

## C. EVALUATION PAR LES MÉTHODES DE LA BIOLOGIE MOLÉCULAIRE

### 1. Nécessité de l'outil moléculaire dans les programmes de gestion des ressources génétiques

a. *Introduction.* Ce chapitre méthodologique est relativement plus développé, du point de vue de l'exposé des techniques, que ceux consacrés à la cytogénétique ou à l'évaluation agronomique (génétique quantitative et planification des expériences). A cela plusieurs raisons:

— Le lecteur concerné par la gestion des ressources génétiques est par sa fonction et ses habitudes de travail plus au fait des méthodes de l'agronomie ou de l'amélioration des plantes que de la biochimie. De ce fait les rappels et les orientations bibliographiques seront parfois de véritables initiations et pas seulement des mises en situation d'outils déjà manipulés.

— L'outil moléculaire, tel qu'il est présenté dans les manuels de biochimie n'a pas la même signification ni le même emploi que dans le domaine des ressources génétiques et de l'amélioration des plantes. Pour les biochimistes l'électrophorèse des protéines est d'abord une méthode de séparation pour repérer des protéines précises dont on cherche ensuite à déterminer des propriétés (poids moléculaire, structure, propriétés fonctionnelles, éventuellement séquence d'acides aminés), l'utilisation des enzymes de restriction pour la digestion des acides nucléiques est destinée à analyser complètement des fragments d'ADN clonés bien déterminés, ou à en préparer le clonage et la purification. En génétique des populations, l'emploi de ces méthodes est apparemment moins ambitieux il s'agit de repérer des différences héréditaires simples entre individus. La simple mise en évidence d'une telle différence, démonstration d'un polymorphisme génétique, est une fin en soi. Malgré la modestie de cet objectif, le généticien rencontrera des difficultés que le biochimiste aurait écartées: modicité quantitative de l'extrait de départ (c'est parfois d'un grain de pollen dont on veut repérer le génotype), démonstration de la stabilité héréditaire de la caractéristique différentielle repérée (donc étude de génétique mendélienne pour montrer par exemple que deux bandes d'isozymes ne diffèrent que par des modifications post-traductionnelles du produit d'un même gène, ou bien que derrière une seule différence (présence/absence) d'un isozyme migrant à un niveau donné plusieurs gènes interviennent...). Autrement dit les problèmes d'interprétation seront nouveaux par rapport à ceux du biochimiste et il est nécessaire d'éclairer autrement les présentations méthodologiques et de leur adjoindre les exemples d'analyse génétique classique appliquée à ces différences moléculaires.

— La biologie moléculaire connaît actuellement une évolution très rapide et les conséquences pour la pratique des ressources génétiques et la diffusion de plus en plus économique de ces moyens d'analyse imposent de présenter plus complètement ces méthodes que l'actualité immédiate ne le justifierait.

— Comme nous allons le montrer l'utilisation de ces méthodes est déjà beaucoup plus avancée qu'il n'y paraît et les premières publications (1966, 1967) concernant les polymorphismes enzymatiques mis en évidence chez l'homme ou la drosophile ont été à l'origine d'une remise en cause très

profonde des convictions de la génétique des populations. La paix des esprits n'est pas encore pleinement acquise et l'hésitation entre l'interprétation neutraliste et **sélectionniste** d'un polymorphisme impose de toujours réfléchir doublement sur les faits en partant de l'un ou l'autre point de vue.

Les méthodes présentées dans ce chapitre concernent soit les outils d'analyse des protéines: électrophorèse d'enzymes ou de protéines de réserve, en conservant précisément les fonctions ou l'activité, électrophorèse bidimensionnelle moyen de repérage de variations «tous azimuts» déployant largement les différences sans pouvoir identifier immédiatement les fonctions, soit les outils d'analyse de l'ADN (et des **ARN**): étude des similitudes globales des génomes (dénaturation, **renaturation**, hybridations moléculaires), repérage d'organisations chromosomiques différentes (coloration différentielle des zones **hétérochromatiques** (cf. paragraphe cytogénétique), hybridation in situ avec une sonde moléculaire d'ADN cloné, analyses de fragments digérés par des enzymes de restriction pour étudier le polymorphisme des sites de coupure pour les ADN **d'organites** cytoplasmiques et même l'ADN nucléaire). D'autres méthodes s'ajouteront avec les progrès du clonage des gènes et la détermination de l'organisation de séquences non codantes intégrées dans les gènes (**introns**) ou reliant des gènes de structures (**espaceurs**)...

*b. Illustrations de l'outil moléculaire déjà présentées dans le manuel.*

Pour faciliter la lecture des descriptions méthodologiques parfois rebutantes et pour encourager l'étude rappelons quelques exemples déjà présentés.

La monographie concernant le riz (tome I) a utilisé les études de polymorphismes enzymatiques pour définir les grands types de riz cultivés et la diversité des formes spontanées, les mêmes données servent de support à l'illustration des méthodes de classification numérique (chapitre Bases de données).

La monographie traitant du café a montré comment la délimitation du complexe d'espèces a permis d'exclure les **Psyllanthus** par exemple (électrophorèse d'enzymes) et préciser l'origine de l'espèce cultivée la plus importante **Coffea arabica** (polymorphisme enzymatique et analyse des ADN **mitochondriaux** par électrophorèse des éléments de digestion par les enzymes de restriction). Ces données permettent d'orienter **convenablement** les stratégies d'amélioration des caféiers par hybridation **interspécifique**.

L'étude par le polymorphisme enzymatique de la diversité et de l'organisation géographique des populations, a été illustrée par l'étude des **clines** des populations de mil du nord Côte-d'Ivoire (Tome I, Mil) et sert de base à la réflexion sur les stratégies d'échantillonnage (chapitre II, échantillonnage).

La caractérisation de l'état des collections et la mise en évidence de dérives aléatoires et systématiques ont été décrits par l'exemple des polymorphismes **d'Adh (F,S)** (chapitre Données de base). Enfin, ce même chapitre présente les paramètres et les concepts de la génétique des populations particulièrement adaptées pour exploiter les données acquises à partir des méthodes de biologie moléculaire.

*c. Utilisations pratiques de l'outil moléculaire par les sélectionneurs.* On peut regrouper ces utilisations sous trois grandes rubriques:

## Utilisation générale de marqueurs génétiques

Les méthodes biochimiques en mettant en évidence une diversité génétique importante cachée derrière l'homogénéité du «type adaptatif» d'une population, ont mis à la disposition des moyens de repérer les génotypes à l'aide des allèles distincts pour plusieurs locus. Rappelons quelques usages de ces marqueurs:

— Possibilité de trier plus efficacement dans une descendance l'état **allélique** d'un locus dont l'étude suppose l'observation tardive des plantes à éliminer (stérilités mâles géniques: il faut éliminer les plantes fertiles avant qu'elles ne puissent entraîner des pollinisations illégitimes) ou l'analyse d'une plante d'un génotype unique dans une famille en ségrégation, dont on voudrait étudier le comportement dans plusieurs conditions (repérage du système génique d'insensibilité à la photopériode pour le contrôle de la floraison (cf. Tome I, le Mil)) ou encore dont l'allèle recherché est récessif et ne peut être décelé directement (des gènes de nanisme que l'on veut conserver à travers l'enchaînement de croisement de retour où le parent récurrent est porteur de l'allèle dominant pour la taille normale). Pour ces trois exemples, l'identification d'un gène marqueur très étroitement lié au locus que l'on veut repérer rend la tâche du sélectionneur possible ou plus rapide.

— L'analyse des taux **d'allogamie** et d'une façon générale des échanges géniques entre populations voisines (cas de la pureté des maïs hybrides réalisés en parcelles en principe isolées des cultures voisines). Il suffit de disposer d'un (ou plusieurs) **gène(s) marqueur(s)** dont un état **allélique** ne peut provenir que des donneurs de pollen extérieurs et de repérer la fréquence de cet allèle dans la population de grains produits par la plante dont on veut mesurer le taux **d'allogamie** ou les variétés dont on veut garantir la pureté.

— Pour des espèces dont la production de graines peut résulter soit d'apomixie soit d'un processus sexué normal (par exemple *Panicum maximum* Vol. I) l'étude de l'hétérogénéité de la descendance à travers quelques marqueurs enzymatiques permet de révéler si la descendance a l'homogénéité d'un clone **apomictique** ou si c'est une famille en ségrégation (produite alors par voie sexuée).

— Au cours des schémas de sélection qui utilisent des croisements en retour pour intégrer l'état **allélique** d'un gène particulier dans un génotype donné (le parent récurrent) on est classiquement contraint de se fier aux probabilités pour atteindre l'**isogénie** souhaitée (chaque croisement en retour diminue en principe de moitié les secteurs du génome encore hétérozygote **donneur/receveur**). Si on dispose de plusieurs marqueurs, dispersés en plusieurs points du génome on peut accélérer le processus **d'isogénéisation** en choisissant parmi les plantes issues du croisement en retour, parmi toutes celles ayant l'allèle du donneur pour le caractère sélectionné (par exemple la résistance verticale à un parasite que l'on inocule) celles qui sont aussi (Déjà homozygotes pour les états **alléliques** du parent récurrent, au plus grand nombre de locus marqueurs étudiés.

— Que ce soit pour des raisons de protection commerciale des obtentions ou pour vérifier qu'une lignée nouvellement introduite est réellement originale, les marqueurs enzymatiques constituent d'excellents éléments pour constituer la fiche signalétique d'une variété ou d'une lignée.

## Connaissance de l'ensemble d'une collection

La diversité génétique des organites cytoplasmiques est indispensable si l'on désire se prémunir contre les risques de **monomorphisme**, longtemps sous-estimé dans ce domaine. La sensibilité à un parasite liée à un état particulier (cytoplasme Texas du maïs et sensibilité à l'*Helminthosporium* race T par exemple) a attiré l'attention des sélectionneurs pour diversifier les origines maternelles de leurs variétés, encore faut-il pouvoir apprécier s'il y a de véritables différences! La mise en évidence par lecture directe d'une diversité génétique permet de choisir beaucoup plus vite les croisements pour lesquels on cherche à mettre en évidence des effets réciproques et à repérer dans les descendances d'éventuelles caractéristiques à hérédité maternelle (stérilité mâle cytoplasmique par exemple).

Les études de distances génétiques entre les familles, ou les populations, d'une collection peuvent permettre d'améliorer les probabilités de découvertes de vigueur hybride en croisant des formes éloignées, car en gros il paraît bien vrai qu'un **hétérosis** important soit souvent associé à la richesse **allélique** des hybrides. Ce repérage de distances génétiques peut être aussi un moyen de trouver plus rapidement des formes hybrides stériles, quand elles sont recherchées (soit parce qu'on produit des métabolites ou des organes dont la production ne sera pas en compétition avec la production **grainière**, soit parce qu'on cherche à exploiter des formes dont on ne veut pas qu'elles puissent être multipliées en dehors du laboratoire d'obtention, etc...).

La mise en évidence d'organisations chromosomiques légèrement différentes (ou particulières par des études d'hybridation in situ avec des sondes d'ADN marqués) permet de prévoir certaines stérilités (petites inversions, translocations) liées à des différenciations structurelles du génome, et donc d'orienter le programme de sélection soit en éliminant immédiatement de nombreuses combinaisons qui ont une bonne probabilité d'être défectueuses soit en imposant immédiatement des recherches visant à restituer l'**homozygotie** structurale (programme de **polyploïdisation**, ou de créations de lignées d'addition ou de substitution).

## Recherches de certaines organisations fonctionnelles bien définies

Bien que l'amélioration des plantes doive se méfier de la visée **d'idéotypes** qui feraient penser qu'une variété puisse être créée par assemblage de structures géniques définies a priori, il existe cependant des possibilités de sélectionner de nouvelles structures bâties autour de l'état **allélique** particulier d'un gène majeur permettant un progrès technologique déterminant, (telle l'acquisition des gènes de nanisme et l'obtention d'une forte résistance à la verse permet d'accroître les apports d'engrais et d'augmenter la densité de semis). Il est possible que certaines caractéristiques biochimiques à hérédité simple puissent ainsi permettre de débloquer une amélioration **variétale** qui plafonne. Ce peut être un gène codant pour une protéine de réserve plus riche en acide aminé important (on sait que pour le maïs d'autres gènes qu'opaque 2 peuvent améliorer la qualité de la **zéine** sans avoir les désavantages fonctionnels de ce gène), ou un gène de structure proche d'un élément de régulation favorable de l'ensemble d'une chaîne métabolique (ce pourrait être le cas de l'allèle **ADH/S** lié à une régulation globale de l'**anaérobiose** ou certaines activités nitrate **réductase**, liés à l'assimilation de l'azote).

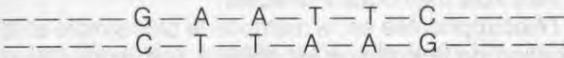
d. *Les travaux d'analyse moléculaire dans les programmes de gestion des ressources génétiques.* Remarquons que les exemples précédents, concernant l'utilisation des données de l'analyse moléculaire sont les sous-produits automatiques d'un bon travail d'inventaire et d'analyse des ressources génétiques.

Les travaux utilisant les méthodes de la biologie moléculaire devront être obligatoirement réalisés, et de façon extensive, dans trois domaines particuliers de la gestion des ressources génétiques:

- Le travail de classification génétique du complexe d'espèces étudié.
- L'étude de la structure d'ensemble des populations, **clines** ou coupages de populations voisines (formes cultivées et spontanées) pour planifier les regroupements ou les sites destinés à assurer des conservations dynamiques).
- L'inventaire et le contrôle des transformations qui ont lieu dans les collections de conservation, par exemple les sélections et les dérives cytoplasmiques, les «mutations» incontrôlables qui ont lieu dans les cultures in vitro, l'évolution au cours des campagnes de rajeunissement des lots de graines (repérage des «pollutions», des mélanges et des pertes) ou des effets de consanguinité des multiplications vivantes isolées.

Ainsi par leur intégration directe dans l'activité d'un centre de ressources génétiques, c'est un des outils de mesure les plus efficaces, et l'importance des applications actuelles et potentielles, les méthodes de la biologie moléculaire sont loin d'être des technologies de luxe, elles forment un des éléments indispensables au fonctionnement des centres de ressources génétiques de haut niveau de responsabilité.

e. *Présentation des méthodes.* Les gènes sont des molécules d'acide désoxyribonucléique (ADN). Rappelons-en succinctement et schématiquement la chimie: une molécule d'ADN est **constituée** d'une séquence de bases **nucléotidiques**, essentiellement au nombre de quatre: Adénine (A), Thymidine (T), Cytosine (C) et **Guanidine** (G). Deux «brins» d'ADN sont associés en une double hélice dans une molécule car à chaque A correspond T, et à chaque C correspond G selon le schéma suivant:



L'association des deux brins complémentaires est réalisée par des liaisons faibles. Leur séparation permet la réplication, chacun des brins reformant son complémentaire. Au cours de la lecture des gènes «structuraux», les molécules d'ADN sont transcrites en molécules homologues d'acide ribonucléique (ARN) qui, après «maturation», seront traduites en un polypeptide. A chaque triplet de bases de l'ADN, correspond un acide aminé (code génétique).

Les molécules d'ADN se trouvent principalement rassemblées dans les chromosomes du noyau de la cellule et en plus petites quantités dans les particules du cytoplasme, **mitochondries** et chloroplastes. Le tableau 22 présente quelques caractéristiques des trois types d'ADN chez les végétaux supérieurs (VEDEL et QUETIER, 1978).

**TABLEAU 22:** Comparaison de quelques propriétés physico-chimiques des ADN chloroplastiques (CP), mitochondriaux (MT), et nucléaires (N) des végétaux supérieurs.

	ADN CP	ADN MT	ADN N
Protéines liées	?	?	Histones
Dénaturation	homogène	homogène	hétérogène
Renaturation	rapide et homogène	rapide et homogène	lente et hétérogène
Répétition (%)	faible	faible	importante (30-80)
Taille de la molécule ou du génome haploïde	85-95 x 10 <sup>6</sup> d (42-48) $\mu$	50 x 10 <sup>6</sup> d (30) $\mu$	10 <sup>12</sup> d plusieurs cm
Hérédité	non mendélienne (maternelle)	non mendélienne (maternelle)	mendélienne

Étudier la variabilité génétique, c'est analyser le polymorphisme des molécules d'ADN. En d'autres termes, il s'agit de savoir, entre deux individus, deux populations ou espèces, quelles sont les bases nucléotidiques qui diffèrent (nombre, proportion, emplacement). Toute molécule d'ADN provient de la réplication d'une molécule d'ADN préexistante (au cours de la division cellulaire et de la méiose). La différenciation de l'ADN selon les organismes résulte des erreurs de réplifications, mutations, qui s'accumulent: substitutions, délétions, addition d'une base ou d'une séquence nucléotidique, réarrangements chromosomiques, etc...

La divergence des ADN portés par deux individus est donc liée au nombre de générations qui les séparent. Il est également probable qu'elle soit liée aux différences de l'environnement habituel des individus puisque, parmi toutes les mutations, sont sélectionnées celles qui procurent un avantage adaptatif dans un environnement donné. L'étude directe des séquences nucléotidiques des ADN est une tâche de longue haleine qu'il est, à l'heure actuelle, impensable de conduire sur un grand nombre de gènes, pour de nombreux individus.

On s'est donc orienté vers trois méthodes indirectes:

— *Étude des protéines:* l'électrophorèse est la méthode la plus simple et la plus utilisée. Une modification de la protéine est reliée à une modification de l'ADN d'un gène structural selon la relation «un gène — un polypeptide». On étudie généralement les protéines enzymatiques car plusieurs dizaines d'entre elles peuvent être facilement isolées spécifiquement, mais aussi les protéines structurales ou de réserves. Elle est également applicable à l'étude des enzymes codés par l'ADN cytoplasmique.

— *Étude de l'hybridation des ADN:* par opposition à l'étude des protéines, cette méthode s'intéresse principalement à la partie redondante (répétitions) des ADN. Elle est donc applicable aux ADN nucléaires.

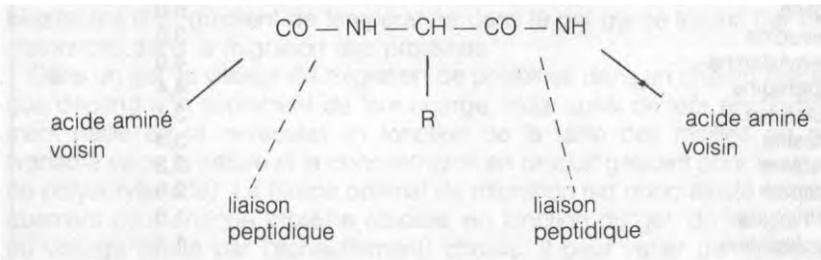
— *Fragmentation spécifique* de l'ADN par les enzymes de restriction: elle est applicable aux petites molécules d'ADN et en particulier aux ADN cytoplasmiques. Indépendamment de l'étude de la variabilité de ces ADN, elle fournit des marqueurs cytoplasmiques (recherche des filiations femelles par exemple).

## 2. Étude des protéines enzymatiques

a. *Le principe.* La séquence de nucléotides d'un gène structural est traduite avec une haute précision en une séquence d'acides aminés constituant une chaîne polypeptidique. Puisqu'une enzyme est constituée par une ou plusieurs chaînes polypeptidiques provenant d'un ou de quelques gènes structuraux, une variation dans la séquence **nucléotidique** se traduit (sauf exception: les cas de substitution redondante de nucléotides) par une substitution, délétion ou addition d'un acide aminé dans la constitution de l'enzyme codée. Une modification de la séquence d'acides aminés d'une enzyme ou d'une protéine structurale est donc directement **représentative** d'une substitution **allélique** au locus codant le polypeptide modifié.

Si plusieurs allèles sont présents et fonctionnels chez un individu (gènes dupliqués, **hétérozygotie**, **polypléidie**), les différentes formes enzymatiques ou les différentes formes de la protéine structurale, codées par les allèles, seront présentes chez cet individu. Chaque allèle est donc **codominant**, à l'exception des mutations ou autres modifications de l'ADN qui suppriment la transcription ou la traduction: les allèles non traduits sont récessifs.

La séquence des acides aminés des protéines est ainsi un phénotype de la structure génique. Malheureusement, comme pour l'ADN, il n'est pas techniquement possible, actuellement, de déterminer la séquence des acides aminés des protéines chez de nombreux individus. Les généticiens se sont donc orientés vers l'utilisation des propriétés physico-chimiques des protéines. Quand ils sont associés en une chaîne polypeptidique, les acides aminés sont sous la forme:



où R est spécifique de l'acide aminé.

Les 20 acides aminés communs dans les protéines (cf. tableau 23) se divisent en trois groupes en fonction de la charge électrique de cette chaîne latérale R:

— seize ont des chaînes latérales non **ionisables**; ils sont électrostatiquement neutres lorsqu'inclus dans une chaîne **peptidique**. Parmi ceux-ci, Histidine, **Cystéine** et Tyrosine ont des chaînes latérales **ionisables** mais ne sont pas ionisés dans les conditions de pH classiquement utilisées en électrophorèse.

— l'**Arginine** et la Lysine ont un groupe ammonium qui est en équilibre entre deux formes, neutre et chargée positivement, selon la concentration en protons (pH) du milieu.

— deux, l'Acide Aspartique et l'Acide Glutamique, ont un groupe R acide carboxylique et sont en équilibre entre des formes neutres et chargées négativement.

Une protéine, combinaison de trois types, aura une charge nette négative ou positive dépendant de la balance des charges en fonction du pH et de l'exposition des radicaux polaires, selon l'environnement et l'association des polypeptides.

Lorsque le pH est abaissé (concentration en protons augmentée), l'équilibre est déplacé vers les charges positives:  $\text{NH}_2$  devient  $\text{NH}_3^+$  et  $\text{COO}^-$  devient  $\text{COOH}$ . Au contraire, si le pH est élevé (on abaisse la concentration en ions hydrogène), la protéine prendra une charge globale négative.

**TABLEAU 23:** Composition en acides aminés de la protéine moyenne» (d'après KING et JUKES, 1969).

ACIDES AMINÉS	% DU TOTAL
<i>- basiques:</i>	
Lysine	7,2
Arginine	4,2
Total	11,4
<i>- neutres:</i>	
Serine	8,1
Leucine	7,6
Glycine	7,4
Alanine	7,4
Valine	6,8
Thréonine	6,2
Proline	5,0
Isoleucine	3,8
Phénylalanine	4,0
Asparagine	4,4
Glutamine	3,7
Tyrosine	3,3
Cystéine	3,3
Histidine	2,9
Méthionine	1,8
Tryptophane	1,3
Total	76,9
<i>- acides:</i>	
Acide aspartique	5,9
Acide glutamique	5,8
Total	11,7

Le pH auquel la protéine n'est pas chargée (**électrostatiquement** neutre) est appelé le point isoélectrique de cette protéine.

Si une mutation **allélique**, à un locus, a pour résultat le remplacement d'un acide aminé d'un groupe par un acide aminé d'un autre groupe, la charge nette de la protéine, de même que son point isoélectrique seront modifiés. Par exemple, la substitution d'un nucléotide dans le codon **AAC** pour donner **AAA** a pour résultat la substitution de la Lysine chargée positivement par l'Asparagine qui est neutre. Le changement peut se faire entre acides aminés chargés négativement et acides chargés positive-

ment: la transition AAG GAG conduit à la substitution de l'Acide Glutamique, chargé négativement, par la Lysine, chargée positivement.

Ces modifications de charge permettent de séparer les protéines et donc de distinguer des produits d'allèles différents du même gène. Les techniques qui permettent cette séparation, basées sur l'utilisation d'un champ électrique, sont l'**électrophorèse**, qui sépare les protéines selon leur charge dans un pH donné, et l'**isoelectrofocusing** (focalisation isoélectrique) qui sépare les protéines selon leur point isoélectrique dans un gradient de pH. L'**électrophorèse**, moins coûteuse, est largement utilisée dans la pratique. C'est elle dont nous décrirons la méthode.

*b. La technique d'électrophorèse.* Un équipement classique d'appareillage pour électrophorèse dans un gel horizontal est essentiellement constitué d'une plaque supportant un gel (amidon, agar, polyacrylamide) constitué avec une solution tampon dont les caractéristiques (nature des ions, force ionique et pH) sont déterminantes pour la qualité de séparation des protéines. Les deux extrémités du gel sont reliées respectivement aux pôles + et - d'un générateur de courant continu. Le champ électrique utilisé est en général une ou quelques dizaines de  $V/cm$ .

L'extrait pour l'**électrophorèse** est introduit dans des «puits» moulés dans le gel, ou par l'intermédiaire de petits bouts de papier (ou autre matière absorbante) imbibés d'extrait et insérés dans une fente du gel.

L'appareillage dans son ensemble est placé dans une chambre froide ou/et un refroidissement est assuré au niveau du gel lui-même (plaques de refroidissement, glace fondante, etc...) pour éviter l'activité des amylases (cas du gel d'amidon), la dénaturation des protéines à la chaleur et l'établissement d'un gradient de température dans le gel qui se traduit par des distorsions dans la migration des protéines.

Dans un gel, la vitesse de migration de protéines dans un champ électrique dépend non seulement de leur charge, mais aussi de leur encombrement (taille de la molécule) en fonction de la taille des mailles du gel (variable selon la nature et la concentration en produit gélifiant pour les gels de polyacrylamide). Le temps optimal de migration est donc ajusté empiriquement pour chaque protéine étudiée, en fonction du gel, du tampon et du voltage (limité par l'échauffement) choisis. Il peut varier de quelques dizaines de minutes à quelques heures.

*c. La technique de révélation enzymatique.* Après la phase de migration, les protéines qui migrent à des vitesses différentes vont se trouver concentrées en différents points du gel. Le problème est alors de visualiser spécifiquement certaines de ces protéines.

Les protéines structurales ou de réserves qui sont en concentration élevées peuvent être marquées directement par un colorant des protéines: noir amidon, bleu de coomassie (protéines de réserves du Blé par exemple; AUTRAN et BOURDET, 1973).

Les enzymes, par contre, nécessitent un système spécifique de révélation qui consiste à immerger le gel dans une solution de révélation qui peut être aussi simplement pulvérisée ou gélifiée par de l'agar et étendue au-dessus du gel de migration. La solution contient un substrat pour cet enzyme et un colorant qui va être fixé ou coloré seulement à la place de l'activité de l'enzyme, faisant apparaître une «bande» parallèle à la ligne d'insertion de l'échantillon. Par exemple, la forme oxydée d'une molécule

est incolore, mais peut se colorer lorsqu'elle est réduite par des électrons transférés à cette molécule quand une enzyme déshydrogénase et son cofacteur se séparent.

Bien que certaines soient plus complexes et nécessitent des intermédiaires variés, le principe est le même pour toutes. Plusieurs enzymes différentes peuvent être visualisées sur le même gel si les conditions de révélation ne sont pas réciproquement exclusives (utilisation d'un substrat non spécifique pour la révélation des **estérases**, des peroxydases ou des **amino-peptidases** par exemple).

d. *Amélioration de la qualité de résolution des électrophorégrammes.* Un **électrophorégramme** peut présenter de nombreuses bandes, en particulier dans les cas suivants:

- révélation non spécifique (protéines de réserves...)
- famille enzymatique révélée avec un substrat non spécifique,
- enzyme présente dans différents compartiments cellulaires et codée par plusieurs systèmes génétiques indépendants,
- enzyme codée par des gènes dupliqués ou plusieurs allèles, avec une association au hasard des produits génétiques.

Le problème se pose alors d'obtenir la meilleure dispersion des bandes sans en perdre aucune. Il faut, par des essais répétés, étudier les combinaisons optimales des facteurs qui interviennent, en limitant au maximum les artefacts in vitro. Les paramètres principaux sont:

- la nature de l'extrait analysé et sa conservation: organe échantillonné et son stade, conditions du broyage et de la conservation éventuelle, etc...
- la nature et la concentration du produit gélifiant,
- la nature du support de l'extrait pour l'insertion dans le gel,
- la composition des tampons de migration (du gel et des bacs: force ionique, pH),
- la vitesse et la durée de la migration,
- la méthode de révélation.

Tous ces facteurs n'étant pas indépendants, de nombreuses combinaisons devront être essayées.

Lorsque l'on compare des **isozymes** codés par des allèles du même gène (**allozymes**), les distances de migration sont relativement voisines. Plusieurs cas se rencontrent mais, très souvent, les **allozymes** sont représentés sur le gel par des bandes régulièrement espacées, ce qui suggère qu'elles diffèrent par une ou des «unités» de charge. Généralement, la bande standard migre à une distance moyenne et les bandes extrêmes sont les plus rarement observées. Trois à cinq bandes sont ainsi le plus souvent distinguées, représentant trois à cinq «états de charge»; MARSHALL et BROWN (1975).

Selon les enzymes, les états de charge sont séparés sur le gel par des distances variables. Dans les meilleurs des cas (grandes différences de migration — 1 cm ou plus —), on peut noter des différences plus subtiles de vitesses de migration (1mm par exemple, dans le cas des **PGI** du riz). Ces variations de vitesse de migration sont interprétées comme des différences de pK (constante de dissociation ou de conformation).

Les molécules ont le même état de charge «potentiel» mais des différences dans la nature, ou la position, des acides aminés mutés entraînent des différences minimales de charge (en fonction du pH) ou une conforma-

tion différente. La séparation d'un nombre plus important d'allèles au même locus peut être obtenue en effectuant des migrations des mêmes extraits, à des pH variés, ou en modifiant la concentration des produits gélifiants.

En définitive, la technique d'électrophorèse est simple. Son pouvoir de résolution dépend cependant de nombreux facteurs qu'il n'est pas toujours facile d'identifier. Les mises au point, pour une application particulière, peuvent se révéler très longues.

e. *Interprétation génétique des zymogrammes.* Dans une première phase de dégrossissage, avant que les études génétiques aient été accomplies, on peut utiliser la lecture directe des zymogrammes pour établir des indices de ressemblance qui souffriront bien entendu des effets de redondance non explorés.

Nous entendons par zymogramme (Z.), la combinaison de bandes observées sur une plaque d'électrophorèse, pour un individu, avec un système de révélation spécifique d'une enzyme ou d'un groupe d'enzymes non spécifiques.

Les bandes de Z. sont la visualisation d'isozymes dont les origines multiples se classent en trois grandes catégories (HARRIS and HOPKINSON, 1976) :

- plusieurs locus de gènes codant pour des chaînes polypeptidiques structurellement différentes de l'enzyme,
- allèles multiples à un locus déterminant des versions structurellement distinctes d'une chaîne polypeptidique particulière,
- modifications ultérieures à la transcription de la structure enzymatique: isozymes secondaires.

Beaucoup de protéines enzymatiques sont multimériques et peuvent combiner des produits primaires de plusieurs gènes (isozyme hétéromère) ou du même gène (isozyme homomère) formant des figures de Z. bien connues.

Des gènes distincts peuvent donner des protéines enzymatiques indistinguables par électrophorèse ou non distinguées avec une technique donnée. KING et OHTA (1975) appellent « électromorphe » une classe d'allèles caractérisée par un phénotype commun en électrophorèse.

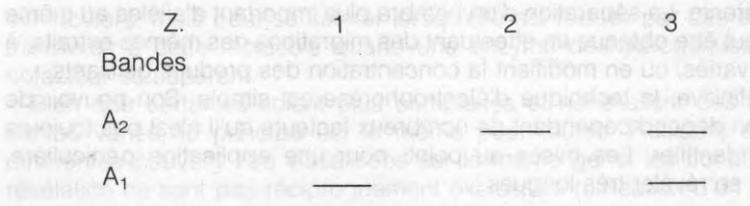
L'interprétation génétique des Z. est basée sur deux approches successives:

- par comparaison de tous les Z. observés, on peut présumer un déterminisme génétique en retenant l'hypothèse explicative la plus simple,
- les hypothèses retenues doivent être confrontées à l'analyse mendélienne de croisements entre des individus présentant des Z. différents ( $F_1$ ,  $F_2$ , B.C.,  $F_3$ ...).

Présentons maintenant quelques exemples d'interprétations basées sur l'observation directe des Z. et leur répartition entre populations ou espèces très apparentées:

- Les Z. observés sont à une ou deux bandes. Trois combinaisons possibles sont observées:

*Hypothèse 1:* les bandes  $A_1$  et  $A_2$  sont codées par des locus différents. Chaque locus présente un « électromorphe » nul. Dans ce cas, le Z.3 (cf. schéma 10) peut être fréquent, même dans une population autogame. Dans une population panmictique, la fréquence de chaque allèle nul est la racine carrée du complément à 1 de la fréquence des Z. à une bande.



**Schéma 10:** Exemple de zymogramme à 2 bandes

*Hypothèse II:* les bandes A<sub>1</sub> et A<sub>2</sub> sont codées par des électromorphes, A<sub>1</sub> et A<sub>2</sub>, du même locus. Le Z.3 représente un hétérozygote A<sub>1</sub>A<sub>2</sub> alors que les deux autres Z. représentent des homozygotes A<sub>1</sub>A<sub>1</sub> et A<sub>2</sub>A<sub>2</sub> ou des hétérozygotes avec un allèle nul récessif. Dans ce cas, le Z.3 doit être rare dans une population autogame. Dans une population panmictique, les fréquences d'électromorphes peuvent être dénombrées comme suit:

$$f A_1 = \frac{2 \times Z1 + Z3}{2n}$$

$$f A_2 = \frac{2 \times Z1 + Z3}{2n}$$

$$f A_2 = \frac{2 \times Z2 + Z3}{2n}$$

$$f A_1 = \frac{2 \times Z2 + Z3}{2n}$$

avec n = nombre d'individus

Z1, Z2, Z3 = nombre de Z.1, 2 et 3 observés.

On peut alors calculer les fréquences de Z. attendues, selon les fréquences de leurs génotypes présumés dans le cadre de la panmixie:

$$f Z.1 = (f A_2)^2$$

$$f Z.2 = (f A_1)^2$$

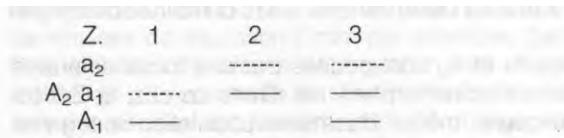
$$f Z.3 = 2 (f A_1 \times f A_2)$$

Les fréquences attendues sont comparées aux fréquences observées par un test du  $\chi^2$  (on regroupe les classes d'effectif inférieur à 5).

Si l'hypothèse génétique est exacte, le test doit être non significatif, mais un léger déficit en hétérozygotes peut s'expliquer par un pourcentage d'autogamie (rencontré chez certaines plantes allogames facultatives) ou par la présence d'un allèle nul récessif à l'état hétérozygote.

Lorsqu'un hétérozygote est représenté par un Z. à 2 bandes, on peut présumer que l'enzyme est un monomère (absence de bandes hybrides). Dans ce cas, on ne s'attend pas à trouver des Z. à plus de deux bandes chez les diploïdes.

— Un cas similaire peut être observé avec des bandes dédoublées (schéma 11):



**Schéma 11:** Zymogramme à bandes dédoublées.

Le même raisonnement peut être appliqué en considérant les bandes  $A_1$  et  $a_1$ , codées par le même électromorphé:  $a_1$  et  $a_2$  seraient des isozymes secondaires. Le chevauchement des bandes  $A_2$  et  $a_1$  (cas fréquemment rencontré) conduit, particulièrement dans le cas de l'hypothèse à deux allèles du même gène, à un Z.3 difficilement distinguable du Z.1.

— Dans le cas d'enzymes di ou polymères, des figures à 3, 4..., 7 bandes..., le même raisonnement s'applique. La présence de bandes hybrides indique que les enzymes sont fonctionnels tant à l'état homomère qu'hétéromère. Les produits homomère sont donc très semblables. Lorsqu'ils sont produits à des locus différents, on parle d'hétérozygotie fixée.

Dans le cas de l'hypothèse à deux locus, quatre produits géniques peuvent être présents chez un individu diploïde. Dans le cas d'un enzyme dimère, dix combinaisons différentes sont possibles. Le chevauchement de certaines bandes conduit cependant généralement à des Z. plus simples: 7 bandes dans le cas des P.G.I. du riz, par exemple. Dans le cas d'une enzyme hexamère (G.D.H. par exemple), avec trois produits géniques différents, 28 combinaisons sont possibles; toutes ne seront pas généralement séparées (pour un exemple de Z. de G.D.H. à 28 bandes, voir cependant l'exemple du pollen de la luzerne DE VIENNE, thèse).

Les cas signalés, et d'autres, ont été rencontrés chez le riz (SECOND et TROUSLOT, 1980). Sur le même Z., plusieurs systèmes génétiques indépendants peuvent se trouver rassemblés. L'élaboration d'hypothèses génétiques simples par comparaison des Z. est alors difficile. L'observation, dans le même complexe spécifique, de populations autogames et allogames (cas des espèces sauvages de riz) peut se révéler grandement utile.

Des espèces diploïdes présentent fréquemment une hétérozygotie fixée à plusieurs locus. En comptant une éventuelle hétérozygotie allélique, on imagine que les Z. obtenus peuvent être complexes. Le tableau 24 résume plusieurs situations d'hérédité avec leurs tests directs. Signalons que le nombre d'analyses nécessaires pour séparer sur une génération  $F_2$  deux hypothèses, 9331 et 8440, peut être beaucoup trop grand et peu satisfaisant d'autant plus que de légères distorsions de ségrégation pourront perturber la lecture. L'étude au moyen d'une deuxième génération d'autofécondation permettra avec un faible nombre, de lever les ambiguïtés: les formes à 1 seule bande (en fréquence 3/16 ou 1/2 suivant les hypothèses), pourront donner des descendance en ségrégations dans le premier cas, jamais dans le second.

Exemple d'analyse mendélienne.

— Exemple du déterminisme des bandes AMC des phosphatases acides (PAC) du riz (PAI, ENDO et OKA... 1975).

Les Z. de PAC du riz sont complexes. Par la comparaison des Z. observés dans différentes espèces, on conclut à l'existence de cinq groupes de bandes contrôlées par quatre à cinq locus au moins (SECOND et TROUSLOT, 1980).

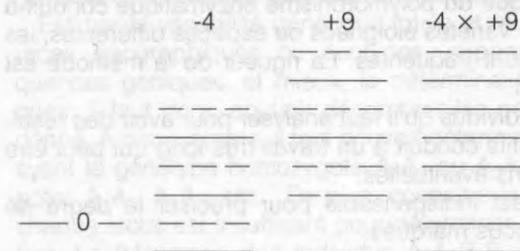
Le groupe de bandes le plus polymorphe est représenté par six bandes (trois faibles et trois intenses) sur la plupart des Z. (elles sont très rarement totalement observées).

Si l'on considère la bande la plus proche de la cathode du groupe, on note huit positions au moins, selon les Z. représentatifs d'individus prélevés



dans différentes espèces du complexe « *Sativa* ». Elles sont symbolisées par la distance standard de migration de -17 à +24 mm.

Un hybride entre deux individus présentant les bandes -4 à +9 a été étudié (schéma 12). L'hybride  $F_1$  montre huit bandes dans le Z.. L'analyse par densitométrie du Z. permet de conclure que des bandes hybrides sont produites entre chaque bande parentale homologue et se confondent avec ces dernières, conduisant à sept bandes intenses et une bande faible.



**Schéma 12:** Zymogramme des phosphatases acides du riz.

Un hybride  $F_1$  entre les mobilités + 4 et + 9 ou - 4 et + 4 montre des Z. à dix bandes (9 bandes intenses et une faible) qui sont expliquées de la même manière.

Les plantes  $F_2$  ségrègent entre les types parentaux et les hybrides, sans exception, généralement selon le rapport 1: 2: 1. La descendance de neuf hybrides différents a été étudié (30 à 193 plantes  $F_2$  selon les cas).

Dans trois cas sur neuf, une déviation significative avec le rapport: 1:2:1, est apparue. Elle peut être raisonnablement expliquée par la sélection gamétique fréquemment rencontrée dans le type d'hybrides étudiés.

On conclut que chaque classe de mobilité est contrôlée par un type allélique (électromorphe) du même locus.

L'hybride  $F_1$ , entre des individus avec et sans bandes, présente une bande avec une intensité réduite à environ la moitié de l'intensité de la bande parentale. La génération  $F_2$  ségrège selon le ratio 3:1 pour la présence/absence: on peut conclure à l'existence d'un électromorphe nul récessif déterminant l'absence de bandes.

— Exemple de l'analyse du déterminisme d'une bande de peroxydase du riz (SHAHI et al., 1969).

Les Z. de peroxydases sont également complexes, avec plusieurs groupes de bandes cathodiques et anodiques. Nous prendrons l'exemple de la bande «4C», cathodique, qui est présente ou absente selon les lignées.

Cinq croisements ont été étudiés entre des plantes avec ou sans la bande. Toutes les  $F_1$  présentent la bande. Au niveau des générations  $F_2$  ou back-cross, les rapports 3: 1 et 1: 1 respectivement, sont observés généralement. Dans un cas sur cinq, il y avait une différence significative avec la proportion attendue. Elle était associée avec une forte stérilité pollinique dans le croisement, et peut être expliquée par la sélection gamétique. La bande «4C» semble donc déterminée par un électromorphe dominant à un locus qui présente un électromorphe nul récessif.

Les auteurs sont cependant allés plus loin en remarquant que, dans une population d'*O. perrenis*, certaines plantes présentaient une bande très

faible. Ils ont donc effectué un croisement **dialèle** entre dix plantes de la même population dont six avaient une bande absente ou très faible. Tous les descendants d'une des plantes (en autofécondation et en croisements) étaient sans bande « 4C ». L'ensemble des résultats s'expliquait par l'intervention d'un autre gène.

## Remarques

— Lorsque l'analyse génétique du polymorphisme enzymatique conduit à l'étude de croisements entre variétés éloignées ou espèces différentes, les **distortions** de ségrégation sont fréquentes. La rigueur de la méthode est alors souvent mise en défaut.

— Le nombre important d'individus qu'il faut analyser pour avoir des résultats statistiquement significatifs conduit à un travail très long qui peut être compromis par des **distortions** éventuelles.

— L'analyse mendélienne est indispensable pour préciser le degré de liaison génétique entre les locus marqués.

f. *Intérêt et critiques de la méthode.* La variabilité **électrophorétique** d'une protéine enzymatique est donc un phénotype pour lequel on peut voir apparaître des différences discrètes, non ambiguës entre génotypes. La méthode est capable de distinguer différents homozygotes, non seulement les uns des autres, mais aussi des hétérozygotes. Comment se situe-t-elle par rapport aux autres méthodes d'évaluation?

En opposition avec l'analyse mendélienne qui repose sur l'existence de caractères simples, déterminés par des différences discrètes au niveau d'un ou quelques gènes, l'observation d'ensemble de populations montre la variabilité quasi-continue du phénotype des individus qui les forment. L'évolution des espèces repose sur l'accumulation progressive de très petits changements dans la morphologie, la physiologie et le comportement. Des espèces apparentées peuvent différer dans leur tolérance ou leur préférence pour des conditions écologiques diverses, dans leur dimension ou leur forme. Il est cependant souvent impossible de distinguer un individu d'une espèce ou d'une autre et, a fortiori, différentes populations d'une même espèce, sur une combinaison de ces caractères, à cause du large recouvrement des phénotypes entre les groupes.

La variation dans ces phénotypes a, de plus, à la fois des composantes environnementales et des composantes **génotypiques**. D'autre part, une simple mutation affectant la taille d'une plante, par exemple, peut avoir des répercussions importantes sur l'aspect de cette plante, sans que le génotype soit profondément modifié dans son ensemble, alors qu'une sélection convergente (adaptative à un milieu naturel ou résultant de la domestication) peut faire ressembler de très près des individus qui ont une base **génotypique** différente.

Si l'on considère les méthodes classiques d'analyse de variabilité des populations, la taxonomie numérique (cf. chapitre V Bases de données) est une méthode puissante de description, basée sur des combinaisons de nombreux caractères qualitatifs ou quantitatifs, qui permet de déterminer des axes principaux de variabilité. Les individus sont ensuite situés sur des plans définis par deux de ces axes. Par elle-même elle ne peut nous renseigner sur la part génétique de cette variabilité et sur son **détermi-**

nisme. Elle permet, grâce aux moyens de calcul, de représenter la variabilité de nombreux individus, **lue** sur de nombreux caractères, et ce n'est qu'en intégrant des caractéristiques génétiquement bien déterminées que les structures révélées prennent une signification évolutive.

L'analyse hiérarchique par famille de descendants permet, par contre, de quantifier la part génétique de la variabilité et de **l'hétérozygotie** globale des individus, mais elle ne nous renseigne pas précisément sur la part du génome concerné par les variations héréditaires.

Estimer la variabilité génétique totale et relative d'espèces, populations, races géographiques ou écotypes, suppose la détermination de fréquences géniques, et mieux, la détermination des fréquences **génotypiques**. Il faut donc pouvoir dénombrer les génotypes. Pour chaque locus donné, A par exemple, il faut pouvoir déterminer la proportion des individus ayant le génotype homozygote  $A_1A_1$ , ou  $A_2A_2$ , et les génotypes hétérozygotes  $A_1A_2$ ,  $A_1A_3$ , etc... De plus, connaître la fréquence **génotypique** pour chaque locus est insuffisant pour déterminer la fréquence de leur association. La fréquence des individus  $A_1A_1B_1B_3$  peut ne pas être le produit simple des fréquences des génotypes  $A_1A_1$  et  $B_1B_3$  prises indépendamment.

Selon LEWONTIN (1974), une technique qui dénombre les génotypes dans des populations doit satisfaire les critères suivants :

- les différences phénotypiques, causées par la substitution d'un atèle par un autre à un locus unique, doivent être détectables comme des différences non ambiguës entre deux individus,
- les substitutions **alléliques** à un locus doivent être distinguables, dans leurs effets, des substitutions **alléliques** à d'autres locus,
- une grande proportion, où toutes les substitutions **alléliques** à un locus doivent être discernables sans confusion l'une de l'autre, indépendamment de l'intensité ou de la nature de leurs répercussions physiologiques,
- les locus observés doivent être un échantillonnage au hasard des gènes, indépendamment de leurs fonctions physiologiques et de la variabilité qu'ils présentent.

Les deux premiers critères impliquent qu'il existe une correspondance univoque entre phénotype et génotype, de telle façon que l'analyse mendélienne ordinaire puisse être effectuée. Il est d'autre part important que les substitutions géniques soient détectables, même en situation hétérozygote, c'est-à-dire qu'il y ait dominance incomplète. Si, par exemple, un allèle dominant est présent avec une fréquence de 0,8 dans une population, et dix allèles récessifs différents avec une fréquence de 0,02 chacun, le phénotype dominant représenterait 96% de la population et aucun homozygote récessif ne serait à la fréquence de 0,1 %. Avec un échantillonnage raisonnable, la variabilité de ce locus serait fortement sous-estimée et, avec un échantillonnage de 30 individus, par exemple, tous seraient probablement du phénotype dominant.

L'analyse de la variabilité, basée sur des caractères visibles, à déterminisme simple (pigmentations, certains cas de forme des grains ou caractéristiques de l'amidon par exemple) répond aux exigences des deux premiers points, bien que, souvent, le phénotype de l'hétérozygote ne soit pas distinguable de celui de l'homozygote dominant. Cependant, les points suivants ne sont pas du tout respectés: il est clair que différentes substitutions **alléliques** auront le même phénotype et que l'échantillonnage des

gènes concernés est loin d'un échantillonnage au hasard. Nous savons par exemple que la domestication a accumulé de telles mutations qui donnent un aspect (rôle esthétique ou permettant leur distinction), un goût ou une qualité particulière de la partie utile de la plante, différents aux variétés, alors que, parallèlement, la base génétique, c'est-à-dire la variabilité génétique globale, a été fortement diminuée par rapport aux ancêtres sauvages. L'analyse de la variabilité au niveau de ces caractères conduirait évidemment à des conclusions allant en sens inverse de la réalité.

L'électrophorèse d'isozymes répond-elle aux exigences de la méthode recherchée?

Le premier point est respecté dans la mesure où l'analyse est effectuée au niveau individuel: le phénotype «mobilité dans un champ électrique et un gel » est généralement non ambigu et ne dépend pas de l'environnement naturel de la plante. Le second point est également généralement respecté, bien que des ambiguïtés puissent se rencontrer dans certains cas: l'absence de bande peut être causée par un allèle « nul » (protéine absente, inactive ou inhibée) ou par un allèle répresséur à un autre locus; certaines enzymes sont codées par plusieurs locus, hétéropolymères, gènes dupliqués, polyploïdisation. Le troisième point est imparfaitement respecté et surtout, on ne connaît pas actuellement quelle proportion de substitutions alléliques est discernable par électrophorèse.

Des estimations, basées sur les différentes substitutions de bases de l'ADN qui conduisent à une modification de la séquence en acides aminés de la protéine (avec une correction fonction de la composition moyenne des protéines en acides aminés) donne environ 1/4 pour la proportion des substitutions qui entraînent une modification de charge de la protéine (MARSHALL et BROWN, 1975). Mais ces calculs sont d'une valeur incertaine. On ne sait pas notamment si toutes les substitutions d'acides aminés ont la même probabilité d'être fixées indépendamment de leur charge, et il est évident que des substitutions en des points différents de la protéine peuvent annuler leurs effets. KING et OTHA (1974) montrent que, si de nombreux acides aminés de la protéine sont libres pour des substitutions de charge, de nombreuses séquences variantes auront la même charge électrophorétique. On ne connaît pas le pouvoir de résolution de la technique entre des molécules différentes ayant la même charge nette. Il dépend de l'expérimentateur (technique employée) et probablement de l'enzyme analysée.

Ainsi, les fréquences d'allèles peuvent être très différentes de la fréquence des classes alléliques distinguables. Dans les grandes populations (qui n'ont pas subi de fortes réductions d'effectifs récemment), ces fréquences dépendent principalement du nombre d'acides aminés libres de muter. Si les électromorphes ne sont pas sélectionnés, le phénotype le plus fréquent sera flanqué de part et d'autre par des phénotypes de moins en moins fréquents

Des données expérimentales mettent en évidence l'hétérogénéité des électromorphes par la sensibilité différentielle à la température des bandes d'électrophorèse homologues.

La technique est simple:

Une série d'échantillons est soumise, pendant un temps fixé (20 mn par exemple), à une gamme de températures (40 à 80° C au plus). Le traitement a lieu avant électrophorèse ou après la migration. La révélation

ultérieure des enzymes montre une chute brutale de l'activité à une certaine température qui peut différer selon les individus présentant le même **zymogramme**. La sensibilité à la température peut s'interpréter comme résultant des liaisons faibles (ioniques ou hydrogènes) qui déterminent la structure tertiaire ou quaternaire de la protéine. Une substitution d'acide aminé peut avoir une influence sur la solidité de ces liaisons, sans modifier la mobilité en électrophorèse.

En conclusion, pour interpréter la variabilité mise en évidence par électrophorèse, il est urgent d'estimer au mieux la proportion des allèles que l'on distingue par cette méthode.

Dans quelle mesure le quatrième point est-il respecté? La constitution du génome, surtout son fonctionnement, sont encore des mystères chez les organismes supérieurs. L'ADN du génome haploïde du tabac représente  $6 \cdot 10^8$  paires de bases **nucléotidiques**. Avec trois paires de bases par codon, et 150 acides aminés par polypeptides, il y a assez d'ADN pour coder environ  $12 \cdot 10^5$  polypeptides. Il est difficile d'admettre que les organismes supérieurs codent 1 million d'enzymes. 10% de ce nombre semble au maximum plausible. Le reste pouvant être, pour une part non fonctionnel (séquences répétées), et pour l'autre part impliqué dans la régulation des synthèses enzymatiques et des protéines structurales. En réduisant notre échantillonnage aux enzymes, nous sommes loin d'un «échantillonnage au hasarda de l'ensemble des gènes. De plus, les enzymes les plus couramment utilisées sont des enzymes solubles, non fixées aux membranes. Encore faut-il, dans la mesure du possible, avoir un échantillonnage le plus varié possible d'enzymes solubles. La plupart des travaux actuels ne s'intéressent qu'à quelques enzymes, dont les **estérases** font souvent partie. Peu de travaux ont été effectués jusqu'ici sur les plantes, et 10 à 20 enzymes au maximum sont étudiées.

g. **Electrophorèse bidimensionnelle**. Une analyse plus complète des polymorphismes de protéine est actuellement en cours de mise au point en génétique de populations de drosophiles et de souris. La méthode proposée par O'FARELL (1975) consiste à séparer les protéines et les révéler sur des plaques de gel de **polyacrylamide**, dans une dimension en fonction de leur point isoélectrique (en utilisant les techniques **d'électrofocusing**) et dans une seconde dimension en fonction de leur poids moléculaire en électrophorèse utilisant le **dodecyl** sulfate de sodium. La détection des protéines peut en théorie être fine (une protéine qui **consistue**  $10^{-4}$  à  $10^{-5}$  % des protéines totales peut être détectée et qualifiée par autoradiographie). Ce système est très efficace pour distinguer des protéines qui ne diffèrent que par un seul changement de charge, mais l'identification précise des fonctions des protéines révélées ne peut être établie directement.

