \end{array} \right.$$

Il pourrait s'agir d'une même addition d'environ 90 bp.

F) m2 → m3 :

$$12) \quad \underline{\text{Ava I}}: \quad 2.42 \quad \text{---} (- / 1^+) \text{---} \rightarrow \quad 2.44 ;$$

\* Il n'y a pas de numérotation de branches pour les bifurcations avec une seule branche latérale (dérivée), l'autre branche restant primitive (sans mutation spécifique);

G) n2 → n3 :

$$13) \left. \begin{array}{l} \text{Eco R I:} \quad 6.39 + 3.70 \\ \text{Ava I:} \quad 3.33 + 1.55 \\ \text{Bam H I:} \quad 5.17 \\ \text{Sma I:} \quad 9.6 \\ \text{Pst I:} \quad 16.3 \end{array} \right\} \xrightarrow{(-/-/1^-)} \left\{ \begin{array}{l} 9.7 , \\ 4.49 , \\ 4.75 , \\ 9.2 , \\ 15.9 ; \end{array} \right.$$

Une délétion d'environ 400 pb qui emmène un site Eco R I et un site Ava I.

H) p2 → p3 :

$$14) \left. \begin{array}{l} \text{Eco R I:} \quad 3.79 \\ \text{Ava I:} \quad 2.71 \\ \text{Hind I:} \quad 6.94 \\ \text{Sma I:} \quad 4.75 \end{array} \right\} \xrightarrow{(-/-/1^+)} \left\{ \begin{array}{l} 3.90 , \\ 2.80 , \\ 7.04 , \\ 4.86 ; \end{array} \right.$$

Une insertion d'approximativement 100 pb.

I) q1 → q2 :

$$15) \left. \begin{array}{l} \text{Eco R I:} \quad 2.73 \\ \text{Ava I:} \quad 2.38 \\ \text{Bst I:} \quad 1.35 \end{array} \right\} \xrightarrow{(-/1^-)} \left\{ \begin{array}{l} 2.69 , \\ 2.35 , \\ 1.33 ; \end{array} \right.$$

Il pourrait s'agir d'une même petite délétion.

$$16) \text{Bst I:} \quad 1.39 \xrightarrow{(-/1^+)} 1.42 ;$$

## b) Deuxième niveau d'intégration

Ce niveau est voisin du premier et il reste encore plusieurs types non regroupés. De fait, la définition de chaque niveau d'intégration représente surtout un choix pragmatique.

J) a1 → centre a2/b/c (id. a2) :

Le centre a2/b/c comprend le type cytoplasmique hypothétique qui réunit les caractères primitifs présents sur l'ensemble des branches subséquentes, ici ce sont

les trois branches conduisant aux types a2, b et c (voir ci-dessus), et le type a2 coïncide avec ce type central hypothétique.

$$17) \begin{array}{l} \text{Ava I:} \\ \text{Sma I:} \end{array} \begin{array}{l} 2.15 \\ 14.6 \end{array} \left. \vphantom{\begin{array}{l} \text{Ava I:} \\ \text{Sma I:} \end{array}} \right\} \xrightarrow{-(1^+ / -)} \left. \vphantom{\begin{array}{l} 1.78 + 0.37 ; \\ 11.4 + 3.17 ; \end{array}} \right\} \begin{array}{l} 1.78 + 0.37 ; \\ 11.4 + 3.17 ; \end{array}$$

K1) centre g/h  $\rightarrow$  g :

$$18) \text{Ava I:} \quad 1.55 \xrightarrow{- / 1^-} 1.54 ;$$

K2) centre g/h  $\rightarrow$  centre h (id. h1) :

$$19) \text{Ava I:} \quad 1.55 \xrightarrow{- / 1^+} 1.56 ;$$

L1) m1  $\rightarrow$  l :

$$20) \text{Ava I:} \quad 2.42 \xrightarrow{- / 1^+} 2.46 ;$$

L2) m1  $\rightarrow$  m4 :

$$21) \text{Ava I:} \quad (0.05 + ) 3.33 \xrightarrow{(1^- / -)} 3.38 ;$$

En principe, une addition ou délétion de 50 pb est suffisamment grande pour se manifester dans les patrons de plusieurs enzymes, comme le montrent beaucoup d'exemples dans cette analyse. C'est pourquoi, en l'absence d'autres indices qu'il s'agisse d'une insertion, recherchés avec trois autres enzymes (Ava I, Sma I, Bst I), nous supposons ici une disparition d'un site au bord d'un fragment d'une taille inférieure à la limite de détectabilité.

L3) m1  $\rightarrow$  centre m2/3 (id. m2) :

$$22) \text{Ava I:} \quad 6.85 + 0.62 \xrightarrow{(1^- / -)} 7.5 ;$$

L4) m1  $\rightarrow$  n1 :

$$23) \text{Eco R I:} \quad 0.75 + 0.4 \xrightarrow{(1^- / -)} 1.15 ;$$

L5) m1  $\rightarrow$  centre n2/3 (id. n2) :

$$24) \text{Eco R I:} \quad 1.00 \xrightarrow{- / 1^+} 1.02 ;$$

M1) centre o/p1  $\rightarrow$  o :

$$25) \text{Eco R I:} \quad 1.76 \xrightarrow{- / 1^-} 1.73 ;$$

$$26) \text{ Ava I: } 2.15 + 0.82 \xrightarrow{-(1^- / 1^+)} 3.00 ;$$

M2) centre o/p1  $\rightarrow$  p1 :

$$27) \text{ Bst I: } 2.55 \xrightarrow{- / 1^+} 2.58 ;$$

### c) Troisième niveau d'intégration

N1) centre a/b/c/d  $\rightarrow$  centre a/b/c (id. a1) :

$$28) \begin{array}{l} \text{Eco R I: } 0.89 \\ \text{Ava I: } 4.49 \\ \text{Hind III: } 3.41 \\ \text{Pst I: } 3.87 \\ \text{Bst E II: } 2.22 \end{array} \left. \vphantom{\begin{array}{l} \text{Eco R I: } 0.89 \\ \text{Ava I: } 4.49 \\ \text{Hind III: } 3.41 \\ \text{Pst I: } 3.87 \\ \text{Bst E II: } 2.22 \end{array}} \right\} \xrightarrow{- / - / 1^-} \left. \vphantom{\begin{array}{l} 0.83, \\ 4.43, \\ 3.36, \\ 3.80, \\ 2.16; \end{array}} \right\}$$

$$29) \text{ Eco R I: } 0.84 \xrightarrow{- / 1^+} 0.86 ;$$

N2) centre a/b/c/d  $\rightarrow$  d :

$$30) \text{ Eco R I: } (0.06 + ) 2.35 \xrightarrow{(1^- / -)} 2.41 ;$$

O1) centre g/h/i  $\rightarrow$  centre g/h :

$$31) \text{ Eco R I: } 3.70 \xrightarrow{(1^+ / -)} 3.05 + 0.65 ;$$

$$32) \text{ Hind III: } 2.67 \xrightarrow{(1^+ / -)} 1.46 + 1.22 ;$$

O2) centre g/h/i  $\rightarrow$  centre i (id. i1) :

$$33) \text{ Eco R I: } 6.12 + 3.70 \xrightarrow{(1^- / -)} 9.8 ;$$

$$34) \text{ Eco R I: } 1.00 + 0.80 \xrightarrow{(1^- / -)} 1.79 ;$$

$$35) \text{ Ava I: } 2 \times 3.07 \xrightarrow{(1^+ / -)} 2 \times 3.00 ( + 2 \times 0.07),$$

probablement une apparition d'un site sur les IR;

$$36) \text{ Ava I: } 2 \times 2.42 \xrightarrow{- / 1^-} 2 \times 2.40 ,$$

peut-être une petite délétion, également sur les IR.

P) centre o/p (id. p1) → centre p2/3 (id. p2) :

- 37) Eco R I: 0.86  $\xrightarrow{(-/1^+)} 0.87$  ;  
 38) Eco R I: 2.42  $\xrightarrow{(-/1^+)} 2.44$  ;  
 39) Bst I: 0.84 + 8.1  $\xrightarrow{(1^-/-)} 8.95$  ;  
 40) Hind III: 7.75  $\xrightarrow{(1^+/-)} 6.6 + 1.17$  ;

#### d) Quatrième niveau d'intégration

Q) centre a-e (id. e1) → centre a/b/c/d :

- 41) Eco R I: 3.70  $\xrightarrow{(-/1^+)} 3.74$  ;  
 42) + 42a) Ava I: 2 x 2.42  $\xrightarrow{(-/2^+)} 2 x 2.44$  ;

R1) centre f/g/h/i → centre f (id. f1)

- 43) Eco R I: 1.33  $\xrightarrow{(1^+/-)} 0.77 + 0.56$  ;  
 44) Ava I: 1.55  $\xrightarrow{(-/1^-)} 1.54$  ;  
 45) Ava I: 4.49  $\xrightarrow{(-/1^-)} 4.46$  ;  
 46) Hind III: 3.41  $\xrightarrow{(-/1^+)} 3.44$  ;

Cette insertion apparaît comme une réversion de la délétion n<sup>o</sup> 107), mais il devrait plutôt s'agir d'une coïncidence sans homologie.

R2) centre f/g/h/i → centre g/h/i :

- 47) Eco R I: 3.79 }  $\xrightarrow{(-/1^+)} \left\{ \begin{array}{l} 3.82 , \\ 2.70 ; \end{array} \right.$   
Ava I: 2.67 }

S1) centre I-p → centre I/m/n (id. m1) :

48)	<u>Eco R I</u> :	2.9	} → (- / - / 1 <sup>-</sup> ) →	}	2.61 ,
	<u>Ava I</u> :	5.1			4.82 ,
	<u>Bst I</u> :	6.94			6.62 ,
	<u>Hind III</u> :	2.44			2.14 ,
	<u>Pst I</u> :	4.66			4.34 ,
	<u>Bst E II</u> :	11.9			11.6 ,
	<u>Sma I</u> :	néant;			

Malgré la confirmation manquante par Sma I, il est hautement probable qu'il s'agit d'une grande délétion d'environ 300 pb.

$$49) \quad \underline{\text{Eco R I}}: \quad (0.05 + ) 6.34 \quad \longrightarrow (1^- / -) \longrightarrow 6.39 ;$$

$$50) \quad \underline{\text{Eco R I}}: \quad 2.29 \quad \longrightarrow (- / 1^-) \longrightarrow 2.25 ;$$

$$51) \quad \underline{\text{Eco R I}}: \quad 0.71 \quad \longrightarrow (- / 1^+) \longrightarrow 0.75 ;$$

A la place des mutations 50) et 51) on pourrait aussi noter une inversion ou une légère transition au lieu de l'addition et de la délétion. Mais pour cette possibilité il ne se trouve aucun indice supplémentaire chez les autres enzymes. Une autre possibilité d'explication pour cette constellation serait une disparition et une apparition d'un site en voisinage.

$$52) \quad \begin{array}{l} \underline{\text{Eco R I}}: \quad 0.69 \\ \underline{\text{Bst I}}: \quad 2.55 \end{array} \quad \left. \vphantom{\begin{array}{l} \underline{\text{Eco R I}}: \quad 0.69 \\ \underline{\text{Bst I}}: \quad 2.55 \end{array}} \right\} \longrightarrow (- / 1^-) \longrightarrow \left. \vphantom{\begin{array}{l} 0.67 , \\ 2.52 ; \end{array}} \right\} \begin{array}{l} 0.67 , \\ 2.52 ; \end{array}$$

S2) centre I-p → centre o/p :

$$53) \quad \underline{\text{Eco R I}}: \quad 2.9 \quad \longrightarrow (1^+ / -) \longrightarrow 1.76 + 1.18 ;$$

$$54) \quad \underline{\text{Eco R I}}: \quad 0.84 \quad \longrightarrow (- / 1^-) \longrightarrow 0.83 ;$$

$$55) \quad \underline{\text{Eco R I}}: \quad 0.80 \quad \longrightarrow (- / 1^+) \longrightarrow 0.81 ;$$

### e) Cinquième niveau d'intégration

T1) centre a-i  $\longrightarrow$  centre a-e (id. e1<sup>\*</sup>) :

$$56) \quad \text{Eco R I:} \quad 2 \times 3.70 \quad \longrightarrow (1^- / -) \longrightarrow 7.4 ;$$

T2) centre a-i  $\longrightarrow$  centre f/g/h/i :

$$57) \quad \text{Bsp X I:} \quad \dots \quad \longrightarrow (1 / -) \longrightarrow 1.10 + \dots ;$$

En absence d'un marqueur spécifique à ce branchement parmi les données des 7 enzymes susdites, nous incluons ici un résultat préliminaire d'une analyse avec une autre enzyme (Bsp X I, isoschizomère stricte de Cla I) au niveau de la tribu des Oryzées. Elle révèle pour les cytoplasmes f,g,h et i, à la 27ième position du patron des bandes, un fragment d'environ 1.10 kb et ceci en exclusivité pour l'ensemble de la tribu (fig. XI dans l'annexe C). La bonne individualisation de cette bande du reste du patron fait penser qu'elle est due à une mutation de site.

U1) centre k-p  $\longrightarrow$  k :

Ici, c'est surtout l'intégration des patrons Eco R I et Ava I qui pose des problèmes complexes. Nous proposons des solutions consistant d'un minimum de mutations en respectant leur cohérence, mais sans vouloir prétendre qu'elles soient les seules possibles.

$$58) \quad \text{Eco R I:} \quad 0.65 + 0.84 \quad \longrightarrow (1^- / -) \longrightarrow 1.48 ;$$

$$59) \quad \text{Ava I:} \quad 2.65 \quad \longrightarrow (1^+ / -) \longrightarrow 1.59 + 1.06 ;$$

$$60) \quad \text{Ava I:} \quad 4.62 \quad \longrightarrow (1^+ / -) \longrightarrow 3.27 + 1.36 ;$$

$$61)^{**} \quad \text{Sma I:} \quad 19.2 \quad \longrightarrow (1^+ / -) \longrightarrow 12.3 + 6.91 ;$$

$$62) \quad \text{Bst I:} \quad 5.30 \quad \longrightarrow (1^+ / -) \longrightarrow 3.74 + 1.56 ;$$

$$63) \quad \text{Bst I:} \quad 6.94 \quad \longrightarrow (1^+ / -) \longrightarrow 4.55 + 2.40 ;$$

$$64) \quad \text{Eco R I:} \quad 0.89 \quad \longrightarrow (- / 1^+) \longrightarrow 0.91 ;$$

$$65) \quad \text{Eco R I:} \quad 2.90 \quad \longrightarrow (- / 1^+) \longrightarrow 2.94 ;$$

\* à noter la primitivité prononcée du type e1

\*\* la mutation n<sup>o</sup> 61) devrait être identique à l'une des deux mutations n<sup>os</sup> 59) ou 60);

$$66) + 67) \text{ Ava I: } 3 \times 0.62 \quad \xrightarrow{(- / 2^-)} \quad 3 \times 0.61 ;$$

$$68) \begin{array}{l} \text{Eco R I: } 1.00 + 0.80 \\ \text{Bst I: } 5.17 + 0.76 \end{array} \left. \vphantom{\begin{array}{l} 1.00 + 0.80 \\ 5.17 + 0.76 \end{array}} \right\} \xrightarrow{(- / - / - / 1^i)} \left\{ \begin{array}{l} 1.17 + 0.64 , \\ 5.23 + 0.70 ; \end{array} \right.$$

- une inversion, le segment retourné portant 1 site Eco R I à 100 pb du centre et 1 site Bst I à 30 pb du centre.

$$69) + 70) \text{ Ava I: } (0.19 + ) 5.47 \quad \xrightarrow{(1^+ + 1^- / -)} \quad 3.20 + 2.46 ;$$

Le reste des mutations proposées ici fait intervenir, dans des constellations changeantes et en combinaison avec d'autres mutations, trois grandes délétions des tailles 0.3 kb, 0.13 kb et 0.07 kb, appelées a, b, c.

$$71) \text{ Ava I: } 3.33 \quad \xrightarrow{(- / - / b)} \quad 3.20 ;$$

$$72) \text{ Ava I: } 2.53 \quad \xrightarrow{(- / - / c)} \quad 2.46 ;$$

$$71) + 72) \text{ Pst I: } 15.0 \quad \xrightarrow{(- / - / b+c)} \quad 14.8 ;$$

$$73) \begin{array}{l} \text{Ava I: } 11.1 + 0.82 \\ \text{Bst E II: } 6.45 \end{array} \left. \vphantom{\begin{array}{l} 11.1 + 0.82 \\ 6.45 \end{array}} \right\} \xrightarrow{(- / - / a)} \left\{ \begin{array}{l} 11.6 , \\ 6.15 ; \end{array} \right.$$

$$71) - 73) \begin{array}{l} \text{Bst I: } 8.95 \\ \text{Hind III: } 7.75 \\ \text{Sma I: } 8.8 \end{array} \left. \vphantom{\begin{array}{l} 8.95 \\ 7.75 \\ 8.8 \end{array}} \right\} \xrightarrow{(- / - / a+b+c)} \left\{ \begin{array}{l} 8.5 ; \\ 7.39 ; \\ 8.4 \end{array} \right.$$

$$74) + 75) \text{ Eco R I: } 3.79 + 2.90 \quad \xrightarrow{(2^+ / - / a)} \quad 3.85 + 1.55 + 0.98 ;$$

Une façon élégante d'expliquer la constellation des fragments Ava I à la base des mutations 69) à 72) pourrait être la suivante:

$$5.47 + 3.33 + 2.53 \quad \xrightarrow{(1^+ / - / - / 1^i)} \quad 2 \times 3.20 + 2 \times 2.46 ;$$

c'est à dire par une importante transition et l'apparition d'un site supplémentaire, puisque les sommes des fragments aux deux côtés sont identiques à la dizaine de paires de bases près. Mais la confrontation de cette version avec les données des autres enzymes n'est pas évidente; au contraire elle ferait tomber l'hypothèse des trois délétions rapprochées, ce qui demanderait au total un nombre de mutations encore plus élevé, en contradiction avec le principe de parcimonie qui guide cette analyse.

Par conséquent, l'hypothèse des trois délétions rapprochées, appelées a, b et c ci-dessus, apparait meilleure. (A ce moment-là, a et b devraient se trouver sur le segment commun aux fragments Pst I-4 et Sal I-1 selon la carte établi par HIRAI (fig. 6), c

sur le segment commun aux fragments Pst I-1 et Sal I-1.)

U2) centre k-p → centre l-p :

$$76) \quad \underline{\text{Bst I}}: \quad 1.39 \quad \text{---}(- / 1^+) \text{---} \rightarrow 1.42 ;$$

### f) Sixième niveau d'intégration

V1) centre k-q → centre k-p :

$$77) + 78) \quad \underline{\text{Eco R I}}: \quad 4 \times 1.45 \quad \text{---}(2^- / -) \text{---} \rightarrow 2 \times 2.9 ;$$

$$79) \quad \underline{\text{Bst I}}: \quad 3.10 \quad \text{---}(- / 1^+) \text{---} \rightarrow 3.12 ;$$

V2) centre k-q → centre q (id. q1) :

La liste ci-dessous se comprend comme proposition parcimonieuse.

$$80) + 81) \quad \underline{\text{Eco R I}}: \quad 2.9 \quad \text{---}(2^+ / -) \text{---} \rightarrow 1.16 + 0.96 + 0.77 ;$$

$$82) \quad \underline{\text{Eco R I}}: \quad 1.83 + 0.89 \quad \text{---}(1^- / -) \text{---} \rightarrow 2.73 ;$$

$$83) \quad \underline{\text{Eco R I}}: \quad 1.38 + 0.48 \quad \text{---}(1^- / -) \text{---} \rightarrow 1.86 ;$$

$$84) \quad \underline{\text{Eco R I}}: \quad 1.38 \quad \text{---}(1^+ / -) \text{---} \rightarrow 0.86 + 0.52 ;$$

$$85) \quad \underline{\text{Bst I}}: \quad 7.9 \quad \text{---}(1^+ / -) \text{---} \rightarrow 7.15 + 0.76 ;$$

$$86) \quad \underline{\text{Bst I}}: \quad (0.13 + ) 3.10 \quad \text{---}(1^- / -) \text{---} \rightarrow 3.23 ;$$

$$87) \quad \underline{\text{Bst I}}: \quad 2.00 + 1.35 \quad \text{---}(1^- / -) \text{---} \rightarrow 3.36 ;$$

$$88) \quad \underline{\text{Sma I}}: \quad (0.3 + ) 9.3 \quad \text{---}(1^- / -) \text{---} \rightarrow 9.6 ;$$

$$89) \quad \underline{\text{Bst E II}}: \quad (0.18 + ) 4.07 \quad \text{---}(1^- / -) \text{---} \rightarrow 4.25 ;$$

$$90) \quad \left. \begin{array}{l} \underline{\text{Eco R I}}: \quad 1.00 \\ \underline{\text{Ava I}}: \quad 5.47 \\ \underline{\text{Sma I}}: \quad 8.8 \end{array} \right\} \quad \text{---}(- / - / 1^-) \text{---} \rightarrow \left\{ \begin{array}{l} 0.90 , \\ 5.37 , \\ 8.7 ; \end{array} \right.$$

$$91) \quad \left. \begin{array}{l} \underline{\text{Eco R I}}: \quad 3.53 \\ \underline{\text{Bst I}}: \quad 4.86 \\ \underline{\text{Sma I}}: \quad 9.3 \\ \underline{\text{Pst I}}: \quad 5.03 \end{array} \right\} \quad \text{---}(- / - / 1^+) \text{---} \rightarrow \left\{ \begin{array}{l} 3.60 , \\ 4.92 , \\ 9.4 , \\ 5.10 ; \end{array} \right.$$

92)	<u>Eco R I:</u>	2 x 1.45	$\longrightarrow (- / 1^-) \longrightarrow$	2 x 1.42 ;
93)	<u>Eco R I:</u>	1.45	$\longrightarrow (- / 1^-) \longrightarrow$	$\left\{ \begin{array}{l} 1.42 , \\ 3.29 ; \end{array} \right.$
	<u>Ava I:</u>	3.33		
94)	<u>Eco R I:</u>	2.9	$\longrightarrow (- / 1^-) \longrightarrow$	$\left\{ \begin{array}{l} 2.86 , \\ 2.00 ; \end{array} \right.$
	<u>Ava I:</u>	2.05		
95)	<u>Eco R I:</u>	0.95	$\longrightarrow (- / 1^+) \longrightarrow$	$\left\{ \begin{array}{l} 0.96 , \\ 0.89 ; \end{array} \right.$
	<u>Ava I:</u>	0.88		
96)	<u>Eco R I:</u>	0.94	$\longrightarrow (- / 1^-) \longrightarrow$	$\left\{ \begin{array}{l} 0.99 ; \\ 2.38 ; \end{array} \right.$
	<u>Ava I:</u>	2.42		
97)	<u>Eco R I:</u>	0.65	$\longrightarrow (- / 1^-) \longrightarrow$	$\left\{ \begin{array}{l} 0.63 , \\ 1.57 ; \end{array} \right.$
	<u>Ava I:</u>	1.59		
98)	<u>Eco R I:</u>	0.56	$\longrightarrow (- / 1^+) \longrightarrow$	$\left\{ \begin{array}{l} 0.59 , \\ 1.57 ; \end{array} \right.$
	<u>Ava I:</u>	1.55		
99)	<u>Bst E II:</u>	2 x 2.41	$\longrightarrow (- / 1^+) \longrightarrow$	2 x 2.45 ;

La délétion n° 92) et l'insertion n° 99) pourraient se trouver sur les IR à proximité l'une de l'autre telle que leurs effets sur les longueurs des fragments coupés par d'autres enzymes se compensent.

### g) Dernier niveau d'intégration

W) centre k-q (Groupe Latifolia)  $\longrightarrow$  centre a-i (Groupe Sativa) :

Rappelons que, pour ce branchement, l'orientation de la notation est fixée arbitrairement.

100)	<u>Eco R I:</u>	6.34	$\longrightarrow (1^+ / -) \longrightarrow$	6.12 ( + 0.22 ) ;
101) + 102)	<u>Ava I:</u>	5.47	$\longrightarrow (1^+ / 1^+) \longrightarrow$	$\left\{ \begin{array}{l} 4.46 + 1.04 ; \\ 7.8 + 1.03 ; \end{array} \right.$
	<u>Sma I:</u>	8.8		

La probabilité que l'apparition du site et la petite insertion soient indépendantes l'une de l'autre est plus grande que celle que ce soit la séquence insérée qui porte le site.

- 103) Ava I: 2.71 }  $\xrightarrow{-(1^+ / -)}$  { 2.42 + 0.28 ;  
Sma I: 9.6 } { 9.3 ( + 0.3 ) ;
- 104) Bst I: 5.17  $\xrightarrow{-(1^+ / -)}$  4.64 ( + 0.53 ) ;
- 105) Bst I: 2.84  $\xrightarrow{-(1^+ / -)}$  1.72 + 1.12 ;
- 106) Pst I: 7.87  $\xrightarrow{-(1^+ / -)}$  7.73 ( + 0.14 ) ;
- 107) Eco R I: 1.14 }  $\xrightarrow{- / 1^-}$  { 1.10 ,  
Hind III: 3.44 } { 3.41 ;
- 108) Eco R I: 1.02 }  $\xrightarrow{- / 1^-}$  { 1.00 ,  
Hind III: 2.44 } { 2.42 ;
- 109) Eco R I: 0.69  $\xrightarrow{- / 1^-}$  0.67 ;

Dans le reste des transformations à décrire, il y a deux grandes délétions d'à peu près 80 ou 90 pb qui interviennent; elles sont désignées par a et b.

- 110) Eco R I: 3.79 }  $\xrightarrow{- / - / a}$  { 3.70 ,  
Sma I: 7.5 } { 7.4 ;
- 111) Eco R I: 0.48 }  $\xrightarrow{- / - / b}$  { 0.4 ;  
Ava I: 2.53 + 0.62 } { 3.07 ;  
Sma I: 9.4 } { 9.3 ;

La délétion b enlève un site Ava I.

- 110)+111) Pst I: 16.3  $\xrightarrow{- / - / a+b}$  16.1 ;
- 112) Ava I: 3.33  $\xrightarrow{(1^+ / - / a)}$  2.67 + 0.58 ;

## h) Remarques sur les coïncidences et réversions

En ce qui concerne les mutations n° 107) (version Hind III) et n° 46), il a déjà été dit que la similitude du type f au type central du Groupe Latifolia (k-q) doit résulter d'une coïncidence. Il est fort probable que les insertions chez les deux ne soient ni les mêmes ni ne se trouvent au même endroit précis.

Il y a quelques coïncidences qui se produisent sur la base des bandes Ava I -

32) (2.44 kb) et Ava I - 33) (2.42 kb). Si, chez les mutations n° 12) et n° 38) il doit s'agir du même type de coïncidence que chez les n°<sup>os</sup> 107) et 46), car la séparation des types invoqués (m3 et p2/3) est bien établie par neuf autres mutations et parce qu'il n'y a pas d'autre indice pour une recombinaison, c'est d'autant plus vrai pour la mutation n° 42) au sein du Groupe Sativa.

