

Polymorphisme enzymatique et organisation génétique des deux espèces cultivées tétraploïdes de cotonnier, *G. hirsutum* et *G. barbadense*

I. Evolution et domestication des deux espèces

C. Bourdon

Laboratoire d'électrophorèse I.R.C.T., C.I.R.A.D., B.P. 5035, 34032 Montpellier Cedex.

RÉSUMÉ

La diversité et la structure génétiques des 2 espèces de cotonnier cultivées allotétraploïdes, *G. hirsutum* et *G. barbadense*, sont étudiées à l'aide du polymorphisme de 6 systèmes enzymatiques, soit 14 locus. L'analyse porte sur des cotonniers sauvages et spontanés des 2 espèces, ainsi que sur un échantillon de variétés cultivées de *G. hirsutum*.

Les cotonniers sauvages se caractérisent par un polymorphisme important, particulièrement au niveau intra-population, à cause d'un taux d'allogamie non négligeable. Chez l'espèce *G. barba-*

dense, les cotonniers du Pérou et des Galapagos apparaissent fortement différenciés.

Les cotonniers spontanés sont très peu variables chez *G. barbadense*, tandis que chez *G. hirsutum*, ils manifestent une importante différenciation entre races géographiques, à laquelle s'ajoutent chez la race Marie-Galante, des phénomènes d'introgressions aussi bien inter qu'intraspécifiques.

L'ensemble de ces résultats, discuté d'un point de vue évolutif, apporte des éléments nouveaux sur les événements ayant jalonné l'histoire et la domestication de ces espèces.

MOTS CLÉS : *Gossypium hirsutum* et *barbadense*, polymorphisme enzymatique, cotonniers sauvages et spontanés, évolution, domestication.

INTRODUCTION

Le genre *Gossypium* L. a une vaste répartition géographique et manifeste une grande diversité à la fois morphologique et cytogénétique puisqu'il est composé d'espèces diploïdes, elles-mêmes subdivisées en 7 génomes (A à G) et d'espèces allotétraploïdes (génome AD). L'électrophorèse protéique et enzymatique offre donc un premier intérêt systématique et évolutif. Deux types essentiels de questions ont été abordés par les équipes ayant utilisé jusqu'à présent ces techniques sur le cotonnier : les filiations entre génomes au niveau diploïde (CHERRY *et al.*, 1970, 1972 ; JOHNSON et THEIN, 1970), la détermination des espèces ancestrales des allotétraploïdes (CHERRY *et al.*, 1971 ; JOHNSON, 1975 ; HANCOCK, 1982). Ces études utilisent la plupart du temps

l'électrophorèse de protéines totales ou d'enzymes non spécifiques (estérases). Seuls des travaux plus récents exploitent le polymorphisme d'une dizaine de systèmes enzymatiques pour évaluer la diversité intraspécifique des espèces diploïdes cultivées *G. herbaceum* et *G. arboreum* (SUTER et PARKS, 1984).

Notre étude se situe plus précisément dans le contexte des ressources génétiques des 2 espèces cultivées allotétraploïdes *G. hirsutum* et *G. barbadense*, où l'analyse du polymorphisme enzymatique vient s'ajouter aux diverses observations effectuées dans le cadre de la Banque de Génotypes de l'I.R.C.T.

MATÉRIEL ET MÉTHODES

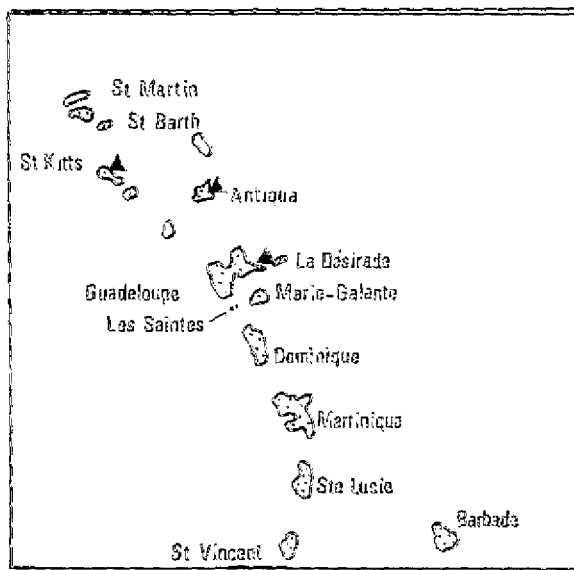
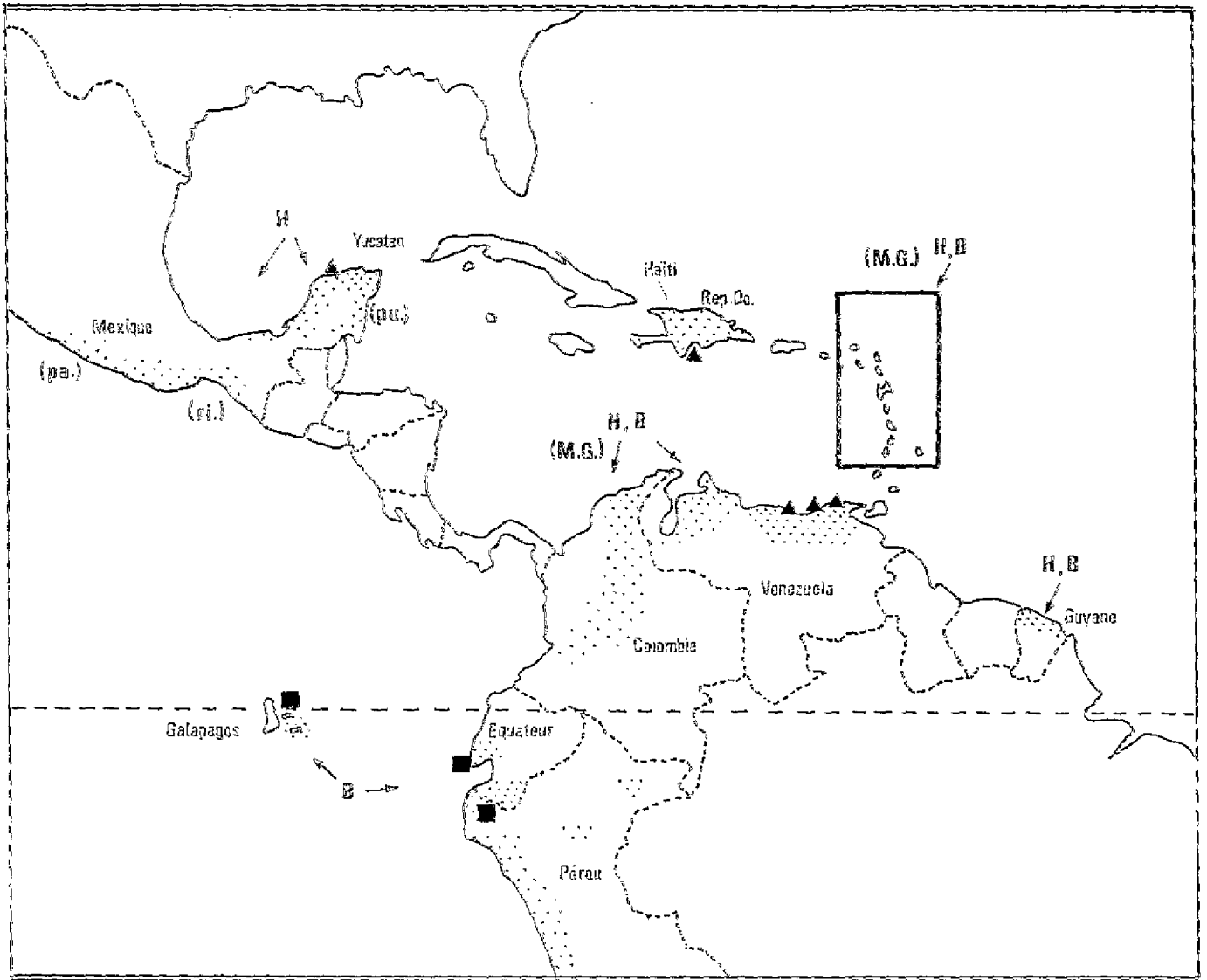
Les collections

Les prospections effectuées sous l'égide de l'I.B.P.G.R. ou C.I.R.P.G. (Conseil International des Ressources Phyto-génétiques) dans la zone de dispersion des 2 allotétraploïdes *G. hirsutum* et *G. barbadense*, l'Amérique Latine, ont réuni plus de 1 000 accessions différentes recouvrant à la fois des formes sauvages et spontanées des 2 espèces.

Les formes sauvages sont issues de populations de tailles très variables, d'une dizaine à un millier de plantes, fréquemment adaptées aux sols halophytes des zones littorales. Au point de vue des caractéristiques de leur fibre et de leur croissance, ces cotonniers représentent des types « primitifs ». Ceux de l'espèce *G. hirsutum* sont trouvés essentiellement autour de la mer des Caraïbes (Mexique, îles de l'Arc Antillais, Vénézuéla), tandis que les cotonniers de l'espèce

G. barbadense poussent au Nord du Pérou (région de Tumbes), en Equateur continental et sur les îles Galapagos où on leur confère selon les auteurs le nom de variété ou d'espèce *darwini* (fig. 1).

A l'état spontané, les cotonniers subsistent souvent isolément sur les bords de route et dans les jardins. Les plantes appartenant à l'espèce *G. barbadense* sont préférentiellement adaptées aux régions humides, mais les formes spontanées des 2 espèces sont en sympatrie sur une large part de leur aire de distribution, la Colombie, le Vénézuéla et l'Arc Antillais. On trouve également l'espèce *G. barbadense* en Amazonie et au Pérou, l'espèce *G. hirsutum* au Mexique. Chez cette dernière, on distingue six races : *morrillii*, *latifolium*, *Marie-Galante*, *punctatum*, *palmeri*, *richmondii*, dont seules les 4 dernières ont été suffisamment collectées (fig. 1).



- | | | | |
|------|---|---|---|
| H | : <i>G. hirsutum</i> | B | : <i>G. barbadense</i> |
| ▲ | : population sauvage wild population | ■ | : population sauvage wild population |
| M.G. | : Marie-Galante | | |
| pu. | : <i>punctatum</i> | | |
| pa. | : <i>palmeri</i> | | |
| ri. | : <i>richmondii</i> | | |

Figure 1
Régions prospectées pour la collecte de cotonniers sauvages et subspontanés des espèces *G. hirsutum* et *G. barbadense* (années 1980-83).

Regions prospected to collect wild cottons and landraces of *G. hirsutum* and *G. barbadense* (1980-83 years).

Echantillonnage

Les accessions à étudier ont été choisies afin d'être les plus représentatives possible de l'aire de répartition des 2 espèces. Au total, 260 accessions ont été analysées par électrophorèse (tabl. 1). Elles se répartissent ainsi :

- 54 cotonniers sauvages issus de 3 populations de *G. barbadense* et 5 de *G. hirsutum* ;
- 139 cotonniers subsponsanés regroupés par région géographique ;
- 67 variétés cultivées de *G. hirsutum*, provenant de plusieurs pays producteurs (pays africains, U.S.A., Chine, U.R.S.S., Bulgarie, Iran, etc.) et appartenant à différents fonds génétiques de départ de sélection (américains : Acala, Old Stormproof... et africains : Allen, Triumph, N'Kourala, H.A.R., ...).

Sauf dans le cas des cultivars quasiment fixés ou une seule graine a été analysée, l'étude de chaque échantillon a porté sur 3 graines afin de comprendre le déterminisme génétique des marqueurs enzymatiques et d'estimer approximativement le niveau d'hétérozygotie des descendance.

Méthodes d'électrophorèse

Les techniques employées (BOURDON, 1984) sont rappelées ici brièvement.

Les extraits sont préparés à partir de graines préimbibées (une graine par analyse) et broyées dans un tampon Tris KCl 0,1 M (pH = 7,3), additionné de cystéine 20 mM. Trois systèmes de migration sont utilisés, 2 sur gel d'amidon à 12,5% (POULIK, 1957 ; SHAW et PRASAD, 1970), l'autre sur gel d'acrylamide, avec un tampon Tris Borate (0,0089 M) EDTA (0,025 M) à pH = 8,3.

Des essais préliminaires portant sur une trentaine de systèmes enzymatiques ont conduit à en retenir 9 en raison de la bonne résolution de leurs zymogrammes. La méthode de migration optimale pour chacun de ces enzymes est exposée dans le tableau 2. Les solutions de révélation utilisées ont été empruntées à :

- TROUSLOT et SECOND (1980) : alcool déshydrogénase (ADH), malate déshydrogénase (MDH), isocitrate déshydrogénase (IDH), leucine aminopeptidase (LAP), phospho-

glucomutase (PGM), phosphoglucose isomérase (PGI), glutamate oxaloacétate transaminase (GOT) ;

- WOODBURY *et al.* (1971) : catalase (CAT) ;
- CARDY *et al.* (1980) : endopeptidase (Endo).