Les premiers résultats obtenus sur les animaux conduisent à réévaluer les résultats acquis par les outils de la génétique des populations décrits dans le chapitre I.

Du point de vue du polymorphisme et de **l'hétérozygotie** observée dans les populations naturelles LEIGH-BROWN et LANGLEY (1979) constatent que par cette technique, l'analyse de 54 locus ne conduit qu'à une estimation du taux **d'hétérozygotie** par locus de seulement 4% avec 6 locus polymorphes (donc beaucoup plus bas que ce qu'on estimait à partir des électrophorèses unidimensionnelles d'enzymes). Cette faible variabilité

correspond à celle observée pour les gènes codant pour des enzymes utilisant un stock étroit de substrats intracellulaires.

C.F. AQUADRO et J.C. AVISE (1981) se sont intéressés aux évaluations de distances génétiques entre espèces de souris. Grâce à l'analyse bidimensionnelle ils ont pu analyser 189 protéines pour 6 espèces. L'ampleur des divergences estimée par l'analyse bidimensionnelle était environ moitié moindre de celle acquise par les analyses par électrophorèse à une seule dimension. Cependant le classement des distances obtenu par les deux techniques était identique.

### 3. Fragmentation des ADN cytoplasmiques par les enzymes de restriction

a. *Le principe.* On connaît des enzymes capables d'hydrolyser l'ADN en des sites particuliers de la molécule. Ce sont les «enzymes de restriction». Le tableau 25 présente la séquence des sites de coupure pour trois enzymes de restriction.

TABLEAU 25: Site de coupure de 3 enzymes de restriction.

Enzyme	Site de coupure spécifique
Eco RI (E. Coli R 13)	— G — A — A — T — T — C — — C — T — T — A — A — G —
Bam I (Bacillus amylolique — faciens)	— G — G — A — T — C — C — — C — C — T — A — G — G —
Sal. I (Streptomyces albus)	— G — T — C — G — A — C — — C — A — G — C — T — G —

En 1976, il a été montré (VEDEL et al., 1976) que cette propriété pouvait être utilisée pour distinguer les ADNcp (chloroplastiques) et les ADNmt (mitochondriaux) de végétaux supérieurs, appartenant, soit à des genres différents, soit à un même genre. Pour cela, les ADNcp et les ADNmt sont isolés sous la forme de longues molécules circulaires (état natif) et traités par une enzyme de restriction. Les fragments de restriction sont ensuite dispersés par électrophorèse sur gel d'agarose. Les électrophorégrammes se révèlent spécifiques des espèces végétales d'où proviennent les ADN chloroplastiques et mitochondriaux.

b. *La méthode (VEDEL et QUETIER, 1978).* Elle est complexe et réservée à un laboratoire de biochimie bien équipé. Les chloroplastes sont isolés à partir de feuilles et les mitochondries, à partir, soit d'organes étolés, soit de cals ou de suspensions cellulaires cultivées à l'obscurité. Toutes les opérations sont effectuées à 4° C. Après broyage dans un tampon, la centrifugation permet d'obtenir un culot enrichi en particules cytoplasmiques. L'action d'une DNase pancréatique élimine l'ADN nucléaire. Les chloroplastes sont alors récupérés dans le culot d'une nouvelle

centrifugation. Les mitochondries sont purifiées par centrifugation dans un gradient de saccharose. Après lyse des particules, l'ADN circulaire est récupéré comme une bande fluorescente en U.V. après une centrifugation dans un gradient de bromure d'éthydiium et de chlorure de cesium. Après purification, on laisse agir les enzymes de restriction. La séparation des fragments de restriction est effectuée par électrophorèse sur gel d'agarose en plaque. Après migration, les fragments sont colorés par le bromure d'éthydiium et l'électrophorégramme est photographié en lumière ultraviolette.

c. *Intérêt et critique de la méthode.* La méthode permet de mettre en évidence les différences entre genres ou entre espèces au niveau des ADN cytoplasmiques. Elle permet de différencier les ADN chloroplastiques des ADN mitochondriaux. Elle a été utilisée par ses auteurs pour:

- analyser la phylogénie du blé,
- étudier l'origine de la stérilité mâle-cytoplasmique chez le blé,
- rechercher l'origine du cytoplasme d'hybrides parasexuels de tabac (originaires de fusion de protoplastes),
- elle est aussi utilisée par F. BERTHOU de l'ORSTOM pour analyser la phylogénèse du caféier (cf. Tome I).

On peut alors penser qu'elle permettra de préciser la carte et la fonction des ADN cytoplasmiques. Elle est précieuse pour préciser l'existence de «populations» d'ADN mitochondriaux qui pourraient être en relation avec des phénomènes d'adaptation et de vigueur hybride chez les végétaux. Enfin, le haut pouvoir de résolution de cette technique permet d'envisager l'étude de la variabilité intraspécifique des ADN cytoplasmiques. Elle reste néanmoins d'utilisation délicate et réservée à l'élucidation de questions bien particulières.

d. *Estimation de la divergence génétique et de la variabilité génétique par les endonucléases de restriction.* ENGELS (1981) montre comment on peut comparer les degrés d'homologie de segment d'ADN d'une région donnée. La proportion de segments non complètement équivalents permet de définir le taux de polymorphisme pour la position d'un site de restriction. Ces méthodes sont cependant perturbées dans la lecture des degrés de divergence (distances génétiques) par le fait que nombre de changements de position de sites ne sont pas seulement dus à des substitutions de bases mais aussi à de petites insertions ou délétions.

## 4. Hybridation de l'ADN

a. *Le principe.* En séparant les deux brins des molécules d'ADN d'un organisme et en les recombinant avec des brins d'ADN d'un autre organisme, un hétéroduplex ou ADN hybride est produit. La quantité d'ADN hybrides peut être facilement mesurée quand l'un des deux brins est marqué radioactivement.

Il est généralement reconnu que la stabilité thermique d'un hétéroduplex ADN/ADN ou ADN/ARN est fonction de la complémentarité des séquences de bases nucléotidiques des deux brins. La température à laquelle 50% de

l'hybride est dissociée est appelée la «température de dissociation thermique moyenne» ( $T_m$ ). La différence des  $T_m$  des hybrides homologues et hétérologues est une mesure de la divergence des séquences de ces ADN. Une différence de  $T_m$  de  $1^\circ\text{C}$  environ est mesurable et correspond environ à 1,5% de substitutions de bases entre deux ADN (BENDICH et Mc CARTHY, 1970). Pour comparer la proximité génétique de différents organismes (différentes plantes représentant des espèces du même genre ou de genres voisins, par exemple), on étudie la stabilité thermique et la proportion de formation des duplex ADN/ADN pour toutes les combinaisons possibles des organismes étudiés, pris deux à deux. La vitesse d'hybridation est fonction du nombre de copies de la même séquence qui sont confrontées. Ces comparaisons mesurent tout d'abord la similitude des séquences répétées d'ADN (ADN redondant qui est en plus grande proportion chez les végétaux supérieurs que chez les animaux).

*b. La méthode.* La méthode est sophistiquée et nécessite l'utilisation de techniques biochimiques spécialisées. Nous la résumerons pour en indiquer les étapes principales (FLAVELL et al., 1978, 1979).

Des plantules sont utilisées (10 à 15 cm) pour l'obtention d'ADN marqué; la culture est effectuée en conditions stériles, en présence de thymidine  $\text{H}^3$  (on vérifiera l'absence d'ADN bactérien étranger dans la préparation). Le matériel végétal est congelé avant la préparation.

Les molécules d'ADN sont purifiées puis fractionnées par sonication (quelques centaines de nucléotides par fragment). Pour sa dénaturation, l'ADN est chauffé jusqu'à  $100^\circ\text{C}$ , pendant 15 mn, puis rapidement refroidi et incubé à  $60^\circ\text{C}$ . Par élution différentielle sur une colonne d'hydroxylapatite, on détermine le pourcentage d'ADN réassocié. Un lavage avec tampon phosphate 0,12 M entraîne l'ADN monobrin. Un lavage avec le même tampon 0,5 M entraîne l'ADN double brin. Les concentrations d'ADN sont lues par absorption dans l'U.V. On peut aussi supprimer l'ADN monobrin par une enzyme spécifique ( $S_1$  nucléase).

Les courbes de réassociation sont établies en fonction du  $Cot$  ( $Cot =$  concentration initiale d'ADN en moles par litre  $\times$  temps d'incubation en secondes). Pour étudier l'hybridation des séquences d'ADN entre différentes espèces, on mélange de l'ADN marqué d'une espèce avec de l'ADN marqué d'une autre espèce en excès (1/3000 par exemple). Le mélange des solutions est dénaturé par la chaleur et incubé à  $60^\circ\text{C}$  pendant 16 heures. Dans ces conditions, l'ADN marqué se réassocie avec l'ADN de l'autre espèce. Par élution sur une colonne, on élimine l'ADN monobrin puis on élève progressivement la température de la colonne par paliers. A chaque paliers, on entraîne l'ADN dénaturé. On peut ainsi tracer les courbes de dénaturation de chaque hybride, en fonction de la température.

Une amélioration de la méthode consiste à utiliser, dans les expériences d'hybridation, une certaine fraction seulement de l'ADN marqué qui représente, par exemple, les séquences les plus répétées. En utilisant un  $Cot$  de réassociation très faible, seules les séquences hautement répétées sont réassociées. On peut les purifier, les marquer in vitro (par «nick translation»), et les considérer comme une sonde pour observer leur hybridation avec différents ADN.

*c. Intérêt et critique de la méthode.* La vitesse de réassociation des ADN homologues fractionnés indique que la plus grande partie du génome est

constituée de séquences répétées (45 à 90% chez les plantes selon FLAVELL et al., 1974). Le rôle, s'il existe, de la plupart de ces séquences répétées est un mystère mais les mutations semblent être considérablement tolérées dans ces portions d'ADN. Ceci les rend particulièrement utiles pour estimer le degré de divergences génétiques entre organismes. La proportion d'hétéroduplex et leur stabilité thermique donnent toutes deux une estimation du degré de divergence entre les ADN.

La méthode a été utilisée pour l'étude de différents groupes d'organismes, tant animaux que végétaux. En ce qui concerne les plantes cultivées, des travaux ont été publiés depuis 1970 pour la comparaison des ADN d'orge, d'avoine, de seigle et de blé (BENDICH et Mc CARTHY, 1970) (FLAVELL et al., travaux cités et antérieurs). En général, le degré d'homologie des séquences répétées entre les espèces est relié à leur similitude phylogénétique déduite des analyses taxonomiques classiques. En d'autres termes, les espèces les plus reliées ont de plus nombreuses familles de séquences répétées en commun.

Le pouvoir de résolution de la technique est cependant suffisant pour distinguer les espèces de genres voisins, voire les espèces du même genre. Il ne semble cependant pas possible de descendre au niveau de la comparaison entre individus. La complication de la méthode ne permet d'ailleurs pas l'étude de nombreux organismes ou la comparaison de populations. Elle doit être réservée à l'étude de problèmes spécifiques de phylogénie ou utilisée comme marqueur dans les manipulations génétiques interspécifiques.



# CHAPITRE IV LA CONSERVATION DES RESSOURCES GÉNÉTIQUES

A. CHARRIER, M. LOURD et J. PERNES



# I. INTRODUCTION

Les recherches sur les ressources génétiques se rattachent à quatre opérations unitaires de base qui sont:

- la collecte du matériel végétal,
- son évaluation génétique et agronomique,
- son stockage et sa conservation,
- l'enregistrement des informations dans une base de données.

Après avoir analysé la collecte et l'évaluation dans les chapitres précédents, nous allons maintenant examiner comment conserver le matériel végétal intéressant pour les travaux présents et futurs d'amélioration des plantes. Ces opérations de recherches ne sont pas indépendantes. Il est bien évident qu'après une prospection, il sera souvent nécessaire de procéder au stockage des graines collectées ou à leur multiplication et à leur introduction dans des collections de plantes vivantes avant d'envisager leur évaluation. Inversement toutes les études se rapportant à la variabilité des populations naturelles ou des cultivars traditionnels et à l'organisation évolutive du complexe d'espèces permettront de déterminer les stratégies appropriées à la conservation des ressources génétiques considérées.

En particulier, depuis une quinzaine d'années, des résultats expérimentaux importants ont été acquis pour certains groupes de plantes cultivées (mil, sorgho, riz, maïs, blé, haricot, pomme de terre, tomate, *capsicum*...). Il est maintenant clair que le matériel végétal intéressant du point de vue des ressources génétiques est constitué non seulement des cultivars traditionnels mais aussi des espèces sauvages et des formes adventices liées aux espèces cultivées. La sauvegarde du potentiel génétique concerne toutes ces sources de matériel végétal. En conséquence, elle peut être assurée soit par la conservation in situ des formes sauvages dans leurs écosystèmes et des formes cultivées dans les zones d'agriculture traditionnelle, soit par la conservation des deux types dans des centres de recherche spécialisés.

Les problèmes posés par la conservation des ressources génétiques sont d'une grande variété et nombre d'entre eux sont loin d'être résolus. D'abord au plan des objectifs à atteindre: doit-on s'en tenir au maintien de l'intégrité génétique ou recréer au contraire les conditions permettant de reproduire les mécanismes évolutifs naturels ou les voies de la *domestication*? Ensuite au plan des stratégies possibles: optera-t-on pour une conservation centralisée et figée dans une grande «banque de gènes» contenant le maximum de diversité génétique du complexe étudié ou pour une conservation dynamique dans plusieurs sites géographiques localisés? Les études expérimentales en ce domaine sont très insuffisantes, certaines d'entre elles devant être conduites à long terme. Enfin au plan des techniques adoptées: choisira-t-on de conserver des plantes entières ou miniaturisées, des organes ou des cellules en culture in vitro, des graines...? Des progrès techniques déterminants ont été réalisés ces dernières années, mais on manque aussi d'informations précises sur leurs applications à la conservation des ressources génétiques. Il faut aussi s'attendre à des découvertes en matière *d'ingénierie* génétique qui pourraient remettre en cause les choix actuels. Nous ne pouvons donc guère

présenter des informations exhaustives sur tous les sujets abordés dans ce chapitre mais plutôt susciter des réflexions et des recherches méthodologiques sur la conservation des ressources génétiques.

Quelle que soit la plante étudiée, les chercheurs chargés de sa conservation auront à choisir parmi les stratégies et les techniques dont nous allons maintenant faire la présentation et discuter l'intérêt, les avantages et les limites.

## II. LES DIFFÉRENTES STRATÉGIES DE CONSERVATION

La structure génétique du matériel végétal protégé se modifie dans le temps et dans l'espace selon les méthodes de conservation utilisées. Son évolution peut être en partie prédite sur la base des données théoriques et expérimentales de la génétique des populations rappelées au chapitre I.

Les dérives génétiques résultant de la conservation doivent être compatibles avec l'objectif de mettre en permanence à la disposition de l'améliorateur de plantes des ressources génétiques utiles à la réalisation de ses objectifs de sélection.

### A. LA MISE EN RÉSERVE DES ÉCOSYSTÈMES

C'est le meilleur moyen de conservation que l'on puisse envisager (FRANKEL, 1970), mais il ne peut être appliqué que de façon limitée.

#### 1. La nature des écosystèmes mis en réserve

On distinguera les écosystèmes naturels peu, ou pas perturbés, des écosystèmes agraires traditionnels.

Les communautés végétales sont en équilibre avec leurs biotopes et conservent leur potentiel évolutif. L'intérêt de leur mise en réserve tient évidemment à leur adaptation écologique et à leur évolution sur une échelle de temps de longue durée, mais aussi à la coexistence d'une grande variété d'animaux, de végétaux, de prédateurs et de parasites évoluant au gré d'interactions fort complexes. Ainsi subsiste-t-il dans le même écosystème naturel, à côté des organismes directement utiles à l'homme, des formes animales et végétales considérées à tort comme «mineures», par manque d'intérêt immédiat ou par méconnaissance.

La conservation des ressources génétiques dans leurs écosystèmes naturels concerne surtout les régions de grande variabilité génétique des plantes cultivées, c'est-à-dire leurs centres d'origine et de diversification. VAVILOV a reconnu dans le Monde huit centres d'origine sur la base de la diversité génétique, de l'endémisme et de la résistance aux maladies. Cette notion est maintenant dépassée; divers types de centres ont été proposés (HARLAN, 1975; JAIN et al., 1975). On reconnaît pour chaque plante

cultivée un ou plusieurs centres de diversité, mais la situation n'est pas toujours très claire (cf. chapitre I). Par exemple pour les mils et les sorghos qui ont une origine diffuse ou multiple très largement dispersée en Afrique du Sénégal à l'Éthiopie, HARLAN parle de «non centre». Dans le cas des riz cultivés, les données expérimentales actuelles permettent de considérer qu'ils résulteraient d'au moins trois domestications indépendantes à partir de formes spontanées africaines et asiatiques d'origine commune; elles ont donné naissance à *Oryza glaberrima* en Afrique, *O. sativa japonica* en Chine et *O. sativa indica* en Inde (SECOND, 1982).

Il existe dans le monde intertropical des écosystèmes ayant en partie conservé leur intégrité comme certains massifs forestiers non exploités abritant des arbres et arbustes sauvages (caféiers et colatiers en Afrique, cacaoyers et hévéas en Amazonie...) ou des savanes herbeuses riches en graminées et légumineuses fourragères (Brésil, Afrique orientale).

Prenons l'exemple du complexe des *Maximae* à reproduction apomictique pour illustrer notre propos (cf. Monographies, chapitre I). Le *Panicum maximum* vit à l'état spontané dans toute la zone tropicale humide d'Afrique. L'étude des populations naturelles de Tanzanie et du Kenya a permis de comprendre l'organisation évolutive du groupe (ERNES, 1972): sa compartimentation résulte de l'apomixie et de la tétraploïdie. Les formes diploïdes sexuées et tétraploïdes apomictiques ne subsistent ensemble que dans cette région de l'aire de répartition où l'on observe une variabilité phénotypique et adaptative maximale. Par contre, les souches de Côte-d'Ivoire exclusivement apomictiques ne présentent que deux types morphologiques liés à des variations écologiques minimales. Seule la conservation des communautés herbeuses à *Panicum* d'Afrique de l'Est serait à retenir en vue de sauvegarder le potentiel évolutif de ce groupe. Un choix plus précis des zones géographiques à protéger est possible d'après les résultats de l'analyse approfondie de la variabilité des populations naturelles prospectées. Par exemple, dans la région de Korogowé en Tanzanie, la présence simultanée de formes sexuées et apomictiques se traduit par un polymorphisme important.

Il existe aussi dans les centres de diversité un équilibre «hôte-parasite» ou « prédateur-proie » en constant réajustement. Cette coévolution peut être très rapide dans certaines conditions: il suffit de quelques dizaines de générations pour passer des fréquences de l'ordre du taux de mutation, à des fréquences proches de l'équilibre. Ces changements de fréquences concernent aussi bien les gènes de résistance aux parasites que les races physiologiques de ces derniers. Par exemple, il existe un équilibre de cette nature entre la rouille orangée due à *Hemileia vastatrix* et son hôte *Coffea arabica*, dans le sud-ouest de l'Éthiopie, alors que cette maladie ravage les plantations d'Amérique établies sur une base génétique trop restreinte.

Dans les zones d'agriculture traditionnelle des pays en voie de développement, les paysans utilisent des cultivars particulièrement bien adaptés au milieu et répondant à certaines préférences. Ces races locales parfaitement identifiées par leurs noms vernaculaires présentent une grande diversité en rapport avec leur domestication et le couplage «formes sauvages — formes cultivées » (HARLAN, 1972). La diversification de ces variétés traditionnelles liée à des adaptations écologiques précises et variées résulte d'une sélection empirique constante en vue de satisfaire les besoins des hommes. Il faut surtout insister sur le fait que les formes spontanées

apparentées aux plantes cultivées sont les sources actives de la variabilité génétique des cultivars traditionnels dans les aires d'origine et de diversification. Le cultivateur, avec ses traditions, tient un rôle capital dans le contrôle de ces transferts géniques qui résulte des pratiques culturales et du choix des semences. Les paysans de ces régions connaissent les formes spontanées apparentées aux céréales cultivées et distinguent les hybrides naturels en les nommant (au Niger, les hybrides sont appelés *Chibras*). Leurs caractéristiques sont connues (morphologie, égrenage, grain) et ils sont parfois consommés. Ils participent surtout à la reproduction (pollinisation), mais n'en constituent jamais la semence. Celle-ci est composée des plus beaux épis du champ, généralement repérés avant la récolte. Chez les plantes *autogames*, cette pratique favorise le choix des plants hybrides issus soit de croisements intra et *intercultivars*, soit de croisements en retour d'hybrides «cultivé x spontané» par le parent cultivé. Bien que les taux de fécondation croisée soient faibles en régime d'autogamie, ils sont souvent renforcés par la présence de plantes mâles stériles dans les variétés traditionnelles (fréquence 1% à 1%). Bien entendu les structures hybrides sont communes chez les plantes *allogames*, mais leur création peut être limitée par la réduction de la fréquence des hybridations «cultivé x sauvage» due à des barrières reproductives de nature diverse (décalage des périodes de floraison, différenciation partielle) et par le «*linkage*» des gènes contrôlant les caractères de domestication.

De ce point de vue, les riz cultivés africains offrent une bonne illustration (cf. Tome I, chap. 3). L'espèce cultivée asiatique *Oryza sativa* d'introduction récente en divers points d'Afrique (Guinée-Bissau, Zanzibar) a donné naissance en 5 à 10 siècles à un grand nombre de cultivars bien adaptés. Depuis 1975, les prospecteurs de l'ORSTOM et de l'IRAT ont rassemblé plus de 1000 échantillons de riz traditionnels dont la diversité concerne essentiellement la panicule et le grain, la durée du cycle en relation avec le climat, les adaptations à différentes pratiques culturales (pluvial, bas-fond, rizière...), la réponse photopériodique, l'utilisation alimentaire... L'espèce *O. sativa* a en partie *surplanté* sur son propre terrain l'espèce *O. glaberrima* domestiquée en Afrique à l'exception de quelques régions particulières comme en Casamance (au Sénégal), les mangroves de Guinée-Bissau, le delta intérieur du Niger au Mali. A cette richesse des cultivars locaux de riz, il faut ajouter les formes adventices (*O. breviligulata*, *O. stapfii*) et parfois sauvages (*O. longistaminata*) qui peuvent subsister en association, dans la même zone de culture, comme au Mali et en Guinée. Cette situation favorise les échanges géniques au sein du complexe d'espèces et les *introgressions* entre formes cultivées et spontanées comme on a pu le montrer.

Rappelons aussi l'exemple bien connu du maïs (*Zea mays*). De façon très schématique, la domestication de cette espèce au Mexique, en Amérique Centrale et Australe (région andine) a donné naissance à une grande richesse de variétés locales. A une échelle plus limitée et dans des conditions comparables à *Oryza sativa*, l'introduction récente du maïs en Afrique a aussi conduit à des cultivars traditionnels bien adaptés. Au Mexique et au Guatemala on notera la présence d'une forme adventice annuelle, le *téosinte* (*E. mexicana*) dont toute l'importance dans l'évolution de *Z. mays* a été reconnue. Cette espèce apparentée se maintient dans les cultures de maïs et à leurs abords, situation très favorable aux brassages géniques entre

deux espèces **allogames** donnant naissance à des hybrides fertiles. On reconnaît dans de nombreuses races de maïs du Mexique des races d'**introgression** de **téosinte** (WILKES, 1967). Ce matériel végétal constitue le fond génétique vital de ce centre de diversité du maïs: on trouve dans les cultures traditionnelles une grande hétérogénéité et une forte **hétérozygotie**.

D'autres exemples de couplages «cultivé-sauvage» accompagnant la diversité des variétés traditionnelles pourraient être cités: le mil au Niger-Mali (BRUNKEN et al., 1977), le blé en Afghanistan (BENNETT, c.p.), le millet en Chine (PERNES, 1981), l'orge au Proche-Orient...

## 2. La protection des écosystèmes et des zones d'agriculture traditionnelle

Leur vulnérabilité face au développement des activités humaines et des technologies modernes est reconnue. Maints exemples sont rapportés par les défenseurs de la nature. Ce phénomène amorcé dans les pays les plus développés atteint maintenant tous les pays, spécialement les régions tropicales d'Asie, d'Afrique et d'Amérique (FRANKEL et STELHE, 1982). Ces régions qui abritent la plupart des richesses biologiques sont la proie des moyens de défrichements puissants. Cette destruction massive des écosystèmes naturels n'est d'ailleurs pas sans danger: désertification, appauvrissement des sols, érosion s'installent en l'absence de mesures d'accompagnement.

Au cours du dernier siècle, le développement d'une agriculture moderne à hauts rendements, en liaison avec l'accroissement de la population mondiale et l'industrialisation, a entraîné le remplacement des cultivars traditionnels par un nombre limité de variétés sélectionnées. Par exemple, pour le riz en Côte-d'Ivoire, où subsistent plusieurs centaines de cultivars locaux, on assiste à leur disparition rapide dans les zones d'agriculture encadrée par la vulgarisation qui préconise leur remplacement par moins de cinq variétés modernes d'*O. sativa* potentiellement plus productives. La disparition de l'espèce africaine cultivée *O. glaberrima* est déjà très avancée. De même, pour le maïs, la zone géographique occupée par la forme adventice *E. mexicana* s'amenuise très rapidement: on estime que l'aire de distribution du **téosinte** au Mexique s'est réduite de moitié depuis le début du siècle avec une accélération au cours des dix dernières années due à l'extension des régions d'élevage et à l'introduction des variétés de maïs-hybride dans un système culturel **agro-industriel**.

La réduction de la base génétique du matériel cultivé se traduit de façon spectaculaire par un risque accru des épidémies: l'**Helminthosporiose** du maïs aux USA en 1970; la Rouille de caféier au Brésil en 1970; la **Pyriculariose** du riz en Afrique.

Nous soumettons à la réflexion du lecteur ces deux citations de HARLAN (1972):

«Les populations variables cultivées dans l'agriculture traditionnelle sont un héritage qui n'a pas de prix pour l'élaboration des programmes agronomiques modernes; il est en train de disparaître sous nos yeux».

«L'agriculture moderne est une entreprise à rendement élevé et à risques élevés.»

La conservation des ressources génétiques naturelles dans des réserves, seul moyen de **préserver** leur évolution dans un environnement changeant doit être favorisée, surtout dans les régions tropicales d'intérêt reconnu. Le modèle de génétique des populations qui correspond à ces réserves est celui d'isolats **biogéographiques**. L'isolement et la sélection sont les principaux facteurs modificateurs des fréquences géniques et de la structure des populations naturelles. Nous manquons singulièrement de données et d'observations sur la dynamique, la taille, la répartition géographique de telles réserves. Leur évolution nécessite des contrôles réguliers et il existe là un vaste champ de recherches conjointes avec les botanistes.

La conservation à long terme des variétés traditionnelles est encore plus difficile à réaliser que celle des écosystèmes naturels. Il faudrait en effet préserver ces régions avec le mode de vie traditionnel des paysans et le maintien de leurs systèmes culturels ancestraux. Quelques propositions concrètes ont été avancées: conservation de parcelles de 1 à 2 ha des variétés de céréales traditionnelles au Proche-Orient (KUCKUCK); mise en défens de bandes de 20 x 5 km dans les zones de culture de maïs au Mexique afin de favoriser **l'introgession** naturelle **maïs-téosinte** (WILKES, 1972). Toutes ces solutions paraissent bien illusoire dans des systèmes d'agriculture en cours de modernisation. Dans les pays en développement, il faut s'attendre à la disparition accélérée de ces variétés de pays, comme cela c'est produit en Europe avant 1950. Afin de ne pas renouveler la même erreur, les organismes de recherches ont adopté des mesures conservatoires. Au début, il s'agissait de collections de plantes vivantes multipliées régulièrement. Maintenant, les chercheurs collectent les variétés traditionnelles et les conservent par stockage de graines en chambre froide avec **réjuvénation** périodique des lots de semences pendant leur pouvoir germinatif. Comme nous le verrons, les collections de variétés maintenues dans ces conditions hors de leur écosystème initial n'en suivent plus les transformations. Ainsi, les allèles des gènes de résistance aux parasites ne seront bientôt plus ceux qui équilibreront les nouvelles races apparues. Il serait certainement préférable d'entretenir les variétés traditionnelles dans leur évolution génétique et écologique synchronisée avec les formes sauvages apparentées résultant de transfert géniques dynamiques et équilibrés (**co-évolution** hôte-parasite).

La conservation à long terme des variétés traditionnelles de pays dans leur écosystème n'étant pas compatible à terme avec une agriculture moderne ne pourrait-elle pas être simulée dans des zones agricoles et des stations agronomiques, sous contrôle scientifique? Il y a dans ce domaine des recherches nouvelles à entreprendre pour mettre au point une méthode de conservation dynamique des ressources génétiques.

## **B. LES COLLECTIONS DE PLANTES VIVANTES**

On distingue classiquement deux types de collections:

- les collections de travail des sélectionneurs,
- les collections pour la conservation des ressources génétiques.

Le premier type évolue en grande partie en fonction des intérêts appliqués du sélectionneur et des schémas de sélection. Elles sont constituées

d'un échantillon très peu représentatif de l'ensemble de la variabilité des espèces étudiées et de leurs connexions évolutives. Ces collections de travail peuvent cependant être plus ou moins riches et diversifiées d'une station à l'autre, bien qu'elles soient souvent développées par échange et introduction de variétés et souches plus ou moins sélectionnées provenant d'autres pays. De toute façon, elles ne permettent ni de considérer dans son ensemble l'organisation génétique d'un complexe d'espèces ni de présenter la sécurité de conservation de la variabilité génétique de ce dernier.

Seules les collections du deuxième type répondent à l'objectif fixé. Elles sont à constituer soit par des introductions d'un autre CRG\* et des collections de travail (choisir les souches originales) soit par des collectes et des prospections axées sur les différentes formes cultivées et les espèces voisines appartenant potentiellement au même complexe d'espèces.

Le matériel végétal récolté sera toujours un échantillon limité\*\*. Les méthodes d'échantillonnage proposées au champ ont pour objectif d'assurer la collecte la plus représentative en rapport avec nos objectifs.

Comment va-t-on conserver ce matériel? Veut-on le fixer définitivement ou cherche-t-on à maintenir son processus évolutif. Autant de questions que l'on doit se poser afin de choisir les types de conservation appropriés en collection. Le stockage à long terme de semences, les collections de plantes à reproduction asexuée placées dans des conditions contrôlées tendent vers une conservation statique du potentiel génétique\*\*\*. Au contraire, les collections d'espèces à reproduction sexuée maintenues de génération en génération, dans un milieu nouveau, vont subir une réorganisation et des dérives sous l'effet de pressions évolutives différentes (effectif restreint, milieu différent, rupture du contact avec les espèces apparentées). Ce type de collection conduit à la destruction des structures génétiques construites au cours des années dans les populations naturelles; c'est-à-dire des états stationnaires acquis, résultant d'équilibre entre: mode de reproduction — pression sélective — organisation des chromosomes.

## 1. Les collections maintenues par multiplication sexuée

Le groupement de populations représentatives des régions de l'aire de distribution du matériel végétal étudié dans une même collection installée en un lieu donné accroît les pressions de sélection naturelle ainsi que les occasions d'hybridation. La structure génétique des plantes conservées en collection dans un nouvel environnement est déterminée par les différences climatiques, édaphiques ou biotiques avec leurs biotopes originaux, par la longueur du cycle et le mode de reproduction, par la compétition, les maladies et les soins techniques donnés à la culture.

---

\*CRG : Centre de Ressources génétiques.

\*\*Les problèmes de maintenance ne peuvent être considérées indépendamment de l'échantillonnage.

\*\*\*On pourrait désigner cette conservation des individus originaux par le terme de «Collection *museum* » de SIMMONDS (1962) quand on parvient à préserver l'intégrité génétique.

Les collections multipliées par voie sexuée peuvent être entretenues soit sous forme d'origines individualisées maintenues en autofécondation ou en isolement, soit sous formes d'un réservoir massai reconduit en fécondation libre.

a. *Collection d'origines individualisées.* Cette méthode donne la possibilité d'étudier individuellement les composants d'un complexe d'espèces, leurs caractéristiques morphologiques, physiologiques, génétiques, cytogénétiques, agronomiques, pathologiques d'intérêt direct pour le sélectionneur et tous les biologistes. L'interaction génotype x milieu est jugée par la répétition des collections de lignées dans différents environnements contrôlés. Ces informations acquises dans le milieu de travail du sélectionneur (relativement différent de l'environnement auquel le matériel est adapté) lui donnent des renseignements précieux bien que les caractéristiques utiles — production, teneur en protéines, résistance aux maladies — soient le plus souvent à déterminisme génétique complexe.

La difficulté du maintien dans les collections d'un grand nombre d'origines sur une surface limitée tient aux changements de la structure génétique induits par la faible taille des échantillons de différentes origines et la fécondation croisée (ALLARD, 1970). Pratiquement, si le nombre d'origines introduites est important, l'isolement de chacune d'elles n'est pas facile à assurer. Avec les plantes autogames, on réalise soit une autofécondation contrôlée, soit une fécondation libre s'en rapprochant. Dans ce cas, il ne faut pas sous-estimer le taux de fécondation croisée résiduelle dont on sait qu'il peut varier dans de larges proportions avec les conditions du milieu, les génotypes et la population d'insectes pollinisateurs. Avec les plantes allogames, l'isolement de chaque origine nécessite des moyens importants en surface et en main-d'œuvre. La fécondation libre c'est-à-dire croisée des individus d'une même variété-population isolée favorise les croisements entre génotypes apparentés. En fait, le mode de reproduction n'influence guère la structure génétique quand il s'agit d'un échantillon de taille réduite: il y a dérive et fixation d'allèles particuliers qu'il s'agisse de plantes allogames ou autogames préférentielles. Le résultat après plusieurs générations de multiplication est une fixation à l'état homozygote d'un grand nombre de gènes. Du fait des gènes délétères, un certain nombre d'origines se maintiennent difficilement. Mais beaucoup de gènes non létaux peuvent être fixés au hasard dans une nouvelle combinaison génétique si le nombre d'origines est suffisamment élevé. Théoriquement, la variabilité génétique est préservée, mais pratiquement on constate une perte au hasard de gènes non létaux.

Par contre, les pressions de sélection du nouveau milieu seront peu importantes au regard des effets de dérive. Pour une espèce autogame conservée en lignées composées de 25 individus, la perte aléatoire de gènes est aussi importante que si l'on imposait une sélection de 1% à chacun des gènes (PERNES, 1975).

Le système de conservation par origines génétiques individualisées entraîne la perte de nombreux allèles quels que soient les efforts déployés pour maintenir la variabilité de la collection dans un milieu peu sélectif. En comparaison, les pertes par sécheresse et maladies sont moins importantes, quoique plus spectaculaires et faciles à apprécier. De plus, le régime d'autofécondation forcée ou de croisements consanguins détruit les

organisations géniques préexistantes. Son seul avantage tient à l'évaluation en facilitant largement le choix de lignées en fonction de leurs performances dans les tests d'aptitude à la combinaison.

En conclusion, les collections vivantes d'origines individualisées s'avèrent une méthode de conservation des ressources génétiques coûteuse en surface et en personnel, aisément applicable aux espèces **autogames** et facilitant le choix du sélectionneur. Le maintien de collections d'espèces à cycle court astreint à des multiplications annuelles ou bisannuelles. Ces difficultés peuvent être largement atténuées par la conservation des graines à long terme (voir ci-après), ce qui restreint le nombre de cycles de multiplication et réduit l'érosion génétique importante due aux multiplications répétées.

b. *Réservoir massai*. Comme nous venons de le voir, la conservation de plusieurs centaines de cultivars ou d'introductions par multiplication individualisée ne peut être pratiquée, surtout pour les espèces **allogames**. Le maintien en autofécondation ou en croisement « Frère x Soeur » des introductions de maïs ou de mil entraîne une perte de vigueur et de gènes, incompatible avec l'objet de la conservation du complexe d'espèce. D'où l'idée de créer une ou plusieurs populations regroupant la variabilité d'un ensemble d'origines différentes **pollinisées** artificiellement ou naturellement. Ce réservoir massai diffère des populations **panmictiques** et se rapporte aux populations composites. Sa valeur initiale dépend de la représentativité des origines par rapport à la variabilité globale du complexe d'espèces.

Dans ce système, la dérive génétique joue un rôle moins important que précédemment et la variabilité n'est pas libérée. Même si certains allèles atteignent des fréquences faibles, ils sont rarement perdus, d'autant que, leur valeur sélective s'accroît souvent dans ce cas. Mais il y aura une réorganisation de cette variabilité suivant les pressions sélectives nouvelles vers une nouvelle cohérence des génomes. L'élimination des gènes désavantageux dans un milieu donné n'est pas inéluctable. On peut réduire le risque de perte d'allèles en plaçant le même réservoir massai dans différents environnements ou en le transférant successivement dans des milieux différents. Pour en juger, on estimera la diversité réelle par l'étude d'un échantillon prélevé dans le réservoir.

La conservation du matériel végétal en réservoir massai s'accompagne d'une adaptation à l'environnement conduisant à un complexe d'espèces **coadaptées**. Il apparaît plus homogène qu'en réalité; le passage par l'autofécondation est nécessaire pour libérer les gènes récessifs.

Beaucoup d'autres facteurs influent la perte de gènes et la recombinaison au sein d'un réservoir massai.. Citons:

- les fréquences géniques,
- les dates de floraison,
- le pourcentage d'autofécondation naturelle,
- le système d'incompatibilité,
- la taille du réservoir,
- la direction du vent,
- les méthodes d'échantillonnage à chaque génération,
- le potentiel individuel de production **grainière**.

BURTON (1976) a étudié l'évolution de la structure génétique de 6 réservoirs isolés de mil (*Pennisetum typhoides*) pendant 3 à 5 générations.

Il constate une réduction de la variabilité phénotypique, la perte de gènes, la dérive du caractère maturité. Pour ce matériel, ces inconvénients pourraient être réduits en initiant un groupement des floraisons par une culture en jours courts à température élevée (proche de 40°), en associant un nombre restreint d'origines, en mélangeant un nombre égal de graines vivantes des individus sélectionnés à chaque génération.

En conclusion, le système de maintien de collections en réservoir massai assure une meilleure intégrité génétique avec des coûts réduits en surface et en **main-d'œuvre** par rapport au système des origines individualisées. On y perd en information et ce système modifie les combinaisons de départ pour donner naissance à de nouvelles structures. Cette réorganisation peut être orientée dans différentes directions en cultivant le même réservoir massai dans des milieux différents (conservation dynamique). Les risques de pertes géniques peuvent être aussi limités dans ce système en le combinant avec la conservation à long terme des graines recueillies sur les premières générations de multiplication du réservoir massai, ce qui réduit le nombre de cycles de multiplications des plantes annuelles surtout.

**ALLARD** (1970) a proposé un système mixte tenant compte des avantages et des inconvénients respectifs de la conservation en « lignée » et en « réservoir massai » : partir d'un grand nombre de lignées, puis, au fur et à mesure de l'obtention des informations, regrouper en réservoir massai les lignées les moins intéressantes ou à diversité restreinte. Par exemple pour 2000 origines d'avoines sauvages, il retient après les premières évaluations 200 origines intéressantes individualisées et regroupe le reste en 30 réservoirs cultivés dans différents milieux. Quand les études sont plus avancées, il garde moins de 100 origines individuelles et 15 populations. Cette conduite de la conservation associe « préservation — évaluation — utilisation ».

## 2. Les collections maintenues par multiplication asexuée

La conservation en collection de souches individualisées reproduites par multiplication végétative ou par graines **apomictiques** permet théoriquement de maintenir la représentation des génotypes introduits initialement. Dans ce type de collections statiques, on atteint le plus haut degré d'intégrité génétique. Comme dans les collections d'origine individualisées maintenues par voie sexuée, il est possible d'acquérir sur chaque souche une information complète qui est directement exploitable par le sélectionneur.

Ce type de collections vivantes concerne des espèces végétales fort diverses appartenant aux arbres forestiers et fruitiers, aux agrumes, aux plantes industrielles comme l'hévéa, les caféiers, le cacaoyer, la canne à sucre, aux plantes florales, aux plantes reproduites par bulbes, tubercules ou rhizomes, etc... De même, les modes de multiplication asexuée concernés sont variés: bouturage, greffage, éclats de souches, bulbes, tubercules, rhizomes, graines **apomictiques**. Vu la diversité des situations rencontrées, nous considérons les problèmes posés par les collections à l'aide de quelques exemples significatifs.

a. *Les collections d'espèces arbustives.* On recourt à la conservation ex situ des populations d'espèces économiquement importantes chaque fois

qu'elles appartiennent à des écosystèmes menacés. La constitution de telles collections d'arbres vivants rencontre des obstacles pratiques et techniques de taille.

Elles sont généralement établies avec des semences provenant d'une prospection ou des semences prélevées dans des collections existantes. Les fréquences géniques des échantillons originaux de semences sont presque intégralement conservées en combinant un ensemble de facteurs externes propres à assurer la survie de la plupart des individus comme le choix des stations écologiques favorables, des soins agronomiques appropriés en pépinière et en plantation, des techniques adéquates de germination, de greffage, etc...

La variabilité génétique ainsi stockée en collection est fortement marquée par l'échantillonnage et la structure de la population d'origine. Faisons remarquer que le mode d'installation d'une collection par semences est basé sur la collecte exclusive des seuls génotypes fructifères. Cette technique d'échantillonnage peut se révéler insuffisante quand elle ne concerne qu'une minorité d'individus. C'est une situation couramment rencontrée chez les arbres (mise à fruit cyclique) surtout dans leur habitat naturel forestier. Il est alors nécessaire d'associer à la récolte des fruits mûrs la collecte sur les plantes stériles d'axes **caulinaires** à greffer et bouturer (cf. l'exemple des caféiers, Tome I, chap. 2), sans perdre de vue les problèmes phytosanitaires liés à l'emploi de matériel végétatif.

Le passage par la multiplication végétative ne doit pas entraîner une restriction notable de la variabilité introduite. Il existe **certe** des souches plus ou moins récalcitrantes au bouturage et des phénomènes d'incompatibilité de greffe; mais il est souvent possible de réussir la multiplication végétative par un choix judicieux du stade et de la nature du fragment prélevé comme du porte-greffe d'une part, par la mise en **œuvre** de moyens techniques variés, d'autre part. En outre, comme nous le verrons plus loin, la reproduction des plantes par multiplication végétative est parfois génératrice de modifications persistantes du type de développement de la plante (**NOZERAN**, 1978).

La taille des plantes arbustives est une contrainte importante de leur installation en collection. Comme il est souhaitable de minimiser la compétition et les interactions défavorables aux génotypes rassemblés, les distances de plantations s'étagent de 2 x 2 m pour les arbustes (2500 plantes à l'ha) à 10 x 10 m pour les plus grands arbres (100 plantes à l'ha). La surface de terrain nécessaire au stockage de différentes origines géographiques devient alors le facteur limitant. **ROCHE** (1975) recommande pour les collections d'arbres forestiers d'affecter de 10 à 30 ha par provenance. Ceci explique que l'on réserve l'implantation de telles collections aux espèces d'une valeur économique certaine et en nombre limité.

La durée de vie des espèces arbustives (20 ans, 50 ans parfois plus de 100 ans) est par contre un avantage incontestable pour le maintien de l'intégrité génétique des collections sur une longue période renouvelable par multiplication végétative. Les risques d'érosion dépendent de l'adaptation des différentes souches rassemblées dans le même milieu et des soins apportés à l'entretien des plantes. Ils peuvent être considérablement réduits par différentes pratiques culturales. Rappelons les moyens mis en **œuvre** dans les collections de caféiers de Côte-d'Ivoire:

- Implantation en 2 lieux favorables, l'un aux espèces d'altitude (*C. arabi-ca* et *C. eugenioides* au mont Tonkoui 1100 m), l'autre aux espèces dites «de basse altitude» (Divo 200m).
- Plantation à Divo dans une forêt primaire aménagée par suppression du sous-bois afin de se rapprocher le mieux possible de l'ambiance écologique d'origine des caféiers.
- Greffage des espèces s'adaptant mal (*C. congensis*) ou peu vigou-reuses sur le porte-greffe *C. canephora* bien adapté aux conditions locales.
- Traitement des maladies et arrosage des jeunes plants en période de sécheresse.
- Reprise par greffe des souches déficientes ou accidentellement endo-magées.

Toutes ces précautions n'éliminent pas les risques climatiques excep-tionnels comme cyclones, tornades, sécheresse. La sécurité des collec-tions vivantes d'arbres n'est donc assurée que par duplication.

Vu le coût élevé de ces opérations, il paraît opportun de s'intéresser à la constitution de collections de plantes miniaturisées grâce à la multiplication végétative in vitro.

*b. Les collections de plantes à tubercules, à rhizomes et à bulbes.* Les différentes variétés de pomme de terre, de manioc, d'ignames, de pavots, de patates douces, de plantes florales à bulbes sont habituellement multipliées par voie végétative. Ces plantes ont souvent en culture un cycle annuel. Aussi pour aisée que soit leur multiplication, les souches en collec-tion doivent être reprises d'année en année. Outre le coût de telles opéra-tions, le risque le plus connu est l'accumulation de viroses entraînant la dégénérescence des souches. Ce risque existe potentiellement pour tous les types de collections reproduites par voie végétative et a été en partie résolu grâce à la culture de méristèmes in vitro (MOREL et MARTIN, 1955) et à la chimiothérapie de matériel végétatif.

Les moyens de stockage des racines et des tubercules tropicaux ne permettent qu'une conservation limitée de 1 ou 2 années entre 2 cam-pagnes de culture (MARTIN, 1975). Des efforts de recherches devraient être déployés pour établir les moyens de conservation à plus long terme de ce type de matériel végétal.

*c. Les collections de graminées fourragères.* Les collections de plantes fourragères implantées et stockées par graines issues de reproduction sexuée se rattachent au premier groupe de collections que nous avons présenté; l'intégrité génétique est alors difficile à préserver, surtout avec les espèces allogames.

Il est au contraire aisé d'atteindre cet objectif en maintenant les collec-tions de graminées fourragères par la voie végétative en utilisant des éclats de souches. Ils seront plantés à des distances de 1m x 1m ou plus, afin de bien individualiser les pieds, d'éviter tout mélange et la compétition et de faciliter l'entretien. Associée à la pérennité de nombreuses espèces fourra-gères, la multiplication végétative permet de ne reconstituer la collection que tous les 3, 5, voire même 10 ans. Cette méthode a surtout le défaut de transmettre les viroses accumulées au cours du temps.

Un autre procédé de multiplication des graminées fourragères permet aussi une reproduction conforme des différents génotypes en collection:

c'est la reproduction par graines d'origine **apomictique**. Elle est très répandue dans la plupart des familles d'Angiospermes mono et dicotylédones, principalement chez les graminées, les composées et les rosacées. VEYRET (1965) a dressé une liste très complète des genres concernés et de leurs différentes modalités d'apomixie.

Cette reproduction asexuée par voie de graines comprend soit l'**embryonie** adventice (**citrus**), soit le passage par le gamétophyte non réduit. Ce dernier mode d'apomixie di **gamétophytique** peut être obligatoire ou facultatif. Dans le premier cas, la descendance obtenue par graines est du type parental et parfaitement homogène. Dans le second cas, il subsiste une sexualité résiduelle qui se traduit dans la descendance par l'apparition de plantes de phénotypes différents de la plante-mère et appelée « hors-types ». C'est ce mode d'apomixie **gamétophytique** facultative que l'on trouve chez le *Panicum* (cf. Tome I).

Pour des collections de conservation de plantes fourragères, la reproduction par graines **apomictiques** est donc intéressante à considérer. Les espèces à reproduction **apomictique** obligatoire peuvent être ainsi multipliées par graines de génération en génération, avec une conservation exacte des génotypes de départ. Ce système est encore exploitable chez les espèces **apomictiques** préférentielles quand le taux de sexualité résiduelle reste suffisamment bas; les quelques plantes « hors-types » apparues sont facilement identifiables et éliminées. C'est la méthode retenue pour régénérer la collection de *Panicum* du centre ORSTOM d'Adiopodoumé (Côte-d'Ivoire). L'intérêt de ce passage par la graine est d'éliminer les viroses qui se sont accumulés au cours de la culture et des cycles de multiplication par éclats de souches.

### III. MOYENS DE CONSERVATION DES RESSOURCES GÉNÉTIQUES

#### A. STOCKAGE DE LONGUE DURÉE DES GRAINES

Chez les espèces végétales, la longévité des graines\* dans les conditions naturelles s'étend de quelques jours (au maximum un mois pour *Theobroma cacao* et *Hevea brasiliensis*), à des dizaines d'années (de l'ordre de 10 ans chez le genre *Gossypium*), et même des centaines d'années (*Mimosa glomerata*), voire un millier d'années (*Nelumbo*). On a coutume de distinguer les graines **microbiotiques** dont la longévité n'excède pas 3 ans, les graines **mésobiotiques**, les plus courantes, dont la durée de vie est au maximum de 15 à 20 années, et enfin les graines **macrobiotiques** qui maintiennent leur pouvoir de germination au-delà de plusieurs dizaines d'années ou de plusieurs siècles (HARRINGTON, 1970; MANGENOT, 1975).

Sous tous les climats, on connaît des espèces qui perdent très tôt leur pouvoir germinatif ou inversement. Cependant, le nombre d'espèces dont

---

\*La longévité des graines est le temps pendant lequel elles restent en état de vie ralentie (**quiescence**) sans cesser de pouvoir germer (MANGENOT, 1975).

les graines ont une faible longévité est plus élevé sous les tropiques humides que dans les régions à hivers marqués (MANGENOT, 1975).

Nous allons considérer les moyens de réaliser artificiellement, au meilleur coût, la conservation des ressources naturelles par le stockage de longue durée de graines. Celui-ci doit avoir pour objectif d'altérer le moins possible leur pouvoir germinatif et leur constitution génétique.

On sait empiriquement conserver les graines des espèces cultivées pour un ou plusieurs cycles culturaux. Le maintien de la longévité des graines est généralement favorisé par une forte déshydratation des semences, par la présence de téguments imperméables (légumineuses), par le stockage en atmosphère sèche et à basse température.

Il convient tout d'abord de distinguer deux types de comportement des graines stockées après avoir subi une déshydratation (ROBERTS, 1972). Les espèces dites « orthodoxes » présentent un allongement de la durée de la viabilité des graines d'autant plus important que l'on stocke des graines contenant moins de 5% d'eau à des températures proches ou inférieures à zéro°C; c'est la situation la plus répandue chez les plantes cultivées. Par contre, les espèces dites « récalcitrantes » ont un comportement inverse: la décroissance de la teneur en eau des graines en-dessous d'un seuil relativement élevé (12 à 31% suivant les espèces) entraîne une réduction de la viabilité. Bien que numériquement moins important, ce groupe contient beaucoup d'espèces à grosses graines parmi les essences forestières des genres *Taxus* et *Quercus*, les arbres fruitiers du genre *Citrus* et d'importantes cultures industrielles tropicales comme l'arbre en caoutchouc (*Hevea brasiliensis*), le palmier à huile (*Elaeis guineensis*), le cacaoyer (*Theobroma cacao*), les caféiers (*C. arabica* et *C. canephora*) le colatier (*Cola nitida*), etc... De même, THOMPSON (1974) distingue les graines généralement de petite taille conservées sèches, des graines souvent de grande taille conservées humides. Cet auteur considère que tous les comportements intermédiaires existent entre ces deux extrêmes.

Du fait du manque d'informations sur les conditions de conservation à long terme de la viabilité des espèces « récalcitrantes », nous allons surtout considérer les principes de conservation des espèces « orthodoxes ». Ce qui ne signifie pas que l'on se désintéresse des premières. Un effort de recherche est au contraire nécessaire pour les espèces « récalcitrantes »; quelques voies d'étude seront envisagées.

## 1. Données de base sur les conditions de stockage

Les 3 facteurs environnementaux à considérer sont la température, la teneur en eau des graines, et, à un moindre degré, la composition de l'atmosphère. Théoriquement, il serait nécessaire de déterminer les relations quantitatives de ces paramètres avec la viabilité observée.

