

II - ETUDE BIBLIOGRAPHIQUE

Elaeis oleifera (H.B.K.) Cortés - CAIAUÉ

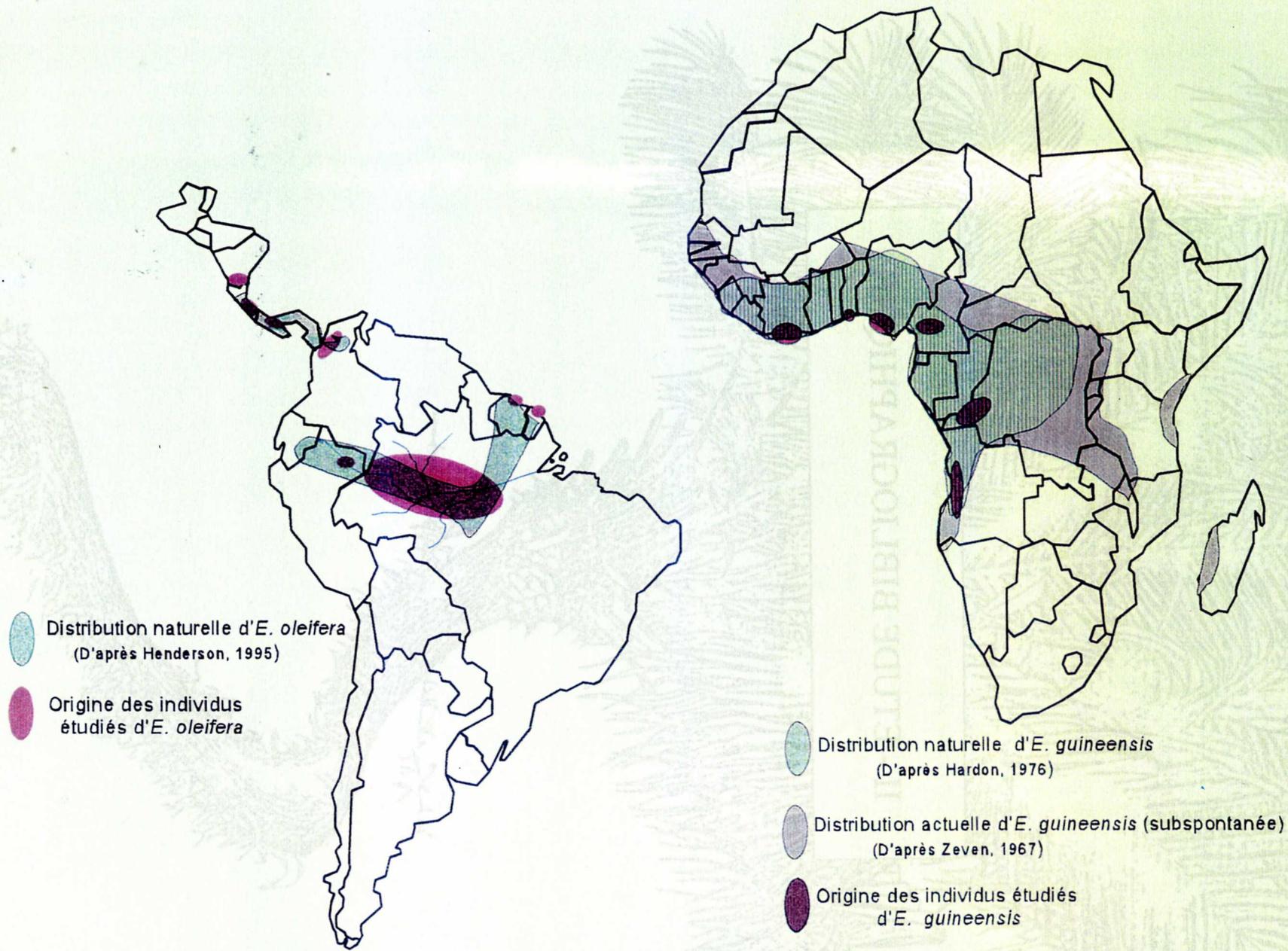


Figure II.1 - Distribution naturelle d'*E. oleifera*, d'*E. guineensis* et origine des individus étudiés de ces deux espèces.

II - ÉTUDE BIBLIOGRAPHIQUE

II.1 - TAXONOMIE ET BOTANIQUE

II.1.1 - Le genre *Elaeis* Jacq.

Le genre *Elaeis* appartient à la famille des Arecaceae (Palmae), tribu *Cocoeae* (*Cocoinae*), sous-tribu *Elaeidinae*. Le genre se rencontre dans les zones tropicales humides de l'Afrique et d'Amérique (Figure II.1). Il comprend deux espèces taxonomiquement bien définies : *E. guineensis* Jacq., le palmier à huile africain, et *E. oleifera* (Kunth) Cortés, le palmier à huile américain. Une autre espèce, *Barcella odora* Traill, qui a été parfois classée dans le genre *Elaeis*, est un palmier également américain, mais très peu connu, dépourvu d'intérêt économique (Bailey, 1940 ; Corley, 1976 ; Hartley, 1988 ; Zeven, 1972).

II.1.1.1 - L'espèce *E. guineensis*, Jacq. - Le palmier à huile africain

L'espèce *E. guineensis* Jacq. doit son nom au botaniste Jacquin (1763) qui a attribué le centre d'origine à la côte guinéenne, en Afrique de l'ouest. Cette région, s'étendant du Sénégal jusqu'à l'Angola (Figure II.1), est considérée par certains auteurs comme l'aire de distribution naturelle d'origine de l'espèce (Hardon, 1976 ; Hartley, 1988). A l'équateur, elle s'étend plus profondément à l'intérieur de la République Démocratique du Congo, ex-Zaïre. Plus récemment elle a été aussi dispersée au Kenya, en Tanzanie et à Madagascar (Figure II.1), où elle se rencontre à l'état subspontané. Le nom générique *Elaeis* dérive du grec *elaion*, qui signifie huile (Hartley, 1988).

Le palmier à huile est une monocotylédone pérenne, monoïque, produisant des inflorescences mâles et femelles, en cycles alternés chez la même plante, induisant un système de reproduction allogame. Les études cytologiques et cytogénétiques montrent que le nombre de chromosomes est de $2n = 32$ (Schwendiman *et al.*, 1982 ; Tan, 1976). Le palmier produit deux types d'huile de composition, d'usage et d'intérêt différents. L'huile de palme est le produit économiquement le plus

Photo II.1a - Fruits mûrs d'*Elaeis guineensis* variété *Tenera*.

A - Mésocarpe d'où est extraite l'huile de palme. L'huile de palme représente environ 24% du poids du régime.

B - Endosperme plus embryon d'où est extraite l'huile de palmiste. L'huile de palmiste représente 5% du poids du régime.

C - Endocarpe ou coque.

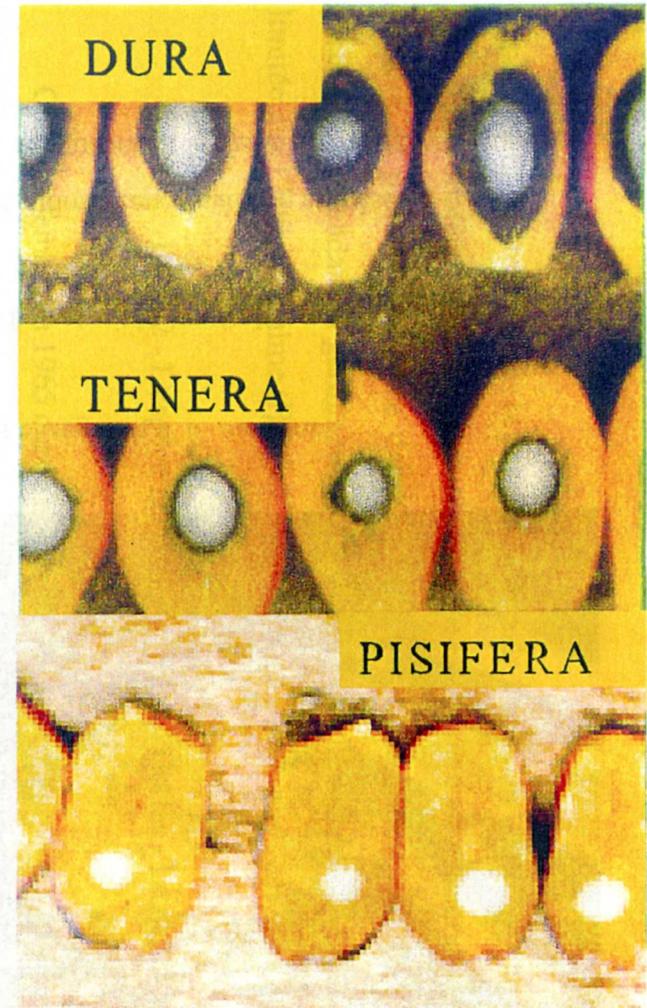


Photo II.1b - Types de fruits d'*E. guineensis* selon l'épaisseur de la coque. La présence/absence de la coque a une hérédité mendélienne simple (monogénique).

Dura - La coque est très épaisse. Etat homozygote $Sh+Sh+$

Pisifera - La coque est absente. Etat homozygote $Sh- Sh-$

Tenera - La coque est fine. Etat hétérozygote $Sh+ Sh-$

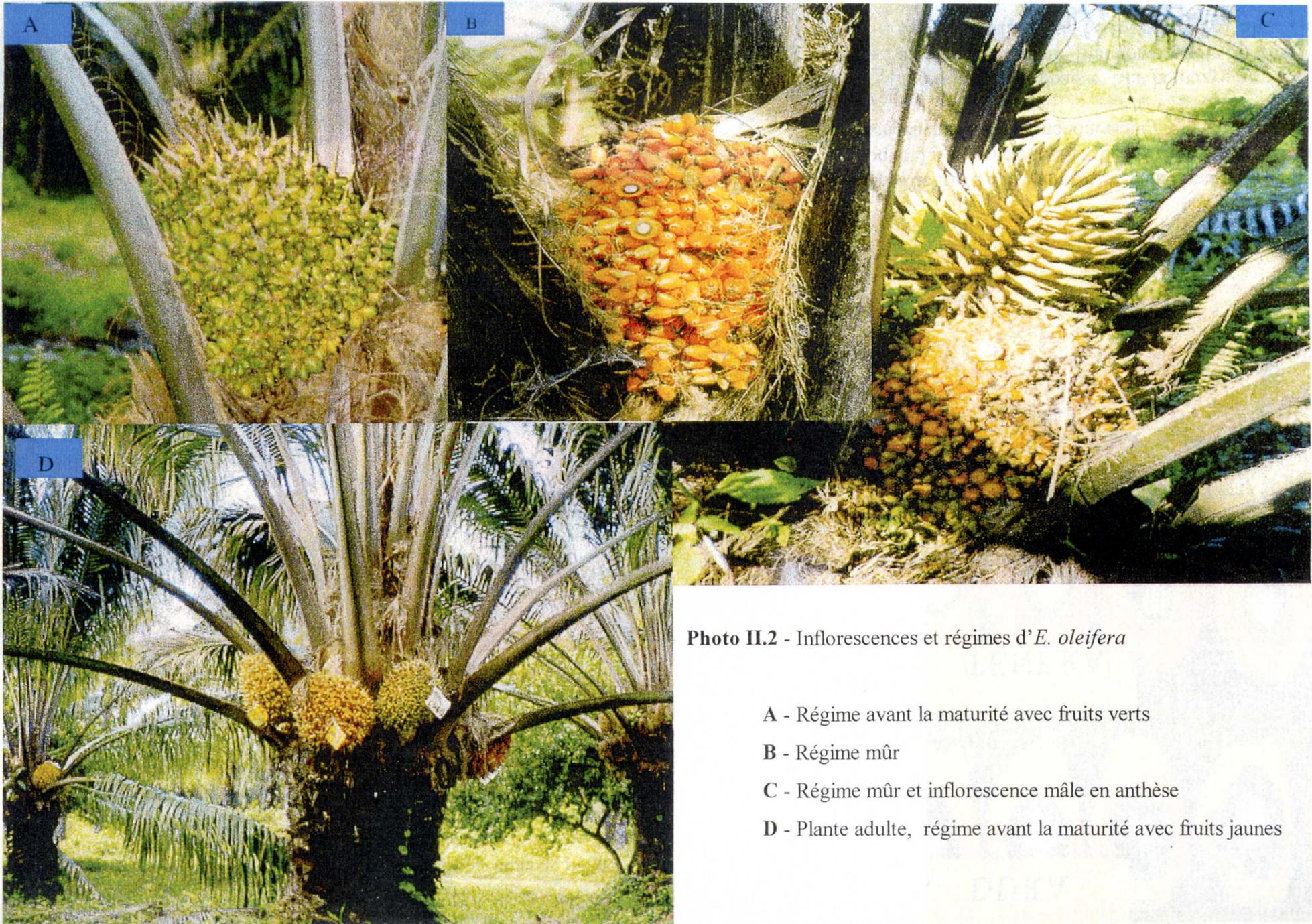


Photo II.2 - Inflorescences et régimes d'*E. oleifera*

- A - Régime avant la maturité avec fruits verts
- B - Régime mûr
- C - Régime mûr et inflorescence mâle en anthèse
- D - Plante adulte, régime avant la maturité avec fruits jaunes

important. Elle se trouve dans la pulpe du fruit mûr, entourant la graine (Photo II.1a). Dans la nature, cette pulpe confère une certaine protection aux graines et facilite leur dispersion par les animaux. L'huile de palmiste, présente dans l'amande, constitue une partie des réserves de la graine, mobilisables pendant la germination de l'embryon ; elle est produite en quantité beaucoup plus limitée que l'huile de palme (Hartley, 1988).