Interprétation génétique des zymogrammes

Sur les 9 systèmes enzymatiques, 2 sont monomorphes quelle que soit l'origine des individus étudiés (PGI, CAT), tandis que les 7 autres présentent des variations à la fois inter et intraspécifiques.

Des hypothèses de contrôle génétique ont été proposées pour ces enzymes polymorphes (excepté la PGM) à partir de l'observation conjointe des zymogrammes d'espèces diploïdes et tétraploïdes et de l'analyse de quelques descendance autofécondées. Chez les allotétraploïdes, l'activité de chaque système enzymatique est gouvernée par au moins 2 locus (A et A' ou B et B'), que l'on suppose localisés chacun sur l'un des génomes constitutifs A et D. Ces hypothèses sont étayées par l'existence d'homologies alléliques entre ces locus et les locus des espèces diploïdes parentales. Cependant, les allèles de 2 locus homéologues codant eux-mêmes fréquemment pour des molécules de mobilités électrophorétiques identiques, il n'est pas toujours possible de distinguer l'expression individuelle des 2 locus. C'est le cas de l'enzyme IDH, pour lequel les variations alléliques des 2 locus doivent être considérées simultanément (locus IDH A, A').

Pour l'ensemble des 6 enzymes (ADH, IDH, MDH, LAP, GOT, Endo), on distingue 14 locus polymorphes présentant un total de 46 allèles (tabl. 2). Les tests d'indépendance réalisés sur 22 couples concernant 8 locus n'ont mis en évidence aucune relation de *linkage*. Les autres couples n'ayant pu être testés, la diversité des espèces allotétraploïdes a, cependant, été estimée à partir des variations alléliques des 14 locus supposés *a priori* indépendants.

Méthodes statistiques

Plusieurs paramètres sont utilisés pour décrire le polymorphisme :

- P, nombre de locus polymorphes ;
- A, nombre d'allèles ;

TABLEAU 1

Description des accessions sauvages, subsponsanées et cultivées analysées pour l'étude de la diversité génétique des 2 espèces allotétraploïdes *G. hirsutum* et *G. barbadense*.

Description of the wild populations, landraces and cultivated varieties analysed to study genetic diversity in *G. hirsutum* and *G. barbadense*.

| Espèce | Type | | | | |
|----------------------|----------------------|-------------------------|--------------------|----|------------|
| | Populations sauvages | Cotonniers subsponsanés | Variétés cultivées | | |
| <i>G. barbadense</i> | Iles Galapagos | 14 | Arc Antillais | 15 | — |
| | Nord Pérou (Tumbes) | 13 | Guyane | 5 | |
| | (2 populations) | | Vénézuéla | 2 | |
| | | | Colombie | 7 | |
| | | | Pérou | 13 | |
| <i>G. hirsutum</i> | Rép. Dominicaine | 3 | Marie Galante | | U.S.A. 27 |
| | St Kitts | 4 | Vénézuéla | 18 | Afrique 23 |
| | Guadeloupe | 4 | Colombie | 12 | Divers 17 |
| | Vénézuéla | 6 | Arc Antillais | 38 | |
| | Mexique (Yucatan) | 10 | <i>punctatum</i> | 14 | |
| | | | (Mexique) | | |
| | | | <i>richmondii</i> | 6 | |
| | | (Mexique) | | | |
| | | <i>palmeri</i> | 8 | | |
| | | (Mexique) | | | |

TABLEAU 2

Méthodes de migration et hypothèses de déterminisme génétique des enzymes.
Nomenclature des locus et des allèles utilisés dans l'étude de la variabilité génétique des espèces
G. hirsutum et *G. barbadense*.

Migration methods and hypotheses of genetic control of enzymes.
List of the loci and alleles used to study genetic variability in *G. hirsutum* and *G. barbadense*.

| Enzyme | Système de migration optimal | Déterminisme génétique | |
|--------|------------------------------|---------------------------------------|---------------------------------|
| | | Locus | Allèles |
| ADH | Amidon : Shaw Prasad (1970) | Adh A | 10, 20 |
| | | Adh B | 5, 10, 20, 30, 40, 50 |
| IDH | Amidon : Shaw Prasad (1970) | Idh A/A' | 10, 20, 30, 60 |
| MDH | Amidon : Shaw Prasad (1970) | Mdh A | 20, 30, 40 |
| | | Mdh B | 1, 2, 3, 4, 5, 6 |
| | | Mdh C | 10, 0 |
| LAP | Amidon : Shaw Prasad (1970) | Lap A | 20, 40, 60*, 0 |
| | | Lap A' | 60, 70, 0 |
| | | Lap B | 30, 40, 0 |
| | | Lap B' | 50, 60, 70 |
| GOT | Acrylamide : TBE | Got A | 10, 20, 30, 40 [†] , 0 |
| | | Got A' | 40, 50, 60, 0 |
| ENDO | Acrylamide : TBE | Endo A | 10, 30 |
| | | Endo A' | 50, 60, 70 |
| PGM | Amidon : Shaw, Prasad (1970) | Déterminisme génétique non interprété | |
| CAT | Amidon : Poulík (1957) | Monomorphe | |
| PGI | Amidon : Poulík (1957) | Monomorphe | |

* Les allèles Lap A⁶⁰ et Got A⁴⁰ sont communs à 2 locus homéologues.

— A int. (°/°), pourcentage d'allèles en fréquence intermédiaire ($0,15 < p_{ij} < 0,85$) avec p_{ij} fréquence de l'allèle i au locus j ;

— G, nombre de combinaisons alléliques différentes sur les 14 locus ;

— Ho, hétérozygotie observée ;

— Ht, diversité génétique de Nei (1973), telle que :

$$Ht = (\sum_j (1 - \sum_i p_{ij}^2)) / r$$

avec r nombre de locus, égal à 13 (IDH excepté).

Selon la formulation de Nei, ce paramètre peut être estimé soit au niveau intra-population (Hs), soit au niveau

de l'espèce (Ht), tel que :

$$Ht = Hs \text{ moy.} + Dst$$

Dst correspondant à la différenciation entre populations.

Les calculs des fréquences alléliques et de l'hétérozygotie sont effectués en tenant compte de la structure biologique différente des cotonniers spontanés (très souvent isolés) et des cotonniers sauvages (en populations) :

— estimations à partir des génotypes maternels déduits pour les cotonniers spontanés (autogamie obligatoire) ;

— estimations directes sur les descendants dans le cas des cotonniers sauvages (existence d'un taux d'allogamie non négligeable).

RÉSULTATS

Polymorphisme des cotonniers sauvages

Toutes origines confondues, les cotonniers sauvages sont polymorphes pour l'ensemble des 14 locus enzymatiques, totalisant ainsi 44 allèles, soit 2,85 allèles par locus (tabl. 3). L'espèce *G. hirsutum* est sensiblement plus variable avec 11 locus polymorphes, 30 allèles et une diversité totale Ht égale à 0,200 contre les valeurs respectives de 7, 25 et 0,179 chez *G. barbadense*.

L'espèce *G. hirsutum* se caractérise également :

— par une diversité génétique intra-population élevée (Hs moy. = 0,161), due à une forte proportion d'allèles en fréquence intermédiaire ;

— une hétérozygotie observée importante (Ho moy. = 0,060), notamment dans le Yucatan (Ho = 0,135).

Malgré l'effectif restreint des échantillons analysés, des différences de variabilité, apparemment fonctions de la taille réelle des populations, apparaissent sans ambiguïté. Ainsi, les 2 populations étendues du Yucatan et de République Dominicaine montrent des niveaux de polymorphisme comparables (7 et 8 locus polymorphes, 22 et 26 allèles) bien que la seconde soit insuffisamment représentée ($n = 3$). A l'opposé, la population de Guadeloupe, extrêmement réduite, apparaît totalement monomorphe.

Dans l'espèce *G. barbadense*, les populations des 2 origines étudiées (les Galapagos et région de Tumbes au Pérou) manifestent des degrés de polymorphisme équivalents mais nettement inférieurs à ceux des populations de *G. hirsutum* (tabl. 3) :

— Hs moy. = 0,060 ;

— hétérozygotie individuelle quasiment nulle (Ho = 0,020).

TABLEAU 3

Polymorphisme enzymatique des cotonniers sauvages chez les espèces *G. hirsutum* et *G. barbadense*.
Enzymatic polymorphism of wild cottons in *G. hirsutum* and *G. barbadense*.

| Espèces et origines géographiques | Paramètres | | | A int. % | Hs | Ht | Ho |
|-----------------------------------|------------|----|----|----------|------------------|-------|------------------|
| | n | P | A | | | | |
| Rép. Dominicaine | 3 | 7 | 22 | 0,43 | 0,167 | 0,200 | 0,021 |
| St Kitts | 4 | 5 | 19 | | | | |
| Guadeloupe | 4 | 0 | 14 | | | | |
| Vénézuéla | 6 | 3 | 17 | 0,25 | 0,075 | | 0,013 |
| Mexique (Yucatan) | 10 | 8 | 26 | 0,63 | 0,242 | | 0,135 |
| Total <i>G. hirsutum</i> | 27 | 11 | 30 | 0,41 | Hs moy. 0,161 | | Ho moy. 0,060 |
| Iles Galapagos | 14 | 3 | 13 | 0,12 | 0,070 | | 0 |
| Tumbes | 13 | 5 | 19 | 0,11 | 0,051 | | 0,03 |
| Total <i>G. barbadense</i> | 27 | 7 | 25 | 0,39 | Hs moy. 0,060 | 0,179 | Ho moy. 0,02 |

n : nombre d'individus ; P : nombre de locus polymorphes ; A : nombre total d'allèles ; A int. % : proportion d'allèles en fréquence intermédiaire ($0,15 < p < 0,85$) ; Ht (Hs) : diversité génétique calculée totale (intra-population) ; Ho : hétérozygotie observée.

Les paramètres A int. %, Ht, Ho sont estimés sur 13 locus (ldh A A' exclu).

Une analyse factorielle des correspondances (AFC) (BENZECRI, 1973), intégrant les variations de fréquences inter et intralocus de 17 allèles ($0,10 < p < 0,90$) appartenant aux 9 locus les plus polymorphes, représente la structure génétique des cotonniers sauvages des 2 espèces (fig. 2a). Elle met en évidence, de façon représentative (54,4 % de la diversité totale sur le 1^{er} plan), une structuration en 3 groupes : l'ensemble de l'espèce *G. hirsutum*, les cotonniers du Pérou, les cotonniers de l'archipel des Galapagos. D'autre part, elle conforte les conclusions précédentes. *G. hirsutum* apparaît très polymorphe avec 19 combinaisons alléliques différentes, tandis que les 2 origines de *G. barbadense* sont peu variables (6 et 4 combinaisons alléliques).

La séparation des 2 espèces sur le plan correspond à une différenciation génétique importante montrée par des déséquilibres de fréquences alléliques à 7 locus (Adh A, Adh B, Mdh A, Mdh B, Lap A, Lap B', Endo A') et l'existence de 16 allèles spécifiques à *G. hirsutum* et de 10 à *G. barbadense*.

A un niveau intraspécifique, aucune structure marquée ne se dégage dans l'espèce *G. hirsutum*. Les cotonniers du Yucatan apparaissent toutefois les plus éloignés de l'autre espèce sur le 1^{er} plan de l'AFC : ils accumulent, en effet, l'ensemble des allèles spécifiques à *G. hirsutum*. En revanche, une forte différenciation génétique, responsable de la diversité totale importante de *G. barbadense* (Ht = 0,179), existe entre les cotonniers des Galapagos et ceux du Pérou : 4 locus (Adh B, Mdh B, Lap A, Lap B') présentent un ou plusieurs allèles spécifiques à chacune de ces origines.

Polymorphisme et origine des cotonniers subspontanés

Espèce *G. barbadense*.

Comparés aux populations sauvages de *G. barbadense*, les cotonniers subspontanés de cette espèce manifestent une nette diminution de leur diversité génétique, essentiellement sensible sur leur taux de polymorphisme (P = 4). La diversité totale également faible (Ht = 0,052) a une valeur identique à la diversité intrapopulation (Hs moy. = 0,051). Cela

traduit une absence de différenciation entre régions géographiques (tabl. 4). D'autre part, ces cotonniers apparaissent génétiquement très proches des populations sauvages du Pérou, étant fixés ou quasiment fixés à chaque locus, pour le même allèle (fig. 2b).

Espèce *G. hirsutum*.

De la même façon que dans l'espèce *G. barbadense*, une nette diminution de la diversité génétique (Hs moy. = 0,084) et de l'hétérozygotie individuelle (Ho moy. = 0,025) est constatée au sein des races subspontanées de *G. hirsutum* (tabl. 4).