Ce sont cependant les mutations n°<sup>os</sup> 3) et 3a), toujours impliquant les deux bandes Ava I - 32) et - 33), qui méritent le soupçon d'être issues d'une recombinaison entre un type indica dérivé (comme a2) et un type japonica (comme e1): le type c rattrape le phénotype ancestral. Le même phénomène est à constater en regard des mutations n° 2) et n° 41), ce qui renforce l'hypothèse de l'origine hybride du type c.

Donc, mis à part cette recombinaison probablement réelle et vu le petit nombre de coïncidences éventuelles, le système développé ici ne souffre pas de contradictions ou ambiguïtés intrinsèques. Sa représentation par un cladogramme est donnée dans la figure 4.

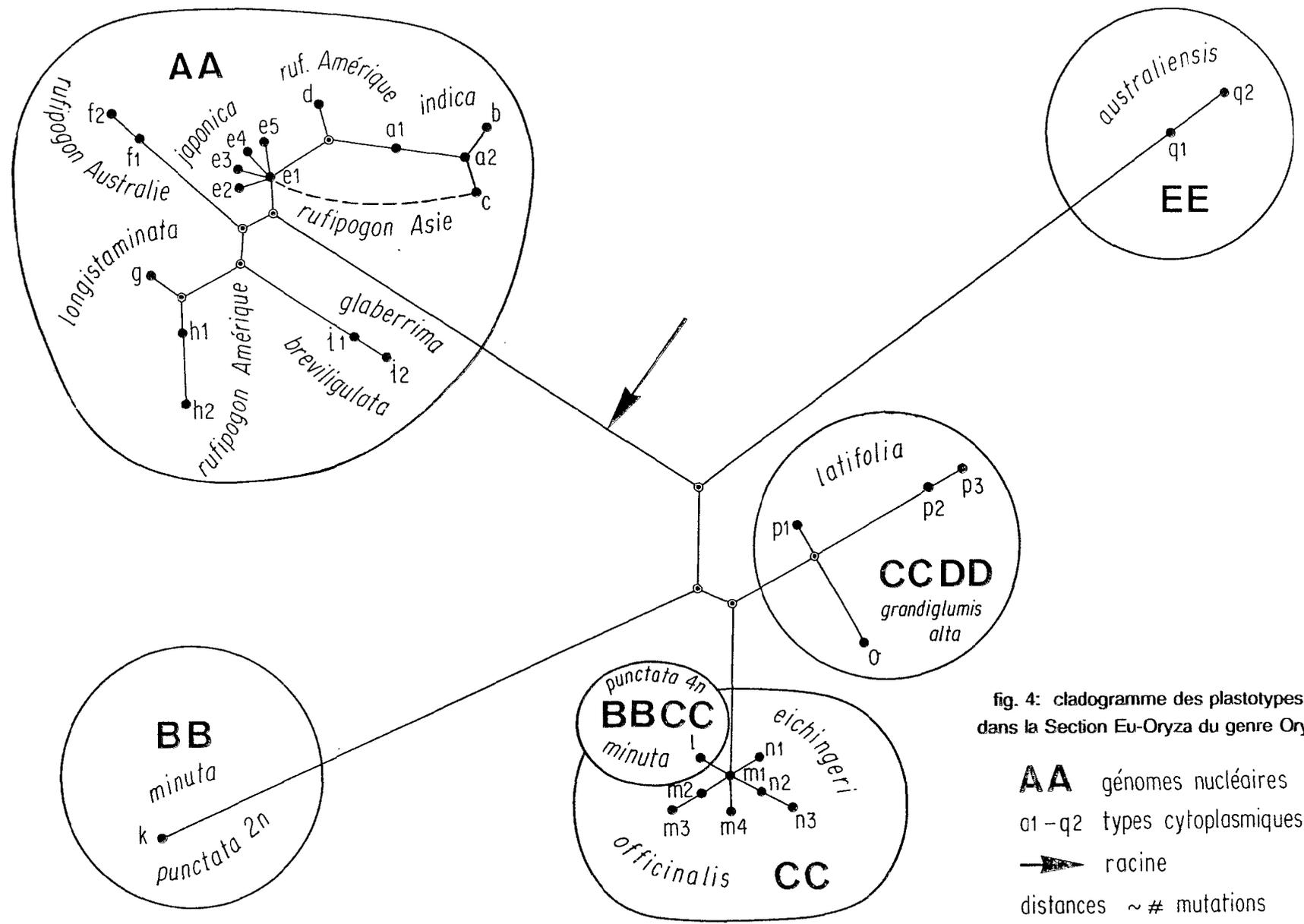
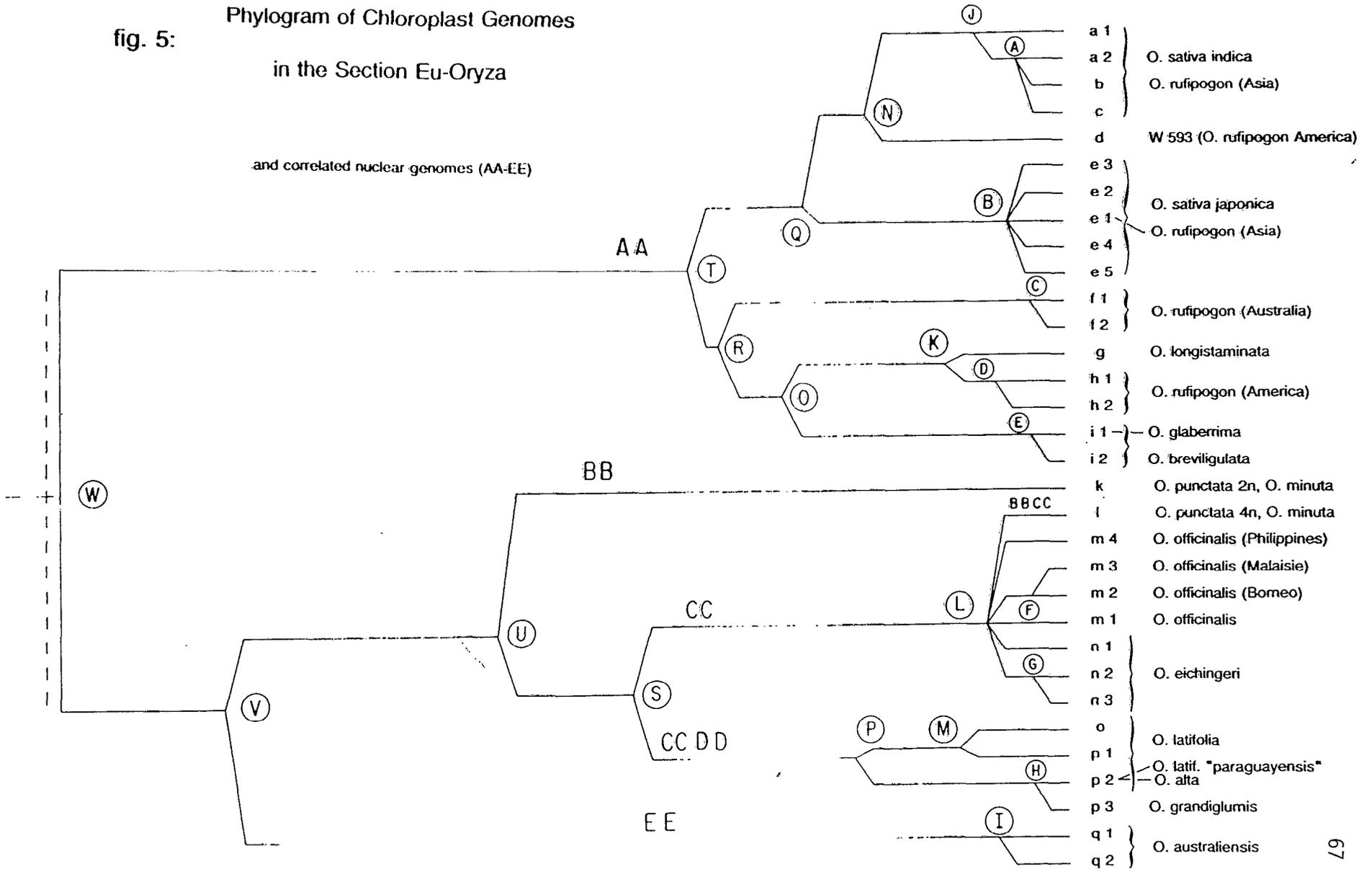


fig. 4: cladogramme des plastotypes dans la Section Eu-Oryza du genre *Oryza*

fig. 5:  
Phylogram of Chloroplast Genomes  
in the Section Eu-Oryza

and correlated nuclear genomes (AA-EE)







## REMARQUES SUR LES RESULTATS

### 1) Comparaisons avec des résultats antérieurs d'analyse des ADNcp de riz

#### a) Détermination des tailles des fragments de restriction

Dans le tableau 10 nous comparons la quantification du patron Pst I de l'ADNcp de la variété Nihon-bare d'O. sativa japonica donnée par ENOMOTO, OGIHARA et TSUNEWAKI (1985) avec celle donnée ci-dessus pour le type e1.

tab. 10. comparaison des quantifications du patron Pst I de l'ADNcp d'Oryza sativa japonica dans ENOMOTO et al. (1985) et dans tab. 8 ci-dessus

bande n <sup>o</sup>	multiplicité x taille de fragment (en kb)	
	selon ENOMOTO et al.	selon tab.8
1)	1 x 19.0	1 x 20
2)	1 x 16.5	1 x 16.1
3)	1 x 15.0	1 x 15.0
4)	1 x 13.6	1 x 14.6
5)	1 x 11.1	1 x 10.7
6)	1 x 10.0	1 x 9.9
7)	2 x 8.8	2 x 8.3
8)	1 x 8.3	1 x 7.73

9)	1 x 5.4	1 x 5.52
10)	1 x 4.9	1 x 5.03
11)	1 x 4.7	1 x 4.66
12)	1 x 4.0	1 x 3.80
13)	1 x 2.2	1 x 2.13
14)	1 x 1.8	1 x 1.85

Pour au moins neuf parmi ces 14 bandes, les différences entre les estimations sont considérables, ce qui pourrait rendre la précision de la deuxième série, attendue à peu de dizaines de bases près, plus ou moins illusoire, au niveau absolu sinon relatif.

Nous avons déjà signalé que, malgré certaines cohérences frappantes lors de l'analyse des mutations de sites, la justification de l'emploi de la deuxième décimale dans notre cas réside dans l'étiquetage des bandes, parfois très denses dans les patrons Eco R I et Ava I.

## b) Interprétations des PLFR

### A) Comparaison avec l'analyse d'ICHIKAWA, HIRAI et KATAYAMA

Notre analyse par Bam H I, voire son isoschizomère Bst I, confirme parfaitement le polymorphisme de la troisième bande parmi les espèces CCDD, décrit par ICHIKAWA, HIRAI et KATAYAMA (1986) lors d'une comparaison de 9 patrons Bam H I au niveau de la section Eu-Oyza, y compris la détermination de la taille du fragment concerné (9.0 kb v 8.95 kb).

Il a cependant échappé à ces auteurs le léger déplacement de cette bande chez les variétés W 1515 et W 1145 d'O. punctata (génome BB) qui est visible sur leur propre photo (comparer leur fig. 4A (page 357) et nos mutations n<sup>os</sup> 71) à 73), version Bst I). Dans la suite de leur étude, la visualisation par "Southern blot" de l'ADNcp, à l'aide d'une sonde extraite d'un riz cultivé et marquée au <sup>32</sup>P, et une faible migration de l'électrophorèse, entraînent une diminution considérable de la définition des bandes qui empêche l'expression de petites différences. Par exemple, pour la variété W 1515 on

reconnaît, sur leur blot, 13 bandes supérieures à 2 kb et 4 bandes inférieures à 2 kb; sur nos gels on en distingue 18 et 9, respectivement.

Loin de confirmer l'identité qu'ils voyaient entre les espèces des génomes BB et CC, nous constatons au contraire une diversification bien prononcée du plastotype k (correspondant au génome nucléaire BB) par rapport à tous les autres plastotypes - au niveau Bam H I comme à n'importe quel autre. En outre, nous ne retrouvons pas la différence entre les espèces de génomes BBCC et CC qu'ICHIKAWA et al. (1986) mentionnent sans donner de précision. Si nous arrivons à distinguer ces deux génomes, ce n'est que par la mutation n<sup>o</sup> 20), une petite insertion révélée par l'enzyme Ava I.

#### B) Comparaison avec l'analyse d'ISHII, TERACHI et TSUNEWAKI

Nous confirmons la première mise en évidence d'une diversité des génomes chloroplastiques parmi les espèces et sous-espèces du riz cultivé par ISHII, TERACHI et TSUNEWAKI (1986), au niveau des patrons Eco R I, Hind III et Pst I.

En particulier, nous confirmons leur observation que les sous-espèces indica et japonica d'O. sativa se distinguent par une délétion chez indica. Il s'agit de la mutation n<sup>o</sup> 28), qui selon ces auteurs se manifeste entre un fragment Pst I de 3.9 kb (japonica) et 3.8 kb (indica) - dans notre estimation 3.87 kb versus 3.80 kb -, et sur un fragment Hind III de 3.7 kb (japonica) et 3.6 kb (indica) - dans notre estimation 3.41 kb versus 3.36 kb. Nous signalons ici que la même délétion s'exprime aussi au niveau Eco R I sur un fragment de taille relativement petite: 0.89 (japonica) contre 0.83 (indica).

En outre nous apercevons, toujours dans les patrons Eco R I, chez indica deux petites insertions par rapport à japonica, ce sont les mutations n<sup>os</sup> 29) et 41), la dernière de ces deux étant même visible sur la photo publiée par ISHII et al. (1986).

L'absence de différences entre O. sativa japonica et O. glaberrima dans les patrons Hind III et Pst I, constatée par ISHII et ses collègues, se vérifie dans notre analyse. Par contre, nos résultats suggèrent de rediscuter la relation entre ces espèces telle qu'elle se présente au niveau Eco R I.

Rappelons que nos types e1 (type central de l'ensemble japonica) et i1 (représentant unique d'O. glaberrima) paraissent séparés par 7 mutations (les n<sup>os</sup> 33) à 36),

47), 56) et 57)), dont 4 sont révélées par Eco R I:

- les 3 mutations de sites

$$n^0 33) \quad 6.12 + 3.70 \longrightarrow 9.8 ,$$

$$n^0 34) \quad 1.00 + 0.80 \longrightarrow 1.79 ,$$

$$n^0 56)^* \quad 7.40 \longrightarrow 2 \times 3.70 \quad \text{et}$$

- la petite insertion

$$n^0 47) \quad 3.79 \longrightarrow 3.82 .$$

D'après ISHII et ses collègues les deux espèces ne se distingueraient que par un seul déplacement d'un site Eco R I, transformant une paire de fragments des tailles 7.3 kb et 5.9 kb en une paire de 10.6 kb et 3.9 kb. Ils estiment négligeable la différence de 1,3 kb entre les deux sommes des tailles respectives, les autres différences passant inaperçues.

Le remplacement de leurs déterminations à 10.6 kb, 3.9 kb, 7.3 kb et 5.9 kb des quatre fragments impliqués dans leur hypothèse par les nôtres à 9.8 kb, 3.82 kb, 7.40 kb et 6.12 kb, respectivement, réduirait à 0.10 kb cet écart. C'est cependant notre connaissance d'autres plastotypes voisins, qui nous fait maintenir notre hypothèse des quatre mutations.

Si sur le plan du génotype nous faisons apparaître deux fragments de la taille 3.70 par la mutation n<sup>0</sup> 56) et disparaître par la mutation n<sup>0</sup> 33), l'un de ces deux ou un troisième de la même taille, respectivement, nous le faisons en présence dans l'analyse d'un phénotype qui représente justement l'état intermédiaire: avec la multiplicité triple de la bande 3.70, la présence de la bande 6.12 et l'absence de la bande 7.40; c'est le type f (existant dans les variants f1 et f2), spécifique aux O. rufipogon d'Australie.

Pour conclure, malgré notre confirmation d'une partie des résultats d'ISHII, TERACHI et TSUNEWAKI, nos investigations sur le même matériel avec les mêmes enzymes de restriction suggèrent une image des relations phylogénétiques parmi les riz cultivés assez différente de celle suggérée par les leurs: Nous voyons une divergence importante entre l'espèce africaine O. glaberrima d'un côté et les deux éco-espèces asiatiques O. sativa japonica et O. sativa indica de l'autre.

---

\* ici, les quatre mutations sont présentées dans le sens O. sativa japonica vers O. glaberrima

## 2) La situation des variétés cultivées mâle-stériles cytoplasmiques

Neuf parmi les 14 plantes mâle-stériles examinées, appartenant à 11 variétés d'*Oryza sativa*, portaient les chloroplastes du type a2, un type dérivé du type a1, qui peut être considéré comme central dans le complexe de la sous-espèce indica. Trois des plantes appartenaient au plastotype a1 même, dont une faisait partie de l'échantillon 103 806 qui dans deux autres cas s'exprimait par le type a2. Par ailleurs, une plante représentait même un mélange moitié-moitié des deux types a1 et a2 (plante n° 84); voir dans l'annexe C les fig IV C et VIII A, pistes 4 et 2, respectivement).

Les plastotypes a1 et a2 se retrouvent parmi les riz sauvages chez *O. rufipogon*. Le type a2 domine largement notre échantillonnage indien (sauf les échantillons particuliers DN et DR, qui représentent des plastotypes japonica). Une dernière variété, Sair-yukoko A, présente le plastotype a, qui n'a pas encore été caractérisé pour le sous-type 1 ou 2.

Tout se présente comme si le cytoplasme de ces *O. rufipogon* d'Inde induisait - probablement par ses mitochondries (cf. MIGNOUNA, VIRMANI, BRIQUET 1987; KADOWAKI et al. 1988) - la stérilité mâle chez ces variétés cultivées. Les pollinisations répétées par la variété mainteneur, portant les chloroplastes du type a1, provoqueraient peu à peu, par d'occasionnelles transmissions paternelles de chloroplastes, une infiltration des cytoplasmes stériles par des chloroplastes originellement liés aux cytoplasmes fertiles. Plus que le dimorphisme de la variété 103 806, c'est l'observation de la chimère susdite chez la variété 103 919 qui va dans ce sens: s'il ne s'agit pas d'un artefact (p.e. par la présence accidentelle de deux plantes dans le même pot, ce que l'on ne peut totalement exclure), la structure mosaïque de cette plante s'expliquerait par un état mixte du zygote dont elle est issue.

Les variétés ci-dessus appartiennent toutes à la sous-espèce indica. A noter que la variété Chinsurah boro II (nous en avons analysé 2 provenances, S1 et 4319), dont le cytoplasme est connu pour induire la stérilité mâle en combinaison avec le noyau de nombreuses variétés de la sous-espèce japonica, présente un plastotype japonica typique.

Les cytoplasmes induisant la stérilité mâle semblent donc spécifiques de chaque sous-espèce. Au niveau des chloroplastes, ils sont peu différents des cytoplasmes normaux respectifs par rapport à la divergence totale des cytoplasmes de l'espèce O. sativa.

### 3) Polymorphisme intravariétal

72 des 320 plantes analysées représentaient des répétitions sur les variétés. Mis à part l'exemplaire mâle-stérile mentionné ci-dessus (plante n° 84) il n'y avait que la plante n° 107) (variété W 135, O. rufipogon d'Inde) qui se retrouvait séparée de deux autres exemplaires de la même variété en exprimant un plastotype différent (a1 au lieu de a2). Bien qu'il puisse s'agir de la reproduction chez une variété sauvage du phénomène d'infiltration liée à la stérilité mâle, tel que décrit ci-dessus, c'est décidément trop peu pour exclure un artefact.

On peut au contraire constater une bonne reproductibilité des résultats à l'intérieur des variétés.

## IMPLICATIONS PHYLOGENETIQUES DE L'ANALYSE DES MUTATIONS

La compatibilité quasiment parfaite des regroupements de plastotypes par les différentes mutations observées, en liaison avec la parcimonie quasiment maximale du cladogramme déduit de ces mutations, suggère fortement que la topologie du cladogramme correspond directement aux relations évolutives entre les plastotypes.

### 1) Transformation du cladogramme en phylogrammes hypothétiques

Ce qui empêche la lecture immédiate du cladogramme en termes phylogénétiques, c'est l'hétérogénéité des longueurs des branches et surtout la position de certains types primitifs sur des branches intérieures de l'arbre, ce qui signifie une coprésence de types ancestraux et récents.

Les propriétés essentielles d'un dendrogramme phylogénétique sont que toutes les extrémités des branches arrivent au même niveau (symbolisant l'époque actuelle) par rapport à une échelle de temps linéaire, et que les branchements d'ordre supérieur (correspondant à nos niveaux supérieurs d'intégration des types) sont placés antérieurement aux branchements subséquents (interdiction de longueurs de branches négatives, auxquelles correspondrait un recul dans le temps pendant l'évolution).

L'approche la plus simple pour déduire un arbre phylogénétique de notre cladogramme serait de traduire chaque pas intégratif (désigné chez nous par une lettre capitale) par une bi- ou poly-furcation, respectivement, en reportant celle-ci sur une échelle de temps relative, directement en proportion du degré de la diversité à intégrer.

Ainsi, tout en conservant la topologie hiérarchisée de notre cladogramme, nous allons identifier les nombres moyens de mutations\* qui séparent deux types ou deux ensembles de types avec le double de leur temps relatif de divergence, et ceci par trois méthodes un peu différentes:

- la première correspondant à l'UPGMA (par groupes de paires non-pondérés),
- la deuxième correspondant à une WPGMA (par groupes de paires pondérés),
- la troisième correspondant à une WPGMA flexibilisée, "fWPGMA".

Voici quelques exemples:

1) la bifurcation D, conduisant du centre h (qui est identique avec h1) aux deux types h1 et h2, est à placer à  $t_D = -1$  (unité de temps), car en moyenne chaque branche subséquente subit  $-t_D = (2 + 0) : 2 = 1$  mutation ;

2) la bifurcation K, conduisant du centre g/h au type g et au centre h, est à placer comme suit:

- a) selon UPGMA: à  $t_K = -1,5$  (unité de temps), car en moyenne chaque branche subséquente subit  $-t_K = (1 + 1 + 1) : 2 = 1,5$  mutations; c'est à dire une mutation spécifique à g, plus une mutation spécifique à l'ensemble h1/h2 (ou au centre h), plus le niveau déjà atteint par la bifurcation D, donc encore une mutation, divisé par le nombre de branches, deux;

- b) selon WPGMA: à  $t_K' = -5/3 = -1.67$  (unité de temps), selon le même calcul, mais en comptant deux fois la branche conduisant au centre h, parce qu'elle porte deux bouts terminaux (h1 et h2) et seulement une fois la branche conduisant au type g, parce que celle-ci ne porte qu'un seul bout terminal (justement g) et en divisant logiquement par  $2 + 1 = 3$ ; donc

$$-t_K' = (2 \times (1 + 1) + 1) : (2 + 1) = 5 : 3 = 1.67 .$$

Les deux premières colonnes du tableau 10 donnent les niveaux de toutes les branchements des phylogrammes hypothétiques calculés selon UPGM et WPGM. La

---

\* ici, tous les types de mutations sont regardés comme équivalents, ce qui semble justifié par l'absence d'une structuration claire dans la distribution de ces types sur les branches du cladogramme, corrélée à la hiérarchie de celui-ci, du moins au-dessus du deuxième niveau d'intégration;

désignation des bifurcations reprend le code du chapitre "analyse des mutations".

tab. 10: niveaux (en unités du temps relatif\* x -1) des bifurcations du phylogramme à déduire du cladogramme, selon trois calculs un peu différents

branchement	UPGM	WPGM	"fWPGM"
A	0,67	0,67	0,67
B	0,8	0,8	0,8
C	0,5	0,5	0,5
D	1,0	1,0	1,0
E	0,5	0,5	0,5
F	0,5	0,5	0,5
G	0,5	0,5	0,5
H	0,5	0,5	0,5
I	1,0	1,0	1,0
J	0,84	1,25	1,25
K	1,5	1,67	1,67
L	1,0	1,13	1,13
M	2,0	2,0	2,0
N	1,9	2,8	2,8
O	4,0	4,0	4,0
P	3,0	3,0	3,0
Q	2,4	2,8	3,7
R	4,75*	4,9*	4,9
S	6,0	6,2	6,2
T	4,6*	4,7*	5,3
U	11,6	7,8	8,0
V	17,7**	11,6	11,8
W	16,8**	13,7	14,1

Malheureusement, ni la version UPGM ni la version WPGM ne sont dessinables, principalement à cause de la longueur négative de la branche TR, signalée par les

\* nous montrerons plus tard que l'unité relative correspond approximativement à un intervalle d'un million d'années;

astérisques dans le tableau:

$$4,6 - 4,75 = -0,15 ; 4,7 - 4,9 = -0,2 .$$

Selon ces calculs, la position de la séparation de f et g+h+i serait antérieure à celle de la séparation de f+g+h+i et a-i, ce qui représenterait une contradiction logique au système cladistique.

La cause de cette irrégularité est le nombre extrêmement faible de mutations atteint par la branche conduisant aux types e. Normalement, c'est justement son intérêt, le calcul par groupes de paires pondérés a comme effet une certaine correction par suppression d'extrêmes. Mais ce principe échoue dans ce cas, car finalement cette branche-là se développe bien et porte 5 bouts de branches.

Si on ramène arbitrairement ce nombre de 5 à 2, l'arbre devient dessinable, comme le montre la colonne "fWPGM" dans le tableau 10. L'effet de la réduction analogue du poids des types m (m1 à m4, O. officinalis) de 4 à 2, également exprimée dans la colonne "fWPGM", est cependant négligeable.

Une telle correction par une utilisation flexibilisée des poids dans les calculs nous paraît justifiée, d'autant que, dans notre cas, la mise en évidence de la diversité des plastotypes e et m fait suite à des nombres de plantes analysées particulièrement élevés, choisies qui plus est dans des grandes variétés d'origines biologiques.

C'est ici le lieu de rappeler que nous avons déjà introduit un autre correctif dans notre analyse. C'est la mutation n° 57), sans laquelle même la troisième version serait à peine dessinable (la longueur de la branche TR se réduirait de 0,6 à 0,1 (unités de mutation)).

La figure 5 montre le dendrogramme phylogénétique hypothétique construit sur la base du cladogramme de la figure 4 à l'aide du calcul flexibilisé du type "fWPGM".

## 2) Interprétabilité du phylogramme

Une telle déduction d'un dendrogramme phylogénétique de l'analyse cladistique d'une série de mutations revient à supposer la présence d'une horloge moléculaire pendant l'évolution, sinon régulière, du moins monotone. Cependant, s'abstraire de la distribution des mutations sur deux branches subséquentes lors de la détermination du niveau temporel d'une bifurcation c'est accepter la possibilité de différences de vitesses - apparentes ou réelles - de cette horloge sur différentes branches et à différentes époques.