II.1.1.2 - L'espèce *E. oleifera* (Kunth) Cortés - Le palmier à huile américain

C'est seulement en 1965 que la nomenclature de cette espèce a été précisée par Wessels Boer, qui a conclu que le binôme proposé par Cortés en 1897 *E. oleifera* (Kunth), Cortés était taxonomiquement le plus correct. Pour cela, il a considéré que les similarités anatomiques entre les deux espèces, *E. oleifera* (Photo II.2) et *E. guineensis*, et la facilité de croisement entre elles, offraient la possibilité d'une classification congénérique (Wessels Boer, 1965). Les études de diversité isoenzymatique et cytogénétique qui ont été réalisées sur les deux espèces et leurs hybrides F1, confortent la classification congénérique actuelle (Rajanaidu et Williams, 1977 ; Schwendiman, Pallares et Amblard, 1982).

L'E. oleifera est présent exclusivement en Amérique tropicale et son aire de distribution (Figure II.1, p.5) couvre une zone qui va du Mexique jusqu'à l'Est du Bassin de l'Amazone au Brésil (Meunier, 1975). Contrairement à son homologue africain, il n'a pas été étudié de façon systématique jusqu'à une époque récente. En l'absence de toute sélection et malgré des productions élevées en régimes, son faible taux d'extraction d'huile limite le rendement en huile par hectare.

L'E. oleifera se rencontre généralement en petits groupes de palmiers constitués de quelques dizaines d'individus à plus de 1.000 dans quelques populations. Il n'a pas été trouvé de plantes isolées dans la forêt ou dans des chablis. Les populations spontanées de cette espèce se situent dans les bas-fonds, dans des zones marécageuses, à proximité de petites rivières, et secondairement dans des zones déforestées par l'homme. Une particularité d'*E. oleifera* est son adaptation aux zones inondables, à tel point que les populations typiquement sauvages sont toujours trouvées dans ces conditions (Barcelos, 1986 ; Blank, 1952 ; Meunier, 1975). Très peu d'informations sont disponibles sur les mécanismes de dispersion et sur l'organisation de la diversité génétique au sein de cette espèce.

Dans le bassin amazonien, au Brésil, les populations d'*E. oleifera* sont fréquemment rencontrées sur des sols d'origine anthropogénique "terra prêta do indio". Ces taches de sols présentent des caractéristiques de bonne fertilité telles qu'une teneur élevée en phosphore, en calcium et en matière organique. Sur ces sites, on trouve une abondance de fragments de poterie, ce qui permet de les associer à une occupation indigène ancienne du lieu. Au cours d'une prospection réalisée sur l'Amazonie brésilienne, 41% (13/32) des populations ont été trouvées sur des sols d'origine anthropogénique (Barcelos, 1986 ; Ooi *et al.*, 1981).

II.1.2 - L'espèce *Barcella odora* Traill

Le genre *Barcella*, représenté par une seule espèce, forme avec le genre *Elaeis*, la sous-tribu *Elaeidinae*. Cette espèce, peu abondante, a une distribution limitée aux aires de savane sableuse au Nord du Rio Negro, en Amazonie brésilienne (Henderson *et al.*, 1995).

II.2 - LE PALMIER À HUILE, SA MISE EN CULTURE ET SON AMÉLIORATION GÉNÉTIQUE

II.2.1 - La domestication du palmier à huile

L'utilisation de l'huile de palme remonte aux Egyptiens (Fridel, 1897 cité par Zeven, 1967). Dans son aire d'origine, l'Afrique intertropicale humide, le palmier à huile fait partie intégrante du paysage et de la culture populaire. Au 15^{ème} siècle, les premiers voyageurs portugais Ca' da Mosto (1435-1460) et Diogo Gomes (1456-1457) ont rapporté la présence du palmier à huile sur la côte guinéenne et l'utilisation de l'huile de palme par ses habitants (Zeven, 1964). Vers 1506-1508, Duarte Pacheco Pereira relate l'existence de palmeraies sur la côte du Liberia et aussi le commerce de l'huile de palme au Nigéria (Crone, 1937 cité par Hartley, 1988).

En Afrique, les populations ont su utiliser les différentes parties de la plante et sa domestication s'est faite en plusieurs étapes (Zeven, 1967). La domestication complète du palmier à huile aurait commencé à l'Est du Nigéria, puis aurait progressé sur le continent africain d'une tribu à l'autre et d'une région à l'autre, par les mouvements migratoires de ces tribus (Zeven, 1967). L'huile de palme est restée d'usage domestique pour les populations locales, principalement en l'Afrique de l'Ouest, jusqu'au début du siècle dernier, quand son commerce a été encouragé par les anglais comme une alternative au commerce des esclaves. C'est au cours du XX^{ème} siècle que la culture industrielle de cette plante a connu un développement intensif (Hartley, 1988).

II.2.2 - L'expansion de la culture du palmier à huile

En Afrique, l'exploitation des palmeraies naturelles a été intensifiée dès le début de ce siècle et vers 1912, quelques petites parcelles ont été établies au Cameroun puis, ensuite au Nigéria et au Congo belge, actuelle République Démocratique du Congo. Vers 1939, 14 000 hectares de plantations de palmiers à huile existaient déjà en Afrique.

Malgré ces efforts, l'expansion en Afrique a été lente et a permis aux pays du sud-est asiatique d'occuper depuis 1935, la position de premiers exportateurs mondiaux d'huile de palme (Hartley, 1988).

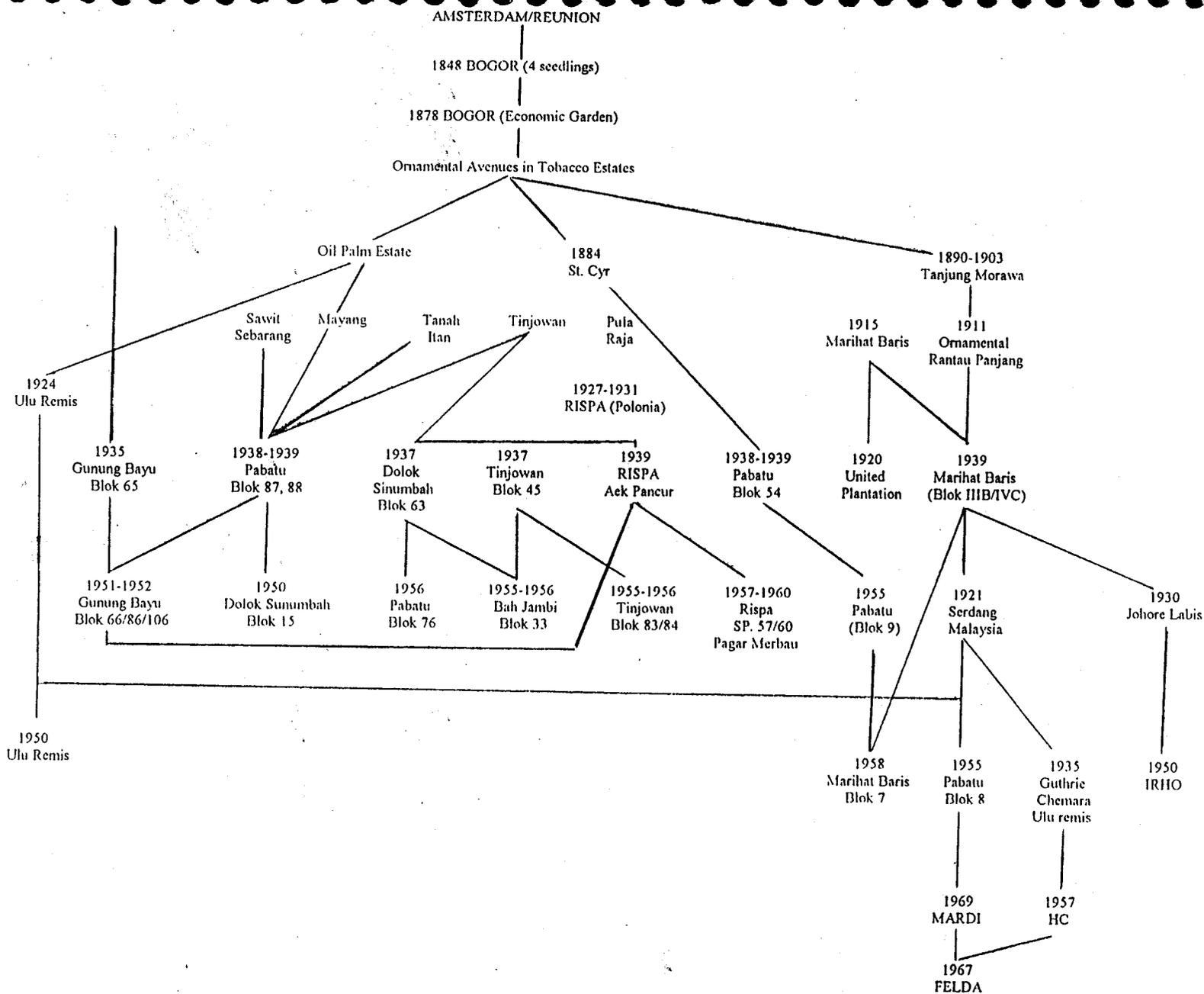
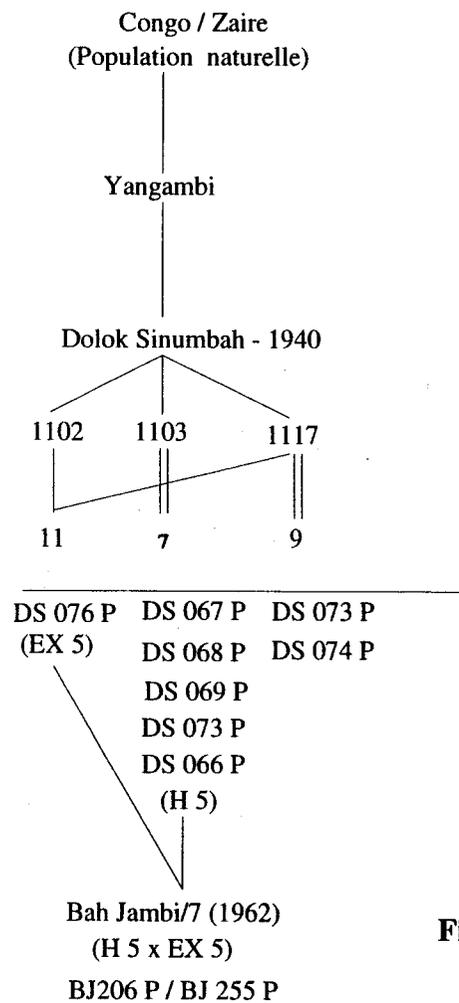
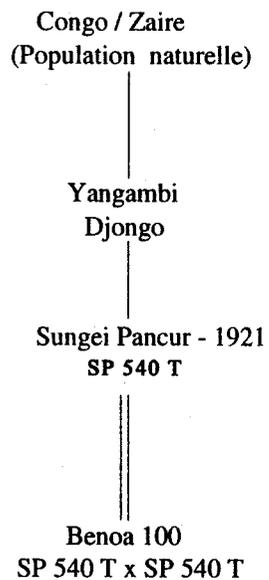


Figure II.2 - Plantations et stations de recherche où les lignées femelles Dura descendant de 4 arbres Deli ont été sélectionnées. Dura Deli est la seule origine de géniteurs femelles d'*E. guineensis* actuellement utilisée pour la sortie variétale du palmier à huile (d'après Rosenquist, 1985).