La race *Marie-Galante* présente toutefois un polymorphisme plus important : 31 allèles, 12 locus polymorphes, Ht égale à 0,145. La valeur de ces paramètres doit sans doute être relativisée par l'effectif de l'échantillon analysé (n = 63) qui a permis la collecte de nombreux allèles rares (11 en fréquence < 0,05). Mais le nombre de combinaisons alléliques présentes dans chacune des régions où se trouve la race *Marie-Galante* (G = 11 à 15) confirme également sa plus grande variabilité vis-à-vis des races du Mexique où ne coexistent jamais plus de 5 combinaisons alléliques. De plus, il apparaît que la Colombie et le Vénézuéla sont les régions les plus polymorphes avec 23 combinaisons alléliques différentes pour 30 individus tandis qu'il existe une combinaison allélique très majoritaire (15 individus sur 38) dans l'Arc Antillais.

L'ensemble des cotonniers subspontanés de *G. hirsutum* manifeste, contrairement à *G. barbadense*, une diversité totale importante (Ht = 0,187). Elle est due à la différenciation existant entre races géographiques, chacune d'elles étant caractérisée par une combinaison allélique majoritaire sur 5 locus ; la nature des allèles observés oppose surtout les races *Marie-Galante* et *punctatum* (tabl. 5).

Une AFC réalisée sur 74 accessions subspontanées de *G. hirsutum* et prenant en compte les variations de 11 allèles aux 7 locus polymorphes ($0,10 < p < 0,90$) illustre parfaitement ces résultats (fig. 3a) :

- la nette différenciation sur l'axe I (41,3 %) entre les 3 races mexicaines et la race *Marie-Galante* ;
- le faible polymorphisme des races *punctatum*, *richmondii* et *palmeri* ;

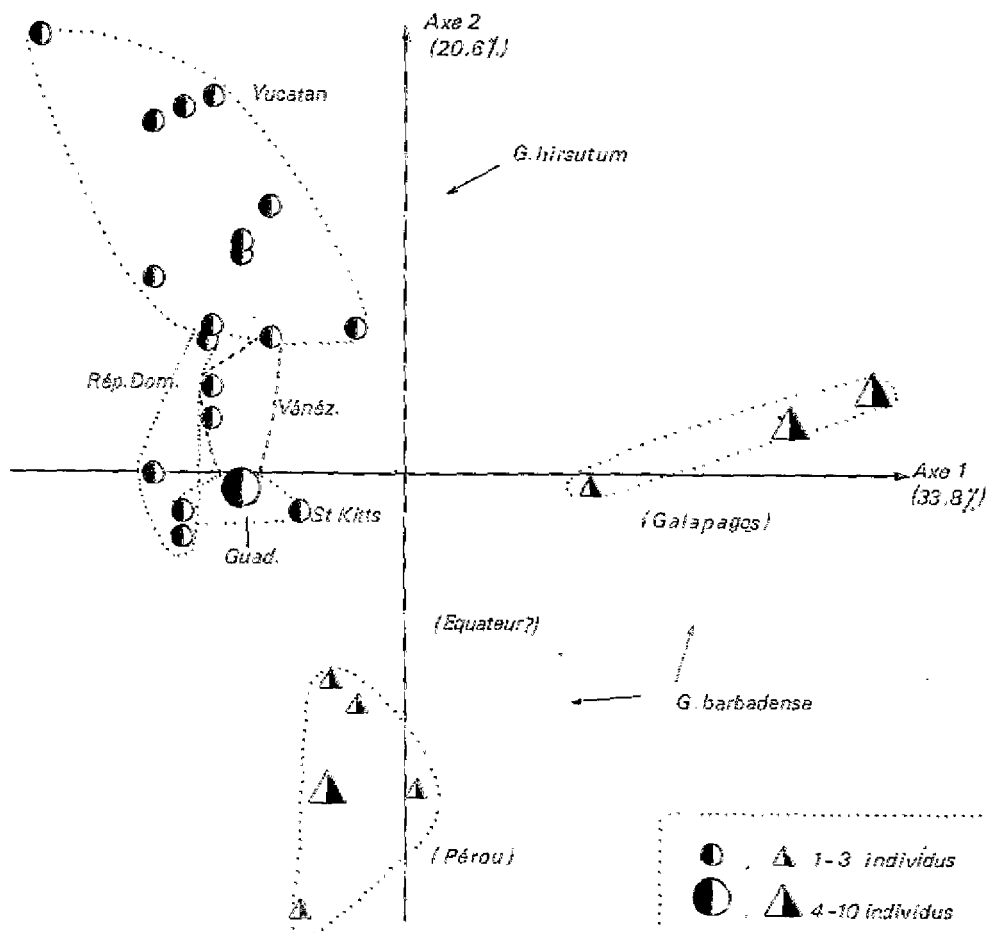


Figure 2a

Organisation génétique des cotonniers sauvages des deux espèces *G. hirsutum* ● et *G. barbadense* ▲. Analyse factorielle des correspondances sur les variations de 17 allèles à 9 locus chez 54 individus.

Genetic organization of the wild cottons of *G. hirsutum* ● and *G. barbadense* ▲. Factor correspondence analysis on the variations of 17 alleles at 9 loci in 54 individuals.

TABLEAU 4
Polymorphisme enzymatique des cotonniers subspontanés chez les espèces *G. hirsutum* et *G. barbadense*.
Enzymatic polymorphism of landraces in *G. hirsutum* and *G. barbadense*.

| Paramètres | n | P | A | A. int. % | Hs | Ht | Ho | G |
|--|-------------------|----|----|-----------|---------------|-------|---------------|-------------------|
| Races ou origines géographiques | | | | | | | | |
| Marie-Galante (Colombie + Vénézuéla + Arc Antillais) | 68 (12 + 18 + 38) | 12 | 31 | 0,30 | 0,145 | | 0,070 | 31 (11 + 14 + 15) |
| <i>punctatum</i> | 14 | 5 | 21 | 0,10 | 0,080 | | 0,030 | 5 |
| <i>richmondii</i> | 6 | 3 | 18 | 0,37 | 0,080 | | 0 | 3 |
| <i>palmeri</i> | 8 | 2 | 17 | 0 | 0,033 | | 0 | 2 |
| Total <i>G. hirsutum</i> | 97 | 12 | 35 | 0,34 | Hs moy. 0,084 | 0,187 | Ho moy. 0,025 | 40 |
| Arc Antillais | 15 | 1 | 15 | 0,14 | 0,038 | | 0,010 | 3 |
| Pérou | 13 | 2 | 16 | 0,13 | 0,047 | | 0,020 | 5 |
| Colombie/Guyane | 14 | 4 | 18 | 0,12 | 0,067 | | 0,010 | 5 |
| Vénézuéla | | | | | | | | |
| Total <i>G. barbadense</i> | 42 | 4 | 18 | 0,125 | Hs moy. 0,051 | 0,052 | Ho moy. 0,013 | 7 |

G : nombre de combinaisons alléliques sur les 14 locus (autres paramètres identiques à ceux du tableau 3).

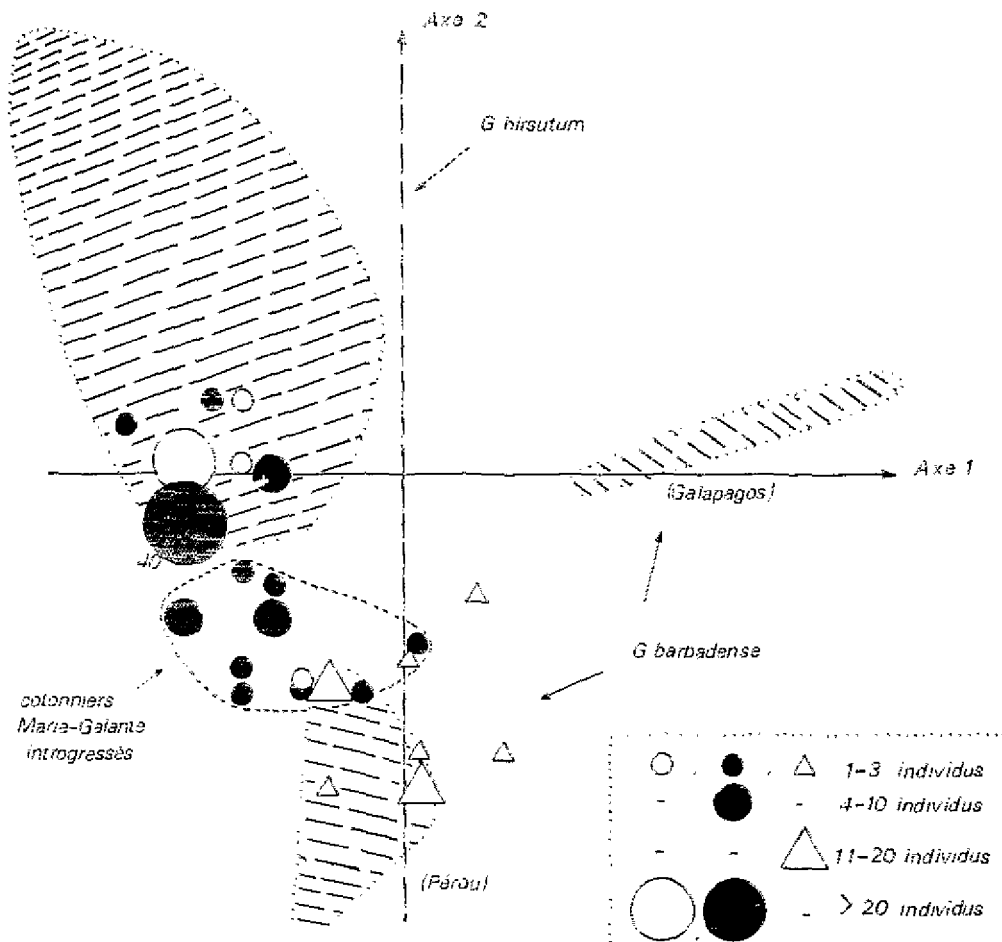


Figure 2b

Origine péruvienne des cotonniers subspontanés de l'espèce *G. barbadense*.

Introgressions génétiques de l'espèce *G. barbadense* chez certains cotonniers de la race *Marie-Galante*.

Projection en individus supplémentaires des cotonniers subspontanés des espèces *G. barbadense* Δ et *G. hirsutum*, races *Marie-Galante* \bullet et mexicaines \circ .

/// : diversité des populations sauvages.

Peruvian origin of *G. barbadense* landraces.

Genetic introgressions of *G. barbadense* species in some *Marie-Galante* cottons.

Projection as additional individuals of the landraces of *G. barbadense* Δ and *G. hirsutum*, *Marie-Galante* \bullet and Mexican races

/// : diversity in wild populations.

TABLEAU 5
 Combinaisons alléliques majoritaires à 5 locus enzymatiques chez les races de *G. hirsutum*.
 Prépondérant alleles combinations in *G. hirsutum* races.

| | Idh A | Mdh C | Lap B | Got A | Got A' |
|----------------------|-------------|-------|-----------------|-------|--------|
| <i>Marie-Galante</i> | 30 30 60 60 | 0 0 | 30 30 140 40 | 20 20 | 60 60 |
| <i>punctatum</i> | 20 20 30 30 | 10 10 | 30 30 | 30 30 | 50 50 |
| <i>palmeri</i> | 20 20 30 30 | 0 0 | 0 0 | 20 20 | 50 50 |
| <i>richmondii</i> | 20 20 30 30 | 10 10 | 30 30 | 20 20 | 50 50 |

— au sein de la race *Marie-Galante*, la grande variabilité des cotonniers d'origine colombienne et vénézuélienne vis-à-vis de ceux de l'Arc Antillais.

La comparaison des cotonniers subspontanés et sauvages est effectuée en projetant ces derniers en individus supplémentaires sur l'AFC (fig. 3a). Leur absence de structure,

leur polymorphisme et leur hétérozygotie aux 5 locus différenciant les races géographiques se traduisent par leur position centrale sur le plan. Les cotonniers sauvages cumulent, en effet, l'ensemble des allèles présents chez les cotonniers subspontanés à ces 5 locus à l'exception de l'allèle Lap B0 spécifique de la race *palmeri*.

La prise en compte de la totalité des allèles observés chez *G. hirsutum* (aux 14 locus) montre cependant que 10 allèles sont absents des populations sauvages et rencontrés exclusivement chez la race *Marie-Galante*, bien qu'en faible fréquence pour 9 d'entre eux ($p < 0,05$).

Trois de ces allèles (Idh A 10, Mdh B 2, Endo A' 50) sont des allèles communs de l'espèce *G. barbadense*. Projétés sur l'AFC représentant la structure génétique des cotonniers sauvages, 19 individus de la race *Marie-Galante*, porteurs d'au moins 1 de ces 3 allèles, se positionnent ainsi de façon intermédiaire entre les cotonniers sauvages de l'espèce *G. hirsutum* et ceux de l'espèce *G. barbadense* (fig. 2b).