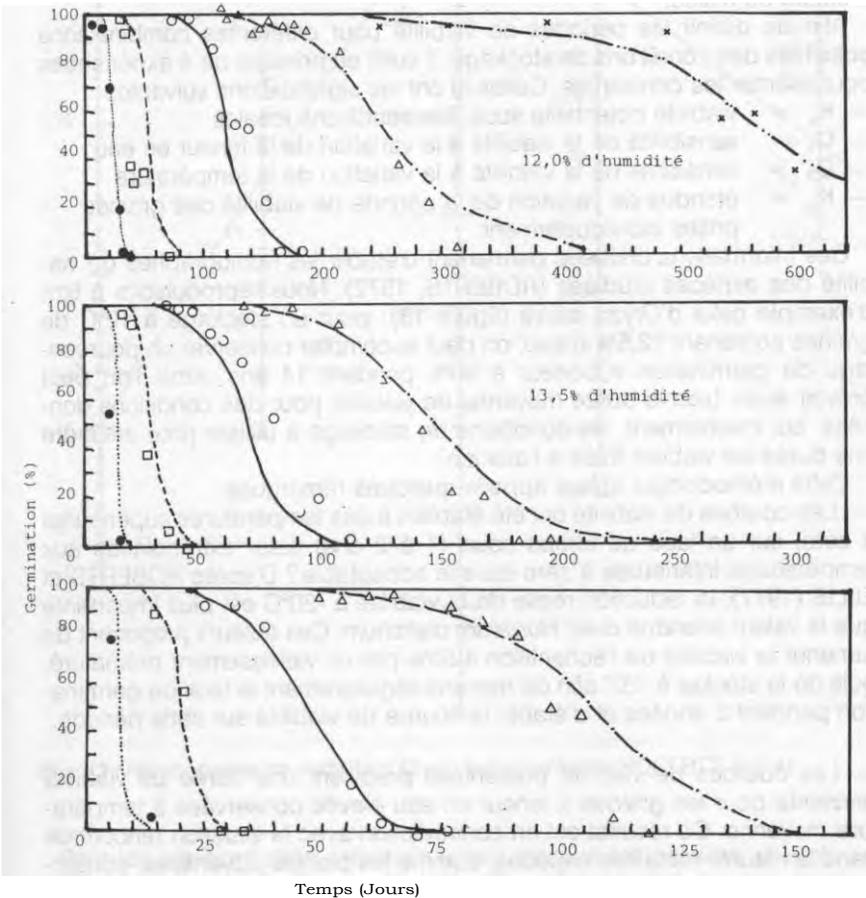
a. *Courbes de viabilité en fonction de la température et la teneur en eau.* Les courbes de viabilité ont été établies sur la base de l'évolution du taux de germination suivi pendant deux années, pour les 9 espèces suivantes: *Triticum aestivum*, *Oryza sativa*, *Vicia faba*, *Pisum sativum*, *Hordeum vulgare* et *distichum*, *Lycopersicon esculentum*, *Allium cepa*, *Lactuca sativa* (ROBERTS, 1972 et 1975; ROBERTS et ELLIS, 1977).

La situation observée chez ces espèces orthodoxes peut être décrite par 3 équations:

— La courbe de viabilité suit la fonction de répartition décroissante de la distribution normale. La figure 12 montre le cas de *Oryza sativa*. La fréquence  $y$  des mortalités individuelles à un instant donné  $p$ , d'une population de graines stockées dans des conditions fixées, suit la distribution normale

$$Y = \frac{1}{2\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{p-\bar{p}}{\sigma}\right)^2}$$

dans laquelle  $\bar{p}$  représente la période de viabilité moyenne et  $\sigma$  l'écart type des mortalités au temps  $p$ . La transformation probit de cette distribution donne une droite de pente négative.



**Fig. 12:** Courbes de survie des graines d'*Oryza sativa* à différentes températures et teneurs en eau (d'après ROBERTS, 1975)

- × 27°C
- △ 32°C
- 37°C
- 42°C
- 47°C

La valeur de  $a$  est directement proportionnelle à la période de viabilité moyenne  $p$ .

$$a = K_0 p \quad \text{avec } K_a = \text{constante.}$$

La relation entre la teneur en eau des graines  $m$  (en % de la matière humide), la température  $t$  (en degré centigrade) et la période de viabilité moyenne  $p$  suit la fonction linéaire suivante:

$$\text{Log } p = K_v - C_1 m - C_2 t$$

avec  
 $K_v, C_1$  et  $C_2 = \text{constante.}$

Pour les espèces étudiées, il n'apparaît pas d'interaction entre les deux facteurs du milieu.

Afin de définir les périodes de viabilité pour différentes combinaisons possibles des conditions de stockage, il suffit en principe de 4 expériences pour estimer les constantes. Celles-ci ont les significations suivantes:

- $K_v$  = viabilité potentielle sous des conditions idéales
- $C_1$  = sensibilité de la viabilité à la variation de la teneur en eau
- $C_2$  = sensibilité de la viabilité à la variation de la température
- $K_a$  = étendue de variation de la période de viabilité des graines prises individuellement.

Ces informations chiffrées permettent d'établir les **nomogrammes** de viabilité des espèces étudiées (ROBERTS, 1972). Nous reproduisons à titre d'exemple celui d'*Oryza sativa* (figure 13): pour un stockage à 10°C de graines contenant 12,5% d'eau, on peut escompter conserver un pourcentage de germination supérieur à 90% pendant 14 ans. Ainsi l'on peut prévoir aussi bien la durée moyenne de viabilité pour des conditions données, ou, inversement, les conditions de stockage à utiliser pour atteindre une durée de viabilité fixée à l'avance.

Cette méthodologie idéale appelle quelques remarques:

— Les courbes de viabilité ont été établies à des températures supérieures à zéro, sur un laps de temps court (1 à 2 ans). Leur extrapolation aux températures inférieures à zéro est-elle acceptable? D'après ROBERTS et ELLIS (1977), la réduction réelle de la viabilité à -20°C est plus importante que la valeur attendue chez *Hordeum distichum*. Ces auteurs proposent de ramener la viabilité de l'échantillon à 50% par un vieillissement prématuré, puis de la stocker à -20° afin de mesurer régulièrement le taux de germination pendant 2 années et d'établir la courbe de viabilité sur cette période.

— Les courbes de viabilité présentées prédisent une durée de viabilité restreinte pour les graines à teneur en eau élevée conservées à température moyenne. Ce résultat est en contradiction avec la situation rencontrée dans la nature: certaines espèces, comme les plantes adventices, conservent pendant plusieurs années leur pouvoir germinatif dans le sol avec des graines imbibées.

VILLIERS (1975) a démontré que le pouvoir germinatif des graines imbibées de *Lactuca sativa* se maintenant au-dessus de 90% pendant 18 mois, c'est-à-dire l'équivalent d'un stockage à 30°C de graines de laitue contenant 50% d'eau.

Période moyenne  
de viabilité, jours  
(axe central)

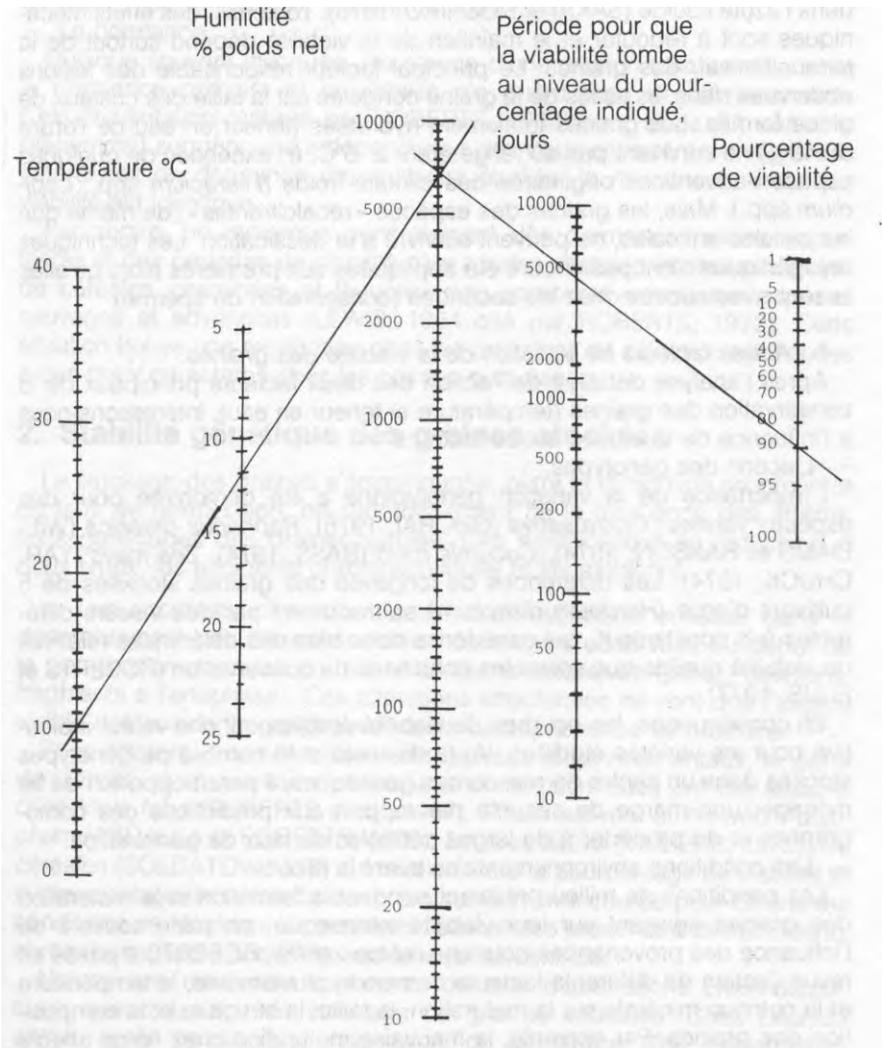


Fig. 13: Nomogramme de viabilité d'*Oryza sativa* (d'après ROBERTS, 1972).

Pour les espèces dites «récalcitrantes» comme les caféiers, des résultats encourageants ont été obtenus avec des graines imbibées conservées entre 15° et 19°C: la viabilité est maintenue à son niveau initial pendant 1 à 2 ans suivant les espèces (VAN DER VOSSEN, 1978; COUTURON, 1979). Les courbes de viabilité pour différentes combinaisons des facteurs du milieu ont été établies. La durée de viabilité est encore insuffisante. D'autres expérimentations sont nécessaires à la mise au point de la conser-

vation pendant une dizaine d'années de graines des espèces « récalcitrantes ».

— La survie des graines à faible teneur en eau de *Medicago sativa*, *Oryza sativa*, *Triticum aestivum* peut même être assurée à très basse température dans l'azote liquide (SAKAI et NOSHIRO, 1975). Toutefois, des effets mécaniques sont à redouter et le maintien de la viabilité dépend surtout de la teneur en eau des graines. Le principal facteur responsable des lésions observées dans les tissus de la graine congelée est la taille des cristaux de glace formés. Les graines totalement hydratées (teneur en eau de l'ordre de 40%) ne survivent pas au réfrigérateur à -5°C, à l'exception de quelques espèces adventices originaires des climats froids (*Hieracium* spp., *Lepidium* spp.). Mais, les graines des espèces « récalcitrantes », de même que les cellules animales, ne peuvent survivre à la dessiccation. Les techniques cryogéniques n'ont pas encore été appliquées aux premières alors qu'elles le sont avec succès chez les secondes (conservation du sperme).

#### b. Autres facteurs de variation de la viabilité des graines.

Après l'analyse détaillée de l'action des deux facteurs principaux de la conservation des graines (température et teneur en eau), intéressons-nous à l'influence de quelques autres facteurs.

— L'action des génotypes.

L'importance de la variation génotypique a été démontrée pour des espèces variées: *Oryza sativa* (GHORAI, 1976), *Raphanus brassica* (WILLIAMS et HANSON, 1974), *Cucumis melo* (BASS, 1974), *Zea mays* (YARCHUCK, 1974). Les différences de longévité des graines stockées de 5 cultivars d'orge (*Hordeum distichum*) se traduisent par des valeurs différentes à la constante  $K_v$  qui caractérise donc bien des différences relatives de viabilité quelles que soient les conditions de conservation (ROBERTS et ELLIS, 1977).

En conséquence, les courbes de viabilité établies ont une valeur indicative pour les variétés étudiées. Vu la diversité et le nombre de génotypes stockés dans un centre de ressources génétiques, il paraît opportun de se ménager une marge de sécurité par rapport aux prédictions des nomogrammes et de procéder à de larges contrôles du taux de germination.

— Les conditions environnementales avant la récolte.

Les conditions de milieu prévalant pendant la formation et la maturation des graines agissent sur leur viabilité intrinsèque; on parle souvent de l'influence des provenances pour une même variété, ROBERTS a passé en revue l'action de différents facteurs comme la pluviométrie, la température et la nutrition minérale sur la maturation, la taille, la structure et la composition des graines. Par exemple, la mauvaise maturation chez l'orge affecte significativement la longévité des graines stockées, par réduction de la valeur de la constante  $K_v$  (ROBERTS et ELLIS, 1977).

— Les conditions de récolte et de conditionnement des graines.

Les conditions de récolte et de préparation des graines modifient aussi la viabilité au stockage. Les traumatismes mécaniques des graines affectent leur teneur en eau; ils sont la porte ouverte aux infections et provoquent des lésions des embryons réduisant largement leur viabilité (MOORE, 1972).

De même, les méthodes de triage des graines et de séchage (en particulier les températures élevées) peuvent affecter qualitativement la viabilité.

— L'environnement gazeux des graines stockées.

Les courbes de viabilité sont calculées pour des graines conservées en boîtes hermétiques. A l'air libre, la durée de viabilité est légèrement réduite; cet effet est peu important dans la pratique. Cette détérioration limitée du pouvoir germinatif des graines en présence d'oxygène peut être évitée en stockant sous atmosphère d'azote ou sous vide.

— La dormance.

Quoique souvent invoquée, l'existence d'une relation fonctionnelle entre la dormance primaire et la viabilité n'a guère reçu de démonstration. L'expérimentation réalisée par ROBERTS (1963) pour le riz (*O. sativa* et *glaberrima*) suggère une indépendance des deux caractères: malgré des différences de dormance des cultivars étudiés, la décroissance de leur viabilité est identique.

Par contre, on remarque généralement des dormances moins importantes et des périodes de viabilité plus courtes chez les espèces cultivées de céréales, graminées et légumineuses comparativement aux espèces sauvages et adventices (LEWIS, 1964 cité par ROBERTS, 1972). Cette situation trouve une explication dans les pressions de sélection appliquées à ces deux caractères chez les plantes cultivées.

## 2. Stabilité génétique des graines stockées

Le stockage des graines s'accompagne, outre la diminution progressive du taux de germination, de l'accroissement de la fréquence des aberrations génétiques des graines viables. Ces modifications se rattachent à deux types, les accidents chromosomiques et les mutations.

a. *Les accidents chromosomiques.* Leur estimation consiste dans le dénombrement des aberrations chromosomiques observées au début de la germination dans les premières divisions mitotiques (ponts, délétions, fragments à l'anaphase). Ces altérations structurales ne sont que l'aspect visible des accidents génétiques variés susceptibles de se produire.

Pour différentes combinaisons des facteurs environnementaux, la perte de viabilité et la quantité d'aberrations chromosomiques ont été établies chez *Vicia faba* (ROBERTS et al., 1967), *Pisum sativum* et *Hordeum distichum* (ABDALLA et ROBERTS, 1969), *Lactuca sativa* (VILLIERS, 1975), les céréales (SOLDATOV, 1976). La relation entre le pourcentage de viabilité et le pourcentage de cellules aberrantes suit la même courbe pour différentes conditions de stockage; la proportion de cellules présentant des accidents chromosomiques s'accroît quand la viabilité diminue.

Notons une diminution importante du taux d'aberrations chromosomiques quand il s'agit du stockage de graines imbibées. Chez *Lactuca sativa*, après un an de conservation, on dénombre 13% d'aberrations chromosomiques chez les graines contenant 5% d'eau et 2% seulement chez les graines imbibées (VILLIERS, 1975; VILLIERS et EDGCUMBE, 1975).

Les accidents chromosomiques dus au vieillissement des graines stockées avec une faible teneur en eau peuvent être maintenus à leur niveau ou même légèrement réduits en poursuivant la conservation après imbibition des graines (VILLIERS, 1975). Cette réversion, bien que démontrée chez *Lactuca sativa*, n'a pas toujours été observée et semble inopérante quand le taux des modifications génétiques et cellulaires dépasse 15%.

b. *Les mutations.* A la germination des graines conservées, on dénombre les modifications génétiques visibles chez les jeunes plantes (forme et taille des organes, coloration...). Après une perte de viabilité des graines stockées de 50%, ABDALLA et ROBERTS (1969) observent 1 à 4% de mutations chez les trois espèces étudiées. Ce taux paraît très élevé et équivaldrait à un traitement des graines fraîches par les rayons X à la dose de 10.000 r.

Des graines de laitue contenant 7% d'eau, conservées pendant 3 mois à 30°C puis mises en germination conduisent à de nombreuses anomalies et à une grande diversité des plantules pour la croissance, la pigmentation, l'enracinement... Au contraire, les graines imbibées conservées plus d'un an produisent des jeunes plantes vigoureuses et homogènes (VILLIERS, 1975).

En conclusion, il y a une bonne corrélation entre la quantité de modifications génétiques et la décroissance de la viabilité d'un lot de semences stockées. On doit donc rechercher les conditions de conservation assurant les pertes minimales de viabilité et une multiplication des lots dès que le taux de germination tombe de 5 à 10% par rapport au taux initial.

La perte de viabilité est due aux modifications du génome, des systèmes enzymatiques et des membranes. L'existence de ces modifications nucléaires et cytoplasmiques dans les graines conservées à sec a été prouvée par observation en microscopie électronique. En l'absence de division cellulaire dans les graines conservées à sec, il y a vieillissement des macromolécules et des structures cellulaires; les accidents ne peuvent que s'accumuler. Les dommages dus au vieillissement résultent des réactions **peroxydasiques** causant la détérioration des lipides et des protéines des membranes. Par contre, le maintien de la division cellulaire et la sélection de nouvelles lignées cellulaires assurent le rajeunissement des graines conservées imbibées (VILLIERS, 1975).

### 3. Aspects pratiques du stockage des graines à long terme

Les objectifs à atteindre par un stockage à long terme des graines sont les suivants:

- réduire les pertes et les modifications génétiques dues à la multiplication sexuée, surtout chez les espèces à cycle de génération court,
- minimiser les altérations génétiques consécutives au stockage,
- utiliser des conditions techniques et du matériel assurant une conservation aux plus faibles coûts.

Afin de concilier «sécurité — intégrité génétique — meilleur coût», le stockage d'espèces «orthodoxes» est réalisé avec des graines à faible teneur en eau (5% environ), à des températures proches de zéro degré (conservation à moyen terme au frigidaire à 4° C) ou inférieures à zéro (conservation à long terme au congélateur à -18°C). Le seuil de viabilité recherché (supérieur à 85 - 90%) permet de déterminer en moyenne, pour les espèces dont le **nomographe** est connu, les différentes combinaisons possibles des facteurs assurant une durée de conservation déterminée. Il ne faut pas pour autant perdre de vue les divers facteurs de variation du taux de viabilité (variation **génotypique**, facteurs extrinsèques comme la maturité, la récolte et la préparation des semences).

Outre le but recherché (conservation à court, moyen ou long terme, le cadre dans lequel se place ce stockage peut revêtir des formes variées:

- conservation soit d'une seule espèce, soit de plusieurs espèces appartenant à une même famille de plantes ou à différentes familles, soit différents groupes de plantes originaires de la même zone écologique,
- structure d'accueil d'un centre de recherche à vocation locale, nationale ou internationale.

Passons en revue les problèmes techniques essentiels du stockage de longue durée des graines.

*a. La teneur en eau des graines.* Elle peut être ajustée à un niveau fixé à l'avance soit par séchage préalable à des températures non létales et conservation dans des containers hermétiques, soit par l'équilibre hygroscopique qui s'établit dans une graine placée dans une atmosphère d'humidité relative contrôlée.

Cette relation d'équilibre varie avec les espèces considérées. Par exemple, pour une humidité relative de 40%, la teneur en eau à l'équilibre des graines de céréales et de légumineuses est de l'ordre de 9 à 12%; celles des graines oléagineuses comme le soja ou l'arachide n'est que de 6 à 7%.

Il faut aussi noter que l'équilibre hygroscopique de la graine avec l'humidité de l'air ambiant n'est pas le même à la désorption (phénomène d'hystérésis). La teneur en eau des graines en désorption est de 1 à 1,5% plus élevé qu'à l'absorption, en moyenne.

Les méthodes de détermination de la teneur en eau des graines ainsi que l'équilibre hygroscopique à différentes humidités relatives d'un grand nombre d'espèces végétales ont été récapitulées par ROBERTS et ROBERTS (1972).

*b. Les modes de conservations utilisés.* La solution adoptée par le laboratoire national japonais pour la conservation des ressources génétiques par stockage de graines est la suivante: séchage pour atteindre une teneur en eau de 4 à 6% et ensachage sous vide; stockage en chambre froide à -10° et -1°C (ITO, 1972; WOLD, 1975). La constitution de petits échantillons au départ permet leur utilisation indépendante; tout sachet ouvert est utilisé à bref délai. Cette méthode est grandement facilitée par l'emploi des sachets plastiques recouverts intérieurement par une couche de papier aluminium; cette sacherie s'est beaucoup développée dans le commerce des semences potagères et florales.

Le laboratoire national de stockage de graines de Fort Collins aux Etats-Unis utilise des chambres à air conditionné: stockage à 4° C avec une humidité relative de l'air de 32% et aussi à -12°C (JAMES, 1972). Les containers ne sont pas scellés ce qui facilite les prises d'échantillons successives; il n'y a pas de risque de réhumidification dans cette atmosphère sèche. La sécurité de ce stockage réside dans la maintenance et la réparation immédiate en cas de panne des systèmes de climatisation.

Le jardin botanique royal de Kew (Angleterre) conserve les graines séchées à 35°C pendant 24 heures, puis dans des piluliers stockés en chambre froide à -10°C ou 2° à 5°C (THOMPSON, 1974).

Ces différents exemples de centres de conservation des graines à long terme nous fournissent les bases techniques d'un tel stockage. Le même résultat est atteint à un niveau plus modeste par différents pays et centres

de recherches en utilisant une salle climatisée, une chambre froide ou des congélateurs organisés suivant les mêmes principes.

Un groupe de travail de l'IBPGR a établi un rapport sur les systèmes de stockage des graines à utiliser et leur prix de revient (IBPGR, 1976). L'investissement de base est constitué par le bâtiment et les systèmes de contrôle de l'environnement (température et humidité). Il ne faut pas oublier que le froid coûte cher et requiert pour la maintenance du personnel qualifié.

Rappelons que la survie des graines peut aussi être assurée à très basse température, dans l'azote liquide. SAKAI et NOSHIRO (1975) conseillent de procéder comme suit:

- ramener la teneur en eau des graines entre 8 à 15%,
- conditionner les graines dans des sacs en aluminium ou dans des tubes plastiques bouchés,
- immerger les containers dans l'azote liquide directement en partant de la température ambiante,
- après le stockage, l'échantillon de graines est replacé à une température de l'air comprise entre 0 à 20°C pour un réchauffement lent.

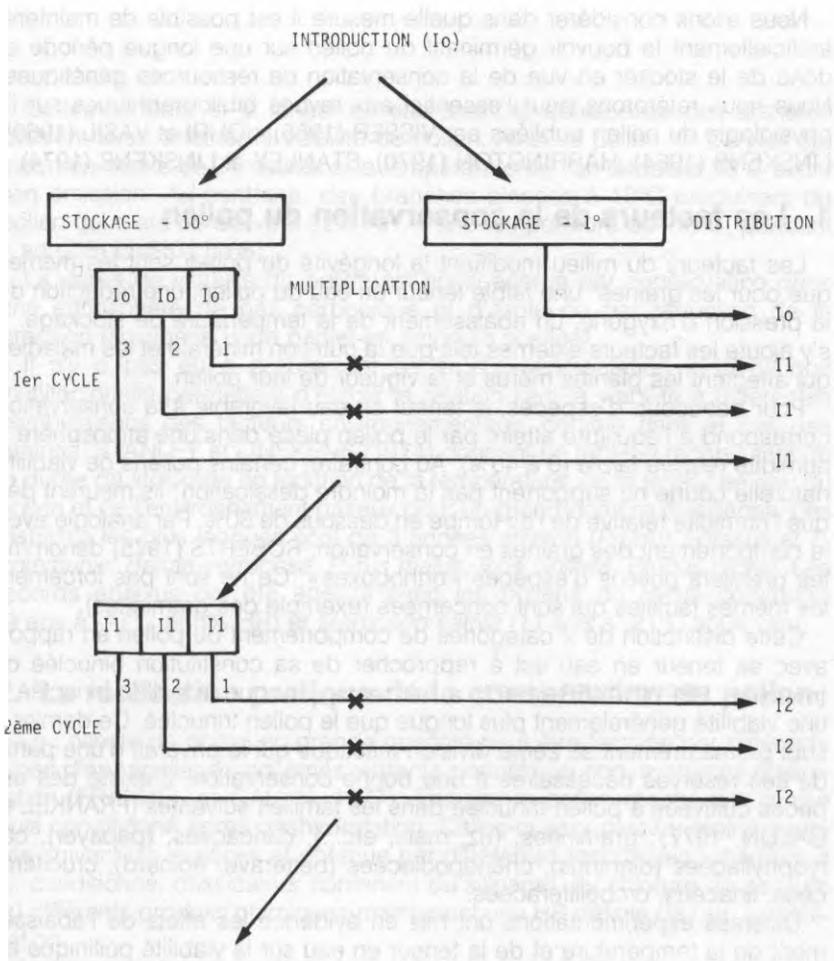
*c. Contrôle du taux de germination.* Un contrôle à intervalles réguliers du taux de germination est indispensable. Par exemple, on effectuera un sondage tous les 5 ans pour des graines dont la période de conservation théorique est de plusieurs dizaines d'années. Les conditions de germination ne sont pas toujours bien définies pour chaque espèce. L'emploi d'étuves de germination permettant le contrôle de la température, de l'humidité et de l'éclairage s'avère indispensable. Cependant le taux de germination estimé en boîte de Pétri est différent de celui obtenu en terre (*Panicum*).

Rappelons aussi que les graines conservées doivent donner naissance à des jeunes plantes normales. Le simple jugement de la germination sur l'émission d'une radicule est insuffisant. Il est nécessaire d'observer l'intégrité des plantules dans les premiers stades de leur développement.

*d. La régénération des lots.* Dès que la viabilité des graines tombe en-dessous de 85%, une multiplication de l'échantillon doit être décidée. Les problèmes posés par la régénération par voie sexuée ont été abordés précédemment; il s'agit d'assurer la plus grande intégrité génétique possible.

La distribution du matériel d'une banque de graines aux utilisateurs est aussi un facteur d'épuisement des lots nécessitant leur multiplication. Il y aura d'ailleurs intérêt à réaliser une multiplication des échantillons initiaux de faible quantité avant toute distribution.

De façon à réduire le nombre de régénérations des graines conservées, ITO (1972) rapporte le système de stockage à 2 voies utilisé au Japon. Le lot de semences reçu initialement est divisé en 2 échantillons. L'un servira à la conservation et à la multiplication; il est stocké dans les meilleures conditions (-10°C). L'autre sera distribué au fur et à mesure des demandes et conservé à moyen terme à -1°C. Le stock de graines à distribuer est renouvelé grâce au lot en conservation long terme. Ce système est résumé au schéma 13.



**Schéma 13:** mode de stockage des grains utilisé au Japon (ITO, 1972).

## B. LE STOCKAGE DE LONGUE DUREE DU POLLEN

La durée de survie du pollen dans les conditions naturelles varie de quelques heures à plus d'une année. Le pollen reste viable 4 heures chez le goyavier, 12 heures chez le cacaoyer, jusqu'à 24 heures chez la plupart des graminées (maïs, riz, blé...). Au contraire, cette viabilité naturelle du pollen atteint 1 année chez les arbres à pollinisation anémophile comme les prunus et le dattier.

Une classification des familles botaniques en rapport avec la longévité naturelle du pollen a été établie par **HOLMAN** et **BRUBAKER** (1922) et rapportée par **HARRINGTON** (1970): durée de vie longue (rosacées, légumineuses, pinacée...), intermédiaire (renonculacées...), courte (graminées...).

Nous allons considérer dans quelle mesure il est possible de maintenir artificiellement le pouvoir germinatif du pollen sur une longue période et donc de le stocker en vue de la conservation de ressources génétiques. Nous nous référerons pour l'essentiel aux revues bibliographiques sur la physiologie du pollen publiées par VISSER (1955), JOHRI et VASIL (1960), LINSKENS (1964), HARRINGTON (1970), STANLEY et LINSKENS (1974).

## 1. Les facteurs de la conservation du pollen

Les facteurs du milieu modifiant la longévité du pollen sont les mêmes que pour les graines: une faible teneur en eau du pollen, une réduction de la pression d'oxygène, un abaissement de la température de stockage. Il s'y ajoute les facteurs externes tels que la nutrition minérale et les maladies qui affectent les plantes mères et la vigueur de leur pollen.

Pour beaucoup d'espèces, la teneur en eau favorable à la conservation correspond à l'équilibre atteint par le pollen placé dans une atmosphère à humidité relative faible (0 à 40%). Au contraire, certains pollens de viabilité naturelle courte ne supportent pas la moindre dessiccation; ils meurent dès que l'humidité relative de l'air tombe en dessous de 50%. Par analogie avec le comportement des graines en conservation, ROBERTS (1975) dénomme les premiers pollens d'espèces « orthodoxes ». Ce ne sont pas forcément les mêmes familles qui sont concernées (exemple des graminées).

Cette distinction de 2 catégories de comportement du pollen en rapport avec sa teneur en eau est à rapprocher de sa constitution binucléé ou trinucléé. BREWBAKER (1967) a fait remarquer que le pollen binucléé a une viabilité généralement plus longue que le pollen trinucléé. Ce dernier a subi prématurément sa 2ème division mitotique qui le priverait d'une partie de ses réserves nécessaires à une bonne conservation. Il existe des espèces cultivées à pollen trinucléé dans les familles suivantes (FRANKEL et GALUN, 1977): graminées, (riz, maïs, etc...), caricacées, (papayer), Caryophyllacées (dianthus), chénopodiacées (betterave, épinard), crucifères, linacées, ombellifères.

Diverses expérimentations ont mis en évidence les effets de l'abaissement de la température et de la teneur en eau sur la viabilité pollinique en fonction du temps. Par exemple, la durée de conservation à 3°C du pollen de *Pyrus communis* avec un taux de germination in vitro supérieur à 60% varie dans de larges proportions avec sa teneur en eau (VISSER, 1955):

— humidité relative (%):	0	10	25	40	55	70	100
— conservation (jours):	42	345	345	207	149	22	0

La combinaison des 2 facteurs est déterminante chez *Vitis vinifera* comme le suggèrent les résultats suivants (MOTI, 1972):

- température ambiante + HR ambiante = viabilité de 10 à 25 jours
- température ambiante + HR 0 à 25% = viabilité de 7 à 8 mois
- température -23°C + HR 0% = viabilité de 11 à 14 mois.

Les données sur la conservation du pollen de l'hybride rose-thé sont du même ordre (VISSER et al., 1977).

Aux températures très basses de -180°C à -271°C obtenues avec les gaz liquéfiés l'aptitude du pollen à germer peut être prolongée. Ainsi, la viabilité du pollen de maïs qui décroît en quelques jours aux températures proches de 0°C, a pu être maintenue une année par un stockage à -196°C avec une teneur en eau de 22% (BARNABAS et RAJKI, 1976).

La température et la teneur en eau avant la déhiscence des anthères peuvent aussi affecter la viabilité du pollen. Ainsi, le pollen de *Corylus sp.* meurt en moins de 24 heures si la température de l'air dépasse 23°C avant son émission. Au contraire, des branches placées à 18°C produisent du pollen qui a été conservé à 92% HR, à une température de -18°C, pendant 1 an (ZIELENSKI, 1968).

La durée de viabilité du pollen est aussi accrue par conservation dans une atmosphère de gaz carbonique et d'azote ou par réduction de la pression d'oxygène (vide partiel en ampoule scellée).

Il n'y a pas eu à proprement parler d'expérimentation systématique multifactorielle permettant d'établir la relation liant la viabilité à différentes combinaisons des facteurs environnementaux, comme dans le cas des graines. STANLEY et LINSKENS (1974) donnent un tableau récapitulatif de la durée de stockage en fonction de la température, de la teneur en eau du pollen et de l'environnement gazeux pour un grand nombre d'espèces. Les viabilités les plus longues sont de 3 années environ (*Prunus cerasus* et *P. communis*, *Betula verrucosa*, *Pyrus malus*, *Vitis vinifera*, *Prunus nigra*). Les records absolus ont été atteints avec les pollens de *Pyrus communis* (9 ans à -17°C et 5% HR) et *Medicago sativa* (11 ans à -21°C sous vide).

## 2. La réalisation pratique de la conservation du pollen

La collecte de pollen en grande quantité peut être facilitée par différents types d'appareillage: systèmes vibrants portatifs ou non, systèmes d'aspiration (FRANKEL et GALUN, 1977). Ensuite, le pollen récolté est tamisé puis conditionné après déshydratation. Celles-ci sera plus ou moins poussée suivant les espèces et obtenue par différentes techniques: chambre à air conditionné, dessiccateur contenant du silicagel, du chlorure de calcium ou différents produits chimiques maintenant une HR définie de l'air, lyophilisation.

La conservation du pollen de nombreuses espèces est assurée après dessèchement, par conditionnement en tube scellé (vide partiel) et stockage au réfrigérateur (4° à 5°C) ou mieux au congélateur (-10° à -35°C). Cette méthodologie est couramment utilisée par les sélectionneurs qui doivent hybrider des géniteurs à floraisons séparées dans le temps et dans l'espace au cours de l'année comme les caféiers (WALYARO et VAN DER VOSSSEN, 1977).

Le pollen desséché par lyophilisation, conditionné en tube scellé, permet une conservation de longue durée à la température ambiante (KING, 1965). Chez le cocotier, BENARD (1973) a maintenu la viabilité 6 mois par cette méthode. Pour retrouver sa faculté germinative, le pollen lyophilisé doit nécessairement être réhydraté avant l'emploi.

Le pollen sec conservé aux très basses températures obtenues avec les gaz liquéfiés n'est pas endommagé. Avec ce procédé, le pollen de noyer a conservé sa viabilité et son pouvoir fécondant 2 ans (FARMER et BARNETT, 1974). Les conditions de la décongélation sont aussi à prendre en compte.

Le stockage du pollen dans les solvants organiques aurait l'avantage de maintenir une HR spécifique et de permettre le transport sans assujettisse-

ment aux appareils de réfrigération. Cette technique donne des résultats très variables avec les solvants et les espèces étudiées ( IWANAMI, 1972). A notre connaissance aucun essai de longue conservation n'a été entrepris avec des solvants peu chers comme l'acétone, le benzène ou l'éther de pétrole.

Il est nécessaire de contrôler régulièrement la viabilité du pollen en conservation de longue durée. On s'intéressera suivant les tests utilisés soit au pouvoir germinatif du pollen, soit à son pouvoir fécondant. Habituellement on se contentera de déterminer le premier bien que du pollen capable de germer in vitro puisse dans certains cas s'avérer incapable de conduire à une fécondation naturelle (STANLEY et LINSKENS, 1974).

Les tests de germination in vitro revêtent différentes modalités, goutte pendante, spot, milieu gélosé... La composition du liquide ou du milieu de germination du pollen influe sur le taux et la croissance des tubes polliniques. Certains sucres (sucrose, raffinose, lactose, maltose), l'adjonction d'acide borique et de chlorure de calcium favorisent la germination du pollen du chou (CHIANG, 1974).

On se contente parfois de tests de viabilité pollinique sans germination. Il s'agit soit de colorations spécifiques des composants du pollen (iodure de potassium, carmin acétique, acridine orange en fluorescence), soit des réactions enzymatiques d'oxydo-réduction avec des sels de tétrazolium, ou le diacétate de fluoresceine (HESLOP-HARRISON, 1970).

Idéalement, la germination in vivo est estimée par la formation de zygotes et de graines issues d'une pollinisation contrôlée. Cette manipulation longue et complexe peut être avantageusement remplacée par une observation en fluorescence de la croissance des tubes polliniques dans le style et jusqu'à sa pénétration dans l'ovule. Cette technique d'étude de l'incompatibilité repose sur la coloration au bleu d'aniline des amas de callose de la paroi du tube qui se colorent spécifiquement en jaune en éclairage UV (MARTIN, 1958).

### 3. Conclusion

La conservation à long terme du pollen présente l'avantage sur les graines d'occuper un moindre volume. Les équipements de stockage et de contrôle des facteurs physiques environnementaux sont de même nature (conditionnement sous vide, appareils de réfrigération ou congélation, etc...). Le coût de la conservation du pollen serait donc avantageux par rapport aux graines.

Les durées de conservation atteintes avec le pollen sont insuffisantes et très inférieures à celles des graines. Nous avons trouvé dans la littérature des viabilités polliniques de 3-4 années, exceptionnellement 9 ans (poirier) et 11 ans (luzerne). La solution de banques de pollen n'est donc pas encore au point. Un important travail expérimental reste à effectuer pour déterminer par espèce, les combinaisons de facteurs permettant d'atteindre des viabilités de plusieurs dizaines d'années.

Néanmoins, les techniques de conservation du pollen rendent de grands services aux généticiens dans la conduite de leurs hybridations contrôlées. Elles sont aussi exploitables par les centres de ressources génétiques pour distribuer, sous un faible volume, du pollen viable destiné aux programmes d'évaluation en croisement ou d'amélioration génétique. Cette utilisation

s'apparente aux techniques classiquement employées pour le sperme des animaux domestiques. L'échange du pollen viable entre stations permettrait de gagner plusieurs années dans le déroulement de recherches concernant des plantes à cycle long transférées par graines ou fragments et d'assurer un transfert sans risque phytosanitaire.

## C. LES TECHNIQUES CLASSIQUES DE LA MULTIPLICATION VÉGÉTATIVE

Point n'est besoin de s'étendre sur les techniques traditionnelles de multiplication végétative largement utilisées en horticulture et en agriculture. Cependant, il nous paraît utile de porter notre attention d'une part sur les progrès récents réalisés dans ce domaine, et d'autre part sur la variabilité du matériel végétal reproduit.

### 1. Les progrès techniques

Nous entendons par là les moyens qui ont permis la multiplication végétative de nouvelles espèces ou de souches récalcitrantes. L'évolution des matériaux est une source importante de progrès: l'emploi de serres, films plastique et de nébulisateurs variés...

Les recherches dans le domaine de la morphogenèse et de la physiologie du développement (substances de croissance, induction de la rhizogenèse, etc...) sont également responsables des succès par greffage ou bouturage.

— Le greffage: les progrès concernant le greffage ont eu lieu dans deux directions, d'une part vers une meilleure compréhension des relations génétiques entre le porte-greffe et le greffon (de l'incompatibilité de greffe jusqu'à des interactions positives), ainsi un bon choix de porte-greffe confère une vigueur et un développement rapide du greffon. D'autre part, en diversifiant et en analysant mieux la nature et l'état des greffons. Ainsi sur l'hévéa la reprise est plus rapide quand on greffe un bourgeon issu d'une aisselle de feuille que d'une aisselle d'écaille. En déterminant bien chez les arbres fruitiers européens la nature du bois (bois vieux ou bien bois de l'année) sur lequel le bourgeon est prélevé et l'état de repos ou d'activité du donneur de greffon (particulièrement dans le cas de plantes à croissance rythmique). Il est devenu possible, sur les caféiers (COUTURON et BERTHAUD, 1981) de greffer directement des embryons; cette technique a permis de sauver des génotypes hybrides interspécifiques pour lesquels les plantes ne pouvaient se développer du fait d'incompatibilité entre le tissu diploïde de l'albumen et le tissu diploïde de l'embryon.

— le bouturage: des analyses de morphogenèse (stade de développement, nature des éléments à bouturer) ont de la même manière permis d'améliorer les techniques de bouturage, l'emploi judicieux de substances de croissance et la réalisation de microenvironnements favorables (bouturage sous brouillard) ont rendu possible des multiplications végétatives de plantes (pins) ou d'organes (feuilles de luzerne ou d'ignames) particulièrement récalcitrants au clonage.

## 2. La reproduction conforme

Quoique la multiplication végétative assure en principe une reproduction fidèle du génome, on n'est pas assuré d'obtenir une plante possédant le même type de développement.

La notion de clone est maintenant beaucoup moins assurée qu'il y a quelques années. Il faut d'abord distinguer le clonage de plantes dont la multiplication végétative était déjà spontanément un moyen de dissémination important de celles pour lesquelles, avant les progrès du greffage ou du bouturage on ne connaissait que des multiplications à partir de graines. Il était difficile de mettre en doute la fidélité du clonage tant qu'on ne pouvait disposer de clones d'origines **morphogénétiques** variées ou de références à une multiplication par graines qui serait équivalente à un clonage.

Sur l'igname (**HAOUSSOU** et **TOURE**, 1982), il est possible d'assurer une multiplication végétative à partir d'organes variés, soit classiquement consacrés à cette multiplication (fragments de tubercules, bulbilles) soit nouvellement mis au service du clonage (bouture d'axes et de feuilles). L'expérimentation démontre que le choix de l'organe utilisé pour le clonage a une incidence remarquable sur la variabilité des plantes obtenues; théoriquement elles ont toutes le même génotype mais les réalisations du développement sont très variées; les fragments de tubercules ont des potentialités différentes et on peut homogénéiser une culture en choisissant convenablement les fragments par référence à leur position dans le tubercule. Les bulbilles conduisent à des reprises plus homogènes.

Sur le *Panicum maximum* (**PERNES** et al., 1970), grâce au mode de reproduction **apomictique** (cf. chapitre Panicum, Tome I) il était possible de comparer des multiplications **clonales** par graines et par boutures (qui pouvaient également être classées et repérées selon la morphogenèse de leurs touffes). L'analyse comparée des clones obtenus par les deux voies (graines, boutures) à partir de génotypes variés conduisait à des lectures différentes de leurs différenciation: certaines variations entre clones d'origine **génotypique** différente s'effaçaient après passage par graines **apomictiques** alors qu'elles étaient indéfiniment entretenues par bouturage.

Ces résultats quantitatifs peuvent être rapprochés des observations de **BANCILLHON** (1972) et **ROUX** (1970) qui démontraient que des boutures prélevées sur des rameaux **plagiotropes** de *Phyllanthus* créaient des clones de plantes qui ne retrouvaient pas un port dressé et ne construisaient pas aisément d'axes **plagiotropes** par contre les boutures prélevées sur un rameau orthotrope permettent l'acquisition d'un clone de plantes à l'architecture normale (**plagiotope**, orthotrope). Des observations comparables ont été faites sur le caféier, sur une légumineuse fourragère du gène *Hedysarum* (**FIGIER**, 1982), et ces différenciations d'axes, non parfaitement reconverties au cours des clonages, se transposent aussi au niveau du système **racinaire** comme l'indiquent des travaux sur le cacaoyer (**VOGEL**, 1975).

On peut penser que ces variations **clonales** résultent de l'entretien par multiplication végétative d'un certain état de fonctionnement (ou différenciation) du génome et qu'elles ne mettent pas en cause l'intégrité du patrimoine génétique ainsi conservé. Certaines réversions (ou différenciations, ou rajeunissements) à force de bouturage ou en repassant par des

multiplications végétatives in vitro en témoignent en partie. Cependant les données les plus récentes de la génétique et les études moléculaires montrent qu'au cours de la différenciation somatique de nombreuses réorganisations de l'ADN nucléaire et des populations d'ADN d'organites font que des cellules de lignées somatiques ne sont pas rigoureusement identiques par leur génome aux cellules aboutissant à la méiose. Dans ce cas les clones obtenus par voie asexuée ne conservent pas l'identité des sélections génétiques habituellement véhiculées par graines. L'ampleur et l'incidence de ces différenciations génomiques au cours du développement ne sont pas encore précisément connues, ce sont des résultats de recherches en cours dont il faudra savoir soigneusement tirer les conséquences pour l'appréciation de la validité des conservations par multiplication végétative des plantes dont ce n'est pas la voie habituellement utilisée pour la multiplication et l'usage agronomique.

## D. LA MULTIPLICATION VÉGÉTATIVE IN VITRO COMME TECHNIQUE DE CONSERVATION DES RESSOURCES GÉNÉTIQUES

Comme nous venons de le voir, la multiplication végétative des plantes supérieures est réalisée traditionnellement avec des fragments d'organes végétaux portant un ou plusieurs bourgeons, ces fragments étant placés dans un environnement contrôlé.

Plus récemment, la multiplication végétative a aussi été réalisée par culture in vitro suivant différentes modalités:

- un simple bouturage faisant démarrer des bourgeons existants,
- des régénérations sur cals produits par une bouture ou une culture de fragments de tissus ou d'organes mis en culture,
- des régénérations à partir de cellules isolées ou de protoplastes mis en culture,
- la prolifération de cellules haploïdes par la mise en culture du gamétophyte mâle ou femelle.

Ces modes de multiplication végétative in vitro, très en vogue depuis une vingtaine d'années, ont été appliqués à un grand nombre d'espèces végétales\*. Les techniques utilisées sont diverses et en rapide évolution. Il suffit pour s'en convaincre de relever le nombre de titres bibliographiques traitant de ce sujet d'actualité. Cette méthode est maintenant appliquée à la multiplication végétative industrielle de quelques espèces horticoles bien que demandant un équipement coûteux et une bonne technicité.

Dans le cadre de ce chapitre, nous n'évoquerons que les problèmes généraux posés par la multiplication végétative in vitro pour porter plus spécialement notre attention sur l'intérêt de cette technique dans la conservation des ressources génétiques.

---

\*VASIL et al., 1979 en fait une revue exhaustive.

## 1. Problèmes généraux de la multiplication végétative in vitro

Le développement in vitro d'un méristème excisé, la néoformation de bourgeons ou d'embryons à partir d'un fragment d'organe ou d'un cal, l'organogenèse d'une plante à partir de cellules isolées, de **protoplastes** ou de grains de pollen sont des événements dépendants de conditions fort complexes d'origine interne (génotypes, organes ou tissus concernés, corrélations internes) et externes (milieu de culture, environnement, équilibres des régulateurs). Envisageons quelques-uns de ces facteurs.

a. *La nature des implants.* Les premières cultures de tissu in vitro ont été obtenues en 1939 à partir des tumeurs de l'hybride *Nicotiana glauca* x *N. langsdorfii* (WHITE, 1939) et des tissus **racinaires** de carotte (GAUTHERET, 1939; NOBECOURT, 1939). Un nouvel essor a été donné aux cultures in vitro par l'emploi des méristèmes apicaux. Utilisés dans un premier temps pour guérir des plantes atteintes de maladies à virus (MOREL et MARTIN, 1952 et 1955), ils ont été ensuite repris dans la multiplication végétative in vitro (MOREL, 1964 et 1975). Ce sont des structures dans lesquelles quelques cellules restées dans un état indifférencié gardent la faculté de se diviser. De ce point de vue, les méristèmes s'identifient à des embryons permanents.

Depuis, beaucoup d'autres types tissulaires provenant d'organes variés ont été multipliés in vitro avec des fortunes diverses; l'intensité et la précocité des néoformations dépendent des tissus considérés. Pour BIGOT (1977), il existe des zones-cibles tissulaires, comme les tissus superficiels, qui conservent une grande aptitude au bourgeonnement. NOZERAN et BANCILHON (1972) conseillent de choisir des portions végétatives les plus proches de l'état de la jeune plante issue de graine, soit des zones proches des racines, soit des zones proches des organes reproducteurs.

Quel que soit le tissu mis en culture, il s'agit toujours d'un fragment d'organe qui faisait partie d'un ensemble intégré et dont l'état physiologique et les corrélations internes influent sur les possibilités de stimulation du fragment considéré. Certaines plantes inférieures comme la fougère *Adiantum* régénèrent des plantes à partir des méristèmes mis en culture sur un milieu simple (**knop** additionné de sucre). Par contre la différenciation de plantes nouvelles à partir d'apex de végétaux supérieurs nécessite certains facteurs et certains équilibres d'acide **gibbérélique**, d'ions K<sup>+</sup>, de **cytokinines** et d'auxines.

Les connaissances acquises à ce jour permettent aussi d'envisager la multiplication végétative par la culture de cellules isolées et de **protoplastes**. Soustraites aux contraintes généralement inhibitrices de leur environnement naturel, certaines cellules sont capables, dans certaines conditions artificielles de milieu, de développer un embryon, de structure semblable à celui d'une graine, qui évoluera en une plante conforme à la plante-mère. On assiste à une véritable « régénération » cellulaire conduisant à une embryogenèse somatique. Cette méthode de multiplication in vitro est souvent illustrée par la prolifération des cellules isolées, issues de cals de *Daucus carota* (HALPERIN et WETHERELL, 1964). La production de plantes haploïdes par culture in vitro du gamétophyte mâle ou femelle

des angiospermes relève du même phénomène; toutefois elle ne fait pas partie de notre propos.

*b. Les milieux de culture.* On utilise les milieux de composition connue et déjà employés avec succès pour d'autres plantes. Nous conseillons aux utilisateurs de se reporter aux ouvrages et articles généraux de MOREL (1965), ABBOTT (1978), REINERT et BAJAJ (1978), GAUTHERET (1959), HUSSEY (1978), STOKES (1974), STREET (1976), TURNER (1971), MURASHIGUE et SKOOG (1962), HELLER (1953), MARGARA (1978), WHITE (1963).

Comme les effets des composants sont différents avec les espèces et les souches étudiées, aucun choix n'est possible a priori; il faut essayer des milieux de formules minérales variées avec différents équilibres d'oligoéléments, de vitamines, d'acides aminés et de substances de croissance (auxines, cytokinines).

En outre, il est souvent nécessaire de changer de milieu de culture plusieurs fois pour réaliser un cycle complet de régénération d'une plante: milieu inducteur de la callogenèse ou de la morphogenèse, milieu de croissance des apex... MURASHIGUE (1974) donne une vue synthétique sur les problèmes rencontrés au cours de la régénération de plantes entières par la culture de tissus in vitro.

D'autres facteurs sont à prendre en considération en rapport avec la stimulation des cultures comme les conditions de croissance de la plante-mère (éclairage et nutrition minérale), la vernalisation préalable à la mise en culture des apex des arbes ligneux des régions tempérées (6 mois à 4°C pour le prunus), la composition spectrale et l'énergie de la lumière utilisée dans la salle de culture in vitro...

*c. Quelques résultats significatifs.* La multiplication végétative du chrysanthème in vitro a été prise comme modèle par BIGOT (1977) pour présenter les avantages et les problèmes posés par les différentes méthodes utilisées qui se résument selon le tableau 26.

**TABLEAU 26:** Différentes méthodes de culture in vitro.

Nature de l'implant	Nature des productions in vitro	Taux de multiplication annuel	variation
apex	axillaires préformés	$9 \times 10^6$	conforme
apex	cal	$9 \times 10^{10}$	non conforme
pédoncules floraux	cal	$1,5 \times 10^6$	conforme
fragments de capitule	axillaires préformés	10 - 15	—

Ces résultats permettent d'attirer l'attention sur les taux de multiplication atteints suivant la nature de l'implant et sur la conformité des plantes obtenues.

L'application de la multiplication végétative in vitro concerne beaucoup d'autres espèces horticoles florales (CARLSON, 1975; HUSSEY, 1978).

Les plantes à bulbes et les orchidées exploitées en horticulture sont réputées difficiles à reproduire. Leur multiplication végétative in vitro apporte une solution satisfaisante pour une exploitation industrielle (STOKES, 1974; CHAMPAGNAT, 1977). Par exemple, les apex des bourgeons dormants d'orchidées donnent en culture in vitro des protocormes de régénération qui, par fragmentation, permettent leur multiplication en grand nombre. De même, la culture in vitro de fragments d'écaillés des bulbes de *Lilium* produit un cal qui, par stimulation, donne naissance à des bulbilles, avec un taux de multiplication de  $6 \times 10^{12}$ .

La mise en culture in vitro des fragments de tiges de vigne est parfaitement au point (GALZY, 1969; FAVRE, 1973; NOZERAN et BANCILHON, 1972). Souvent le premier implant s'enracine difficilement; par contre, les portions d'axes caulinaires à une feuille produits par les implants primaires donnent très facilement des racines. Ce simple bouturage en tube a été réussi même pour des souches de vigne considérées comme récalcitrantes au bouturage traditionnel; il a suffi dans ce cas de rechercher la température optimale. Le taux de multiplication annuel atteint par la vigne est de  $12^6$  boutures. Les jeunes plantes obtenues par culture in vitro possèdent les mêmes caractéristiques que les plantes issues de graines (phyllotaxie rayonnante, absence de vrilles).