La vitesse relative de l'horloge, correspondant à une branche ou à un segment de l'arbre, se calcule facilement par la division du nombre des mutations spécifiques à cette branche ou à ce segment par leur longueur, celle-ci étant égale à la différence des niveaux temporels des branchements qui délimitent la branche ou le segment. Dans la figure 5 A ces vitesses relatives sont portées sur toutes les branches supérieures au premier niveau d'intégration. Elles varient de 0 à environ 2,5 avec une moyenne proche de 1.

### a) Justification des branches de vitesse 0

Il ne faut pas penser que la présence dans notre diagramme d'une branche qui n'est pas caractérisée par des mutations spécifiques, donc qui exprime une vitesse relative de l'horloge moléculaire égale à zéro, soit plus trompeuse que l'absence de cette branche dans le cladogramme de base.

Regardons comme exemple la branche QB qui conduit du centre a-e au centre e. On voit que fusionner la polyfurcation B avec la bifurcation Q n'aurait pas de sens, parce que:

- on ne peut pas descendre le niveau de Q en dessous de celui de N, et
- remonter B signifierait un rallongement simultané de quatre branches à la fois,

portant chacune 1 mutation seulement, tout en conservant l'axe central sans aucune mutation. Evidemment, la solution qui conserve le noeud B sur le niveau bas, correspondant dans une interprétation normale à sa diversité subséquente, et qui n'admet qu'une seule branche avec une évolution très faible ou nulle, en termes d'accumulation de mutations, satisfait au mieux le principe de la parcimonie: elle demande le minimum de coïncidences sur les irrégularités de marche de l'horloge.

### **b) Fiabilité des relations évolutives entre les plastotypes**

La variabilité du taux de mutations sur notre phylogramme pourrait-elle remettre en cause la topologie de celui-ci?

Nous pouvons nous rassurer que ce n'est pas à craindre par une considération abstraite de quelques

### **Aspects de la génétique de la transmission des ADN chloroplastiques.**

Dans les populations naturelles des plantes, les molécules d'ADNcp se trouvent organisées dans toute une hiérarchie de populations. Par quelques exemplaires ils occupent les nucléïdes, dont il y a environ une douzaine par chloroplaste. Comme les cellules du riz sont petites, en correspondance avec leur faible contenu en ADNn, le nombre de chloroplastes dans une cellule parenchymateuse ne devrait pas dépasser quelques douzaines. Une plante issue de reproduction végétative peut consister en plusieurs clones de cellules, de même que les populations entières sont composées de clones de plantes. Finalement ces populations peuvent être intégrées en méta-populations.

L'absence générale de recombinaisons, la ségrégation végétative perpétuée des allèles cytoplasmiques et le faible taux de mutations de l'ADNcp ont pour conséquence une constante autopurification rapide et efficace des populations à tous les niveaux par simple dérive stochastique.

Regardons comme un modèle simple une population de taille constante  $n$ , passant d'une génération à la suivante par automultiplication de ses éléments. Supposons

que chaque élément possède  $f$  chances équivalentes de se propager vers la prochaine génération ( $f > 1$ ). Pour donner un exemple, au niveau de la reproduction des plantes annuelles dans une population donnée,  $f$  peut être considéré comme étant égal au nombre moyen de graines fertiles par plante. Dans une situation inter-populations,  $f$  est lié au taux de migration tel que défini par TAKAHATA et MARUYAMA (1981).

Si par exemple  $f = 3$ , par le simple hasard il y a des éléments qui ont trois ou deux descendants au lieu d'un seul qui correspond à la moyenne statistique; logiquement il y en a donc qui n'en ont aucun - voilà la perte de clones par dérive stochastique.

Appelons  $p$  le taux de cette perte par génération. Evidemment ce taux est fonction de  $f$ . Il est  $p = 0$  si  $f = 1$ . Mais 1 n'est pas une valeur réaliste pour  $f$ . Par exemple, on suppose sans problème que  $f = 2$  pour les processus de réplication corrélés avec la division cellulaire, puisqu'il n'y a pas de mécanisme cellulaire connu qui puisse éviter la répartition aléatoire des copies moléculaires sur les organites filles ou des chloroplastes sur les cellules filles (cf. BIRKY 1983).

Il est possible de déduire une formule théorique exprimant la relation entre  $f$  et  $p$ , mais elle est complexe et peu claire. Quelques vérifications simples par calculs de probabilités permettent cependant d'estimer que  $p > 0,2$  pour  $f = 2$  ou  $f > 2$  et  $n = 4$  ou  $n > 4$ :

$$\begin{array}{lll} f = 2, & n = 2 \Rightarrow & p = 1/6 = 0,167 ; \\ f = 2, & n = 3 \Rightarrow & p = 1/5 = 0,2 ; \\ f = 2, & n = 4 \Rightarrow & p = 3/14 = 0,214 ; \\ f = 3, & n = 3 \Rightarrow & p = 1/6 = 0,167 ; \\ f = 3, & n = 4 \Rightarrow & p = 6/19 = 0,316 . \end{array}$$

On peut donc conclure que la perte moyenne de clones par génération (appelée  $p$ ) dépasse facilement les 20 % , si  $f$  n'est pas inférieur à 2 .

Comme estimation de l'intervalle  $t_{\text{fix}}$  nécessaire pour la fixation d'un allèle dans une population nous allons déterminer l'âge moyen  $a$  (en générations) du clone ancestral le plus jeune commun à l'ensemble des éléments de la population actuelle ( $t_{\text{fix}} = a$ ):

Remontons au delà des  $n$  éléments ou clones présents à la génération 0 ;

- à la génération -1 nous retrouvons les ancêtres de tous les  $n$  éléments parmi seulement  $n(1-p)$  éléments, puisque, lors de la production de la génération 0,  $np$  des éléments de la génération -1 n'arrivaient pas à se reproduire, et ceci par le simple hasard; ainsi:

- à la génération -i nous retrouvons l'ensemble des ancêtres parmi seulement  $n(1-p)^i$  éléments; et

- à partir de la génération -a ce n'est qu'un seul élément qui est l'ancêtre commun de tous les éléments de la génération 0.

$$\begin{aligned} \text{Alors} \quad n(1-p)^a &= 1; \\ \Rightarrow \quad a &= t_{\text{fix}} = \log(1/n) : \log(1-p). \end{aligned}$$

Une estimation de ce que cela signifie pour les intervalles caractéristiques de la fixation des allèles de l'ADN chloroplastique aux différents niveaux d'organisation est portée sur le tableau 11.  $t_{\text{gen}}$  désigne l'intervalle de temps auquel correspond une génération au niveau en question, par exemple  $t_{\text{gen}} = c$  pour les niveaux cellulaires et subcellulaires, avec  $c$  égal à la durée du cycle cellulaire.

tab.11: intervalles nécessaires pour la fixation des allèles de l'ADNcp aux différents niveaux par dérive stochastique neutre (surestimations)

niveau	n	f	p	$t_{\text{gen}}$	$t_{\text{fix}}$
nucléide	10	2	0,2	1 c	11 c
chloroplaste	20	2	0,2	1 c	14 c
cellule	50	2	0,2	1 c	18 c
clone cellulaire		(automatiquement)			0
plante *	300 *	?	0,05	1 c	112 c
population	$10^4$	20	0,4	1 an	18 ans
population	$10^5$	20	0,3	1 an	33 ans
population	$10^6$	20	0,2	1 an	62 ans
méta-population	$10^3$	1,...	0.001	1 an	6.900 ans

\* pour une plante issue de reproduction végétative, le nombre maximal possible de clones cellulaires est égal au nombre de cellules de son méristème d'origine; une plante issue d'une graine ne représente qu'un seul clone cellulaire (tout cela en absence de mutations);

Retenons de ces estimations grossières et superficielles que la fixation ou non d'un allèle mutant de l'ADNcp se décide sans problème pendant une génération, c'est à dire 1 an, au niveau de la plante, ce qui explique bien la rareté de chimères cytoplasmiques. La fixation au niveau de la population se déroule à l'échelle de quelques dizaines d'années et éventuellement celle au niveau d'une méta-population, reposant sur un faible échange entre populations, dans à l'échelle du millénaire ou de la dizaine de milliers d'années.

Comme la présente analyse est basée sur l'observation d'à peu près 220 sites de restriction, correspondant en tout à environ 1.300 pb , et comme on doit chez les Graminées s'attendre à un taux de mutations de  $1.1 \times 10^{-9}$  par pb par an selon CURTIS et CLEGG (1984) ou jusqu'à  $1.6 \times 10^{-9}$  par pb par an selon WOLFE, LI et SHARP (1987), toujours en substitutions synonymes pour des gènes codants, tous les intervalles de fixation d'allèles calculés ci-dessus, même celui concernant les méta-populations, sont négligeables par rapport à l'intervalle de temps de divergence qui correspond à une mutation détectée: Un tel intervalle correspondrait à un demi million d'années ou plus pour des substitutions de bases (car, en moyenne, sous le taux susdit de  $1,1 \times 10^{-9}$  substitutions par site par an, un ensemble de 1.300 nucléotides accumule maximalelement  $1,3 \times 10^3 \times 1,1 \times 10^{-9} \times 10^{-6} = 1,43$  mutations par Ma; voir plus loin pour un calcul plus précis).

Donc, la probabilité de la présence simultanée de deux ou plusieurs plastotypes à l'intérieur d'une même communauté de reproduction doit être très faible, quel que soit son niveau, sauf si elle est maintenue par un mécanisme spécifique. Logiquement, notre arbre porte donc des branches d'une "épaisseur" remarquablement faible et ainsi devrait bien refléter la phylogénie des chloroplastes de riz, et par cela celle des riz, là où ceux-ci n'ont pas connus d'introgessions, voire d'échanges de cytoplasmes.

### c) Le calibrage de l'échelle de temps et le scénario d'évolution

A priori il y a deux approches de calibrage possible. La première serait d'appliquer sans hésiter les valeurs pour les taux de mutations de l'ADNcp cités ci-dessus, la deuxième de rechercher des points de repère dans le développement

connu de la paléo-géographie. Pour notre problème, cette dernière approche repose essentiellement sur l'hypothèse que toutes les migrations importantes des riz, notamment les migrations intercontinentales, ne se réalisaient que par la voie terrestre et à l'intérieur des zones climatiques adéquates, jusqu'à l'intervention des transports humains.

#### i) L'estimation du taux de mutations attendu

La première des approches susdites est compliquée par le fait que l'estimation à  $1,1 \times 10^{-9}$  par site par an du taux de substitutions de sites ne concerne que les sites pour lesquels on suppose une neutralité maximale: les sites dits synonymes qui sont neutres par rapport à la translation en acides aminés, et dont, pendant l'évolution, les modifications ne sont soumises à aucune contrainte sélective, tout du moins au niveau de la fonction des protéines encodées. Cette estimation est donc proche de la valeur maximale possible de ce que l'on peut attendre comme taux de substitution.

En ce qui concerne les sites ou séquences qui sont, par spécificité des nucléotides, liés à des fonctions soumises aux contraintes de la sélection, ceux-ci évoluent logiquement à un taux plus faible: une partie, en principe même la plupart, des mutations est constamment éliminée. Ainsi, le taux de substitutions chez les sites codants descend à une moyenne de

$$1/16 \times 1,1 \times 10^{-9} = 0,07 \times 10^{-9} \quad (\text{ou de } 1/16 \times 1,6 \times 10^{-9} = 0,1 \times 10^{-9}, \text{ respectivement) substitutions par site par an (spsa),}$$

d'après les observations de WOLFE et al. (1987). Pour les séquences codantes, on peut donc réestimer le taux de mutations global comme suit, en tenant compte de la dégénération du code génétique.

Puisqu'on sait que dans le code génétique, les 64 codons possibles n'ont que 24 significations distinctes, en ignorant les fréquences réelles des différentes combinaisons, on pourrait estimer la fréquence de synonymie " $f_s$ " chez les séquences codantes à environ

$$f_s = (64 - 24) : 64 = 0,6... ;$$

appelons  $f_s$  l'"indice de synonymie";  $f_s$  est égal à 1 pour les sites synonymes et les séquences neutres, et égal à 0 pour les sites et les séquences sans synonymie.

Une estimation plus réaliste peut être déduite de la comparaison des séquences de six gènes chloroplastiques entre le blé et le maïs par WOLFE et al. (1987): ils observent un rapport de 934 sites synonymes pour 3.114 sites non-synonymes, ce qui donne un indice de synonymie de

$$f_S = 934 : ( 934 + 3.114 ) = 0,23 .$$

Appelons  $t_S$  le taux de substitution sur les sites synonymes,  $t_A$  le taux de mutations non-synonymes et  $t_C$  le taux global de substitution sur les séquences codantes. Ce dernier, d'après les références citées pour les Graminées, s'estime alors à environ (en spsa):

$$\begin{aligned} t_C &= f_S \times t_S + ( 1 - f_S ) \times t_A = 0,23 \times 1,1 \times 10^{-9} + 0,77 \times 0,07 \times 10^{-9} \\ &= 0,3 \times 10^{-9} . \end{aligned}$$

L'ADN chloroplastique est saturé en séquences codantes (cf. SHINOZAKI et al. 1986), et il semble raisonnable que la partie non-codante ne dépasse pas les 30 % du génome total.

Pour une séquence non-codante mais remplissant une fonction régulatrice importante, on ne peut pas exclure qu'elle arrive à un indice de synonymie (tel que défini ci-dessus) assez bas. Malheureusement, nous ne disposons pas d'informations sur le taux de mutations de la partie non-codante de l'ADNcp dans son intégralité. Cependant, d'après les données de WOLFE et al. (1987), la réduction du taux de substitution de bases dans les séquences non-codantes par rapport aux substitutions synonymes chez les séquences codantes paraît plutôt négligeable. Ceci pourrait bien correspondre à une surestimation pour les séquences non-codantes mais, nous considérerons donc que 30 % du génome chloroplastique est non codant et évolue par substitutions de base au taux  $t_S$ .

Cependant ces mêmes données révèlent une réduction globale du taux de mutation sur les IR d'un facteur 4 environ. Comme les IR constituent à peu près un tiers de la molécule de notre ADNcp, le taux total se voit baissé par un facteur

$$"r_{IR}" = 1 \times 2/3 + 1/4 \times 1/3 = 0,75 .$$

En nous basant sur ces indications nous estimons le taux de substitutions de sites (" $t_M$ ") attendu pour l'ADNcp total des Graminées à

$$\begin{aligned} t_M &= r_{IR} \times ( 0,3 \times t_S + 0,7 \times t_C ) = \\ &= 0,75 \times ( 0,3 \times 1,1 + 0,7 \times 0,3 ) \times 10^{-9} = \end{aligned}$$

$$= 0,75 \times (0,33 + 0,21) \times 10^{-9} = 0,4 \times 10^{-9} \text{ spsa .}$$

## ii) Calibrage du phylogramme par le taux de mutations attendu

La présence de différences de vitesses relatives de l'horloge moléculaire sur notre phylogramme est évidente: nous les avons justifiées en même temps que les branches de "vitesse" 0. Ces différences atteignent un facteur entre 3 et 4 sur les 4 branches les plus importantes du phylogramme (V-W-T, V-U, V-g, U-k). Notons cependant qu'elles peuvent être due en partie à une marche intrinsèque de l'horloge différente pour les différentes branches mais que dans tous les cas une variance statistique importante les affecte aussi.

Pour un premier test de l'hypothèse nulle selon laquelle les différences observées seraient uniquement dues au hasard, nous avons admis que la distribution du nombre total de mutations sur les branches principales des Groupes Sativa et Latifolia, respectivement, suivait une loi de Poisson de moyenne égale à la moyenne observée. La statistique ne permet de rejeter cette hypothèse qu'au seuil  $\alpha = 0,1$  pour le Groupe Sativa (la branche conduisant au complexe Japonica ne porte qu'une mutation), mais au seuil  $\alpha = 0,01$  pour le Groupe Latifolia avec les branches conduisant à O. punctata et O. australiensis ayant accumulé environ le double de mutations de celles conduisant aux complexes des génomes CC et CCDD, respectivement.

On peut noter que ces taux maximaux concernent les deux espèces qui sont particulièrement exposées au soleil tropical en zone aride, alors que les autres représentants du Groupe Latifolia sont plus souvent à l'ombre en zone humide. Ceci permet pour les premières de soupçonner une certaine influence des UV mutagènes. Mais peut-être pourrait-on également invoquer qu'elles sont annuelles alors que les espèces des deux branches "lentes" sont pérennes.

Les deux exceptions recevant donc une explication plausible, l'hypothèse conservatrice qui identifie la moyenne de la vitesse relative de notre horloge avec le taux de mutations attendu,  $t_M$ , paraît raisonnable et nous l'acceptons pour la suite.

Le fait que notre horloge (relative) soit déterminée sur la base de toutes les mutations observées et pas seulement des mutations de sites, demande cependant que soit apporté un facteur de correction. Comme facteur de correction nous allons

appliquer la fréquence des mutations de sites (en identifiant celles-ci avec les substitutions de bases) observées sur la totalité des mutations observées, et ceci en excluant les branches des deux premiers niveaux d'intégration, sur lesquelles les mutations de sites se trouvent sous-représentées. Le facteur de correction, "f<sub>c</sub>", s'estime ainsi à

$$f_c = 42 : (40 + 42) = 0.51 .$$

Notre calibrage définitif se présente donc ainsi: Une "unité de mutation" telle que nous l'observons correspond à 0,51 substitution de site sur l'ensemble des 1.300 sites observés. Donc, comme  $t_M = 0,4 \times 10^{-9}$  spsa, notre "unité de mutations", "U<sub>M</sub>", correspond à un intervalle de

$$U_M = 0,51 : (0,4 \times 10^{-9} \times 1.300) \text{ années} = 10^6 \text{ a} = \underline{1 \text{ Ma}} .$$

### iii) Comparaison des deux approches de calibrage

Dans la suite nous comparons les niveaux de certains branchements de notre phylogramme situés dans le temps par rapport au cadre du scénario d'évolution des riz de la section Eu-Oryza proposé par SECOND (1985a).

1<sup>ier</sup> repère: la divergence d'O. australiensis (génomome E) (V)\*:

SECOND (1985a) met en relation l'estimation du temps de divergence d'O. australiensis sur la base des données isozymiques avec l'époque de la collision entre les plaques continentales australasienne et sud-est asiatique, datée à -15 Ma, et calibre ainsi son horloge isozymique.

Ce même repère pourrait bien correspondre aussi à la séparation du plastotype g du reste du Groupe Latifolia, puisque nous nous retrouvons vers environ -12 Ma (selon WPGM) ou -18 Ma (selon UPGM, comp. tab. 11).

2<sup>ième</sup> repère: la séparation du génomome B (U):

Sur la même échelle de temps, le branchement conduisant au type k, marquant les diploïdes de l'espèce O. punctata et certains tétraploïdes d'O. minuta, se situerait entre -8 et -12 Ma.

---

\* les lettres capitales entre parenthèses reprennent le code pour les branchements définis antérieurement

L'isolement pourrait ici s'expliquer par l'apparition d'une barrière d'échange dans le paléo-environnement, cette fois-ci la descente dans les latitudes des climats tempérés et l'ouverture de la Mer Rouge entre l'Afrique et l'Eurasie.

3<sup>ème</sup> repère: la séparation des complexes "Indica" et "Japonica"

D'après le scénario d'évolution du Groupe *Sativa* proposé par SECOND (1985a), la séparation des complexes "Indica" et "Japonica" parmi les ancêtres d'*Oryza sativa* fut provoquée par la montée de la chaîne himalayenne en barrière à la migration il y a 2 ou 3 Ma.

Selon nos calculs actuels, notre bifurcation Q qui correspond à cette séparation tombe sur un niveau de 2,4 , 2,8 ou 3,7 Ma (comp. tab. 11).

Donc, le calibrage de notre horloge ADNcp par le taux de mutation attendu pour les Graminées s'avère remarquablement compatible avec ces trois autres calibrages indépendants:

- celui par le repère de la divergence du type g, spécifique à O. australiensis (génomme E), et de la collision du continent australien avec le subcontinent sud-est-asiatique,

- celui par le repère de la diversité du type k, spécifique au génome B, et de l'isolation du continent africain ainsi que

- celui par la diversité entre le type e et le type a1 et l'apparition de l'Himalaya comme barrière de migration.

Ce calibrage correspond également bien à la datation proposée par SECOND (1985c) pour la divergence d'O. glaberrima liée à la persistance d'une possibilité de migration entre l'Asie et l'Afrique plus tardive pour une espèce du Groupe Sativa adapté aux zones arides que pour une espèce du Groupe Latifolia, primitivement adapté à la forêt humide: selon nos calculs cette divergence date de 5 Ma, selon SECOND (1985c) d'environ 3 Ma.

Par contre, la divergence entre les plastotypes g (O. longistaminata), f (forme australienne d'O. rufipogon) et i (O. glaberrima et O. breviligulata) est, dans nos résultats, comparable et correspondrait à une séparation de l'ordre de 5 Ma. Ceci est en contradiction avec les estimations basées sur l'horloge isozymique. On peut voir ici la

probabilité d'une substitution nucléo-cytoplasmique entre l'ancêtre d'O. longistaminata et O. breviligulata\* d'une part et d'autre part entre une introduction d'O. rufipogon en Australie au cours de l'ère Tertiaire et une introduction beaucoup plus récente durant les glaciations, comme documenté par G. SECOND (comm. pers.). Un exemple d'une telle substitution cytoplasmique peut également être vu dans O. rufipogon d'Inde de l'Ouest que LOLO (1987) a trouvé indistinguable d'O. glaberrima au niveau isozymique mais qui présente un plastotype appartenant au complexe Japonica dans nos analyses (échantillons DN 9 et DN 13).

Notons qu'au total, dans le Groupe Sativa, au cours des substitutions nucléo-cytoplasmiques il y a un biais en faveur du cytoplasme le plus récemment parvenu d'Asie, lieu d'origine du genre Oryza. Ceci pourrait expliquer l'absence de branchement proche de la racine du système Eu-Oryza dans le Groupe Sativa qui présente cependant une distribution géographique comparable à celle du Groupe Latifolia.

Un même biais se retrouve au niveau des riz cultivés où le complexe des plastotypes japonica se retrouve fréquemment associé à un génotype nucléaire indica alors que la réciproque n'a pas été observée.

#### iv) Conséquences du calibrage

Dans la suite nous allons nous servir de notre calibrage pour essayer de situer d'autres branchements de notre arbre phylogénétique dans l'évolution par rapport au temps absolu.

#### Groupe Latifolia - Groupe Sativa (W):

La position de la racine du système Eu-Oryza entre le Groupe Sativa (génome A) et le Groupe Latifolia (génomes B à E) peut être considérée comme bien établie: ce n'est pas seulement notre estimation quantitative lors de la déduction du phylogramme du cladogramme qui le suggère; à part le regroupement évident des types appartenant à la section Eu-Oryza, il y a également l'observation de deux fragments dérivés spécifiques à l'ensemble du Groupe Latifolia parmi les résultats préliminaires de notre ana-

\* cette possibilité d'échange génétique "fossilisé" avait déjà été présumée par SECOND (1985c, voir Fig. 5)

lyse par Bsp X I au niveau de la tribu des Oryzées.

Selon nos estimations, la division de cette racine se serait passée environ 14 ou 17 Ma avant notre ère, sans que l'on puisse de manière évidente faire coïncider cette date avec un évènement du paléo-environnement.

types l-n - types o/p:

C'est la divergence de l'ensemble des plastotypes o/p, caractéristique des allo-tétraploïdes CCDD et correspondant, d'après notre calibrage, à un temps de divergence d'à peu près 6 Ma, qui suggère fortement qu'il s'agit du cytoplasme accompagnant le génome DD. Et c'est leur structure interne qui suggère son origine en Amérique. En effet, on devrait estimer à plusieurs Ma le temps nécessaire pour le développement d'une structure telle que la montre le phylogramme pour les espèces CCDD (selon nos calculs presque 3 Ma; comp. le niveau de la bifurcation P selon tab. 10). Toute importation du génome DD en Amérique postérieure à ce délai aurait déjà dû importer de la diversité avec elle. Or nous n'avons retrouvé aucun représentant de ces plastotypes en Asie.

Donc, si on ne veut pas supposer une accélération extraordinaire du taux de mutations depuis quelques centaines d'années, on arrive à exclure l'intervention de l'Homme lors de l'introduction de cette série de riz en Amérique. Même lors de l'allo-tétraploïdisation du génome hypothétique DD par le génome CC l'intervention de l'Homme ne paraît pas vraisemblable, car elle affecte plusieurs espèces (surtout O. latifolia et O. grandiglumis) en partie subdivisées et portant différents cytoplasmes.

Dans ce contexte, la description par BRÜCHER (1977) d'un riz présumé diploïde, trouvé dans le Chaco humide du Paraguay et appartenant à la série Latifolia mérite attention, puisqu'il pourrait bien s'agir d'un représentant du génome DD natif.

Si l'on conserve notre postulat selon lequel les migrations intercontinentales des riz ne se sont effectuées que par voie terrestre avant l'intervention de l'Homme, notons que cette datation à -6 Ma de la divergence entre le génome DD présumé américain et le génome CCDD asiatique implique une époque relativement récente pour la migration d'une espèce tropicale entre le Nouveau et l'Ancien Monde. Le postulat est-il erroné, y a-t-il eu possibilité de migration vers l'Amérique à la fin du Tertiaire pour des plantes tropicales ou encore, notre date de divergence est-elle sous estimée?

L'hypothèse que les cytoplasmes o et p correspondent au génome nucléaire DD semble d'autant plus raisonnable, que chez les autres allotétraploïdes analysés, appartenant au génome BBCC, l'identification du parent maternel ne pose aucun problème. Notons à ce sujet que les deux plastotypes correspondant aux génomes BB et CC se retrouvent dans l'espèce O. minuta qui a donc une origine diphylétiq. Cependant, les BBCC portant le cytoplasme CC montrent le type ! qui diffère du type central primitif, m1, par une petite insertion, ce qui indique d'ailleurs une bonne isolation génétique de ces tétraploïdes par rapport à leurs parents diploïdes, du moins du côté CC.

Si jusqu'ici les possibilités d'interprétation de notre phylogramme hypothétique sont restées simples, c'est grâce au fait que ses cinq branches principales coïncident précisément avec les cinq génomes nucléaires connus dans la section Eu-Oryza. De plus, le nombre de mutations impliquées dans les analyses sur ce niveau bien intégré semble suffisant pour permettre des estimations d'échelle du temps absolu.

Mais plus nous descendons vers les niveaux d'intégration bas, et plus les distances entre types ou groupes de types se rapprochent de l'unité de mutation, moins il devient raisonnable de comparer des taux de mutations: a priori, quand il n'y a plus de hiérarchie dans les relations cladistiques, il n'est plus possible de distinguer des mutations toutes récentes de celles qui datent peut-être de deux millions d'années ou plus.

Néanmoins, et pour finir, spéculons un peu sur quelques indices pour des implications éventuelles de nos informations sur l'ère la plus récente.