Polymorphisme des variétés cultivées

Avec 4 locus polymorphes pour 2 allèles chacun, les 67 variétés cultivées forment un ensemble peu diversifié.

En considérant les différentes combinaisons alléliques existantes, on dénombre 10 génotypes dont 2 largement majoritaires présents chez 47 variétés (annexe 3). La répartition des cultivars selon ces génotypes montre, d'autre part, une hétérogénéité relativement plus grande du groupe africain (8 génotypes) par rapport à l'ensemble des variétés américaines (5 génotypes). On ne peut cependant différencier les divers *pools* d'origine des variétés (Allen, Acala, Triumph, N'Kourala, H.A.R., etc.).

La composition allélique des cultivars constitue un sous-ensemble de celle des races mexicaines. Cette proximité génétique est bien illustrée par la projection des 10 génotypes trouvés chez les variétés sur l'AFC représentant la diversité des cotonniers subspontanés (fig. 3b).

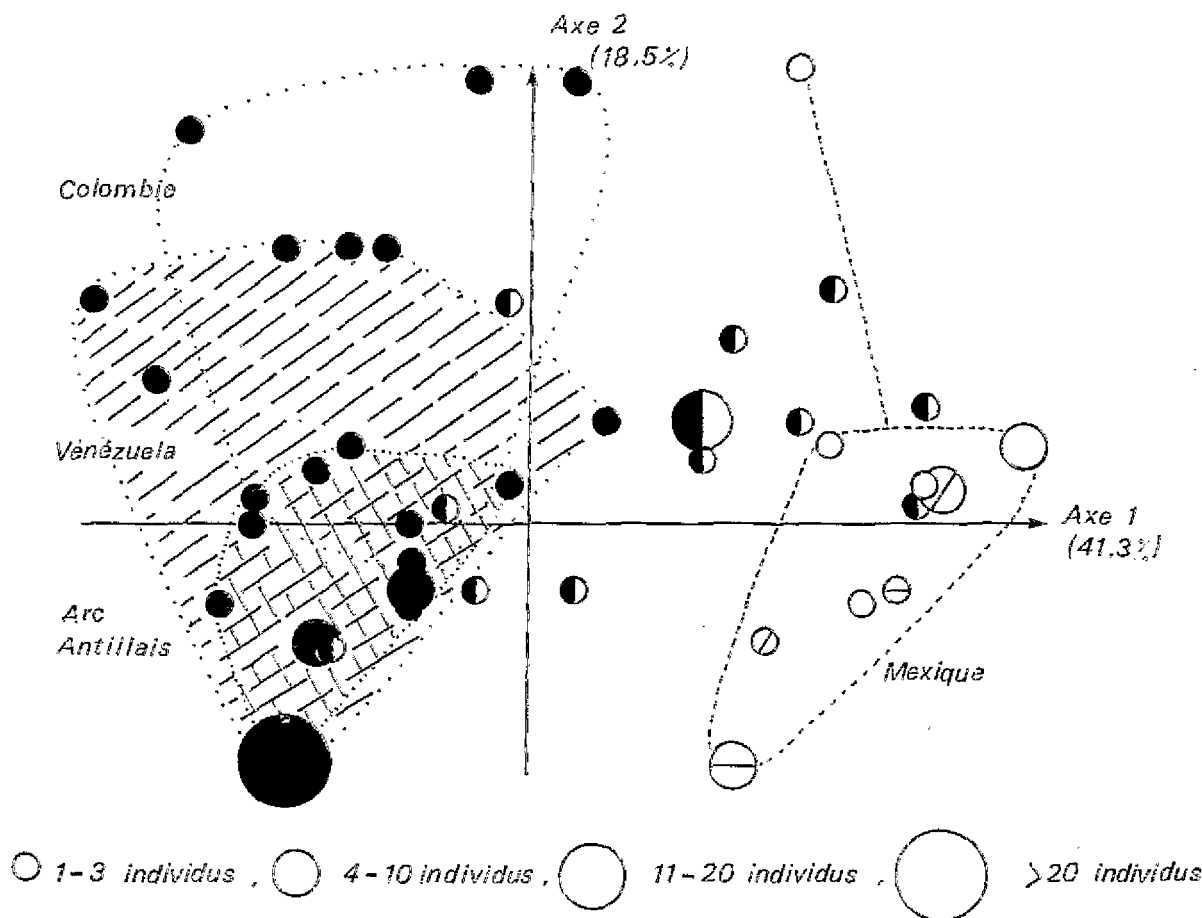


Figure 3a

G. hirsutum : — polymorphisme et différenciation génétique des races subspontanées (*Marie-Galante* ●, *punctatum* ○, *palmeri* ⊖ and *richmondii* ⊕) ; — comparaison avec le polymorphisme des cotonniers sauvages ●.

Analyse factorielle des correspondances sur les variations de 11 allèles à 7 locus chez les races subspontanées (96 individus). Projection en individus supplémentaires des cotonniers sauvages ●.

G. hirsutum : — polymorphism and genetic differentiation of landraces (*Marie-Galante* ●, *punctatum* ○, *palmeri* ⊖, and *richmondii* ⊕) ; — comparison with the polymorphism of wild cottons ●.

Factor correspondence analysis on the variations of 11 alleles at 7 loci in landraces (96 individuals). Projection of wild cottons as additional individuals.

l'intermédiaire du courant de Humboldt, comme c'est le cas pour une grande partie de la flore équatorienne. Quant aux cotonniers du Nord du Pérou, leur situation sur la zone marginale de la distribution des cotonniers expliquerait l'importante dérive génétique qu'ils ont subie. Dans ces conditions, les populations d'Equateur non étudiées ici devraient se situer logiquement sur l'APC des cotonniers sauvages, en position intermédiaire entre les cotonniers du Pérou et ceux des Galapagos (fig. 2a).

La domestication : origine des cotonniers spontanés

En montrant que la différenciation génétique entre *G. hirsutum* et *G. barbadense* préexistait au niveau des populations sauvages, l'électrophorèse confirme que la domestication du cotonnier a eu lieu de façon indépendante dans chacune des 2 espèces comme l'indiquaient les données archéologiques et historiques (STEPHENS, 1973).

Espèce *G. barbadense*.

Deux hypothèses sont proposées par STEPHENS (1973) pour expliquer l'origine des cotonniers cultivés de l'espèce *G. barbadense* :

— une domestication unique à partir des populations sauvages de l'Equateur et du Pérou où ils se seraient répandus et d'où ils auraient diffusé vers le Bassin Amazonien à travers la Cordillère des Andes par la brèche du Maraïon ;

— deux domestications indépendantes, l'une au Pérou, l'autre en Amazonie à partir de populations sauvages dont on suppose l'existence dans une région située à la frontière de la Bolivie et du Paraguay.

Or, si la collection de cotonniers spontanés de l'espèce *G. barbadense* manifeste une diversité phénotypique importante touchant essentiellement aux caractères de la graine (libre ou soudée) et de la fibre (coloration blanche, kaki, brune, acajou), son polymorphisme enzymatique est, en revanche, très faible et ne permet de dégager aucune structure génétique de l'ensemble. D'autre part, les cotonniers spontanés apparaissent génétiquement très semblables à ceux des populations sauvages du Pérou, étant quasiment fixés pour les allèles les plus fréquents dans cette région.

Ces résultats corroborent donc plutôt la première hypothèse avancée par STEPHENS. De plus, ils semblent montrer, jusqu'à plus ample information par l'analyse des populations équatoriennes, que les cotonniers spontanés dérivent uniquement du pool génétique déjà restreint du Nord du Pérou.

Le contraste frappant qui apparaît ainsi entre variations enzymatiques et variations morphologiques, observé également sur les races d'amarante en Inde (JAIN *et al.*, 1980) montre que la sélection humaine a pu réaliser une grande diversification phénotypique tandis que le polymorphisme enzymatique a évolué parallèlement à cette sélection, essentiellement sous l'action de la dérive génétique.

Espèce *G. hirsutum*.

Plusieurs hypothèses sont également formulées à propos de l'origine des races spontanées de *G. hirsutum*. Certains admettent l'idée d'un centre unique de domestication de l'espèce au Mexique, dans le golfe du Tehuantepec. Diverses données archéologiques ou historiques établissent clairement que le cotonnier y a été domestiqué il y a environ 5 000 ans ; toutes les races spontanées seraient alors issues des ancêtres de la population sauvage actuelle du Yucatan. Selon d'autres, la race *Marie-Galante* aurait été domestiquée de façon indépendante en Colombie en se basant sur la découverte d'anciens établissements humains datés de 3 000 ans avant J.C. et sur la grande variabilité phénotypique observée dans ce pays (STEPHENS, 1973).

D'un point de vue enzymatique, la différenciation génétique des 4 races caractérisées chacune par une combinai-

son allélique sur 3 locus et mettant essentiellement en évidence un clivage entre la race *Marie-Galante* et les races mexicaines, est peu compatible avec une origine commune de l'ensemble des cotonniers spontanés. Cependant, l'hypothèse selon laquelle la race *punctatum* est la relique d'une population de transition entre la population sauvage ancestrale et les autres races spontanées (ANO et SCHWENDEMAN, 1983) reste acceptable si on ne considère que les races mexicaines. La nature peu polymorphe de *punctatum* semble alors étonnante, mais on ne peut exclure que les cotonniers observés actuellement soient les descendants d'un ensemble autrefois beaucoup plus variable. Quant à la race *Marie-Galante*, elle serait issue d'un processus de domestication géographiquement indépendant, mais dont la localisation reste imprécise dans l'état actuel des données.

Une telle différenciation génétique entre les races mexicaines et la race *Marie-Galante* n'était pas prévisible à partir des observations phénotypiques du matériel spontané. En effet, les données des prospections montrent que la distinction des races géographiques n'est pas évidente, exception faite de la race *palmeri* aux feuilles laciniées.

La race *Marie-Galante* est la seule à manifester sur le plan génétique un polymorphisme important sous la forme de nombreux allèles rares. Bien qu'on ne puisse négliger l'influence des biais d'échantillonnage, sa répartition beaucoup plus vaste et son état parfois quasi-spontané et parapatric des populations sauvages peuvent également expliquer sa diversité génétique. Plusieurs arguments suggèrent, en effet, que la race *Marie-Galante* a subi ou subit encore des introgressions génétiques ponctuelles en provenance des cotonniers sauvages de *G. hirsutum* :

— l'existence des 7 allèles spécifiques de *Marie-Galante*, mais que l'on peut supposer exister chez les cotonniers sauvages sans avoir été détectés en raison de leur faible fréquence ;

— la présence, très ponctuelle chez *Marie-Galante* (3 individus), de 2 allèles trouvés, au contraire, communément dans les populations sauvages ;

— au Venezuela, la position géographique de 4 des 6 individus porteurs de ces allèles, soit au voisinage de populations sauvages (région de Piritu), soit dans des régions où l'existence de tels cotonniers est suspectée (Péninsule de Paraguana, Ile de Margarita) ;

— la détection plus récente de 2 allèles normalement caractéristiques des cotonniers sauvages chez 11 % des cotonniers *Marie-Galante* de la République Dominicaine et d'Haïti où se trouve une population sauvage.

Il apparaît sans ambiguïté qu'une part de la variabilité de la race *Marie-Galante* résulterait aussi d'hybridations avec l'espèce *G. barbadense* ; 3 allèles spécifiques de cette dernière ont été observés en fréquences peu élevées chez *Marie-Galante*. Plusieurs auteurs ont d'ailleurs déjà discuté de l'occurrence de ces introgressions génétiques. Certains vont jusqu'à proposer une origine hybride des cotonniers *Marie-Galante* par croisements entre des cotonniers *G. barbadense* spontanés et des formes sauvages de *G. hirsutum* du Venezuela (ANO *et al.*, 1983) ou des Antilles (STEPHENS, 1967). D'autres s'en tiennent à des croisements plus ponctuels entre la race *Marie-Galante* déjà différenciée et des plants de *G. barbadense* lors de cultures traditionnelles où les 2 espèces étaient voisines.

Les résultats présents concordent avec cette dernière hypothèse d'introgressions secondaires, les allèles introgressifs étant toujours trouvés en faible fréquence chez *Marie-Galante*.

Actuellement, bien que l'espèce *G. barbadense* et la race *Marie-Galante* soient toujours sympatriques en Colombie, au Venezuela et dans l'Arc Antillais, il ne semble pas que ces croisements se produisent encore du fait de l'abandon de leur culture au profit des variétés modernes et de leurs exigences écologiques propres trop différentes. Quoiqu'il en soit, l'ensemble de ces observations montrant le caractère doublement introgressé de la race *Marie-Galante*, met

en valeur l'intérêt de ce matériel pour les ressources génétiques du cotonnier.