Pisifera Dolok Sinumbah



Pisifera AVROS



Pisifera La Mé

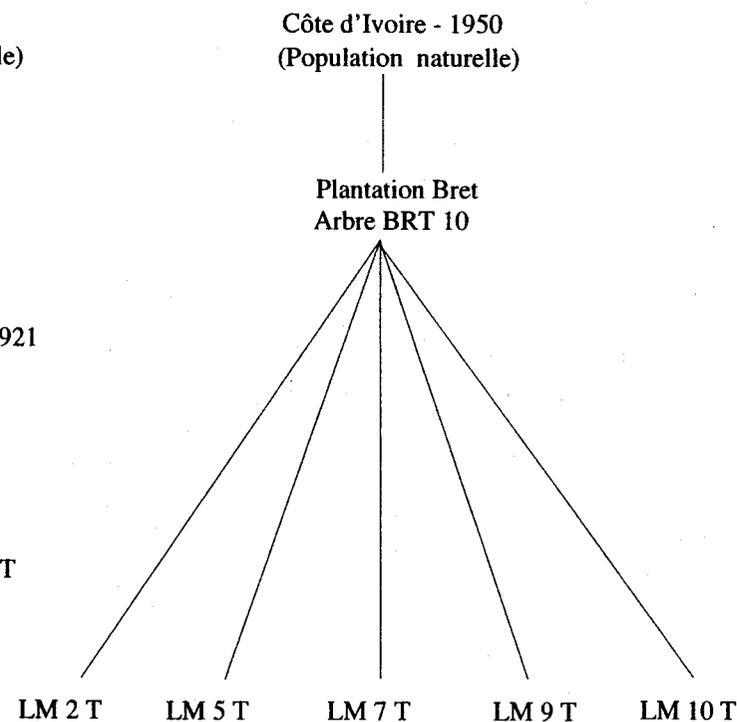


Figure II.3 - Origine et base génétique de géniteurs mâles Pisifera d'*Elaeis guineensis* actuellement utilisés pour la sortie variétale du palmier à huile. (D'après Rosenquist, 1985)

Le palmier à huile est arrivé en Amérique au 17^{ème} siècle, avec les esclaves. Au Brésil, une palmeraie sub-spontanée de près de 20 000 hectares a été établie, dans la région de Bahia. La première plantation industrielle sur le continent américain a été réalisée en 1943 au Honduras. Vers la fin des années cinquante, la culture était déjà installée au Venezuela, en Colombie, en Equateur, au Costa Rica et au Nicaragua. Au Brésil, les premières plantations industrielles ont été faites dans les années soixante à Bahia, et au Pará en Amazonie.

L'introduction du palmier à huile en Asie du sud-est a eu lieu en 1848 sous forme de 4 plantes, reçues de l'Hortus Botanicus d'Amsterdam et plantées au jardin botanique de Buitenjorg (=Bogor) à Java (Hartley, 1988). Les descendances produites par ces 4 plantes sont à l'origine de l'exploitation économique du palmier à huile à grande échelle dans le sud-est asiatique. Le palmier a été utilisé d'abord à des fins ornementales dans les plantations de tabac à partir de 1884. C'est seulement vers 1911 en Indonésie et 1912 en Malaisie, que l'on a vu apparaître les premières plantations importantes de palmier à huile (Hartley, 1988). Un quart de siècle plus tard, environ 100.000 ha étaient en exploitation en Indonésie et 30.000 ha en Malaisie, ces deux pays représentaient alors 40% des ventes du marché mondial. Aujourd'hui, ces deux pays comptent, ensemble, plus de 4.000.000 hectares en production (Hartley, 1988 ; Oil World, 1997).

Au cours de cette étape d'expansion du palmier à huile dans le Sud-est Asiatique, les plantations ont été réalisées avec du matériel issu d'arbres-mères sélectionnés dans les premières plantations faites à partir des graines produites à partir de quatre plantes d'origine imprécise, baptisé "Deli" (Figure II.2), du nom de la province de Sumatra où avait été faites les premières plantations.. Ces quatre arbres Deli sont, aujourd'hui, à l'origine de la plupart des lignées "Dura" femelles, utilisées pour la création variétale dans tous les centres de production de semences de palmier à huile du monde. L'origine des lignées mâles utilisées actuellement pour produire la grande majorité des semences est aussi peu diversifiée (Figure II.3). On peut donc constater que la culture du palmier à huile est basée actuellement sur des semences (Tableau II.1) provenant d'une base génétique relativement étroite (Hartley, 1988 ; Rajanaidu, 1985b ; Simmonds, 1993 ; Yuan *et al.*, 1996).

Tableau II-1 - Production mondiale de semences de palmier à huile en 1997 : quantité par pays producteur et base génétique

Pays	Quantité de semences (millieus)	Géniteur femelle Dura Deli %	Géniteur mâle - %	
			SP/BG/Ex5	BRT 10
Indonésie	61	100	70	10
Malaisie	50	100	60	
Papouasie Nouvelle Guinée	10	100	100	
Costa Rica	12	100	75	
Côte d'Ivoire	6	100		100
Bénin	2	100		100
Cameroun	1	100		100
Autres	8	100		12
Totaux	150	100	59	11

Adapté de Rajanaidu et al., 1995 ; Yuan et al., 1996 ; Noiret, communication personnelle

Dura Deli = Descendants de seulement 4 plantes introduites en Indonésie en 1873

SP/BG/Ex5 = SP540/Bangun/Ex5, Descendants de 3 arbres sélectionnés au Congo/Zaire

Brt 10 = Les géniteurs mâles *pisifera* sont obtenus de descendance par autofécondation ou croisements entre plantes *tenera* LM 2 T, LM 5 T, LM 7 T et LM 10 T sélectionnées à La Mé / Côte d'Ivoire. Tous ces géniteurs ont pour ancêtre commun l'arbre BRT 10 (Figure II.3).

II.2.3 - L'amélioration génétique du palmier à huile

L'expansion de la culture du palmier a été soutenue par un important effort de recherche agronomique avec de gros progrès sur l'augmentation des rendements. L'augmentation de la productivité du palmier à huile a été de 315% entre 1951 et 1991, en prenant pour base les productions obtenues sur les premières plantations faites avec les variétés de type "Dura". Cette augmentation vient des améliorations sur la gestion de l'utilisation des engrais et des progrès du potentiel génétique des semences (Rajanaidu and Jalani, 1996). On est parvenu à des rendements remarquables en comparaison avec d'autres oléagineux, en produisant dans les conditions écologiques les plus favorables, en moyenne dix fois plus d'huile par hectare et par an que le soja (Escobar, Sterling and Peralta, 1996 ; Hee, 1996 ; Pamin *et al.*, 1996 ; Rajanaidu and Jalani, 1996 ; Sharma and Pau, 1996 ; Weng and Shamsuddin, 1996 ; Yuan, Peng and Weng, 1996 ; Yuan and Weng, 1996).

Augmenter le rendement d'huile totale (huile de palme plus huile de palmiste) à l'hectare, est le but principal de tous les programmes d'amélioration génétique du palmier à huile. Cela a consisté en l'obtention d'un matériel présentant une productivité élevée, d'un produit de bonne qualité, assurant une durée raisonnable d'exploitation des palmiers (25 ans), avec une production régulière, et une réduction des coûts de production.

Chez le palmier à huile, il existe certains caractères à hérédité mendélienne simple. Le caractère de ce type le plus important économiquement est "l'épaisseur de la coque" (Photo II.1b, p. 3). Il est contrôlé par un couple d'allèles co-dominants Sh+ et Sh-, les arbres produisant des fruits de type *dura* (à coque épaisse) sont de génotype Sh+Sh+ alors que le génotype Sh-Sh- correspond aux arbres produisant des fruits de type *pisifera* (sans coque) et présentant une stérilité femelle marquée. Le génotype hétérozygote Sh+Sh- correspond aux arbres produisant des fruits de type *tenera* à coque mince. Ce dernier présente une fertilité normale et une teneur en huile plus élevée, donc plus de valeur économique. Les individus de type *Dura*, utilisés comme géniteurs femelles en croisement avec les individus type *pisifera*, produisent des descendants 100% *tenera*. L'origine Deli ne comporte que des individus type *dura* (Gascon, 1989).

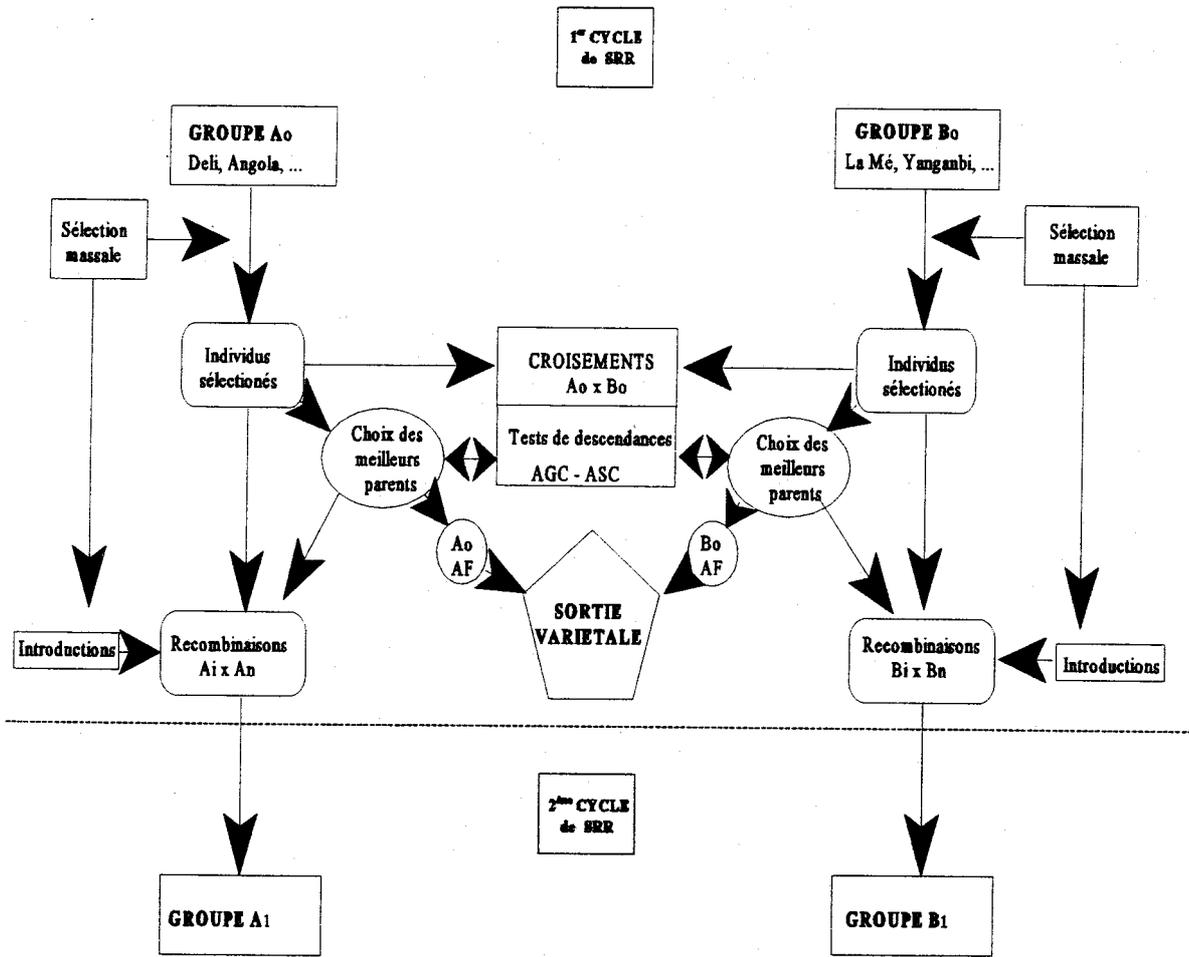


FIGURE II.4 - Schéma de sélection récurrente réciproque adaptée au palmier à huile (D'après Gascon, 1989).

Elaeis guineensis



Elaeis oleifera



Hybride F1



Rétrocroisement
Hybride F1 x *E. guineensis*



Photo II.3 - L'hybridation interspécifiques *E. oleifera* x *E. guineensis*. Les hybrides F1 et les rétrocroisements fournissent des régimes avec de nombreux fruits fertiles

Pour le palmier à huile, comme pour plusieurs autres espèces allogames, un effet d'hétérosis sur la productivité a été détecté très tôt (Gascon *et al.*, 1969 ; Hardon, 1970). Les premiers résultats ont montré que les hybrides entre origines les plus différenciées en terme de nombre et de poids moyen de régimes produisent beaucoup plus. Ceci a servi de base pour la définition du schéma d'amélioration par sélection récurrente réciproque (Figure II.4) adoptée par l'IRHO en 1957 (Gascon et de Berchoux, 1964). Les origines de palmier à huile ont été réparties en deux groupes portant des caractères complémentaires pour les composantes de rendement, avec un groupe constitué par l'origine Deli comme géniteur femelle *Dura* et l'autre groupe constitué par d'autres origines africaines notamment La Mé/Côte d'Ivoire et Yangambi/Gongo comme géniteur mâle *pisifera* (Meunier et Gascon, 1972). Cette stratégie, utilisée avec de légères modifications par les programmes d'amélioration génétique de cette espèce, a permis l'obtention de matériel hautement productif (Escobar, Sterling and Peralta, 1996 ; Hee, 1996 ; Pamin *et al.*, 1996 ; Rajanaidu and Jalani, 1996 ; Sharma and Pau, 1996 ; Weng and Shamsuddin, 1996 ; Yuan, Peng and Weng, 1996 ; Yuan and Weng, 1996).