Parmi les 13 mutations figurant au premier niveau d'intégration, il y a 4 importantes insertions et délétions, dont on peut bien mettre en doute le caractère neutre; il y a d'autre part peu de substitutions de sites. Est-ce que ceci ne pourrait pas indiquer, qu'à court terme il y a un taux fort élevé de mutations fixées, pour la plupart faiblement délétères, qui, à long terme, sont éliminées par la sélection, celle-ci ne laissant qu'une trace filtrée de ce qui se passe à court terme? Une telle trace serait forcément constituée d'une bonne partie de mutations effectivement neutres, d'une partie plus faible de mutations avantageuses (ou épisodiquement avantageuses) et, bien sûr, encore d'une partie de mutations faiblement délétères.

ZURAWSKI et CLEGG (1987) proposent encore une autre explication possible à

un tel phénomène de décroissance du taux relatif des insertions / délétions par rapport aux substitutions de sites avec l'augmentation de la distance taxonomique: selon ces auteurs, il s'agirait d'un effet de saturation causé par de rapides reprises du même type d'insertions / délétions toujours aux mêmes endroits, associés à certaines petites séquences répétées dispersées. Mais, dans notre cas, le petit nombre de réversions ou coïncidences que nous avons observé ne nous paraît pas très suggestif dans ce sens.

La recombinaison sexuelle accidentelle de l'ADNcp peut-elle être un phénomène plus fréquent qu'on ne l'admet couramment, passant inaperçu d'abord par l'identité des plastotypes des parents et ensuite par un caractère délétère des recombinants (toujours haploïdes!) véritablement hétérogènes?

Est-ce peut-être une accumulation rapide de mutations dans les niches perturbées qui pourrait expliquer le fait que tous les variants dérivés du plastotype m1, le plastotype primitif central des espèces CC (O. officinalis en Asie, O. eichingeri en Afrique), se trouvent en milieu perturbé par l'Homme, alors que toutes les variétés situés dans leur milieu originel (la forêt humide en Asie) expriment le type primitif?

Ou encore, plus concrètement, les variants du plastotype japonica (e2 à e4), peuvent-ils être spécifiques des variétés cultivées, ou étaient-ils préexistants parmi les parents sauvages, comme il semble que ce soit le cas pour les quatre types du complexe indica (a1, a2, b, c)?

La différenciation du type h d'O. rufipogon américain peut-elle être aussi récente, que l'on puisse continuer à croire qu'il soit arrivé en Amérique par la main de l'Homme?

## CONCLUSION

L'analyse à échelle élargie du polymorphisme des longueurs des fragments de restriction de l'ADN chloroplastique dans la section Eu-Oryza du genre Oryza, nous a permis de vérifier la valeur phylogénétique de ce marqueur, en particulier pour la détection des relations évolutives parmi des éléments distingués par des divergences importantes. Tant la topologie du phylogramme que les estimations de temps de divergence renforce donc la robustesse du scénario d'évolution proposé par SECOND (1985a) sur la base de la diversité isozymique. En définitive, dans l'état d'imprécision de la datation moléculaire, ce sont donc les connaissances sur le paléo-environnement qui peuvent permettre de préciser les époques de divergence vers lesquelles pointe l'horloge moléculaire.

Comme on sait, les allèles cytoplasmiques déterminent ou codéterminent qualitativement et quantitativement nombre de caractères morphologiques, anatomiques et physiologiques des plantes, liés à leur croissance, à leur reproduction et à leur embryogénèse (cf. TSUNEWAKI, 1980). Vu cependant la rapidité et l'efficacité du processus d'autopurification par simple dérive stochastique des génomes cytoplasmiques dans les communautés de reproduction, voire dans les systèmes artificiels d'hybridation, la protection contre l'érosion des ressources génétiques cytoplasmiques demande une attention et des efforts particuliers constants.

De là s'explique un intérêt pratique immédiat du simple marquage de différents plastotypes, en particulier parmi les riz cultivés. Notre étude en a présenté déjà dix parmi ceux-ci. L'aptitude et la simplicité relative de la technique exposée devraient permettre d'en mettre en évidence davantage en même temps qu'elle pourrait faciliter la mise en oeuvre de techniques plus fines, telles que par exemple le séquençage direct.

## ANNEXE A

### Liste des plantes analysées dans la section Eu-Oryza

#### GRUPE SATIVA

*O. sativa* (classées sous les groupes 1 à 6 d'après GLASZMANN)

##### classe 1:

##### - testeurs indica

1)	108	Peh Kuh	a 1
2)	108	"	a 1
3)	108	"	a 1
4)	S 38	"	a 1
5)	S 38	"	a 1

##### - testeurs japonica

6)	521	Kisshin	e 1
7)	521	"	e 1
8)	521	"	e 1
9)	521	"	e 1
10)	521	"	e 1
11)	Balila	"	e 1

##### - autres

12)	S 2	Hong xie nuo	a 1
13)	S 5	Liuzhou bayoya zao	a 1
14)	S 6	Tau Reng Chau	a 1
15)	S 14	IR 28	a -
16)	S 18	Hu nan hsien	a 1
17)	S 24	L 78 - 9148	b
18)	S 29	MS 52-3	e 1
19)	S 30	H.105 (T 6-802)	a -
20)	S 31	Segadis	a 1
21)	S 32	Vary vato	e 1
22)	S 34	Sixa (T 5-337)	a 1
23)	S 37	D 52-37	a 1
24)	S 39	Chuh tu	a 1
25)	S 40	Taichung native 1	a 1
26)	23 717	Dao Pao	c
27)	47 520	cut 389	a 1

28)	Akai-Mai		a 1
classe 2:			
29)	S 1	Chinsurah Boro II	e 1
30)	S 1	"	e 1
31)	S 22	2 LS 102	a 1
32)	S 22	"	a 1
33)	S 23	3 LS 102-4	a 1
34)	S 23	"	a 1
35)	BS 20		a 1
36)	BS 20		a 1
37)	(N) 4 319	Chinsurah Boro II	e 1
38)	6 246	DA 28	e 1
39)	27 509	Baran Boro	e 1
40)	27 516	Jagli Boro	e 1
41)	32 301	Gerdeh	e 3
classe 3:			
42)	6 541	Bhadoia 233	e 1
classe 4:			
43)	27 588	Rayada	e 1
44)	27 591	"	e 1
classe 5:			
45)	33 191	Kaukkyi Meedon	e 1
46)	58 881	Basmati Lamo	e 1
47)	58 881	"	e 1
48)	72 008	Sadri	e 1
classe 6:			
49)	S 4	Sunbonnet	e 1
50)	S 8	Jaguary	e 1
51)	S 9	Nan Ching Shiang Tao	e 1
52)	S 10	Fujisaka 5	e 1
53)	S 12	Canabong bong	e 1
54)	S 13	Texas Fortuna	e 1
55)	S 15	Vista	e 1
56)	S 16	Siampang	e 1
57)	S 19	M 2.2 (T 2.325)	e 1
58)	S 26	Iguape Cateto	e 1
59)	S 27	Cigalon	e 1
60)	S 35	IRAT 13	e 1
61)	S 35	" (à sec)	e 1

62)	S 36	Moroberekan	e 1
63)	S 41	Taichung 65	e 1
64)	AS 37	Latsika	e 2
65)	YS 309		e 1
66)	23 729	Hawm Om	e 4
67)	23 744	Khao Nuan	e 1
68)	Basmati		e 1
69)	Cigalon		e 1
70)	Hokkai 241		e 1
71)	Silewah		e 1

## intermédiaire 1/6:

72)	S 7	Basilanon	e 1
-----	-----	-----------	-----

## encore incertains:

73)	S 11	Khao Nok	e 1
74)	S 20	Mack Fay Deng	e 1
75)	S 21	500 A (M 2.I)	a -
76)	32 306	Gharib	e 3

## cytoplasmes stériles

77)	Sairyukoku A		a -
78)	103 435	chen-shan 97 A	a 1
79)	103 804	Er-chiu-nan A	a 2
80)	103 806	V 20 A	a 1
81)	103 806	V 20 A	a 2
82)	103 806	V 20 A	a 2
83)	103 871	Gang-yi-ya-ai-zao A	a 2
84)	103 919	V 41 A	a 1+2
85)	104 052	Lian-tang zao A	a 1
86)	104 054	Er-jiu-ai A	a 2
87)	IR 46 829 A		a 2
88)	IR 46 830 A		a 2
89)	IR 46 830 A		a 2
90)	IR 54 752 A		a 2

## adventice

91)	EX 56	Pururé	e 1
92)	EX 56		e 1

*O. glaberrima*

93)	G 1	Sénégal	i 1
-----	-----	---------	-----

94)	GG 1 = G 1	"	i 1
95)	G 1	"	i 1
96)	G 2	"	i 1
97)	G 2	"	i 1
98)	G 3	Mali	i 1
99)	G 6	Mutant alutineux	i 1
100)	G 8	Côte d'Ivoire	i 1
101)	G 8	"	i 1
102)	Glab Haïti	Haïti	i 1
103)	CGI x DR 38 (hybr. glab x ruf)		i 1

*O. rufipogon*

## Inde

104)	W 133	e 1 ?
105)	W 135	a 2
106)	W 135	a 2
107)	W 135	a i
108)	W 1534	a 2
109)	W 1669	a 2
110)	W 1669	a 2
111)	W 1681	a 2
112)	W 1681	a 2
113)	W 1681	a 2
114)	W 1686	a 2
115)	DN 9	e 5
116)	DN 13	e 5
117)	DN 39	e 1 ?
118)	DN 41	?
119)	DR 38	e 1 ?

## SE Asie

120)	W 555	Ceylon	a -
121)	W 574	Malaisie	a -
122)	W 593	"	c
123)	W 593	"	c
124)	W 593	"	c
125)	W 593	"	c
126)	W 1692	Thaïlande	b
127)	W 1699	"	b

## Chine

128)	W 1652	e 1
129)	W 1653	a 2

130)	W 1654	e 1
131)	W 1655	e 1
132)	W 1663	e 1
133)	W 1827	e 1
134)	W 1939	a 1
135)	W 1944	e 1
136)	W 1945	e 1
137)	W 1949	a 1
138)	W 1949	a 1

## Australie

139)	W 1627	f 3
140)	W 1627	f 3
141)	W 1629	f 1
142)	W 1629	f 1
143)	W 1629	f 1
144)	OR 7	f 1
145)	OR 10	f 1
146)	OR 31	f 1
147)	OR 39	f 1
148)	OR 41	f 1
149)	OR 42 (60°)	f 1
150)	OR 51-2	f 2
151)	OR 54-3	f 1

## Amérique

152)	W 612	e 1
153)	W 612	e 1
154)	W 1185	h 1
155)	W 1185	h 1
156)	W 1185 x W 708 hybride	h 1
157)	W 1185 x W 708 "	h 1
158)	100 970 = W 1185	h 1
159)	W 1187	d
160)	W 1191	h 2

*O. longistaminata* (obk = hybride naturel *O. longist.* x *O. sativa*)

161)	AL 2	Madagascar	g
162)	BL 25-4	Guinée Bissau	g
163)	CL 7-2	obk Sénégal	g
164)	EL 15-9	Tanzanie	g
165)	EL 16	"	g
166)	EL 16	"	g

167)	EL 89-12	obk	"	g
168)	UL 12		Cameroun	g
169)	UL 12-2		"	g
170)	UL 12-6		"	g
171)	UL 12-6		"	g
172)	UL 16-1		"	g
173)	UL 16-6	obk	"	g
174)	UL 16-6	obk	" (60°)	g
175)	UL 16-a12	obk	"	g
176)	UL 12B		"	g
177)	WL 02		Botswana	g
178)	ZL 14		Zambie	g

*O. breviligulata*

179)	B 5	= TB 82	Tchad	i 1
180)	B 5	"	"	i 1
181)	B 5	"	"	i 1
182)	B 7	= UB 35	Cameroun	i 1
183)	B 7	"	"	i 1
184)	B 9	= WB 01	Botswana	i 2
185)	B 9	"	"	i 2
186)	B 10	= IB 03	Côte d'Ivoire	i 1
187)	B 10	"	"	i 1
188)	B 11	= NB 02	Niger	i 1
189)	B 13	= SB 304	Sénégal	i 1
190)	B 14	= SB 321	"	i 1
191)	B 14	"	"	i 1
192)	B 15	= ILB 05	Côte d'Ivoire	i 1
193)	ZB 11		Zambie	i 1 ?

## GROUPE LATIFOLIA

## généme BB

*O. punctata*

194)	W 1515	Tanzania	k
195)	W 1515	"	k
196)	W 1515	"	k
197)	W 1515	"	k
198)	W 1515	" (40°)	k
199)	W 1515	" (40°)	k
200)	W 1515	"	k
201)	W 1586	Tchad	k
202)	W 1590	Cameroun	k
203)	W 1592	"	?
204)	W 1593	Nigeria	k
205)	TP 43	Tchad	k
206)	TP 43	"	k

## généme BBCC

*O. punctata*

207)	W 15		l
208)	W 43		l
209)	W 1023	Ghana	l
210)	W 1024	"	l
211)	W 1145		l
212)	W 1408	Nigeria	l
213)	W 1408	"	l
214)	W 1408	"	l
215)	W 1408	" (40°)	l
216)	W 1409	"	l
217)	IP 24	Côte d'Ivoire	l ou m-
218)	IP 27	"	m l
219)	IP 27	"	m l

*O. minuta*

## minuta 1

220)	103 865	Philippines	k
221)	103 881	"	k

## minuta 2

222)	W 45		l
------	------	--	---

223)	W 1213	Philippines	1
224)	W 1331	"	1
225)	W 1331	"	1
226)	W 1344	"	1

*O. malampuzhaensis*

227)	W 1159	Inde	1
228)	W 1159	"	1

*O. officinalis*

229)	W 2	Thaïlande	(BBCC)	1	
230)	W 611	Burma	(BBCC)	1	il pourrait
231)	W 1269	Brunei	(BBCC)	1	s'agir
232)	W 1343	Philippines	(BBCC)	1	d'erreurs
233)	W 1361	Borneo	(BBCC)	1	de collection
234)	W 1361	" (40*)	"	1	
235)	W 1361	"	"	1	

## génomme CC

*O. officinalis*

236)	100 180	Malaisie		m -
237)	100 181	Burma		m -
238)	101 121	Philippines		m -
239)	101 399	Vietnam		m -
240)	102 384	Indonesie		m -
241)	104 314	Proc		m -
242)	CH 83-3	Chine		m -
243)	D 04	Inde	(forêt)	m 1
244)	D 04	"		m 1
245)	W 6	Sri Lanka	(forêt)	m -
246)	W 12	Inde		m -
247)	W 65	Thaïlande	(forêt)	m 1
248)	W 66	"	(forêt)	m 1
249)	W 67	"	(forêt)	m -
250)	W 565	Malaisie		m 3
251)	W 614	Burma	(forêt)	m 1
252)	W 1131	Inde	(forêt)	m 1
253)	W 1149			m -
254)	W 1198	Philippines	(voisin rizière)	m -

255)	W 1199	"	"	m -
256)	W 1203	"	"	m 4
257)	W 1205	"	"	m 4
258)	W 1240	Borneo		m 2
259)	W 1251	"	(mare)	m 2
260)	W 1252	"	(bord de route)	m 2
261)	W 1260	"	(voie ferrée)	m 2
262)	W 1268	Brunei	(mare bord de route)	m 2
263)	W 1271	"	(proche habitation)	m 2
264)	W 1273	"	(bord de route)	m 2
265)	W 1275	"	(canal irriguant)	m -
266)	W 1277	Sarawak	(bord de route)	m -
267)	W 1278	"	"	m 2
268)	W 1303	Philippines	(rizière)	m -
269)	W 1306	"	(canal irriguant)	m 1
270)	W 1306	"		m -
271)	W 1315	"	(forêt vierge)	m -
272)	W 1318	"		m 2
273)	W 1334	"	(sous cocotiers)	m 1
274)	W 1830	Chine		m -
275)	W 1830	"		m 1
276)	W 1831	Borneo		m -
277)	W 1930	Thaïlande	(verger près forêt)	m 1
278)	W 1934	"	"	m 1
279)	W 1941	"	"	m 1
280)	W 1942	"	"	m -
281)	W 1950	Chine		m 1
282)	W 1950	= 104 618		m 1

## O. eichingeri.

283)	IP 7	Côte d'Ivoire	n 3
284)	IP 7	"	n 3
285)	1 IP 2	"	n 2
286)	2 IP 2	"	n 2
287)	W 1519	Ouganda	n 1
288)	W 1525	"	n 1
289)	W 1526	"	n 1
290)	W 1526	"	n 1

*O. latifolia*

291)	W 46		o
292)	W 542	Mexico	p 1
293)	W 542	"	p 1
294)	W 553		p 1
295)	W 1168	Cuba	o
296)	W 1168	*	o
297)	W 1168	"	o
298)	W 1168	"	o
299)	W 1176	Guatemala	o
300)	W 1539	Panama	p 2
301)	100 165	Guatemala	o
302)	100 169		p 1
303)	100 172		p 1
304)	102 481	Nicaragua	o

*O. alta*

305)	W 17		p 2
306)	101 395		p 2

*O. grandiglumis*

307)	W 613	Brésil	p 3
308)	W 1247	"	p 3
309)	W 1483	"	p 3

*O. paraguayensis*

310)	W 1144		p 2
311)	W 1144		p -

## généme EE

*O. australiensis*

312)	W 0008		q 1
313)	W 0008		q 1
314)	W 1562		q 2
315)	W 1562		q 2
316)	W 1628		q 1
317)	W 1639		q 1
318)	W 1639		q 1
319)	W 1639		q 1
320)	QA 27		q 2

## ANNEXE B

Liste des plastotypes déterminés dans la section Eu-Oryza

### GROUPE SATIVA

a 1 :

*O. sativa*

classe 1:

- testeurs indica

1)	108	Peh Kuh
2)	108	"
3)	108	"
4)	S 38	"
5)	S 38	"

- autres

12)	S 2	Hong xie nuo
13)	S 5	Liuzhou bayoya zao
14)	S 6	Tau Reng Chau
16)	S 18	Hu nan hsien
20)	S 31	Segadis
22)	S 34	Sixa (T 5-337)
23)	S 37	D 52-37
24)	S 39	Chuh tu
25)	S 40	Taichung native 1
27)	47 520	cut 389
28)	Akai-Mai	

classe 2:

31)	S 22	2 LS 102
32)	S 22	"
33)	S 23	3 LS 102-4
34)	S 23	"
35)	BS 20	
36)	BS 20	

cytoplasmes stériles

78)	103 435	chen-shan 97 A
80)	103 806	V 20 A
85)	104 052	Lian-tang zao A

*O. rufipogon*

107)	W 135	Inde
134)	W 1939	Chine
137)	W 1949	"
138)	W 1949	"

a 1/2 :

## mâle-stérile

84)	103 919	V 41 A
-----	---------	--------

a 2 :

*O. sativa* - cytoplasmes stériles

79)	103 804	Er-chiu-nan A
81)	103 806	V 20 A
82)	103 806	V 20 A
83)	103 871	Gang-yi-ya-ai-zao A
86)	104 054	Er-jiu-ai A
87)	IR 46 829 A	
88)	IR 46 830 A	
89)	IR 46 830 A	
90)	IR 54 752 A	

*O. rufipogon*

105)	W 135	Inde
106)	W 135	"
108)	W 1534	"
109)	W 1669	"
110)	W 1669	"
111)	W 1681	"
112)	W 1681	"
113)	W 1681	"
114)	W 1686	"
129)	W 1653	Chine

a - :

*O. sativa*

15)	S 14	IR 28
19)	S 30	H.105 (T 6-802)

75) S 21 500 A (M 2.I)  
 77) Sairyukoku A (mâle-stérile)

*O. rufipogon*

120) W 555 Ceylon  
 121) W 574 Malaisie

b :

*O. sativa* , classe 1

17) S 24 L 78 - 9148

*O. rufipogon*

126) W 1692 Thaïlande  
 127) W 1699 "

c :

*O. sativa indica* :

26) 23 717 Dao Pao

*O. rufipogon*

122) W 593 Malaisie  
 123) W 593  
 124) W 593  
 125) W 593

d :

*O. rufipogon*

159) W 1187 Amérique

e 1 :

*O. sativa*- testeurs *japonica*

6)	521	Kisshin	classe 1
7)	521	"	"
8)	521	"	"
9)	521	"	"

10)	521	"	"
11)	Balila	"	"
- autres			
18)	S 29	MS 52-3	classe 1
21)	S 32	Vary vato	"
29)	S 1	Chinsurah Boro II	"
30)	S 1	"	"
37)	(N) 4 319	Chinsurah Boro II	"
38)	6 246	DA 28	"
39)	27 509	Baran Boro	"
40)	27 516	Jagli Boro	"
42)	6 541	Bhadooia 233	classe 3
43)	27 588	Rayada	classe 4
44)	27 591	"	"
45)	33 191	Kaukkyi Meedon	classe 5
46)	58 881	Basmati Lamo	"
47)	58 881	"	"
48)	72 008	Sadri	"
49)	S 4	Sunbonnet	classe 6
50)	S 8	Jaguary	"
51)	S 9	Nan Ching Shiang Tao	"
52)	S 10	Fujisaka 5	"
53)	S 12	Canabong bong	"
54)	S 13	Texas Fortuna	"
55)	S 15	Vista	"
56)	S 16	Siampang	"
57)	S 19	M 2.2 (T 2.325)	"
58)	S 26	Iguape Cateto	"
59)	S 27	Cigalon	"
60)	S 35	IRAT 13	"
61)	S 35	" (à sec)	"
62)	S 36	Moroberekan	"
63)	S 41	Taichung 65	"
65)	YS 309		"
67)	23 744	Khao Nuan	"
68)	Basmati		"
69)	Cigalon		"
70)	Hokkai 241		"
71)	Silewah		"
72)	S 7	Basilanon	intermédiaire 1/6
73)	S 11	Khao Nok	
74)	S 20	Mack Fay Deng	

91) EX 56 Pururé  
92) EX 56 "

*O. rufipogon*

104) W 133 Inde (?)  
117) DN 39 " (?)  
119) DR 38 " (?)  
128) W 1652 Chine  
130) W 1654 "  
131) W 1655 "  
132) W 1663 "  
133) W 1827 "  
135) W 1944 "  
136) W 1945 "  
152) W 612 Amérique  
153) W 612 "

e 2 :

*O. sativa japonica*

51) AS 37 Latsika classe 6

e 3 :

*O. sativa japonica*

41) 32 301 Gerdeh classe 2  
76) 32 306 Gharib classe 6 (?)

e 4 :

*O. sativa japonica*

66) 23 729 Hawm Om classe 6

e 5 :

*O. rufipogon*

115) DN 9 Inde  
116) DN 13 "

f 1 :

*O. rufipogon*      Australie

139)	W 1627	
140)	W 1627	
141)	W 1629	
142)	W 1629	
143)	W 1629	
144)	OR 7	
145)	OR 10	
146)	OR 31	
147)	OR 39	
148)	OR 41	
149)	OR 42	(60°)
151)	OR 54-3	

f 2 :

*O. rufipogon*      Australie

150)	OR 51-2	
------	---------	--

g :

*O. longistaminata*      (obk = hybride naturel *O. longist.* x *O. sativa*)

161)	AL 2		Madagascar
162)	BL 25-4		Guinée Bissau
163)	CL 7-2	obk	Sénégal
164)	EL 15-9		Tanzanie
165)	EL 16		"
166)	EL 16		"
167)	EL 89-12	obk	"
168)	UL 12		Cameroun
169)	UL 12-2		"
170)	UL 12-6		"
171)	UL 12-6		"
172)	UL 16-1		"
173)	UL 16-6	obk	"
174)	UL 16-6	obk	"      (60°)
175)	UL 16-a12	obk	"
176)	UL 128		"
177)	WL 02		Botswana
178)	ZL 14		Zambie

h 1 :

*O. rufipogon* Amérique  
 154) W 1185  
 155) W 1185  
 156) W 1185 x W 708 hybride  
 157) W 1185 x W 708 "  
 158) 100 970 = W 1185

h 2 :

*O. rufipogon* Amérique  
 160) W 1191

i i :

*O. glaberrima*

93) G 1 Sénégal  
 94) GG 1 = G 1 "  
 95) G 1 "  
 96) G 2 "  
 97) G 2 "  
 98) G 3 Mali  
 99) G 6 Mutant alutineux  
 100) G 8 Côte d'Ivoire  
 101) G 8 "  
 102) Glab Haïti  
 103) CGI x DR 38 (hybr. glab x ruf)

*O. breviligulata*

179) B 5 = TB 82 Tchad  
 180) B 5 " "  
 181) B 5 " "  
 182) B 7 = UB 35 Cameroun  
 183) B 7 " "  
 186) B 10 = IB 03 Côte d'Ivoire  
 187) B 10 " "  
 188) B 11 = NB 02 Niger  
 189) B 13 = SB 304 Sénégal  
 190) B 14 = SB 321 "  
 191) B 14 " "  
 192) B 15 = ILB 05 Côte d'Ivoire  
 193) ZB 11 Zambie (?)

i 2 :

*O. breviligulata*

184)	B 9	=	WB 01	Botswana
185)	B 9		"	"

## GROUPE LATIFOLIA

k :

*O. punctata* BB

194)	W 1515	Tanzania	
195)	W 1515	"	
196)	W 1515	"	
197)	W 1515	"	
198)	W 1515	"	(40°)
199)	W 1515	"	(40°)
200)	W 1515	"	
201)	W 1586	Tchad	
202)	W 1590	Cameroun	
203)	W 1592	"	
204)	W 1593	Nigeria	
205)	TP 43	Tchad	
206)	TP 43	"	

*O. minuta* 1

220)	103 865	Philippines
221)	103 881	"

l :

*punctata* BBCC

207)	W 15	
208)	W 43	
209)	W 1023	Ghana
210)	W 1024	"
211)	W 1145	
212)	W 1408	Nigeria

213)	W 1408	"	
214)	W 1408	"	
215)	W 1408	"	(40°)
216)	W 1409	"	

*O. minuta* 2

222)	W 45		
223)	W 1213	Philippines	
224)	W 1331	"	
225)	W 1331	"	
226)	W 1344	"	

*O. malampuzhaensis*

227)	W 1159	Inde	
228)	W 1159	"	

*O. officinalis* BBCC

229)	W 2	Thaïlande	(BBCC)	
230)	W 611	Burma	(BBCC)	il pourrait
231)	W 1269	Brunei	(BBCC)	s'agir
232)	W 1343	Philippines	(BBCC)	d'erreurs
233)	W 1361	Borneo	(BBCC)	de collection
234)	W 1361	"	(40°)	"
235)	W 1361	"	"	"

## \* 1 :

*O. punctata* BBCC

218)	IP 27	Côte d'Ivoire	
219)	IP 27	"	

*O. officinalis*

243)	D 04	Inde	(forêt)
244)	D 04	"	
247)	W 65	Thaïlande	(forêt)
248)	W 66	"	(forêt)
251)	W 614	Burma	(forêt)
252)	W 1131	Inde	(forêt)
269)	W 1306	Philippines	(canal irriguant)
273)	W 1334	"	(sous cocotiers)
275)	W 1830		
277)	W 1930	Thaïlande	(verger près forêt)
278)	W 1934	"	"