Sélection récente : polymorphisme des variétés cultivées

L'ensemble des variétés de cotonnier cultivées actuellement dans le monde a été sélectionné à partir d'introductions régulières de la race *latifolium* du Mexique et du Guatemala vers les Etats-Unis. L'amélioration cotonnière effectuée par l'I.R.C.T. en Afrique a également utilisé des apports de la race *punctatum* ramenée au XVIII^e siècle par les Portugais. On distingue ainsi en Afrique 3 fonds génétiques anciens, Allen, Triumph et N'Kourala, et un plus récent, H.A.R., créé à partir de croisements interspécifiques entre des variétés de *G. hirsutum* et les 2 espèces diploïdes *G. arboreum* (génome A) et *G. raimondii* (génome D).

L'influence de la sélection pratiquée depuis le XVIII^e siècle sur la diversité génétique du matériel variétal ne pourrait être correctement évaluée à l'aide du polymorphisme enzymatique que si l'on pouvait étudier celui des cotonniers spontanés *latifolium* dans leur lieu d'origine

(Mexique et Guatemala). On peut cependant admettre que cette race a la même origine que les autres races mexicaines et que son apparition s'est également accompagnée d'une dérive génétique importante.

La faible diversité des variétés cultivées mise en évidence par l'analyse enzymatique (4 locus polymorphes) reste donc cohérente avec celle de l'une des quelconques races mexicaines étudiées (entre 2 et 5 locus polymorphes). Il semble ainsi que la sélection récente n'ait pas occasionné de forte réduction de la variabilité totale, mais qu'en revanche, les brassages liés à tous les croisements réalisés ont détruit toute structure initiale éventuelle du matériel puisqu'aucune différenciation n'est observée entre les divers fonds génétiques (Allen, Triumph, N'Kourala). Par ailleurs, il n'a pas été possible de relier la diversité un peu plus importante des cultivars africains aux programmes d'hybridations interspécifiques réalisés en Côte-d'Ivoire. En effet, aucun apport génétique spécifique du matériel H.A.R. n'est détecté, les variétés issues de ces programmes ne se distinguant pas des variétés de fonds génétiques purement *G. hirsutum*. Cette absence de caractérisation peut en fait s'expliquer par la composition allélique très voisine des espèces diploïdes *G. arboreum* et *G. raimondii* avec celle de l'espèce *G. hirsutum* (BOURDON, 1984).

CONCLUSION

Tout en confirmant les caractéristiques génétiques déjà connues des 2 espèces de cotonniers cultivés (structure allo-polypléide, différenciation interspécifique, domestication indépendante de chaque espèce, régime de reproduction à autogamie prédominante chez les formes cultivées, base génétique réduite des variétés modernes), l'électrophorèse apporte des données nouvelles sur leur organisation génétique :

- les populations sauvages, notamment celles de *G. hirsutum*, sont très polymorphes et semblent présenter un taux d'allogamie non négligeable. Des conclusions précises sur l'histoire évolutive de chaque espèce nécessiteraient la poursuite d'un certain nombre d'analyses :

- d'une part, sur les populations d'Equateur de *G. barbadense* afin de vérifier que les cotonniers sauvages de cette région peuvent être à l'origine des populations très différenciées du Pérou et des Galapagos,

- d'autre part, sur toutes les populations sauvages de *G. hirsutum* dans la mer des Caraïbes pour juger de la validité des hypothèses de dispersion par les courants marins.

- la diversité génétique des cotonniers spontanés montre que ceux de l'espèce *G. barbadense* (peu polymor-

phes et non structurés) sont issus d'un seul processus de domestication, tandis que, chez *G. hirsutum*, les races mexicaines et *Marie-Galante* (très différenciées) ont des origines différentes :

- cette organisation génétique tirée de l'analyse enzymatique contraste avec les observations morphologiques mettant en évidence une grande diversification phénotypique chez *G. barbadense* et une distinction parfois difficile des races géographiques de *G. hirsutum* :

- la race *Marie-Galante* présente un statut tout à fait particulier, avec un polymorphisme pouvant s'expliquer en partie par des hybridations à la fois inter et intra-spécifiques, situation qui met en valeur l'intérêt de ce matériel en termes de ressources génétiques.

L'impact de ces résultats sur les collectes, la conservation et l'utilisation des ressources génétiques seront discutés ultérieurement. Dès à présent, la possibilité de mettre en évidence un polymorphisme de l'ensemble du matériel, une diversité structurée de façon différente chez les formes sauvages et spontanées et de relier ces données à des processus de domestication, souligne l'intérêt d'utiliser le polymorphisme enzymatique pour l'établissement de stratégies dans le cadre des ressources génétiques.

ANNEXE 1

Fréquences alléliques aux 14 locus enzymatiques chez les cotonniers sauvages des espèces *G. hirsutum* et *G. barbadense*.
Allelic frequencies at 14 enzymatic loci in *G. hirsutum* and *G. barbadense* wild cottons.

| | <i>G. barbadense</i> | | | <i>G. hirsutum</i> | | | | | | Total |
|------------------------|----------------------|----------------|-------|----------------------------|-----------------|----------------|--------------------------------|-----------|-------------------|-------|
| | Iles des Galapagos | Tumbes (Pérou) | Total | République Dominicaine (1) | Saint Kitts (1) | Guadeloupe (1) | Arc Antillais moyenne pondérée | Vénézuéla | Yucatan (Mexique) | |
| • Adh-A ¹⁰ | 0 | 0 | 0 | — | — | — | 0 | 0,333 | 0,283 | 0,205 |
| • A ³⁰ | 1 | 1 | 1 | P | P | P | 1 | 0,667 | 0,717 | 0,795 |
| • Adh-B ⁵ | 0 | 0 | 0 | — | — | — | 0 | 0 | 0,217 | 0,072 |
| • B ¹⁰ | 0,929 | 0 | 0,482 | — | — | — | 0 | 0 | 0 | 0 |
| • B ²⁰ | 0,071 | 1 | 0,518 | P | P | P | 0,955 | 1 | 0,583 | 0,346 |
| • B ⁴⁰ | 0 | 0 | 0 | P | — | — | 0,045 | 0 | 0,200 | 0,082 |
| • Idh-A ¹⁰ | 0 | 0,120 | 0,058 | — | — | — | 0 | 0 | 0 | 0 |
| • A ²⁰ | 0 | 0 | 0 | — | — | — | 0 | 0 | 0,058 | 0,019 |
| • A ³⁰ | 1 | 0,880 | 0,942 | P | P | P | 0,910 | 1 | 0,792 | 0,901 |
| • A ⁵⁰ | 0 | 0 | 0 | P | P | — | 0,090 | 0 | 0,150 | 0,080 |
| • Mdh-A ³⁰ | 1 | 1 | 1 | P | P | — | 0,273 | 0 | 0,667 | 0,313 |
| • A ⁵⁰ | 0 | 0 | 0 | P | P | P | 0,727 | 1 | 0,333 | 0,687 |
| • Mdh-B ¹ | 0 | 0,796 | 0,383 | P | P | P | 1 | 1 | 1 | 1 |
| • B ² | 0 | 0,204 | 0,093 | — | — | — | 0 | 0 | 0 | 0 |
| • B ⁴ | 0,143 | 0 | 0,075 | — | — | — | 0 | 0 | 0 | 0 |
| • B ⁵ | 0,500 | 0 | 0,259 | — | — | — | 0 | 0 | 0 | 0 |
| • B ⁶ | 0,357 | 0 | 0,185 | — | — | — | 0 | 0 | 0 | 0 |
| • Mdh-C ¹⁰ | 1 | 1 | 1 | P | P | P | 0,727 | 1 | 1 | 0,909 |
| • C ⁰ | 0 | 0 | 0 | P | P | — | 0,273 | 0 | 0 | 0,071 |
| • Lap-A ²⁰ | 1 | 0,037 | 0,536 | — | — | — | 0 | 0 | 0,267 | 0,089 |
| • A ⁴⁰ | 0 | 0,963 | 0,464 | P | P | P | 1 | 0,750 | 0,183 | 0,644 |
| • A ⁶⁰ | 0 | 0 | 0 | — | — | — | 0 | 0,250 | 0,483 | 0,244 |
| • A ⁰ | 0 | 0 | 0 | — | — | — | 0 | 0 | 0,067 | 0,023 |
| • Lap-A ⁶⁰ | 0 | 0 | 0 | — | — | — | 0 | 0,083 | 0 | 0,028 |
| • A ⁷⁰ | 1 | 1 | 1 | P | P | P | 1 | 0,917 | 1 | 0,972 |
| • Lap-B ³⁰ | 1 | 1 | 1 | P | P | P | 0,773 | 1 | 1 | 0,924 |
| • B ⁴⁰ | 0 | 0 | 0 | P | P | — | 0,227 | 0 | 0 | 0,076 |
| • Lap-B ⁵⁰ | 1 | 0 | 0,518 | P | P | P | 0,727 | 1 | 0,833 | 0,353 |
| • B ⁶⁰ | 0 | 0,037 | 0,018 | P | P | — | 0,273 | 0 | 0,167 | 0,147 |
| • B ⁷⁰ | 0 | 0,963 | 0,464 | — | — | — | 0 | 0 | 0 | 0 |
| • Got-A ²⁰ | 1 | 0,992 | 0,991 | P | P | P | 1 | 1 | 0,817 | 0,939 |
| • A ³⁰ | 0 | 0 | 0 | — | — | — | 0 | 0 | 0,183 | 0,061 |
| • A ⁴⁰ | 0 | 0,018 | 0,009 | — | — | — | 0 | 0 | 0 | 0 |
| • Got-A ⁴⁰ | 0 | 0 | 0 | P | — | — | 0,136 | 0 | 0 | 0,045 |
| • A ⁵⁰ | 0 | 0 | 0 | P | — | — | 0,091 | 0 | 0,367 | 0,152 |
| • A ⁶⁰ | 1 | 1 | 1 | P | P | P | 0,773 | 1 | 0,633 | 0,803 |
| • Endo-A ¹⁰ | 0,143 | 0 | 0,074 | — | — | — | 0 | 0 | 0 | 0 |
| • A ³⁰ | 0,857 | 1 | 0,926 | P | P | P | 1 | 1 | 1 | 1 |
| • Endo-A ⁵⁰ | 1 | 1 | 1 | — | — | — | 0 | 0 | 0 | 0 |
| • A ⁶⁰ | 0 | 0 | 0 | P | P | P | 1 | 1 | 1 | 1 |

• Allèles pris comme variables dans l'analyse factorielle des correspondances (AFC 1).

(1) Ces 3 populations sont caractérisées simplement par leurs présences alléliques en raison de leur effectif restreint (P = allèle présent). Des fréquences moyennes pondérées ont été calculées pour l'ensemble de l'Arc Antillais.

ANNEXE 2

Fréquences alléliques aux 14 locus enzymatiques chez les cotonniers spontanés des espèces *G. hirsutum* (races *Marie-Galante*, *punctatum*, *palmeri* et *richmondii*) et *G. barbadense*.Allelic frequencies at 14 enzymatic loci in landraces of *G. hirsutum* (*Marie-Galante*, *punctatum*, *palmeri* and *richmondii*) and *G. barbadense*.