L'amélioration d'autres caractères tels que la résistance aux maladies, la vitesse de croissance et la qualité de l'huile (Tableau II.2) , n'a pas connu d'avancées aussi marquées, surtout à cause de la faible variabilité de ces caractères au sein des origines de l'espèce *E. guineensis*. Les sélectionneurs s'intéressent donc à l'*E. oleifera*, qui présente de telles caractéristiques. Il a été montré la possibilité de l'exploiter par la voie d'hybridation inter spécifique et de rétrocroisements pour l'amélioration du palmier à huile (Meunier and Hardon, 1976).

II.2.4 - L'hybride *E. oleifera* x *E. guineensis*

Par hybridation entre les deux espèces (Photo II.3) on obtient des descendance nombreuses et viables (Meunier and Hardon, 1976 ; Hardon, 1969). L'hybride F1 présente une valeur intermédiaire entre celle de ses deux parents pour les caractères agromorphologiques (Tableau II.3). Cet ensemble de caractéristiques fait de l'hybride un matériel tout à fait prometteur, mais jusqu'à maintenant, les hybrides connus ne sont pas directement exploitables commercialement en F1, sauf à accepter des rendements au minimum 30% inférieurs à ceux des variétés commerciales d'*E. guineensis* (Amblard *et al.*, 1995).

Tableau II.2 - Composition en acide gras des principales huiles végétales

Huile	Indice d'iode	Indice de Saponification	Acides gras saturé	Acide gras insaturé		
				C18:1 oléique	C18:2 linoléique	C18:3 linoléique
Arachide	89-98	185-197	18	54	24	traces
Colza	94-106	170-186	3	12	13	9
Coton	103-111	197-197	23	17	54	traces
Lin	170-185	180-195	9	23	20	48
Olive	80-85	185-203	11	76	7	traces
Tournesol	125-136	188-194	11	16	72	traces
Maïs	117-123		13	26	55	traces
Soja	130-142	185-196	15	25	52	8
Coprah (<i>C. nucifera</i>)	8-10	246-260	80	5	1	traces
Palmiste (<i>E. guineensis</i>)	16-23	242-260	78	15	tr	traces
Palme (<i>E. guineensis</i>)	44-58	192-210	50	32-36	8-11	0,4
Palme (<i>E. oleifera</i>)	81-84		24	64-68	2-14	0,6
Palme (Hybride <i>E.o. x E.g.</i>)	60-72		36	48-51	12-14	0,4

D'après: Hardon, 1969 ; Meunier & Hardon, 1976 ; Péhault, 1986 ; Hartley, 1988.

Tableau II.3- Valeurs de différents caractères agronomiques chez les espèces *E. guineensis*, *E. oleifera* et leurs hybrides

Caractères	<i>E. guineensis</i>	<i>E. oleifera</i>	Hybrides
Croissance du stipe en hauteur (cm/an)	30 - 75	5	15 - 25
Résistance à la pourriture du coeur/mortalité (%)	75	0	≤ 1
Résistance à la Fusariose (indice)	58 - 141	0 - 400	0 - 150
Résistance au <i>Ganoderma</i> / mortalité (%)	10 - 70	ND	≤ 3
Résistance au <i>Coelaenomenodera elaeidis</i> -mortalité de larves (%)	26 - 46	ND	39 - 89
Résistance au <i>Leptopharsa gibbicularina</i> - mortalité de larves (%)	nulle	60%	60%
Degré d'insaturation de l'huile (%)	40 - 60	60 - 83	62 - 69
Taux d'huile sur pulpe sèche (%)	67 - 76	35 - 49	59 - 68
Taux d'huile sur régime (%)	18.3 - 25.5	1.7 - 4.4	3.8 - 17
Taux d'extraction d'huile industriel (%)	20 - 24	≤ 9	8.9 - 18.8

Source: Hardon 1969 ; Hardon & Tan 1969 ; Macfarlane et al. 1975 ; Meunier 1975 ; Vallejo et Cassalet 1975 ; Meunier et al. 1976, 1979 ; Meunier et Hardon 1976 ; Rajanaidu et al. 1979 ; Renard et al. 1980 ; Ooi e al. 1981; Hartley 1988 ; Rajanaidu et al. 1983 ; Barcelos et al. 1985 ; Le Guen et al. 1991 ; Amblard et al. 1995.

Degré d'insaturation de l'huile - Proportion des acides gras présentant des doubles et triples liaisons entre les atomes de carbone en substitution à des atomes d'hydrogène. Sur les acides gras totaux présents dans l'huile de palme les acides insaturés représentent environ la moitié : l'acide oléique C18:1 avec environ 40%, l'acide linoléique C18:2 avec environ 10% et des traces d'acide linoléique C18:3 et d'acide palmitoléique C16:1.

Fusariose - Maladie fongique létale (*Fusarium oxysporum*, f. *elaeidis*) qui cause d'importants dégâts sur les plantations du palmier à huile sur le continent africain. Dernièrement cette maladie a été aussi identifiée sur une plantation au Brésil.

Ganoderma - Maladie fongique létale (*Ganoderma sp*), présente sur les plantations africaines et surtout asiatiques.

Pourriture du coeur - Maladie létale, de cause inconnue, qui produit de très importants dégâts sur les plantations de palmier à huile du continent américain.

Coelaenomenodera sp - Insecte mineur de folioles du palmier à huile, provoque des dégâts par la réduction de l'aire foliaire dans les plantations d'Afrique et d'Amérique.

Leptofarsa gibbicularina - Insecte piqueur étroitement lié à des dégâts foliaires provoqués par des champignons sur des plantations de palmier à huile en Amérique.

Tableau II.4- Caractères biologiques et cytologiques chez les espèces *E. guineensis*, *E. oleifera* et leurs hybrides

Caractères	<i>E. guineensis</i>	<i>E. oleifera</i>	Hybrides
Fertilité pollinique <i>in vitro</i> (%)	>70	>70	20 ¹ - 44 ²
Sacs embryonnaires normaux (%)	93,6	70,4	21,4 ¹ - 43,2 ²
Fleurs fécondables (%)	100	100	49,7 - 81,5
Fécondation des sac embryonnaires (%)	25 - 60	25 - 60	0 - 40
Nombre de chromosomes 2N	32	32	32
Nombre de monovalents I	0,310	0,146	-
Nombre de bivalents II	15,516	15,683	15,714
Nombre de tétravalents IV	0,234	0,122	0,143
Chiasma par cellule	29,860	30,440	29,095
Chiasma par bivalent	1,867	1,910	1,815

¹Hybride avec l'*E. oleifera* d'origine Colombie

²Hybride avec l'*E. oleifera* d'origine Brésil

Source: Hardon et Tan, 1969 ; Arnaud, 1973, 1980 ; Tan, 1976 ; Schwendiman, 1982 ; Baudouin, 1983.

Cette perte de rendement en huile vient de quelques caractères défavorables de l'hybride au niveau de la F1. Elle résulte à la fois d'une mauvaise formation des régimes liée à un problème de stérilité partielle et aussi d'une faible teneur en huile des fruits. Plusieurs études cytologiques et cytogénétiques, notamment celles de Tan (1976) et de Schwendiman *et al.* (1982), montrent une grande homologie entre les génomes d'*E. guineensis* et d'*E. oleifera*, avec le même nombre de chromosomes ($2n=32$). La méiose chez l'hybride présente en règle générale le même comportement que chez les deux parents (Tableau II.4) avec un faible nombre d'univalents et de très rares multivalents, ce qui peut expliquer l'obtention facile d'hybrides entre ces deux espèces..

Jusqu'à présent, plusieurs populations d'*E. oleifera* originaires du Nord de la Colombie et de l'Amérique Centrale, deux populations du Brésil et une population du Surinam ont été évaluées en croisement avec l'espèce africaine. Il a été constaté des descendance plus ou moins fertiles entre les hybrides F1, avec une très importante variabilité génétique pour ces caractères (Tableau II.4) selon les origines d'*E. oleifera* utilisées (Baudouin, 1983 ; Schwendiman *et al.*, 1983). Les hybrides F1 *E. oleifera* x *E. guineensis* obtenus à partir des populations d'*E. oleifera* originaires du Surinam et du Brésil sont en moyenne plus fertiles et plus productifs que ceux obtenus avec les origines de Colombie et d'Amérique Centrale (Amblard *et al.*, 1995 ; Arnaud, 1980 ; Baudouin, 1983 ; Lubis *et al.*, 1987 ; Schwendiman, Pallares and Amblard, 1982 ; Schwendiman *et al.*, 1983).

En considérant l'existence d'un grand nombre de populations d'*E. oleifera* déjà en collection, provenant de l'Amazonie brésilienne, on peut supposer l'existence de variabilité génétique tant entre ces populations qu'au sein des populations. Une telle diversité génétique peut permettre l'obtention d'hybrides présentant un comportement encore plus intéressant que ceux étudiés, ce qui renforce l'intérêt de mieux connaître la diversité génétique des populations brésiliennes d'*E. oleifera*.

II.3 - RESSOURCES GÉNÉTIQUES

II.3.1 - Origine et dispersion des espèces d'*Elaeis*

L'origine géographique de *E. guineensis* reste encore imprécise. L'Afrique de l'Ouest, plus particulièrement la Guinée et l'est du Nigéria, a été proposée par Zeven (1972) comme centre d'origine et de diversification de l'espèce. Cette origine est fondée sur la découverte de pollen fossile semblable à celui du palmier à huile actuel, dans le delta du Niger, dans des sédiments datant du Miocène et renforcée par plusieurs évidences historiques et culturelles (Zeven, 1964). Les découvertes plus récentes de pollen dans des sédiments datant de l'Eocène en Guinée (Zaklinskaya et Prokof'yev, 1971), et d'autres du Quaternaire, de l'Holocène au Cameroun (Maley et Brenac, 1998), confirment la présence continue de cette espèce en Afrique, depuis environ 50 millions d'années.

Une origine américaine pour le palmier à huile a été proposée par Cook (1942), à partir des arguments suivants :

a) Les études paléontologiques montrent que les palmacées ont une longue histoire géologique, qui peut être suivie depuis le début du Crétacé, il y a 100 millions d'années. Elles ont leur origine à l'Ouest du Gondwana, correspondant actuellement à l'Amérique du sud. Avec d'autres monocotylédones, elles se sont déplacées vers le nord et ensuite de l'est vers l'ouest. Des troncs pétrifiés ont été retrouvés datant de tout le Tertiaire. Dans les études et classifications des palmacées, on considère que les espèces les plus anciennes sont originaires de l'Amérique du sud (Moore et Uhl, 1982 ; Tomlinson, 1990).

b) Au Brésil, le genre *Elaeis* cohabite avec la grande majorité des espèces de la tribu *Cocoeae* (*Cocoinae*), exclusivement américaine, sauf pour deux espèces *E. guineensis* et *Jubaeopsis caffra*, qui présentent une distribution extra-Américaine.

La répartition actuelle de ces espèces sur les deux continents peut résulter de plusieurs événements paléogéographiques et paléoclimatiques à différentes échelles temporelles. Ces

Tableau II.5 - Paramètres de diversité génétique estimés sur des espèces pérennes

Espèce	Marqueur	Effectif étudié	Nombre de locus	Allèles par locus	Diversité génétique de Nei He
<i>Elaeis guineensis</i> ¹	Isoenzymes	51	15	2,22	0,262
<i>E. oleifera</i> ²	Isoenzymes	124	14	2,04	0,271
<i>Theobroma cacao</i>	Isoenzymes ³	327	12	1,62	0,170
	RFLP ^{3a}	155	43	1,78	0,280
<i>Hevea brasiliensis</i> ⁴	Isoenzymes	92	11	4,45	0,516
<i>Coffea canephora</i> ⁵	Isoenzymes	471	9	3,50	0,150
<i>Coffea congensis</i> ⁵	Isoenzymes	133	12	1,91	0,180
<i>Coffea stenophylla</i> ⁵	Isoenzymes	23	7	2,00	0,065
<i>Coffea humilis</i> ⁵	Isoenzymes	59	7	3,00	0,380
<i>Coffea liberica</i> ⁵	Isoenzymes	120	7	3,10	0,210
<i>Cordia alliodora</i> ⁶	Isoenzymes	550	11	1,72	0,126
<i>Carapa procera</i> ⁷	Isoenzymes	225	15	1,70	0,200
<i>Argania spinosa</i> ⁸	Isoenzymes		9	3,60	0,238
Arbres forestiers tropicaux ⁹	Isoenzymes	38 ^a	21	1,87	0,191
Arbres forestiers pérennes ⁹	Isoenzymes	191 ^a	18	2,22	0,177

nombre d'espèces étudiées de différents genres

Source: ¹Ghesquière, 1983 ; ²Ghesquière et al., 1987 ; ³Lanaud, 1987 ; ^{3a}Lerceteau et al., 1997 ; ⁴Besse et al., 1994 ; ⁴Chevallier, 1988 ; ⁵Berthaud, 1986 ; ⁶Chase et al., 1995 ; ⁷Doligez et Joly, 1997 ; ⁸ElMousadik et Petit, 1996 ; ⁹Hamrick, et al., 1992, ⁹Loveless, 1992.