279)	W 1941	"	"
281)	W 1950	Chine	
282)	W 1950	= 104 618	

## E 2 :

*O. officinalis* CC<sub>2</sub>

258)	W 1240	Borneo	
259)	W 1251	"	(mare)
260)	W 1252	"	(bord de route)
261)	W 1260	"	(voie ferrée)
262)	W 1268	Brunei	(mare bord de route)
263)	W 1271	"	(proche habitation)
264)	W 1273	"	(bord de route)
267)	W 1278	Sarawak	(bord de route)
272)	W 1318		

## E 3 :

*O. officinalis*

250)	W 565	Malaisie	
------	-------	----------	--

## E 4 :

*O. officinalis*

256)	W 1203	Philippines	(voisin rizière)
257)	W 1205	"	"

## I ou E - :

*O. punctata* BBCC

217)	IP 24	Côte d'Ivoire	
------	-------	---------------	--

*O. officinalis*

236)	100 180	Malaisie
237)	100 181	Burma
238)	101 121	Philippines
239)	101 399	Vietnam
240)	102 384	Indonesie
241)	104 314	Proc
242)	CH 83-3	Chine

317)	W 1639
318)	W 1639
319)	W 1639

q 2 :

*O. australiensis*

314)	W 1562
315)	W 1562
320)	OA 27



245)	W 6	Sri Lanka	(forêt)
246)	W 12	Inde	
249)	W 67	"	(forêt)
253)	W 1149		
254)	W 1198	Philippines	(voisin rizière)
255)	W 1199	"	"
265)	W 1275	Brunei	(canal irriguant)
266)	W 1277	Sarawak	(bord de route)
268)	W 1303	Philippines	(rizière)
270)	W 1306	"	(canal irriguant)
271)	W 1315	"	(forêt vierge)
274)	W 1830	Chine	
276)	W 1831	Borneo	
280)	W 1942	Thaïlande	(verger près forêt)

n 1 :

*O. eichingeri*

287)	W 1519	Ouganda
288)	W 1525	"
289)	W 1526	"
290)	W 1526	"

n 2 :

*O. eichingeri*

285)	1 IP 2	Côte d'Ivoire
286)	2 IP 2	"

n 3 :

*O. eichingeri*

283)	IP 7	Côte d'Ivoire
284)	IP 7	"

o :

*O. latifolia*

291)	W 46	
295)	W 1168	Cuba
296)	W 1168	"

297)	W 1168	*
298)	W 1168	"
299)	W 1176	Guatemala
301)	100 165	Guatemala
304)	102 481	Nicaragua

p 1 :

*O. latifolia*

292)	W 542	Mexico
293)	W 542	*
294)	W 553	
302)	100 169	
303)	100 172	

p 2 :

*O. alta*

305)	W 17	
306)	101 395	

*O. latifolia*

300)	W 1539	Panama
------	--------	--------

*O. paraguayensis*

310)	W 1144	
311)	W 1144	

p 3 :

*O. grandiglumis*

307)	W 613	Brézil
308)	W 1247	*
309)	W 1483	"

q 1 :

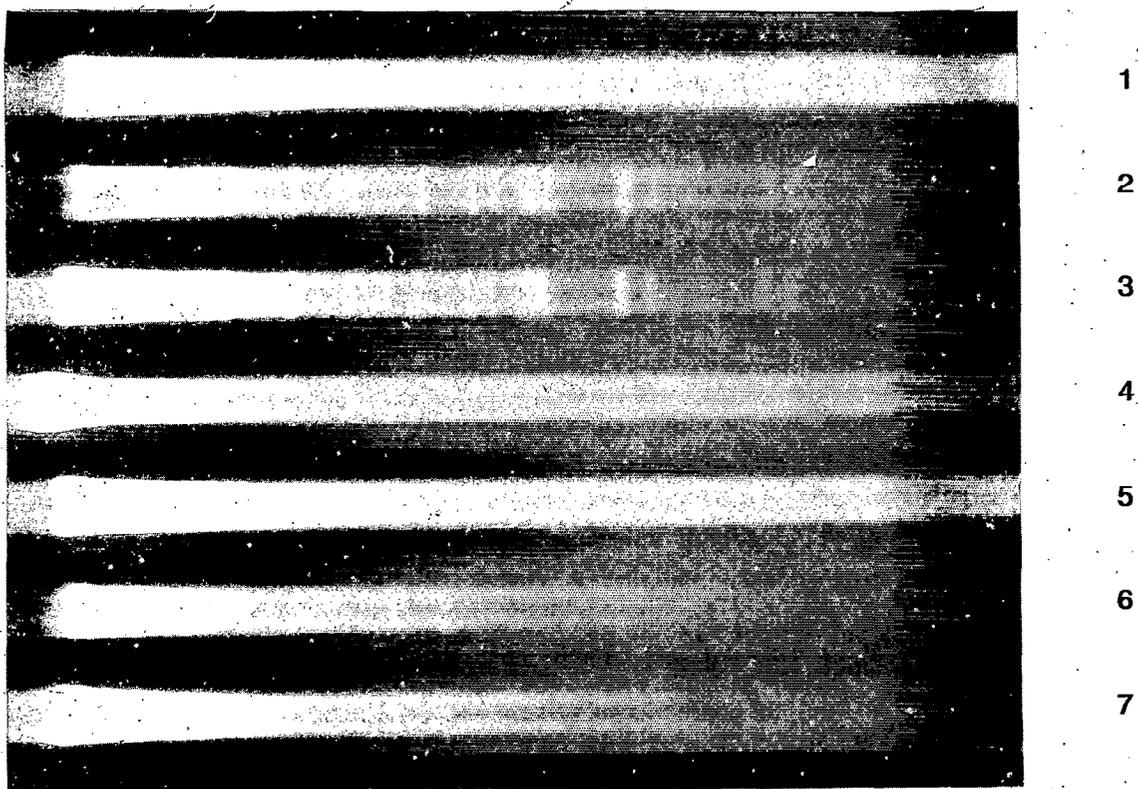
*O. australiensis*

312)	W 0008	
313)	W 0008	
316)	W 1628	

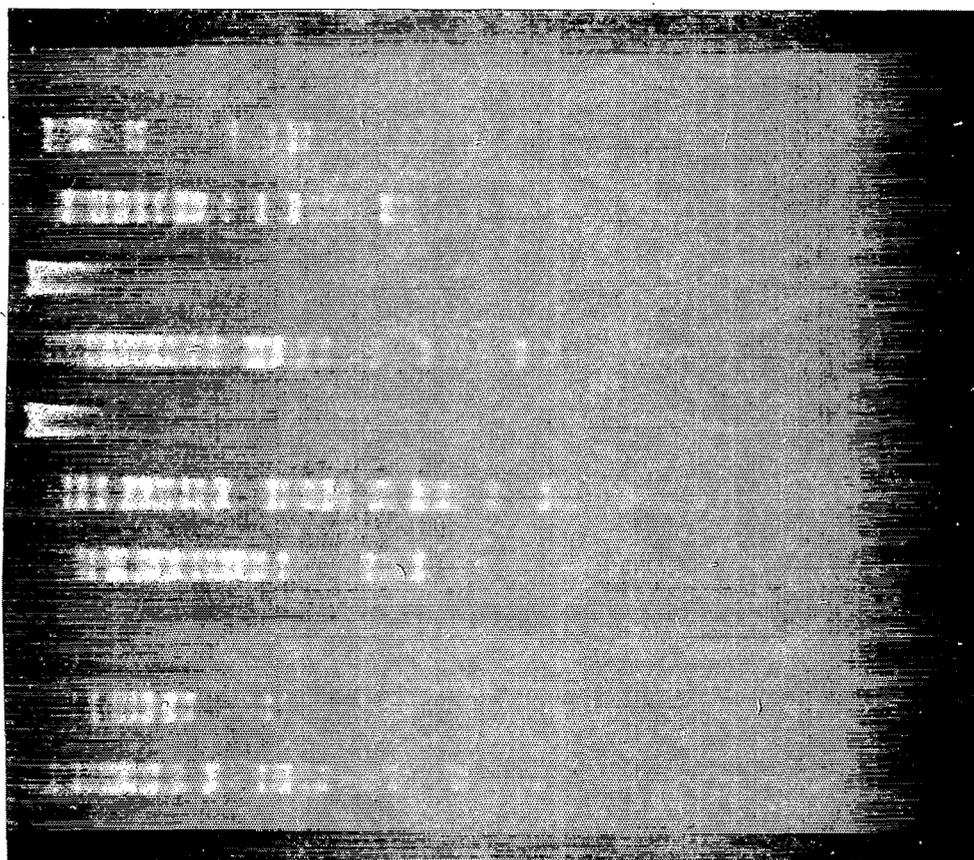
## ANNEXE C

### Résultats d'électrophorèse d'ADN chloroplastique

Fig. 1: ADNcp de riz extrait par la méthode de BOWMAN et DYER (1982)



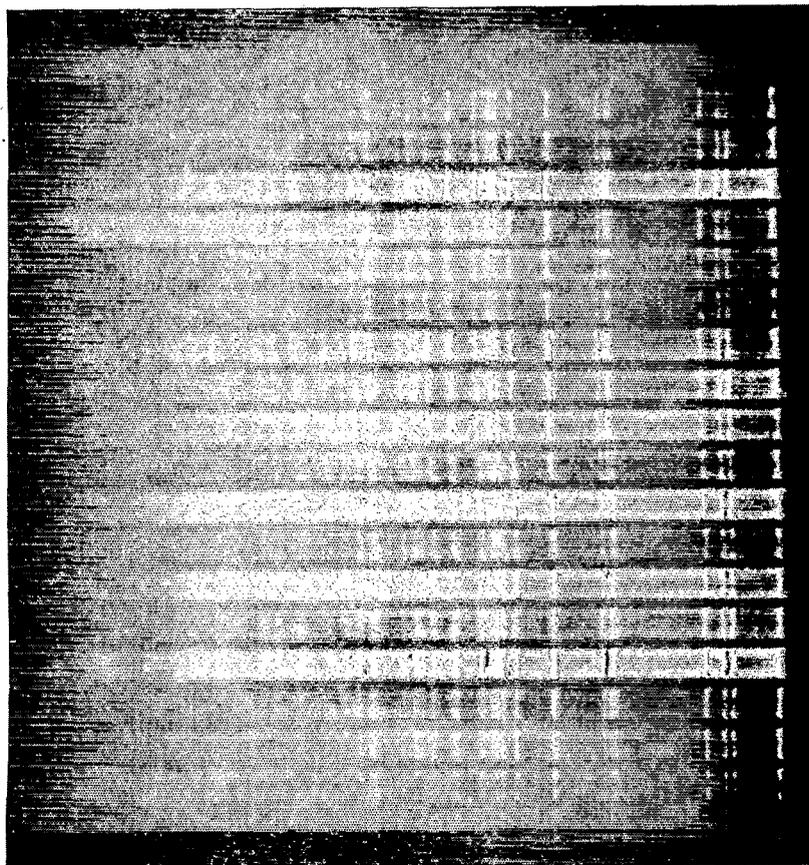
1: W 1652	} x Bam H I	5: W 1652	} x Pst I
2: TB 82		6: TB 82	
3: 521		7: 521	
4: W 1652 non digéré			

Fig. 11: ADNcp digéré par *Ban H I*

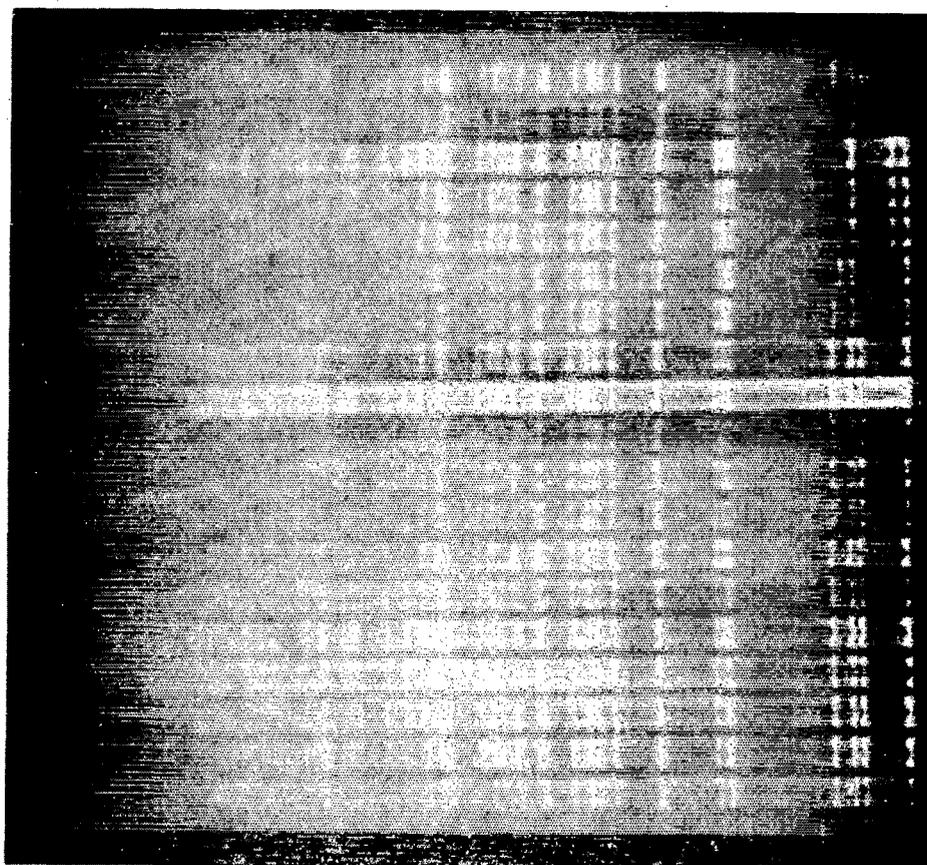
- |                              |                                    |
|------------------------------|------------------------------------|
| 1: <i>Thymus vulgaris</i>    | 7: <i>Asphodelus fistulosus</i>    |
| 2: <i>Gossypium australe</i> | 8: <i>Avena sterilis</i>           |
| 3: <i>Arum maculatum</i>     | 9: <i>Plantago lanceolata</i> *    |
| 4: <i>Quercus ilax</i>       | 10: <i>Scilla Lilio-Hyacinthus</i> |
| 5: <i>Centaurea collina</i>  | 11: <i>Eichhornia crassipes</i>    |
| 6: <i>Iris sp.</i>           |                                    |

\* Ici, l'ADN est dégradé par le broyage. N. Hochlander (dans le même laboratoire) obtient actuellement de l'ADNcp de cette espèce en bon état.

Fig. III A+B: ADNcp (Eu- *Oryza* ) x Eco R I en TAE



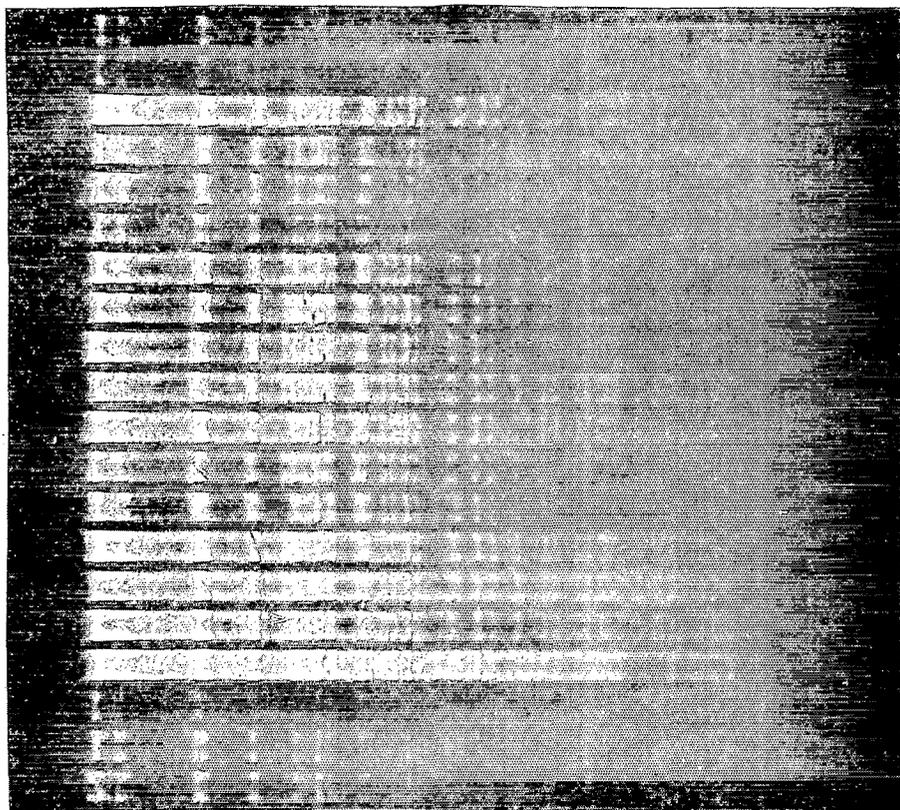
- S 38
- S 37
- S 27
- S 36
- S 35
- S 35
- S 34
- S 32
- S 31
- S 30
- S 29
- S 26
- S 24
- S 23
- S 22
- S 21
- S 20
- S 19
- S 18
- S 16



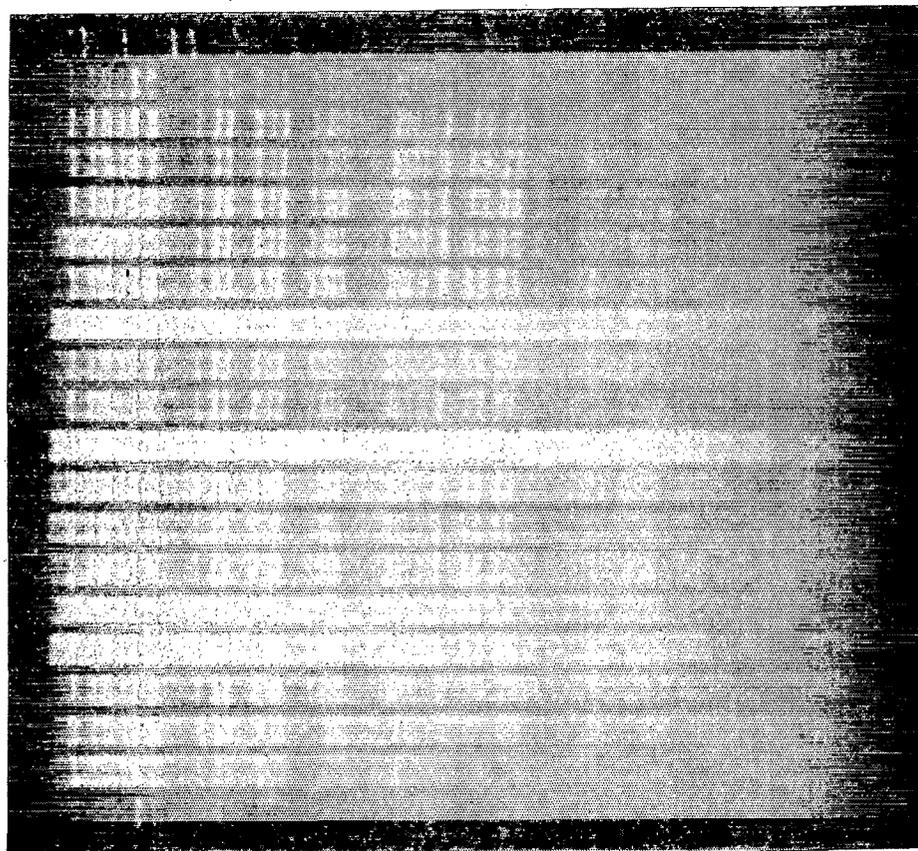
- BS 20
- B 13
- B 7
- G 6
- G 1
- W 1628
- W 1627
- Hokkai 241
- 4319
- W 1827
- EX 56
- 58 881
- S 4
- W 1949
- W 593
- IR 46 330
- S 37
- 108
- S 24
- (marqueur)

Fig. III C: patrons ADNcp x Eco R I en TBE

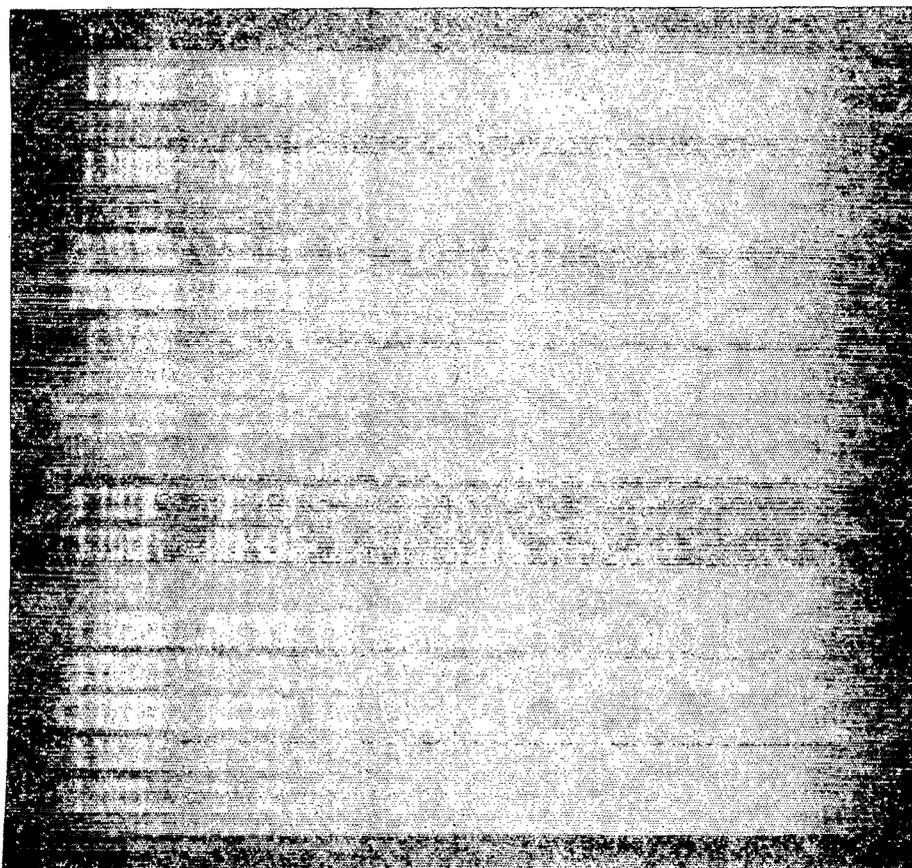
Fig. IV A: patrons Ava I



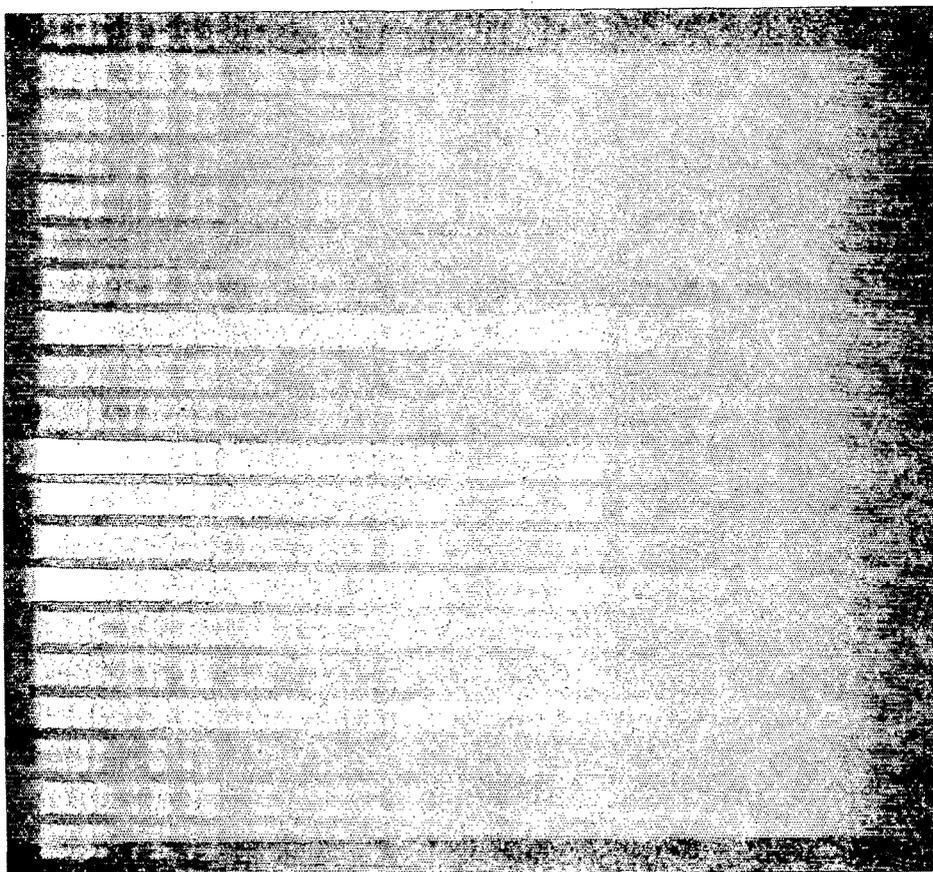
- W 2
- D 04
- W 1159
- W 1331
- W 1318
- W 1271
- W 1344
- W 1213
- W 45
- W 1409
- W 542
- IP 27
- W 1515
- W 1408
- W 1145
- W 15
- TP 43
- W 1593
- W 1592
- W 43



- X Ava I :
- (marqueur)
- G 8
- B 14
- UL 16-6
- W 1185
- W 1627
- W 1187
- S 35
- 521
- W 1163
- W 1652
- W 593
- 104 054
- 103 806 (n°81)
- S 40
- S 38
- W 1949
- W 1653
- S 24
- W 1699

Fig. IV B+C: patrons ADNcp x *Ava I*

2 IP 2  
 W 1361  
 W 1260  
 W 66  
 W 2  
 W 1181  
 W 1159  
 W 45  
 W 43  
 W 1639  
 103 881  
 W 1515  
 W 1168  
 W 1592  
 W 1144  
 101 395  
 W 1483  
 W 1539  
 W 542  
 S 10



W 593  
 W 593  
 W 1534  
 DN 13  
 DN 9  
 W 1686  
 W 1681  
 W 1669  
 W 135 (n=107)  
 EX 56  
 IR 54 752  
 IR 46 830  
 IR 46 830  
 IR 46 829  
 104 054  
 104 052  
 103 919 (cybr)  
 103 871  
 103 804  
 V 20 A (n=82)