| | <i>G. barbadense</i> | | | | <i>G. hirsutum</i> | | | | |
|-----------------------|----------------------|---------------------------|-------|-------|----------------------|------------------|-------------------|----------------|-------|
| | Arc Antillais | Colombie Guyane Venezuela | Pérou | Total | <i>Marie Galante</i> | <i>punctatum</i> | <i>richmondii</i> | <i>palmeri</i> | Total |
| • Adh-A ¹⁰ | 0 | 0 | 0 | 0 | 0,024 | 0 | 0 | 0 | 0,006 |
| A ²⁰ | 1 | 1 | 1 | 1 | 0,976 | 1 | 1 | 1 | 0,994 |
| Adh-B ²⁰ | 1 | 1 | 1 | 1 | 0,312 | 1 | 1 | 1 | 0,953 |
| B ³⁰ | 0 | 0 | 0 | 0 | 0,012 | 0 | 0 | 0 | 0,003 |
| • B ⁴⁰ | 0 | 0 | 0 | 0 | 0,164 | 0 | 0 | 0 | 0,041 |
| B ⁵⁰ | 0 | 0 | 0 | 0 | 0,012 | 0 | 0 | 0 | 0,003 |
| Idh-A ¹⁰ | 0 | 0,036 | 0 | 0,012 | 0,016 | 0 | 0 | 0 | 0,004 |
| A ²⁰ | 0 | 0 | 0 | 0 | 0 | 0,5 | 0,5 | 0,5 | 0,375 |
| A ³⁰ | 1 | 0,964 | 1 | 0,988 | 0,508 | 0,5 | 0,5 | 0,5 | 0,502 |
| A ⁶⁰ | 0 | 0 | 0 | 0 | 0,476 | 0 | 0 | 0 | 0,119 |
| Mdh-A ²⁰ | 1 | 1 | 1 | 1 | 0,933 | 1 | 1 | 1 | 0,997 |
| A ⁴⁰ | 0 | 0 | 0 | 0 | 0,012 | 0 | 0 | 0 | 0,003 |
| • Mdh-B ¹ | 0,5 | 0,571 | 0,385 | 0,435 | 0,639 | 0,929 | 1 | 1 | 0,892 |
| B ² | 0,5 | 0,429 | 0,615 | 0,515 | 0,349 | 0,071 | 0 | 0 | 0,105 |
| B ³ | 0 | 0 | 0 | 0 | 0,012 | 0 | 0 | 0 | 0,003 |
| • Mdh-C ¹ | 1 | 1 | 1 | 1 | 0,150 | 0,929 | 0,833 | 0,125 | 0,509 |
| C ⁰ | 0 | 0 | 0 | 0 | 0,850 | 0,071 | 0,167 | 0,875 | 0,491 |
| Lap-A ²⁰ | 0 | 0,143 | 0,077 | 0,073 | 0 | 0 | 0 | 0 | 0 |
| A ⁴⁰ | 1 | 0,857 | 0,923 | 0,927 | 1 | 1 | 1 | 1 | 1 |
| Lap-A ⁷⁰ | 1 | 1 | 1 | 1 | 0,976 | 1 | 1 | 1 | 0,994 |
| Lap-A ⁹⁰ | 0 | 0 | 0 | 0 | 0,024 | 0 | 0 | 0 | 0,006 |
| • Lap-B ³⁰ | 1 | 1 | 1 | 1 | 0,513 | 1 | 1 | 0 | 0,628 |
| B ⁴⁰ | 0 | 0 | 0 | 0 | 0,437 | 0 | 0 | 0 | 0,122 |
| B ⁰ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0,250 |
| Lap-B ⁵⁰ | 0 | 0 | 0 | 0 | 0,008 | 0,036 | 0,167 | 0 | 0,052 |
| B ⁶⁰ | 0 | 0 | 0 | 0 | 0,992 | 0,964 | 0,833 | 1 | 0,947 |
| B ⁷⁰ | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 |
| Got-A ¹⁰ | 0 | 0,071 | 0 | 0,024 | 0 | 0 | 0 | 0 | 0 |
| A ²⁰ | 1 | 0,929 | 1 | 0,976 | 0,966 | 0,357 | 1 | 1 | 0,831 |
| A ³⁰ | 0 | 0 | 0 | 0 | 0,017 | 0,643 | 0 | 0 | 0,165 |
| A ⁰ | 0 | 0 | 0 | 0 | 0,017 | 0 | 0 | 0 | 0,004 |
| • Got-A ⁵⁰ | 0 | 0 | 0 | 0 | 0 | 0,358 | 1 | 1 | 0,714 |
| A ⁶⁰ | 1 | 1 | 1 | 1 | 0,996 | 0,071 | 0 | 0 | 0,267 |
| A ⁷⁰ | 0 | 0 | 0 | 0 | 0,004 | 0,071 | 0 | 0 | 0,019 |
| Endo-A ³⁰ | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Endo-A ⁵⁰ | 1 | 1 | 1 | 1 | 0,077 | 0 | 0 | 0 | 0,019 |
| A ⁶⁰ | 0 | 0 | 0 | 0 | 0,923 | 1 | 0,5 | 0,875 | 0,825 |
| A ⁷⁰ | 0 | 0 | 0 | 0 | 0 | 0 | 0,5 | 0,125 | 0,156 |

• Allèles pris comme variables dans l'analyse factorielle des correspondances sur les cotonniers spontanés de *G. hirsutum* (AFC 2).(1) Locus montrant la plus forte différenciation entre les races de *G. hirsutum*.

ANNEXE 3

Classification des 67 variétés cultivées d'après leurs génotypes aux 4 locus enzymatiques polymorphes
(Lap B' : allèles 50, 60 ; Mdh B : allèles 1, 2 ; Got A' : allèles 50, 60 ; Mdh C : allèles 10, 0).

Classification of the 67 cultivated varieties according to their genotypes at 4 polymorphic enzymatic loci.

| Génotypes | 20 20 11 20 | 20 20 10 20 | 20 20 02 20 | 11 20 02 20 | 11 20 10 20 | 02 20 20 20 | 02 20 02 20 | 20 02 20 20 | 02 02 20 20 | 02 20 10 02 |
|---------------|---|---|-------------------------|-------------|---|---|-------------|------------------|-------------|-------------|
| Variétés | | | | | | | | | | |
| Africaines : | DZA 71-39 Allen common IRCO 3028 Banda 2 Allen ancien Samaru 25C UK 54 | N'Kouria-14-E ₁ F 280 gl Reba B 50 | 3 BJA 592 Reba W 296 | — | Pan F; 3492 SRI F; 71 Bouale fregé Allen 159 Allen 333-57 HAR 444-2-63 HAR 444-2-66 Bou 76 | BAR 67 A 66-29 BP 4-68 | DPMA 61 | J 193 CA 64-4 | — | — |
| Variétés | | | | | | | | | | |
| Américaines : | — | Stoneville 20 Deltapine 61 Acala Del Cerro Coler C.Q. Pronto Stoneville 7A Stoneville 213 Acala Roja Stoneville 7A s.o. LAG-75-26 gl | — | — | Deltapine SL Acala ancien Hopicala | Acala 1517-BE Acala 531 Coler 417 Delcot 277 Mac Nair 1032 Stoneville 5A YCA-H-1332-1 Triumph Big B. | — | C.A.M.D.-H-75-C | — | Dixie King |
| Autres : | | | | | | | | | | |
| — | Sahel 50 C 15-82 S 1622 Jendouba Type 0 Ipeaco SL 71 Ya Pong Hlan Jiang Su Hlan | Reba P 279 | Chine 7 ^o 5 | — | — | Yaramine 50 137 F Y 37 Turkménie Copai 68 Chirpan 433 Chirpan 432 Andaluca Sivunia | Alep 1 | — | — | — |

0 : allèle absent
1 : allèle présent à l'état hétérozygote
2 : allèle présent à l'état homozygote

RÉFÉRENCES BIBLIOGRAPHIQUES

- ANO, G. ; SCHWENDIMAN, J., 1982. — Rapport de mission au Mexique sur la préservation des ressources génétiques du cotonnier. *F.A.O. I.B.P.G.R.*, AGR-PR 3 11, 34 p.
- ANO, G. ; SCHWENDIMAN, J. ; FERSING, J. ; LACAPE M., 1982. — Les cotonniers primitifs *G. hirsutum* race yucatanense de la Pointe des Châteaux en Guadeloupe et l'origine possible des cotonniers tétraploïdes du Nouveau Monde. *Cot. Fib. trop.*, 37, 4, 327-332.
- ANO, G. ; FERSING, J. ; LACAPE M., 1983. — Les cotonniers de l'île de Marie-Galante. *Cot. Fib. trop.*, 38, 2, 201-205.
- BOURDON, C., 1984. — Différenciation génétique inter et intraspécifique dans le genre *Gossypium* L. : Le polymorphisme enzymatique chez des espèces diploïdes et tétraploïdes de cotonnier. *Thèse de 3^e cycle, Univ. Paris XI, Centre d'Orsay, France*, 135 p.
- CARDY, B.J. ; STUBER, C.W. ; GOODMAN, M.M., 1980. — Techniques for starch gel electrophoresis of enzymes from maize (*Zea mays* L.). *Inst. Statistics Mimeogr. Ser.*, 1317, 1-31.
- CHERRY, J.P. ; KATTERMAN, F.R.H. ; ENDRIZZI, J.E., 1970. — Comparative studies of seed proteins of species of *Gossypium* by gel electrophoresis. *Evolution*, 24, 431-447.
- CHERRY, J.P. ; KATTERMAN, F.R.H. ; ENDRIZZI, J.E., 1971. — A comparative study of seed proteins of allopolyploids of *Gossypium* by gel electrophoresis. *Can. J. Genet. Cytol.*, 13, 155-158.
- CHERRY, J.P. ; KATTERMAN, F.R.H. ; ENDRIZZI, J.E., 1972. — Seed esterases, leucine aminopeptidases and catalases of species of the genus *Gossypium*. *T.A.G.*, 12, 213-226.
- DE VIENNE, D., 1984. — Limites et perspectives des marqueurs moléculaires. *Le sélectionneur français*, 33, 35-46.
- ENDRIZZI, J.E. ; TURCOTTE, E.L. ; KOHEL, R.J., 1985. — Genetics, cytology and evolution of *Gossypium*. *Adv. in Genetics*, 23, 271-375.
- FRYXELL, P.A., 1967. — The interpretation of disjunct distributions. *Taxon*, 16, 316-324.
- HANCOCK, J.F., 1982. — Alcohol dehydrogenase isozymes in *Gossypium hirsutum* and its putative diploid progenitors. The biochemical consequences of enzyme multiplicity. *Pl. Syst. Evol.*, 140, 141-149.
- JAIN, S.K. ; WU, L. ; VAIDYA, K.R., 1930. — Levels of morphological and allozyme variation in Indian amaranths : a striking contrast. *J. Heredity*, 71, 283-285.
- JOHNSON, B.L. ; THEIN, M.M., 1970. — Assessment of evolutionary affinities in *Gossypium* by protein electrophoresis. *Amer. J. Bot.*, 57, 9, 1031-1092.
- JOHNSON, B.L., 1975. — *Gossypium palmeri* and a polyphyletic origin of the New World cottons. *Bull. Torrey Bot. Club.*, 102, 340-349.
- NEI, M., 1973. — Analysis of gene diversity in subdivided populations. *Proc. Nat. Acad. Sci. US*, 70, 3321-3323.
- POULIK, M.D., 1957. — Starch gel electrophoresis in a discontinuous system of buffers. *Nature*, 180, 1477-1478.
- SHAW, C.R. ; PRASAD, R., 1970. — Starch gel electrophoresis of enzymes, a compilation of recipes. *Biochem. Genetics*, 4, 297-320.
- STEPHENS, S.G., 1967. — Evolution under domestication of the New World cottons (*Gossypium* spp.) *Ciencia e cultura*, 19, 1, 118-134.
- STEPHENS, S.G. ; PHILLIPS, L.L., 1972. — The history and geographical distribution of a polymorphic system in New World cottons. *Biotropica*, 4, 2, 49-60.
- STEPHENS, S.G., 1973. — Geographical distribution of cultivated cottons relative to probable centers of domestication in the New World. *Genes, Enzymes and Populations*, 16, 239-254.
- SUITER, K.A. ; PARKS, C.R., 1984. — Genetic control and mode of inheritance of allozyme variation in Old World cotton : *G. arboreum* L. and *G. herbaceum* L., *Am. J. Bot. Abstr.*, 71, 5, (1), 91.
- TROUSLOT, P. ; SECOND, G., 1980. — Technique d'électrophorèse en gel d'amidon appliquée à l'étude de quatorze enzymes de riz. In : *Electrophorèse d'enzymes de riz (Oryza spp.) Travaux et documents O.R.S.T.O.M.*, 120, Paris, 38 p.
- WOODBURY, W. ; SPENCER, A.K. ; STAHMANN, M.A., 1971. — An improved procedure using ferricyanide for detecting catalase isoenzymes. *Anal. Biochem.*, 44, 301-305.

Enzymatic polymorphism and genetic organization of two cotton tetraploid cultivated species, *G. hirsutum* and *G. barbadense*

I. Evolution and domestication of both species

C. Bourdon

SUMMARY

The genetic diversity and structure of the two allotetraploid cultivated cotton species, *G. hirsutum* and *G. barbadense*, are studied with the polymorphism of six enzymatic systems that is 14 loci. The analysis is made on wild cottons and landraces of both species, as well as on a sample of cultivated varieties of *G. hirsutum*.

Diversity in wild cottons is very large, especially within a same population, because of a non-negligible allogamy rate. As far as

G. barbadense is concerned, the cottons in Peru and Galapagos Archipelago seem strongly differentiated.

In *G. barbadense* landraces are not very variable (only in *G. hirsutum* they show a large differentiation among geographical races) to which are added, in Marie-Galante, both inter and intra-specific introgression phenomena.

These results are discussed from the evolutionary standpoint supplying new insight on the events that marked the history and domestication of these species.

INTRODUCTION

The genus *Gossypium* L. has a wide geographical distribution and shows a large morphological and cytogenetic variety since it is made up of diploid species, subdivided themselves into 7 genomes (A to G) and of allotetraploid species (genome AD). Protein and enzymatic electrophoresis therefore offers a primary systematic and evolutionary advantage. Two major types of questions were examined by the teams who have so far used these techniques on cotton: the genetic relationships between genomes at the diploid level (CHERRY *et al.*, 1976, 1972; JOHNSON and THEIN, 1970) and the determination of the ancestral species of allotetraploids (CHERRY *et al.*, 1971; JOHNSON,

1975; HANCOCK, 1982). Most of the time, these studies use the electrophoresis of total proteins or non specific enzymes (esterases). Only more recent works use the polymorphism of about ten enzymatic systems to estimate the intraspecific diversity of *G. herbaceum* and *G. arboreum* (SUTTER and PARKS, 1984).