événements peuvent aller de la dérive des continents à l'échelle géologique la plus éloignée pour expliquer la spéciation et l'existence de deux espèces, aux dernières glaciations et à la théorie des refuges forestiers pour expliquer la forte diversité au sein de l'espèce américaine. Plus récemment après les dernières glaciations, des phénomènes climatiques tels que de fortes sécheresses, ont limité les re-colonisations avec le maintien de groupes isolés et ont empêché le rétablissement du flux génique entre ces groupes, ce qui expliquerait le maintien d'une telle diversité. Des événements encore plus récents, comme l'activité humaine, ont affecté la dispersion de l'espèce africaine et la faible structuration de sa diversité sur le continent africain. Ces mêmes facteurs doivent aussi s'appliquer à l'espèce américaine (Blondel, 1995 ; Lavallée, 1996 ; Maley, 1996 ; Servant *et al.*, 1997 ; Zeven, 1967).

II.3.2 - Organisation de la diversité génétique des espèces de plantes pérennes

Les résultats d'analyse par isoenzymes de la diversité des espèces pérennes en général montrent que les espèces forestières sont extrêmement polymorphes, avec plus de 50% des locus polymorphes, et un nombre moyen d'allèles par locus variant de 1,68 pour les Angiospermes, à 1,83 pour les Gymnospermes (Tableau II.5). Les arbres forestiers présentent une diversité génétique moyenne estimée par l'indice de Nei de l'ordre de 0,143 (Angiospermes) à 0,151 (Gymnospermes), et le niveau de diversité intra-population est compris entre 30% à 40%. Ces résultats ont été obtenus sur 213 espèces d'arbres forestiers appartenant à 54 genres, avec une taille de population qui varie de quelques individus à plusieurs milliers et un régime de reproduction proche de l'allogamie stricte (Hamrick *et al.*, 1992 ; Kremer *et al.*, 1994).

Chez les espèces pérennes tropicales, les données bibliographiques disponibles indiquent que la diversité génétique de ces espèces analysée par isoenzymes est similaire à celle des espèces pérennes de zone tempérée, avec une diversité génétique élevée, dont la majeure partie est due à une variation intra-population (Hamrick, Godt and Sherman-Broyles, 1992 ; Loveless, 1992). Les espèces pérennes tropicales, très allogames, présentent une structuration géographique de la diversité, probablement due à des facteurs spatiaux, écologiques et phénologiques. En moyenne, seulement 11% de la variabilité dans les espèces tropicales est dû à la différence entre populations, alors que ces populations présentaient en moyenne 40% de locus polymorphes, et une diversité génétique

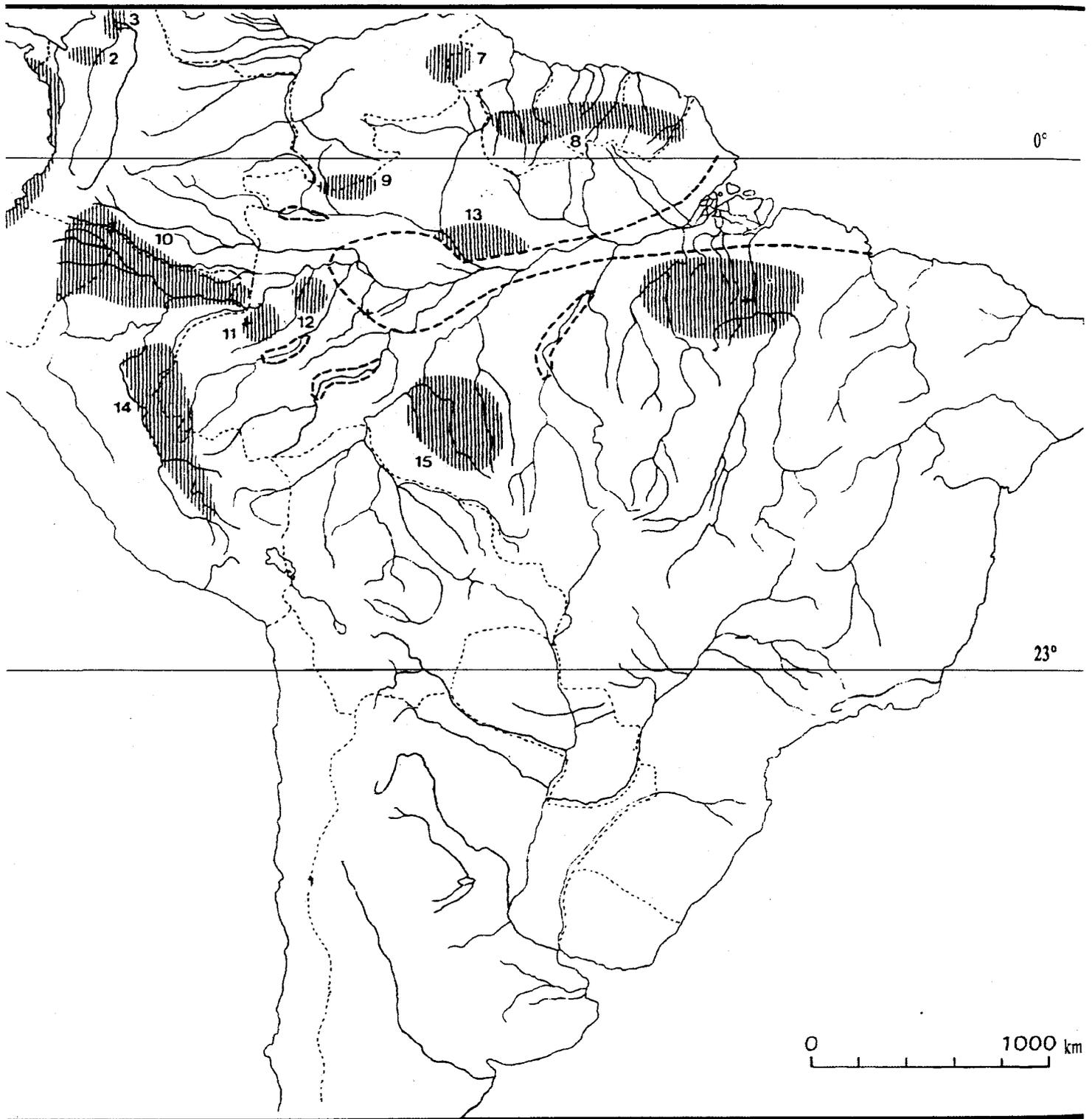


Figure II.5 - Localisation de refuges forestiers dans le bassin amazonien après les dernières glaciations du quaternaire (Whitmore & Prance, 1987), établie à partir de données sur la répartition de quatre familles d'arbres à feuilles caduques : 1 - Chocó ; 2 - Nechi ; 7 - Imataca ; 8 - Guyana ; 9 - Imerí ; 10 - Napo ; 11 - Olivença ; 12 - Tefé ; 13 - Manaus ; 14 - Estero ; 15 - Rondonia ; 16 - Belém/Xingu.

Wycherley, 1977). Ces deux centres de diversité ont été associés à deux refuges forestiers, correspondant à la région du haut Rio Negro et de Manaus respectivement (Besse, 1993 ; Whitmore and Prance, 1987). D'autres refuges comme ceux de l'est du Pérou/Acre et de Rondonia permettent d'expliquer les divergences intra spécifiques et même la spéciation chez l'hévéa (Besse, 1993).

L'évaluation de la diversité génétique de l'*Hevea brasiliensis* a été faite à l'aide des marqueurs enzymatiques (Chevallier, 1988) et des marqueurs RFLP (Besse *et al.*, 1994 ; Luo *et al.*, 1995 ; Seguin *et al.*, 1997). Ces deux types de marqueurs ont révélé une même structuration génétique, avec une organisation de la variabilité génétique suivant les différents bassins versants de la région. Une telle structuration peut s'expliquer par un mode de reproduction préférentiellement allogame et suggère un flux migratoire par voie de graines, fortement dépendant du réseau fluvial (Besse, 1993 ; Besse *et al.*, 1994 ; Seguin, Rodier-Goud and Lespinasse, 1997).

Les estimations de la diversité génétique sur l'*Hevea brasiliensis*, $H_e = 0,383$ à $0,545$ pour les données RFLP (Besse *et al.*, 1994) et $H_e = 0,458$ à $0,604$ pour les données isoenzymes (Chevallier, 1988), avec 2,90 à 4,45 allèles par locus révélés par isoenzyme, indiquent des valeurs beaucoup plus élevées que celles observées pour les arbres forestiers en général (Hamrick, Godt and Sherman-Broyles, 1992) ou plus spécifiquement pour les espèces pérennes tropicales (Loveless, 1992).

Une autre espèce amazonienne, le cacaoyer (*Theobroma cacao* L.) a aussi été étudiée. Les résultats des analyses de la diversité génétique à l'aide de marqueurs morphologiques (Engels, 1983), enzymatiques (Lanaud, 1987) et moléculaires (Figueira *et al.*, 1992 ; Laurent *et al.*, 1994 ; Lerceteau *et al.*, 1997), sont en concordance et révèlent une diversité structurée selon les 3 types, Forastero, Criollo, et la forme intermédiaire Trinitario, hybride entre les deux groupes précédents. Les résultats suggèrent que les deux groupes principaux Forastero et Criollo, séparés par la barrière andine, auraient subi des événements de différenciation indépendants, les Forasteros du côté amazonien et les Criollos en Amérique Centrale (Cuatrecasas, 1964). Au niveau intra-groupe, le Forastero présente une légère différenciation, avec deux groupes, le Forastero haut amazonien, lequel possède une très haute diversité et le Forastero bas amazonien, y compris les origines du Venezuela et de la Guyane française (Laurent, Risterucci et Lanaud, 1994). La structuration observée par les marqueurs morphologiques a aussi été confirmée par les marqueurs moléculaires, mais

aucune structuration géographique n'a pu être identifiée à l'intérieur de ces groupes.

II.3.4 - L'amélioration du palmier à huile et les ressources génétiques

La production de semences commerciales du palmier à huile est réalisée avec une base génétique étroite, malgré les vastes ressources génétiques en collection disponibles pour les programmes d'amélioration. Ces ressources génétiques sont en général peu évaluées et très coûteuses à entretenir, comme pour plusieurs autres cultures (Karp et Edwards, 1997 ; Kresovich *et al.*, 1997 ; Tanksley et McCouch, 1997). Pour l'élargissement de la base génétique de cette culture, des variétés primitives ou populations semi-sauvages d'*E. guineensis* trouvées en Afrique tropicale humide, considérées comme le centre d'origine de l'espèce (Zeven, 1964), ont été collectées. Une espèce sauvage apparentée, l'*E. oleifera*, largement disponible en Amérique tropicale a fait aussi l'objet d'un important effort de prospection et de collecte (Escobar, 1981 ; Hartley, 1988 ; Meunier, 1975 ; Meunier, 1976 ; Rajanaidu, 1985a ; Rajanaidu, 1985b). Malgré tous ces efforts et de graves problèmes de maladie présents dans la culture d'*E. guineensis*, les programmes d'amélioration n'ont pas encore été capables d'exploiter la diversité génétique présente dans les collections de ces deux espèces, *E. guineensis* et *E. oleifera* pour les sorties variétales du palmier à huile (Rajanaidu, 1994 ; Simmonds, 1993 ; Yuan and Weng, 1996).