Les agrumes comme beaucoup d'espèces fruitières et forestières sont propagés par bouturage et greffage. Leur multiplication in vitro quoique relativement plus difficile à réaliser, donne maintenant des résultats encourageants. Le bouturage des citrus mis au point par BOUZID et LASRAM (1971) consiste à cultiver in vitro des fragments d'entre-nœuds de jeunes tiges d'arbres âgés de 2 ans au moins donnant un cal porteur de bourgeons néo-formés; ceux-ci donnent des axes qui s'enracinent très facilement. La multiplication in vitro des orangers est aussi réalisée par la culture de cellules séparées (BUTTON et BOTHA, 1975): à partir d'ovules non fécondés on obtient des cals qui sont dissociés par une enzyme pectolytique; les cellules isolées et étalées sur un milieu approprié donnent naissance directement à de nombreux embryons.

On peut se faire une idée des travaux concernant la multiplication végétative in vitro des plantes ligneuses fruitières forestières et industrielles dans les mises au point récentes de ABBOTT (1978), BONGA (1977), PIERIK (1975), SECKINGER et Mc COWN (1978), TSAI-YING CHENG (1978).

## 2. Problèmes relatifs à la conservation des ressources génétiques

L'intérêt de la multiplication végétative in vitro comparativement aux autres techniques appliquées à la conservation des ressources génétiques est à considérer du point de vue de la durée et de la capacité de stockage, des pertes et des modifications génétiques dues à la culture in vitro, de la sécurité et du coût de cette technique, etc...

a. *Capacités de stockage et de multiplication.* La culture in vitro de méristèmes, de tissus et de cellules produit une nouvelle plante de petite taille, à faible développement, contenue dans un tube à essai. Ainsi est-il possible de rassembler sur quelques mètres carrés de paillasse plusieurs milliers de génotypes. Il suffit de comparer aux surfaces de terrain nécessaires pour entretenir le même nombre de souches, surtout s'il s'agit de plantes arbustives, pour apprécier la capacité de stockage offerte par la culture in vitro. Par exemple, GALZY (1969) a maintenu une collection de 800 cultivars de vigne sur 2 m<sup>2</sup> de paillasse, pendant 15 ans, à raison d'un repiquage annuel de boutures d'axes caulinaires portant une feuille.

La capacité de stockage des collections in vitro dépend aussi de la durée de survie des plantes et des cals en tube. La croissance d'abord rapide, décroît jusqu'à s'annuler avec l'épuisement du milieu de culture. Le maintien de la plante en vie ralentie est ainsi possible pendant un temps plus ou moins important. On admet en général que la viabilité des souches est conservée en pratiquant 1 ou 2 repiquages par an.

Quelques rares résultats expérimentaux permettent d'envisager l'allongement de la survie d'une culture et, par voie de conséquence, la réduction du nombre de transferts nécessaires. Par exemple, le développement des cultures de tissus de chrysanthème a été bloqué à 6°C pendant 6 mois avant de reprendre la multiplication; le potentiel morphogène de la souche a été conservé pendant 4 années (EARLE et LANGHANS, 1974). Les premiers essais de conservation de souches cellulaires végétales dans l'azote liquide ont été un succès: après 10 mois à -196°C pour la carotte, il subsistait 70% de cellules vivantes qui retrouvaient un rythme normal de division après quelques repiquages (WITHERS et STREET, 1975; BAJAJ, 1976).

Les facteurs favorables à la conservation de cellules vivantes dans l'azote liquide se résument à (HENSHAW, 1975): un rythme de refroidissement lent (2°C par minute), des cellules en phase de multiplication exponentielle, l'emploi d'un agent protecteur du froid (5% de glycérol ou 10% de diméthylsulfoxyde), un retour rapide à la température ambiante. Un appareil de congélation programmé a été mis au point par HENSHAW (1975). Les résultats de la congélation de cultures cellulaires c'est-à-dire le pourcentage de cellules viables après traitement varient avec les espèces étudiées: *Acer pseudoplatanus* (HENSHAW, 1975), soja, lin. L'estimation de cette viabilité cellulaire peut être établie par coloration vitale.

b. *La stabilité génétique des cultures de tissu in vitro.* La reproduction des plantes par la multiplication végétative est conforme si le clone issu de culture in vitro possède le même stock héréditaire que la plante-mère et le même type de développement intégré que la plante issue de graine (NOZERAN et BANCILHON, 1972). On connaît beaucoup d'exemples de modifications persistantes conservées lors de la multiplication végétative bien que les plantes portent le même stock génétique. Citons le cas de la différenciation orthotrope et plagiotrope des axes caulinaires de *Phyllanthus* (BANCILHON et al., 1963), de *Coffea arabica* (CARVALHO et al., 1950), du cacaoyer (CHARRIER, 1969)... Ces variants stabilisés et multipliés par voie végétative sont souvent exploités en horticulture.

Après ce rappel des cas de modification persistante du fonctionnement de l'information héréditaire, envisageons la stabilité génétique des cultures in vitro au niveau nucléaire.

Chez la plupart des espèces végétales étudiées, on sait que le degré de ploïdie des différents tissus et cellules différenciées varie dans de larges proportions: c'est le phénomène d'endopolyploïdie. De ce fait, le niveau de ploïdie des explants d'un individu donné peut varier. De plus, il est courant d'observer aussi une ou plusieurs réplifications de l'ADN au cours de la première phase de développement du cal. Seules les lignées cellulaires méristématiques restent diploïdes.

A cette exception près, peu d'espèces végétales conservent en culture in vitro leur niveau diploïde au cours de la différenciation cellulaire. On peut citer le cas de *Crepis capillaris* dont le cal de tissu foliaire est resté au niveau diploïde pendant une année; puis la polyploïdisation s'est installée avec une fréquence croissante au cours du temps ( D'AMOTO, 1975).

L'instabilité chromosomique constatée ne se limite pas à des variations du stock complet des chromosomes (nombre euploïde). Elle porte aussi sur les nombres chromosomiques à raison d'un ou plusieurs chromosomes (nombre aneuploïde). Cette situation a été bien étudiée chez *Nicotiana tabacum* par MURASHIGUE et NAKANO (1967).

Pour cette espèce, les explants d'extrémités de tige à  $2n : 4x = 48$  chromosomes donnent naissance à des cals présentant les nombres chromosomiques indiqués dans le tableau 27.

**TABLEAU 27:** Evolution du nombre chromosomique des cellules d'un cal.

Durée de culture (années)	Nombre de mitoses dénombrées	Nombre de cellules du cal			
		à 48	à 96	à 192	aneuploïde
0	53	25	28	—	—
1	15	—	4	3	8
6	12	—	—	—	12

Ce tableau indique clairement l'accroissement du degré et de la fréquence des polyploïdes et des aneuploïdes avec l'âge de la culture in vitro. D'autres exemples de même nature ont été développés chez les hybrides de *Saccharum*.

De même, des changements chromosomiques structuraux, comme les translocations, sont à noter dans les cultures in vitro. Dans les cals de *Crepis capillaris*, SACRISTAN (1971) a décrit plusieurs types de modifications. L'induction et le maintien de ces différentes manifestations de l'instabilité chromosomique dans les cellules en mitose sont souvent liés à la composition du milieu de culture. Par exemple, la culture sur le milieu de STRIGEMURA d'un fragment de pericycle diploïde des racines de pois conduit à la prolifération de cellules diploïdes. La simple adjonction de kinétine à ce milieu entraîne la prolifération de cellules tétraploïdes. Cette situation pourrait donc être exploitée comme moyen de sélection artificielle de lignées cellulaires possédant un stock chromosomique défini.

C'est le même principe de sélection des types cellulaires sur un milieu de culture de composition choisie qui est appliqué aux problèmes de phyto-

pathologie (AUBERTIE et al., 1976). C'est ainsi que BEHNKE (1979) a obtenu des plantes résistantes par régénération de cals de jeunes feuilles de *Solanum tuberosum* cultivés sur le milieu de LINSMAIER et SKOOG modifié et additionné de filtrats des cultures de 4 pathotypes de *Phytophthora infestans*.

Outre la nature de l'implant initial et la composition du milieu de culture, on constate que la fréquence de cellules aneuploïdes et polyploïdes s'accroît avec l'âge de la culture in vitro. D'ailleurs, la perte de l'aptitude à l'organogénèse au cours de l'entretien des cultures est en partie liée à l'accroissement du degré de ploïdie.

Malgré l'existence de cellules polyploïdes et aneuploïdes en culture in vitro on constate que les plantes régénérées peuvent être dans certains cas toutes diploïdes (exemple: carotte, blé, riz) et posséder dans d'autres cas des nombres chromosomiques variés (tabac, chou, canne à sucre). On peut d'ailleurs rencontrer ces deux comportements pour une même espèce. Ainsi pour deux génotypes différents de *Nicotiana tabacum* placés dans les mêmes conditions de culture in vitro, l'un donne exclusivement des plantes néoformées diploïdes, l'autre des plantes néoformées polyploïdes.

c. *Sécurité et protection phytosanitaire.* Les techniques de culture in vitro offrent des solutions intéressantes aux problèmes de protection phytosanitaire en général, et à ceux qui se posent dans le cas des ressources génétiques en particulier.

Rappelons que c'est grâce aux cultures de méristèmes in vitro que MOREL et MARTIN ont recréé des souches sans virus de dahlia en 1952 et de pomme de terre en 1955 à partir de clones virosés. Depuis, cette méthode a été largement appliquée à de nombreuses espèces ainsi débarrassées des viroses accumulées au cours de la multiplication végétative classique (MARTIN, 1977): oeillet, fraisier, framboisier, lys, iris, orchidée, chrysanthème, canne à sucre, bananier, caladium, colocasia, manioc, etc... Cette solution élégante a pu être associée de façon efficace aux traitements par la thermothérapie de certaines viroses (GALZY, 1969) et des maladies à mycoplasmes (MARCHOUX et al., 1970).

De plus, pendant la culture in vitro en conditions stériles, les problèmes de contamination par les maladies et les parasites sont écartés. La perte de souches par des attaques parasitaires sont tout à fait minimes.

La multiplication végétative in vitro procure donc une grande sécurité du point de vue phytosanitaire et permet d'envisager maintenant le transfert de matériel végétatif qui était généralement prohibé des échanges internationaux.

### 3. Conclusion

La technique des cultures de tissus in vitro paraît bien adaptée à la conservation des ressources génétiques. Les principaux avantages sont:

- des besoins en surface de laboratoire très réduits par rapport aux cultures en champs,
- une méthode de maintien simple par repiquage quand le milieu de culture est épuisé,
- le faible risque de perte des génotypes stockés en laboratoire par rapport aux aléas rencontrés en culture,

- l'absence de contamination par l'insecte, les champignons, les bactéries et les virus,
- la possibilité de multiplier **végétativement** à un grand nombre d'exemplaires les souches stockées.

Malgré les progrès énormes des dix dernières années, il ne faut pourtant pas éluder les difficultés rencontrées dans la culture in vitro. Pour toute nouvelle espèce, l'expérimentation des conditions de culture et le choix de l'implant sont à reprendre. Même cette étape franchie, on sait que les conditions d'organogénèse in vitro diffèrent aussi avec les souches. De ce fait, il est techniquement fort difficile de maîtriser la multiplication in vitro de l'ensemble des espèces et génotypes représentant un complexe d'espèces en collection. En outre, il faut être assuré de la stabilité génétique des cultures in vitro dans le cadre de la conservation des ressources génétiques. C'est le cas de la culture de méristèmes et de **microboutures**. MOREL (1975) propose d'utiliser le terme « **mériclones** » pour désigner les clones dérivés de la culture in vitro des méristèmes. Par contre, les cultures in vitro réalisées avec les autres tissus ou cellules non **méristématiques** conduisent à une grande variation de leur contenu nucléaire (**polyploïdie**, **aneuploïdie**, aberrations chromosomiques variées). Cette importante variabilité génétique des cultures est **certe** un réservoir de variation pour la recherche et l'amélioration, mais elle est incompatible avec la stabilité génétique recherchée pour conserver les collections.

En conclusion, la conservation la plus aisée des ressources génétiques pour le long terme reste le stockage des graines. Mais pour les plantes qui ne peuvent être propagées aisément par graines pour des raisons variées, la multiplication végétative in vitro des méristèmes et des **microboutures** est une technique d'avenir qui offre l'avantage d'une grande sécurité sur le plan phytosanitaire.

## E. LES TRANSFERTS DE RESSOURCES GÉNÉTIQUES ET LA MISE EN QUARANTAINE

### 1. Généralités

Les programmes de conservation de ressources génétiques et d'amélioration des plantes cultivées comportent toujours des échanges de matériel végétal collecté dans la nature ou provenant d'autres stations de recherches.

Toute introduction dans un pays de matériel vivant entraîne un risque de transfert et de dissémination de parasites et ravageurs susceptibles, dans certaines conditions, de créer une épidémie aux conséquences économiques graves. Rappelons l'apparition du *Phylloxera* en Europe due à l'introduction des vignes américaines vers 1860 ou de la **bactériose** en Afrique Centrale consécutive au transfert du manioc brésilien vers 1970, ou encore la rouille américaine du maïs due à *Puccinia polysora* introduite vers 1945 en Afrique. Le risque encouru par le pays d'accueil ne se limite pas à l'introduction primaire du parasite par une voie connue ou fortuite. Il est aussi dans l'introduction de nouvelles races de parasites entraînant déjà des dommages aux cultures.

On distingue habituellement 4 catégories de matériel végétal en regard du contrôle exercé vis-à-vis des parasites et ravageurs susceptibles d'être hébergés par le matériel transféré.

— Le matériel végétal prohibé: le risque d'introduire une maladie d'un pays étranger est si évident que l'importation est interdite, même aux services officiels. C'est surtout le cas des pays qui ne disposent pas de stations d'isolement efficaces pour protéger leurs principales productions agricoles. Le transit du matériel végétal par un pays tiers d'un autre continent où cette culture n'est pas pratiquée, est alors la seule condition satisfaisante.

— Le matériel végétal introduit en station de quarantaine: le pays d'accueil tient à se protéger des maladies sévissant dans le pays d'origine du matériel. Cette admission en quarantaine est le plus souvent réservée aux services publics. Les stations de quarantaine sont pourvues d'infrastructures d'isolement et d'un personnel spécialisé. Nous reviendrons plus en détail sur ces différents aspects de la quarantaine.

— Le matériel végétal soumis à une surveillance phytosanitaire: quand le matériel végétal provient d'une région où les maladies et parasites constituant un **risque** n'existent pas, la protection du pays d'accueil se limite à une inspection du matériel introduit et à une surveillance des premières cultures par les services de la protection des cultures. Les services publics et les institutions privées peuvent procéder à de telles introductions.

— Le matériel végétal soumis à une entrée contrôlée: toutes les espèces non concernées par les 3 catégories précédentes peuvent être importées sans restriction après une inspection et un traitement éventuel du matériel à son entrée. Ce matériel est accessible au grand public et peut entrer dans le circuit commercial.

Dans le cadre de la constitution de centres de ressources génétiques, du matériel végétal est couramment transféré de son lieu de collecte vers un pays différent pour mise en collection et étude. Les stations de quarantaine permettent donc de se protéger du risque d'introduire de nouveaux parasites ou de nouvelles souches de parasites. Ce risque apparaît d'autant plus important qu'il peut s'agir de matériel végétal d'origine spontanée ou adventice mal connu, provenant de régions et d'écosystèmes peu explorés. Dans de nombreux cas, les régions prospectées correspondent aussi aux aires d'origine probables des parasites spécifiques et recèlent de ce fait une diversité extrême des génotypes **pathogènes**\*

Considérons maintenant l'organisation et le fonctionnement des services de quarantaine.

## 2. Organisation et fonctionnement de la quarantaine

a. *Bases légales.* La réglementation des échanges internationaux de plantes repose sur la convention internationale pour la protection des végétaux approuvée à la sixième conférence de la FAO par 44 pays adhérents du monde entier (Rome 1951). Souvent, elle est complétée par

---

\*Cette situation accroît certes le risque d'introduction de nouvelles races du parasite mais elle permet aussi de collecter un matériel végétal intéressant car il a déjà subi la pression de sélection de ce pathogène (Voir chapitre I).

des accords intergouvernementaux au niveau régional et par les décrets du Ministère de l'Agriculture de chaque Etat.

Les services officiels de la protection des végétaux de chaque Etat sont en mesure de communiquer aux intéressés les conditions et la réglementation régissant l'introduction de matériel végétal. Dans le cas de collectes de plantes sauvages dans un but scientifique, les dérogations aux restrictions ou interdictions d'importation peuvent être accordées dans la mesure où le matériel demeure en quarantaine sous le strict contrôle de spécialistes et en dehors des circuits commerciaux.

*b. Localisation des quarantaines.* Idéalement, ces stations devraient se situer hors des zones de culture des espèces concernées afin d'éviter tout risque de contamination. Par exemple, pour les plantes d'origine tropicale, il peut être envisagé d'installer la quarantaine dans un pays tempéré ou méditerranéen en dehors de toute culture de la plante. Les conditions d'environnement ont peu de chance de permettre le développement des parasites pour le cas peu probable où ils seraient accidentellement propagés. De même, dans ce nouveau milieu, les espèces tropicales ne peuvent se développer normalement qu'en serre.

Dans certains pays ou groupes de pays, les services de quarantaine sont centralisés en un même lieu comme la station d'introduction de l'USDA dans le Maryland ou celle de Muguga au Kenya qui fonctionnait pour le groupe régional Ouganda-Tanzanie-Kenya ou celle d'Antananarivo à Madagascar... De telles stations de quarantaine accueillent des espèces variées nécessitant des conditions de cultures très différentes. Il faut donc disposer d'infrastructures importantes et de personnel très polyvalent.

Ces inconvénients sont limités par la création de petites unités d'isolement décentralisées, avec une spécialisation par groupe de plantes, comme en Grande-Bretagne. Dans ce cadre, le matériel végétal peut même être confié directement aux Instituts spécialisés.

*c. Principes généraux.* Dans tous les cas, les principes fondamentaux de la surveillance en quarantaine demeurent les mêmes: isolement total, surveillance permanente, maintien et multiplication du matériel dans des conditions optimales de milieu. La réalisation pratique et les infrastructures nécessaires dépendent pour l'essentiel du type de plantes (encombrement, durée du cycle), de leurs exigences climatiques, de la nature du matériel végétal introduit (graines, jeunes plants, matériel végétatif), de leurs parasites et ravageurs spécifiques.

Dans le pays d'origine, la collecte du matériel doit être faite sur des individus sains, avec observation de l'état sanitaire d'ensemble. Ce matériel végétal, après désinfection superficielle, est expédié en container fermé, accompagné d'un certificat phytosanitaire. Dans le pays d'accueil, le permis d'importation du matériel aura été obtenu préalablement au transfert et les services de quarantaine auront prévu et se seront dotés des moyens matériels nécessaires à la réception immédiate du colis. De tels échanges internationaux demandent en effet une coordination parfaite des actions et des services de sorte que le matériel végétal vivant ne risque pas de perdre sa capacité à régénérer de nouvelles plantes. On se reportera à l'exemple du transfert des ressources génétiques du genre *Coffea* (Cf. Tome I, chap. II).

La serre de quarantaine doit présenter le maximum de garanties dans sa conception pour éviter tout échange avec l'extérieur. D'une manière générale, la serre doit être étanche aux insectes, alimentée par de l'air conditionné, filtré, et conçue de telle sorte que les eaux d'arrosage ne soient pas évacuées à l'extérieur. L'accès à la serre ne doit pas être direct, mais se faire par l'intermédiaire d'un sas ou d'une pièce indépendante dans laquelle peuvent être disposés des bacs pour la stérilisation de l'outillage et un incinérateur. A ces conditions propres à la quarantaine, doivent s'ajouter les caractéristiques techniques permettant la régulation des facteurs d'environnement: température, hygrométrie, éclairage selon les besoins du type de plantes récoltées.

Un personnel hautement spécialisé dans les domaines de la phytopathologie, la virologie, la **nématologie**, la bactériologie et la génétique assure le contrôle de la quarantaine. Il doit être secondé par un important personnel technique assurant les observations de routine, la multiplication et la culture du matériel végétal. A ce stade on ne peut pas se permettre de perdre des échantillons souvent représentés par un nombre restreint de plantes du fait des surfaces limitées de serre d'isolement. Il est aussi nécessaire que les services de quarantaine possèdent des laboratoires de recherche ou soient reliés à de tels laboratoires orientés vers l'étude des agents des maladies.

Le rôle de ces spécialistes n'est pas seulement de diagnostiquer les maladies mais d'examiner les problèmes particuliers présentés par un matériel nouveau, de suggérer tout ce qui doit être fait pour déterminer le plus précisément possible les risques potentiels. Ils doivent s'assurer en outre que toutes les précautions nécessaires sont prises pour éviter l'évasion de germes pathogènes. Une bonne pratique culturale est tout aussi importante pour assurer le développement convenable des plantes pendant toute la durée de l'isolement. Cette condition est primordiale pour la sauvegarde et la multiplication d'un matériel précieux, mais aussi pour l'établissement de diagnostics sûrs. Si les plantes souffrent de mauvaises conditions physiologiques, la mise en évidence de maladies d'origine parasitaire peut devenir aléatoire.

d. *Surveillance phytosanitaire.* Les règles de surveillance en quarantaine sont fonction de la nature du matériel végétal introduit d'une part, des maladies et parasites concernés d'autre part.

L'introduction de graines est toujours préférée à celle de matériel végétatif. Les semences sont facilement manipulées, leur désinfection externe étant relativement aisée. Cependant, il faut redouter les parasites internes transmis par la semence — champignon, virus —. Seule l'observation constante du semis à la reproduction, pendant 1 ou 2 cycles, permettra de mettre en évidence l'existence de ces parasites internes.

L'introduction de jeunes plantes pose le problème des parasites de racines qu'une stérilisation superficielle n'élimine pas comme les nématodes, les champignons du système vasculaire. Lorsqu'elle est possible, la greffe sur des porte-greffes sélectionnés au préalable permet alors de supprimer et de **détruire** les systèmes **racinaires**. On manipulera essentiellement des axes végétatifs à bouturer ou à greffer après désinfection superficielle. L'observation permettra ensuite de déceler les infections d'origine interne et les infections latentes si la durée de la quarantaine est

suffisamment longue. Le matériel est conservé pendant cette période sous forme de clones individualisés. La même méthodologie est à utiliser pour les plantes à tubercules, à racines, à bulbes.

Parmi tous les parasites et ravageurs susceptibles d'être hébergés par le matériel végétal, les plus difficiles à détecter sont sans doute les virus. La simple observation est le plus souvent insuffisante car l'apparition de symptômes de maladie virale n'est pas une indication suffisante de la présence de virus. Aussi il est nécessaire de procéder à des tests de transmission et d'inoculation sur des plantes témoins pour déterminer la présence de pathogènes d'origine virale.

La durée de la quarantaine est bien entendu fonction du type de plante et de sa vitesse de croissance. Dans tous les cas elle doit être suffisamment longue pour permettre la multiplication du matériel originel et si possible pour assurer une surveillance sur la totalité du cycle végétatif de la plante.

Enfin, tout individu montrant des symptômes quelconques d'infection doit être analysé afin d'effectuer un diagnostic précis de l'agent infectieux avant traitement approprié ou destruction par incinération si aucun traitement n'est possible, la sauvegarde du matériel étant un objectif prioritaire.

### 3. Conclusion

D'un point de vue phytosanitaire, l'introduction de matériel végétal sous forme de graines est considérée comme la méthode idéale. Nous avons vu qu'elle n'est pas sans risques, mais, surtout, elle n'est pas universelle. En effet, si l'on se place du point de vue des ressources génétiques, la multiplication végétative est dans certains cas le seul moyen de transférer le matériel végétal. Il en est ainsi de nombre d'espèces cultivées intéressantes multipliées par voie végétative mais aussi des espèces se reproduisant par graines quand on tient à reprendre le même génotype ou à collecter des plantes stériles au cours de la prospection de formes sauvages. Il faut donc considérer que le transfert de matériel végétatif est absolument nécessaire au déroulement des programmes des centres de ressources génétiques.

Il est maintenant aisé de combattre la tendance originelle des stations de quarantaine qui consistait à proscrire impérativement le matériel végétatif jugé plus dangereux que les graines, essentiellement à cause des viroses accumulées et transmises au cours de la multiplication végétative. En effet, les progrès techniques réalisés au cours des 10 dernières années dans le domaine de la multiplication végétative *in vitro* en font un outil de choix dans la protection phytosanitaire des échanges. Rappelons les avantages de cette technique développée précédemment: multiplication en milieu stérile à l'abri des parasites; récupération des génotypes atteints par une virose ou toute affection interne par la régénération de nouvelles plantes à partir d'une cellule ou d'une portion d'organes indemne — les méristèmes et les extrémités d'axes végétatifs; stockage sous un très faible volume d'un grand nombre d'individus; pouvoir multiplicateur de la méthode très important. Il est donc temps de changer d'attitude vis-à-vis du matériel végétal transféré sous sa forme végétative et de doter les stations de quarantaine des moyens techniques modernes nécessaires à la multiplication végétative *in vitro*.

Il semble enfin nécessaire de se demander si les systèmes de protection phytosanitaire mis en oeuvre ne sont pas d'une efficacité limitée du fait de l'accroissement des échanges internationaux du monde moderne? Répondre à cette évolution par un renforcement du protectionnisme et des réglementations ne paraît pas être une attitude constructive. Une réflexion sur ce thème doit considérer les conditions de développement des épidémies catastrophiques. Celles-ci sont déclenchées en présence du parasite par la conjonction de facteurs écologiques favorables et d'un matériel végétal réceptif peu diversifié. De ce point de vue, l'agriculture moderne et industrielle accroît le risque de telles épidémies par la culture sur de grandes surfaces de quelques variétés à base génétique restreinte. Cette situation favorise la propagation d'une nouvelle maladie ou la sélection de nouvelles races du parasite. Est-il besoin de rappeler l'explosion de l'*helminthosporiose* du maïs aux USA en 1970 où toutes les variétés possèdent le même cytoplasme «Texas», ou encore le développement spectaculaire de la rouille orangée du caféier en Amérique depuis 1970 sur des cultivars de *C. arabica* dont la pauvreté génétique est bien connue.

De telles épidémies constituent un drame économique et social lors de leur apparition. Mais il est clair qu'il résulte pour une part de l'imprévoyance et de l'inconséquence des services de l'agriculture et de la recherche. Nous sommes en mesure d'apporter une solution partielle à ce problème par une meilleure exploitation de la diversité génétique du matériel végétal. On sait en particulier qu'il existe un équilibre entre l'hôte et ses parasites stricts dans les zones d'origine et de diversification. La collecte de la diversité génétique et la conservation de ces ressources génétiques apportent donc à moyen terme une réponse adaptée à l'accroissement de la dissémination des parasites du fait du développement des déplacements et des échanges internationaux. La stricte observation des règles de protection sanitaires lors des introductions du matériel végétal ne doit pas pour autant être abandonnée.



CHAPITRE V  
LES BASES DE DONNÉES  
ET LEUR EXPLOITATION  
STATISTIQUE

E. Nguyen Van et J. Pernès



L'étude de l'organisation des bases de données des ressources génétiques renvoie à tous les aspects de la gestion des banques de gènes. Ce chapitre pourrait être un abrégé de l'ensemble du manuel. C'est plutôt le *vade mecum* des connaissances biologiques nécessaires aux informaticiens désireux d'adapter les bases de données générales aux problèmes particuliers propres aux ressources génétiques; ce n'est surtout pas un document informatique car son objectif principal est de permettre au lecteur du manuel et aux spécialistes des ressources génétiques de bien saisir en quoi et comment l'ordinateur et ses servitudes sont indispensables à acquérir et à accepter. Cette compréhension devrait assurer une bonne sérénité face aux labeurs répétitifs, fastidieux du recueil des données, sérénité nécessaire à l'obtention d'informations rigoureuses et **répétables**, qui vaillent d'être interprétées et transmises!

# I. PRINCIPES

## A. L'IMPORTANCE DES EFFECTIFS

Les collections de conservation des ressources génétiques concernent des dizaines de milliers d'échantillons (ou entrées) pour chaque complexe d'espèces. Lorsque ces collections sont dispersées sur le terrain d'expérimentation soit pour l'évaluation soit pour la multiplication ou le rajeunissement des semences ce sont des dizaines d'individus par entrée qui sont observés ou récoltés. Les collections de riz de l'IRRI affichent 70.000 entrées conservées, le CYMMIT possède plus de 15.000 populations de maïs en chambres froides, les collectes de mil réalisées en Afrique par les équipes ORSTOM ont réuni plus de 5.000 variétés-populations (cf. vol. I).

Pour chaque entrée plusieurs dizaines de caractéristiques doivent être notées ou mesurées, puis répertoriées. Cet ensemble de données doit être enregistré et stocké avec fiabilité, de façon utilisable et aisément communicable. Des informations nouvelles ou des modifications doivent pouvoir être apportées sans remettre en cause l'ensemble des informations acquises.

Ainsi, avec 100 caractères observés pour chacune des 70.000 entrées c'est 7 millions de données qu'il faut stocker dans des fichiers, étiqueter (ou numéroter) correctement de façon que chacune puisse être connue (et éventuellement modifiée si des renseignements nouveaux doivent être consignés), et rendue accessible dans des délais compatibles avec une utilisation fréquente et diverse. Les nombres deviennent encore plus impressionnants lorsqu'on sait que l'on recherche simultanément des informations sur plusieurs caractères. Cette « combinatoire » conduit très vite à des opérations tellement nombreuses (bien que chacune très simple: rechercher une donnée dans un fichier constitué de tableaux à double entrées bien indexés et coordonnés) que même les ordinateurs les plus puissants imposeraient des délais de recherche impraticables. Face à cette situation il faut le concours de deux stratégies:

— D'une part des recherches en informatique mettent au point des méthodes et des programmes de gestion des bases de données les meilleurs possibles compte-tenu des systèmes informatiques (ordinateurs disponibles) et de la nature des fichiers à gérer. Nous n'analyserons pas ici les différentes approches retenues par les informaticiens sauf si elles imposent une présentation particulière des données (fichiers en réseaux hiérarchisés ou non, en liste inverse, systèmes relationnels, etc...).

— D'autre part, et c'est surtout ce que nous présenterons dans ce chapitre, il faut que le gestionnaire des ressources génétiques organise ses priorités parmi les informations dont il a besoin, qu'il choisisse certaines utilisations privilégiées pour que l'organisation informatique des fichiers leur soit adaptée, qu'il accepte certains compromis entre les coûts de stockage et de recherche de l'information, les risques de perte ou de dégradation (coût des systèmes de protection) et le coût de la mise à la disposition des données à des utilisateurs non informaticiens (langages les plus immédiatement clairs).

La clarté maximale nécessaire pour placer correctement les problèmes informatiques doit donc être acquise sur:

- la nature des données,
- leur utilisation,
- les types d'utilisateurs de ces données.

La nature des données sera étudiée de façon détaillée dans le paragraphe «les listes de descripteurs», qu'il suffise ici de noter qu'il s'agit pour chaque échantillon (variété, population, lignée, clone, etc...) de la collection de l'ensemble des valeurs prises par les différents caractères observés (accompagné d'informations sur les conditions de l'observation) et de renseignements concernant l'origine de l'échantillon lui-même (numérotations, provenance, **pédigree**, etc...).

## B. UTILISATIONS DES DONNÉES

### 1. Simple recensement

Il est bon de posséder pour chaque entrée de la collection une fiche signalétique la décrivant le plus complètement possible. Cette utilisation ne conduit qu'au deux exigences suivantes:

- un stockage des fichiers peu encombrant, économique et sûr,
- une consultation rapide.

### 2. Gestion rationnelle de la collection

Le problème des nombres initialement posé est encore plus dramatique lorsqu'il s'agit de la conservation des échantillons eux-mêmes (coût des chambres froides, risque d'erreur et modification des structures génétiques dus aux renouvellements — rajeunissement des semences par exemple). Aussi cherche-t-on à réduire le plus possible le nombre d'échantillons à conserver sans perdre la diversité génétique, c'est-à-dire à faire la chasse à toutes les duplications qui conduisent à stocker sous des numéros différents ou des noms différents la même variété. Cette identification des duplications sera efficace et saine si le nombre d'observations effectué est important (liste longue des descripteurs) et si l'outil informatique permet d'analyser correctement le fichier pour regrouper toutes les entrées identiques pour l'état des caractères enregistrés dans la liste des descripteurs.

Mais une bonne gestion des collections demandera souvent plus d'audace et de pertinence que le simple repérage des duplications. Quand il s'agit de gérer des populations de plantes **allogames** annuelles il faut constituer des populations réservoirs pas très nombreuses et d'effectifs importants (cf. chapitre conservation). Ces populations réservoirs doivent présenter une certaine homogénéité qui permet d'assurer la pérennité d'un type **variétal** bien défini avec des caractéristiques **phénologiques** données tout en protégeant un polymorphisme sous-jacent important, ce que les cultures traditionnelles ont su nous transmettre à travers tous les cultivars analogues (cf. par exemple les cultivars du mil, vol. I). Les méthodes statistiques et informatiques doivent permettre la constitution de ces populations réservoirs dès qu'une analyse génétique sérieuse des échantillons a été systématiquement réalisée (évaluation des collections en champ, et études biochimiques fines).

L'analyse des états des caractères consignés dans les listes de descripteurs de chaque entrée conduira à des classifications de degrés de finesse variés à la fois pour obtenir une vue d'ensemble de l'organisation du complexe et établir des stratégies de regroupement pour les conservations dynamiques de populations ou la protection de sites naturels typiques.

### 3. Recherches de caractéristiques utiles

*a. Recherche directe.* L'organisation de bases de données devra permettre de donner des informations telles que (exemple pour le riz): quelles sont toutes les variétés naines à grains rouges et longs résistantes au *Pericaria oryzae* disponibles dans la collection. La possibilité d'obtenir rapidement et économiquement des réponses à un tel questionnaire est très attractive pour le questionneur. Cette recherche directe a été en général l'objectif prioritaire des conceptions de bases de données de ressources génétiques, elle sera d'autant plus économique et efficace que la collaboration entre les sélectionneurs et le responsable de la base de données aura permis de dégager parmi la liste des descripteurs les caractères dont la recherche immédiate est la plus fréquente et la plus importante.

*b. Recherche par un processus d'évaluation hiérarchisée.* Nous avons vu (cf. chapitre évaluation génétique) que les caractéristiques agronomiques les plus importantes (par exemple les délais de floraison, la résistance à une (ou des) *race(s) locale(s)* d'un parasite) ne peuvent être nullement obtenues à partir d'évaluations centralisées acquises dans un nombre restreint de situations écologiques. Chaque sélectionneur, dans la zone et en fonction des besoins pour lesquels il travaille, doit organiser ses propres évaluations agronomiques. Celles-ci ne sont compatibles avec ses moyens de recherche que si, à chaque cycle de culture, elles ne concernent qu'un nombre limité d'entrées (disons de l'ordre des centaines pour une céréale). Il faut donc que le sélectionneur soit guidé dans son choix pour établir les échantillons qu'il introduira dans sa collection d'évaluation. Une tâche prioritaire des gestionnaires des banques de gènes centralisées est donc de fournir non pas un catalogue peu structuré mais des descriptions aussi synthétiques que possible des collections (sous forme de classifications, distances génétiques, organisation écologique et génétique des grands groupes *variétaux*) pour que l'agronome puisse organiser rationnellement son échantillonnage et apprécier progressivement la composition de la banque de gènes. Le dernier paragraphe présentera le type de travail taxonomique que la base de données doit rendre possible, cela suppose que la liste des descripteurs comprenne beaucoup de caractères dont l'analyse génétique a déterminé la clarté héréditaire (faible ambiguïté de lecture phénotypique du génotype, marqueurs précis du génome).

*c. Recherches dans le cadre d'un réseau décentralisé d'évaluation.* Nous venons de voir cette primauté de l'évaluation agronomique locale; cette observation à deux conséquences: 1. des banques de données locales doivent être facilement constituées et posséder un grand degré d'autonomie dans la lecture, l'entrée et la modification des informations; 2. le profil agronomique d'une variété (d'un cultivar) naîtra de la réunion des informations décrivant ses comportements dans des conditions écologiques ou

agronomiques variées, il faut pouvoir développer facilement les échanges d'information, ou de bases de données, complète entre « banques locales ». La constitution de tels réseaux conduit à une informatique développant les petits systèmes (micro informatique) avec des stockages de données très aisément échangeables et lisibles (on trouvera avec les bases G.D.M. une illustration de ce type d'organisation). Plus que d'un principe informatique lié aux dimensions des ordinateurs utilisables, il s'agit d'une conception de l'organisation des banques de gènes: *banques de gènes centrales* aux gros moyens de stockages des échantillons et de traitement de données contrôlant intégralement tous les aspects des ressources génétiques ou *réseaux de conservations et d'évaluation décentralisés* des ressources génétiques, le bureau central dans un tel réseau veillant seulement à ce que les responsabilités déléguées localement soient correctement assumées (cf. chapitre organisation des centres de ressources génétiques). En fait ces deux types de conceptions ne sont pas exclusifs l'un de l'autre et doivent être complémentaires et compatibles et donc par voie de conséquences les bases de données centrales et locales doivent être *informatiquement* compatibles également.

## 4. Evolution des collections

a. *Décisions pour la gestion des collections.* Deux types d'informations doivent être très aisément accessibles et régulièrement affichés:

— L'un concerne la gestion des stocks, et ce sont des indications très simples concernant les quantités et semences conservées à court, moyen et long terme, leur âge et leur taux de germination. L'évolution de ces quantités permet d'établir les dates des actions de rajeunissement (« *réjuvenation* » c'est-à-dire des mises en culture pour *refabriquer* des graines de meilleur taux de germination ou en quantité plus importante, ou pour des conservations en culture in vitro et remises de plants en conditions et cultures normales avant de *refabriquer* des clones). Les problèmes informatiques ne concernent que des enregistrements et des lectures périodiques et des étiquetages d'avertissements.

— L'autre concerne la création des populations de conservation, le regroupement des éléments dupliqués, le repérage de l'identité génétique approximative de toutes les descendances issues par séries d'autofécondations de mêmes lignées pures initiales. La réalisation de ces groupes d'identité génétique approximative repose sur la conjonction des renseignements ayant conduit aux calculs de distances génétiques entre les diverses entrées, comme nous avons vu précédemment, et des enregistrements généalogiques consignants toute l'histoire des multiplications (autofécondations, croisements frère x *sœur*, endogamie, etc...). Ces enregistrements généalogiques posent des problèmes aux informaticiens qui doivent connecter des fichiers concernant des individus (parents) et des familles (leurs descendants) compte-tenu que certains individus, mais pas tous, se retrouvent en tant que parents d'autres familles qu'ils ont produites. Il faut pouvoir sans ambiguïté vérifier les caractéristiques des individus-parents de façon à garantir que l'enchaînement généalogique des campagnes de multiplication n'a pas entraîné de changements génétiques profonds dus à des erreurs d'étiquetage, des mélanges, des pollinisations illégitimes, des mutations importantes, etc... Cette tâche évidente est loin d'être organisée

de façon efficace et sûre dans les banques de gènes à responsabilité mondiale les plus importantes!

b. *Appréciation de l'évolution génétique des collections.* Chaque cycle de multiplication de graines est l'occasion de transformation des polymorphismes des populations conservées. Toutes les variétés traditionnelles, que le mode de reproduction soit **allogame**, **autogame** ou **apomictique** sont polymorphes. Les multiplications modifient la structure reproductive de la population (endogamie avec des effectifs différents, **consaguinité** systématiquement contrôlée), les pressions de sélection (conditions de milieu généralement très soigneusement contrôlées, choix différents des géniteurs) sont différentes. Le gestionnaire de la collection doit pouvoir enregistrer et consulter toutes les informations qui de cycle en cycle de multiplications permettent de prévoir théoriquement l'ampleur des transformations des populations conservées. Les évaluations génétiques successives de ces mêmes populations permettront d'apprécier la réalité des dérives et les éventuelles anomalies par rapport aux dérives attendues théoriquement acceptables. C'est une tâche importante des bases de données de ressources génétiques que d'être accompagnées de programmes d'analyses statistiques nécessaires pour apprécier lucidement les situations de conservation. Des études de polymorphisme enzymatique de populations de maïs ou de mil conservés en endogamie nous ont montré que la raréfaction des polymorphismes était beaucoup plus intense qu'on ne s'y était attendu!

La préparation des données et la définition des caractères des listes de descripteurs en termes de population sont ainsi indispensables. La plupart des listes normalisées actuellement publiées n'entrent pas dans le détail des paramètres propres à la génétique des populations, or ce sont des populations que nous lèguent les collecteurs de variétés traditionnelles ou d'écotypes spontanés, ce sont des populations et non des individus qui possèdent les propriétés d'adaptation, c'est toute la richesse des populations que les conservations doivent préserver. Bien évidemment les bases de données accompagnant les conservations dynamiques in situ devraient comporter tout cet arsenal de paramètres propres aux populations.

## **C. LES UTILISATEURS DES DONNÉES**

De nombreux problèmes informatiques à résoudre ont pour origine la pratique des utilisateurs eux-mêmes.

### **1. Nature et forme de l'information**

L'économie et la simplicité apparente plaident en faveur d'un codage systématique de la plupart des données. Dans certains cas, et après consultations soigneuses des spécialistes du complexe d'espèces étudié, certains caractères peuvent être transmis ou exprimés sous forme de codes universellement acceptables. Dans de nombreux cas cependant,

malgré des définitions apparemment précises des conditions d'observations, des désaccords pratiques finiront par mettre en cause le sens même des caractères codés. D'autre part la consultation des fichiers exigera souvent une clarté immédiate de l'information. Beaucoup d'observations importantes ne pourront être consignées que sous formes de commentaires. La valeur de ces commentaires est souvent grande, particulièrement aux yeux des spécialistes mais leur coût d'enregistrement et de stockage est élevé, leur consultation systématique et leur exploitation globale ne sont pas simples. Les mesures directes n'ont de sens que si les conditions d'observation et les données statistiques générales de l'expérimentation les accompagnent ainsi que des références à des témoins stables accessibles à tous. Il faut organiser les espaces disponibles dans les fichiers en sachant que pour une entrée donnée beaucoup d'états des caractères de la liste proposée n'auront pu être décrits. L'absence d'information est très fréquente et ne doit pas être source d'erreur. Il ne faut pas pousser l'observateur à rapporter une information erronée pour ne pas laisser un blanc! D'autant plus que certains logiciels de bases de données organisent la gestion des «blancs» dans les fichiers pour diminuer le coût de stockage. C'est la nécessité d'accéder à beaucoup d'informations à la fois en peu de temps et surtout le fait d'obtenir des informations partielles et d'avoir à réajuster en permanence les fichiers qui obligent à certain type d'organisation et de gestion des données, et donc à un certain type de support, en l'occurrence des disques magnétiques (accès direct à l'information, rapidité d'accès). Par contre l'échange d'information peut et doit se faire sur un support compatible à la fois avec la banque centrale et les banques locales d'où l'utilisation de disquettes, cartes perforées, etc...

En dehors des corrections qu'il faut pouvoir réaliser partiellement et sans risque, l'organisation des bases de données de ressources génétiques doit être conçue pour permettre la plus grande facilité de mise à jour des listes de caractères. Les outils de l'évaluation évoluent sans cesse: des tests de sensibilité *in vitro* sur des races bien définies de parasites remplaceront les notes de résistance approximatives établies sur une collection cultivée en un milieu plus ou moins infesté; les méthodes biochimiques apporteront des lectures nouvelles (caractéristiques génétiques, critères d'appréciation de sensibilité à la photopériode), des propriétés nouvelles importantes seront indispensables à consigner (compositions nutritives, relations symbiotiques avec des fixateurs d'azote atmosphérique, etc...). Les spécialistes trouveront de nouveaux accords pour mieux définir un caractère dont l'appréciation était initialement par trop empirique. D'autres notations considérées d'abord d'un grand intérêt se verront reléguées au second plan.

Des pages entières d'information doivent pouvoir être supprimées, modifiées, ajoutées sans que le bel **ordonnement** ou l'exploitation de la base de données soit à reprendre.

Cependant il ne faut pas oublier l'objectif de gestion des données réelles et ne pas tenter de suivre les progrès de l'informatique qui sans cesse créent de nouveaux algorithmes au risque de n'en utiliser aucun car le suivant sera plus performant! Il faut savoir accepter un système de base de données fonctionnel et que l'on exploite régulièrement en sachant que d'autres seront ultérieurement mieux agencés. Il vaut mieux une base de données ouverte aux progrès biologiques mais un peu désuète dans son économie et son fonctionnement qu'une future base **informatiquement**

supérieure mais dont l'organisation remettra en cause tout l'enregistrement du travail biologique déjà effectué!

## 2. Protections, duplications, distribution de l'information

Ce sous-titre en lui-même suffit au rappel élémentaire évident des propriétés de toute base de données. La distribution totale de l'information n'est plus réalisable sous forme d'impressions de « listing » et il faut échapper à l'impérialisme d'une banque centrale seule capable d'accéder à l'ensemble des données. Il faut organiser un découpage de l'information de façon que des blocs entiers puissent être distribués et que ces blocs aient un sens relativement à l'ensemble des données parce qu'ils seront situés par rapport à des ensembles interprétés (positions dans des généalogies, caractéristique d'un groupe situé dans une ou des classifications). Ces consignes de protections, duplications, distributions ont aussi une incidence sur le choix des matériaux supports de l'information (disques, bandes, disquettes, fiches perforées, etc...).

## II. LES LISTES DE DESCRIPTEURS

Toutes les informations relatives à un échantillon de la collection seront consignées sur une fiche dont nous allons étudier les différentes rubriques. La liste de ces rubriques et le détail de chaque définition des informations **consignables** constitue la liste des descripteurs. Pour un échantillon donné chaque descripteur de la liste est représenté par un état particulier, que cet état soit une valeur d'un code, un numéro, une mesure directe ou un commentaire ou l'indication que l'information manque.

### Structure génétique de l'échantillon

Le terme d'échantillon recouvre des réalités biologiques variées, ce peut être un individu, un clone, un lot de graines représentatif d'une variété, d'une lignée pure, d'une population, ou encore une population définie dans un écosystème donné. Pour un même matériel biologique, par exemple des caféiers, les rubriques de la liste des descripteurs peuvent être différentes suivant la nature de l'échantillon: dans un verger où chaque arbre est identifié et suivi les notations concernent précisément l'individu; dans une population naturelle (où tous les arbres ne sont pas repérés, où la population en tant que telle survit malgré la disparition de certains arbres, l'apparition aux stades adultes de plantes qui n'était jusque là pas en production), des paramètres très statistiques tenteront d'apprécier globalement la population (âge moyen, hétérogénéité, disposition, densité, dates moyennes de floraison et sa dispersion, distance approximative des populations voisines, etc...); derrière des intitulés comparables des descripteurs ce ne seront pas exactement les mêmes notations qui seront réalisées pour un individu ou une population. Ces remarques soulignent cependant une

difficulté, ou une nécessité de clarification, car les collections de conservation contiennent des entités biologiques très variées. Une collection de mil contient des « lignées » (qui ont été extraites de populations définies ou non, avec une histoire de leur obtention précisée ou non), des populations spontanées, des variétés traditionnelles, des formes hybrides intermédiaires entre les variétés cultivées et les formes sauvages, des populations réservoir créées spécialement pour la conservation et dont divers sous-ensemble ont été distribués et entretenus dans des écologies différentes... De nombreux descripteurs seront communs à ces différentes catégories mais d'autres demanderont des lectures particulières propres à chaque type d'échantillon. Ainsi ne faudrait-il pas parler ou établir des listes de descripteurs du mil mais des listes propres aux lignées de mil stabilisées et d'autres listes propres aux cultivars, etc...

Après ces remarques soulignant l'importance de la structure génétique de l'échantillon qui sera décrit étudions les diverses rubriques que cette liste des descripteurs tente de clarifier et de formaliser.

## A. CARACTÉRISATION, NUMÉROTATION DE L'ÉCHANTILLON\*

Ces numérotations doivent être aussi complètes que possible, certaines doivent être des clés d'accès direct (voir ce mot dans le paragraphe gestion informatique des données) particulièrement le numéro d'inscription dans le catalogue de base; il faut pouvoir reconstituer le mieux possible l'historique de l'échantillon avant son entrée dans la banque centrale, ces éléments étant des moyens importants de vérification de la conformité de l'échantillon avec la structure génétique originale qu'il représente. Les autres numérotations d'enregistrements doivent être agencées pour que l'évolution génétique des ressources génétiques conservées puisse être correctement décrite et analysable: toutes les informations concernant les processus de multiplication et de rajeunissement des semences (« **rejuvenation** »), toutes les connexions entre les fichiers — individus et les fichiers — familles doivent être assurées pour que de proche en proche le **pédigree** complet de tout individu puisse être extrait automatiquement (sur demande) depuis le premier échantillon, introduit dans la banque, dont il dérive. (voir annexe).

---

\*Les remarques contenues dans tous les paragraphes suivants ne constituent qu'un accompagnement succinct des extraits de la liste de descripteurs du mil de l'IBPGR que nous avons pris comme exemple. La liste complète comprend plus de 100 rubriques non comprises les rubriques définissant les conditions de l'observation.

## B. INFORMATIONS ACQUISES LORS DES COLLECTES

Ces données sont souvent difficilement **codifiables** mais il ne faut pas hésiter à consacrer une part importante des places dans les fichiers pour rapporter des commentaires entiers. Ce sont des observations uniques qu'il ne faut pas risquer d'appauvrir. Nous donnons à la fois l'intitulé des rubriques et la présentation des fiches de **préenregistrement**. (voir annexe).

## C. INFORMATIONS GÉNÉTIQUES, BASES D'ÉLABORATION DE TAXONOMIES, OBSERVATIONS QUALITATIVES À HÉRÉDITÉ BIEN PRÉCISÉE, CARACTÉRISATIONS MORPHOLOGIQUES TRÈS **RÉPÉTABLES** DANS DES MILIEUX VARIÉS

Nous avons vu dans le chapitre consacré aux méthodes d'évaluation comment des informations doivent faciliter les choix des échantillons destinés aux évaluations agronomiques locales faites par les observateurs. Les enregistrements sont le plus souvent réalisés sous forme de codes arbitraires non ordonnés, les stades de développement des plantes au moment de l'observation sont rigoureusement précisés. (voir annexe).

La liste des descripteurs qui nous sert d'exemple manque actuellement de deux types d'informations:

- les rubriques détaillées (gène à gène) concernant les données enzymatiques de l'analyse par électrophorèse,
- une appréciation précise de l'hétérogénéité des populations mises en observation (sauf, comme nous le verrons pour le paragraphe suivant évaluation agronomique, en ce qui concerne la floraison).

Cette appréciation précise peut être obtenue en introduisant:

- pour les caractères mesurés, sans déterminisme génétique simple ou élucidé, les notations importantes: variances résiduelles pour l'ensemble de la collection dans l'expérimentation d'évaluation et variance propre à l'échantillon (ou un code définissant son degré de supériorité par rapport aux variances résiduelles environnementales de référence),
- pour les caractères à déterminisme génétique lisible, les indices classiques de la génétique des populations: fréquences géniques, indices de diversité pour résumer une série de polymorphismes, taux **d'hétérozygotie** moyens, etc... (cf. chapitre I).

## D. INFORMATIONS DE L'ÉVALUATION AGRONOMIQUE

Les exemples tirés de la liste des descripteurs du mil de l'IBPGR sont assez clairs (voir annexe). Remarquons (en prolongement des **commen-**

taires du paragraphe précédent) que les données statistiques du plan d'expérience effectivement réalisé manquent, d'où la faiblesse des informations concernant les données numériques observées sur l'échantillon (telles le tallage, la précocité, les hauteurs, les longueurs, les largeurs, les poids et volume des grains, etc...).

Les descripteurs des sensibilités aux parasites sont de maniements très délicats soit parce qu'aucune inoculation contrôlée n'est encore mise au point (il faut se référer alors à des témoins de sensibilité qui permettront d'évaluer l'importance de l'infestation naturelle au moment de l'expérimentation), soit parce que les inoculations contrôlées (quand elles sont possibles à grande échelle dans un milieu hors zone de culture) ne représentent encore que très partiellement les conditions d'une infestation réelle (population de races où d'éléments polymorphes pour l'agressivité, situations écologiques particulières, multiplicité d'attaques). Les sensibilités aux attaques d'insectes sont encore plus difficiles à noter (voir annexe).