Our study is more precisely related to the genetics resources of *G. hirsutum* and *G. barbadense*, where the analysis of enzymatic polymorphism is involved as complement to the observations made within the framework of I.R.C.T. Genotype Bank.

MATERIALS AND METHODS

Collections

The collecting missions carried out under the aegis of I.B.P.G.R. (International Board for Plant Genetic Resources) in the distribution area of *G. hirsutum* and *G. barbadense*, i.e. latin America, collected more than 1,000 different accessions covering both wild types and landraces of these species.

The wild types derive from populations with highly variable sizes, from about ten to one thousand plants, frequently adapted to the halophyte soils of the coast areas. As far as fiber properties and growth are concerned, these cottons are « primitive » types. *G. hirsutum* cottons are mainly found around the Caribbean sea (Mexico, West Indies, Venezuela) while *G. barbadense* cottons grow in northern Peru (Tumber area), continental Ecuador and in the Galapagos Archipelago where they are given, according to the authors, the name of *darwini* variety or species (Fig. 1).

Landraces are often found isolated on road sides and in gardens. *G. barbadense* plants are preferably adapted to humid areas but the landraces of both species are in sympatry in most of their distribution range, Colombia, Venezuela and West Indies. *G. barbadense* is also found in Amazonia and Peru and *G. hirsutum* in Mexico. There are six different races in *G. hirsutum*: *morrilli*, *latifolium*, *Marie-Galante*, *punctatum*, *palmeri*, *richmondii*, of which only the last four were sufficiently collected.

Sampling

The accessions to study were chosen so as to be as representative as possible of the distribution range of both species. A total of 260 accessions were analysed by electrophoresis (Table 1). They are distributed thus:

- 54 wild cottons derived from 3 and 5 populations of *G. barbadense* and *G. hirsutum* respectively;
- 139 landraces grouped together by geographical region;
- 67 cultivated varieties of *G. hirsutum* originating from various producing countries (Africa, U.S.A., China, U.R.S.S., Bulgaria, Iran, etc.) and belonging to various original genetic pools, either American (Acala, old Storm-proof, etc.) or African (Allen, Triumph, N'Kourala, H.A.R., etc.).

Apart from all the cultivars (almost fixed) where only one seed was analysed, three seeds of each sample were examined to study the genetic control of enzymatic markers and estimate approximately the heterozygosity level of progenies.

Electrophoresis methods

The techniques used (BOURDON, 1984) are briefly reviewed here. Extracts are prepared with pre-impregnated seeds

(one seed per analysis) crushed in a buffer Tris KCl 0.1 M (pH = 7.3) added with cystein 20 mM. Three migration systems are used, two on starch gel at 12.5% (POULIN, 1957; SHAW and PRASAD, 1970), one on acrylamide gel, with a buffer Tris Borate (0.0089 M) EDTA (0.025 M) at pH = 8.3.

Preliminary tests on around thirty enzymatic systems led to choose nine of them, because of the good resolution of their banding patterns. The optimal migration method for each enzyme is shown in Table 2. The staining solutions used were borrowed from:

- TROUSLOT and SECOND (1980), alcohol dehydrogenase (ADH), malate dehydrogenase (MDH), isocitrate dehydrogenase (IDH), leucine aminopeptidase (LAP), phosphoglucosmutase (PGM), phosphoglucose isomerase (PGI), glutamate oxaloacetate transaminase (GOT);
- WOODBURY *et al.* (1971), catalase (CAT);
- CARDY *et al.* (1980), endopeptidase (Endo).

Genetic interpretation of banding patterns

Two of the 9 enzymatic systems are monomorphic, whatever the origin of the individuals studied (PGI, CAT) whereas the other seven exhibit inter and intraspecific variations.

Hypotheses of genetic control were proposed for these polymorphic enzymes (except PGM) on the basis of the joint observation of the banding patterns of diploid and tetraploid species and the analysis of several selfed progenies. Among allotetraploids, the activity of each enzymatic system is governed by at least 2 loci (A and A' or B and B') that are supposed to be each localized on one of the constituent genomes A and D. These hypotheses are supported by the existence of allelic homologies between these loci and those of the parental diploid species. However, as the alleles of two homeologous loci themselves often code for molecules with identical electrophoretic mobilities, it is not always possible to distinguish the individual expression of the 2 loci. This applies to the enzyme IDH, for which the allelic variations of the 2 loci should be simultaneously considered (locus IDH A/A').

For the six enzymes (ADH, IDH, MDH, LAP, GOT, Endo), 14 polymorphic loci are distinguished, showing a total of 46 alleles (Table 2). The independence tests carried out on 22 pairs showed no linkage relation. As the other pairs could not be tested, the diversity in allotetraploid species was however estimated on the basis of the allelic variations of the 14 presumed independent loci.

Statistical methods

Various parameters are used to study genetic diversity:

- P, number of polymorphic loci;
- A, number of alleles;
- A int. (%), percentage of alleles in intermediate fre-

quency ($0.15 < p_{ij} < 0.85$) with p_{ij} frequency of the i th allele at the j th locus ;

— G , number of different allelic combinations on the 14 loci ;

— H_o , observed heterozygosity ;

— H_t , genetic diversity of NEI (1973), such as :

$$H_t = (\sum_j (1 - \sum_i p_{ij}^2)) r,$$

with r number of loci, equal to 13 (IDH being excepted).

According to NEI's formulation, this parameter can be estimated either at the intrapopulation level (H_s) or at the species level (H_t) so that :

$$H_t = \text{mean } H_s + Dst,$$

Dst corresponding to the differentiation between populations.

Allelic frequencies and heterozygosity are calculated by taking into account the different biological structure of landraces (very often isolated) and wild cottons (in populations) :

— estimates with deduced female genotypes for landraces (compulsory autogamy) ;

— direct estimates on progenies for wild cottons (existence of a non-negligible allogamy rate).

RESULTS

Polymorphism of wild cottons

All origins grouped together, wild cottons are polymorphic for all the 14 enzymatic loci, totalling thus 44 alleles, i.e. 2.85 alleles per locus (Table 3). *G. hirsutum* is more variable with 11 polymorphic loci, 30 alleles and a total diversity H_t equal to 0.200 against the respective values of 7, 25 and 0.179 for *G. barbadense*.

G. hirsutum is also characterized by :

— a high intra-population genetic diversity (mean $H_s = 0.161$) due to a high proportion of alleles in intermediate frequency ;

— a high observed heterozygosity (mean $H_o = 0.060$), especially in Yucatan ($H_o = 0.135$).

Despite the small number of samples analysed, differences in variability appear clearly, apparently depending on the real size of the populations. So, the two widespread populations in Yucatan and the Dominican Republic show comparable levels of polymorphism (7 and 8 polymorphic loci, 22 and 26 alleles), although the second is insufficiently represented ($n = 3$). On the opposite, the population of Guadeloupe, extremely small, seems totally monomorphic.

In *G. barbadense*, the populations of the two origins studied (Galapagos Archipelago and Tumbes region in Peru) show equivalent levels of polymorphism, that are lower than those of *G. hirsutum* populations (Table 3) :

— mean $H_s = 0.060$;

— practically nil individual heterozygosity ($H_o = 0.020$).

A factor correspondence analysis (BENZECRI, 1973) is made on the basis of the between and within locus variations of frequencies of 17 alleles ($0.10 < p < 0.90$) belonging to the nine most polymorphic loci, showing the genetic structure of the wild cottons of the two species (Fig. 2a). It clearly points out (54.4 % of total diversity on the first plane) the existence of 3-group structure : the whole species *G. hirsutum*, the cottons from Peru and those of the Galapagos Archipelago. Also, it confirms the previous conclusions. *G. hirsutum* appears to be very polymorphic with 19 different allelic combinations while the two origins of *G. barbadense* are little variable (6 and 4 allelic combinations).

The separation of the two species on the plane corresponds to a large genetic differentiation shown by disequilibria of allelic frequencies at 7 loci (Adh A, Adh B, Mdh A, Mdh B, Lap A, Lap B', Endo A') and by the existence of 16 and 10 specific alleles to *G. hirsutum* and *G. barbadense* respectively.

At the intraspecific level, no marked structure appears in *G. hirsutum*. The cottons from Yucatan seem however the most distant from the other species on the first plane of the factor correspondence analysis : all the alleles specific to *G. hirsutum* are accumulated in this population. On the other hand, a strong genetic differentiation, responsible for the large total diversity of *G. barbadense* ($H_t = 0.179$), exists between the cottons of the Galapagos and those of

Peru : 4 loci (Adh B, Mdh B, Lap A, Lap B') possess one or several alleles specific to each of these origins.

Polymorphism and origin of landraces

G. barbadense

As compared with *G. barbadense* wild populations, landraces show a sharp decrease in their genetic diversity, mainly noticeable on their polymorphism rate ($P = 4$). Total diversity is low ($H_t = 0.052$) and equivalent to within-population diversity (mean $H_s = 0.051$). It expresses absent differentiation between geographic areas (Table 4). Also, these cottons seem genetically very close to the wild populations of Peru, being fixed or almost fixed at each locus, for the same allele (Fig. 2b).

G. hirsutum

As in *G. barbadense*, a sharp decrease in genetic diversity (mean $H_s = 0.034$) and individual heterozygosity (mean $H_o = 0.025$) is observed in *G. hirsutum* landraces (Table 4).

However, *Marie-Galante* shows a larger genetic diversity : 31 alleles, 12 polymorphic loci, $H_t = 0.145$. The value of these parameters should certainly be considered in view of the sample size ($n = 68$) which allowed many rare alleles to be collected (11 in frequency < 0.05). But the number of allelic combinations present in each region where *Marie-Galante* is found ($G = 11$ to 15) also confirms its higher variability as compared with the races of Mexico where never more than 5 allelic combinations co-exist. In addition, Colombia and Venezuela seem to be the most polymorphic regions with 23 different allelic combinations for 30 individuals, while there is a preponderant allelic combination (15 individuals out of 38) in the West Indies.

Contrary to *G. barbadense* all the landraces of *G. hirsutum* show a large total diversity ($H_t = 0.187$). It is due to the differentiation existing between geographic races, being each of them characterized by a preponderant 5-locus genotype ; moreover, *Marie-Galante* and *punctatum* are the most different as they have no allele in common.

A factor correspondence analysis made on 74 accessions of *G. hirsutum* landraces and taking into account the variations of 11 alleles at the 7 polymorphic loci ($0.10 < p < 0.90$) is a perfect illustration of these results (Fig. 3a) :

— the clear-cut differentiation on axis 1 (41.3 %) between the 3 Mexican races and *Marie-Galante* ;

— the low polymorphism of *punctatum*, *richmondii* and *palmeri* races ;

— the large variability within *Marie-Galante* of the cottons from Colombia and Venezuela as compared with those from the West Indies.

Landraces and wild cottons are compared by projecting the latter as additional individuals on the factor correspondence analysis (Fig. 3a). Their absent structure, their poly-

morphism and their heterozygosity at the 5 loci differentiating the geographic races are expressed by their central position on the plane. Wild cottons possess all the alleles present in landraces at these 5 loci except allele Lap B₁ specific to *palmeri* race.

Taking into account all the alleles observed in *G. hirsutum* (at the 14 loci) shows however that 10 alleles are absent in wild populations and exclusively found in *Marie-Galante*, although rarely for 9 of them ($p < 0.03$).

Three of these alleles (Idh A 10, Mdh B2, Endo A'59) are common in *G. barbadense*. Projected on the factor correspondence analysis representing the genetic structure of wild cottons, 19 individuals of *Marie-Galante*, carriers of at least one of these three alleles, are thus placed in intermediate position between *G. hirsutum* and *G. barbadense* wild cottons (Fig. 2b).

DISCUSSION

As the generally neutral behaviour of enzymatic alleles towards selection is now almost unanimously admitted (DE VIENNE, 1984), enzymatic polymorphism is a useful tool to study the evolutionary history of populations (drifts and founder effects due to size reductions, migrations...).

Pre-domestication evolution : dispersion of wild cottons

Despite disagreements among authors (ENDRIZZI *et al.*, 1985), it seems admitted that allotetraploid species have a monophyletic origin and derive from the ancestor of the present wild species *G. mustelinum* discovered in north-eastern Brazil (ANO *et al.*, 1982). Electrophoresis shows here that the phenomena of speciation having led to the appearance of *G. hirsutum* and *G. barbadense* were accompanied by a large genetic differentiation.