Un des motifs de cette faible utilisation de ressources génétiques du palmier à huile, est en partie dû à la faible caractérisation de ces collections, ajouté au fait que l'introgession de gènes nouveaux provenant des populations sauvages ou des espèces apparentées, a comme conséquence une réduction sévère des rendements. Les rétrocroisements successifs, qui permettent d'ajouter de nouvelles caractéristiques à des génotypes déjà plus performants, sont trop lents et coûteux pour pouvoir être utilisés sur des espèces pérennes comme le palmier à huile (Frankel et Brown, 1985). On constate que pour toutes les espèces, comme pour le palmier à huile, la grande majorité des ressources génétiques stockées dans les collections n'apporte que très peu de contribution à la sortie variétale actuelle.

II.3.5 - “Core collection”

II.3.5.1 - Considérations générales

Par définition, une “core collection” est un ensemble réduit d’accessions, issu d’une collection existante. Ce groupe d’individus est choisi de façon à représenter la partie la plus importante de la diversité génétique présente dans la collection de base (Brown, 1995 ; Brown, 1989 ; Frankel et Brown, 1984). L’établissement d’une “core collection” a pour objectif principal de faciliter et promouvoir l’utilisation des ressources génétiques contenues dans les collections (Brown, 1989 ; Hamon *et al.*, 1997). Cette utilisation serait rendue possible par la caractérisation plus fine d’un effectif plus réduit de plantes, la “core collection”, représentative de la diversité génétique de la collection de base.

Fréquemment, de nombreuses collections voient leur taille augmenter de manière considérable, rendant leur caractérisation très difficile et posant même de sérieux problèmes au niveau de leur conservation. Du fait de cette caractérisation insuffisante, ces collections sont très faiblement utilisées au sein des programmes d’amélioration, et l’allocation des moyens financiers nécessaires à leur entretien est de plus en plus difficile à obtenir. La constitution d’une core collection permettrait d’alléger de tels problèmes de caractérisation insuffisante et de faible utilisation de ces collections.

Le principe de constitution d’une “core collection” est : 1) l’élimination des “redondances” génétiques et, 2) la réduction des similarités génétiques entre les individus choisis. Il est alors possible de représenter la variabilité génétique présente dans la collection originale, par un nombre restreint d’individus (Brown, 1995).

Différentes stratégies sont envisageables pour la constitution d’une “core collection”, essentiellement en fonction du niveau de connaissance de l’espèce étudiée. Les considérations ne sont pas les mêmes pour une espèce cultivée ou pour un de ses parents sauvages, ces derniers présentant, en général, une diversité génétique moins connue et structurée différemment. Les objectifs de base d’une telle “core collection”, sont d’assurer une représentation de la diversité

allélique i) la plus complète possible, ii) avec le minimum d'individus (Brown, 1989). La première stratégie proposée a été celle de l'échantillonnage au hasard. Cette proposition est basée sur la théorie formulée en considérant les populations en équilibre génétique pour les allèles neutres (Ewens, 1972). Selon cette théorie, le choix au hasard d'au moins 10% des individus d'une population, permettra avec une probabilité de 95%, de capturer au moins 70% des allèles existants dans cette population (Brown, 1989). D'autres stratégies ont été postérieurement proposées, pour assurer et même améliorer l'efficacité de la "core collection", selon le degré d'information existant sur la collection de base. Pour une synthèse voir Hamon *et al.* (1997).

Les stratégies d'échantillonnage, en plus des "données de passeport", peuvent prendre en considération les marqueurs agromorphologiques et les marqueurs neutres. Ces stratégies sont classées selon le modèle adopté lors de la réalisation de l'échantillonnage. Les principaux modèles d'échantillonnage dont on dispose sont : a) le modèle des "allèles neutres" avec l'échantillonnage au hasard ; b) la structuration hiérarchique basée sur des caractéristiques majeures telles que l'origine écogéographique ou des caractères phénotypiques spécifiques ; c) l'échantillonnage basé sur les estimations de paramètres génétiques évalués sur des données quantitatives ou qualitatives (Bataillon *et al.*, 1996 ; Brown, 1989 ; Crossa *et al.*, 1993 ; Gepts, 1997 ; Schoen et Brown, 1997 ; Spagnoletti Zeuli et Qualset, 1993 ; Van Hintun, 1997) et une stratégie globale d'échantillonnage (Noirot *et al.*, 1996). L'efficacité de ces stratégies est hautement dépendante de la qualité d'information existant sur la collection de base. Les spécialistes insistent sur l'utilité des marqueurs moléculaires pour aider à maximiser la richesse allélique d'une "core collection" en améliorant le rapport entre la diversité capturée et le nombre d'individus choisis (Bhat *et al.*, 1997 ; Bonierbale *et al.*, 1997 ; Gepts, 1997 ; Lanaud et Lebot, 1997 ; Schoen and Brown, 1997).

II.3.5.2 - Intérêt de la constitution d'une "core collection" pour le palmier à huile

Des efforts considérables ont été faits sur les vingt dernières années, pour la mise en collection d'un nombre très important d'individus, couvrant l'aire de distribution de l'espèce *E. guineensis* et *E. oleifera*.

En prenant comme exemple la collection d'*E. guineensis* de l'Institut de recherche sur le

palmier à huile de Malaisie (PORIM), celle-ci comporte plus de 60 000 individus et occupe plus de 400 hectares de surface (Rajanaidu, 1994). Plusieurs autres collections de palmier à huile, moins grandes que celle du PORIM, sont disponibles dans d'autres centres de recherche tout particulièrement en Indonésie, en Côte d'Ivoire, au Nigéria et au Brésil.

Le palmier à huile est une espèce pérenne, pour le moment multipliée exclusivement par graine. Un hectare de collection ne comporte que 143 individus. La caractérisation agromorphologique d'une collection représente une activité continue sur une période longue, jusqu'à la phase adulte de la plante, environ huit ans après la mise en champ.

La constitution d'une "core collection" pour le palmier à huile devient fortement intéressante, si l'on considère les particularités de cette espèce, et surtout le faible niveau actuel d'utilisation des ressources génétiques existantes en collection. L'apport des techniques de marquage moléculaire pour l'évaluation et la caractérisation de la diversité génétique des collections est devenu courant, permettant d'accroître le niveau d'information existant sur ces collections. Cette étape est essentielle pour la constitution d'une "core collection" ayant une représentativité correcte de la collection de base et permettrait de promouvoir l'utilisation plus efficace de ces ressources génétiques.

II.4 - LES GÉNOMES DES PLANTES ET LES MARQUEURS MOLÉCULAIRES

Les végétaux supérieurs (eucaryotes), parmi lesquels se situent les espèces cultivées comme le palmier à huile, possèdent dans chacune de leurs cellules trois génomes avec toutes les fonctions nécessaires à leur expression : le génome nucléaire, le génome mitochondrial et le génome chloroplastique. Les estimations pour le nombre de gènes codés par l'ensemble de ces génomes peuvent varier énormément, entre 20 000 et 50 000, selon l'espèce végétale (Bennett et Leitch, 1995 ; Dean et Schmidt, 1995 ; Forgeois *et al.*, 1994 ; Lapitan, 1992).

Les trois génomes présentent chacun leurs propres caractéristiques, telles que taux de mutation, vitesse d'évolution et type de transmission (hérédité maternelle ou non). Les techniques de marquage moléculaire permettent d'accéder à ces génomes, de caractériser les différences entre individus même les plus proches et de mettre en évidence des liaisons entre individus ou groupes

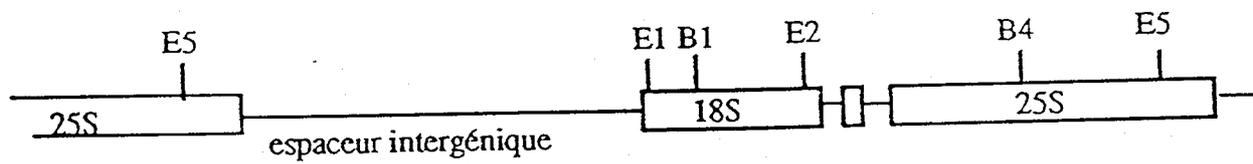


Figure II.6 - Schéma d'organisation de l'ADNr de radis (d'après Delseny et al., 1990) montrant l'espaceur intergénique (IGS), les régions de l'ADNr très conservées codant 18S et 25S et la localisation des sites de restriction pour les enzymes *Eco* RI (E) et *Bam* HI (B).

d'individus (Dean and Schmidt, 1995 ; Forgeois, Sourdille and Bonjean, 1994 ; Lapitan, 1992).

II.4.1 - Le génome nucléaire

La majorité de l'ADN cellulaire est constitué par l'ADN nucléaire, contenu dans le noyau, sous forme de plusieurs filaments constitués d'une molécule linéaire de plusieurs dizaines de millions de nucléotides, et condensés lors des divisions cellulaires, sous forme d'unités appelées chromosomes (Forgeois, Sourdille and Bonjean, 1994).

L'ADN génomique nucléaire peut être classé en trois catégories de séquences (Forgeois, Sourdille and Bonjean, 1994) :

- les séquences uniques ou faiblement répétées (1 à 10 copies), qui constituent la partie codante du génome. Ces séquences sont en général bien conservées entre les génomes d'espèces apparentées. Le polymorphisme détecté par RFLP avec les sondes de ADNc apparaît en général à proximité de ces séquences codantes, soit dans les introns, soit dans les séquences flanquantes.

- les séquences moyennement répétées, de 10 à 10 000 copies, y compris quelques gènes connus comme ceux codant pour les ARN ribosomiaux et les histones.

- les séquences hautement répétées (10 000 à 10 millions de copies), avec des fonctions très peu connues (Mighell *et al.*, 1997). Elles constituent la plus grande partie du génome dans certaines espèces. Ces séquences sont beaucoup plus variables (même au niveau intra population) que les séquences codantes.

Les gènes nucléaires codant pour les ADN ribosomiaux (ADNr) constituent une famille multigénique et sont présents en de nombreux exemplaires dans le génome des plantes supérieures. Ils sont organisés en unités constituées par une séquence codant pour les ARN ribosomiaux 18S, 5S et 25S et un espaceur inter génique IGS, situé entre les séquences codant pour les ARN 25S et 18S (Figure II.6). Les IGS sont constitués de sous unités longues d'environ 100 à 400 pb et répétées en tandem. Le nombre de copies de ces unités varie selon l'espèce de 200 à 22.000 par génome haploïde, et sont localisées au niveau des organisateurs nucléolaires (Hemleben *et al.*, 1988 ; Rogers et Bendich, 1987). Les gènes ribosomiaux sont sujets à une évolution concertée, homogénéisée par crossing-over inégal (Dover, 1988 ; Smith, 1976).

Un gène d'ADNr est donc constitué d'une séquence hautement répétée dans le génome ce qui est un avantage pour la révélation des RFLP, par rapport à des séquences uniques comme les ADNc. D'autre part, le génome des plantes comprend souvent plus d'un locus d'ADNr ce qui augmente encore le taux de polymorphisme détecté entre individus.

Les séquences codantes ont une taille constante d'environ 6 kb et, du fait de l'évolution concertée des gènes ribosomiques, elles sont très conservées chez toutes les plantes. Les variations de séquence au niveau des régions codantes entraînent la disparition de certains sites de restriction et peuvent donc être révélées par RFLP. Ces variations sont plus fréquemment détectées au niveau interspécifique ou plus en amont dans l'échelle évolutive (Delseny *et al.*, 1990 ; Zimmer *et al.*, 1988). Ces séquences nucléotidiques sont très conservées au niveau intra-spécifique, mais montrent des variations caractéristiques entre les espèces ou entre les groupes en cours de spéciation. Ce type de variation est donc plus informatif pour la phylogénie que les variations de l'IGS. C'est pourquoi, plusieurs études de phylogénie ont été faites en considérant de telles variations, dans différentes espèces (Cordesse *et al.*, 1990 ; Delseny *et al.*, 1990 ; Jorgensen *et al.*, 1987 ; Springer *et al.*, 1989 ; Zimmer, Jupe and Walbot, 1988).

La présence de séquences très conservées encadrant des régions très variables est une autre raison pour laquelle l'ADNr a été très utilisé pour l'analyse de la diversité. Les RFLP de l'ADNr peuvent être révélés à l'aide de sondes hétérologues provenant d'espèces évolutivement très éloignées, aussi éloignées que les monocotylédones et les dicotylédones. Ces sondes peuvent s'hybrider en conditions stringentes sur les séquences 18S et 25S et permettent ainsi de révéler des RFLP de l'IGS adjacent. Dans l'espaceur, les variations de séquences sont beaucoup plus fréquentes qu'au niveau des séquences codantes et elles peuvent être révélées même au niveau intra spécifique. L'IGS présente des variations importantes tant au niveau de la séquence nucléotidique que de sa longueur qui peut varier de 1 à 12 kb (Hemleben *et al.*, 1988 ; Rogers and Bendich, 1987). Les variations intra-spécifiques de longueur de l'IGS sont le plus souvent dues à des variations du nombre de sous unités (Rogers and Bendich, 1987). Pour les études de diversité, les informations sur la variation de la taille de IGS sont fort utiles et peuvent même révéler une structuration au niveau de populations naturelles. Cette région est hautement homogène à l'intérieur de groupes et présente des variations caractéristiques entre des groupes mêmes proches (Dover, 1988 ; Zimmer,

Jupe and Walbot, 1988). La vitesse d'évolution et d'homogénéisation des IGS implique qu'ils deviennent rapidement spécifiques d'un groupe de plantes isolées géographiquement (Besse *et al.*, 1993 ; Cordesse, Second and Delseny, 1990 ; Glaszmann *et al.*, 1990 ; Lanaud *et al.*, 1992 ; Laurent *et al.*, 1993b) et peuvent constituer de bons marqueurs de l'histoire évolutive des complexes d'espèces. De plus, vu le mode d'évolution particulier de l'ADNr, son analyse par RFLP est complémentaire de l'étude RFLP du génome nucléaire par les ADNc.