Ces extraits démontrent assez nettement comment une évaluation agronomique initiale est illusoire et ces descripteurs doivent surtout servir de point de repère pour chaque utilisateur réalisant son propre schéma d'évaluation agronomique locale.

L'utilisation de ces notations est aussi très difficile du fait de l'hétérogénéité des enregistrements: notations qualitatives, codes ordonnés, mesures. Certaines données tant quantitatives que qualitatives doivent être cependant intégrées dans les études taxonomiques sinon les groupements que constituent les classifications auront un niveau d'abstraction par trop décourageant.

### **III. GESTION INFORMATIQUE DES DONNÉES**

L'organisation du stockage des données dans le domaine des ressources génétiques pose de nombreux problèmes du fait de la très grande diversité des informations à collecter, et aussi de la difficulté à les réutiliser. Suivant le type de plantes, le mode de conservation, de multiplication, les caractères observables seront différents. Le choix du codage des caractères dépend de l'utilisation que l'on voudra en faire, ce qui n'est pas toujours possible de définir parfaitement au moment de la création du système informatique. Pourtant pour une meilleure utilisation des informations collectées il est indispensable d'avoir défini au départ ce que l'on attend d'un tel système parce que toute sa logique va dépendre du but que l'on veut atteindre. Ce n'est que lorsque ces buts seront parfaitement déterminés que l'on pourra espérer réutiliser au mieux les informations acquises. Actuellement il n'y a pas de système satisfaisant pour la gestion des données en ressources génétiques, ni de mode de stockage uniforme.

## A. DONNÉES DE BASE

### 1. Problèmes généraux

Parmi les problèmes que pose la création d'une banque de données en ressources génétiques, certains ont été largement étudiés et résolus. D'autres sont par contre à peine abordés ou traités de façon incomplète du fait de leur complexité et surtout parce que les banques de données ont été conçues comme des systèmes de stockage d'informations mais pas du tout en vue de leur utilisation dynamique.

a. *Nature des données.* 3 types d'informations sur les plantes collectées ou en collections sont nécessaires pour exploiter les données. Le sous-chapitre précédent a commenté ce propos et l'exemple concret de la liste des descripteurs du mil n'est que repris plus systématiquement ici.

— Un système d'indexation pour le repérage des plantes. Ce peut être un numéro d'accès qui doit être unique permettant le repérage immédiat de l'individu recherché, ou un code donnant aussi un certain nombre d'informations par exemple: type de plante, année de récolte ou d'obtention, ordre d'obtention.

— Les caractéristiques et propriétés des plantes indexées, appelées généralement *descripteurs*. Ils sont de deux sortes: les descripteurs de terrain et ceux résultant d'études ultérieures.

Les descripteurs de terrain sont les plus faciles à définir mais pas forcément les plus faciles à collecter. Ce sont par exemple:

Les références	de la prospection, l'année de prospection, le nom des prospecteurs, le ou les pays prospectés.
Les noms	d'espèce, de sous-espèce, vernaculaire,
La localisation	le nom du village, la distance par rapport à une ville sur un grand axe, l'ethnie cultivant cette plante (si c'est une plante cultivée).
La position géographique	la longitude, la latitude, l'altitude.
La topographie	montagne, plaine, colline, etc...
Mode de prospection	en grenier, au marché, en champ...
Mode de prélèvement	en grain, en bouture...
Phénologie	stade du développement des plantes à la récolte.

Etat sanitaire présence de maladies au moment de la récolte.

Lieu du stockage

Ceci n'est pas une liste exhaustive mais un exemple.

*Les descripteurs «Résultat d'études ultérieures».* Ils peuvent être très divers en fonction de la plante étudiée, de ses caractéristiques et des problèmes qu'elle pose. Il est très difficile de définir à priori une liste de caractères à observer, certains peuvent être communs à tous les types de plantes, par exemple: le nombre chromosomique, les résultats d'analyse enzymatique, la précocité; d'autres ne pourront être que spécifiques à un type de plantes donné, par exemple: étude du système d'incompatibilité, étude de la sensibilité à la longueur du jour, mode de multiplication.

En plus de la difficulté à définir les descripteurs à noter, se pose le problème de leur codage. Il y a eu plusieurs essais de standardisation, par exemple: à la conférence d'Izmir en Turquie en avril 1972, KONZAK (USA) a proposé un modèle sur les descripteurs de terrain. L'IRRI donne un autre exemple pour les descripteurs de type agronomique pour le riz, nous venons de voir celui couvrant le mil, l'IBPGR utilise un ensemble de descripteurs très complet pour la pomme de terre comportant des descripteurs de type terrain, morphologique, taxonomique, agronomique mais ils restent très spécifiques de la plante observée.

*Les relations de parenté ou les processus de multiplications.* Lorsque les plantes entrent dans un processus de multiplication ou de croisement, il faut repérer les liens de parenté existant entre les plantes par un système d'indexation qui différencie les croisements des autres modes de multiplication (végétative, culture de protoplastes, d'anthers...) et qui permettent de reconstituer les généalogies.

*b. Volume des données.* Le volume des données, dans le cas de la constitution d'une banque centrale de ressources génétiques, ne peut que s'accroître continuellement. A chaque retour de prospection et de collecte ou après chaque introduction il y aura de nouvelles informations à traiter, de même qu'après les résultats d'évaluation. La nature des supports physiques dépend de leur fréquence d'utilisation. Pour les données peu utilisées il faut les stocker sur des bandes magnétiques (en vue d'archivage), pour les autres le stockage doit être de type permanent à accès direct (disques). Différents modèles de fichiers sont à constituer en fonction du type de plantes recensées et des problèmes à traiter. Il faut prévoir le découpage des fichiers principaux en sous-fichiers en vue d'échange d'informations entre la banque centrale et les banques locales.

*c. Programmation: Utilisation des informations.* Trois problèmes se posent pour réutiliser les informations contenues dans la base de données.

— *Recherche des éléments de la collection répondant aux propriétés recherchées.* C'est la fonction la plus généralement réalisée dans les différents centres de ressources génétiques à l'aide de TAXIR/EXIR par exemple (cf. description plus loin) ou par des programmes écrits dans ces centres. Exemple: rechercher toutes les plantes possédant un caractère donné.

— *Reconstitution des généalogies.* Lorsque la conservation des ressources génétiques est dynamique il est indispensable de mesurer l'évolution du matériel à travers les différents procédés de multiplication. Cette

reconstitution est assez facile à programmer lorsqu'il s'agit de la généalogie acquise par l'enchaînement d'autofécondations.

Le problème le plus difficile à résoudre résulte de l'utilisation de modes de conservation qui échappent aux autofécondations ou hybridations simples, soit parce qu'elles sont impossibles (ex: auto-incompatibilité), soit parce qu'elles sont destructrices des structures génétiques (cas des variétés traditionnelles de plantes), soit parce qu'il ne s'agit plus à proprement parler de généalogie (ex.: **androgenèse**, multiplication végétative).

— L'exploitation *des données*. Il s'agit généralement d'analyse des distances et d'établissement de classification. L'étude des variabilités se fait par des analyses en composantes principales ou en correspondances, on peut chercher aussi à établir des corrélations entre composants du milieu et les propriétés des plantes. Les données doivent être rendues directement accessibles à ces programmes par l'écriture d'une interface d'extraction compatible avec les bibliothèques statistiques.

## 2. Vocabulaire général

Définissons quelques termes souvent utilisés en gestion de bases de données.

### a. *Vocabulaire de base*

— Base de données: (BD) ensemble des données proprement dites.

— Système de gestion de bases de données (SGBD): ensemble de logiciels permettant de décrire, mémoriser, manipuler, traiter les données (voir CODASYL, ADABAS).

— Banque de données: base de données avec son système de gestion de base de données.

— Logiciels: ensemble de programmes permettant de gérer un système informatique ou ses applications. Il existe 2 sortes de logiciels:

Logiciels de base: ce sont des systèmes d'exploitations, des traducteurs de langage qui servent à gérer la machine (ordinateur),

Logiciels d'applications: ce sont des ensembles de programmes qui permettent la gestion et le traitement des données.

— Programmes d'application: programmes permettant la manipulation *et/ou* le traitement de données.

— Système intégré de gestion: base de données plus SGDB, plus programmes d'application.

### b. *Les fonctions*

— Description des données: mise au point de la structure logique *du/ou* des fichiers de la base.

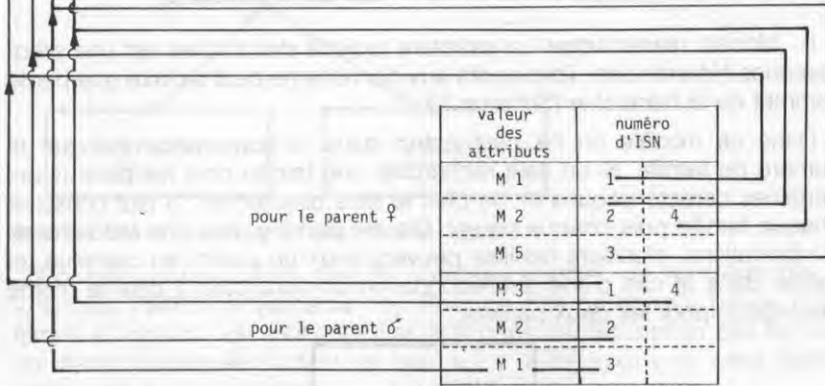
— Mémorisation des données: mise sur support physique des données dans le format et le mode d'accès choisis.

— Manipulation des données: fonction introduisant de nouvelles données, établissant de nouveaux liens, supprimant ou modifiant des données ou des liens.

— Mise à disposition des données: fonction permettant de présenter les données auxquelles on a eu accès sous une forme directement exploitable soit par l'utilisateur (résultat direct d'une consultation d'un fichier) soit en vue d'un traitement.

n° ISN* donné par ADABAS		Parent ♀	Parent ♂	Hauteur	Précocité
1	Famille 1	M 1	M 3	1,80	42
2	Famille 2	M 2	M 2	1,50	50
3	Famille 3	M 5	M 1	2,10	80
4	Famille 4	M 2	M 3	2,20	60

Si on déclare Parent ♀ et Parent ♂ « Descripteur »\*\* (= clé d'accès) alors on a 6 listes inverses de créées (qui doivent être enregistrées en mémoire et sont communicables individuellement et complètement sur simple demande).



**Schéma 14:** Exemple de modèle en liste inverse: ADABAS

\*ISN : Identification logique de l'enregistrement

\*\*Descripteur: a deux sens suivant que l'on est dans le contexte ADABAS ou non:

— sens large: tout attribut d'une entité

— sens strict ADABAS: clé d'accès au fichier

### 3. Modèles de SGBD

(Dans l'ordre historique d'évolution des méthodologies informatiques).

a. *Modèle en listes inverses.* Tous les attributs\* peuvent être définis comme clé\*\* d'accès dans le fichier. Ce système est géré par un très grand nombre d'index, un pour chaque valeur de chaque clé. Chaque valeur permet ainsi d'établir une liste «inverse» contenant tous les numéros des enregistrements qui la possède.

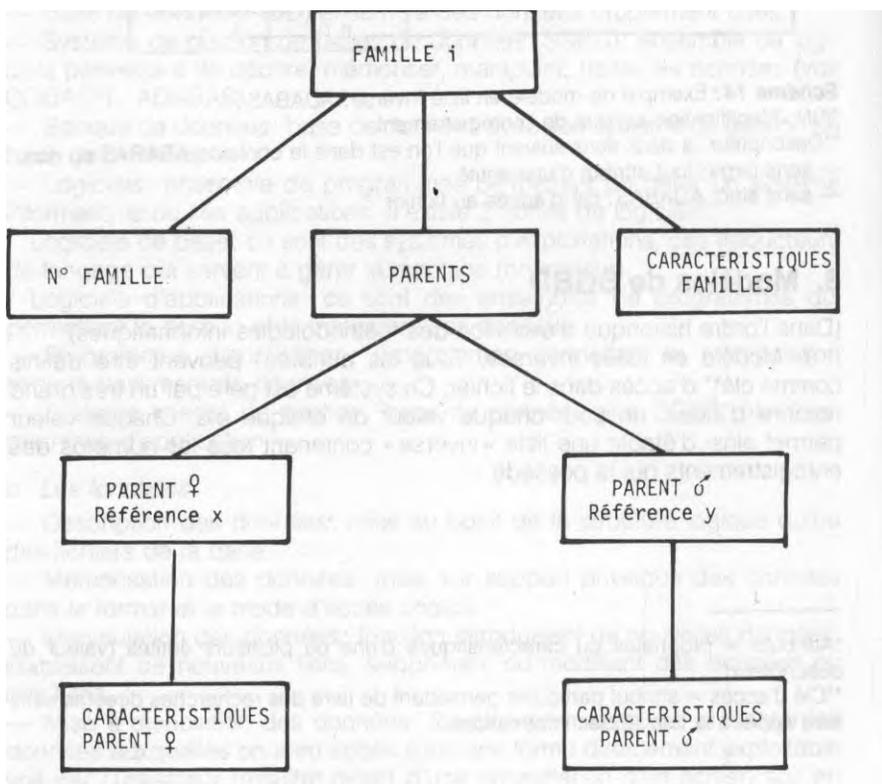
\*Attributs = propriétés ou caractéristiques d'une ou plusieurs entités (valeur du descripteur).

\*\*Clé d'accès = attribut particulier permettant de faire des recherches directes sans faire appel à la totalité des informations.

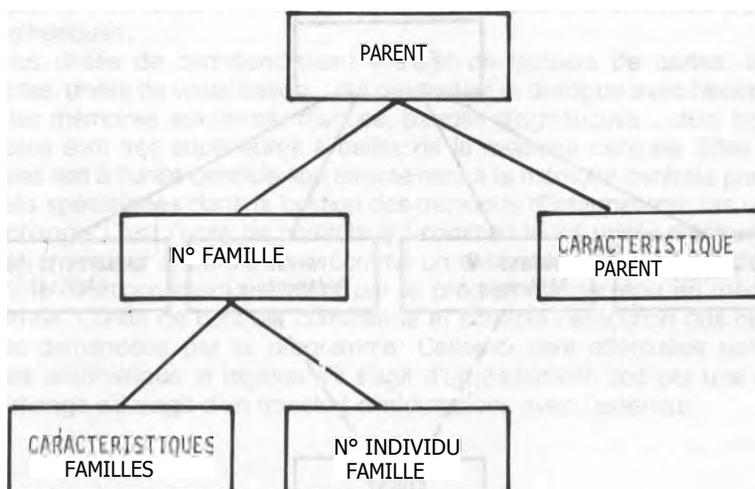
En fait la gestion des index est beaucoup plus complexe, le schéma 14 est donné pour faciliter la compréhension. ADABAS (Adaptable Data Base) est un système de gestion de base de données basé sur les listes inverses utilisant 2 techniques pour réduire la place mémoire occupée et compenser celle occupée par les listes inverses. 1) les descripteurs n'occupent pas une place fixe (ils sont délimités par des séparateurs), les descripteurs vides n'occupent pas de place. 2) les données sont comprimées (suppression des zéros et des blancs non significatifs) d'où un gain de place pouvant aller jusqu'à 50%. Il possède deux autres avantages: rapidité d'exécution, possibilité d'adapter le logiciel à son problème. On peut accéder à n'importe quelle information sans hiérarchie entre elles.

b. *Modèle hiérarchique.* La structure logique des fichiers est une arborescence hiérarchisée. Tout accès aux données ne peut se faire que par le sommet de la hiérarchie (Schéma 15).

Dans ce modèle on ne peut entrer dans l'arborescence que par le numéro de famille. Si on veut rechercher une famille dont les parents ont certaines caractéristiques on ne peut le faire directement. Il faut consulter chaque famille puis chaque parent. D'autre part il y aura une redondance d'informations, plusieurs familles peuvent avoir un parent en commun, et même dans le cas d'une autofécondation on aura côte à côte la même description pour les deux parents.



**Schéma 15:** Modèle hiérarchique: famille



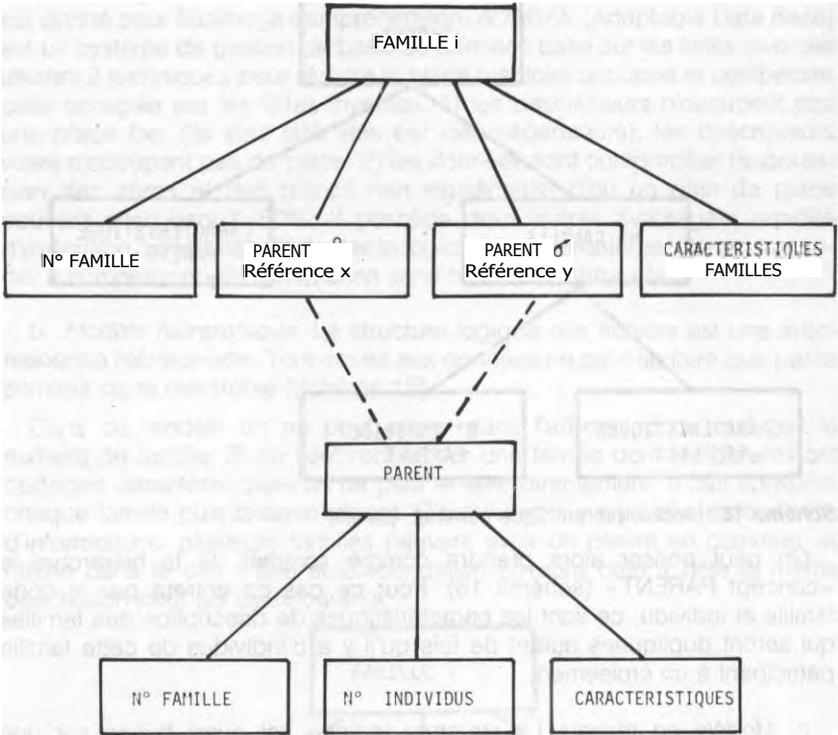
**Schéma 16:** Modèle hiérarchique sommet: parent

On peut penser alors prendre comme sommet de la hiérarchie le « concept PARENT » (schéma 16). Pour ce cas on entrera par le code famille et individu, ce sont les caractéristiques de description des familles qui seront dupliquées autant de fois qu'il y a d'individus de cette famille participant à un croisement.

c. *Modèle en réseau.* La structure logique est aussi basée sur une arborescence, mais on peut créer des liens logiques entre les caractères. De plus l'accès aux données peut se faire de façon ascendante ou descendante. Les modèles en réseau ont été mis au point pour éviter d'une part les duplications des systèmes hiérarchiques et d'autre part pour pouvoir entrer à n'importe quel niveau de la hiérarchie (Schéma 17).

Il n'y a plus duplication des informations pour un parent participant à plusieurs croisements puisque le fichier famille ne contient que l'information qui permet de retrouver les parents et leur description dans un autre fichier. D'autre part on peut faire des recherches directes dans le fichier-parents sans passer nécessairement par le fichier-famille, et retourner du fichier-parent vers le fichier-famille. D'où un gain d'espace (pas de duplications) et de temps de recherche (il n'est pas nécessaire d'entrer par le sommet de la hiérarchie).

d. *Modèle relationnel.* Il a pour objectif d'étudier et de normaliser les transformations de tableaux (calcul relationnel) et repose sur la théorie élémentaire des relations. Ainsi toutes les données seront considérées comme des ensembles de tables qu'il convient de transformer pour en exploiter l'information. *CODASYL* est un système de gestion de bases de données en réseau ayant la particularité d'accepter plusieurs relations logiques pour un même réseau. Ceci rend la programmation très complexe car on doit préciser sur quelle relation logique on travaille. *DML* utilise un langage informatique particulier le Data Manipulation Language. Ce n'est pas un langage en soi, mais une extension du Fortran (langage hôte).



**Schéma 17:** Modèle en réseau

#### 4. Définition succincte d'un ordinateur

Un système informatique (ordinateur) est composé:

- d'une mémoire centrale (contenant programmes et données),
- d'une unité centrale de traitement (exécutant les programmes),
- et plusieurs unités d'entrées-sorties (exécutant les échanges avec l'extérieur).

a. *L'unité centrale.* Elle est composée de l'unité de contrôle et de l'unité arithmétique et logique qui exploitent respectivement les deux types d'informations contenues dans la mémoire centrale, les instructions du programme (informations traitantes) et les données (informations traitées).

— L'unité de contrôle extrait de la mémoire centrale l'instruction suivante à traiter. Elle analyse cette instruction et établit les connexions correspondantes dans l'unité arithmétique et logique. Elle extrait de la mémoire centrale les données sur lesquelles porte l'instruction, déclenche leur traitement dans l'unité arithmétique et logique et éventuellement range les résultats dans la mémoire centrale.

— L'unité arithmétique et logique effectue sur les données qu'elle reçoit les traitements commandés par l'unité de contrôle.

b. *Unités d'échange et unités périphériques.* Il existe deux sortes d'unités périphériques:

— les unités de communication: il s'agit de lecteurs de cartes, imprimantes, unités de visualisation... qui permettent le dialogue avec l'extérieur,  
— les mémoires auxiliaires: disques, bandes magnétiques... dont les capacités sont très supérieures à celles de la mémoire centrale. Elles sont reliées soit à l'unité centrale soit directement à la mémoire centrale par des unités spécialisées dans la gestion des transferts d'informations: les unités d'échange. C'est l'unité de contrôle qui commande les unités d'échange.

Un ordinateur apparaît donc comme un assemblage d'unités distinctes dont le fonctionnement est dicté par le programme contenu en mémoire centrale. L'unité de contrôle commande et contrôle l'exécution des opérations demandées par le programme. Celles-ci sont effectuées soit par l'unité arithmétique et logique s'il s'agit d'un traitement, soit par une unité d'échange s'il s'agit d'un transfert d'informations avec l'extérieur.

## 5. Notion de langage

(Quelques définitions complémentaires)

— Un langage de programmation est un langage utilisé pour décrire un programme à l'ordinateur. Ce terme ne s'applique qu'aux langages autres que le langage machine ou le langage assembleur.

— Langage évolué: langage de programmation proche de la formulation logique ou mathématique. Il doit être traduit dans le langage machine par un programme appelé compilateur. A une instruction du langage correspondent plusieurs instructions-machine.

— Langage assembleur (ou langage symbolique): c'est un langage formé par les instructions propres à un ordinateur, écrites sous forme symbolique, c'est-à-dire facilement lisible par l'homme. A une instruction ou langage assembleur correspond une instruction en langage machine et une seule. Cette transformation se fait grâce à un programme appelé assembleur, avant d'être exécutée en machine.

— Langage machine: langage défini par la liste des instructions propres à un ordinateur et leur représentation interne, sous forme binaire. C'est le langage directement exécutable par l'ordinateur.

— Langage hôte: langage évolué qui sert de support à un autre langage non autonome.

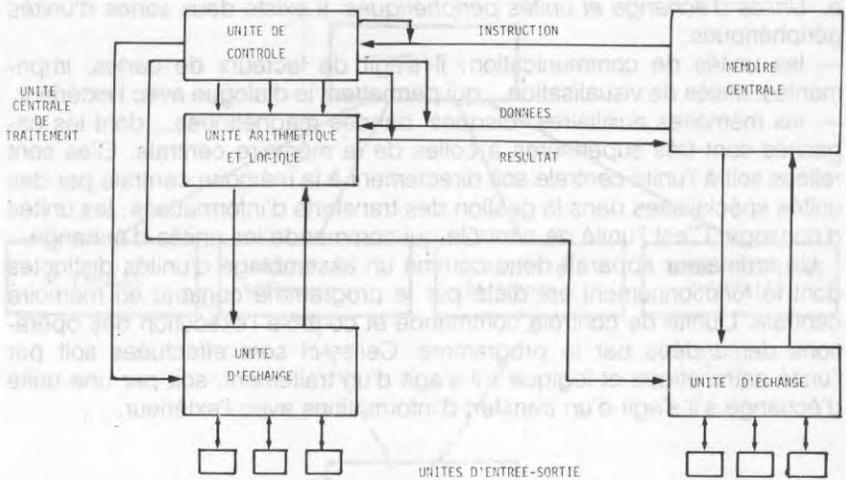
— Compilateur: programme traduisant en langage machine, un programme écrit en langage évolué.

— Assembleur: programme traduisant en langage machine un programme écrit en langage symbolique (ou langage assembleur).

## 6. Structure logique et physique

— Structure logique: inter-relation des données dans une base telles qu'elles sont perçues par les utilisateurs.

— Structure physique: inter-relation des données dans une base telles qu'elles sont effectivement stockées.



**Schéma 18:** Fonctionnement d'un ordinateur

## B. PROBLÈMES À RÉSOUDRE POUR INSTALLER UNE BANQUE DE DONNÉES EN RESSOURCES GÉNÉTIQUES

Pour constituer une banque de données dans le cadre des ressources génétiques quatre points principaux doivent être traités:

- choix des descripteurs et de leurs états possibles,
- types de traitements désirés sur ces données,
- choix du système de gestion de bases de données (SGBD) en fonction de deux critères précédents,
- diversité des utilisateurs (novice ou expert en informatique, isolé ou en groupe, diversité des problèmes due aux plantes différentes).

### 1. Choix du SGBD

Il n'existe pas actuellement de systèmes parfaitement au point et directement transposable pour constituer une banque de données dans le cadre des ressources génétiques. La constitution d'une solide base centrale exige la connexion à un gros ordinateur et l'existence d'une équipe de spécialistes se consacrant aux problèmes de bases de données, ainsi qu'une équipe de statisticiens.

Dans ce cas on peut utiliser par exemple le SGBD (ADABAS) c'est un modèle en listes inverses permettant la création de réseaux (qui en particulier peut être utilisé pour les reconstitutions généalogiques). Il semble assez bien adapté aux problèmes de ressources génétiques. Le système anglais est Codasyl, leur projet d'utilisation d'un modèle relationnel est orienté d'une manière radicalement différente de celui exploité par ADABAS. Les programmes seront absolument inutilisables directement d'un système à

l'autre. Les systèmes TAXIR/EXIR constituent un type de base de ressources génétiques déjà en fonction. Si la connexion à un gros ordinateur n'est pas possible, ou si l'on veut organiser un réseau d'utilisateurs plus ou moins interdépendants, on peut opter pour un SGBD du type GDM.

## 2. Choix des descripteurs

Certaines sources d'informations sont très utiles et même quelquefois directement transposables. C'est le cas du choix des descripteurs et de leur codage (telles les listes du BIRP ou du CSIRO).

## 3. Traitement des données — Programmation

*a. Entrée.* L'entrée des données doit être aussi souple que possible et pourra utiliser les différents supports disponibles: consoles, cartes, rubans, bandes, disques, cassettes...

*b. Stockage.* Le mode de stockage doit correspondre à l'utilisation qu'on veut faire des données. On peut proposer les principes suivants:

— Mode archive: stockage à long terme des données qu'on n'a pas l'intention d'utiliser dans l'immédiat mais dont on prévoit l'utilisation dans un futur lointain: stockage sur bandes (le moins onéreux).

— Stockage permanent: ce sont les fichiers permanents sur disques sur lesquels sont stockés les données d'utilisation fréquentes.

— « Espace travail »: ou espace utilisé au cours du fonctionnement de la base, ce sont les données en exploitation, plus les fichiers temporaires nécessaires aux traitements en cours. Ces fichiers sont détruits en fin d'exploitation.

C'est à l'utilisateur de faire passer ses données d'un type de fichier à un autre suivant l'évolution de son problème.

*c. Validité des données.* (programmes ou ensemble des programmes analysant la qualité des données). Ces programmes qui sont évidemment indispensables n'ont pas encore été établis sur une base générale. Chacun a mis au point sa propre procédure pour les opérations suivantes:

- Détection des données erronées,
- Détection des données redondantes,
- Correction des erreurs.

La mise au point de programmes cohérents de ce genre exige des actions spécifiques.

*d. Procédure d'accès et d'utilisation des données stockées dans la base.* Ces accès peuvent être selon les types suivants:

— *Recherches conditionnelles:* Ce sont les recherches destinées à fournir la liste des numéros d'enregistrements ayant des propriétés définies a priori. Le système doit prévoir des recherches sur une clé simple ou sur des clés multiples avec une relation logique entre elles. Les clés peuvent être numériques, alphabétiques ou les deux à la fois. Seuls les SGBD comportent cette possibilité. Tous les autres modes de gestion de fichiers classiques ne peuvent faire des recherches que sur un seul critère.

— *Recherches séquentielles*: Quand les données sont rangées selon un ordre logique, les recherches séquentielles ont pour but de retrouver les enregistrements entre deux valeurs d'un attribut.

Par exemple: retrouver tous les croisements obtenus entre deux dates données.

— *Recherches itératives*: (cas de la généalogie). Recherche d'une donnée dont un ou plusieurs des attributs va servir de clé d'accès pour la recherche suivante.

Une procédure très importante concerne *l'adaptation des données aux analyses multivariées* (taxonomie numérique décrite en particulier dans le sous-chapitre suivant). Il faut que les données soient le plus directement accessibles aux bibliothèques statistiques. Ceci suppose les opérations suivantes:

— *Transfert des données*: De la banque centrale vers la banque locale ou vice versa à partir des fichiers existants en créant des sous-fichiers.

— *Destruction des données*: en cas d'erreur ou de redondance, sans risque pour les autres données.

— *Couplage des données aux bibliothèques de programmes*: Ces problèmes n'ont pas été résolus de façon directe en général, il entraîne des manipulations de sous-fichiers.

## 4. Diversité des utilisateurs

La plupart des utilisateurs ne seront pas des informaticiens de formation. Certains n'auront aucune connaissance de l'informatique, d'autres quelques notions mais probablement très peu seront des informaticiens confirmés. Il est donc indispensable que l'entrée des données soit simple, de même que l'exploitation de la base par les programmes utilisateurs. Pour cela le gestionnaire de la banque centrale aura un gros problème de simplification des appels à la base.

Un gros effort de vulgarisation du système de gestion adopté et de son mode d'utilisation s'impose pour qu'il soit facilement exploitable par n'importe quel utilisateur, son mode d'emploi devra être très simple et souple pour l'entrée des données. Des séminaires d'information pour vulgariser la base de données et des démonstrations seront à prévoir. Pour le choix des descripteurs et de leur codage il est très important de travailler en relations avec des équipes parfaitement rodées d'un point de vue pratique et scientifique sur les problèmes de collections et d'évaluations.

## C. ÉTUDE DE CERTAINS SYSTÈMES

### 1. Étude d'EXIR

a. *Présentation*. EXIR est un système de gestion de base de données s'approchant des modèles relationnels, basé sur la théorie des ensembles. EXIR range les données sous une forme compactée, le système les retrouve rapidement par calcul. La théorie des ensembles donne un outil simple, puissant et pratique pour manipuler des informations. Un ensemble est une collection d'éléments (ou membres) (par exemple: un genre de 50

espèces est un ensemble de 50 éléments (ou membres)). La structure des éléments d'un ensemble et les ensembles eux-mêmes sont hiérarchisés entre eux. Les premiers travaux d'algèbre logique ont été publiés vers le milieu du XIX<sup>e</sup> siècle par Georges BOOLE. Ces concepts ont été étendus à d'autres contextes et sont connus sous les noms d'algèbre de BOOLE, théorie des ensembles, calcul proportionnel, arithmétique logique, etc... Dans le contexte de la manipulation des données l'algèbre de BOOLE est utilisée pour combiner ou modifier des ensembles, pour en former de nouveaux. Ces manipulations peuvent se faire sur les variables de son propre choix et non pas sur la totalité. C'est une propriété fondamentale d'EXIR. Le stockage des données est basé sur l'appartenance ou non à un ensemble donné. Par exemple si on s'intéresse à la couleur des fleurs on pourra poser la question «cette fleur appartient-elle à la classe des fleurs à couleur rose?» si oui EXIR code 1, si non 0. Ceci pour toute la liste des caractères collectés ou étudiés. Il existe un moyen de réduire le nombre de 0 ou de 1 à stocker en mémoire (forme compactée) mais les informations de n'importe quel membre d'un ensemble pour n'importe quel état de descripteur donné seront toujours restitués sous une forme claire.

Le langage d'interrogation des fichiers est un langage Booléen. Le résultat d'une interrogation donne un sous-ensemble. On peut rechercher des éléments d'un ensemble possédant des caractères donnés, ou au contraire ne les ayant pas, ou bien à la fois sur la présence de certains caractères et sur l'absence d'autres. Ceci est exécuté très facilement et rapidement par la machine car il s'agit en fait d'opération sur des chaînes de bits\*. Par exemple la fonction négation (absence) consiste simplement à transformer tous les zéros en uns et vice versa.

Une nouvelle version (3.0) d'EXIR vient d'être mise au point. C'est celle que nous présenterons mais auparavant nous allons exposer les problèmes que posaient les anciennes versions 2.4 et 2.5 pour illustrer des problèmes pratiques que les informaticiens ont dû résoudre et comment certains sont résolus dans la dernière version 3.0.

b. *Problèmes non résolus par les versions 2.4 et 2.5.* Les versions 2.4. et 2.5 étaient spécifiques d'un ordinateur, respectivement CDC et IBM, c'est la seule différence entre elles. La liste des problèmes s'exprimera par l'enchaînement de leurs définitions.

— Recherche récursive (ou itérative): c'est la recherche d'un élément d'un ensemble dont un (ou plusieurs) attribut doit servir de clé d'accès pour la recherche suivante (cas de la généalogie).

— Création de plusieurs fichiers dans une même base: dans le cas de l'étude de la généalogie il est nécessaire de créer 2 fichiers l'un décrivant «les familles» résultant des croisements, l'autre décrivant les parents ayant participé à ces croisements afin de pouvoir rechercher des familles dont les parents avaient certaines caractéristiques.

— Souplesse d'entrées/sorties, compatibilité avec d'autres systèmes: dans l'ancienne version les entrées étaient assez simples mais les sorties étaient propres au système EXIR et obligeaient à écrire une interface pour

---

\*bit = abréviation de «Binary digit». C'est la plus petite unité d'information d'un ordinateur qui n'a que 2 valeurs possibles 0 ou 1.

les rendre compatibles avec les bibliothèques d'analyses statistiques **multi-variates**.

— « Place mémoire » : ce genre de système exige une place en mémoire très importante pour les raisons suivantes:

- Les programmes eux-mêmes sont très grands, ils occupent donc déjà une grande place mémoire,
- Des dictionnaires décrivant les descripteurs sont créés afin de vérifier toute nouvelle entrée de données,
- Des tables de **codage/décodage** sont créées. En effet, pour que les résultats soient sous forme codées à l'entrée des données il faut que le système enregistre les correspondances entre codes et états de descripteurs. Les mêmes tables sont réutilisées pour la sortie en clair des résultats.
- Enfin, plus les résultats d'une recherche contiennent d'éléments plus ils occupent de place mémoire. Il faut donc faire attention lors d'une interrogation que l'on ne sélectionne pas la totalité d'un fichier ou les 3/4!

La description du système nous dit «La seule limite dans la gestion de vos données est la taille mémoire de votre ordinateur» ce qui veut bien dire qu'il est impossible de l'utiliser sur un système même moyen.

— Compression des données. Place occupée par le stockage: Les versions 2.4 et 2.5 stockaient les données telles qu'on les entrait sans aucune compression en cas d'absence d'un descripteur. Si on réservait 100 caractères pour une zone commentaire, par exemple, et que pour un élément donné il n'y avait pas de commentaire particulier cette zone restait à blanc et on perdait donc 100 caractères. La place occupée par le fichier occupait la place réelle réservée au niveau de sa description.

— Descripteurs à zones multiples variables: pour un caractère donné il peut y avoir, par exemple, un changement entre l'âge juvénile et l'âge adulte (cas de certaines enzymes). Si nous savons que ce changement est rare il est préférable, au lieu de créer deux descripteurs, l'un décrivant l'âge juvénile et l'autre l'âge adulte, de déclarer ce descripteur à «zones multiples variables» c'est-à-dire que, au cas où ce changement existe on aurait un descripteur à 2 états, sinon 1 descripteur à 1 seul état. Ceci est parfaitement réalisé par **ADABAS** et permet de gagner dans certains cas la place de stockage. **EXIR** ne le réalise pas.

— Coûts d'une base de données **EXIR**: les 3 derniers problèmes exposés augmentent considérablement le coût d'une base. L'absence de compression des données fait croître le coût de stockage, le coût d'une manipulation d'une base étant fonction du temps de manipulation et de la place mémoire réservée. Si le système de gestion est important le coût d'utilisation est élevé rien que par l'appel du logiciel. Il est très important d'essayer de diminuer ces coûts en réduisant les fichiers et les logiciels et en rendant ces derniers le plus performant possible pour réduire les temps de manipulations. Malheureusement le système devient alors hermétique à des non informaticiens et l'on sera conduit à accepter un compromis entre performance et accès à tout utilisateur.

*c. Présentation de la nouvelle version 3.0. Exemple d'utilisation. Amélioration qu'elle apporte par rapport à 2.4 et 2.5.* La nouvelle version est unique et ne dépend plus du type d'ordinateur. Elle se compose de 3 parties (selon le diagramme ci-dessous):

Module de chargement  
EXIR:  
Données :

choisit un système (CDC, IBM...)  
créée ou gère la base  
la base elle-même  
dans le système choisi.

Le **LOAD MODULE** (ou module de chargement) permet de sélectionner le système choisi (IBM, CDC...). Une fois la sélection faite il appelle **EXIR** dans la version désirée et celui-ci crée ou gère les données dans ce système.

L'inconvénient de cette nouvelle version est son exigence en taille mémoire: 96 à 135 k\* suivant le système utilisé (il faut savoir que certains **microordinateurs** n'ont que 64 k de mémoire centrale). C'est un problème général: plus on veut faciliter la tâche des utilisateurs plus le logiciel devient grand.

— **Entrées des données**: l'entrée des données est assez simple. Il reste une difficulté au niveau de la définition des données qui est en fait liée à la nécessité de connaître parfaitement au départ l'utilisation que l'on veut faire de la base (si on pense par exemple qu'il y aura des calculs statistiques à exécuter). Une fois que l'on a bien défini les zones de type «NUMÉRIQUE», «**CODE**», ou «**NOM**» l'entrée des données devient très simple. Il existe une série de commandes permettant de créer et de contrôler la création de la base ainsi que d'en faire la mise à jour. L'utilisateur n'a plus à gérer les positions. Il suffit de mettre des virgules après chaque descripteur.

d. *Création et contrôle de la banque.* Décrivons l'enchaînement des opérations prévues.

— **DEFINE DESCRIPTORS\*\***: cette commande informe **EXIR** du nombre, des noms, et du type des descripteurs qui caractérisent les données. Cette définition permet à **EXIR** de contrôler chaque nouvelle entrée de données. Il existe 6 options possibles de descriptions qui concernent 3 catégories de descripteurs:

- ceux dont tous les états possibles sont connus à la création de la base, ou qui sont **manipulables davantage**s sous forme de code que de noms.
- ceux dont les états sont strictement numériques et ont une valeur finie.
- ceux dont tous les états ne sont pas connus à la création de la base, mais dont le maximum est un nombre fini donné.
- ceux qui ne peuvent être traités que comme des chaînes de caractères (commentaires).

— **DEFINE ITEMS\*\*\***: cette commande permet de définir la source d'entrée des données et leur format. Elle peut être exécutée aussi souvent que nécessaire et à n'importe quel moment de la vie d'une base de données.

---

\*k = 1025 octets. Unité de quantité de mémoire d'un ordinateur. octet = élément d'information de 8 bits.

\*\***Define descriptors** : définition des descripteurs.

\*\*\***Define items**: définition d'un enregistrement.

La source d'entrée peut être: des cartes,  
une bande,  
un disque,  
un clavier.

Le format peut être: fixe,  
libre,  
ou le même que celui défini à la création de la base  
ou lors de l'utilisation précédente (**SAME\***)

Si la description est acceptée par **EXIR** il répond par un point d'interrogation sinon il signale à partir de quel descripteur il y a une erreur. Lorsque l'entrée des données est terminée, on répond au point d'interrogation par **END OF ITEMS\*\***.

— **EDIT ITEMS**: cette commande permet de faire éditer toutes les données afin de les vérifier avant de les écrire dans la banque (car jusqu'à ce stade de travail les données n'existent que sous forme de fichier temporaire).

— **RESEQUENCE DESCRIPTORS\*\*\***: permet de changer l'ordre des descripteurs défini dans la commande **DEFINE DESCRIPTORS**. On peut le faire de deux façons, soit en donnant l'ancien numéro du descripteur et le nouveau, soit en donnant son nom et le nouvel ordre qui lui est attribué.

— **ORIGINAL DESCRIPTORS\*\*\*\***: cette commande est utilisée lorsqu'au cours d'un travail on doit changer plusieurs fois l'ordre des descripteurs. **ORIGINAL DESCRIPTORS** redonne les descripteurs dans l'ordre donné à la création, et non pas l'ordre courant. Ceci évite les erreurs au moment de la renumérotation des descripteurs.

— **CONTROL VOCABULARY\*\*\*\*\***: fait imprimer toutes les définitions des descripteurs. On peut ainsi vérifier une dernière fois la banque avant de l'écrire.

— Compression des données: Elle se fait automatiquement, sans aucune gestion de la part de l'utilisateur. Dès qu'une zone est vide (blanc pour l'alphabétique, zéro pour le numérique) la zone est compactée, cela veut dire qu'au lieu de mettre toute une zone blanche, le système compte le nombre de blancs qu'il devrait y avoir et enregistre simplement ce nombre. On gagne ainsi une très grande place. Une fois la banque créée et contrôlée on l'écrit par la commande **WRITE DATA BANK\*\*\*\*\***. Maintenant la banque existe réellement sous forme **EXIR**.

— Correction et mise à jour de la banque: si des erreurs ont été détectées on peut les corriger par 4 commandes qui permettent soit de modifier l'état d'un descripteur par un enregistrement donné, soit de détruire des enregistrements complets ou des états de descripteurs, soit de rajouter de nouveaux descripteurs.

e. *Manipulation et interrogation de la base.* Une fois la banque créée et corrigée. 7 commandes permettent de la manipuler ou de l'interroger\*\*\*\*\*.

---

\*Même définition des enregistrements qu'à la création.

\*\*Fin de définition des enregistrements.

\*\*\*Changement de séquence des descripteurs.

\*\*\*\*Séquence d'origine des descripteurs.

\*\*\*\*\*Contrôle du vocabulaire.

\*\*\*\*\*Enregistrer la base de données.

\*\*\*\*\*Combien, imprimer, générer, statistiques, déchargement des enregistrements, blocage des enregistrements, déblocage des enregistrements.

HOW MANY  
PRINT  
GENERATE  
STATISTICS  
DUMP ITEMS  
DEACTIVE ITEMS  
REACTIVE ITEMS

— **HOW MANY** est la plus simple de ces commandes. Elle compte simplement le nombre d'enregistrements répondant à cette question. On peut enchaîner une série de **HOW MANY**.

— Pour avoir l'impression du résultat obtenu on intercale **PRINT** entre chaque question, ou à la fin de la liste de questions.

· **GENERATE** exécute le même travail que **PRINT** mais au lieu de donner une liste des enregistrements sur imprimante il crée un fichier sur bande ou sur disque.

— **STATISTICS** ressemble à **GENERATE** mais ne crée des fichiers qu'à partir de descripteurs de type numérique. Ces fichiers pourront être utilisés pour des calculs statistiques à partir de bibliothèques telles que **SPSS**, **BMDP**... Ce qui n'était pas possible dans les anciennes versions. Il n'y a donc plus aucun problème de compatibilité entre **EXIR** et toutes ces bibliothèques.

— **DUMP ITEMS**. Cette commande crée des fichiers sur cartes, banques ou disques qui sont les images-cartes des enregistrements. On peut sélectionner les descripteurs et les écrire dans le format que l'on désire.

— **DEACTIVATE ITEMS**. Cette commande enlève temporairement des enregistrements de la banque sans les détruire. Quelquefois il peut être plus économique d'enlever une partie des données de la banque pour accélérer une recherche, si l'on sait d'avance quels enregistrements n'ont aucune chance d'être sélectionnés.

— **REACTIVATE ITEMS**. Réintroduit dans la banque les enregistrements enlevés par la commande **DEACTIVATE**.

f. *Problèmes encore en suspens.* Il y a une très nette amélioration d'**EXIR**. Son emploi est devenu souple et simple. Mais il reste 3 problèmes non résolus dont deux d'une très grande importance par rapport aux problèmes que nous avons à résoudre.

— Descripteurs à zones multiples variables. Il faut encore créer un descripteur pour un état que l'on sait être rarement réalisé. Mais ce problème peut être considéré comme secondaire puisqu'il s'agit d'un problème de stockage et non pas de gestion de base, à condition que la base soit de taille moyenne.

— Création de deux fichiers dans une même base et recherche récursive. Ces deux problèmes sont par contre d'une très grande importance car ils rendent la gestion de la généalogie extrêmement difficile à réaliser, à tel point que l'équipe **EXIR** ne pense pas l'aborder tant que **EXIR** existe sous cette forme. Si l'**IS/GR** n'a pas encore donné de solution pour la **généalo-**

gie c'est d'une part parce qu'il n'y avait pas ou peu de demande, mais aussi et surtout parce que le logiciel n'était pas adapté à ce problème (ce pourquoi GDM est mieux placé).

## 2. Gestion de données en ressources génétiques. (par le système GDM)\*

a. *Présentation.* GDM est un système conçu pour créer et gérer des données dans le cadre de la gestion des ressources génétiques, lorsque le volume des données devient trop grand pour être géré manuellement, mais qu'on ne possède pas de moyens financiers suffisants pour s'équiper d'un gros ordinateur et des logiciels nécessaires. Sa conception a été rendue possible du fait de la diminution spectaculaire du prix des micro-ordinateurs et simultanément de l'accroissement de leur taille mémoire et de leur puissance de calcul. Dans le cas de la création d'une banque centrale et de banques locales, il permet à chaque banque locale de créer et gérer ses propres données pour les traitements simples. Lorsque le traitement demande plus de place mémoire, (cas d'analyses *multivariates*), il peut être effectué au niveau de la banque centrale sur un gros ordinateur, bien que l'on commence à voir apparaître des programmes d'analyses *multivariées* pour micro-ordinateur, ce qui laisse penser que très bientôt ce problème n'existera plus pour des fichiers de taille moyenne.

b. *Fichiers GDM.* Un fichier GDM est composé d'enregistrements qui contiennent chacun les données concernant un seul individu, chaque «champ» ou donnée étant indépendant l'un de l'autre. En terme GDM chaque *champ* est un *descripteur* (c'est-à-dire une clé d'accès). GDM associe automatiquement un numéro à chaque enregistrement par ordre d'entrée. Ce numéro n'est pas un descripteur et l'utilisateur ne peut y avoir accès. Pour compacter le fichier, les données sont enregistrées sous forme d'une «chaîne de caractères», la virgule est le marqueur de fin de descripteur. La figure suivante montre un exemple de fichier, tel qu'il est enregistré sur disquette.

---

\* GDM: gestion des données en ressources génétiques ( *germplasm data management* )

Exemple pour le GDM :

\*FILE.GF4.TXT,DES,9,10,22,TEXT,2  
 CARACTERISTICAS DE MAZORCA Y GRANO DEL DEPARTAMENTO

9 NUMERO DE COLECCION  
 10 LARGO DE MAZORCA  
 11 ANCHO DE MAZORCA  
 12 NUMERO DE HILERAS  
 13 ANCHO DE PEDUNCULO DE MAZORCA  
 14 LARGO DE PEDUNCULO DE MAZORCA  
 15 NUMERO DE MUDOS  
 16 LARGO DE GRANO  
 17 ANCHO DE GRANO  
 18 ESPESOR DE GRANO  
 19 TEXTURA DE GRANO  
 20 COLOR DE GRANO  
 21 FORMA DE GRANO  
 22 DEPRESION DE GRANO  
 (REC, 1,1001,12,0,5,1,0,11,0,0,0,10,8,8,7,4,8,2,0,1,0,(REC), 2,1002,10,8,4,3,9,6,10,8,19,0,0,11,9,10,5,5,0,1,1,MDS,3,(REC),  
 3,1003,13,5,4,4,10,6,11,9,0,10,7,10,2,6,0,1,9,1,MDS,2,(REC), 4,1006,12,8,5,0,11,0,11,8,0,12,8,10,8,5,4,1,0,5,MDS,2,(REC), 5,  
 1007,13,6,4,2,19,6,10,6,0,11,0,10,4,3,6,1,8,MDS,MDS,D,(REC), 6,1008,12,0,4,2,8,8,9,1,0,12,1,11,1,5,1,0,5,MDS,3,(REC), 7,1009,  
 13,2,4,3,9,9,10,9,0,10,6,10,3,6,0,1,9,1,5,3,(REC), 8,1011,14,4,4,8,11,4,11,6,0,13,0,10,5,5,8,1,5,8,MDS,4,(REC), 9,1012,14,6,5,  
 1,0,8,15,0,0,12,6,10,5,5,1,5,MDS,MDS,7,(REC), 10,1013,14,1,5,2,12,8,13,6,0,12,9,10,4,5,5,1,6,8,MDS,1,4,(REC), 11,1014,13,9,  
 5,2,12,6,13,6,0,0,12,8,10,7,5,8,1,2,(REC), 12,1015,14,9,5,11,6,13,8,0,12,5,10,9,5,8,1,7,MDS,MDS,9,(REC), 13,1016,14,  
 9,4,8,0,11,9,0,11,5,10,7,6,1,1,9,1,MDS,5,(REC), 14,1017,15,6,4,9,11,8,13,0,0,12,7,10,7,5,6,2,0,1,1,5,(REC), 15,1018,14,6,4,7,1,  
 1,3,12,5,0,0,12,3,10,7,5,6,1,3,14,MDS,(REC), 16,1020,12,2,5,5,12,8,13,5,0,0,12,8,11,5,6,3,1,0,5,0,1,(REC), 17,1021,13,4,5,2,13,  
 8,11,8,0,12,5,10,6,5,7,1,0,5,MDS,1,3,(REC), 18,1022,13,5,3,8,11,1,10,0,0,0,11,9,4,5,4,1,7,1,MDS,1,(REC), 19,1023,13,7,5,0,11,  
 9,14,2,0,12,4,10,3,5,9,1,1,MDS,18,9,(REC), 20,1024,14,4,4,5,3,1,2,13,7,0,12,9,10,5,5,7,1,0,MDS,MDS,1,(REC), 21,1025,11,6,4,4,9,  
 4,12,1,0,12,4,10,5,5,9,2,0,1,MDS,0,(REC), 22,1026,9,8,4,9,10,3,12,0,13,8,11,0,5,4,1,0,MDS,14,1,16,1,(REC), 23,1027,10,4,4,5,10,  
 0,11,2,0,0,1,1,10,2,6,0,1,7,MDS,9,3,(REC), 24,1028,14,2,4,7,10,4,11,8,0,12,5,11,0,5,8,1,8,1,MDS,3,(REC), 25,1029,10,3,5,1,10,9,  
 11,4,0,19,4,11,4,5,4,1,0,5,MDS,1,4,(REC), 26,1030,12,3,4,9,10,6,10,9,0,12,8,10,7,5,6,1,MDS,9,1,(REC), 27,1031,11,9,4,5,9,5,12,  
 0,0,12,4,10,4,5,4,1,2,18,5,(REC), 28,1032,13,1,5,10,6,12,3,9,7,2,12,8,10,7,5,7,1,1,5,MDS,7,(REC), 29,1033,8,9,5,3,0,12,6,0,  
 0,13,8,11,4,5,8,1,0,5,MDS,1,7,(REC), 30,1034,12,6,4,4,9,4,11,8,0,0,13,0,10,7,5,4,1,4,1,MDS,2,(REC), 31,1035,10,6,4,7,0,10,1,0,0,1,  
 1,10,5,5,7,1,0,5,MDS,(REC), 32,1036,11,4,7,10,0,11,3,0,0,12,8,10,9,5,3,1,0,1,MDS,1,1,(REC), 33,1037,13,6,4,3,10,0,12,2,0,0,11,  
 7,10,2,5,8,1,8,MDS,1,2,(REC), 34,1038,10,4,5,0,11,12,0,14,3,10,4,5,0,1,2,1,3,1,4,(REC), 35,1039,9,2,4,7,10,5,9,0,12,2,0,0,13,5,10,  
 4,5,1,1,0,5,MDS,1,1,(REC), 36,1040,9,4,4,6,10,3,9,8,0,0,12,7,10,1,5,3,1,0,MDS,MDS,6,(REC), 37,1041,11,7,4,2,9,4,10,0,0,0,12,0,10,  
 5,3,3,2,0,1,1,5,(REC), 38,1042,17,0,5,0,6,13,3,13,0,0,11,2,9,6,5,5,2,0,6,1,5,1,3,(REC), 39,1043,16,2,5,2,13,7,0,0,10,6,10,1,4,5,0,  
 0,0,0,EOF

Le fichier s'appelle GF4.TXT, ses descripteurs sont numérotés de 9 à 22, les descripteurs du fichier occupent deux lignes. On voit ceci dans la phrase:

— FILE, GFILE4.TXT, DES, 9, TO, 22, TEXT, 2

Ensuite GDM imprime la description du fichier GFILE4.TXT sur deux lignes puis le numéro et le nom des descripteurs.