The dissemination, concomitant or subsequent to this speciation, of *G. hirsutum* wild populations along the coasts of Venezuela, the Caribbean and Yucatan (Mexico) would be due to the sea streams running alongside the northern coast of South America and entering the West Indies Sea and the Mexico Gulf (STEPHENS, 1957 ; ANO *et al.*, 1982). This hypothesis is also supported by the excellent seed tolerance to immersion in salt water without loss of germination power (FRYNELL, 1967).

The six populations analysed by electrophoresis are relatively undifferentiated and five of them, especially those of Yucatan and the Dominican Republic, exhibit a large genetic diversity. The heterozygosity rate of the plants also indicates that there is in these populations a strong allogamy rate, which is consistent with the frequent exerted stigma of wild cottons. This polymorphism is properly reflected at the morphological level : the large within-population phenotypical diversity and the characters common to all the populations (brown fuzz, exerted stigma, slow growth, bushy aspect, etc.) led to group them under the name of *yucatanense* race (ANO *et al.*, 1982). All these characteristics seem little compatible with a dispersion process involving strong genetic drifts. The monomorphic population of Guadeloupe is the only one which seems to originate in a dissemination by sea streams. While it can be imagined that the repeated supply of seeds by the sea in one region is sufficient to introduce populations of wild cottons on a long term basis, the nature of the polymorphism observed in Yucatan and in the Dominican Republic would rather indicate that these populations are the relics of a distribution of the species that used to be larger in the past. The analysis of an ancestral population of *G. hirsutum* or *G. mustelinum* in Brazil is however lacking to interpret the levels of polymorphism observed in the different populations.

Polymorphism of cultivated varieties

With 4 polymorphic loci for 2 alleles each, the 67 varieties form an almost steady whole. By considering the different existing allelic combinations, 10 genotypes are counted, 2 of which are preponderant in 47 varieties (annex 3). The distribution of the cultivars according to these genotypes also shows the African group (3 genotypes) is relatively more heterogeneous than the American varieties (5 genotypes). The various original pools (Allen, Acala, Triumph, N'Kourala, H.A.R., etc.) cannot be differentiated.

The allelic composition of the cultivars is a subset of that of the Mexican races. This genetic proximity is properly illustrated by the projection of the 10 genotypes found in the varieties on the factor correspondence analysis representing diversity in landraces (Fig. 3b).

In *G. barbadense*, the large genetic differentiation found between the populations of Peru and those of the Galapagos Archipelago and the relatively low genetic diversity existing in each of these regions are so many arguments for a secondary origin of these populations. On the other hand, the size, phenotypical diversity and geographical position of the populations of Ecuador seem to indicate they would be much more polymorphic and could be the ancestral pool the cotton plants of the Galapagos Archipelago and northern Peru would derive from. While the localization of these populations as compared with those of *G. mustelinum* remains difficult to explain, the later dispersion of the cottons towards the Pacific islands could be achieved through the Humboldt stream, as it occurred for most the Ecuadorian flora. As far as the cottons of northern Peru are concerned, their location on the marginal cotton distribution area would explain the considerable genetic drift they experienced. Under these conditions, the populations of Ecuador which were not studied here should logically be placed on the factor correspondence analysis of wild cottons, in an intermediate position between the cottons of Peru and those of the Galapagos Archipelago (Figure 2a).

Domestication : origin of landraces

By showing that the genetic differentiation between *G. hirsutum* and *G. barbadense* pre-existed at the level of wild populations, electrophoresis confirms that cotton domestication occurred in an independent way in each of the two species as indicated by the archeological and historical data (STEPHENS, 1973).

G. barbadense

Two hypotheses are proposed by STEPHENS (1973) to explain the origin of *G. barbadense* cultivated cottons :

— a single domestication from the wild populations of Ecuador and Peru where they would have spread and from where they would have disseminated through the Andes Cordillera by the breach of Marañon ;

— two independent domestications, one in Peru and one in Amazonia from wild populations the existence of which is supposed in a region located on the border of Bolivia and Paraguay.

Whereas the collection of *G. barbadense* landraces exhibits a considerable phenotypical diversity mainly affecting seed properties (free or joined) and fiber characteristics (white, khaki, brown or mahogany colour), its enzymatic polymorphism is very low and does not allow the genetic structure of the whole to be drawn. Besides, the landraces

seem to be genetically identical to those of the wild populations of Peru, being almost fixed for the most frequent alleles in this region.

These results therefore seem to support the first hypothesis proposed by STEPHENS. They also seem to show that, until further information is available by the analysis of the Ecuadorian populations, landraces only derive from the already limited genetic pool of northern Peru.

The striking contrast appearing thus between enzymatic and morphological variations, also observed on amaranth races in India (JAIN *et al.*, 1980) shows that human breeding could achieve a large phenotypical diversification whereas enzymatic polymorphism evolved simultaneously, mainly under the action of genetic drift.

G. hirsutum

Various hypotheses are also proposed about the origin of *G. hirsutum* landraces. Some admit the idea of an only domestication centre in Mexico, in the Gulf of Tehuantepec. Various historical or archaeological data clearly establish that cotton was domesticated there about 5,000 years ago; all the landraces therefore would derive from the ancestors of the present wild population of Yucatan.

Others think that *Marie-Galante* was independently domesticated in Colombia, on the basis of the discovery of old human establishments dated 3,000 years B.C. and of the large phenotypical variability observed in this country (STEPHENS, 1973).

From the enzymatic standpoint, the genetic differentiation of the 4 races each characterized by a 5-locus genotype and mainly stressing a division between *Marie-Galante* and the Mexican races, is little compatible with a common origin of all the landraces. However, the hypothesis according to which the race *punctatum* is the relic of a transitional population between the ancestral wild population and the other landraces (ANO and SCHWENDIMAN, 1982), remains acceptable if only the Mexican races are considered. The little polymorphic nature of *punctatum* seems then surprising but it cannot be excluded that the cottons observed today are the progenies of a whole which used to be much more variable in the past. As far as *Marie-Galante* is concerned, it could be derived from a geographically independent process of domestication, the localization of which remains imprecise with the present data available.

Such a genetic differentiation between the Mexican races and *Marie-Galante* was not predictable from the phenotypical observations on landraces. The data of the collecting missions show that the distinction between geographical races is not obvious, except for *palmeri* race with lacinate leaves.

From the genetic standpoint, *Marie-Galante* is the only race to exhibit a large genetic diversity under the form of many rare alleles. Although the effect of sampling biases cannot be neglected, its wider distribution and sometimes almost spontaneous and parapatric state of wild populations can also explain its genetic diversity. Several arguments seem to indicate that *Marie-Galante* has experienced or still experiences limited genetic introgressions from *G. hirsutum* wild cottons:

- the existence of 7 specific alleles of *Marie-Galante*, but which can be supposed in wild cottons without being detected because of their low frequency;

- the rare occurrence in *Marie-Galante* (3 individuals) of 2 alleles that are on the opposite frequently found in wild populations;

- in Venezuela, the geographical position of 4 of the 6 individuals carrying these alleles, either near wild popula-

tions (region of Piritu) or in regions where the existence of such cottons is suspected (Paraguana peninsula, Margarita Island);

- the more recent detection of 2 alleles normally characteristic of wild cottons in 11 % of the *Marie-Galante* cottons in the Dominican Republic and Haiti where a wild population is found.

It clearly appears that part of the variability of *Marie-Galante* would also result from hybridizations with *G. barbadense*; 3 alleles specific of the latter were observed in relatively low frequencies in *Marie-Galante*. Several authors have already discussed the occurrence of these genetic introgressions. Some even propose a hybrid origin of *Marie-Galante* cottons by crosses between *G. barbadense* landraces and wild forms of *G. hirsutum* from Venezuela (ANO *et al.*, 1983) or the West Indies (STEPHENS, 1967). Others only propose more limited crosses between the already differentiated race *Marie-Galante* and *G. barbadense* plants when they were traditionally grown in close plots.

The present results are in agreement with this latter hypothesis being introgressive alleles always found in low frequency in *Marie-Galante*.

Today, although *G. barbadense* and *Marie-Galante* are still sympatric in Colombia, Venezuela and West Indies, it does not seem that these crosses still occur because these varieties have been replaced by modern varieties and because their specific ecological requirements are highly different. However, these results showing the doubly introgressed nature of *Marie-Galante* emphasize the advantages of this material for cotton genetic resources.

Recent selection : polymorphism of cultivated varieties

All the cotton varieties grown today in the world were selected from regular introductions of *latifolium* race from Mexico and Guatemala to the United States. The cotton improvement works carried out by I.R.C.T. in Africa also used supplies of *punctatum* race brought back by the Portuguese in the XVIIth century. So, there are in Africa three old genetic pools, *i.e.* Allen, Triumph and N'Kourala, and a more recent one, H.A.R., resulting from interspecific crosses between varieties of *G. hirsutum* and *G. arboreum* (Genome A) and *G. raimondii* (Genome D).

The incidence of the breeding carried out since the XVIIIth century on the genetic diversity of the variety material could be properly estimated with enzymatic polymorphism only if that of *latifolium* landrace could be studied in its place of origin (Mexico and Guatemala). It can however be admitted that this race has the same origin as the other Mexican races and that its appearance was also accompanied with a strong genetic drift.

The low diversity of the cultivated varieties shown by the enzymatic analysis (4 polymorphic loci) therefore is consistent with one of any of the Mexican races studied (between 2 and 5 polymorphic loci). It seems thus that recent breeding did not cause any sharp reduction in total variability but that, in return, the mixings connected with all the crosses carried out have destroyed any possible initial structure of the material since no differentiation is observed between the various genetic pools (Allen, Triumph and N'Kourala). Besides, it was not possible to link the slightly higher diversity of the African cultivars and the programmes of interspecific hybridizations carried out in the Ivory Coast. No specific genetic supply of the H.A.R. material is detected, as the varieties derived from these programmes do not distinguish themselves from the varieties of purely *G. hirsutum* genetic pools. This absent characterization can in fact be explained by the very close allelic composition of *G. arboreum*, *G. raimondii* and *G. hirsutum* (BOURDON, 1984).

CONCLUSION

While confirming the already known genetic characteristics of the 2 species of cultivated cottons (allopolyploid structure, interspecific differentiation, independent domestication of each species, system of reproduction with predominant autogamy in cultivated forms, reduced genetic basis of the modern varieties), electrophoresis brings new data on their genetic organization :

— wild populations, especially those of *G. hirsutum*, are very polymorphic and seem to have a non-negligible allogamy rate. Precise conclusions on the evolutionary history of each species would require a certain number of analyses be carried out :

- first, on the Ecuadorian populations of *G. barbadense* as to check whether the wild cottons of this region can be the origin of the very differentiated populations of Peru and the Galapagos Archipelago,

- second, on all the wild populations of *G. hirsutum* in the Caribbean sea to estimate whether the hypotheses of dissemination by sea streams are valid.

— the genetic diversity of landraces shows that those of *G. barbadense* (low polymorphism rate and absent struc-

ture) are derived from a single process of domestication while in *G. hirsutum*, the Mexican and Marie-Galante races (very differentiated) have different origins :

— this genetic organization drawn from the enzymatic analysis contrasts with the morphological observations showing a large phenotypical diversification in *G. barbadense* and a sometimes difficult distinction between *G. hirsutum* geographical races :

— *Marie-Galante* has a very particular status, since its genetic diversity can partly be explained by both inter and intraspecific hybridizations. Such a situation emphasizes the advantage of this material as far as genetic resources are concerned.

The incidence of these results on the collection, conservation and utilization of genetic resources will be discussed later. As from now, the possibility of showing polymorphism in the whole material, a differently structured diversity in wild forms and landraces, and of connecting this data with processes of domestication, emphasizes the advantage of using enzymatic polymorphism to establish strategies within the framework of genetic resources.

RESUMEN

La diversidad y la estructura genética de las dos especies de algodónes cultivadas aotetraploides, *G. hirsutum* y *G. barbadense*, están estudiadas por medio del polimorfismo de 6 sistemas enzimáticos, sea 14 loci. El análisis está realizado sobre algodónes salvajes y subespontáneos de las dos especies y sobre una muestra de variedades cultivadas de *G. hirsutum*.

Los algodónes salvajes se caracterizan por un polimorfismo importante, sobre todo a nivel intrapoblación, a causa de una tasa de alogamia no despreciable. En la especie *G. barbadense*, los

algodones de Peru y de los Galápagos parecen fuertemente diferenciados.

Los algodónes subespontáneos son muy poco variables en *G. barbadense* mientras que, en *G. hirsutum*, manifiestan una diferenciación importante entre razas geográficas, a la cual se añaden en la raza *Marie-Galante*, fenómenos de introgresiones tanto inter como intra específicas.

Estos resultados, discutidos desde un punto de vista evolutivo, aportan elementos nuevos sobre los acontecimientos que ocurrieron la historia y la domesticación de estas especies.