II.4.2 - Les génomes cytoplasmiques

La majorité des Angiospermes présentent une hérédité maternelle stricte pour les génomes chloroplastique et mitochondrial. Des exceptions existent et elles sont bien plus importantes pour le génome chloroplastique, avec plus de 40 cas déjà décrits. La majorité de ces exceptions concerne une hérédité biparentale occasionnelle. Pour le génome mitochondrial, seulement 6 cas d'hérédité maternelle non stricte ont été enregistrés pour le moment. Les Gymnospermes présentent une hérédité paternelle et biparentale à la fois pour ces génomes, avec des espèces qui présentent une hérédité maternelle pour les mitochondries et une hérédité paternelle pour les chloroplastes, comme certains pins (Latta et Mitton, 1997 ; Mogensen, 1996 ; Reboud et Zeyl, 1994 ; White, 1990)

La caractéristique présentée par les génomes cytoplasmiques, d'être hérité uniparentalement sans recombinaison pour la majorité des plantes, fait que ces génomes sont très utilisés pour les études de phylogénie, de flux géniques et de dispersion par voie de graines. Le fait que ces génomes soient très conservés, limite beaucoup leur utilité pour les études intraspécifiques, surtout pour le génome chloroplastique.

Cependant, de nouvelles techniques de biologie moléculaire, telles que les RFLP sur le génome chloroplastique et mitochondrial et le séquençage du gène chloroplastique *rbcL* codant pour la synthèse de l'enzyme ribulose-1-5-biphosphate carboxylase/oxydase, permettent d'accéder directement à ces génomes, et de dégager les différences entre individus même les plus proches. Ces techniques permettent la réalisation des études de liaisons entre individus ou groupes d'individus sur plusieurs niveaux taxonomiques (Dean and Schmidt, 1995 ; Dumolin-Lapègue *et al.*, 1997 ; Hilu et Stalker, 1995 ; Lapitan, 1992 ; Luo *et al.*, 1995).

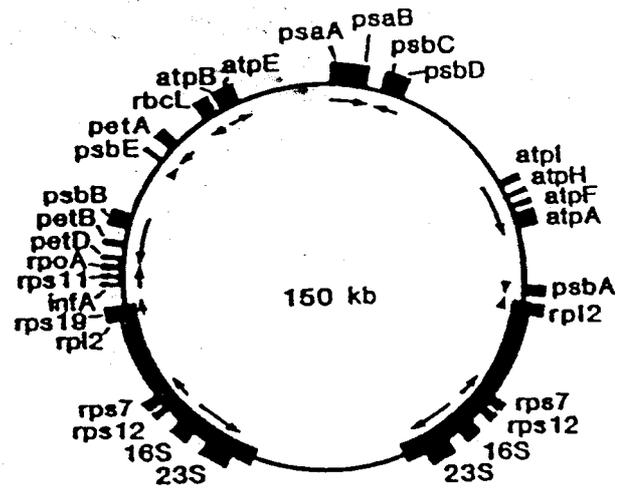


Figure II.7 - Organisation des gènes chloroplastiques d'épinard (Palmer, 1987).

II.4.2.1 - Le génome chloroplastique

Le génome chloroplastique des végétaux est constitué de molécules d'ADN circulaires (Figure II.7) avec une taille variable selon l'espèce (Birky, 1988 ; Palmer, 1987). Une cellule foliaire typique contient quelques milliers de molécules d'ADN chloroplastique génétiquement identiques, alors qu'une cellule non foliaire ne contient que le dixième de cette valeur. Chaque chloroplaste renferme entre 20 et 900 copies du génome chloroplastique, qui sont groupées en plusieurs structures cytologiques, les nucléoïdes. Chaque copie du génome chloroplastique est constituée de 100 à 220 kpb. La plupart des plantes à fleurs présentent une organisation similaire de ce génome (Bendich, 1987 ; Palmer, 1987 ; Palmer, 1990).

La molécule circulaire de l'ADN chloroplastique est typiquement constituée d'une grande et d'une petite région codantes à séquence unique, séparées par deux régions identiques, aussi codantes mais répétées et inversées. Cette région répétée inversée est présente chez toutes les plantes étudiées, sauf chez un groupe de légumineuses, où cette région a été perdue par délétion. Une fréquence de recombinaison intramoléculaire très élevée entre les deux régions répétées a été vérifiée sur l'épinard et sur le blé (Ogihara *et al.*, 1988 ; Palmer, 1987). La caractérisation complète du génome chloroplastique de plusieurs plantes cultivées montre la présence de plus d'une centaine de gènes codant pour les polypeptides et plus d'une trentaine de gènes codant pour différents ARNr (Sugiura, 1993).

La plus grande partie de la variation de taille du génome chloroplastique d'une espèce à l'autre est due à des variations de taille de la région répétée inversée. Les variations dans la complexité de ce génome sont surtout dues à des mutations de longueur de séquence, par l'addition ou délétion de séquences. La substitution de nucléotides est relativement faible, seulement la moitié du taux de substitution du génome nucléaire, mais trois fois plus que le taux de substitution du génome mitochondrial (Palmer, 1987 ; Weil, 1988 ; Wolfe *et al.*, 1987).

Le génome chloroplastique présente une évolution très lente, en comparaison avec les génomes nucléaire et mitochondrial. Il varie très peu en taille, il garde encore une ancienne région répétée, il présente un ordre primitif pour presque toutes les plantes et il accumule très lentement les

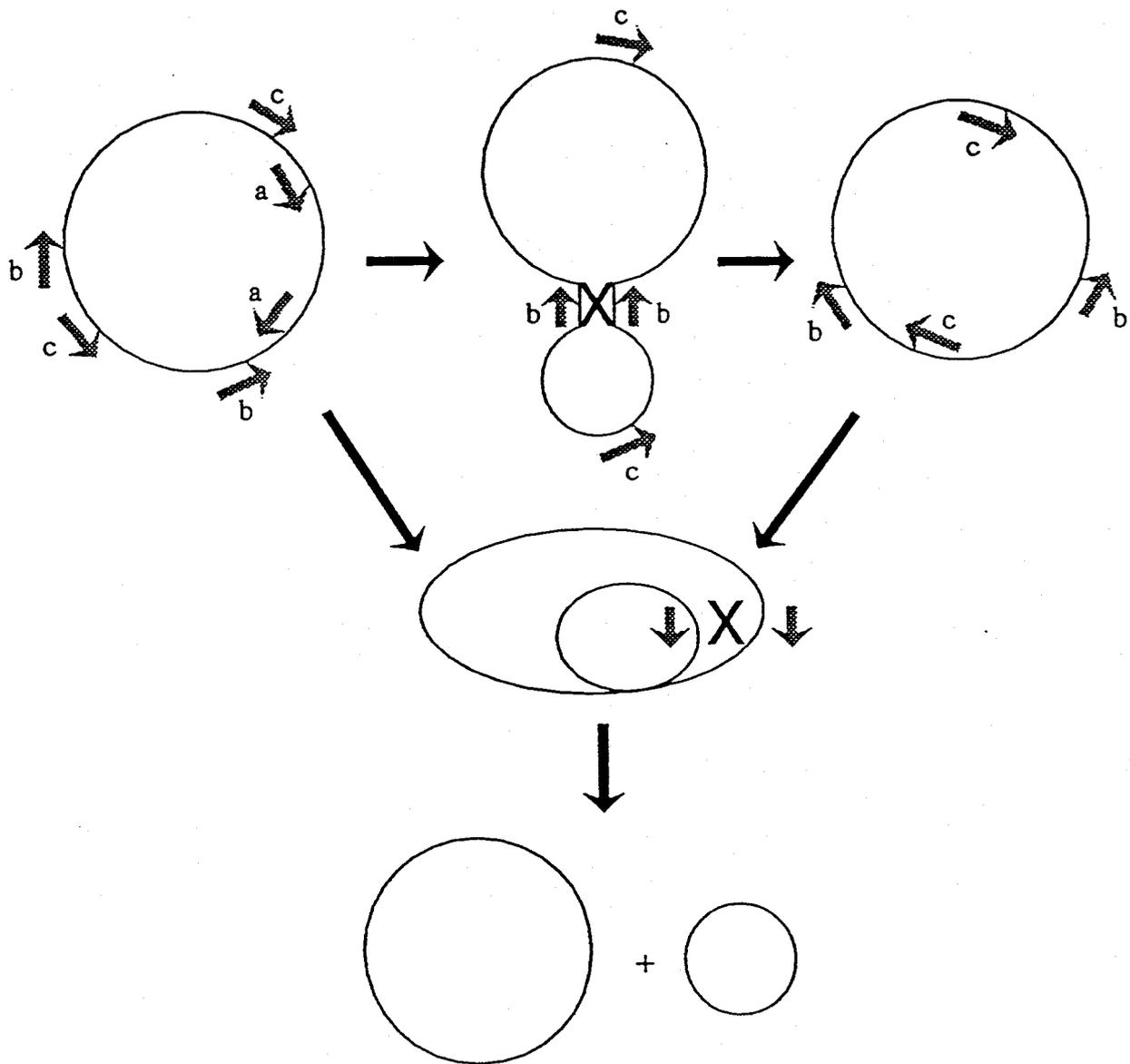


Figure II.8 - Mécanismes de recombinaison de molécules d'ADN mitochondrial avec la génération de molécules différentes (Quétier et al., 1985). Les combinaisons entre les deux séquences **b** et **c** inversées (flip-flop), donne des molécules de même taille avec l'inversion du segment entre ces deux répétitions inversées. La recombinaison entre deux répétitions non inversées **a** (loop-out) donne deux molécules subgénomiques de tailles différentes équivalentes à la distance qui sépare les deux répétitions.

mutations ponctuelles (Ogihara, Tarachi and Sasakuma, 1988 ; Weil, 1988 ; Wilson, 1990 ; Wolfe, Li and Sharp, 1987). Ces caractéristiques, ajoutées à d'autres telles que l'hérédité maternelle pour la plupart des espèces, l'abondance et la facilité de l'isoler à partir de tissu foliaire et la petite taille du génome, permettent que les variations du génome chloroplastique soient couramment utilisées pour les études de différenciation et de phylogénie à différents niveaux taxonomiques (Close *et al.*, 1989 ; Dally et Second, 1990 ; Doyle *et al.*, 1990 ; Green *et al.*, 1986 ; Waugh *et al.*, 1990 ; Baurens *et al.*, 1997 ; Breiman *et al.*, 1991 ; Downie, 1997 ; Doyle *et al.*, 1997 ; Havey, 1991 ; Latta and Mitton, 1997 ; Luo *et al.*, 1995 ; Morton *et al.*, 1997 ; Pax *et al.*, 1997 ; Petit *et al.*, 1993 ; Song et Osborn, 1992 ; Terauchi *et al.*, 1992).

II.4.2.2 - Le génome mitochondrial

L'ADN du génome mitochondrial des végétaux se présente sous la forme d'une population de molécules circulaires organisées à partir d'une molécule maîtresse, propre à chaque espèce (Quétier et Vedel, 1977 ; Palmer, 1988). Ces molécules circulaires sont plus ou moins hétérogènes et les mitochondries peuvent aussi être constituées de molécules linéaires (Bendich, 1993 ; Leaver et Gray, 1982). La taille du génome mitochondrial varie beaucoup plus que la taille du génome chloroplastique entre espèces végétales, mais beaucoup moins que celle du génome total : de 150 kb pour des espèces des genres *Solanum* et *Brassica* jusqu'à 2 500 kb pour des espèces du genre *Cucumis*. Le génome mitochondrial comporte très peu de séquences répétées par rapport aux autres génomes (Quétier et Vedel, 1977 ; Palmer, 1988).