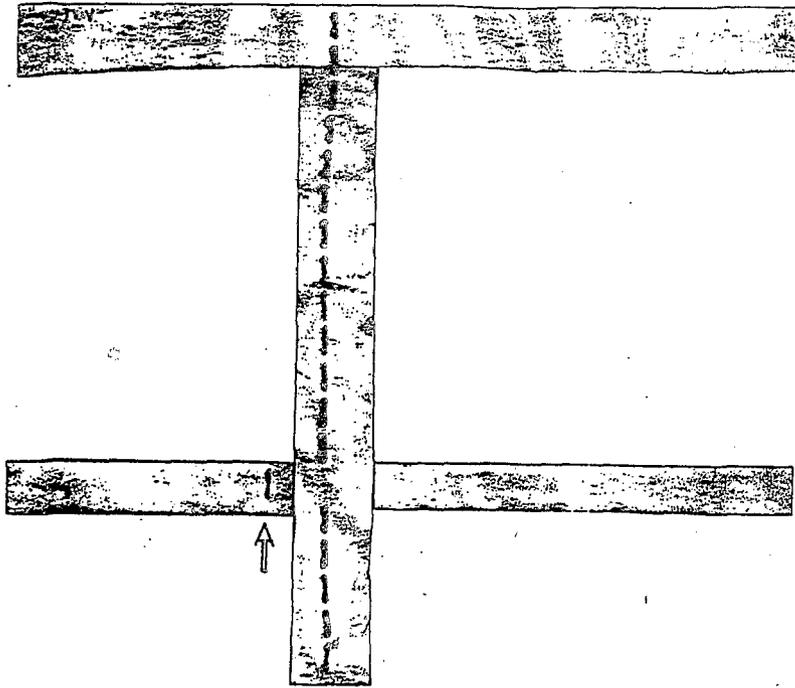
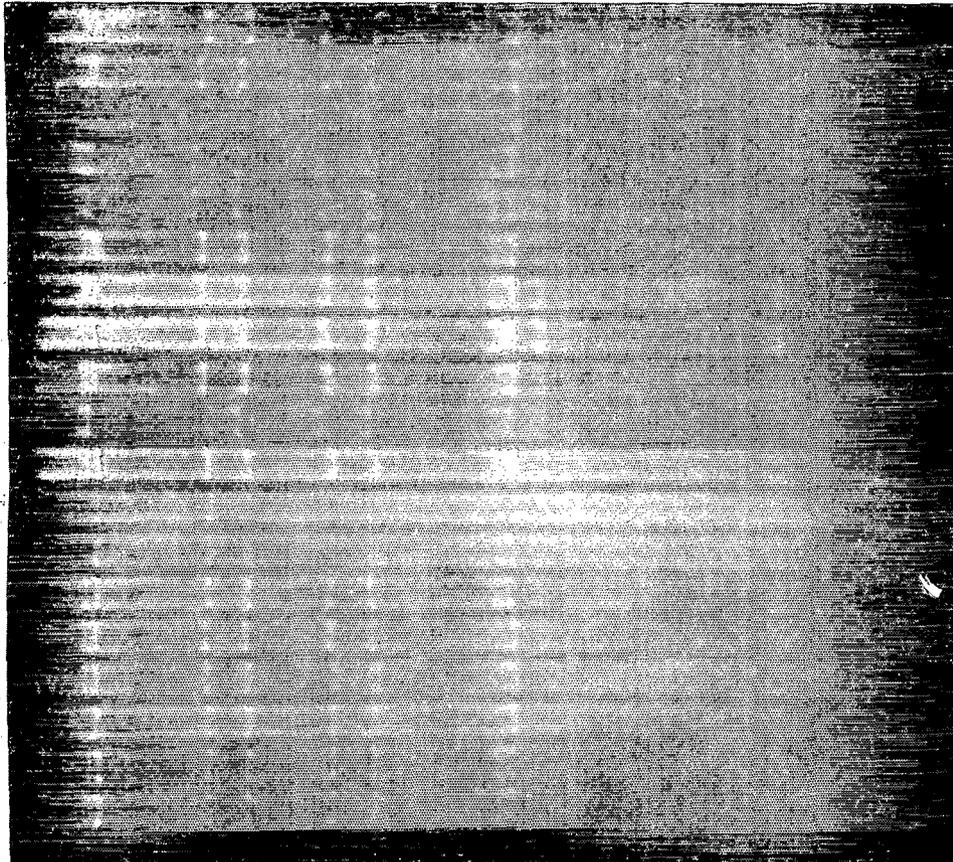


Fig. IV D:

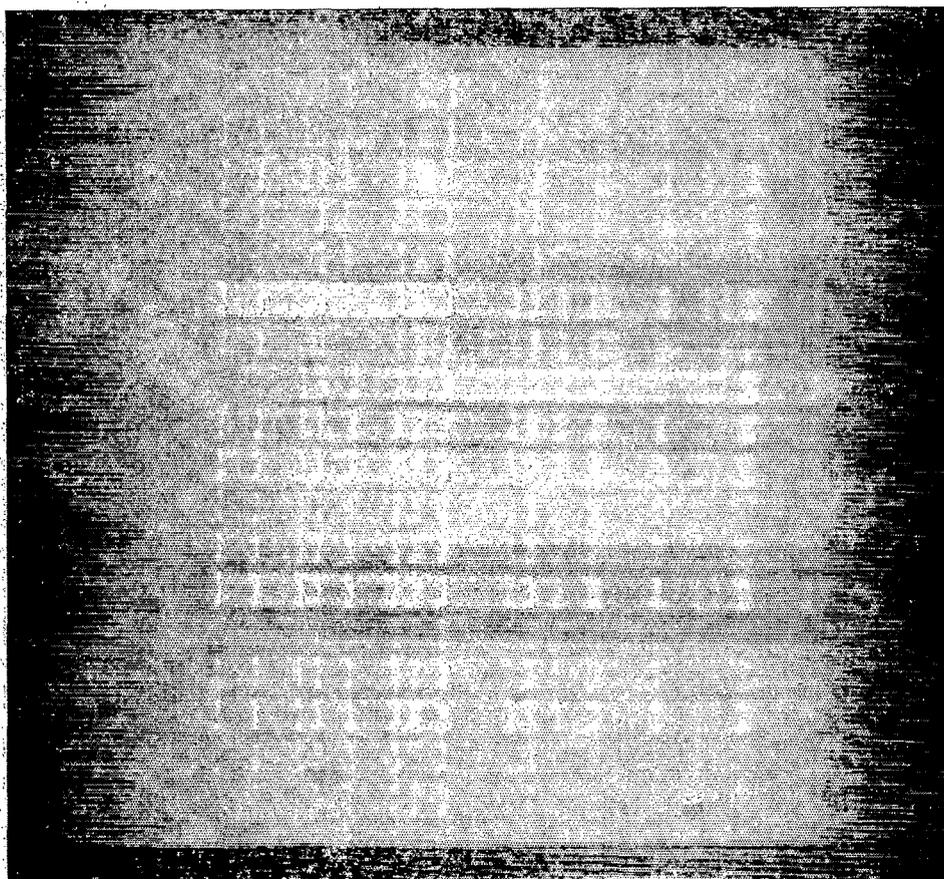
hybridation du gel *Ava I* de la Fig. IV A avec une sonde froide spécifique du gène de la cytochromeoxidase; différent: W 1187; (hybridation effectuée par ZHANG S.H.)

Fig. V: patrons de restriction ADNcp x *Bst E II*

plastotypes:

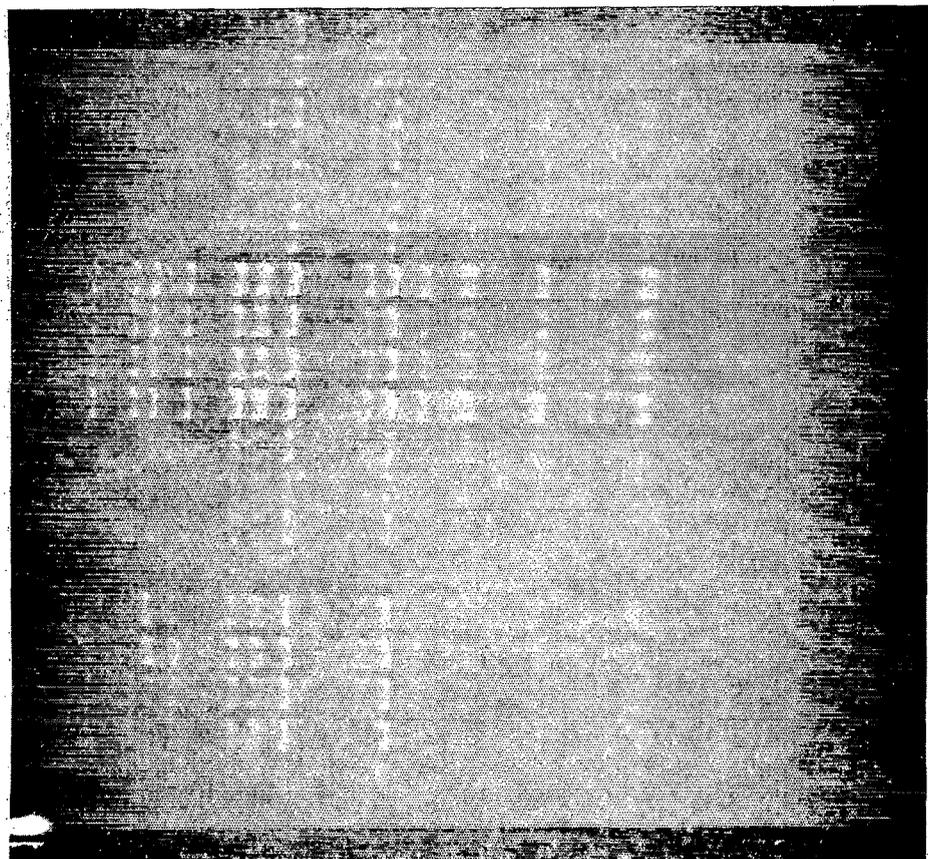
q 2  
n 2  
n 1  
m 1  
m 4  
m 2  
l  
o  
p 3  
k  
i 1  
h  
g  
f  
e 1  
c  
b  
a 2  
a 1

Fig VI A+B: patrons de restriction ADNcp x Bam H I



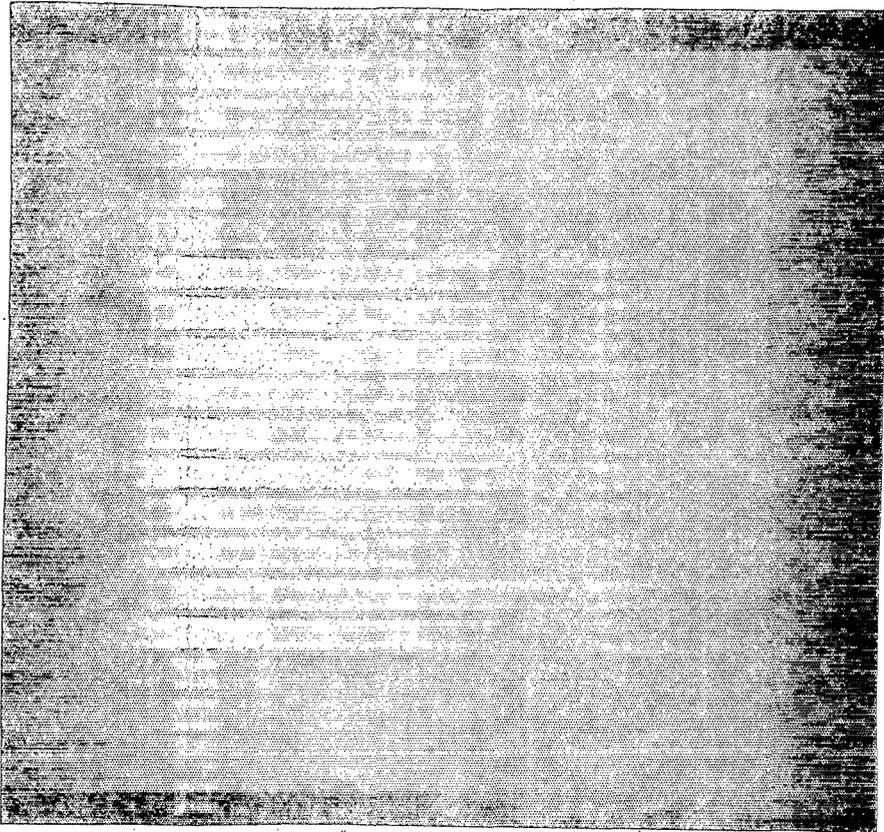
- (marqueur)
- 108
- S 11
- G 6
- W 1168
- W 1592
- W 542
- W 1515
- W 1408
- W 43
- W 1213
- W 1950
- W 1830
- W 1361
- W 1334
- W 1260
- W 66
- W 65
- W 2
- D 04

plastotypes:

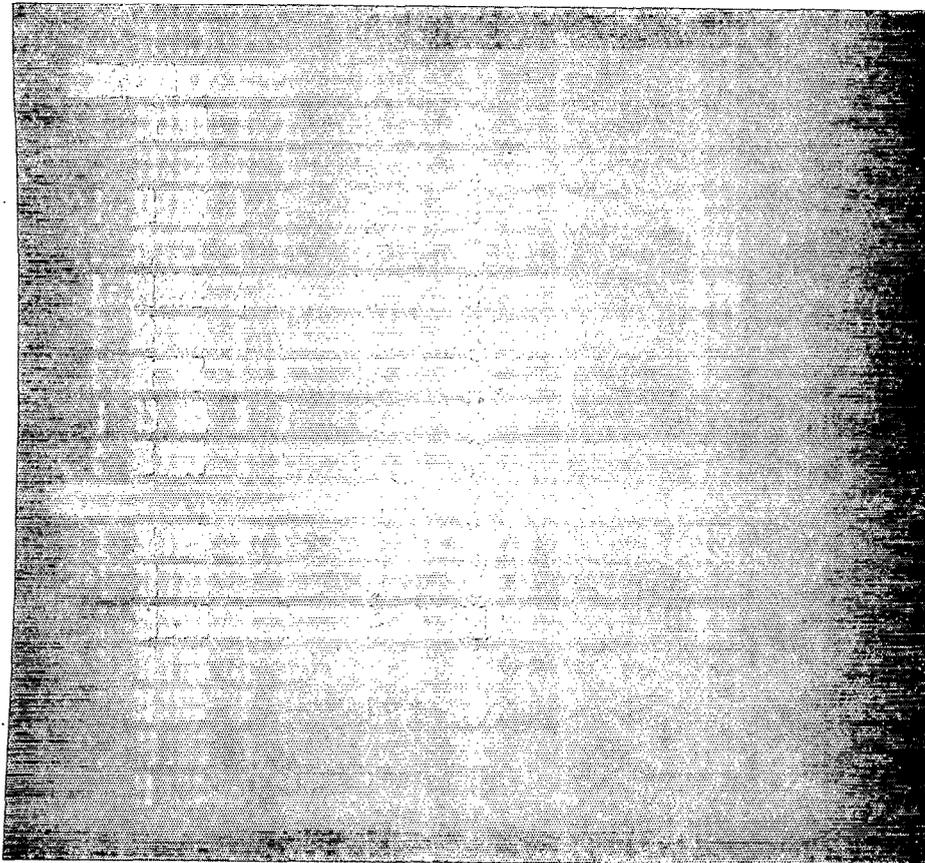


- q 1
- q 2
- n 3
- n 2
- n 1
- 1 (n°210)
- 1 (n°212)
- 1 (n°224)
- 1 (n°227)
- 1 (n°223)
- m 1
- m 2
- m 3
- m 4
- o
- p 1
- p 2
- p 3
- k

Fig. VII A+B: patrons de restriction ADNcp x *Hinc* III



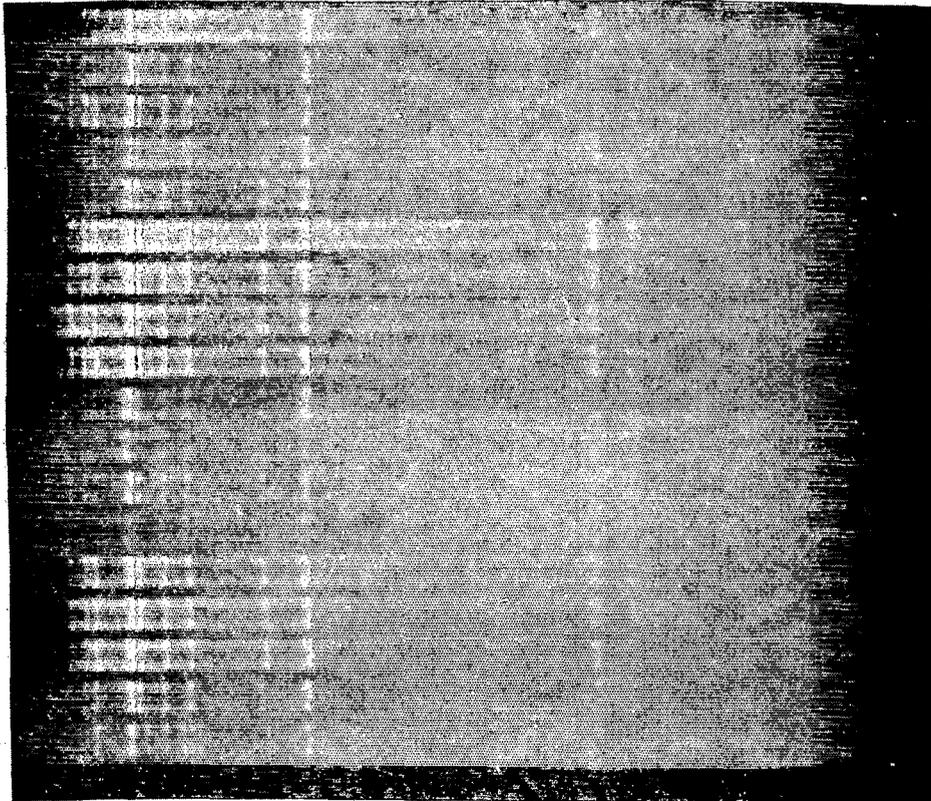
- EX 56
- YS 309
- Akai Mai
- BS 20
- B 15
- B 14
- UL 12-6
- W 1945
- W 1409
- W 1408
- W 15
- I IP 2
- W 1331
- W 1273
- W 1199
- W 1144
- W 1108
- W 1029
- W 1029
- W 1027



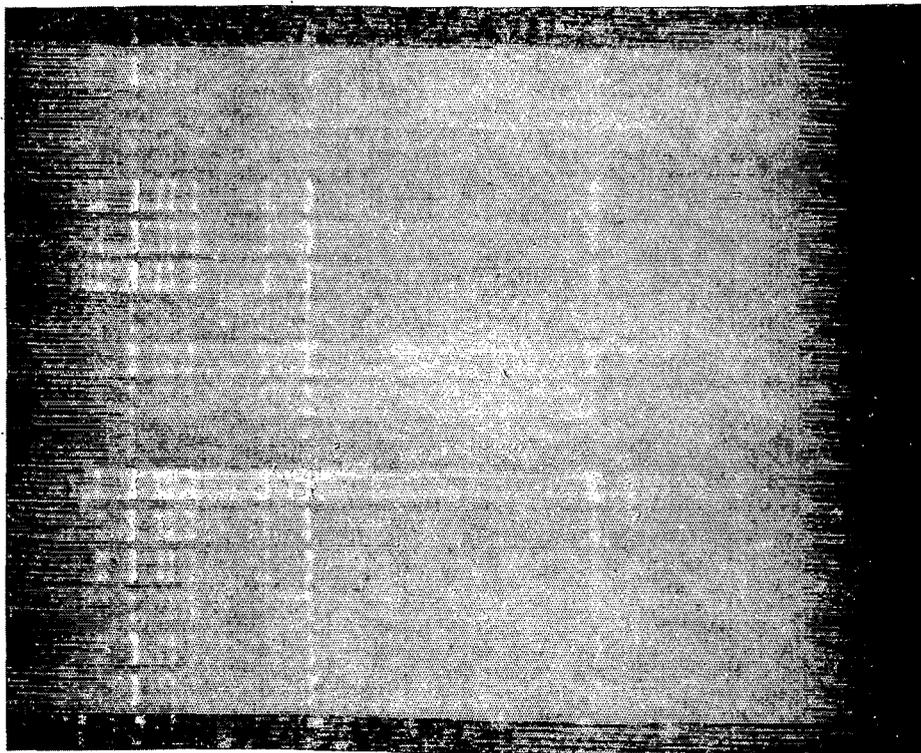
plastotypes

- m<sub>3</sub>
- (q)
- m<sub>2</sub>
- m<sub>1</sub>
- m<sub>0</sub>
- m<sub>2</sub>
- l
- o
- pp<sub>3</sub>
- k
- ii<sub>1</sub>
- (h)
- g
- f<sub>1</sub>
- e<sub>1</sub>
- c
- b
- a<sub>2</sub>
- a<sub>1</sub>

Fig. VIII A+B: patrons de restriction ADNcp x Sma I



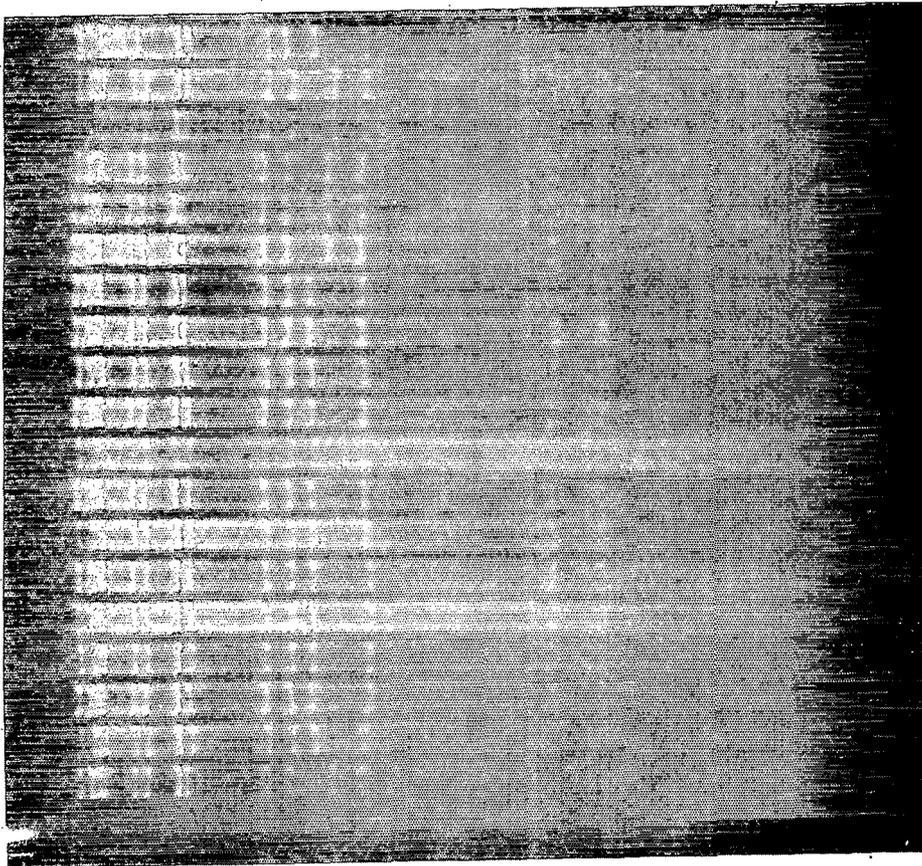
- plastotypes
- q 1
- n 3
- n 2
- n 1
- 1 (n°210)
- 1 (n°212)
- 1 (n°224)
- 1 (n°227)
- 1 (n°233)
- m 1
- m 2
- m 3
- m 4
- o
- p 1
- p 2
- p 3
- k
- (marqueur)



- i 2
- i 1
- h 1
- g (n°178)
- g (n°161)
- f 1
- d
- e 5
- e 4
- e 3
- e 2
- e 1
- e
- b
- a 2
- (n° 84)
- a 1
- (marqueur)