(REC),	18,	1022
annonce un nouvel enregistrement	n° enregistrement (inaccessible par l'utilisateur)	n° d'accession

13.6, 3.8, 11.1, 10,0, D,D 11.1, 9.4, 5.4, 1.7, 1, MDS, 1, est la liste des valeurs des descripteurs pour l'enregistrement. EOF indique la fin de fichier (End of File).

GDM est à la fois un hardware" et un software\*\*.

— Le hardware se compose de 3 parties:

- 1 mémoire de 48 k (peut varier de 16 k à 64 k),
- 1 écran,
- 1 imprimante de 132 colonnes.

Ce système peut marcher de façon autonome (banques locales) ou être connecté à un gros ordinateur par téléphone (banque centrale). L'enregistrement des données se fait sur disquettes, ce qui facilite l'échange de données d'une banque à l'autre.

c. *Fonctions existantes.* Certaines fonctions sont réalisées directement par le système, d'autres par des logiciels. Ils sont écrits en langage BASIC\*\*\* étendu et peuvent être utilisés sur un système d'exploitation CP/M. GDM peut être installé facilement sur n'importe quelle machine avec un S-100 bus\*\*\*\* supportant le système d'exploitation\*\*\*\*\* CP/M. et avec plus de difficultés sur d'autres systèmes.

---

"HARDWARE = mot anglais signifiant «quincaillerie» et par analogie définit tout ce qui est matériel informatique. Il s'oppose à Software.

\*\*SOFTWARE = mot anglais fabriqué par opposition à Hardware pour désigner tout ce qui est immatériel en informatique. Il comprend les programmes et la documentation permettant de faire fonctionner un ordinateur. On distingue le «software» de base (traducteur de langage, système d'exploitation) et le «software» d'application (les programmes de gestion, calcul, etc...).

\*\*\*Basic : langage évolué, ayant la particularité d'occuper peu de place mémoire, d'être facile à apprendre et à utiliser. C'est souvent un langage destiné à l'enseignement.

\*\*\*\*Bus : Abréviation de ligne omnibus. Ensemble de fils permettant l'interconnexion de plusieurs organes digitaux. Organe digital: partie de l'ordinateur permettant le stockage des informations (ex. les registres de la mémoire centrale).

\*\*\*\*\*Système d'exploitation: ensemble homogène de programmes conçus pour travailler les uns avec les autres, se répartissant en programmes de traitement et programmes de contrôle. C'est cet ensemble de programmes qui assure la gestion de l'ordinateur et en donne son type. Ici le type est CP/M (control program for microprocessor) d'autres peuvent être DOS, OS/VS, etc... CP/M est essentiellement l'interface logiciel entre l'utilisateur et le système. Il fournit un ensemble limité de commandes ainsi que des programmes utilitaires facilitant l'utilisation du système. Il a été réalisé pour occuper une faible quantité de mémoire.

On peut réaliser les fonctions suivantes:

- Création de fichier (DE1)
- Contrôle des données (DE2A)
- Edition des erreurs de données (DE2B)
- Ajouts de nouveaux descripteurs (JOIN)
- Ajouts de nouveaux enregistrements (PIP)
- Mise à jour des fichiers: correction d'erreurs, retrait d'enregistrements (Z-tel).
- Interrogation des fichiers (QS). Cette fonction permet la création de fichiers intermédiaires résultants de l'interrogation (fonction facultative). Elle permet d'extraire du fichier les enregistrements et les descripteurs utiles. Le mode d'interrogation ressemble à celui d'EXIR, il est aussi basé sur l'algèbre de BOOLE.
- Il existe aussi quelques programmes statistiques (STAT 1 à STAT 4) mais cette fonction est encore peu développée.
- STAT1 réalise des fonction de statistiques descriptives (moyenne, variance, régression linéaire, corrélation entre 2 descripteurs).
- STAT2 régression multiple par étape.
- STAT3 analyse de variance à 1 facteur (modèle randomisé).
- STAT4 analyse de variance à 2 facteurs (bloc, complètement randomisé).
- Destruction de fichiers (ERASE).
- Editions de fichiers:
  - Soit tels qu'ils existent (TYPE)
  - Soit sous forme de tableau (RW) sur imprimante ou sur écran.
- Il permet la création de descripteurs à zones multiples variables (par «MDS» multiple `descriptor state`: lorsqu'on déclare un descripteur «MDS» le système accepte plusieurs états pour un même descripteur et les range dans des fichiers secondaires).
- Une autre fonction (SC) permet la gestion des stocks et doit être précisément définie en fonction des besoins que l'on a.
- PASPORT est une fonction qui transforme les données du système GDM dans un autre modèle spécifié par l'utilisateur, par exemple pour rendre les données compatibles entre deux `microordinateurs` de système d'exploitation différents.

d. *Extensions à court terme des logiciels.* La connexion avec un `gros` système est réalisée seulement pour CDC (par ce que l'IS/GR est connecté à un CDC) et IBM, toute base créée de façon autonome est ainsi `connectable` tant sur IBM que CDC pour des traitements d'analyses `multivariées` par exemple.

— Logiciel d'extraction de généalogie. Mr McMILLAN étudie actuellement des propositions de logiciels adaptables à GDM susceptibles de rendre la gestion de la généalogie plus `facile`\*. Le problème d'extraction de généalogie est difficile à résoudre à grande échelle de façon satisfaisante.

— Adaptation d'EXIR à GDM. L'IS/GR étudie la possibilité d'étendre EXIR à GDM et pense pouvoir le réaliser en utilisant le langage BASIC.

---

\*Nos travaux sur le SGBD ADABAS nous permettent l'étude complète de généalogies sur un IBM 370/168 (cf. exemple des extractions de généalogies).

### 3. Comparaison des deux logiciels EXIR et GDM

EXIR est un logiciel qui semble ne plus pouvoir beaucoup évoluer parce que l'équipe qui le travaillait est disséminée et les personnes restantes ne peuvent pas s'attaquer à un si gros problème, ou du moins pas dans l'immédiat. C'est un logiciel coûteux à l'utilisation et exigeant beaucoup de place mémoire. Il s'est toutefois nettement amélioré par rapport aux précédentes versions quant à la souplesse d'utilisation.

GDM est peu exigeant en mémoire et d'une très grande facilité d'emploi pour non informaticiens. Le stockage ne coûte que le prix des disquettes (5 dollars la disquette). Il est limité à l'heure actuelle à 50 descripteurs mais le logiciel est transformable pour plus de descripteurs. C'est encore un logiciel en pleine évolution. Ce système qui peut marcher de façon complètement autonome par rapport à un gros ordinateur laisse penser qu'il est parfaitement adapté à l'organisation des réseaux de banques locales gérant elles-mêmes leurs données. La banque centrale ne serait indispensable que pour de gros traitements et la création des fichiers centraux (soit sur système EXIR, ou CODASYL ou ADABAS). L'échange des données peut se faire par courrier en envoyant les disquettes par poste. L'utilisation de ce système ne demande pas de formation informatique aux utilisateurs. Seul le gestionnaire de la banque centrale a besoin de ces connaissances pour adapter ou créer de nouveaux logiciels en fonction des besoins.

## IV. ORGANISATION TAXONOMIQUE DU COMPLEXE ÉTUDIÉ

Bien que cela ne corresponde pas encore à la pratique habituelle des grandes «banques de gènes», nous avons montré que pour organiser une évaluation agronomique décentralisée il était indispensable de synthétiser au mieux les données acquises par l'évaluation génétique. Des bases de données inexploitées ou *ininterprétées* sont des capitaux «vitrines» qui ne fructifient pas. Des gestionnaires responsables auront donc à cœur de fouiller leurs bases de données et de mener les expérimentations complémentaires nécessaires pour pouvoir donner une idée précise des richesses génétiques dont ils ont la charge. Nous allons montrer comment les bases de données peuvent être analysées pour fournir des classifications utiles des complexes d'espèces conservés.

### A. LES GRANDS ENSEMBLES ÉTABLIS PAR L'ÉTUDE GÉNÉTIQUE QUALITATIVE

Les données, souvent enregistrées sous forme numérique, ne devront pas faire oublier que les grands compartiments des complexes d'espèces sont d'abord établis par des études biologiques qui n'exigent pas de traitement statistique particulier. Ces travaux, indispensables pour définir

des coupures raisonnables entre des grands groupes font appel aux points de repères déjà établis par la taxonomie botanique classique, les observations cytogénétiques, les études de biologie florale et les repérages de barrières reproductives entre groupes, les observations **ethno** et **écogénétiques** visant au repérage et à la définition des formes spontanées (sauvages et adventices) et des formes cultivées, etc...

Le traitement statistique des données visera seulement à évaluer les distances génétiques entre les groupes qui auront été définis à l'aide des méthodes précédentes. Ces évaluations permettent de hiérarchiser les compartiments du complexe en précisant leur degré de connexion évolutive, ceux entre lesquels les barrières reproductives témoignent d'une différenciation génétique profonde (spéciations **allopatriques** en cours) et ceux pour lesquels les barrières reproductives sont la conséquence et l'expression d'une organisation adaptative coordonnée entre des formes complémentaires unies par des flux de gènes contrôlés (cf. chapitre I). En particulier ces hiérarchies guideront le sélectionneur pour planifier et maîtriser ses travaux d'hybridations **interspécifiques**.

## **B. RECHERCHE DE CLASSIFICATIONS PLUS FINES CONSTRUITES SUR DES DONNÉES GÉNÉTIQUES PONCTUELLES**

Cette recherche de classification peut paraître absurde à priori puisqu'il s'agit d'établir des groupes dans un ensemble d'éléments pour lequel aucune coupure biologique nette n'est manifeste, puisqu'alors elle aurait été traitée dans le paragraphe précédent. On doit en fait traiter d'une diversité génétique plus ou moins continue ou recombinaisonnée et les limites de groupes (contours plus ou moins flous) seront en partie arbitraires. Les travaux de sélection (hybridation et fixation des types nouveaux recombinaisonnés) auront le plus souvent pour conséquence la création d'intermédiaires entre des groupes qui auront été définis. Ainsi les classifications que nous tenterons d'établir ne seront ni complètement rigoureuses ni définitives. Elles sont cependant le moyen de décrire une collection d'une façon beaucoup plus assimilable et compréhensible que par un tableau immense des distances génétiques constituées deux à deux entre toutes les entrées de la collection. **Biologiquement** elles mettront en évidence une organisation partiellement structurée de la variabilité génétique disponible ce qui correspond en profondeur à des lois vraisemblables non encore établies de la génétique des populations qui font pressentir que les associations cohérentes bien fonctionnelles des états **alléliques** pour des groupes de gènes sont en nombre limité. C'est tout ce que les vocabulaires de «groupes de gènes **coadaptés**», «structures de déséquilibre **gamétiques**», «**supergènes**», «**linkats**», etc... tentent de cerner. Avant d'illustrer les traitements **multivariés** des données, résumons les objectifs et les difficultés.

Deux objectifs principaux sont les suivants:

— *Faciliter la vision d'ensemble de la collection.* Il s'agit là de guider l'échantillonnage pour permettre des évaluations agronomiques **décentralisées** (en repérant d'abord les grands types d'organisation, peut-on établir

une typologie?) et d'orienter le sélectionneur en lui révélant des sous-ensembles entre lesquels l'ordre de grandeur des distances génétiques aura été calculé (recherche d'hétérosis, recherche de structures complémentaires ou nouvelles).

— *Apprécier immédiatement le degré d'originalité de toute nouvelle introduction.* Cet objectif est le plus simple à réaliser dès qu'une certaine classification a été proposée, il consiste à établir des règles d'attribution à un groupe, il simplifiera considérablement les tâches du sélectionneur, toujours à l'affût des lignées nouvelles qu'il introduit sans connaître en général les détails de leur obtention génétique. Beaucoup d'évaluations agronomiques décevantes pourront être ainsi évitées.

Rappelons les difficultés rencontrées, il n'existe probablement pas une seule classification capable de mettre l'accent sur les grands éléments d'organisation génétique sous-jacents. Conséquence de l'histoire évolutive du complexe: origines multiples des formes cultivées, suivies de confrontation, création de blocs géniques coadaptés variés préservés par des limitations des recombinaisons génétiques, etc... les classifications ne seront pas univoques, les contours des groupes seront le plus souvent sans coupures nettes. Le travail génétique tend à recréer des types intermédiaires et à déformer les typologies.

En conséquence il faudra accepter des «côtes mal taillées» et provisoires. L'absence d'absolu biologique unique définit à l'apport statistique un rôle de soutien du jugement pratique et non pas celui d'une méthode capable de révéler et définir une vérité unique qui le plus souvent n'existera pas.

## C. ILLUSTRATION DES MÉTHODES D'ANALYSE DES DONNÉES ET DE TAXONOMIE NUMÉRIQUE.

Nous avons volontairement choisi un exemple très simple, ne comportant qu'un seul type de données (des analyses enzymatiques par électrophorèse) et une diversité biologique sous-jacente assez nettement structurée (les riz cultivés asiatiques et africains) pour que le jeu d'attribution d'individus à des groupes ne paraisse pas trop arbitraire.

### 1. Description des données utilisées

L'exemple choisi utilise des données tirées d'une étude plus vaste de G. SECOND (1982). Il s'agit de décrire l'ensemble des riz cultivés, à partir d'un échantillonnage des collections mondiales représentatif de la diversité des grands types connus: *Oryza glaberrima* (riz africain), *Oryza japonica* et *O. sativa* (riz asiatique) et des formes difficiles à cataloguer dans ces grands types (certaines ont été désignées *javanica*). A ces riz cultivés on a adjoint un échantillonnage des formes spontanées annuelles (sauvages et adventices) africaines: *O. breviligulata* (cf. le chapitre riz, vol. I). Soixante lignées sont ici analysées à partir d'une liste de descripteurs limités à 40 locus codant pour des enzymes étudiées par électrophorèse. Suivant ces enzymes étudiés, les états possibles des différents caractères (ici le nombre d'allèles présents dans la collection de 60 lignées) variaient de 1 (absence de polymorphisme) à 4.

## 2. Représentations d'ensemble

Les analyses en composantes principales et en **correspondances\***, construites respectivement par diagonalisation d'une matrice de corrélation ou d'une matrice de distance  $x^2$  entre lignées ont pour but de représenter le nuage multidimensionnel (40 dimensions: 40 locus) des points représentatifs des 60 lignées à travers les projections sur des plans successifs déterminés par les vecteurs propres de la diagonalisation. Les axes constitués par ces vecteurs ont les propriétés suivantes: ils sont classés par ordre décroissant des valeurs propres (éléments de la trace de la matrice diagonale équivalente à la matrice de départ), ils sont indépendants entre eux. Géométriquement cela revient à observer la dispersion des lignées en considérant les directions dans lesquelles le nuage est le plus étiré. Ainsi la projection sur le plan des deux premiers axes correspond au déploiement le plus important du nuage, rapporté à deux dimensions indépendantes (corrélation nulle entre l'ensemble des couples (abscisse, ordonnée) calculée sur les 60 lignées). Les valeurs propres, converties en pourcentage de leur somme évaluant en quelque sorte la part de la dispersion qui est représentée par la projection du nuage sur un seul axe. La différence entre 100 et la somme des pourcentages due aux deux axes décrivant un plan donné révèle l'étendue de la dispersion non décrite par ce plan et permet de douter de la réalité des agglomérations qui sur ce plan peut rapprocher des points peut-être très dispersés dans l'espace (dans des directions définies par les autres axes).

Les graphiques de la figure 14 montrent la disposition des points, on peut être tenté de voir à travers ces représentations trois noyaux bien condensés avec quelques rares éléments de transition. On peut constater que dans les deux représentations Analyse en composantes principales (fig. 14) et Analyse en correspondance (fig. 15) les voisinages entre points sont comparables.

Les **dendogrammes** (fig. 16) arborescences constituées par regroupements successifs deux à deux de toutes les lignées, pourront être réalisés en utilisant divers indices de ressemblance (dont la distance  $X^2$  de l'analyse en correspondance). Les traits horizontaux raccordant des lignées entre elles ou des ensembles constitués par des regroupements successifs se situent au niveau des ressemblances moyennes entre ces deux ensembles (ou au niveau de la ressemblance moyenne des éléments à l'intérieur du super-ensemble constitué par le regroupement des deux ensembles précédents), ressemblances lues sur l'axe vertical. L'examen de la figure 16 peut suggérer des partitions en groupes différents suivant la valeur de l'indice de ressemblance au-dessus duquel on refuserait de considérer que les éléments regroupés constituent un ensemble d'une homogénéité acceptable. Ainsi une coupure au niveau de ressemblance  $r_1 = 0,2$  conduirait à définir 5 groupes; une coupure au niveau  $r_2 = 0,3$  proposerait une partition en 3 groupes mais pour lesquels quelques éléments paraissent mal intégrables, par exemple les n° (S1, S22, S23, S7, S39, B19, S36, S35,

---

\*Nous renvoyons bien entendu aux ouvrages spécialisés pour la description rigoureuse des méthodes (BENZECRI et al., 1973)

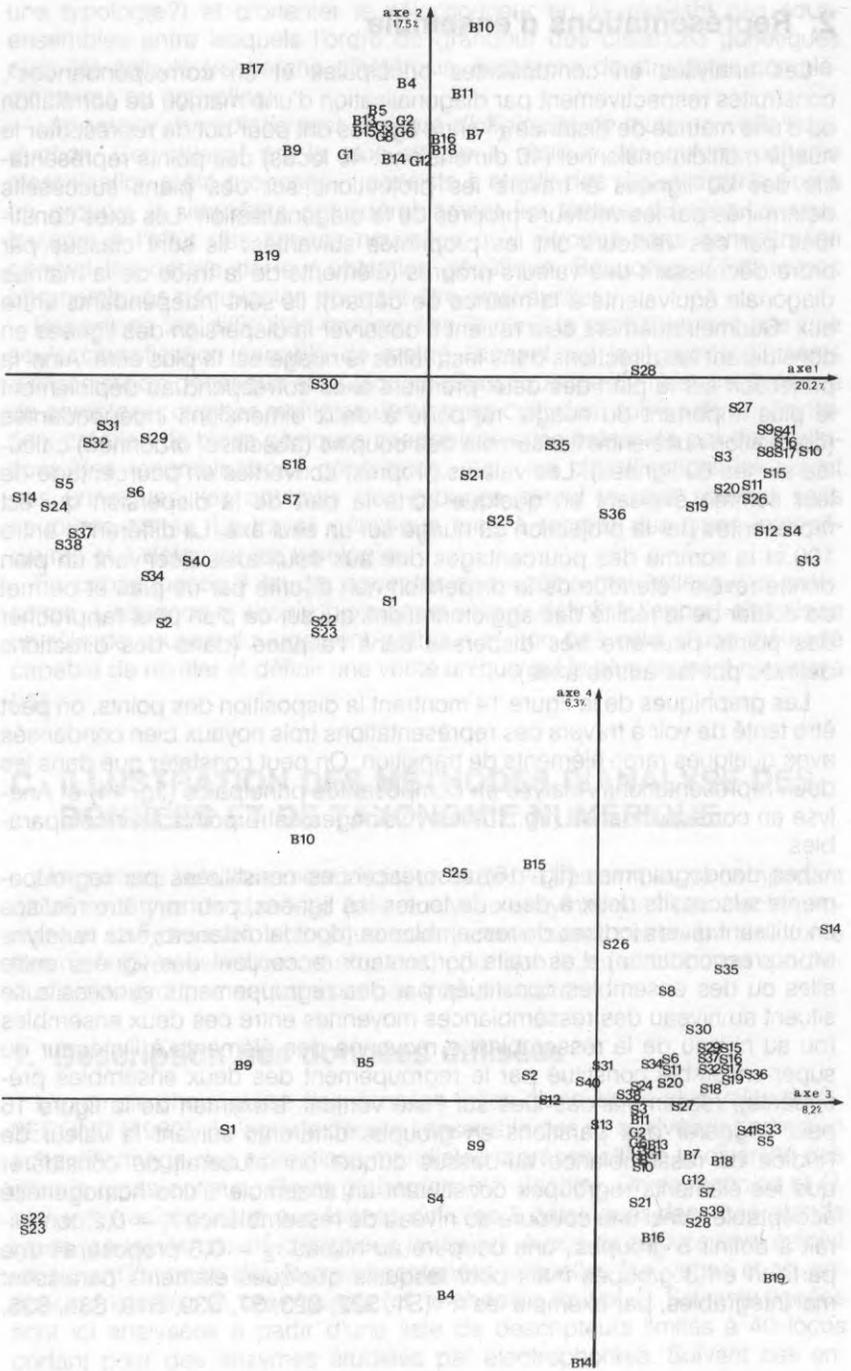
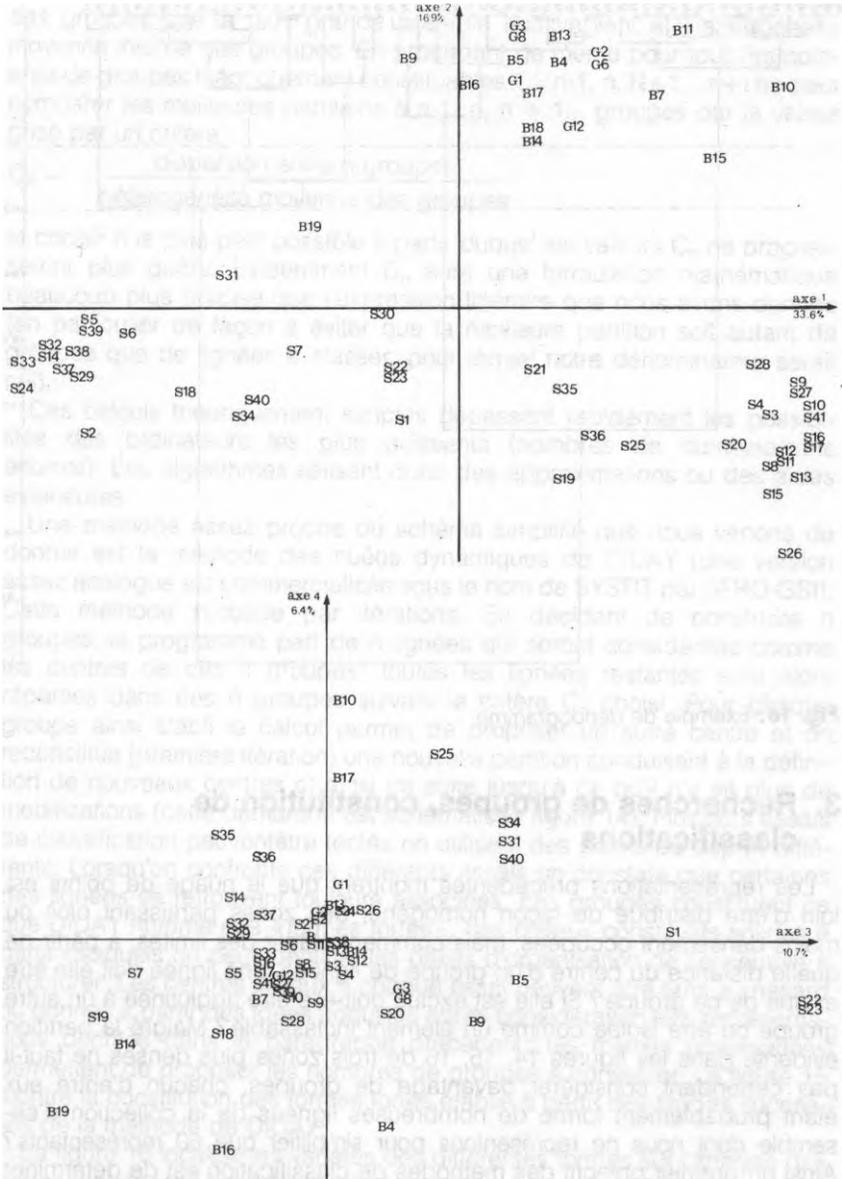
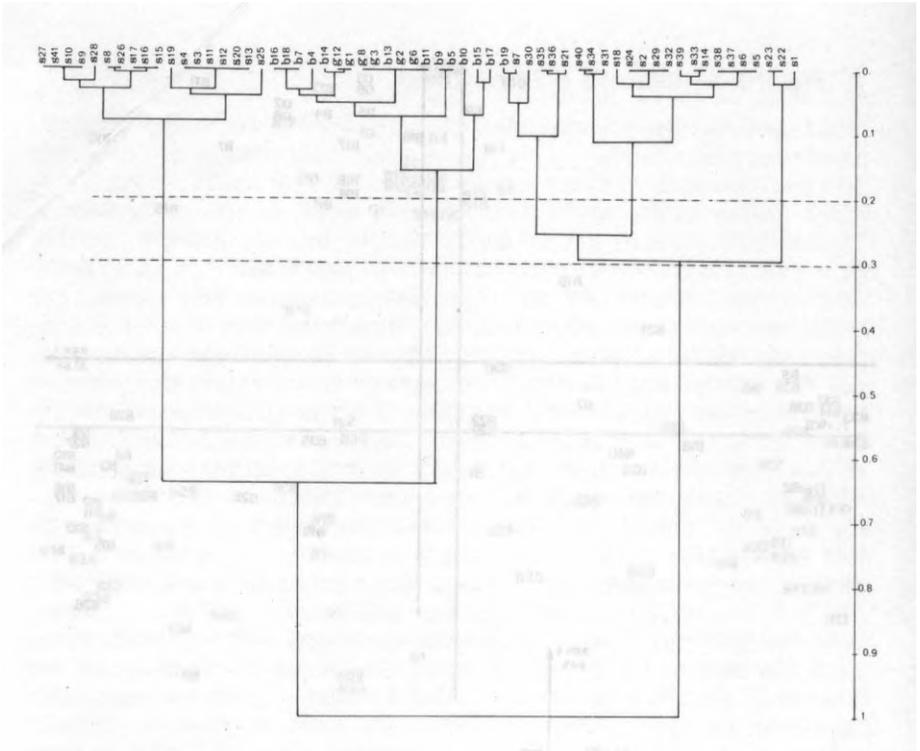


Fig. 14: Exemple d'analyse en composantes principales.



**Fig. 15:** Analyse en correspondance

S31) dans le groupe 3 qui paraît de ce fait très hétérogène. Remarquons que ces mêmes numéros paraissent être des éléments de transition entre noyaux plus denses dans les représentations des plan (1, 2) des analyses en composantes principales et en correspondances et étaient dans certains cas remarquablement excentriques dans les représentations du plan (3, 4).



**Fig. 16:** Exemple de dendrogramme

### 3, Recherches de groupes, constitution de classifications

Les représentations précédentes montrent que le nuage de points est loin d'être distribué de façon homogène, des zones paraissant plus ou moins densément occupées, mais comment établir des limites, à partir de quelle distance du centre d'un groupe de lignées, une lignée doit elle être exclue de ce groupe? Si elle est exclue doit-elle être agglutinée à un autre groupe ou être isolée comme un élément inclassable? Malgré la partition évidente dans les figures 14, 15, 16 de trois zones plus denses ne faut-il pas cependant considérer davantage de groupes, chacun d'entre eux étant probablement formé de nombreuses lignées de la collection d'ensemble dont nous ne représentons pour simplifier que 60 représentants? Ainsi un premier objectif des méthodes de classification est de déterminer ces groupes. L'ambition la plus haute serait que la méthode elle-même décide le nombre de groupes qu'il faut constituer. Une solution plus modeste, plus empirique mais plus économique est de chercher à constituer les groupes (définir leur composition en lignées) en nombres définis à priori par l'analyste: «répartir les 60 lignées en 3 groupes». Le deuxième objectif est de caractériser les groupes, de définir les caractéristiques principales qui ont rapproché les lignées et qui leur confèrent un certain type.

Un moyen de définir les groupes est le suivant: en constituant toutes les partitions et associations possibles des lignées en n groupes, choisir la

classification pour laquelle une mesure telle que la dispersion entre centres des groupes soit la plus grande possible relativement à l'hétérogénéité moyenne interne des groupes. En procédant de même pour tous les nombres de groupes théoriquement **constituables** 1... n-1, n, n+1... n+t on peut comparer les meilleures partitions à n-1, n, n + 1... groupes par la valeur prise par un critère

$$C_n = \frac{\text{dispersion entre } n \text{ groupes}}{\text{hétérogénéité moyenne des groupes}}$$

et choisir n le plus petit possible à partir duquel les valeurs  $C_n$  ne progresseront plus guère. Evidemment  $C_n$  aura une formulation mathématique beaucoup plus précise que l'expression littéraire que nous avons donnée (en particulier de façon à éviter que la meilleure partition soit autant de groupes que de lignées à classer, pour lequel notre dénominateur serait nul).

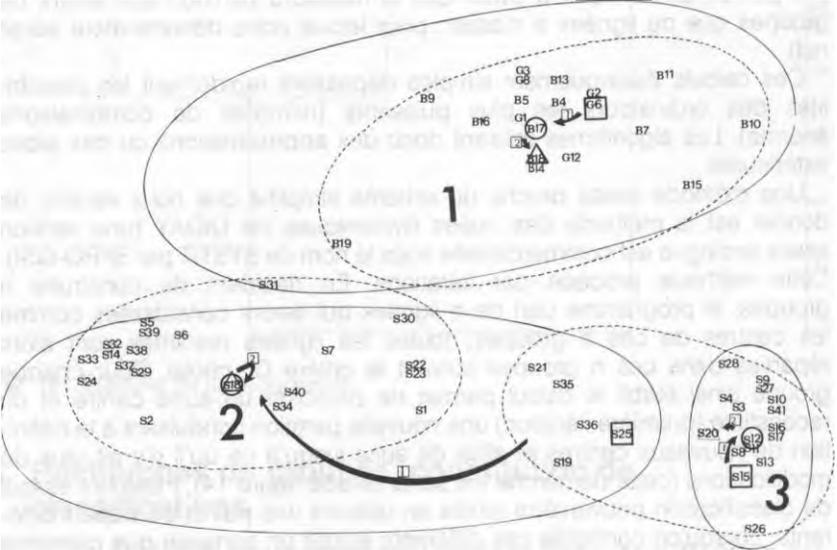
Ces calculs théoriquement simples dépassent rapidement les possibilités des ordinateurs les plus puissants (nombres de combinaisons énorme). Les algorithmes utilisent donc des approximations ou des aides extérieures.

Une méthode assez proche du schéma simplifié que nous venons de donner est la méthode des nuées dynamiques de DIDAY (une version assez analogue est commercialisée sous le nom de SYSTIT par SFRO-GSI). Cette méthode procède par itérations. En décidant de construire n groupes, le programme part de n lignées qui seront considérées comme les centres de ces n groupes; toutes les lignées restantes sont alors réparties dans ces n groupes suivant le critère  $C_n$  choisi. Pour chaque groupe ainsi établi le calcul permet de proposer un autre centre et on reconstitue (première itération) une nouvelle partition conduisant à la définition de nouveaux centres et ainsi de suite jusqu'à ce qu'il n'y ait plus de modifications (cette démarche est schématisée figure 14). Plusieurs essais de classification **peuvent être** tentés en utilisant des points de départ différents. Lorsqu'on confronte ces différents essais on constate que certaines des lignées se retrouvent toujours associées, ces groupes constituent ce que DIDAY nomme des « formes fortes », des noyaux constitutifs solides à partir desquels on peut identifier les bases d'organisation de l'ensemble à structurer. Les centres initiaux à chaque essai peuvent être tirés au hasard ou donnés à partir des idées acquises après considération des représentations décrites dans le paragraphe précédent; les mêmes observations permettent de proposer les nombres de groupes à constituer. La figure 15 montre la constitution des formes fortes établie à partir des tirages successifs de la méthode de DIDAY.

La figure 16 localise les variétés des différentes formes sur l'analyse des correspondances et montre les différents caractères associés à ces formes fortes ce qui permettra l'ébauche d'une typologie dont la description plus complète est réalisée à partir de la fréquence moyenne des états des différents caractères dans chaque noyau (tableau 28). Ce tableau constitue la typologie relative à chaque forme forte pour différents caractères.

## 4. Critères d'affectation des individus à des groupes (classement) et utilisation de la classification

Ainsi entre les formes fortes des éléments plus ou moins classés gravitent dans ces constellations. Si l'on veut simplifier la partition et répartir tous les individus entre les 3 catégories, on peut pratiquer des analyses discriminantes\* ou pour des variables qualitatives des méthodes telles que DISQUAL (commercialisée par SFRO-GSI). Nous allons procéder de la façon suivante:



**Fig. 17:** Schéma illustratif pour les nuées dynamiques

- trois lignes tirées au hasard comme centre regroupements proposés, puisqu'il faut tant bien que mal tout classer
- 1 nouveau centre proposé à la fin de ce classement
- centre pour la première itération
- nouveau regroupement proposé
- 2 changement de centre à l'issue de cette itération
- A nouveaux centres constitués

La nouvelle itération n'ayant pas de multiples propositions, la partition est achevée.

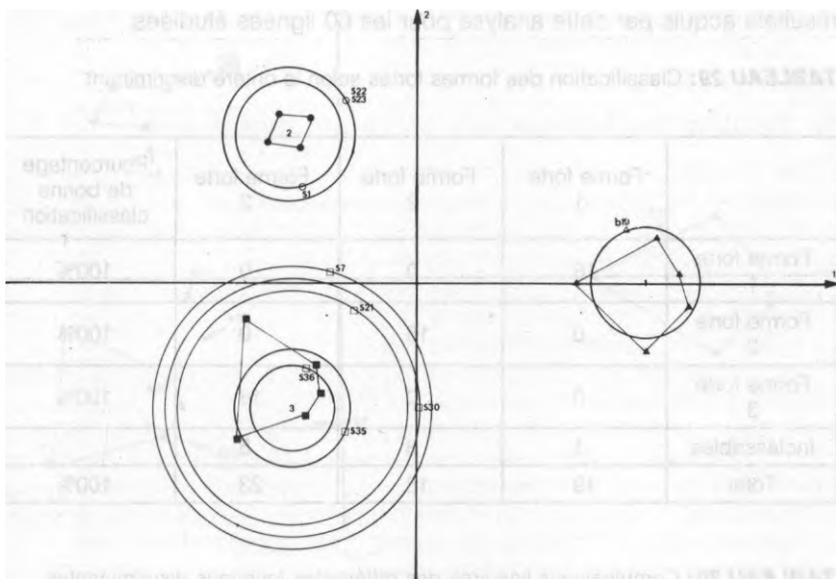
\*Pour les méthodes statistiques voir les livres spécialisés.



**TABLEAU 28:** Typologie de chaque forme forte pour différents caractères.

CHARACTÈRES	FORME FORTE 1	FORME FORTE 2	FORME FORTE 3
CA	1	1.65	1.95
EB	0.95	0.80	1
	0.26	1.40	0.09
RD	1.05	0.90	0.43
RR	1.05	1.30	0.52
	1	1.70	0.24
EG	1	0.15	0.95
EH	0.84	0	0.86
	1	0.25	1.95
EJ	0	1.85	0.65
A	1	0.65	0.47
S	1	1.20	1.14
E	1.10	1.15	1.19
A	1.15	1	1
AT	1	1.05	1.76
A	1	1.10	1
DH	1	1	1.04
AD	1.31	1.95	1.04
AC	0.68	0.90	0.04
AE	1.37	1	1
GA	1.16	1.05	1.95
GB	1.42	1.75	1
DB	1.10	1	0.80
A	1.26	1.65	1.19
EA	1.05	1	1
EB	1.26	3.30	3.04
EC	0.94	1	0.80
ED	1	1.05	1
EE	1.05	2	2

— Connaissant les lignées appartenant à chacune des formes nous allons procéder à une vérification de la fiabilité des attributions en posant la question suivante: sachant que les lignées n° B16, B18, B7, B4, B14, G12, G1, G8, G3, B13, G2, G6, B11, B9, B5, B10, B15, B17 appartiennent au groupe 1, les lignées n° S40, S34, S31, S18, S24, S2, S29, S32, S39, S33, S14, S38, S37, S6, S5 au groupe 2, les lignées n° S27, S41, S10, S9, S28, S8, S26, S17, S16, S15, S19, S4, S3, S11, S12, S20, S13, S25 au groupe 3 est-il possible que par la seule analyse de l'état des différents caractères de trouver une règle pour les affecter correctement à ces groupes? L'utilisation de la règle (en fait des valeurs prises par des fonctions discriminantes: combinaisons linéaires des différents états des caractères) conduit à proposer lignées n°. i → groupe 1, lignée n°. j → groupe 2, etc... On peut dénombrer les attributions correctes et le pourcentage de bonnes classifications, ce sera pour nous une bonne manière d'apprécier la qualité de la classification proposée au paragraphe précédent. Cette démarche est complémentaire de la précédente, car on peut toujours imaginer que des groupes puissent être constitués mais cela n'implique pas qu'ils aient une existence ou une signification suffisamment tangible pour que les attributions d'individus à ces groupes puissent être réalisées avec fiabilité.



**FIG. 20:** Projection de l'analyse discriminante dans le plan des axes 1 et 2.

— Une fois cette classification jugée acceptable, on voit que l'on dispose alors d'un critère d'attribution d'une lignée quelconque à ces différents groupes, car les fonctions discriminantes constituent de véritables clés taxonomiques. L'application de cette clé à toutes les lignées permet d'affecter chacune d'entre elles à l'un des trois groupes. Les valeurs prises par les fonctions discriminantes sont, dans certaines méthodes, les distances de chaque lignée au centre de chacun des groupes. Ces valeurs permettent donc d'apprécier la proximité ou l'éloignement de toute lignée par rapport aux noyaux de références. On peut voir ainsi que certaines lignées «flottent» loin de tout centre et échappent réellement à toute attribution précise. On constate que le centre de groupe le plus proche de S22, S23 est le groupe 2 mais à une distance 8,02 qui est très largement supérieure aux distances représentées à l'intérieur du groupe 2\*. De façon analogue le centre du groupe le plus proche de S30 est 3 mais la distance de S30 au centre du groupe 3 est 8,79 donc très largement supérieure encore au maximum de distance à l'intérieur du groupe 3. Ainsi S22, S23, S30 ne méritent pas réellement d'être intégrés à l'un des trois groupes 1, 2, 3. Il en est de même pour S7 et S35. Par contre S36 est à une distance 3,94 du centre du groupe 3 et B19 a une distance du centre du groupe 1 de 3,99 et s'intègrent donc bien respectivement à ces groupes.

Les projections sur le plan des 2 variables canoniques de l'analyse discriminante mettent bien en évidence la situation de ces différents points (figure 20).

\* Dans le groupe 1, le plus hétérogène les distances n'excèdent pas 6,73 dans le groupe 2, 5,53 et dans le groupe 3, 2,24.

Les tableaux 29 et 30, les figures 21 et 22 résument les différents résultats acquis par cette **analyse** pour les 60 lignées étudiées.

**TABLEAU 29:** Classification des formes fortes selon le critère discriminant

	Forme forte 1	Forme forte 2	Forme forte 3	Pourcentage de bonne classification
Forme forte 1	18	0	0	100%
Forme forte 2	0	15	0	100%
Forme forte 3	0	0	18	100%
Inclassables	1	3	5	
Total	19	18	23	100%

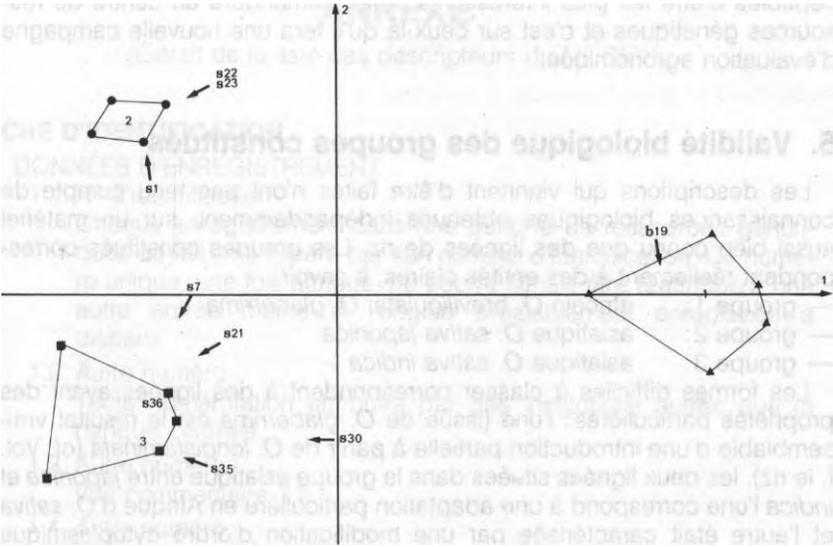
**TABLEAU 30:** Combinaisons linéaires des différentes fonctions discriminantes.

EH	25.97285	15.79875	34.68935
EJ	51.21442	115.68530	70.56705
CAT	0.05696	13.11389	42.45151
PGA	1.44416	-12.18014	22.19351
PGB	2.64415	1.32783	-12.63643
GOTB	7.26697	-15.21945	-20.49678
PEB	15.25392	38.19882	35.73387
PEC	55.03374	80.07568	23.51375
PEE	37.20390	112.64781	161.9906
CONSTANT	-144.14919	-426.31982	-386.32739

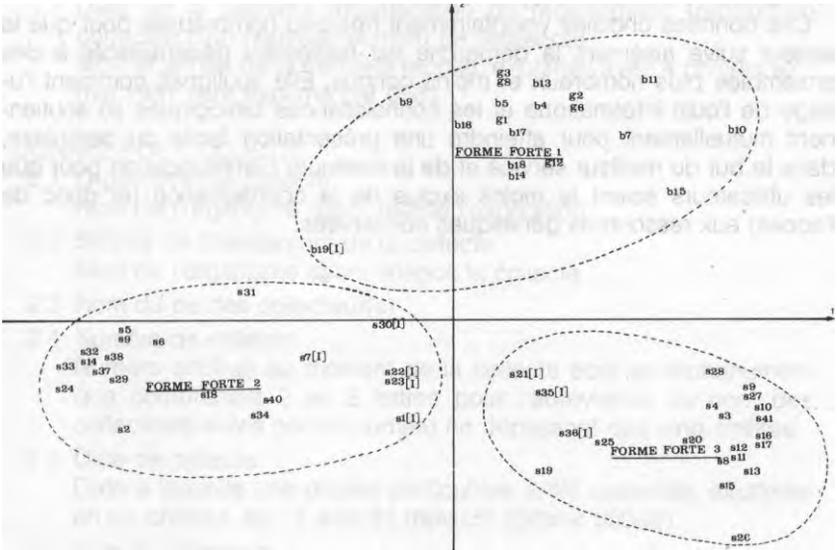
— L'utilisation de ces méthodes est pratiquement de grande valeur.

- D'une part la clé fournie avec les moyens utilisés permet après analyse **électrophorétique** de situer dans cet ensemble toute nouvelle lignée introduite.

- D'autre part, lors d'une évaluation agronomique et pour savoir à quelles propriétés correspondent ces groupes dans les conditions écologiques particulières une analyse analogue peut être faite où chaque lignée est appréciée pour tout autre caractère (floraison, dimensions, résistances, etc...). Sur ces données l'analyse discriminante s'interrogera pour savoir si les caractères agronomiques permettent de retrouver les groupes constitués avec une bonne fiabilité et si oui, en mettant en évidence les groupes qui peuvent être identifiés par des états utiles des caractères, le sélectionneur retiendra dans la collecte mondiale les groupes d'échantillons sus-



**Fig. 21:** Attribution des individus aux différents groupes par les combinaisons linéaires des différentes fonctions discriminantes.  
(plan des 2 variables canoniques, la flèche désigne le sens de l'affectation).



**Fig. 22:** Regroupement autour des formes fortes des 60 lignées étudiées.

ceptibles d'être les plus intéressants. Il les demandera au centre de ressources génétiques et c'est sur ceux-là qu'il fera une nouvelle campagne d'évaluation agronomique.

## 5. Validité biologique des groupes constitués

Les descriptions qui viennent d'être faites n'ont pas tenu compte de connaissances biologiques obtenues indépendamment, sur un matériel aussi bien connu que des lignées de riz. Les groupes constitués correspondent réellement à des entités claires, à savoir:

- groupe 1: africain *O. breviligulata*, *O. glaberrima*
- groupe 2: asiatique *O. sativa japonica*
- groupe 3: asiatique *O. sativa indica*

Les formes difficiles à classer correspondent à des lignées ayant des propriétés particulières: l'une (issue de *O. glaberrima* est le résultat vraisemblable d'une introduction partielle à partir de *O. longistaminata* (cf. Vol. I, le riz), les deux lignées situées dans le groupe asiatique entre *japonica* et *indica* l'une correspond à une adaptation particulière en Afrique d'*O. sativa* et l'autre était caractérisée par une modification d'ordre cytoplasmique conduisant à des phénomènes de stérilité particulière.

Deux informations importantes résultent de cette classification simple: on ne distingue pas fondamentalement (en terme de distance génétique) les formes spontanées et cultivées africaines; les formes dites *javanica* ne se distinguent génétiquement guère des *japonica* et paraissent plus des *japonica* adaptées au sud-est asiatique que des formes stabilisées recombinantes intermédiaires entre *japonica* et *indica*. Nous renvoyons au Vol. I pour que le lecteur puisse utiliser ces informations dans le cadre de l'analyse du complexe d'espèces des riz.

Ces données choisies volontairement très peu nombreuses pour que le lecteur suive aisément la démarche est facilement généralisable à des ensembles plus nombreux et moins connus. Elle soulignait comment l'usage de l'outil informatique et les connaissances biologiques se soutiennent mutuellement pour atteindre une présentation facile du complexe, dans le but du meilleur service et de la meilleure communication pour que les utilisateurs soient le moins exclus de la connaissance (et donc de l'accès) aux ressources génétiques conservées.

# ANNEXE

(Extrait de la liste des descripteurs du Mil BIRP)

## FICHE D'IDENTIFICATION

### 1. DONNÉES D'ENREGISTREMENT

#### 1.1 N° d'identification

Chaque enregistrement dans une banque de ressources génétiques de Mil sera **répéré** par son numéro d'identification. Ce numéro unique, une fois attribué, ne pourra jamais être **réattribué** à une autre entrée même si l'origine préalablement enregistrée a disparu.

#### 1.2 Autre numéro

Toute numérotation attribuée par des institutions autres que la banque centrale.

#### 1.3 Autre numéro

Voir commentaire 1.2

#### 1.4 Autre numéro

Voir commentaire 1.2

#### 1.5 Nom commun et local et groupe ethnique

Nom donné par les cultivateurs à un cultivar particulier dans une région et nom du groupe ethnique.

#### 1.6 Pedigree (Généalogie)

Noms ou codes attribués par les spécialistes de la culture à une origine en définissant brièvement les parents et le schéma d'obtention génétique.

#### 1.7 Date de la dernière multiplication ou régénération (production contrôlée d'un nouveau numéro de semence).

Mois et année de la dernière récolte, exprimés avec 4 chiffres ex.: mai 1981 sera exprimé par 0581.

### 2. DONNÉES DE COLLECTE

#### 2.1 Organisme de collecte

Nom de l'organisme ayant réalisé l'expédition

#### 2.2 Source de financement de la collecte

Nom de l'organisme ayant financé la collecte

#### 2.3 Nom du ou des **collecteur(s)**

#### 2.4 Numéro de collecte

Numéro attribué au moment de la collecte écrit en alphanumérique comprenant 2 ou 3 lettres pour l'abréviation du nom des collecteurs suivis par un numéro ne dépassant pas cinq chiffres.

#### 2.5 Date de collecte

Date à laquelle une origine particulière a été collectée, exprimée en six chiffres, ex.: 5 avril 81 transcrit comme 050481

#### 2.6 Nom du donateur

Nom de la personne de l'institution ou de l'organisme ayant donné une origine particulière

#### 2.7 Origine échantillon

Origine à partir de laquelle un échantillon a été obtenu:

CC Champ de cultivateur

SC **Echantillon** de semence du cultivateur

EG **Echantillon** de grenier

EM **Echantillon** sur un marché

IN Institution

AO Autre origine

## 2.8 Type d'origine (structure génétique de l'origine)

OI Origine indigène authentique, semence originale non sélectionnée.

PN Pollinisation n fois. Origine indigène authentique non sélectionnée, mais dont la semence a été augmentée au cours de « N » pollinisations sous sac successivement.

AN Origine indigène authentique non sélectionnée dont la semence a été augmentée par n autofécondations successives.

SN Origine initialement authentique, maintenant subdivisée d'après l'observation de quelques caractères tels que ceux de la chandelle, du grain, le lot de graines ayant été augmenté au cours de multiplications.

MN Population réservoir constituée en mélangeant différents cultivars de même dénomination, ex.: différentes origines de « Zongo » sont mélangées pour constituer la population « Zongo » et les semences ont été accrues au cours de n multiplications.

LN Lignée obtenue après n autofécondations.

LS Lignée de sélectionneur, obtention expérimentale.

IN Origine inconnue, même après que des tentatives aient été faites pour une identification du type, aucune information n'était utilisable.

## 2.9 Statut «taxonomique» au moment de la collecte.

CV Variété cultivée.

CS Variété principalement **cultivée** dans laquelle on peut trouver des **shibras**\* (ou N'Douls).

FI Forme intermédiaire entre sauvage et cultivé ( **shibras** ou N'doul).

FS Forme ou espèce sauvage.

## 2.10 Pays d'origine.

Pays où l'échantillon a été pour la première fois collecté. Utiliser l'abréviation de 3 lettres telle qu'elle est définie par l'office statistique des Nations Unies.

## 2.11 Province ou état.

## 2.12 Site de prélèvement

Nombre de kilomètres et direction à partir du plus proche village ou d'un repère géographique permanent sur la carte.

---

\*Forme intermédiaire dont l'origine présumée est un croisement entre *Pennisetum* cultivé et sauvage.

2.13 Altitude

indiquée en mètres par rapport au niveau de la mer.

2.14 Latitude

Exprimée en degrés et minutes avec les suffixes N ou S pour Nord ou Sud respectivement. ex.: 12°45'N

2.15 Longitude

Exprimée en degrés et minutes avec les suffixes E ou W pour Est ou Ouest respectivement. ex.: 09°20'E.

2.16 Climat

Climat de la localité où l'origine particulière a été collectée. Utiliser le système de classification des climats mondiale de TROLL basé sur les grands groupes de classification en relation avec l'évapotranspiration potentielle.

V1 Climat tropical humide avec de 9 1/2 à 12 mois de saison des pluies sans interruption, courte forêt tropicale humide **semipervirens** et forêts de transition **semi-décidues**.

V2 Climat tropical chaud à humidité estivale avec 7 à 9 mois 1/2 de saison des pluies. Forêts et savanes herbacées humides.

V2A Climat tropical à saison humide hivernale avec 7 à 9 mois 1/2 d'humidité. Forêts de transition serai-décidues.

V3 Climat tropical humide et sec avec 4 1/2 à 7 mois de pluie. Forêts et savanes sèches.

V4 Climat tropical sec avec 2 à 4 mois 1/2 de saison de pluie. Forêts à épineux et plantes succulentes et savanes.

V4A Climat tropical sec avec mois humides d'hiver.

V5 Climat tropical **semi-désertique** et désertique avec moins de 2 mois de pluie.

2.17 Pluviométrie

Classée en 4 types d'après les moyennes annuelles enregistrées en mm.

1 moins de 450 mm

2 451-650 mm

3 651-900 mm

4 au-dessus de 900 mm

2.18 Pratiques culturales

1 irriguées

2 pluviales

3 inondées

4 en repiquage

2.19 Système de culture

1 culture pure

2 culture associée

2.20 Sol

1 Sableux, sable et alluvions

2 Alluvions et limon

3 Limoneux siliceux (argile et limon)

4 Très fortement organique

5 Autre (à spécifier)

2.21 Topographie

1 Bas-fond

- 2 Plaine inondable
- 3 Plaine de basse altitude
- 4 Vallonnée
- 5 Colline
- 6 Colline découpée
- 7 Profondément découpée
- 8 Montagneux
- 9 Autre (à spécifier)

### 2.22 Remarques

Informations complémentaires sous forme de commentaires donnés par le ou les **prospecteur(s)** et non enregistrés ailleurs.