La molécule principale est constituée par des séquences uniques, avec deux régions ayant des séquences répétées inversées et une région de séquence dupliquée (Figure II.8). Le taux de substitution de bases dans le génome mitochondrial est 3 fois inférieur à celui du génome chloroplastique et environ 6 fois moins que pour le génome nucléaire (Wolfe, Li and Sharp, 1987). Toutes les espèces de plantes étudiées présentent ces gènes dans le même ordre (Lejeune *et al.*, 1988 ; Oda *et al.*, 1992 ; Palmer, 1990). Les variations de la complexité par réarrangement de l'ordre des gènes, dues tant à un niveau élevé d'inversions qu'aux délétions au sein des séquences répétées, expliquent l'évolution du génome mitochondrial, alors qu'il présente une très faible variation sur les séquences primaires (Palmer, 1988 ; Palmer, 1990 ; Palmer, 1991).

Il présente une hérédité maternelle, pour la plupart des plantes supérieures, sans recombinaison, ce qui permet la fixation rapide d'un génotype au sein d'une population ou groupe d'individus, amenant une rapide différenciation entre les groupes (Lejeune *et al.*, 1988). Les études du polymorphisme mitochondrial sont utilisées pour évaluer les relations inter et intra espèces (D'Hont *et al.*, 1987 ; Laurent *et al.*, 1993a ; Luo et Boutry, 1995).

II.4.3 - La quantité d'ADN nucléaire chez les plantes

La quantité d'ADN nucléaire peut varier de façon importante entre des espèces très proches (Dean and Schmidt, 1995), malgré un nombre constant de chromosomes (Bennett and Leitch, 1995), comme chez les caféiers (Barre *et al.*, 1996 ; Cros *et al.*, 1995), le maïs et le sorgho (Bennett and Leitch, 1995 ; Laurie et Bennett, 1985), *Medicago sp* (Blondon *et al.*, 1984), *Helianthus sp* (Michaelson *et al.*, 1991 ; Natali *et al.*, 1993) et le soja (Bennett and Leitch, 1995). De telles variations intra-spécifiques sont assez fréquentes. Des variations intra espèces de 4% à 288% ont été détectées chez un ensemble d'Angiospermes (Bennett and Leitch, 1995).

La proportion de séquences répétées dans le génome de certaines espèces peut atteindre 60% à 80%, comme pour le maïs ou pour le blé (Flavell *et al.*, 1974). Cette proportion peut augmenter considérablement au sein des génomes de grande taille, où la partie non codante peut représenter jusqu'à 90% et même plus, de l'ADN (Flavell, 1980). Les variations de cette partie non codante de l'ADN sont responsables des variations de la taille du génome, au niveau intra spécifique (Bennett and Leitch, 1995 ; Flavell *et al.*, 1974). La partie codante de l'ADN est d'une grande stabilité et présente une forte homologie entre espèces et même à des niveaux taxonomiques plus élevés (Charlesworth *et al.*, 1994 ; Dean and Schmidt, 1995 ; Lapitan, 1992).

La variation de la taille du génome est corrélée à divers caractères phénotypiques, au niveau chromosomique, nucléaire, cellulaire, tissulaire et même de l'organisme, tels que la taille et le volume des chromosomes, la durée du cycle de la mitose et de la méiose, la longueur de génération, la radiosensibilité et aussi à des facteurs de l'environnement tels que le stress et le climat (Bennett and Leitch, 1995). La comparaison de la quantité de l'ADN a été utile à plusieurs études cytotaxonomiques et d'évolution (Ohri *et al.*, 1994).

La quantité d'ADN chez le palmier à huile a déjà été étudiée, et 4 estimations différentes ont été trouvées : 2 pg ont été trouvés par la méthode de Feulgen pour le noyau diploïde (Jones *et al.*, 1982) ; 4,0 à 4,8 pg pour le 4C, sans mention de la méthode (Bennet *et al.*, 1997) ; et finalement $3,76 \pm 0,09$ pg par la méthode de cytométrie en flux (Rival *et al.*, 1997). Deux de ces estimations ont été faites en vue de vérifier la stabilité du génome au cours du processus de multiplication par culture de tissu (Jones *et al.*, 1982). Aucune référence n'a été trouvée sur l'évaluation de la quantité d'ADN dans le genre *Elaeis*, pour vérifier l'existence d'une variation inter ou intra spécifique pour ce caractère. Une telle variation de la taille du génome, au niveau inter et intra-espèce au sein du genre *Elaeis*, pourrait expliquer en partie les problèmes de stérilité partielle des hybrides entre les deux espèces, et qui varient selon l'origine des génotypes impliqués dans les croisements.

Plusieurs méthodes peuvent être utilisées pour estimer la quantité d'ADN des noyaux des cellules de végétaux supérieurs, comme la réalisation d'analyses chimiques ou d'analyses microdensitométriques (Bennett and Leitch, 1995). Actuellement, la technique de cytométrie en flux est de plus en plus répandue (Bennett and Leitch, 1995) du fait de sa précision et de sa facilité de mise en oeuvre (Galbraith *et al.*, 1983).

II.4.4 - L'intérêt des marqueurs moléculaires

L'utilisation de marqueurs agromorphologiques tant pour la caractérisation des ressources génétiques que pour l'amélioration des plantes reste limitée du fait de la très forte influence de l'environnement sur les marqueurs quantitatifs, de la subjectivité parfois des différents états pour les marqueurs qualitatifs et surtout de la faible disponibilité de marqueurs efficacement utilisables.

Pour le palmier à huile (*E. guineensis* Jacq.) par exemple, le seul caractère morphologique non influencé par l'environnement et utilisé pour caractériser des groupes d'individus, est l'épaisseur de la coque (Hartley, 1988). Ce caractère ne permet que la gestion de la généalogie de quelques origines et n'est observable que pendant la phase reproductive de la plante, c'est à dire au minimum 30 mois après la germination de la graine, lors de l'apparition des premiers régimes.

Le développement des marqueurs moléculaires a fait suite à celui des marqueurs biochimiques (isoenzymes). Ils ont été développés vers la fin des années 70, avec les recherches sur le génome humain. Aujourd'hui ils sont disponibles (RFLP, RAPD, AFLP, SSR, STS, CAP) en grand nombre (de Vienne, 1998).

Ces marqueurs constituent des balises directement placées sur le génome et permettent une meilleure caractérisation des populations ainsi que des études plus précises des relations et de l'histoire évolutive entre ces populations. Plusieurs approches et méthodes de la biologie moléculaire sont couramment utilisées pour étudier la variabilité génétique des plantes. Le polymorphisme de séquences des génomes cytoplasmiques et du génome nucléaire est utilisé pour des études de phylogénie et de diversité génétique. Ce polymorphisme de séquence peut être révélé par différentes méthodes. Il peut être propre à une région spécifique du génome, et l'analyse du génome ou des régions plus conservées est plus appropriée aux études de phylogénie, alors que l'analyse des régions non codantes, présentant des taux de mutations *a priori* plus élevés, est plus appropriée aux études de diversité génétique au niveau des espèces ou des individus (Crozier, 1990 ; Dean and Schmidt, 1995 ; Lapitan, 1992 ; Palmer, 1987).

L'utilisation des marqueurs moléculaires présente de nombreux avantages pour la gestion et l'utilisation de ressources génétiques, surtout pour les espèces pérennes (Ayad *et al.*, 1985). Le fait que les marqueurs moléculaires soient en majorité neutres vis-à-vis des influences environnementales, explique que ces marqueurs ont connu un important essor (Beckmann et Soller, 1983 ; Helentjaris, 1991 ; Rafalski *et al.*, 1991 ; Tanksley *et al.*, 1989 ; Vos *et al.*, 1995 ; Waugh et Powell, 1992), notamment les marqueurs RFLP et PCR appliqués aux espèces pérennes, avec pour objectif :

- l'analyse de la diversité génétique et la connaissance de la structuration de la variabilité génétique pour la gestion des ressources génétiques.
- l'identification des variétés et la protection variétale.
- l'évaluation des génotypes pour des caractères très influencés par le milieu ou d'évaluation longue et coûteuse.
- la caractérisation de matériel inconnu pour son utilisation dans les programmes d'amélioration.

- le suivi des transferts de gènes majeurs dans les programmes de sélection généalogique, rétrocroisements ou transformations génétiques.
- le contrôle de qualité dans les programmes d'hybridation et de production de semences commerciales.

II.4.4.1 - Les marqueurs RFLP

Les marqueurs RFLP (Restriction Fragment Length Polymorphism), comme les autres marqueurs moléculaires, révèlent directement le polymorphisme de l'ADN. Cette technique a été utilisée pour la première fois pour l'étude des adénovirus (Grodziker et al., 1974). Mais ce sont les études de la carte génétique humaine qui ont donné à cette technique son essor (Botstein *et al.*, 1980), puis elle a été utilisée à l'étude des végétaux (Burr *et al.*, 1983 ; Tanksley *et al.*, 1988).

La technique est basée sur la mise en évidence de variations du génome, générées par des mutations ponctuelles, délétions, inversions et translocation qui sont réparties sur l'ensemble du génome et entraînent des variations des sites de reconnaissance des enzymes de restriction ou endonucléases. L'action de ces enzymes va donner des fragments de taille différente, en nombre variable selon la complexité du génome. Pour les génomes plus complexes, ces fragments seront révélés spécifiquement par hybridation avec d'autres fragments d'ADN servant alors de sondes. Cette méthode de révélation dite Southern, due au nom de son inventeur, a été décrite pour la première fois en 1975 (Southern, 1975).

Les sondes servant à révéler le polymorphisme de l'ADN peuvent être de différentes origines et de différents types. Elles peuvent être soit homologues soit hétérologues, selon qu'elles proviennent de l'espèce analysée ou non. Les sondes d'ADN complémentaire (ADNc) sont les plus utilisées, elles sont issues de la partie codante du génome et correspondent le plus souvent à des séquences uniques ou faiblement répétées. Les sondes ADNc sont obtenues par transcription reverse *in vitro* des ARN messagers.

Les sondes d'ADN génomique sont constituées simplement d'un morceau d'ADN issu de n'importe quelle partie du génome. Elles peuvent correspondre à des séquences uniques, répétées,

ou englober les deux. Dans une banque de sondes d'ADNc de palmier à huile, les sondes obtenues à partir de l'ADN total possèdent environ 76% de sondes uniques ou faiblement répétées et celles de l'ADN total enrichi en ADN nucléaire sont constituées de 90% de sondes uniques ou faiblement répétées (Billot, communication personnelle).

Les marqueurs RFLP sont des marqueurs neutres, donc indépendants du milieu et du stade de développement du tissu; en nombre élevé, ils sont stables et hérités comme des facteurs mendéliens co-dominants (Evola *et al.*, 1986). Les RFLP sont très utilisés pour les études des relations génétiques de différents niveaux pour diverses espèces (Bachmann, 1992 ; Miller et Tanksley, 1990 ; Tanksley *et al.*, 1989), au sein desquelles plusieurs espèces tropicales comme le cacaoyer (Laurent, Risterucci and Lanaud, 1993a ; Laurent, Risterucci and Lanaud, 1994 ; N'Goran *et al.*, 1994), l'hévéa (Besse *et al.*, 1994 ; Besse *et al.*, 1993 ; Luo and Boutry, 1995) le cocotier (Lebrun *et al.*, 1995) et le palmier à huile (Barcelos *et al.*, 1997 ; Jack *et al.*, 1995 ; Shah *et al.*, 1994).

Par ces caractéristiques, les RFLP sont des marqueurs puissants et robustes ; cependant c'est une technique coûteuse et lourde à mettre en oeuvre. L'utilisation de produits radioactifs, rend la technique plus efficace, mais implique plus de contraintes de manipulation.

II.4.4.2 - La PCR et les AFLP

La méthode PCR est devenue l'une des techniques de base de toute la biologie moléculaire. En jouant notamment sur les amorces, plusieurs méthodes de détection du polymorphisme particulièrement performantes et efficaces ont été développées (RAPD, AFLP, AP-PCR, DAF, SCAR, STS, VNTR, VNDR, SSR).

Une des plus récentes et des plus élégantes méthodes, l'AFLP (Amplified Fragment Length Polymorphisme), est fondée sur la mise en évidence conjointe du polymorphisme de site de restriction, la base fondamentale du RFLP, et l'hybridation et l'amplification sélective à partir d'amorces arbitraires, point de base de la technique PCR. Cette technique combine donc la robustesse des RFLP avec la performance de la PCR. L'AFLP se présente comme très efficace pour la détection de diversité même au niveau intra population, avec une couverture très ample du

génomique et l'évaluation de plusieurs locus en un seul essai. Pour le moment les marqueurs AFLP sont traités comme des marqueurs hérités de façon dominante (Vos *et al.*, 1995 ; Paul *et al.*, 1997 ; Second *et al.*, 1996), (Hill *et al.*, 1996 ; Maughan *et al.*, 1996).