Fig. IX: patrons de restriction ADNcp x *Pst* I

plastotypes



q 2

n 2

n 1

m 1

m 2

l

o

p 3

k

i 1

h

g

e 1

c

b

a 2

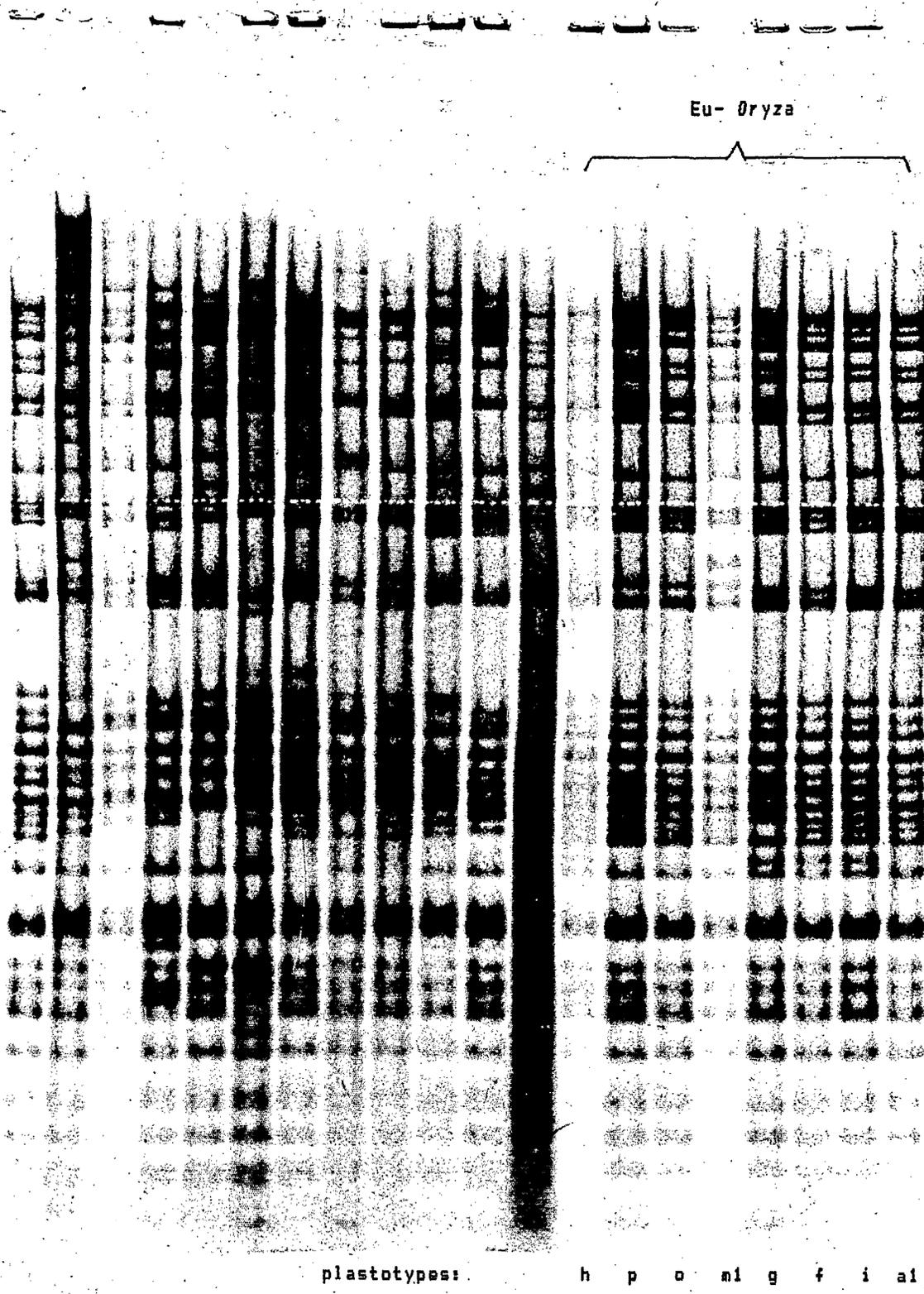
a 1

f

m 3

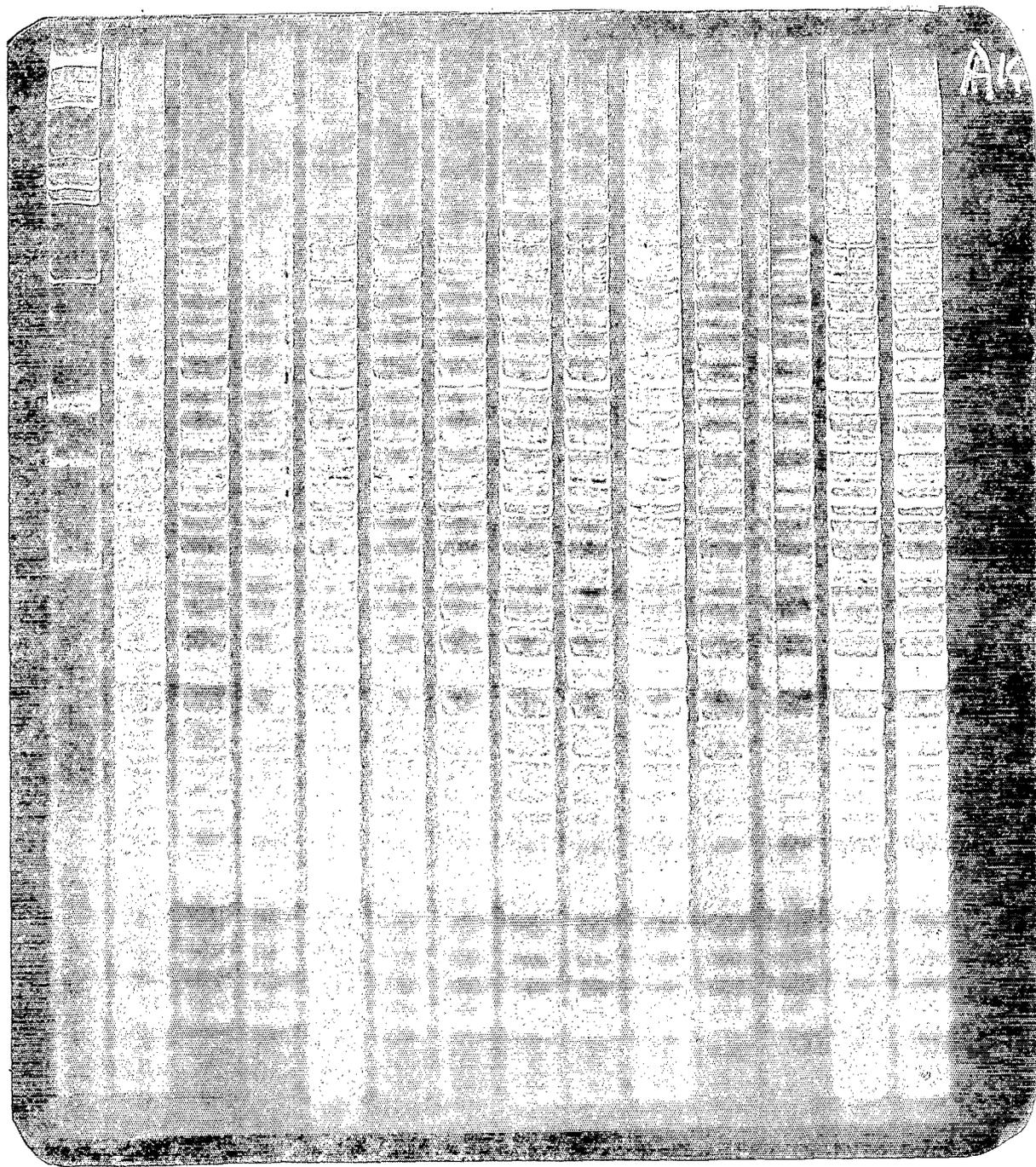
(marqueur)

Fig. X: patrons de restriction ADNcp x Bsp X I (iso Cla I)



plastotypes: h p o n l g f i a l

Fig. XI A: patrons de restriction ADNcp x Ksp I



(lambda)

S 24

W 1699

104 054

W 1653

108

521

W 1827

W 593

6 6

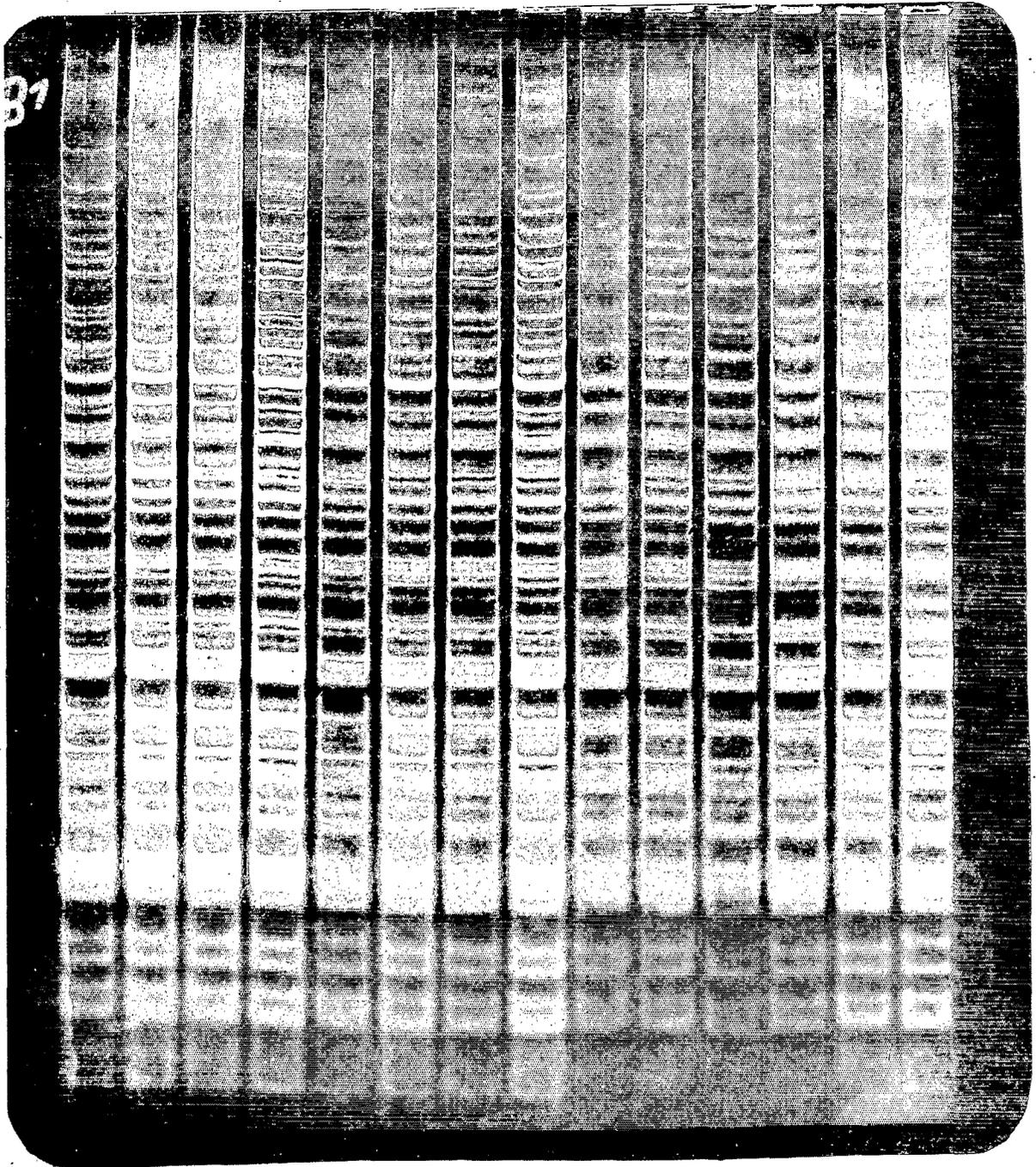
B 14

W 1187

W 1185

UL 128

Fig. XI B: patrons de restriction ADNcp x Hsp I



101 395

W 1539

W 1168

W 1592

W 1515

W 43

W 1271

W 1181

W 2

W 1024

W 1525

2 IP 2

IP 7

W 8

## ANNEXE D

Restitutions dessinées des patrons de restriction d'ADNcp dans la Section Eu-Oryza

fig. XII: patrons de restriction par Eco R I de l'ADNcp du riz (Section Eu-Oryza)

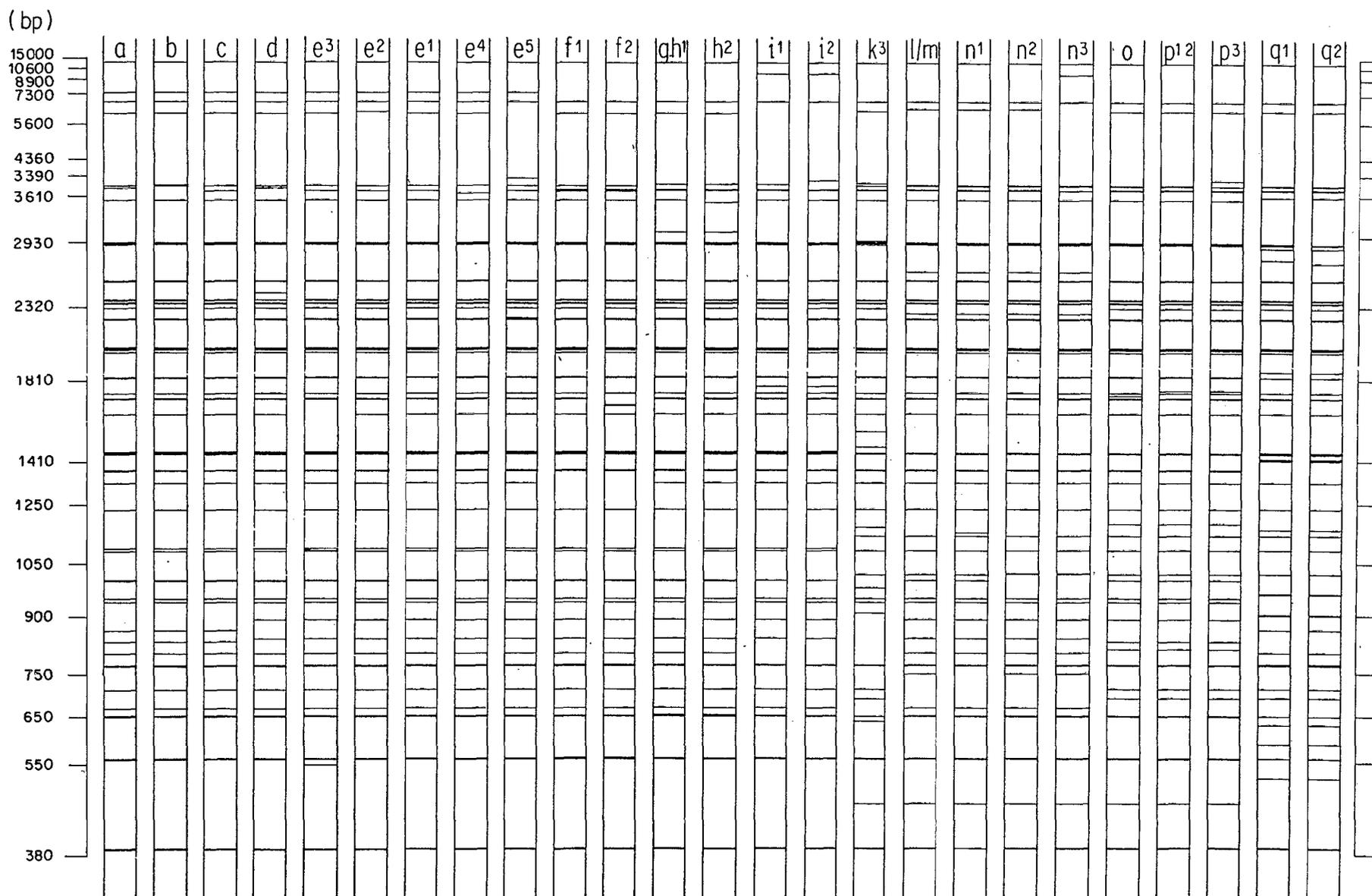


fig. XIII: patrons de restriction par Ava I de l'ADNcp du riz (Section Eu-Oryza)

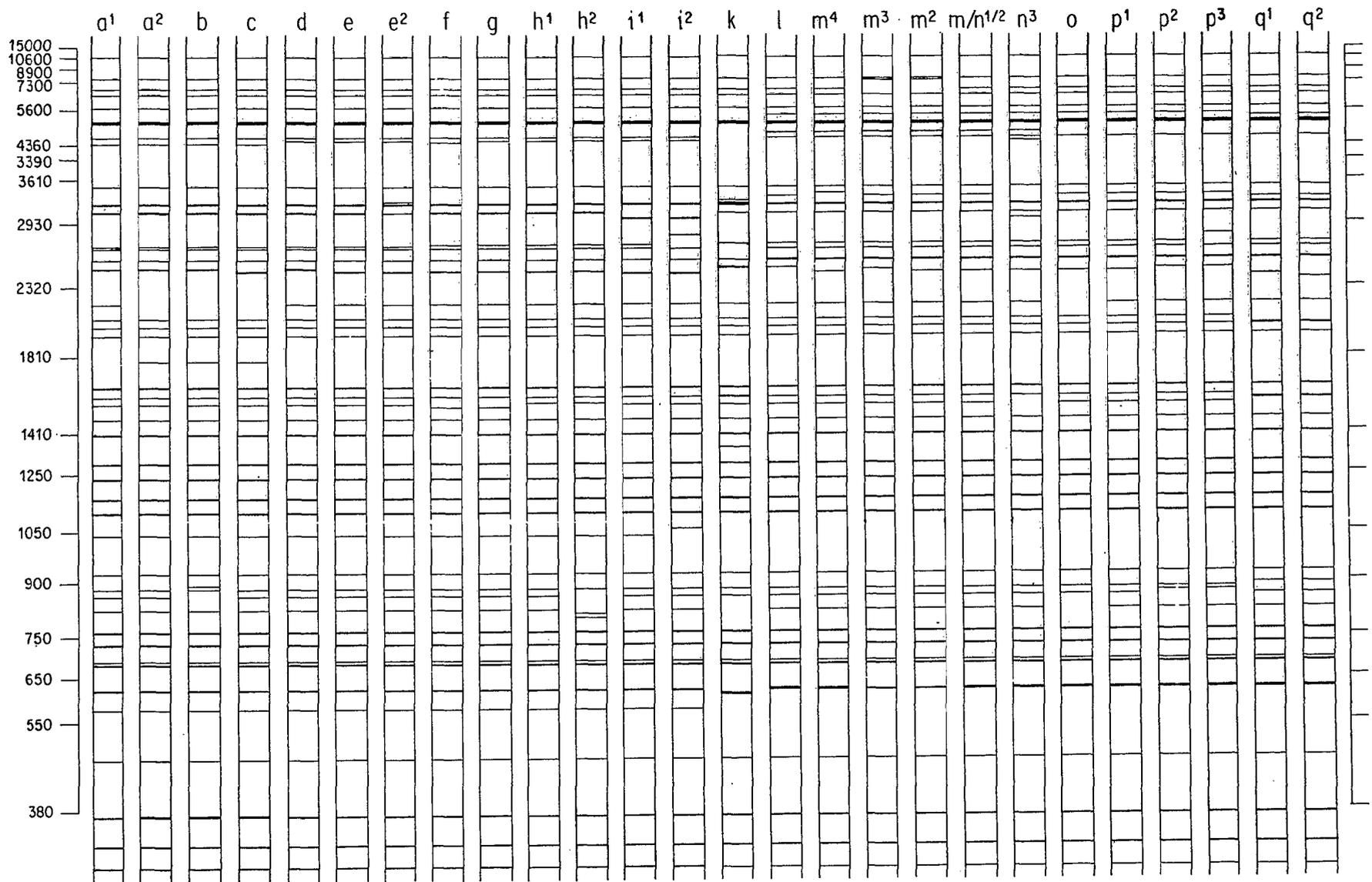


fig. XIV: patrons de restriction de l'ADNcp du riz (Section Eu-Oryza)

a) par Bam H I

b) par Hind III

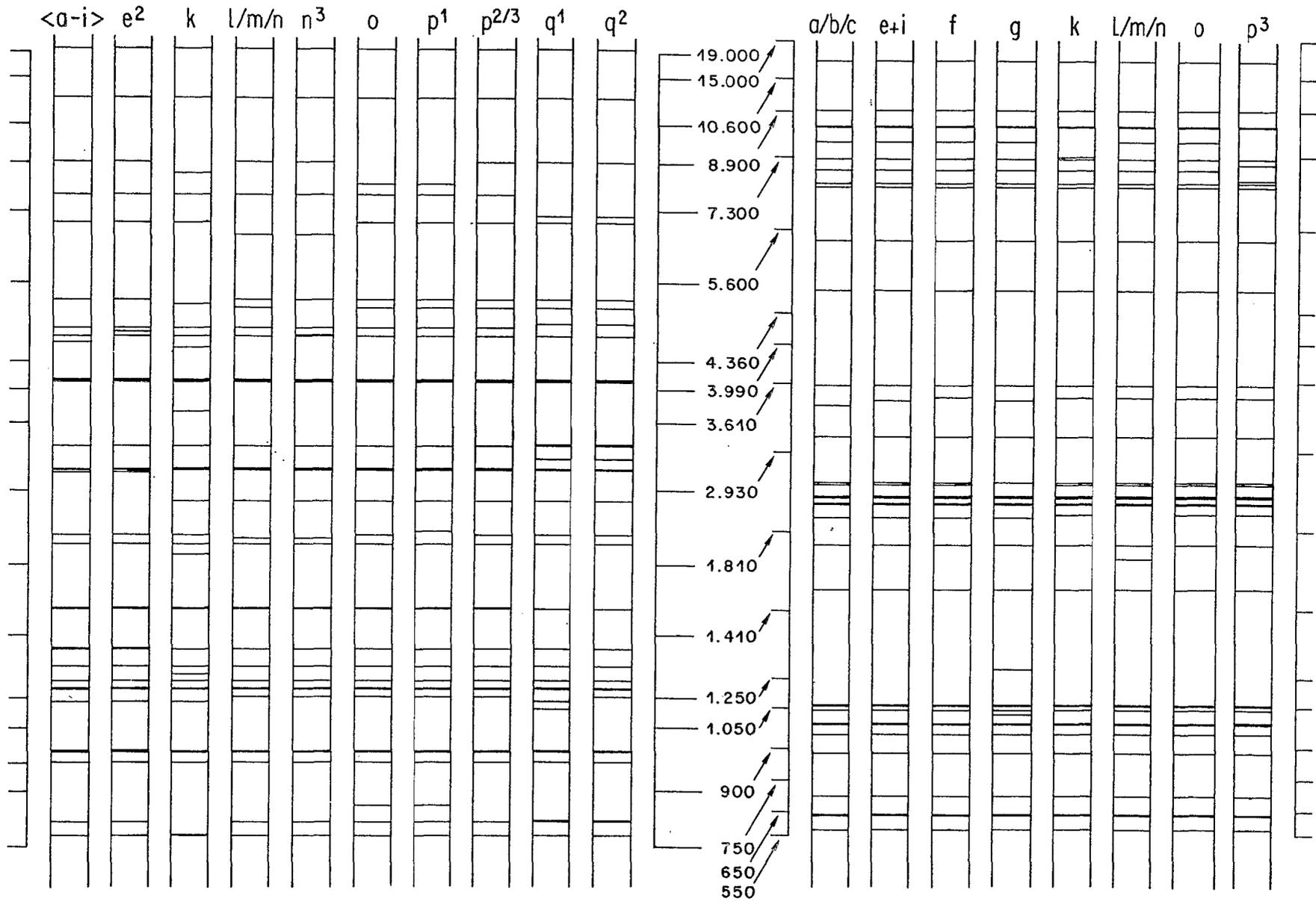
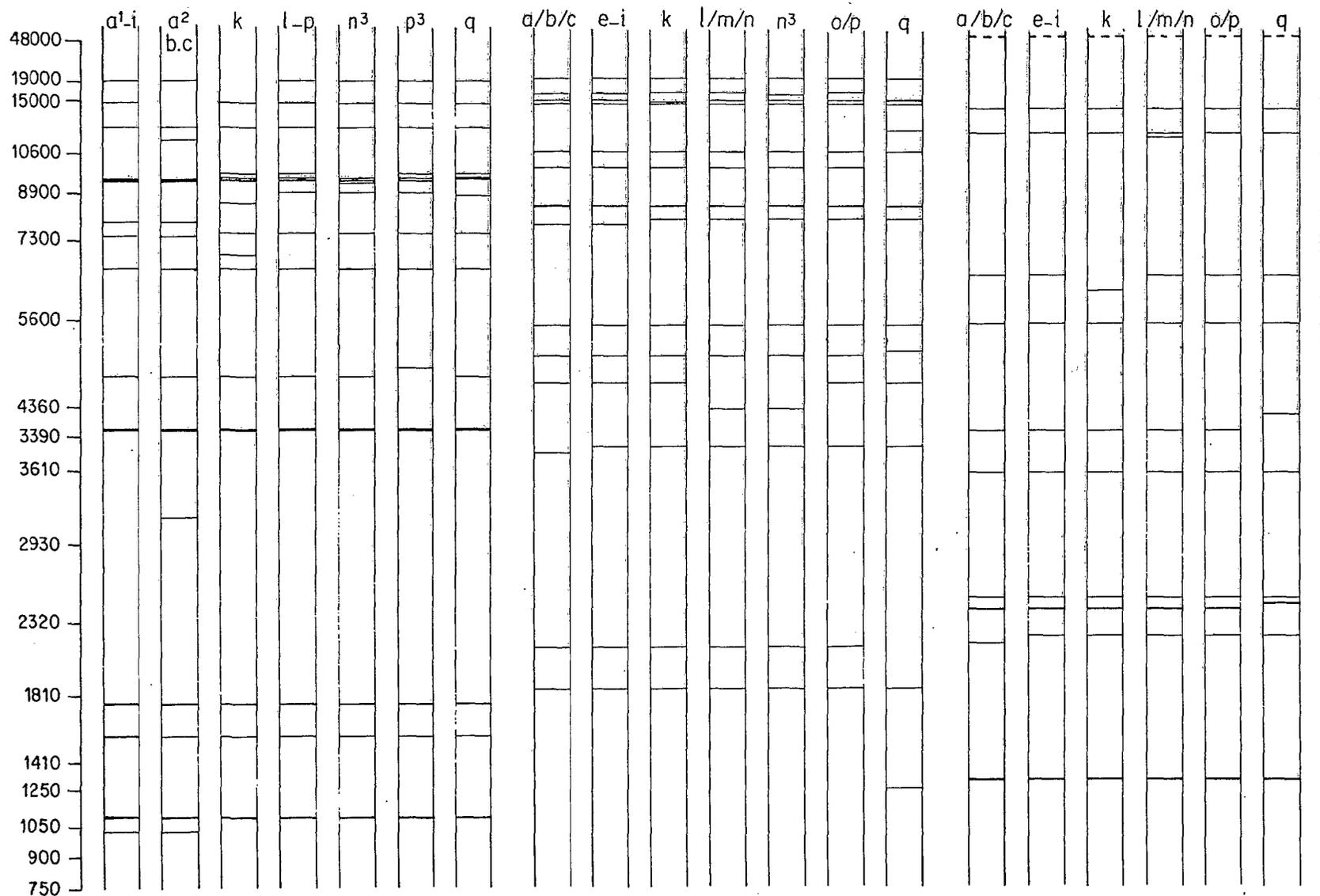


fig. XV: patrons de restriction de l'ADNcp du riz (Section Eu-Oryza)

a) par Sma I

b) par Pst I

c) par Bst E II



## ANNEXE E

### Protocole de la digestion de l'ADNcp par l'enzyme Msp I\* suivie par marquage radioactif au <sup>35</sup>S et électrophorèse en gel polyacrylamide à gradient de densité continu

#### 1) Composition des gels d'électrophorèse :

composants	solution A	solution B
acrylamide** 30% - bisacrylamide** 0,8% (p/v)	7,9 ml	20,0 ml
saccharose 36% (p/p)	2,6 ml	2,4 ml
TBE 10x	6,6 ml	6,0 ml
TEMED**	70 µl	50 µl
eau	48,7 ml	31,5 ml
persulfate d'ammionium** 12%	110 µl	35 µl

à partir des solutions A et B, les gels (épaisseur 0,7 mm, longueur 20 cm) sont coulés en gradient continu (de concentration en acrylamide de 10% à environ 4%) à l'aide d'un mélangeur et support (Pharmacia) entre des plaques de verre (les quantités indiquées sont suffisantes pour une paire de gels);

\* le même protocole est valable pour Hpa II

\*\* produits Biorad

2) composition des essais de restriction (à 0°C) :

solution d'ADNcp + tampon TE :	2,4	µl	(ADN: 0,1 µg) ,
tampon d'incubation 10x :	0,7	µl ,	
enzyme de restriction (10 U/µl) :	0,5	µl ,	
RNAse (2mg/ml) :	0,3	µl ,	
eau bidestillée stérile :	3,1	µl ;	

3) incuber à 37°C pendant 2 h , puis refroidir à 0°C;

4) pour le marquage radioactif, rajouter :

- enzyme de Klenow (5 U/µl; Boehringer)	:	0,1	µl ,
- solution de dC5'-( <sup>35</sup> S)TP (650 Ci/mmol, 10 mCi/ml; Amersham):	:	0,05	µl ;
- eau bidestillée stérile	:	0,95	µl

incuber pendant 30 min à température ambiante;

refroidir dans la glace, mêler avec 0,9 µl de tampon de charge;

5) électrophorèse: 40' à 50 V , puis 5 h à 190 V ,  
en TBE recirculée dans un bac vertical (Pharmacia);

6) séchage des gels (sècheur Biorad);

7) autoradiographie des gels séchés pendant 5 jours.