## FICHE DE DESCRIPTION MORPHOLOGIQUE

3. OBSERVATION MORPHOLOGIQUE Stade de croissance au moment de l'observation  
 DESCRIPTEURS ET CODE
- 3.1 **Épaisseur** de la tige à la récolte  
**Épaisseur** mesurée en millimètres sans les graines entre le Sème et 4ème nœud à partir du sommet.
- 3.2 Forme de l'épi à maturité
- CYL Cylindrique: épaisseur de l'épi plus ou moins uniforme tout au long de sa longueur
- CON Conique: épaisseur maximum à la base, diminuant graduellement vers l'apex
- FUS Fusiforme: épaisseur de l'épi maximum au milieu, diminuant graduellement vers les extrémités
- MAS En massue: épaisseur maximum à l'extrémité apicale diminuant graduellement vers la base.
- CHA En forme de bougie: forme intermédiaire entre cylindrique et conique, l'épi est cylindrique sur les 3/4 de sa longueur et l'épaisseur diminue progressivement vers l'extrémité apicale.
- HAL En hauteur: épi caractéristique de type **souna-malien**. L'épaisseur maximum est à la base, elle diminue graduellement jusqu'au 2/3 de l'épi, puis **réaugmente** légèrement.
- LAN Lancéolé: intermédiaire entre fusiforme et conique. L'épaisseur maximum est proche du milieu, et l'épi s'aminçait davantage vers l'apex que vers la base.
- OBL **Ob lanceolé** par opposition au type lancéolé: l'épaisseur qui est maximum au milieu diminue davantage vers la base.
- GLB Globulaire: presque sphérique. La longueur de l'épi n'est jamais supérieure au double du diamètre.
- 3.3 Caducité de l'épillet/Comportement à maturité au battage
- 1 Égrenage spontané
  - 2 Égrenage après contact
  - 3 Non caduque et battage facile
  - 4 Non caduque et battage difficile

- 3.4 Longueur des soies Stade grain pâteux
- 1 Les soies en dessous du niveau de l'extrémité apicale de la graine
  - 5 La longueur des soies dépasse de 0 à 2 cm l'extrémité de la graine.
  - 7 Les soies dépassent le sommet de la graine de plus de 2 cm
- 3.5 Enveloppement de la graine à maturité
- 2 Grain nu
  - 5 Intermédiaire
  - 7 Enveloppé
- 3.6 Forme de la graine
- OB Obtus
  - LN Lancéolé
  - EL Elliptique
  - HG Hexagonal
  - GB Globulaire
- 3.7 Couleur de la graine
- La couleur de la graine est notée après battage, pour l'identification de la couleur de référence, voir les tableaux de couleur de Munsell.
- IV Ivoire
  - CM Crème
  - JA Jaune
  - GR Gris
  - GF Gris foncé
  - GB Gris brun
  - BR Brun
  - PO Pourpre
  - PF Pourpre foncé

## ÉVALUATION PRÉLIMINAIRE

### 4. DONNÉES DE L'ÉVALUATION AGRONOMIQUE

- 4.1 Lieu de l'évaluation
- Nom du lieu où l'évaluation a été faite.
- 4.2 Date de semis
- Jour, mois, année où le semis a eu lieu. Ex.: 15 juin 1981 sera enregistré 150681.
- 4.3 Densité de culture à maturité
- Estimation du nombre de plantes au m<sup>2</sup>
- 4.4 Vigueur précoce 18 jours après la levée
- Notée après **démariage** pour éviter l'effet dû au nombre de plantes.
- 3 Faible
  - 5 Intermédiaire
  - 7 Élevé
- 4.5 Tallage
- 4.5.1 Port du tallage à épiaison

- 3 Érigé (cylindrique)  
 5 Intermédiaire  
 7 Procombant
- 4.5.2 Nombre total de talles à maturité  
 Le nombre total d'épis mûrs ou non noté au moment de la récolte. Tout axe y compris l'axe principal portant un épi est considéré comme une talle.
- 4.5.3 Talles productives à maturité  
 C'est le nombre d'épis qui portent des graines au moment de la récolte et qui contribuent donc au rendement. Les épis immatures ne sont pas comptés.
- 4.5.4 Tallage nodal (aérien) à maturité  
 0 Absent  
 3 Modéré  
 7 Abondant
- 4.6 Délai de floraison à floraison  
 Nombre de jour compris entre la levée et la date où 50% des plantes ont fleuri. La floraison est définie par l'apparition des stigmates sur l'épi principal.
- 4.7 Étalement de floraison (entre plante) à floraison  
 FC Floraison continue de durée courte, inférieure à 7 jours.  
 FL Floraison continue longue, de durée supérieure à 7 jours.  
 FD Floraison discontinue 2 ou plusieurs groupes de floraison.
- 4.8 Synchronisation de la maturité des épis à la récolte  
 N Non synchrone  
 S Synchrone
- 4.9 Réponse de restauration de fertilité (pour la stérilité mâle cytoplasmique de type A1) à la récolte  
 1 Maintenue (non restauration)  
 2 Restauration partielle (toutes les plantes ne libèrent qu'une faible quantité de pollen)  
 3 Restauration complète  
 4 Ségrégation pour la restauration.
- 4.10 Potentiel de production fourragère en vert à floraison  
 On tient compte globalement du tallage et de l'abondance du feuillage.  
 3 Faible  
 5 Intermédiaire  
 7 Bon
- 4.11 Hauteur de la plante Stade grain pâteux  
 Mesurée en centimètres du sol à l'extrémité de l'épi.
- 4.12 Exertion de l'épi à maturité  
 Les observations sont prises sur la talle principale en centimètres. Elle est mesurée par la distance entre la ligule de la feuille paniculaire et la base de l'épi.  
 ENN Exertion négative, N en centimètres  
 EPN Exertion positive, N en centimètres
- 4.13 Mesure de l'épi  
 4.13.1 Longueur de l'épi Stade grain pâteux

Mesurée en centimètres, de la base à l'apex de l'épi sur la talle principale

- 4.13.2 **Epaisseur** de l'épi  
Diamètre maximum de l'épi non compris les soies, mesurées en millimètres
- 4.13.3 Densité de l'épi à maturité  
3 Lâche  
5 Intermédiaire  
7 Compact
- 4.14. Graines
- 4.14.1 Poids de graines par épi Après récolte  
Poids des graines en grammes à hygrométrie de 12%
- 4.14.2 Poids de graines Après récolte  
Poids de 1000 graines en grammes à 12% d'hygrométrie
- 4.14.3 Volume de la graine **Après récolte**  
Volume de 1000 graines exprimé en centimètres cubes, en plongeant les graines dans l'alcool.
- 4.15 Texture de l'albumen Après récolte  
3 Principalement cornée  
5 Partiellement cornée  
7 Farineux
- 4.16 Albumen jaune Après récolte  
0 Non  
+ Oui
- 4.17 Potentiel de production à maturité  
En considérant le nombre d'épis, leur taille et leur densité, le nombre et la taille des graines en comparaison avec un témoin.  
3 Faible  
5 Intermédiaire  
7 Élevé
- 4.18 Jugement d'ensemble sur la plante à maturité  
Note indicatrice de l'intérêt agronomique global  
3 Médiocre  
5 Moyen  
7 Bon
- 4.19 Sensibilité à la verse à maturité  
3 Faible  
5 Moyenne  
7 Élevée
- 4.20 Sensibilité à la photopériode à la floraison  
Concerne la floraison dans son aptitude à être influencée par la longueur du jour.  
3 Très sensible  
5 Partiellement sensible  
7 Insensible

## CARACTÉRISATION ET ÉVALUATION COMPLÉMENTAIRES

### 5. CARACTÉRISATION COMPLÉMENTAIRE

- 5.6.5 Type d'aristation à maturité
- 1 Mono-aristé — court
  - 3 Mono-aristé — long
  - 5 Poly-aristé — clairsemé
  - 7 Poly-aristé — dense

5.7 Epillet Stade grain pâteux

5.7.1 Couleur de la glume

- 3 Clair
- 5 Intermédiaire
- 7 Sombre

5.7.2 Couleur des anthères Avant déhiscence de l'anthere

- 1 Crème
- 3 Jaune léger
- 5 Jaune
- 7 Brun
- 9 Pourpre

5.7.3 Pigmentation des stigmates à floraison

- 1 Absence
- 2 Présence

5.7.4 Nombre de fleurs par épillet Stade sortie des stigmates  
Nombre et type de fleurs par épillet au milieu du rachis en disséquant les fleurs.

- 1 Une fleur parfaite seulement
- 3 Deux fleurs: 1 parfaite, 1 stérile
- 5 Deux fleurs: 1 parfaite, 1 mâle
- 7 Deux fleurs: 2 parfaites
- 9 Plus de deux fleurs parfaites

### 7. DONNÉES D'ÉVALUATION DE LA SENSIBILITÉ AUX MALADIES

Les notations ne devraient être faites qu'à partir d'expériences planifiées avec des variétés témoins.

7.1 Lieu de l'évaluation (nom du site)

7.2 Date de l'installation

Jour, mois, année où l'implantation de l'évaluation a été faite. ex. 15 juin 1981: 150681

7.3 Mildiou (*Sclerospora graminicola* (Sacc.) Schroet)

Stade grain pâteux

- 1 Absence de symptôme
- 2 Symptôme seulement sur les talles aériennes
- 3 Symptômes sur les talles basales mais plus de 50% des épis sont normaux
- 4 Symptômes sur toutes les talles principales et moins de 50% de talles normales.
- 5 Symptômes sur toutes les tiges au point qu'il n'y a aucun épi productif. Les plantes peuvent être mortes bien antérieurement ne laissant en place seulement que des pailles sèches ou des traces.

- 7.4 Ergot (*Claviceps fusiformis* (Lov.)) Stade grain pâteux  
 On enregistre à l'aide de dessins de référence le taux moyen en pourcentage de fleurs infectées sur 10 épis.
- 1 Pas de symptôme
  - 2 Moins de 5% de grains sont sclérotiques
  - 3 6 à 10% de grains sont sclérotiques
  - 4 11 à 20% de grains sont sclérotiques
  - 5 Plus de 20% de grains sont sclérotiques
- 7.5 Charbon (*Tolyposporium penicillariae* (Bref.)) Stade grain pâteux  
 Enregistrer le pourcentage moyen de fleurs infectées sur 10 épis inoculés, ensachés à l'aide de figures de références.
- 1 Pas de symptôme
  - 2 Moins de 5% de grains charbonneuses
  - 3 de 6% à 10% de grains charbonneuses
  - 4 de 11 à 20% de grains charbonneuses
  - 5 Plus de 20% de grains charbonneuses
- 7.6 Rouille (*Puccinia penniseti* Zimm.) à la floraison  
 Caractériser les 4 dernières feuilles lorsque 50% des pieds de la parcelle sont fleuris et indiquer si l'affection a lieu aussi sur les grains en utilisant un suffixe additionnel G si nécessaire.
- 1 Pas de symptôme
  - 2 Quelques tâches clairsemées
  - 3 Des tâches sur 10% de la surface foliaire
  - 4 Des tâches sur 11 à 25% de la surface foliaire
  - 5 Plus de 25% de la surface foliaire recouverte de tâches.

8. DONNÉES D'ÉVALUATION SUR LA SENSIBILITÉ AUX PARASITES (principalement insectes) Ces notations ne devraient être faites qu'à partir d'expérimentations appropriées avec des variétés témoin convenables.

Les attaques d'insectes sont notées de 1 à 9 où:

- 3 Dommages inférieurs à ceux montrés par le témoin
- 5 Dommages égaux à ceux montrés par le témoin
- 7 Dommages supérieurs à ceux montrés par le témoin

8.1 Site d'évaluation.

8.2 Date d'évaluation.

8.4 Mouches des tiges

8.4.1. *Atherigona apporoximata*

8.4.2. Autres (à spécifier)

8.5 Mineuses des tiges

8.5.1 *Chilo partellus*

8.5.2 *Sesamia inferens*

8.5.3 *Acigona ignefusalis*

8.5.4 Autres (à spécifier)

8.6 Chenilles velues

8.6.1 *Amsacta sp.*

8.6.2 *Estigeme lactinea*

8.6.3 Autres (à spécifier)

8.9 Sauterelles



CHAPITRE VI  
CENTRES DE RESSOURCES  
GÉNÉTIQUES  
ET FORMATION DES  
PERSONNELS DE GESTION

J. Pernès



# I. INTRODUCTION

Rares sont les plantes cultivées dont l'exploitation n'a pas largement dépassé les frontières de leurs zones d'origine. Un pays n'exploite par son agriculture qu'une minorité de plantes dont il soit géographiquement un centre d'origine ou de diversification; une grande part des plantes qui y sont cultivées sont introduites à partir d'autres continents. Cette constatation élémentaire montre à l'évidence que la plupart des ressources génétiques, dans leur entretien dynamique légué par l'histoire de la domestication des plantes, échappe aux utilisateurs. Il en résulte deux approches complémentaires: soit une organisation internationale coordonnée gérant les ressources génétiques considérées comme des biens de l'humanité, **inappropriables** par des groupes d'intérêt limités, soit chaque pays essaie de se constituer ses propres réserves (ses banques de gènes) à partir d'expéditions organisées visant des collectes systématiques, ou par des échanges. Chacune des approches a ses faiblesses: la première, toutes celles bien connues des efforts internationaux qui ne sont jamais réellement universels, toujours susceptibles d'être dominés par les groupements d'intérêt les plus puissants et les plus avancés, et forcément limités dans les moyens d'action à financement collectif; la deuxième, exclut de façon régulière, l'approvisionnement à des ressources génétiques gérées dynamiquement dans les zones de plus grande diversité et réduit l'action à des «banques de gènes» limitées à des échantillonnages ponctuels et soumis à des dérives génétiques considérables.

Cependant il faut agir et vite, l'urgence ayant abondamment été créée depuis plus de 20 ans (**FRANKEL\***, 1974; **HARLAN\*\***, 1972...), les ressources génétiques disparaissant rapidement de par la transformation des paysages (catastrophes écologiques, modification des structures agraires) et l'efficacité redoutable de l'amélioration des plantes dont les excès sont imposés par l'accroissement de la population mondiale. Force nous est cependant de constater dans chaque pays la triple raréfaction génétique des agricultures: moins d'espèces cultivées (impérialisme de quelques cultures amenées à un niveau de productivité et de mécanisation rentable), moins de variétés cultivées par espèce (malgré parfois la richesse trompeuse des catalogues **variétaux**, les variétés ne sont souvent que des doubles légèrement modifiés d'un **idéotype** unique bien ajusté aux contraintes technologiques et commerciales), moins de polymorphisme génétique interne aux variétés (pour des raisons commerciales, il est plus facile d'assurer la multiplication et la protection de structures **variétales** simples et reproductibles).

A long terme la survie de nos ressources génétiques ne viendra que du renversement de cette tendance aux triples raréfactions, renversement qui ne sera rendu possible que par de nouveaux principes d'organisation de nos sociétés, en donnant de la valeur à la diversité et à la sécurité plus qu'à la productivité. On pourrait donner des plus values commerciales aux développements de cultures nouvelles, imposer des contraintes d'inscription aux catalogues **variétaux** pour n'admettre de nouvelles variétés à un

---

\*Introduction au XIII<sup>e</sup> congrès de génétique

\*\*The **genetics of disaster**

niveau de productivité donné que si leur constitution génétique (lisible par les généalogies et les méthodes **électrophoréliques**) est suffisamment différente de celle des variétés inscrites et suffisamment polymorphe (variabilité génétique cachée derrière un phénotype convenable pour ses exploitations modernes). Il ne s'agit là que d'un premier volet des «lutttes contre les 3 raréfactions» le second, plus profond et plus efficace passera par une nouvelle délégation de la création **variétale** aux cultivateurs eux-mêmes, reconduisant et sélectionnant des variétés-populations polymorphes et originales. Les sociétés de production de semence auraient alors une importance accrue dans un rôle d'encadrement et de conseil et dans leur travail de création et d'introduction de géniteurs et de populations sources qui très rapidement sortiraient du ghetto des stations pour être sélectionnés par des « paysans experts » eux-mêmes. Ce point de vue, qui fait des ressources génétiques et de l'amélioration des plantes l'affaire de tous pourra paraître utopique à ceux qui n'ont pas eu l'occasion de s'émerveiller devant le savoir-faire et la sagacité des paysans héritiers de tous les « **domesticateurs** des plantes», qu'il s'agisse des cultivateurs traditionnels de maïs et de haricots du Mexique et du **Guatemala**, des paysans chinois **diversificateurs** des blés, créateurs du millet, du riz ou du soja, des paysans africains gérant les mils, les sorghos et de multiples légumes, etc... Pour notre part nous mettrons cependant l'espoir du côté de cette utopie; mais pour conforter «les réalistes» et travailler au présent, et au futur proche, nous décrivons les organisations internationales actuelles  les éléments nécessaires à la constitution de centres de ressources génétiques nationaux ou régionaux.

## II. ORGANISATION MONDIALE DES RESSOURCES GÉNÉTIQUES ET LES PRINCIPAUX CENTRES DE CONSERVATION

Nous ne pouvons mieux présenter le conseil international des ressources **phytogénétiques** (C.I.R.P. sigle français pour I.B.P.G.R.) qu'en reprenant l'intervention de son secrétaire exécutif J.T. WILLIAMS\*, faite à la conférence internationale sur les ressources génétiques des cultures (1981). Cet exposé esquissera l'histoire, la vocation et les orientations de ce bureau.

### A. HISTORIQUE ET VOCATION

Le conseil international des ressources **phytogénétiques** (CIRP) est une organisation scientifique internationale autonome sous l'égide du groupe international pour la recherche en agriculture (CGIAR). Le CIRP a été établi

---

\*International conference on crop genetic resources, avril 1981.

par le CGIAR en 1974 et son secrétariat exécutif est nommé par la FAO. La fonction de base du CIRP telle qu'elle a été définie par le groupe consultatif est de promouvoir un réseau international de centres de ressources génétiques pour faire avancer les collectes, conservations, documentations, évaluations et l'utilisation des ressources génétiques des plantes et par là contribuer à élever le niveau de vie et le bien-être des peuples du monde.

Le réseau global qui a été envisagé par le CIRP avait pour but principal de rendre disponible les ressources génétiques à tous les sélectionneurs du monde quand elles sont nécessaires pour leurs programmes tant actuels que futurs. Il est salutaire de rappeler que jusqu'à la décade précédente environ, les scientifiques et les sélectionneurs (pour la majorité venant de pays développés) parcouraient de façon non coordonnée les régions de diversité génétique des plantes cultivées pour remplir les stocks qui constituaient les bases génétiques de leurs programmes. Le matériel était collecté, évalué, une partie utilisée et la plupart éliminé. Les régions où l'agriculture était primitive étaient considérées comme des réservoirs inépuisables de races adaptées localement que l'on pouvait échantillonner à volonté quand le besoin s'en faisait sentir. Ce fut au cours des années 1960 que l'alarme fut donnée par nombre de groupes actifs de scientifiques intéressés à l'agriculture qui ont mis en lumière les menaces concernant les ressources génétiques des plantes. Les variétés traditionnelles commençaient à être éliminées dans de nombreuses parties du monde; on les menait à l'extinction au profit de variétés sélectionnées à haut rendement. Les populations spontanées qui sont également un matériel de base pour les programmes d'amélioration actuels étaient aussi en train de se perdre. La tendance à constituer une agriculture à haut niveau de technologie les menaçait de même. En même temps on a pris conscience de la nécessité d'élaborer l'amélioration des plantes sur des bases génétiques plus larges. Il y avait une conscience grandissante que les plantes cultivées, sélectionnées à partir d'une base génétique étroite, ne bénéficieraient pas d'une protection contre les maladies équivalente à celle conférée par la multitude des phénotypes d'une culture traditionnelle. Pour mettre un terme à cette destruction un programme global a été envisagé. Il s'agit de prendre en compte toute cette diversité génétique et d'agir pour la conserver.

Des efforts ont commencé au début des années 1970 pour traduire ce projet dans la réalité et 1974 vit la naissance du CIRP. On pourrait penser que le progrès depuis lors a été directement proportionnel aux fonds dont le CIRP disposait. Il n'en a pas été strictement ainsi cependant, car de nombreux pays et aussi un certain nombre d'organismes, ont financé parallèlement des travaux de ressources génétiques. Le Bureau, avec un secrétariat situé à la FAO, et financé par elle, agit comme un catalyseur; son travail est essentiellement un travail d'encouragement. Dans certains cas il a aidé des efforts nationaux à long terme, fondé sur des collections réunies de nombreuses années auparavant pour qu'ils deviennent une partie du programme international. Il encourage un transfert de technologie des pays développés vers les pays en voie de développement et essaie de stimuler les activités partout où elles sont nécessaires et chaque fois qu'elles peuvent être menées à bien.

Une des premières tâches du bureau a été d'établir des priorités par culture et par région. Un certain nombre de faits importants sont issus de ce premier travail. Même quand l'origine et l'évolution d'une culture particu-

lière est bien connue, l'organisation réelle de sa variation et de sa distribution dans les champs est loin d'être claire et les vitesses d'érosion génétique avancées sont fréquemment purement spéculatives. Aux vues de tels problèmes le bureau a mis en place des groupes de travail pour étudier des cultures particulières et recevoir des avis sur le déroulement des interventions en direction de ces cultures. Depuis ces premières rencontres (1976), des actions ont été entreprises ou accélérées sur trente cultures ou groupes de cultures majeures. En 1980 le bureau était capable de développer un plan global d'actions. Celui-ci sera reconsidéré chaque année si nécessaire. Le tableau 31 donne un exemple des priorités définies pour 1981.

Le blé sera pris comme exemple pour illustrer les problèmes qui sont apparus. Une lacune majeure concernait l'ampleur et le but des collections existantes. On ne connaissait pas combien de matériel spontané elles contenaient, bien que l'on puisse soupçonner cependant que la plupart des échantillons étaient des variétés récentes ou des lignées en cours de sélection. On ne connaissait pas non plus l'ampleur des duplications entre collections connues ni la diversité taxonomique des échantillons. Cependant une enquête achevée en 1980 a montré que beaucoup d'espèces n'étaient que pauvrement représentées dans beaucoup de collections majeures. En 1970 on pensait qu'il y avait plus de 250.000 échantillons dans les collections de blé alors que l'enquête a montré qu'il n'y en avait pas plus de 150.000. A partir de cette enquête et d'autres, on pouvait conclure qu'aucune culture majeure n'avait été collectée convenablement bien que certaines collections bien conçues soient en cours de réalisation pour certaines d'entre elles. Une mention spéciale doit être faite pour les grandes collections de l'Union Soviétique et des Etats-Unis. Nombre des premières collections ont été constituées par des collectionneurs et des chercheurs plutôt que pour la conservation des ressources génétiques.

Bien qu'il puisse être valable d'organiser le programme global sur une base **phytogéographique**, pour des raisons de commodités pratiques, le bureau a utilisé une approche régionale: 14 régions, chacune constituée en groupes de pays adjacents.

**TABLEAU 31:** Priorités globales par plantes.

plantes	Priorité globale 1	Priorité globale 2		Haute priorité régionale
Céréales	Blé	Sorgho Eleusine Orge	Mil à chandelles Millet Riz	Maïs Quinoa
Légumineuses	<i>Phaseolus</i> Haricots	Cacahuètes Soja Niebe  Haricot ailé	Pois chiche <i>Vigna radiata</i> <i>V. mungo</i> <i>V. aconitifolia</i> <i>V. umbellata</i>	<i>Vicia faba</i> Lentille Lupin
Racines et Tubercules	Manioc Patate douce	Pomme de terre		Igname Taro et
Oléagineux		Huile de palme Cocotier Crucifère oléagineuse		
Plantes à fibre		Coton		
Fruits farineux		Banane plantain		Jacquier
Plantes à sucre		Betteraves Canne à sucre		
Breuvages	Café	Cacao (variétés criollo)		
Fruits tropicaux et subtropicaux	Mangues	Banane Agrumes		Avocat Durion <i>Lansium</i> Ramboutan <i>Annona passiflora</i>
Légumineuses	Tomate Choux	<del>Amaranthé</del> Oignon  Concombres Aubergine	Gombo  Piment Radis	Gourde amer <i>Sechium</i> Artichaut <i>Spinacia</i> <i>Cucumis</i>
Arbres		Arbres pour le combustible et stabilisation de l'environnement		

Le tableau tient compte des efforts et des résultats acquis les années précédentes.

Dans l'ensemble l'idée d'un centre régional au service de plusieurs pays n'était pas largement acceptée, on préférait soutenir des programmes nationaux.

Depuis 1976 le bureau a organisé des missions de collectes dans de nombreuses parties du monde pour le blé, le riz, le sorgho, les mils et les millets, le maïs, les haricots, l'arachide, le vigna, les bananes, le coton, le cocotier et la betterave. Ensuite des expéditions ont été réalisées pour des programmes régionaux. En 1979, par exemple, le bureau et la FAO ont financé des missions de collectes pour les céréales dans 26 pays, pour les légumes dans 11, pour les tubercules dans 7, pour les fruits dans 5 et pour les plantes fourragères dans 5.

Lors de la dernière conférence technique MARSHALL et BROWN avaient recommandé que la stratégie d'échantillonnage pour les céréales se dirige principalement vers les allèles communs localement en tentant d'inclure dans les échantillons collectés au moins une copie de chacun des allèles qui avait une fréquence supérieure à 0,05 dans la population.

Un échantillonnage à mailles grossières est d'habitude suivi par un échantillonnage plus fin, le principe étant de collecter de 50 à 100 plantes différentes par site, sur autant de sites que possible et pour une diversité caractéristique du milieu. Cependant, les rapports ont montré que les échantillonnages n'ont été faits que le long des grands axes routiers et que l'intérêt principal était de rapporter des plantes qui paraissaient utiles plutôt qu'une variabilité génétique représentative.

En ce qui concerne la conservation des collections on a fait la distinction entre collections de base maintenue à -18° C pour la conservation à long terme et les collections actives maintenues à environ 0° C pour le stockage à moyen terme.

Lorsqu'une enquête sur les possibilités de stockage a été menée en 1975, il est apparu que 8 instituts seulement au monde avaient des chambres réfrigérées pour stocker des graines. Ce nombre était de 20 en 1978. Ce nombre a acquis ces dernières années un point tel que le bureau a été capable d'initier l'organisation d'un réseau global de collections de base pour sauvegarder à perpétuité les graines des cultures majeures. Actuellement (1981) le réseau comprend 17 centres de ressources génétiques pour 19 cultures, (cf. tableau 32). En 1985 le réseau devrait être complet pour les céréales majeures, les légumineuses à graines et les légumes. Cependant la distinction entre collection de base et collection active n'a pas encore été assimilée par le grand nombre et il n'y a pas encore de réseaux nettement développés de centres actifs associés aux collections de bases. Des moyens considérables devraient être nécessaires dans le futur pour obtenir davantage de chambres réfrigérées pour les collections de base et en accroître le personnel. En ce qui concerne la sécurité et la disponibilité des échantillons on a reconnu que toutes les collections devraient avoir au moins une duplication.

En ce qui concerne les évaluations et les informations sur les collections, on pensait il y a une décade, que dès que les collections auraient été effectuées elles seraient automatiquement évaluées et les données obtenues informatisées. Ceci n'a pas eu lieu et dans beaucoup de collections du matériel valable est insuffisamment décrit. Une des réalisations majeures a été d'élaborer et de publier des listes de descripteurs susceptibles d'être utilisées au niveau international. Actuellement environ trente cultures ont été traitées ainsi.

**TABLEAU 32:** Réseau de centres de ressources génétiques pour 19 cultures.

CÉRÉALES		
RIZ	<i>Oryza sativa - indica</i>	IRRI, Los Banos, Philippines
	<i>javanica</i> <i>japonica</i>	IRRI, Los Banos, Philippines NIAS, Tsukuba, Japon
	Formes méditerranéennes, formes d'Amérique du sud tempérée et types intermédiaires des USA (plus des doubles d'autres centres)	NSSL, Fort Collins, USA
	Espèces sauvages	IRRI, Los Banos, Philippines
	Formes africaines	IITA, Ibadan, Nigeria
BLE	Espèces cultivées	VIR, Leningrad, URSS CNR, Institut de Ressources Génétique, Bari, Italie
	Espèces sauvages de <i>Triticum</i> et <i>Aegilops</i>	Plant Germplasm Institute Univ. de Tokyo, Japon
MAIS	Matériel du Nouveau Monde	NSSL, Fort Collins, USA
	Matériel asiatique	NIAS, Tsukuba, Japon TISTR, Bangkok, Thaïlande
	Matériel européen	VIR, Leningrad, URSS Braga, Portugal (pour le matériel méditerranéen)
SORGHO	Cultivé et sauvage	NSSL Fort Collins, USA ICRISAT, Hyderabad, Inde
MILLETS	Cultivés et sauvages	NSSL, Fort Collins, USA
	<i>Pennisetum</i> spp.	PGR, Ottawa, Canada
		ICRISAT, Hyderabad, Inde
	<i>Eleusine</i> spp	ICRISAT, Hyderabad, Inde
	Millets mineurs des Indes	PGRC, Addis, Ababa, Ethiopie
	<i>Eragrostis</i> spp.	ICAR, New Delhi, Inde
	<i>Panicum miliaceum</i>	PGRC, Addis Ababa, Ethiopie
	<i>Setaria italica</i>	ICRISAT, Hyderabad, Inde
ORGE	Cultivées et sauvages (collection globale)	PGR, Ottawa, Canada
	Matériel européen	Nordic Genebank, Lund, Suède
	Matériel africain	PGRC, Addis, Ababa, Ethiopie
	Matériel asiatique	NIAS, Tsukuba, Japon
AVOINE	Cultivées et sauvages	PGR, Ottawa, Canada Nordic Genebank, Lund, Suède
CULTURES INDUSTRIELLES		
BETTERAVES SUCRIÈRES ET AUTRES		Genebank, FAL, Braunschweig Völkensrodé, FRA
LÉGUMINEUSES		
HARICOTS	Matériel du Nouveau Monde (toutes espèces mais principalement <i>P. vulgaris</i> , <i>P. coccineus</i> , <i>P. lunatus</i> )	CIAT, Cali, Colombia (double au NSSL, Fort Collins, USA)

Dans l'ensemble l'idée d'un centre régional au service de plusieurs pays n'était pas largement acceptée, on préférait soutenir des programmes nationaux.

Depuis 1976 le bureau a organisé des missions de collectes dans de nombreuses parties du monde pour le blé, le riz, le sorgho, les mils et les millets, le maïs, les haricots, l'arachide, le vigna, les bananes, le coton, le cocotier et la betterave. Ensuite des expéditions ont été réalisées pour des programmes régionaux. En 1979, par exemple, le bureau et la FAO ont financé des missions de collectes pour les céréales dans 26 pays, pour les légumes dans 11, pour les tubercules dans 7, pour les fruits dans 5 et pour les plantes fourragères dans 5.

Lors de la dernière conférence technique MARSHALL et BROWN avaient recommandé que la stratégie d'échantillonnage pour les céréales se dirige principalement vers les allèles communs' localement en tentant d'inclure dans les échantillons collectés au moins une copie de chacun des allèles qui avait une fréquence supérieure à 0,05 dans la population.

Un échantillonnage à mailles grossières est d'habitude suivi par un échantillonnage plus fin, le principe étant de collecter de 50 à 100 plantes différentes par site, sur autant de sites que possible et pour une diversité caractéristique du milieu. Cependant, les rapports ont montré que les échantillonnages n'ont été faits que le long des grands axes routiers et que l'intérêt principal était de rapporter des plantes qui paraissaient utiles plutôt qu'une variabilité génétique représentative.

En ce qui concerne la conservation des collections on a fait la distinction entre collections de base maintenue à -18° C pour la conservation à long terme et les collections actives maintenues à environ 0° C pour le stockage à moyen terme.

Lorsqu'une enquête sur les possibilités de stockage a été menée en 1975, il est apparu que 8 instituts seulement au monde avaient des chambres réfrigérées pour stocker des graines. Ce nombre était de 20 en 1978. Ce nombre a acquis ces dernières années un point tel que le bureau a été capable d'initier l'organisation d'un réseau global de collections de base pour sauvegarder à perpétuité les graines des cultures majeures. Actuellement (1981) le réseau comprend 17 centres de ressources génétiques pour 19 cultures, (cf. tableau 32). En 1985 le réseau devrait être complet pour les céréales majeures, les légumineuses à graines et les légumes. Cependant la distinction entre collection de base et collection active n'a pas encore été assimilée par le grand nombre et il n'y a pas encore de réseaux nettement développés de centres actifs associés aux collections de bases. Des moyens considérables devraient être nécessaires dans le futur pour obtenir davantage de chambres réfrigérées pour les collections de base et en accroître le personnel. En ce qui concerne la sécurité et la disponibilité des échantillons on a reconnu que toutes les collections devraient avoir au moins une duplication.

En ce qui concerne les évaluations et les informations sur les collections, on pensait il y a une décade, que dès que les collections auraient été effectuées elles seraient automatiquement évaluées et les données obtenues informatisées. Ceci n'a pas eu lieu et dans beaucoup de collections du matériel valable est insuffisamment décrit. Une des réalisations majeures a été d'élaborer et de publier des listes de descripteurs susceptibles d'être utilisées au niveau international. Actuellement environ trente cultures ont été traitées ainsi.

**TABLEAU 32:** Réseau de centres de ressources génétiques pour 19 cultures.

CÉRÉALES		
RIZ	<i>Oryza sativa</i> - <i>indica</i> <i>javanica</i> <i>japonica</i> Formes méditerranéennes, formes d'Amérique du sud tempérée et types intermédiaires des USA (plus des doubles d'autres centres) Espèces sauvages Formes africaines	IRRI, Los Banos, Philippines IRRI, Los Banos, Philippines NIAS, Tsukuba, Japon NSSL, Fort Collins, USA IRRI, Los Banos, Philippines IITA, Ibadan, Nigeria
BLE	Espèces cultivées  Espèces sauvages de <i>Triticum</i> et <i>Aegilops</i>	VIR, Leningrad, URSS CNR, Institut de Ressources Génétique, Bari, Italie Plant Germplasm Institute Univ. de Tokyo, Japon
MAIS	Matériel du Nouveau Monde Matériel asiatique  Matériel européen	NSSL, Fort Collins, USA NIAS, Tsukuba, Japon TISTR, Bangkok, Thaïlande VIR, Leningrad, URSS Braga, Portugal (pour le matériel méditerranéen)
SORGHO	Cultivé et sauvage	NSSL Fort Collins, USA ICRISAT, Hyderabad, Inde
MILLETS	Cultivés et sauvages <i>Pennisetum</i> spp.  <i>Eleusine</i> spp.  Millets mineurs des Indes <i>Eragrostis</i> spp. <i>Panicum miliaceum</i> <i>Setaria italica</i>	NSSL, Fort Collins, USA PGR, Ottawa, Canada ICRISAT, Hyderabad, Inde ICRISAT, Hyderabad, Inde PGRC, Addis, Ababa, Ethiopie ICAR, New Delhi, Inde PGRC, Addis Ababa, Ethiopie ICRISAT, Hyderabad, Inde ICRISAT, Hyderabad, Inde
ORGE	Cultivées et sauvages (collection globale) Matériel européen Matériel africain Matériel asiatique	PGR, Ottawa, Canada Nordic Genebank, Lund, Suède PGRC, Addis, Ababa, Ethiopie NIAS, Tsukuba, Japon
AVOINE	Cultivées et sauvages	PGR, Ottawa, Canada Nordic Genebank, Lund, Suède
CULTURES INDUSTRIELLES		
BETTERAVES SUCRIÈRES ET AUTRES		Genebank, FAL, Braunschweig Völknerode, FRA
LÉGUMINEUSES		
HARICOTS	Matériel du Nouveau Monde (toutes espèces mais principalement <i>P. vulgaris</i> , <i>P. coccineus</i> , <i>P. lunatus</i> )	CIAT, Cali, Colombia (double au NSSL, Fort Collins, USA)

*P. acutifolius*)

Matériel européen

Genebank, FAL, Braunschweig  
Völkeroode, RFA

Espèces sauvages

Université Gembloux, Belgique

CAJAN  
(Pois de  
Pigeon)

ICRISAT, Hyderabad, Inde

CACAHUETE

ICRISAT, Hyderabad, Inde  
INTA, Pergamino, Argentine

POIS  
CHICHE

ICRISAT, Hyderabad, Inde

NIEBE

ITA, Ibadan, Nigeria

POIS

Nordic Genebank, Lund, Suède

HARICOT

IPB, Los Banos, Philippines

AILE

TISTR, Bangkok, Thaïlande

#### TUBERCULES

POMME

DE

TERRE

Espèces sauvages et cultivées

CIP, Lima, Pérou

#### LÉGUMES

AMARANTHE Collection globale  
Collection d'Asie du Sud-Est

NSSL, Fort Collins, USA  
IPB, Los Banos, Philippines

OIGNON Collection globale  
Collection asiatique

NVRS, Wellesbourne, UK  
NIAS, Tsukuba, Japon

PIMENT Collection globale  
POIVRE, Collection globale  
Collection d'Asie du Sud-Est

CATIE, Turrialba, Costa Rica  
IVT, Wageningen, Pays-Bas  
IPB, Los Banos, Philippines

AUBERGINE Collection globale  
Collection du Nouveau Monde  
Collection d'Asie du Sud-Est

IVT, Wageningen, Pays-Bas  
NSSL, Fort Collins, USA  
IPB, Los Banos, Philippines

TOMATE Collection globale  
Collection asiatique

CATIE, Turrialba, Costa Rica  
NSSL, Fort Collins, USA  
IPB, Los Banos, Philippines

CRUCIFÈRES *Brassica oleracea*

NVRS, Wellesbourne, UK  
IVT, Wageningen, Pays-bas  
NVRS, Wellesbourne, UK

Légumes et fourrages:

*B. campestris*, *B. juncea*,

*B. napus*

Légumes et fourrages

*B. napus*

Crucifères, Colza, etc...

*B. campestris*, *B. juncea*,

*B. napus*,

*Sinapis alba*, *B. carinata*

Genebank, FAL, Braunschweig  
Völkeroode, RFA

PRG, Ottawa, Canada

Genebank, FAL, Braunschweig

Völkeroode, RFA

PGRC, Addis Ababa, Ethiopie

Genebank, FAL, Braunschweig

Völkeroode, RFA

NVRS, Wellesbourne, UK

*Raphanus*

Parents sauvages

Univ. **Politechnique**, Madrid,  
ESPAGNE

Collection de l'Asie de l'Est

Tohoku Univ. Sendai, Japon  
NIAS, **Tsukuba**, Japon

AUTRES  
LÉGUMES

Espèces de l'Asie du Sud-Est

IPB, Los Banos, Philippines

En ce qui concerne l'information nécessaire aux centres de ressources génétiques une distinction doit être faite maintenant entre les données d'enregistrement (**passport data**) qui permettent l'identification des échantillons, les données de description (**characterization data**), se référant aux caractères **hautement héritable**s qui s'expriment facilement et peuvent être facilement reconnus dans tous les milieux, et les données d'évaluation préliminaire qui incluent un certain nombre de caractères additionnels jugés souhaitables pour des cultures particulières. Au-delà de ce stade l'évaluation est clairement la tâche des sélectionneurs.

En ce qui concerne le stockage de données le bureau considère que n'importe quel système de base de données pourrait être utilisé pour s'adapter aux exigences locales. L'échange des données entre banques de gènes pourrait avoir lieu par «listing», bandes ou disquettes, la seule exigence étant qu'elles puissent être lisibles ou facilement transcrites sous forme lisible par celui qui les reçoit.

En ce qui concerne la formation dans le domaine des ressources génétiques le bureau devrait continuer, au moins jusqu'en 1985, à financer les cours sur la conservation des ressources génétiques de plantes initiés par le professeur **HAWKES** à l'université de Birmingham (Angleterre).

Le bureau a aussi organisé des cours d'enseignement technique tels que l'identification des espèces de blé, la technologie de semences par des responsables de banques de gènes et les méthodes de collectes.

## **B. RECOMMANDATIONS DE LA CONFÉRENCE TECHNIQUE (ROME, 1981)**

Nous ne mentionnons à titre d'illustration que quelques points particuliers.

### **1. Collectes**

Le **CIRP** demande à tous les organismes de faire en sorte que les collectes d'espèces locales en danger et de variétés traditionnelles soient toujours une activité propre au projet de développement et d'amélioration des cultures.

De développer davantage de missions de collectes des parents sauvages des variétés cultivées.

Que les collectes dans les plantations mixtes ou dans les systèmes des cultures associées soient faites de telle sorte qu'elles permettent la conservation des combinaisons intéressantes.

Que l'on développe une gamme de techniques de collecte pour faire face aux besoins des prospecteurs car des techniques d'échantillonnages différentes doivent être utilisées pour des cultures différentes et dans des milieux différents.

## 2. Cultures spéciales

Encourager des programmes concernant des espèces d'intérêts particuliers tels que les plantes traditionnelles et médicinales.

## 3. Travaux de Gestion

### a. *Conservation et régénération*

- Des chambres froides additionnelles devraient être équipées pour renforcer le réseau international.
  - Le CIRP devrait soutenir des études pour déterminer les principes permettant de développer des méthodes standard de régénération particulièrement pour les cultures tropicales et les espèces allogames.
  - Le CIRP initie une enquête sur la dormance des semences des ancêtres sauvages des plantes cultivées et sur les techniques nécessaires pour la lever.

b. *Conservation in vitro*. Pour permettre l'utilisation des techniques in vitro des conservations la recherche devrait être intensifiée sur les points suivants:

- amélioration des techniques spécifiques pour les cultures pour lesquelles la propagation in vitro a été développée à un degré tel qu'il puisse être réaliste maintenant d'essayer d'appliquer ces techniques au matériel des banques de gènes,
- des études de base sur les cultures pour lesquelles il y a eu peu de succès,
- cryopréservation pour tout type de matériel végétal afin d'en établir les premiers principes.

c. *Évaluation et utilisation*. Il faudrait activer les travaux de caractérisation et d'évaluation dans les banques de gènes et transmettre les découvertes aux utilisateurs potentiels aussi rapidement que possible.

Le CIRP devrait stimuler des travaux destinés à transférer des caractères utiles des espèces sauvages vers les lignées de sélections pour développer l'utilisation par les sélectionneurs de ces caractères utiles.

d. *Documentation et informatisation*. L'accent doit être mis davantage sur l'amélioration des échanges d'information entre les centres de ressources génétiques et sur le développement de l'information en retour provenant des utilisateurs de ressources génétiques des plantes.

## 4. Quarantaine

Tous les échanges devraient avoir lieu par le canal du service national des quarantaines.

Des laboratoires nationaux ou régionaux devraient être créés par les gouvernements pour accélérer les passages dans les quarantaines.

Il faudrait envisager des investigations des instituts régionaux de recherche pour l'étude des pathogènes et des parasites portés par les constituants des ressources génétiques, y compris les espèces sauvages et les formes spontanées apparentées aux cultivars.

Des initiatives de recherches devraient être prises pour utiliser les techniques de culture in vitro, pour assainir les plantes échangées pour qu'elles puissent répondre aux exigences de quarantaine particulièrement en ce qui concerne les virus.

## C. ORGANISATION

Le CIRP organise ses groupes de consultants:

soit par comités spécifiques régionaux:

- Programme Sud-ouest Asiatique,
- Programme Sud-est Asiatique,
- Programme méditerranéen,
- Programme Afrique de l'Ouest;

soit par comités spécifiques par cultures:

- Maïs,
- Phaseolus,
- Riz,
- Sorgho, mil, millet,
- Blé.

Des comités particuliers peuvent être mis en place temporairement:

• Comités spéciaux pour le stockage des graines, Groupes de travail et de consultation pour les cultures particulières:

- Orge,
- Agrumes,
- Canne à sucre,
- Vigna,

Enfin le financement international en 1981 était acquis à partir des contributions directes des six pays faisant l'objet du tableau 33.

## III. SCHÉMAS D'ORGANISATION DE CENTRES DE RESSOURCES GÉNÉTIQUES

L'organisation des centres de ressources génétiques est conçue selon approximativement deux tendances: l'une à vocation centralisatrice très lourde et établie autour de laboratoires ayant des moyens propres importants, particulièrement en ce qui concerne les chambres de conservation des semences, le traitement informatique des données et les terrains d'expérimentation agronomique; l'autre tente de déléguer le plus possible les conservations et les évaluations dans le cadre paysan, dans des ré-

serveurs ou parcs nationaux ou régionaux, dans chaque station d'amélioration des plantes. Le seul élément de centralisation est alors un « bureau des ressources génétiques » assurant la coordination des informations en veillant à la qualité des travaux dont les responsabilités ont été déléguées.

**TABLEAU 33:** Contributions des différents pays au financement de l'IBPGR en 1981

	Équivalence en US \$
Allemagne	133.215.00
Australie	81.473.00
Banque Mondiale (IBRD)	200.003.75
Belgique	93.030.47
Canada	150.100.20
Danemark	47.789.73
Espagne	49.908.50
Etats-Unis d'Amérique	800.000.00
France	66.659.72
Grande-Bretagne	259.138.00
Italie	50.000.00
Japon	500.000.00
Norvège	107.223.48
Pays-Bas	185.000.00
Suède	<u>175.342.46</u>
TOTAL	2.898.884.31

Les centres s'organisant selon le premier type sont les plus nombreux, tels l'IRRI (conservation internationale des riz), l'institut VAVILOV (URSS) pour toutes les cultures, l'institut de TSUKUBA (Japon) à vocation multiple également, le centre régional pour les pays méditerranéens de BARI (Italie).

La banque de gènes du Nord (pays scandinaves) s'oriente plutôt vers le 2ème type; par la force des choses les quelques entreprises ponctuelles initiées en France seraient aussi de ce type, ainsi que le système de multiplication des ressources génétiques utilisé en Hongrie.

L'organisation chinoise, à l'échelle pratiquement d'un continent, vise une sorte de compromis entre bureau central et centres régionaux importants (provinces), avec un site de conservation à long terme central ( SINKIANG); le réseau A.C.C.T. se propose de soutenir l'effort de chaque pays tout en assurant une coordination et des facilités pour la gestion des données, l'effort des collectes et les échanges.

Illustrons ces types d'organisation.

## A. ORGANISATION CENTRALE

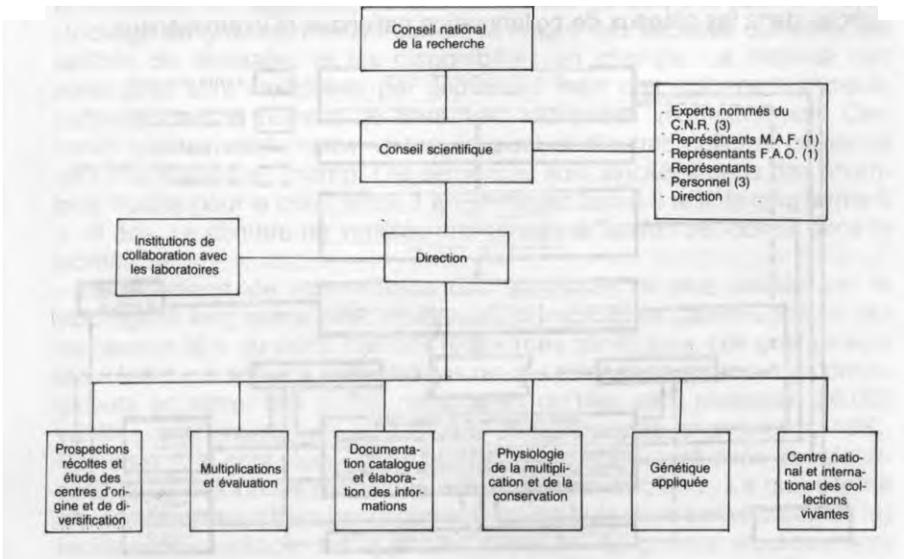
### 1. Centre de ressources génétiques de BARI

En Italie, en automne 69, le conseil national de la recherche a proposé la transformation de l'institut de BARI en un centre national de la recherche pour la conservation des ressources génétiques des espèces cultivées qui

intéressent la région méditerranéenne, et particulièrement les cultures majeures en Italie, organiser des expéditions de collectes dans les zones d'origine primaires et dans les centres de diversification pour évaluer et pour distinguer les matériels récoltés et aussi organiser des programmes de recherches.

a. *L'organisation.* Elle est représentée dans le schéma 19 a, b. Le CNR BARI est une institution du centre de la recherche dotée d'un personnel approprié, destinée à définir les projets de recherches, à contribuer à la formation du personnel scientifique et technique et à développer des rapports de collaboration avec des institutions scientifiques italiennes et étrangères. Le laboratoire est dirigé par un directeur et un conseil scientifique, et les activités de recherche sont distribuées à travers divers départements. Le conseil scientifique est composé d'experts nommés par le centre national de la recherche, le ministère de l'agriculture et des forêts, de représentants de la FAO, de représentants du personnel. Le personnel permanent comprend actuellement 8 docteurs, 6 en agriculture et 2 en biologie, 8 techniciens supérieurs (agriculture, chimie), des auxiliaires techniques et administratifs. Le laboratoire est implanté aux abords de l'université d'agriculture de Bari et possède également des laboratoires en propre comprenant des instruments nécessaires à la recherche (analyseurs d'azote et d'acides aminés, microscopes, spectrophotomètres, scintillateurs, centrifugeuses, chromatographes en phase gazeuse...).

Le laboratoire de ressources génétiques dispose de moyens de conservation. D'une part des chambres de conservation pour des périodes inférieures à 10 années (1°C, 30% d'hygrométrie), pour stocker le matériel destiné à la multiplication, l'évaluation et la distribution aux sélectionneurs. D'autre part, pour la conservation à long terme le laboratoire dispose d'autres chambres à dimensions plus réduites conditionnées à -10°C et 30% d'hygrométrie relative. Pour l'observation, l'étude et le renouvellement des collections le laboratoire, en collaboration avec la faculté d'agriculture,



**Schéma 19a:** Organigramme des laboratoires du CNR BARI

dispose de serres de champs d'expériences à BARI et en d'autres points de la péninsule.

Pour mieux garantir la sauvegarde des espèces cultivées d'intérêt pour la zone méditerranéenne et d'importance particulière pour l'agriculture italienne et pour en assurer une plus grande utilisation, le laboratoire publie régulièrement des rapports concernant aussi les résultats de toutes études intéressant l'amélioration d'une espèce donnée et décrit les résultats concernant les méthodes de cultures, de multiplication et d'évaluation les plus appropriées pour les conditions locales.

b. *But et activités du laboratoire.* Les activités principales du laboratoire sont l'exploration, la récolte des ressources génétiques, l'étude des centres de diversification des plantes concernant le laboratoire, la multiplication et l'évaluation du matériel mis en collection, l'étude des processus physiologiques dans les semences au cours de la conservation, enfin la détermination de l'ampleur et de la nature de la variabilité génétique des caractères agronomiquement intéressants.

Les collections de ressources génétiques constituées jusqu'à présent par le laboratoire comprennent des semences de formes spontanées ou des variétés anciennes, des variétés locales, des écotypes obtenus au moyen d'échange avec des institutions similaires ou directement récoltés par le laboratoire au cours de voyages et d'expéditions. De telles explorations ont eu lieu en Italie méridionale et insulaire, en Afrique du Nord avec la participation de la FAO, en Ethiopie. Comme conséquence de ces activités, le laboratoire conserve une des plus grandes collections de blés tétra-ploïdes sauvages et cultivés (environ 6.000 échantillons, provenant d'URSS, USA, Argentine, Kenya, Allemagne, Israël, France, Italie ou récoltés par le laboratoire en Ethiopie, Algérie et Italie), une ample collection de blés hexaploïdes (environ 8.000 échantillons), de vesse (environ 1.500 origines dont la plus grande partie vient d'Italie et de Turquie).

Les objectifs actuels du laboratoire sont d'augmenter le nombre d'espèces dans les réseaux de collaboration nationaux et internationaux.