## REFERENCES BIBLIOGRAPHIQUES

- BANKS JA, BIRKY CW JR (1985) Chloroplast DNA diversity is low in a wild plant, Lupinus texensis. Proc Natl Acad Sci USA 82: 6950-6954
- BIRKY CW JR (1983) Relaxed Cellular Controls and Organelle Heredity. Science 222: 468-475
- BOWMAN CM, BONNARD G, DYER TA (1983) Chloroplast DNA variation between species of Triticum and Aegilops. Location of the variation on the chloroplast genome and its relevance to the inheritance and classification of the cytoplasm. Theor Appl Genet 65: 247-262
- BOWMAN CM, DYER TA (1982) Purification and analysis of DNA from wheat chloroplasts isolated in nonaqueous media. Anal Biochem 122: 108-118
- BRÜCHER H (1977) Morpho-genetic study of a perennial wildrice (Oryza sp.) from the Paraguayan Chaco. Sabrao Journal 9(2): 86-90
- CLEGG MT, BROWN AHD, WHITFIELD PR (1984) Chloroplast DNA diversity in wild and cultivated barley: implications for genetic conservation. Genet Res 43: 339-343
- CLEGG MT, RAWSON JRY, THOMAS K (1984) Chloroplast DNA variation in pearl millet and related species. Genetics 106: 449-461
- CURTIS SE, CLEGG MT (1984) Molecular evolution of chloroplast DNA sequences. Molec Biol Evol 1(4): 291-301
- DOEBLEY J (1987) Evolution of the Zea chloroplast genome. Evol Trends Plants 1(1): 3-6
- ENOMOTO S, OGIHARA Y, TSUNEWAKI K (1985) Studies on the origin of crop species by restriction endonuclease analysis of organellar DNA. I. Phylogenetic relationships among ten cereals revealed by the restriction fragment patterns of chloroplast DNA. Jpn J Genet 60: 411-424
- HERRMANN RG, POSSINGHAM JV (1980) The plastid DNA - the plastome. In Reinert J (ed.) pp 45-96
- HIRAI A, ICHIKAWA H, IWATSUKI N, SUGIURA M (1985) Rice chloroplast DNA: a physical map and the location of the genes for the large subunit of ribulose 1,5-bisphosphate carboxylase and the 32 kd photosystem II reaction center protein. Theor

Appl Genet 70: 117-122

- HOLWERDA BC, JANA S, CROSBY WL (1986) Chloroplast and Mitochondrial DNA variation in Hordeum vulgare and Hordeum spontaneum. Genetics 114: 1271-1291
- ICHIKAWA H, HIRAI A, KATAYAMA T (1986) Genetic analyses of Oryza species by molecular markers for chloroplast genomes. Theor Appl Genet 72: 353-358
- I.N.R.S., Paris: fiches toxicologiques n<sup>os</sup> 17 et 151
- ISHII T, TERACHI T, TSUNEWAKI K (1986) Restriction endonuclease analysis of chloroplast DNA from cultivated rice species, Oryza sativa and Oryza glaberrima. Jpn J Genet 61: 537-541
- JACOBI G (ed.) (1974) Biochemische Cytologie der Pflanzenzelle - ein Praktikum. Thieme Stuttgart, 197 pp
- JACQUART P, HEIM G, ANTONOWICS J (eds.) (1985) Genetic differentiation and dispersal in plants. NATO Workshop Proceedings Series. Springer, 452 pp
- KADOWAKI K, OSUMI T, NEMOTO H, HARADA K, SHINJYO C (1988) Mitochondrial DNA polymorphism in male-sterile cytoplasm of rice. Theor Appl Genet 75: 234-236
- LEHVÄSLAIHO H, SAURA A, LOKKI J (1986) A sensitive method for restriction enzyme analysis of higher plant chloroplast DNA and its application to grass tribus Festuceae. Poster au Symposium International sur la Biologie Moléculaire des Plantes, Strasbourg, France 8-11/7/1986
- LEHVÄSLAIHO H, SAURA A, LOKKI J (1987) Chloroplast DNA variation in the grass tribe Festuceae. Theor Appl Genet 74: 298-302
- LOLO OM (1987) Analyse comparée de la structure génétique d'Oryza rufipogon Griff. en Inde de l'Ouest: un exemple d'introduction suivie d'introggression génétique en Inde d'un riz d'origine africaine. Thèse, Université de Paris-Sud, Centre d'Orsay, 125 pp
- MANIATIS T, FRITSCH EF, SAMBROCK J (1982) Molecular Cloning - a laboratory manual. Cold Spring Harbor Laboratory, 530 pp
- MIGNOUNA H, VIRMANI SS, BRIQUET M (1987) Mitochondrial DNA modifications associated with cytoplasmic male sterility in rice. Theor Appl Genet
- PALMER JD (1985) Chloroplast DNA and molecular phylogeny. BioEssays 2: 263- 267
- PERL-TREVES R, GALUN E (1985a) The Cucumis plastome: physical map, intrageneric variation and phylogenetic relationships. Theor Appl Genet 71: 417-429

- REINERT J (ed.) (1980) Chloroplasts. Springer, Berlin, 240 pp
- RICKWOOD D, HAMES BD (ed) (1982) Gel electrophoresis of nucleic acids - a practical approach. IRL press, Oxford
- SECOND G (1985a) Relations évolutives chez le genre *Oryza* et processus de domestication des Riz. Etudes et Thèses, ORSTOM, Paris 180 pp
- SECOND G (1985b) Geographic origins, genetic diversity and the molecular clock hypothesis in the *Oryzeae*. In Jacquart P,...(eds.) (1985)
- SECOND G (1985c) Evolutionary relationships in the *Sativa* group of *Oryza* based on isozyme data. Genet Sel Evol 17-1: 89-114
- SHINOZAKI K, OHME M, TANAKA M, WAGASUKI T, HAYASHIDA N, MATSUBAYASHI T, ZAITA N, CHUNWONGSE J, OBOKATA J, YAMAGUCHI-SHINOZAKI K, OHTO C, TORAZAWA K, MENG BY, SUGITA M, DENO H, KAMOGASHIRA T, YAMADA K, KUSUDA J, TAKAIWA F, KATO A, TOHDOH N, SHIMADA H, SUGIURA M (1986) The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. The EMBO Journal 5(9): 2043-2049
- SOKAL RR, SNEATH PHA (1963) Principles of numerical taxonomy. W.H. Freeman, San Francisco
- SOUTHERN EM (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol 98: 503-517
- SYTSMAN KJ, GOTTLIEB LD (1986) Chloroplast DNA evolution and phylogenetic relationships in *Clarkia* sect. *Peripetasma* (Onagraceae). Evolution 40(6): 1248-1261
- TAKAHATA N, MARUYAMA T (1981) A mathematical model of extranuclear genes and the genetic variability maintained in a finite population. Gen Res 37: 291-302
- TERACHI T, OGIHARA Y, TSUNEWAKI K (1984) The molecular basis of genetic diversity among cytoplasmic of *Triticum* and *Aegilops*. Genetics 108: 681-695
- TSUNEWAKI K (ed) (1980) Genetic Diversity of the Cytoplasm in *Triticum* and *Aegilops*. Japanese Society for the Promotion of Science, Tokyo, 290 pp
- VEDEL F, QUETIER F, BAYEN M (1976) Specific cleavage of chloroplast DNA from higher plants by *EcoRI* restriction nuclease. Nature 263: 440-442
- VEDEL F, QUETIER F, DOSBA F, DOUSSINAULT G (1978) Study of wheat phylogeny by *EcoRI* analysis of chloroplastic and mitochondrial DNAs. Plant Science Lett 13: 97-102
- VEDEL F, LEBACQ P, QUETIER F (1980) Cytoplasmic DNA variation and relationships in

cereal genomes. *Theor Appl Genet* 58: 219-224

ZURAWSKY G, CLEGG MT (1987) Evolution of higher-plant chloroplast DNA-encoded genes: Implications for structure-function and phylogenetic studies. *Ann Rev Plant Physiol* 38: 391-418

ZURAWSKY G, CLEGG MT, BROWN AHD (1984) The nature of nucleotide sequence divergence between barley and maize chloroplast DNA. *Genetics* 106: 735-749

## BIBLIOGRAPHIE

- ADAMS J, ROTHMAN ED (1982) Estimation of phylogenetic relationships from DNA restriction patterns and selection of endonuclease cleavage sites. Proc Natl Acad Sci USA 79: 3560-3564
- ATCHINSON BA, WHITFIELD PR, BOTTOMLEY W (1976) Comparison of chloroplast DNAs by specific fragmentation with EcoRI endonuclease. Mol Gen Genet 148: 263-269
- BAATOUT HM, MARRAKCHI M, MATHIEU C, VEDEL F (1985) Variation of plastid and mitochondrial DNAs in the genus Hedysarum. Theor Appl Genet 70: 577-584
- BARD SA, MILTON PG (1969) Studies on spinach chloroplast and nuclear DNA using large-scale tissue preparations. Plant Physiol 44: 377-384
- BECKMANN JS, SOLLER M (1986) Restriction fragment length polymorphisms in plant genetic improvement. In: MIFLIN BJ (ed.) Oxford Surveys of Plant Molecular & Cell Biology, Vol. 3
- BELL G (1988) Recombination and the immortality of the germ line. J evol Biol 1: 67-82
- BERTHOU F, MATHIEU C, VEDEL F (1983) Chloroplast and mitochondrial DNA variation as indicator of phylogenetic relationships in the genus Coffea L. Theor Appl Genet 65: 77-84
- BIRKY CW JR (1978) Transmission genetics of mitochondria and chloroplasts. Ann Rev Genet 12: 471-512
- BOHNERT HJ, LÖFFELHARDT W (1982) Cyanelle DNA from Cyanophora paradoxa exists in two forms due to intramolecular recombination. FEBS Lett 150: 403-406
- BONNARD G (1981) Une étude de la variabilité cytoplasmique du blé et d'espèces voisins. Mémoire de DEA, Cambridge / Rennes
- BOWMAN CM, DYER TA (1986) The location and possible evolutionary significance of small dispersed repeats in wheat ct DNA. Curr Genet 10: 931-941
- BUTTERFASS T (1980) The Continuity of Plastids and the Differentiation of Plastid Populations. In: REINERT J (ed.) pp 29-44
- CAVALIER-SMITH T (1970) Electron microscopic evidence for chloroplast fusion in zygotes of Chlamydomonas reinhardtii. Nature 228: 333-335

- DAY A, ELLIS THN (1984) Chloroplast DNA deletions associated with wheat plants regenerated from pollen: possible basis for maternal inheritance of chloroplasts. *Cell* 39: 359-368
- DEBRY RW, SLADE NA (1985) Cladistic analysis of restriction endonuclease cleavage maps within a maximum-likelihood framework. *Syst Zool* 34: 21-34
- DOVER GA (1987) DNA turnover and the molecular clock. *J Mol Evol* 26: 47-58
- DYER TA (1985) The chloroplast genome and its products. *Oxford Surveys of Plant Molecular and Cell Biology* Vol 2
- EDELMAN M, HALLICK RB, CHUA N-H (eds) (1982) *Methods in chloroplast molecular biology*. Elsevier / North Holland, Amsterdam, 259 pp
- FELSENSTEIN J (1985a) Confidence limits on phylogenies with a molecular clock. *Syst Zool* 34: 152-161
- FELSENSTEIN J (1985b) Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* 39: 783-791
- FITCH WM, MARGOLIASH E (1967) Construction of phylogenetic trees. *Science* 155: 279-284
- FLUHR R, AVIV D, GALUN E, EDELMAN M (1984) Generation of heteroplastidic *Nicotiana* cybrids by protoplast fusion: analysis for plastid recombinant types. *Theor Appl Genet* 67: 491
- FLUHR R, EDELMAN M (1981) Conservation of sequence arrangement among higher plant chloroplast DNAs: Molecular cross hybridization among the Solanaceae and between *Nicotiana* and *Spinacia*. *Nucleic Acids Res* 9: 6841-6853
- GLEBA YY, KOLESNIK NN, MESHKENE IV, CHEREP NN, PAROKKONY AS (1984) Transmission genetics of the somatic hybridization process in *Nicotiana*. I. Hybrids and cybrids among the regenerates from cloned protoplast fusion products. *Theor Appl Genet* 69: 121
- GHESEQUIERE A, SECOND G (1983) Polymorphisme enzymatique et évolution d'*Oryza sativa* en Afrique. Réunion de la Société de Zoologie (Paris, Mai 1983): Electrophorèse et Systématique
- GORDON KHJ, GROUSE EJ, BOHNERT HJ, HERRMANN RG (1982) Physical mapping of differences in chloroplast DNA of the five wild-type plastomes in *Oenothera* subsection *Euoenothera*. *Theor Appl Genet* 61:373-384
- GOTOH O, HAYASHI J-I, YONEKAWA H, TAGASHIRA Y (1979) An improved method for esti-

mating sequence divergence between related DNAs from changes in restriction endonuclease cleavage sites. *J Mol Evol* 14: 301-310

GRAY MW (1983) The bacterial ancestry of plastids and mitochondria. *Bioscience* 33: 393-399

HILU KW (1983) The role of single-gene mutations in the evolution of flowering plants. *Evol Biol* 16: 97-128

HILU KW (1985) Trends of variation and systematics of Poaceae. *Taxon* 34(1): 102-114

HILU KW, WRIGHT K (1982) Systematics of Gramineae: a cluster analysis study. *Taxon* 31(1): 9-36

HOSAKA K, OGIHARA Y, MATSUBAYASHI M, TSUNEWAKI K (1984) Phylogenetic relationship between the tuberous Solanum species as revealed by restriction endonuclease analysis of chloroplast DNA. *Jpn J Genet* 59: 349-369

KEMBLE RJ (1987) A rapid, single leaf, nucleic acid assay for determining the cytoplasmic organelle complement of rapeseed and related Brassica species. *Theor Appl Genet* 73: 364-370

KIRK JTO, TILNEY-BASSET RAE (1978) *The Plastids - Their Chemistry, Structure, Growth, Inheritance*. Elsevier, Amsterdam

KIMURA M (1987) Molecular evolutionary clock and the neutral theory. *J Mol Evol* 26: 24-33

KUNG SD, ZHU YS, CHEN K, SHEN GF, SISSON VA (1981) Nicotiana Chloroplast Genome. II. Chloroplast DNA alteration. *Mol Gen Genet* 183: 20-24

KUNG SD, ZHU YS, SHEN GF (1982) Nicotiana chloroplast genome. III. Chloroplast DNA evolution. *Theor Appl Genet* 61: 73-79

LEBACQ P, VEDEL F (1981) Sal I restriction enzyme analysis of chloroplast and mitochondrial DNAs in the genus Brassica. *Plant Sci Lett* 23: 1-9

LI W-H (1981) Simple method for constructing phylogenetic trees from distance matrices. *Proc Natl Acad Sci USA* 78: 1085-1089

LÖFFELHARDT W (1987) Engineering Chloroplasts: Prospects and limitations. *Physiol Plantarum* 69: 735-741

MACINTYRE RJ (ed.) (1985) *Molecular evolutionary genetics*. Plenum press, New York, 610 pp

MALIGA P, FEJES E (1985) Physical evidence for recombination of chloroplast DNA in a

somatic hybrid of Nicotiana tabacum and Nicotiana plumbaginifolia. In: **Freeling M** (ed.) *Plant Genetics*, A.R. Liss Inc., New York

- MEDGYESY P, FEJES E, MALIGA P** (1985) Interspecific chloroplast recombination in a Nicotiana somatic hybrid. *Proc Natl Acad Sci USA* 82: 6960-6964
- MEDGYESY P, PAY A, MARTON L** (1986) Transmission of paternal chloroplasts in Nicotiana. *Mol Gen Genet* 204: 195-198
- MIKAMI T, KISHIMA Y, SUGIURA M, KINOSHITA T** (1984) Chloroplast DNA diversity in the cytoplasms of sugar beet and its related species. *Plant Sci Lett* 36: 231-235
- NAGL W** (1986) Molecular phylogeny. In **RAUP DM**, ... pp 223-232
- NEI M, LI W-H** (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc Natl Acad Sci USA* 76: 5269-5273
- NEI M, TAJIMA F** (1983) Maximum likelihood estimation of the number of nucleotide substitutions from restriction sites data. *Genetics* 105: 207-217
- NG NQ, CHANG TT, WILLIAMS JT, HAWKES JG** (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
- OGIHARA Y, TSUNEWAKI K** (1984) The diversity of chloroplast DNA among Triticum and Aegilops species. *Proc VI Int Wheat Genet Symp*: 407-413
- OHTA T** (1987) Very slightly deleterious mutations and the molecular clock. *J Mol Evol* 26: 1-6
- OKA HI** (1974) Experimental studies on the origin of cultivated rice. *Genetics* 78: 475-486
- OKA HI** (1988) *Origin of Cultivated Rice*. Elsevier, Amsterdam, 270 pp
- PALMER JD** (1982) Physical and gene mapping of chloroplast DNA from Atriplex triangularis and Cucumis sativa. *Nucleic Acids Res* 10: 1593-1605
- PALMER JD** (1984) Chloroplast DNA exists in two orientations. *Nature* 301: 92-93
- PALMER JD** (1985a) Comparative organization of chloroplast genomes. *Ann Rev Genet* 19: 325-354
- PALMER JD** (1985b) Isolation and structural analysis of chloroplast DNA. *Methods Enzymol* 118: 167-186
- PALMER JD** (1985c) Evolution of chloroplast and mitochondrial DNA in plants and algae. In **MACINTYRE RJ** (ed.) (1985), pp 131-240

- PALMER JD, THOMPSON WF (1982) Chloroplast DNA rearrangements are more frequent when a large inverted repeat is lost. *Cell* 29: 537-550
- PALMER JD, JORGENSEN RA, THOMPSON WF (1985) Chloroplast DNA variation and evolution in Pisum: patterns of change and phylogenetic analysis. *Genetics* 109: 195-213
- PALMER JD, NUGENT JM, HERBON L (1987a) Unusual structure of geranium chloroplast DNA: A triple-sized inverted repeat, extensive gene duplications, multiple inversions, and two repeat families. *Proc Natl Acad Sci USA* 84: 769-773
- PALMER JD, OSORIO B, ALDRICH J, THOMPSON WF (1987b) Chloroplast DNA evolution among legumes: Loss of a large inverted repeat occurred prior to other sequence rearrangements. *Curr Genet* 11: 275-286
- PALMER JD, SHIELDS CR, COHEN DB, ORTON TJ (1983) Chloroplast DNA evolution and the origin of amphidiploid Brassica species. *Theor Appl Genet* 65: 181-189
- PALMER JD, STEIN DB (1986) Conservation of chloroplast genome structure among vascular plants. *Curr Genet* 10: 823-833
- PALMER JD, SINGH GP, PILLARY DTN (1983) Structural and sequence evolution of three legume chloroplast DNAs. *Mol Gen Genet* 190: 13-19
- PALMER JD, THOMPSON WF (1982) Chloroplast DNA rearrangements are more frequent when a large inverted repeat is lost. *Cell* 29: 537-550
- PALMER JD, ZAMIR D (1982) Chloroplast DNA evolution and phylogenetic relationships in Lycopersicon. *Proc Natl Acad Sci USA* 79: 5006-5010
- PELLETIER G, PRIMARD C, VEDEL F, CHETRIT P, REMY R, ROUSSELLE, RENARD M (1983) Intergeneric Cytoplasmic Hybridization in Cruciferae by Protoplast Fusion. *Mol Gen Genet* 191: 244-250
- PERL-TREVES R, ZAMIR D, NAVOT N, GALUN E (1985) Phylogeny of Cucumis based on isozyme variability and its comparison with plastome phylogeny. *Theor Appl Genet* 71: 430-436
- PLATNICK NI, FUNK VA (1985) (ed) *Advances in Cladistics*. Columbia University Press, New York, 190 pp
- RAUP DM, JABLONSKI D (1986) *Patterns and Processes in the History of Life*. Dahlem Konferenzen 1986. Springer, Berlin, Heidelberg
- SAITOU N, NEI M (1987) The neighbour-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4(4): 406-425

- SALTZ Y, BECKMAN J (1981) Chloroplast DNA preparation from Petunia and Nicotiana. Plant Mol Biol Newslet 2: 73-74
- SALTZ Y, HERRMANN RG, PELEG N, LAVI U, IZHAR S, FRANKEL R, BECKMANN JS (1984) Physical mapping of plastid DNA variation among eleven Nicotiana species. Theor Appl Genet 69: 1-14
- SCOTT NS, TIMMIS JN (1984) Homologies between nuclear and plastid DNA in spinach. Theor Appl Genet 67: 279-288
- SEARS BB (1980) Elimination of plastids during spermatogenesis and fertilization in the plant kingdom. Plasmid 4: 233-255
- SECOND G (1982) Origin of the genetic diversity of cultivated rice (Oryza spp.): study of the polymorphism scored at 40 isozyme loci. Jpn J Genet 57: 25-57
- SECOND G (1983a) The pattern of interrelationships among strains of the Sativa group (AA genome) of Oryza studied at 24 isozyme loci.
- SECOND G (1983b) A differential rate of genome divergence evidenced between two species groups in genus Oryza L. Comm. to the Int. Symp. on chromosome research, Calcutta, 23-26 Dec. 1983
- SECOND G (1984a) The study of isozymes in relation to the distribution of the genus Oryza in the palaeoenvironment and the subsequent origin of cultivated rice. In Whyte RO (ed.) The evolution of the east asian environment Vol II: Palaeobotany, Palaeozoology and Palaeoanthropology, pp 664-681
- SECOND G (1984b) Different rates of genome divergence presumed between two species groups in the genus Oryza. Nucleus 27(1,2): 44-48
- SECOND G (1985) A new insight into the genome differentiation in Oryza L. through isozymic studies. In SHARMA AK, ... (eds.) (1985), pp 45-78
- SECOND G (1986) La domestication en régime autogame: exemple des Riz (Oryza spp.). Bull Soc bot Fr 133: Actual bot 35-44
- SECOND G, TROUSLOT P (1980) Electrophorèse d'enzymes de Riz (Oryza spp.). Travaux et Documents 120, ORSTOM Paris
- SHARMA AK, SHARMA A (eds.) (1985) Advances in chromosomes and cell genetics. Oxford and IBH Publishing Co. New Delhi
- STEARNS DB, PALMER JD (1984) Extensive and widespread homologies between mitochondrial and chloroplast DNA in plants. Proc Natl Acad Sci USA 81: 1946-1950

- STEBBINS GL (1956a) Taxonomy and the evolution of genera, with special reference to the family Gramineae. *Evolution* 10: 235-245
- STEIN DB, PALMER JD, THOMPSON WF (1986) Structural evolution and flip-flop recombination of chloroplast DNA in the fern genus Osmunda. *Curr Genet* 10: 835-41
- SYTSMA KJ, GOTTLIEB LD (1986) Chloroplast DNA evidence for the derivation of the genus Heterogaura from a species of Clarkia (Onagraceae). *Proc Natl Acad Sci USA* 83: 5554-5557
- TAKAHATA N, SLATKIN M (1984) Mitochondrial gene flow. *Proc Natl Acad Sci USA* 81: 1764-1767
- THORPE JP (1982) The molecular clock hypothesis - Biochemical evolution, genetic differentiation and systematics. *Ann Rev Ecol Syst* 13: 139-168
- TIMMIS JN, SCOTT NS (1983) Sequence homology between spinach nuclear and chloroplast genomes. *Nature* 305: 65-67
- TSUNODA S, TAKAHASHI N (1984) *Biology of Rice*. Elsevier, Amsterdam, 390 pp
- UPHOLT WB (1977) Estimation of DNA sequence divergence from comparison of restriction endonuclease digest. *Nucleic Acids Res* 4: 1257-1265
- VAUGHN KC, DEBONTE LR, WILSON KG, SCHAEFFER GW (1980) Organelle alteration as a mechanism for maternal inheritance. *Science* 108: 196-198
- VEDEL F, QUETIER F, CAUDERON Y, DOSBA F, DOUSSINAULT G (1981) Studies on Maternal Inheritance in Polyploid Wheats with Cytoplasmic DNAs as Genetic Markers. *Theor Appl Genet* 59: 239-245
- WEISSBACH A, WEISSBACH H (eds.) (1985) *Plant molecular biology - methods in enzymology*. Academic Press, Orlando, Florida
- WILLIAMS BL, WILSON K (<sup>2</sup>1984) *Methoden der Biochemie*. Thieme Stuttgart / New York, 365 pp
- WOBUS U (ed.) (1981) *Isolierung, Fraiktionierung und Hybridisierung von Nukleinsäuren: eine Einführung und methodische Anleitung*. Akademie-Verlag Berlin (DDR), 229 pp

## REMERCIEMENTS

Mlle. COULAZOU du C.R.O.U.S. et ses collègues m'ont réservé un accueil bien chaleureux et généreux à Montpellier,

comme cela est vrai pour M. P. JACQUART, le personnel, les chercheurs et les collègues étudiants au Centre Emberger du C.N.R.S..

Tout le long de mon séjour au C.N.R.S., les chercheurs du C.R.B.M. m'ont aidé de nombreuses façons, ainsi que ceux de la station de l'INRA à St. Christol-les-Alès où j'ai reçu un accueil très favorable et généreux lors d'un court stage en décembre 1987.

MM. L. THALER et M. DELSENY ont bien voulu accepter de présider et participer au Jury.

M. T. DYER m'a fait profiter sans compter de son savoir-faire et de ses conseils. En plus, comme M. A. CHARRIER, il acceptait généreusement la charge du rapport.

Et, bien sûr, pendant trois années c'est M. G. SECOND qui me guidait avec un instinct infallible de la bonne voie.

Je tiens à les remercier tous !

### errata

page 110: toutes les variétés classe 1 sont en fait des classes 2 sauf S 29 et S 32

pages 101 et 108: pour "W 1949", il doit s'agir d'une erreur d'étiquetage, car W 1949 est un Q. officinalis

### précision

page 97: Dao Pao (n° 23 717) est un riz pluvial thaïlandais

## Résumé

Une méthode non-aqueuse améliorée pour la préparation d'ADN chloroplastique (ADNcp) à partir de matériel foliaire séché de plantes adultes a permis une analyse des PLFR de l'ADNcp du riz à l'aide de 10 endonucléases de restriction. 320 plantes, représentant 248 variétés et 15 espèces de la Section Eu-Oryza du genre Oryza, furent analysées.

Un phylogramme hypothétique complet de 32 types de plastomes, dont 10 parmi les riz cultivés, fut déduit d'une analyse cladistique de 112 mutations décelées, calibré à l'aide de données bibliographiques sur les taux de mutation de l'ADNcp chez les Graminées, et confronté à certains repères paléo-géographiques.

Les résultats apportent des informations nouvelles, en bonne partie compatibles avec les données existantes au niveau des génomes nucléaires et leur interprétation évolutive

**Mots clés:** riz - chloroplastes - ADN - extraction non-aqueuse - PLFR - analyse des mutations - cladistique - évolution - phylogénie - horloge moléculaire - genre Oryza - Section Eu-Oryza

## Summary

An improved non-aqueous method for the preparation of chloroplast DNA (cpDNA) from dried leaf material of adult plants permitted a RFLP analysis of rice cpDNA with 10 restriction endonucleases. 320 plants, representing 248 varieties and 15 species of the Section Eu-Oryza of the genus Oryza have been investigated.

A complete phylogram of 32 types of chloroplast genomes, 10 of them present in cultivated rices, has been deduced from a cladistic analysis of 112 precisely characterized mutations, calibrated by using bibliographic data on the cpDNA mutation rates of Gramineae, and confronted with certain paleo-geographic reference marks.

The results give new informations, in general compatible with the existing data at the level of nuclear genomes and their interpretation in terms of evolution.

**Key words:** rice - chloroplasts - DNA - non-aqueous extraction - RFLP - mutation analysis - cladistics - evolution - phylogeny - molecular clock - genus Oryza - Section Eu-Oryza