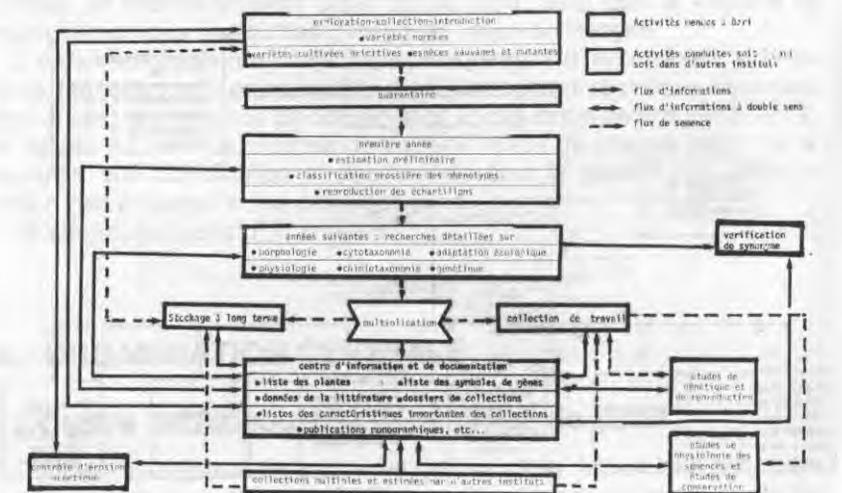


Schéma 19b: Organisation des activités du CNR BARI

## 2. La banque de ressources génétiques de TAIWAN

Les informations résumées ici concernent les résultats d'une enquête relative aux organisations de maintenance des semences, au nombre de variétés introduites et maintenues, aux méthodes de collectes, aux objectifs de maintenance, à la quantité minimum des semences stockées, aux méthodes de stockages, à la viabilité des semences et aux systèmes d'enregistrement informatique. Au cours des dix dernières années passées 432 espèces correspondant à un total de 11.507 lignées ont été introduites pour les espèces telles que le riz, diverses plantes vivrières, des cultures industrielles, des plantes médicinales, fourragères, horticoles, etc... Le soja, l'arachide, le maïs, la pomme de terre, diverses plantes vivrières collectées à Taiwan, représentent 38,5% (4.469 variétés) des introductions totales suivies par le riz 13,7%, les cultures industrielles 7,7%, les légumes 4,8% et diverses autres cultures 2,2%. A Taiwan les stocks génétiques sont obtenus par le canal du centre international d'échange des semences de TARI (Taiwan Agricultural Research Institute) ou directement à partir d'organisations de recherche des pays étrangers ou encore par échange de stocks génétiques entre différentes organisations de Taiwan, soit, enfin à partir de collectes directes aux champs. 65,2% (31.083 variétés) ont été directement introduites à partir de différents pays du monde et 23,6% (11.225 variétés) ont été transférées entre organisations de Taiwan.

*a. Organisation assurant la maintenance des ressources génétiques.* La plupart des organisations de recherches agricoles traitant de l'amélioration des cultures maintiennent une certaine quantité de stocks génétiques. Une vingtaine d'organisations de recherche à Taiwan constitue ainsi les institutions majeures de maintenance des ressources génétiques. 13 organisations maintiennent le riz, 7 les principaux légumes et les arbres fruitiers, 6 le soja. Les plantes textiles, le thé et le tabac sont chacun maintenu par une seule organisation.

*b. Méthodes de conservations des ressources génétiques.* Les durées de stockage de graines diffèrent suivant la nature des espèces cultivées, les facilités de stockage, et les disponibilités en champs. La majorité des ressources sont multipliées par semences mais des cultures comme la patate douce, la pomme de terre sont multipliées *végétativement*. Certaines plantes médicinales, le thé, le *cisal* et d'autres cultures pérennes sont maintenues au champ. Les semences sont stockées dans des chambres froides pour le court terme 1 an, le moyen terme 3 ans, le long terme 5 à 10 ans. Le nombre de variétés maintenues à Taiwan est donné dans le tableau 34.

La technique de maintenance des semences la plus utilisée est le stockage à long terme avec multiplication fractionnée (tableau 35) ce qui représente 68% du stock total des ressources génétiques. Les graines sont stockées dans des chambres froides ou des réfrigérateurs dans des *dessiccateurs* en verre, des boîtes métalliques ou des sacs plastique. 26.000 variétés sont maintenues à 5°C et 4.300 dans des chambres à 10°C. Moins de 10% sont stockées en réfrigérateurs et 6% sont dans des *dessiccateurs* ou des boîtes métallique à température ambiante. La quantité de graines stockées diffère beaucoup suivant les buts de la conservation et les facilités de stockage, les quantités minimum de graines stockées sont données dans le tableau 36.

**TABLEAU 34:** Nombre des variétés ou lignées de ressources génétiques maintenues par 20 organisations.

Principales ressources génétiques	Nombre total de variétés	
Riz	6.520	13.68
Arachide	965	2.02
Patate douce	1.004	2.10
Pomme de terre	930	2.95
Soja	14.004	29.38
Blé, Orge, Avoine, triticale	2.270	4.76
Sorgho	209	0.43
Colza	33	0.06
Mais	1.219	2.55
Plantes à fibre	2.533	5.32
Plantes médicinales	253	0.53
Plantes de couvertures, engrais vert, fourrages	225	0.47
Sesame	48	0.10
Autres cultures spéciales	36	0.07
Tabac	492	1.03
Thé	460	0.96
Tournesol	202	0.42
Haricot mungo	5.010	10.51
Haricot ajuki	229	0.48
Haricot ailé	41	0.08
Fève commune	20	0.04
Niebe	154	0.32
Autres légumineuses	105	0.22
Tomate	4.342	9.21
Légumes	627	13.04
Raisin	130	0.27
TOTAL	47.651	100.00

c. *Enregistrement et système de traitement des données des ressources génétiques.* Un système d'enregistrement et de traitement des données est essentiel pour l'utilisation des ressources génétiques. Au cours des années passées on a utilisé la méthode conventionnelle d'enregistrement par fichiers-cartes ou en stockant les informations dans des catalogues de semences. Depuis 1978 le système de traitement de données informatisé par les banques de gènes a été adopté par TARI. Un système de codage alphanumérique est maintenant utilisé. Cependant on est en train d'organiser le changement vers un système de codage décimal, pour faciliter, dans un futur proche, la comparaison des données imprimées avec celles d'autres organisations. Les informations suivantes sont incluses:

**TABLEAU 35: Méthodes** de maintenance de ressources génétiques à Taiwan.

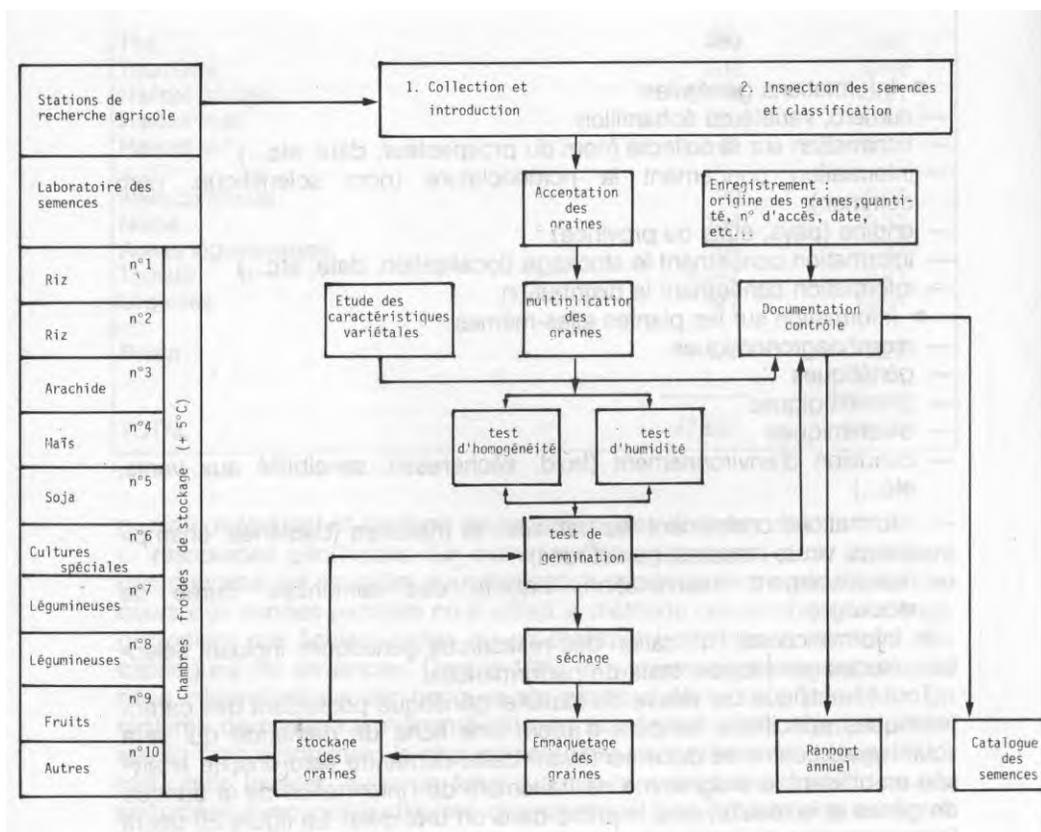
Méthodes	Nombre de variétés ou de lignées	%
1. multiplication des semences chaque année	7.311	16.6
2. multiplication des semences tous les deux ans	5.499	12.5
3. multiplication des semences tous les trois ans	6.751	15.3
4. multiplication annuelle et stockage à long terme	20.667	46.8
5. propagation végétative	2.438	5.5
6. Entretien de plantes vivantes (vergers)	89	0.2
7. stockage à long terme (4-5 ans)	1.366	3.1
TOTAL	44.121	100.0

- Informations générales:
  - numéro, variété ou échantillon
  - information sur la collecte (nom du prospecteur, date, etc...)
  - information concernant la nomenclature (nom scientifique, nom commun)
  - origine (pays, états ou province)
  - information concernant le stockage (localisation, date, etc...)
  - information concernant la distribution.
- Information sur les plantes elles-mêmes:
  - **morphoagronomiques**
    - génétiques
    - physiologiques
    - biochimiques
    - condition d'environnement (froid, sécheresse, sensibilité aux vents, etc...)
    - informations concernant les parasites et maladies (bactéries, champignons, virus, insectes, nématodes)
    - rajeunissement (**réjuvénation**, viabilité des semences, durée de stockage).
  - Information sur l'utilisation des ressources génétiques incluant sélection, études génétiques, tests de performances.

Tout scientifique qui désire du matériel génétique possédant des caractéristiques spécifiées remplira d'abord une fiche de demande qui sera soumise au centre de documentation. Cette demande sera ensuite analysée en utilisant le programme de traitement de l'information de la banque de gènes et le résultat sera imprimé dans un bref délai. La figure 20 décrit l'organigramme de fonctionnement de la banque de gènes du TARI.

**TABLEAU 36:** Quantité minimum de graines stockées

Quantité de graines	Nombre de variétés ou de lignées	%
1. moins de 5.g.	174	0.6
2. moins de 10 g.	1.502	5.0
3. moins de 50 g.	8.986	30.7
4. moins de 100 g.	4.586	15.7
5. moins de 200 g.	823	2.8
6. plus de 200 g.	1.674	5.7
7. Autres	11.493	39.3
TOTAL	29.238	100.0



**Fig. 23:** Organigramme du fonctionnement de la banque de gènes de TARI.

### **3. Organisation de l'enregistrement et de la multiplication des semences à l'institut de TSUKUBA (Japon).**

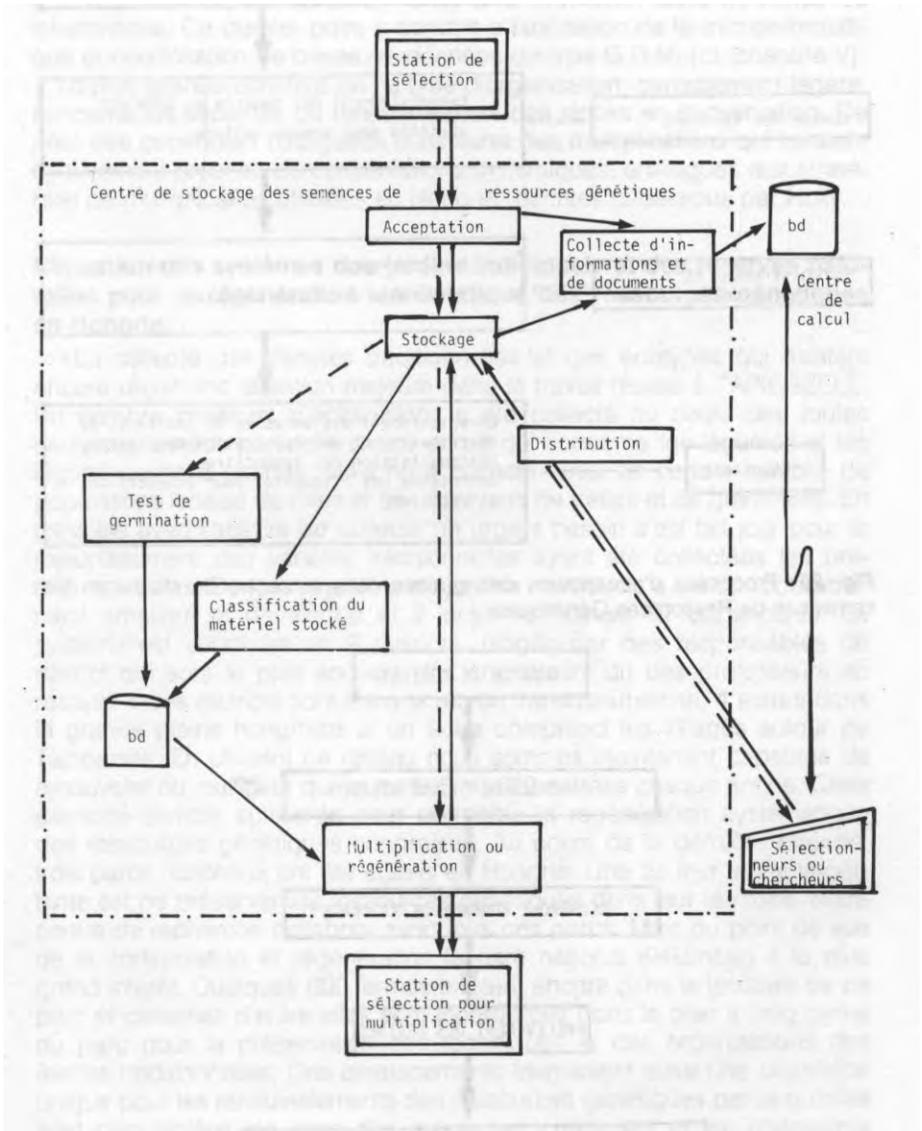
La gestion des ressources génétiques des principales plantes à graines cultivées est officiellement prise en charge par le G.S.S.C. de Tsukuba (Germplasm Seed Storage Center). La figure 24 donne les analyses de routine des graines destinées à la conservation des ressources génétiques qu'il s'agisse de nouvelles variétés, de lignées produites dans les stations de sélection, ou d'échantillons introduits ou collectés. En général elles ont été évaluées pour des caractères agronomiques importants dans chaque station, mais parfois les données sont insuffisantes pour qu'elles puissent être enregistrées comme informations de ressources génétiques. L'enregistrement systématique de données avec une standardisation des descripteurs et de leurs états est nécessaire et a été discuté récemment pour plusieurs espèces. La figure 25 montre le processus d'acceptation des graines avec les différents enregistrements réalisés au moment de la réception. La base de données est mise à jour grâce à l'information contenue dans la carte de format spécial attachée à tout lot de graines introduit et on imprime des listes de variétés classées dans l'ordre d'enregistrement, dans l'ordre alphabétique des variétés en fonction de l'année de récolte ou de la localité. Certaines de ces listes sont utilisées au moment de la distribution de semences, des tests de germination ou de la multiplication des semences. Les étapes des actions entreprises pour répondre aux demandes de graines venant des chercheurs ou des sélectionneurs sont décrites dans la figure 26. Après expédition la base de données est corrigée et des listes de statistiques montrant le mode de distribution sont imprimées quand c'est nécessaire. La base de données a aussi été créée séparément pour chaque espèce cultivée et comprend actuellement plusieurs milliers d'entrées. Cependant elle couvrira plus de 30.000 entrées de données en stockage et sera ouverte à tous les chercheurs ou sélectionneurs du ministère de l'agriculture, des forêts et des pêches.

## **B. ORGANISATIONS EN RÉSEAU**

Ce type d'organisation est imposé lorsque plusieurs pays d'une même région sont intéressés par la conservation des ressources génétiques des mêmes plantes. C'est le cas de la banque de gènes de la région du nord qui associe le Danemark, la Finlande, l'Islande, la Norvège et la Suède. L'intérêt de cet exemple réside dans les consignes proposées pour la mise en place toute récente de cette banque (création 1979, non encore complètement opérationnelle en 1982).

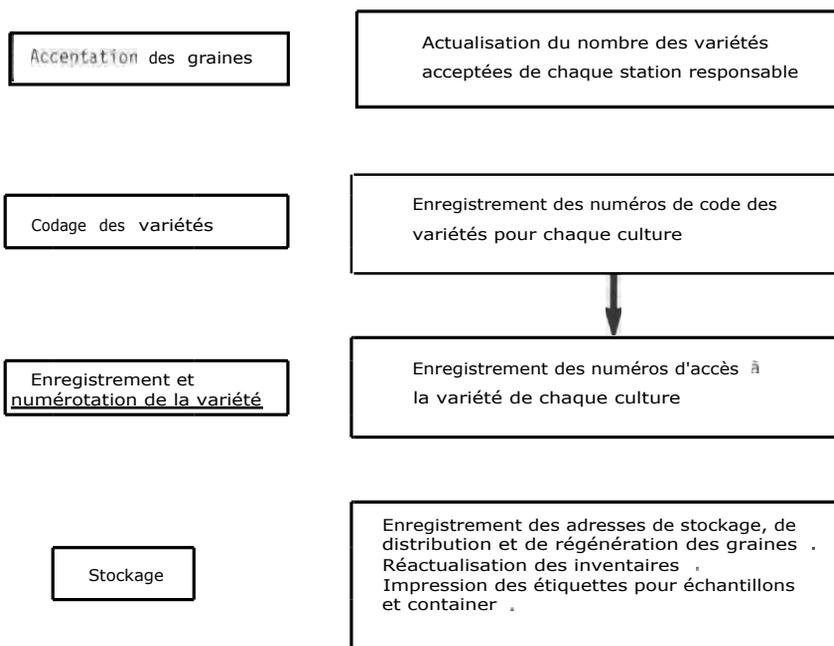
Le but de la banque de gènes du Nord est de conserver et de décrire la variation génétique des matériels agricoles et horticoles qu'ils soient cultivés ou spontanés et qui correspondent aux zones climatiques des contrées nordiques. La banque de gènes sera une institution de service pour les sélectionneurs et autres spécialistes des plantes des pays du Nord et sera en étroite collaboration avec ces groupes. Par principe le matériel stocké et les informations seront librement utilisables. La banque de gènes

n'assurera pas le stockage de toutes les collections, certains types de plantes seront maintenus ailleurs et seules les données informatisées seront gérées par la banque de gènes. La capacité de stockage de la banque de gènes est prévue pour être de l'ordre de 50.000 échantillons, chaque échantillon comprenant 15 à 20.000 graines. Il y aura des chambres de stockage pour des matériels venant d'autres banques de gènes, aussi bien que des possibilités de développement futur si besoin est. La température dans les chambres froides sera comprise entre  $-18^{\circ}\text{C}$  et  $-20^{\circ}\text{C}$ . Tous les échantillons seront **desséchés** jusqu'au degré hygrométrique requis dans une salle de séchage spéciale et seront stockés dans des conteneurs étanches. Les chambres froides de la banque de gènes devaient être mises en service en 1981. Les tests de viabilité de tous les échantillons de semences seront établis en **connection** avec le stockage du matériel. Plus tard des tests de viabilité périodiques seront effectués tous les 4 ou 5 ans. Environ 10.000 tests de viabilité seront effectués chaque année. La banque de gènes ne disposera pas de laboratoire propre pour ces tests de viabilité. Ces tests seront par exemple réalisés par l'institut suédois de tests et de certifications de semences contre rémunération. Quand la viabilité de la semence des échantillons ou quand la quantité de semences impose un renouvellement ou un accroissement, celui-ci sera effectué directement à la banque de gènes ou dans tout autre institut du Nord convenablement situé. Ce travail a été programmé très soigneusement de façon à éviter les dérives génétiques, les croisements entre populations et diminuer les attaques parasitaires. Toute l'information utilisable concernant l'origine, les caractéristiques, etc... des échantillons individuels sera organisée d'une façon uniforme et compréhensible. Le système de documentation sera applicable **internationalement**.

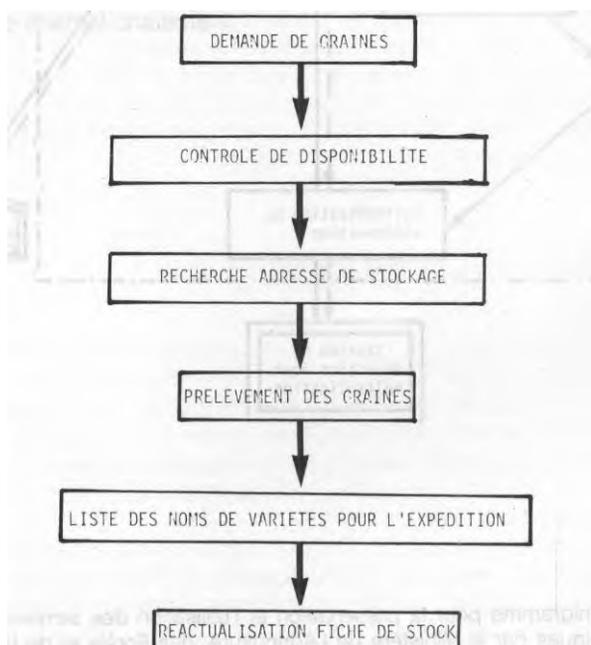


**Fig. 24:** Organigramme pour la préservation et l'utilisation des semences de ressources génétiques par le Ministère de l'Agriculture, des Forêts et de la Pêche au Japon.

Circuit des informations ———→  
 Circuit des semences. - - - - -→



**Fig. 25:** Processus d'acceptation des graines dans le centre de stockage des semences de Ressources Génétiques.



**Fig. 26:** Enregistrement des demandes de graines.

Les priorités retenues sont les délégations de conservation dans toute institution de la région en limitant le plus possible les conservations centralisées, l'ampleur des conservations dynamiques sous forme de réserves (particulièrement pour les petits fruits) et la circulation facile de toutes les informations. Ce dernier point a conduit à l'utilisation de la microinformatique et l'exploitation de bases de données du type G.D.M. (cf. chapitre V).

La plus grande difficulté de ce type d'organisation, centralement légère, concerne les sécurités du renouvellement des stocks en conservation. Ce peut être cependant l'obligation d'instaurer des multiplications qui seraient relativement proches de conservations dynamiques, analogues aux stratégies de multiplication utilisées en Hongrie, décrites ci-dessous par Holly.

### **Utilisation des systèmes des jardins individuels et des réserves naturelles pour la régénération isoclimatique des ressources génétiques en Hongrie.**

«La collecte des variétés traditionnelles et des écotypes qui existent encore reçoit une attention majeure dans le travail réalisé à TAPIOSZELE. Un nombre croissant d'échantillons a été collecté au cours des toutes dernières années particulièrement en ce qui concerne les légumes et les légumineuses à graines. Mais on a collecté aussi un certain nombre de populations locales de maïs et des écotypes de trèfles et de graminées. En parallèle avec l'activité de collecte un urgent besoin s'est fait jour pour le rajeunissement des variétés traditionnelles ayant été collectées les premières années. Donc un système de jardins individuels a été considérablement amélioré et développé et il inclut maintenant 87 participants. Le système est subdivisé en 9 districts, dirigés par des responsables de district qui sont le plus souvent des chercheurs ou des professeurs en retraite. 4 des districts sont dans la région transdanubienne, 4 autres dans la grande plaine hongroise et un autre comprend les villages autour de Taposzele. En utilisant ce réseau nous sommes maintenant capables de renouveler ou multiplier quelques 500 ou 600 entrées chaque année. Cette capacité semble suffisante pour permettre la régénération systématique des ressources génétiques hongroises. Au cours de la dernière décennie, trois parcs nationaux ont été établis en Hongrie. Une de leur tâche importante est de préserver les ressources génétiques dans leur territoire. Notre centre de recherche collabore avec tous ces parcs. Mais du point de vue de la conservation et régénération le parc national Kiskunság a le plus grand intérêt. Quelques 600 fermes existent encore dans le territoire de ce parc et certaines d'entre elles sont incorporées dans le plan à long terme du parc pour la préservation des techniques et des organisations des fermes traditionnelles. Ces emplacements fournissent aussi une possibilité unique pour les renouvellements des ressources génétiques parce qu'elles sont bien isolées les unes des autres territorialement et les traitements chimiques (c'est-à-dire l'application des pesticides et des fertilisants) sont strictement réglementés même au voisinage dans ce qui est appelé les zones tampons. Ces ressources naturelles peuvent servir deux objectifs principaux: renouveler et multiplier les variétés traditionnelles, les vieilles variétés améliorées et sélections locales qui sont originaires de conditions écologiques analogues; conduire des expériences pour comparer les effets des techniques de conservation dynamique et statique sur la structure génétique de certaines populations cultivées variables.

Nous n'avons que trois années d'expérience dans notre collaboration avec les parcs nationaux et d'autres réserves naturelles, mais il semble très clair qu'ils peuvent mettre à notre disposition les sortes d'habitats cultivés extrêmes qui disparaissent rapidement sur les terres cultivées par les techniques modernes. Les parcs nationaux peuvent donc nous aider à introduire une plus grande diversité écologique dans notre système de renouvellement des ressources génétiques et il pourrait en résulter un maintien plus efficace de la diversité génétique de nos collections de variétés **traditionnelles\***. » (d'après **HOLLY**, 1981).

Ces approches qui n'ont pas la séduction (trompeuse?) de gros centres lourds, seront peut-être les seules capables de faire que la gestion des ressources génétiques soit l'affaire du plus grand nombre possible; ainsi des vergers conservatoires bien recensés coordonnant des bonnes volontés» non spécialistes (amateurs) ne doivent pas être négligés et les parcs nationaux et régionaux doivent être non seulement des réserves naturelles mais aussi des vitrines suscitant l'intérêt pour des réalisations très dispersées de recherche et d'entretien de variétés traditionnelles. En France de petites amorces sont ainsi constituées par des travaux sur la conservation des pommiers de la région des landes et les **expérimentations** sur les millets, à l'écomusée de **Marquèze** (parc régional des landes de Gascogne).

## **IV GRANDS ÉLÉMENTS D'ORGANISATION ET VOCATIONS DES CENTRES DE RESSOURCES GÉNÉTIQUES**

Les exemples précédents ont illustré les éléments de base d'organisation des Centres de Ressources Génétiques dont l'organisation dépend des tâches qui leur incombent, les responsabilités locales ou centrales devant être définies pour susciter les moyens nécessaires:

- Collectes et introductions, planification des collectes,
- Quarantaines,
- Conservation: statique accompagnée de semis (rajeunissement des stocks) ou de plantations périodiques et dynamique (déléguée et décentralisée)
- Evaluation
- Enregistrement et documentation
- **Echanges** de matériel et renouvellement
- Formation des responsables,

---

\*A une question posée à Monsieur **HOLLY** par Sir O. **FRANKEL** sur la maintenance des populations de maïs en Hongrie, Monsieur **HOLLY** répondit que les parcelles de maïs comprenaient 4 à 5.000 plantes en culture associée avec 400 cucurbitacées. La plupart des races de pays des autres céréales ont disparu du fait de l'utilisation de variétés améliorées, mais le maïs avait été collecté comme tel il y a plus de 20 ans et régénéré. Certains types locaux pourraient encore être trouvés dans des habitats extrêmes aux sols très pauvres.

- Connexion avec des laboratoires spécialisés,
- Connexion des centres entre eux: réseaux nationaux et régionaux.

Les laboratoires spécialisés seront connectés avec ce centre et graviteront autour des 4 éléments de base:

- Chambres de conservations, chambres froides et de cultures in vitro,
- Surfaces d'expérimentation agronomiques, vergers de conservation, parcelles de multiplication (rajeunissement des semences) et d'évaluation, zones de quarantaine,
- Centre de calcul et de traitement de l'information,
- Services de conditionnement et d'expédition des échantillons échangés et introduits.

Ces centres établiront des réseaux de stations d'observations et de multiplications et des sites de conservations dynamiques.

Les spécialistes appartiendront à des disciplines scientifiques variées et devront être largement polyvalents, nombre d'entre eux devront être d'une grande mobilité tant pour l'organisation et la réalisation des prospections que pour suivre, stimuler et encadrer toutes les conservations déléguées auprès de groupes décentralisés. Ces disciplines scientifiques sont:

- génétique et amélioration des plantes,
- phytopathologie et protection des cultures (évaluation et quarantaine),
- informatique et statistique, planification des expériences,
- physiologie des semences et des cultures in vitro,
- biochimie,
- des équipes d'agronomies, des organisateurs de réseaux, d'enquête et de surveillance,
- prospecteurs de terrain ayant des formations assez polyvalentes.

Administrativement les centres de ressources génétiques doivent être sous la responsabilité directe des plus hautes instances gouvernementales (premier ministre ou président de la République, suivant les types d'organisation) pour que les connexions évidentes avec l'agriculture (en particulier l'amélioration des plantes) ne risquent pas de mettre en péril le fonctionnement du centre par des compromis budgétaires entre les impératifs à court terme de l'agronomie et la sauvegarde à long terme d'un patrimoine collectif. Il faut nourrir le mieux possible l'amélioration des plantes mais en toute indépendance.

## V. FORMATION ET ENSEIGNEMENT

A tous les niveaux l'exigence principale pour la gestion des ressources génétiques est le sens des responsabilités vis-à-vis de la collectivité. Ce travail à long terme ne peut être apprécié par de simples critères quantitatifs; un ensachage réalisé trop tard lors d'une campagne de renouvellement des graines pour des lignées de maïs ne sera révélé que 8 ans plus tard si tel est le rythme des campagnes de rajeunissement des stocks; quelle énergie faut-il mettre pour sauver une variété traditionnelle qui paraît bien misérable et de peu de valeur aux yeux d'un sélectionneur? Des

spécialistes chinois avaient raconté qu'au cours des dernières années de la révolution culturelle il leur avait été interdit de multiplier leurs collections de ressources génétiques, ce travail étant jugé comme relevant d'une préoccupation de conservateur de musée empreinte de nostalgie passéiste. Malgré les risques que cela représentait ils allaient en cachette, de nuit, poser quelques sacs pour assurer quand même un minimum de maintenance de leurs variétés. Sur des populations de maïs ces autofécondations augmentaient brutalement et **irréversiblement** le niveau de consanguinité (cf. chapitre I, effet des fluctuations d'effectifs), mais tout n'était pas perdu.

Quelle que soit la catégorie du travailleur que l'on formera, c'est d'abord cette responsabilité et cet engagement moral qu'il faut faire sentir et rendre vivants; des tâches subalternes mal faites ou négligées détruisent tous les efforts ultérieurs. Une erreur d'étiquetage, un mélange non détecté, une hybridation mal contrôlée et tout l'arsenal de chambres de conservation sophistiquées, de bases de données informatisées ne sont qu'illusion. La haute technologie ne remplacera jamais la conscience.

Outre la simple difficulté de détection des fautes, et ce sera aux responsables d'organiser de multiples contrôles et vérifications pour s'en prémunir au mieux, il faut considérer les situations de compromis où le jugement personnel sera déterminant. Nous avons déjà souligné l'aspect paradoxal de la conservation dynamique qui impose de sagaces sélections, on conserve un mouvement évolutif mais pas une «somme de gènes». Seuls des techniciens bien motivés peuvent conduire ce travail.

Ainsi, quel que soit le niveau auquel on s'adresse, toute formation après avoir bien montré les risques que l'humanité encoure par la disparition de nos ressources génétiques, devra consacrer l'essentiel à l'étude de la variabilité génétique des plantes cultivées, et on pourrait par exemple utiliser pour ce faire les canevas du chapitre I (données de base) du présent manuel et mettre en évidence les 7 points que développe Alice M. EVANS à l'Université de Cambridge:

- L'acte de domestication et ses conséquences,
- La mutation comme source ultime de toute variation génétique,
- Le rôle des formes adventices dans le développement des cultures,
- Les effets de la migration et de la dispersion des cultures vers de nouvelles aires,
- Les conséquences des systèmes reproductifs naturels des espèces cultivées,
- La nature des changements génétiques majeurs qui ont lieu au cours du développement des plantes cultivées,
- Les dangers de l'utilisation limitée de la variabilité génétique en amélioration des plantes et en agronomie.

A cette communauté de base près, il faut considérer les différents cadres de formation possible: s'agit-il d'initiation ou de sensibilisation de futurs responsables administratifs de haut niveau aux méthodes ou aux moyens nécessaires, ou de formation pratique pour des travailleurs de terrain, prospecteurs ou observateurs et mainteneurs de collection, ou d'un enseignement spécialisé complet pour des ingénieurs ou des chercheurs de haut niveau. Nous donnerons des exemples de ces divers types d'enseignement, déjà effectivement réalisés, pour préciser les matières et les durées. Une des grandes complexités de l'enseignement résulte d'une

diversité particulièrement grande des domaines de la biologie pour lesquels une bonne compréhension est indispensable.

## A. INITIATION AUX ASPECTS THÉORIQUES ET PRATIQUES DE FUTURS RESPONSABLES

Un stage de ce genre a été organisé par l'I.A.C.C.T. au Congo (Brazzaville) en 1980; il dura deux semaines, réunissant des représentants de plusieurs pays africains. Il était destiné à établir de premiers contacts et de premières bases pour organiser pratiquement une gestion régionale africaine des ressources génétiques. Ce stage suivait une première conférence, organisée à Abidjan où les principes généraux d'une telle organisation collective africaine des ressources génétiques avaient été définis. Les travaux de ce stage d'initiation aux études et à la constitution des ressources génétiques des plantes» ont été les suivants:

- 10 séances de conférences (matinées),
- 10 séances de travaux dirigés.

Conférences:

1. Importance et urgence des constitutions et des évaluations des Ressources Génétiques, étapes de travail (constitution, évaluation, conservation), relations avec l'amélioration des plantes.

2. Notions de complexes d'espèces et leur organisation. Données de base de génétique des populations.

3. Domestication et origine des plantes cultivées. Rôle de ces connaissances pour la constitution des ressources génétiques.

4. Principes et indications générales sur les méthodes de collecte.

5. Evaluation agronomique et évaluation génétique. Signification, difficultés. Constitution des descripteurs. Problèmes des tests de résistance.

6. Méthodes d'évaluation génétique:

— les données biochimiques (analyses enzymatiques par électrophorèse, étude des acides nucléiques),

— les données cytogénétiques (méioses, caryotypes et méthodes de banding).

7. Méthodes de classification numérique.

8. Problèmes et méthodes de conservation (graines, multiplication végétative, cultures in vitro, pollens, réserves, simulation des cultures traditionnelles).

9. Organisation des centres de Ressources Génétiques.

Travaux dirigés: Illustration des méthodes et initiation à l'emploi des ordinateurs. Organisation des Ressources Génétiques pour le Manioc au Congo.

1. Principes et vocabulaires des bases de données. Problèmes des nombres et occupations matérielles des mémoires. Première discussions sur les descripteurs.

2. Codage et enregistrement des données. Exemples concrets de descripteurs. Constitution de descripteurs pour le manioc.

3. Bases de l'interprétation statistique **multivariée**.

4. Réunion préparatoire sur le thème des collectes du manioc au Congo, recherches d'information.

5. Etablissement d'un programme de collecte du manioc au Congo, descripteurs de terrain, matériel végétal à collecter.
6. Exemple de données acquises par électrophorèse. Représentations statistiques. Discussion.
7. Exemples de données acquises par cytogénétique.
8. Exemple d'étude des relations entre formes sauvages et formes cultivées (interprétation et analyses de données expérimentales sur le mil).
9. Organisation des collectes du manioc au Congo. Constitution des équipes de collectes. Evaluation des moyens.
10. Organisation de centres de conservation de ressources génétiques du manioc.

## B. ENSEIGNEMENT TECHNIQUE

L'IBPGR (CIRP) a organisé, surtout pour du personnel anglophone, plusieurs stages de formation pratique. A titre d'exemple nous donnons ci-dessous les programmes et les niveaux de base demandés pour un stage de six semaines.

Le but général du cours était de fournir du personnel correctement formé pour la conservation des ressources génétiques au niveau national en Afrique. Les candidats devaient posséder un brevet d'agriculture ou de sciences naturelles et être impliqués dans la recherche agronomique, dans des instituts nationaux reconnus. Une expérience étendue comme techniciens dans la recherche agronomique et une bonne culture générale pouvaient remplacer les diplômes dans des cas particuliers. Cet enseignement a eu lieu dans le cadre de l'IITA (Nigéria). Le cours couvrait tous les aspects pratiques majeurs de la conservation des ressources génétiques. Des conférences introductives et des démonstrations étudiaient la diversité génétique et expliquaient son importance pour le développement futur de l'agriculture. Les cours mettaient cependant l'accent principalement sur l'enseignement pratique dans les domaines suivants:

Techniques aux champs:

planification des prospections,

lecture des cartes,

contact avec les responsables administratifs et les fermiers,

méthodes pratiques d'échantillonnage,

récupération d'informations aux champs,

spécimen d'herbier et matériel à multiplication végétative,

réalisation d'une collecte de deux semaines dans le Nigéria.

Activités concernant les banques de gènes:

stockage des semences,

test de germination et de teneur en eau des semences,

évaluation des ressources génétiques et recueil d'informations,

renouvellement, multiplication («*rejuvenation*»),

maintenance des plantes à multiplication végétative,

gestion des bases de données de ressources génétiques au niveau

technique,

quarantaine

distribution des semences.

## C. ENSEIGNEMENT UNIVERSITAIRE ET DE RECHERCHE DE NIVEAU SUPÉRIEUR

Le département de biologie végétale de l'université de Birmingham (U.K.) a organisé depuis 1975 un enseignement à plein temps d'une année (*master of science*), soutenu par une aide financière de l'IBPGR (BIRP).

Le cours est organisé pour des candidats possédant l'équivalent d'une maîtrise ès sciences en biologie, botanique, génétique, amélioration des plantes ou agricultures. Il était possible de considérer aussi des candidatures d'étudiants bien qualifiés dans certains autres domaines. Ces cours sont destinés aux écologistes des plantes cultivées, aux responsables de conservation de ressources génétiques, de l'introduction des plantes, aux sélectionneurs et aux responsables de quarantaine.

Le cours comprenait neuf mois de conférences, séminaires et travaux pratiques suivis d'un stage de trois mois et de la soumission d'une thèse courte. Le cours commence chaque année fin septembre.

Matières enseignées:

— Introduction.

Revue générale des travaux de conservation génétique dans le monde, collaboration internationale, rôle de la FAO, du BIRP et d'autres organismes.

— Origines des plantes cultivées.

- taxonomie,
- origines et domestication des plantes cultivées,
- revue historique,
- techniques morphogéographiques, cytogénétiques, chimio-taxonomiques, archéologiques, et linguistiques,
- rôle des espèces sauvages et adventices à l'origine des plantes cultivées.

— Plantes d'intérêt économique particulier.

Botanique et production d'alcaloïdes de parfums et d'huiles essentielles, botanique des plantes à épices, organisation de la production alimentaire et des prix.

— Méthodes taxonomiques.

Le concept de caractères, critères morphologiques, anatomiques, paléontologiques, embryologiques et cytologiques, méthodes numériques, systématiques, biochimiques: chromatographie des composés secondaires, électrophorèse et méthode immunochimique pour l'étude de protéines.

— Génétique des populations.

Fréquences génétiques et effets des mutations, sélections, migration et effectifs des populations. Les sources de variation, les types de variation, les méthodes d'études, les systèmes de reproduction, le maintien et la libération de la variation, les types de sélections, polymorphisme, spéciation, évolution en action dans les populations spontanées.

— Variation dans les plantes cultivées et spontanées.

Hybridation, introgression et flux de gènes, reproduction sexuée et asexuée, mécanismes d'isolements, polyploïdie et ses conséquences en taxonomie et évolution. Variations morphologiques et physiologiques. Adaptation aux facteurs climatiques et édaphiques. Etude expérimentale des clones et des écotypes. Microévolution et sélection naturelle.

— Systèmes agraires.

Principaux types de systèmes dans le monde, leurs caractéristiques écologiques et leur développement historique.

— Agroécologie, climatologie et productivité.

Distribution mondiale des climats, de la végétation naturelle et des cultures; propriétés physiques des plantes et des surfaces foliaires; absorption de l'énergie solaire et ses conversions par la végétation; microclimatologie des types climatiques principaux; instrumentations; exemples de manipulation agricole des microclimats; productivité des écosystèmes.

— Pathologie des plantes.

Méthode de contrôle et de diagnostics des maladies dues aux champignons, virus, insectes et nématodes. Base génétique de la résistance. Effet des mécanismes de dispersion du pathogène, de l'environnement et des pratiques culturales.

— Statistiques pour la biologie.

Résumé des données; probabilités; distributions; test de signification et analyse de variance; plan d'expériences; régression. L'accent est en permanence mis sur les applications pratiques de la statistique à la biologie.

— Traitement des données.

Introduction aux langages informatiques tel que le basic et à la programmation. Utilisation et limitation des ordinateurs y compris des microordinateurs.

— Gestion de l'information.

Communication dans les ressources génétiques. Systèmes d'information et de documentation. Codage des descripteurs et états des descripteurs. Création et gestion de base de données. Système internationaux.

— Prospection.

Distribution de la variabilité. Structures de populations et méthodes d'échantillonnage. Stratégie et tactiques des prospections.

— Conservation.

Centres de ressources génétiques. Banques de gènes. Stations d'introduction et jardins botaniques. Stockage des semences et de pollen. Culture de tissus et multiplication végétative. Collection de base et collection de travail. Centres nationaux, régionaux et internationaux. Quarantaine.

— Utilisation.

Exemples d'utilisation des ressources génétiques dans un choix de plantes cultivées prépondérantes.

— Amélioration des plantes.

Objectifs de sélection. Contrôle et exploitation de la variabilité génétique. Sélection et évolution. Sélection pour la résistance aux parasites et aux maladies. Production et maintenance des nouvelles variétés.

— Ressources génétiques forestières.

Méthodes de collecte, test de provenance et sélection.

Le diplôme de «master of science» sanctionne le succès à un examen (écrit, oral et thèse) concernant l'ensemble des cours. Dans certains cas des étudiants peuvent être autorisés à poursuivre des recherches pendant deux ou trois ans pour obtenir le Ph. D. (équivalent d'une thèse de Sème cycle).

Un enseignement comparable a été initié en France en 1979 dans le cadre d'une option particulière du D.E.A. d'amélioration des plantes de l'université de Paris XI (Orsay), une formation complémentaire de deux ans pourra être assurée soit dans un laboratoire du CNRS de la région pari-

sienne (G.P.D.P. Gif sur Yvette), soit en collaboration avec l'ORSTOM à Abidjan (Côte-d'Ivoire). Les deux années complémentaires permettent l'obtention d'une thèse de Sème cycle. Des enseignements, des stages, visites et conférences suivent les mêmes thèmes que ceux décrits pour l'université de Birmingham.

Cet enseignement se déroule dans le cadre de la première année de l'enseignement de Paris XI comme une option du D.E.A. (diplôme d'études approfondies) d'amélioration des plantes. Les bases générales d'amélioration des plantes couvrent les différents aspects suivants: génétique quantitative, base et principes de sélection; morphogenèse et problèmes particuliers aux plantes à multiplication végétative, aspect physiologique de l'amélioration des plantes, phytopathologie. Puis pour l'option particulière de ressources génétiques des compléments sont données en génétiques des populations, génétique de base (cytogénétique et génétique moléculaire) et génétique du développement. Le cours de génétique des populations met l'accent sur les thèmes principaux:

- polymorphisme des populations naturelles et variétés cultivées (concepts, mesures et techniques modernes (moléculaires)),
- distances génétiques et études des phylogénies,
- paramètres de déséquilibre gamétique et sa signification pour l'étude des populations subdivisées,
- domestication des plantes et organisation des complexes d'espèces.

Après avoir suivi ces enseignements et avoir été reçu à différents examens de contrôle, les étudiants entreprennent un stage et suivent également des enseignements complémentaires concernant:

- les ressources génétiques et les discussions sur de nouvelles méthodes en amélioration des plantes destinées à obtenir des variétés d'un type nouveau avec une variabilité génétique plus large,
- l'érosion génétique (avec des exemples), connections et interactions entre plantes cultivées, formes sauvages et adventices,
- la signification des barrières reproductives et de différentes formes de spéciation et méthodes d'hybridation interspécifiques,
- l'étude sur la conservation de la variabilité génétique.

L'expérimentation de stage a lieu en général dans le laboratoire de génétique et physiologie du développement des plantes C.N.R.S. à Gif-sur-Yvette. Les travaux sont pour l'essentiel dirigés autour des complexes d'espèces du mil à chandelles (*Pennisetum*) et des millets (*Setaria*). Quatre thèmes principaux ont ces dernières années fait l'objet de stages:

- Etude des isozymes comme marqueurs génétiques, analyse de polymorphismes enzymatiques et distances génétiques,
- déterminisme génétique de la sensibilité à la photopériode pour la floraison, conséquences pour les stratégies d'introduction des ressources génétiques exotiques,
- organisation des banques de données des ressources génétiques pour les complexes de *Setaria* et de Mil: évaluation des collections, codage et établissement de descripteurs; organisation des bases de données sur microordinateur et gros système; couplage aux bibliothèques de programmes de traitement statistiques,
- Etude des changements génétiques consécutifs aux cultures in vitro.

Une partie des expérimentations se déroulent pendant la période d'été en stations expérimentales au champ dans le Sud Ouest de la France,

pendant la période hivernale les plantes sont cultivées en phytotron. Au cours de la 2<sup>ème</sup> année deux possibilités peuvent être offertes aux candidats, soit la poursuite de leurs travaux de thèse dans la région parisienne, soit dans le cadre d'une association de formation avec l'ORSTOM, étudier en Côte-d'Ivoire les problèmes de collectes, conservation et constitution des ressources génétiques sur le riz, le café, le mil et une plante fourragère le *Panicum maximum*.

Cependant tout ne s'apprend pas à l'université, en laboratoire et en station de recherche. Ce sera notre conclusion sur l'enseignement et pour l'ensemble du manuel: il faut apprendre à découvrir, avec respect, la richesse des patrimoines que nous transmettent les paysans. Ce n'est possible qu'au travers de longues heures ou journées d'enquêtes menées avec curiosité et sympathie, sans formulaires stéréotypés, sur les lieux mêmes où les ressources génétiques sont vivantes: au champ ou aux abords des greniers. Seuls les prospecteurs, qui ont nourri leur intelligence et leur sensibilité au contact de quelques maîtres paysans encore de ce monde, seront marqués de l'empreinte génératrice de gestionnaires responsables des ressources génétiques.

## RÉFÉRENCES BIBLIOGRAPHIQUES

- ABBOTT A.J. 1978. Practice and promise of micropropagation of woody species. *Acta Horticultural*, 79: 113-127.
- ABDALLA F.H. and ROBERTS E.H., 1969. The effects of temperature and moisture on the induction of genetic changes in seeds of barley, broad beans and peas during storage. *Ann. Bot.* 33: 153-57.
- ALLARD R.W., 1970. Population structure and sampling methods. In *Genetic Resources in Plants*, Ed. O.H. Frankel and E. Bennett: 97-118.
- ANDERSON E., 1949. *Introgressive hybridization*. Wiley, J. & Sons. New York Chapman Hall, London.
- ANTONOVICS J., 1968. Evolution in closely adjacent plant populations. V. The evolution of self-fertility. *Heredity* 23: 219-238.
- ANTONOVICS J., 1968. Evolution in closely adjacent plant populations. VI. Manifold effects of gene flow. *Heredity* 23: 507-524.
- AQUADRO C.F. and AVISE J.C., 1981. Genetic divergence between rodent species assessed by using two-dimensional electrophoreses. *Proc. Nat. Acad. US* 78: 3784-3788.
- AUTRAN J.C. et BOURDET A., 1973. Nouvelles données révélant l'hétérogénéité électrophorétique des gliadines du grain de blé en vue d'une identification variétale. *C.R. Heb. A.C.* 277: 2081-2084.
- BAJAJ Y.P.S., 1976. Regeneration of plants from cell suspension frozen at -20, -70 and -196°C. *Phy. Plant.* 37: 263-268.
- BANCILHON L., 1971. Contribution à l'étude taxonomique du genre *Phyllanthus* (Euphorbiacées). *Boissiera*, Vol. 18.
- BARNABAS B. and RAJKI E., 1976. Storage of maize (*Zea mays* L.) pollen at -196°C in liquid nitrogen. *Euphytica* 25: 747-752.
- BASS L.N., 1974. Response of seed of 27 *cucumis melo* cultivars to three storage conditions. *Hort. Science*, 9, 2: 157.
- BENDICH, A.J. and McCARTHY B.J., 1970. *Proc. Nat. Acad. Sci. USA*, 65: 349-356. *Proc. Nat. Acad. Sci. USA*, 65: 545-565.
- BENZECRI et al., 1973. La Taxonomie Tome I, L'analyse des correspondances, Tome 2, Ed. Dunod Paris.
- BIGOT C., 1977. Multiplication végétative «in vitro» par néoformation de bourgeons et d'embryons somatiques. In: *Multiplication végétative et rhizogenèse*, Ed. GAGNAIRE MICHARD J., LEMAIRE F. et RIEDACKER A., Tome 5, 1ère partie: 51-103.
- BLACK W., MASTENBROEK C., MILLS W.R. et PETERSEN L.C., 1953. A proposa for an international nomenclature of races of *Phytophthora infestans* and of genes controlling immunity in *Solanum demissum* derivatives, *Euphytica* 2: 173-178.
- BOGYO T.P., PORCEDDU E., PERRINO P., 1980. Analysis of sampling strategies for collecting genetic material. *Economic Botany* 34, 2: 160-174.
- BONGA J.P., 1977. Application of tissue culture in Forestry. In *Plant Cell, Tissue and Organ culture*. Ed. Reinert and Bajaj, Springer Verlag: 93-108.
- BOUZID S. et LASRAM M., 1971. Utilisation de culture in vitro pour l'obtention de clones de *Citrus* homogènes et de bon état sanitaire. VIII° Congrès Inter. *Agrum. Médit.* 2: 1-6.
- BREWBAKER J.L., 1967. The distribution and phylogenetic significance of

- binucleate and trinucleate pollen grains in the Angiosperms. *Amer. J. Bot.*, 54 (9) : 1069-1083.
- BROWN A.H.D., 1978. Isozymes, plant population genetic structure and genetic conservation. *Theor. Appl. Genet.* 52: 145-157.
- BROWN A.H.D., 1979. Enzyme polymorphism in plant population. *Theor. Pop. Bio.* 15, 1: 1-42.
- BRUNKEN J., de WET J.M.J. and HARLAN J.R., 1977: The morphology and domestication of pearl millet. *Economic Botany* 31.
- BURSON B.L., 1981. Cytogenetic relationships between *Paspalum jurgensii* and *P. intermedium*, *P. vaginatum*, and *P. setaceum* var. *ciliatifolium*, *Crop Science* 21 (4) : 515-519.
- BURTON G.W., 1976. Gene loss in pearl millet germplasm pools. *Crop Science* 16: 251-255.
- BUTTON J. and BOTH C.E.J., 1975. Enzymatic maceration of citrus callus and the regeneration of plants from single cells, *J. of Exp. Botany* : 26: 723-729.
- CARLSON P.S., 1975. Crop improvement through techniques of plant cell and tissue culture. *Bioscience*, 25, 11: 747-749.
- CARVALHO A., KRUG C.A. et MENDES J.E.T., 1950. O dimorfismo dos ramos em *Coffea arabica* L. *Brangantia* 10: 151-159.
- CAUDERON Y., 1981. Hybridations interspécifiques et amélioration des plantes. I. Evolution des voies d'étude des relations entre espèces. *C.R. Acad. Agric. Fr.* 67 (12): 1001-1012.
- CAVALLI-SFORZA L.L., BODMER W.F., 1971. The Genetics of human population. W.H. Freeman.
- CHARRIER A., 1978. La structure génétique des caféiers spontanés de la région malgache (*Mascarocoffea*) et leurs relations avec les caféiers (*Eucoffea*). Mémoires ORSTOM, Paris, 87: 223 p.
- CHIANG M.S., 1974. Cabbage pollen germination and longevity. *Euphytica* 23: 579-584.
- CHINNAPPA C.C., 1970. Interspecific hybrids of *Coffea canephora* and *C. liberia* *Genetica* 41: 141-146.
- CHU Y.E. and OKA H.I., 1970. Introgression accross isolating barriers in wild and cultivated *Oryza* species. *Evolution* 24, 2: 344-355.
- CLIFFORD B.C., 1972. The histology of rare-non-specific resistance to *Puccinia hordei* Oth. in barley. *Proc. and Eur. Mediterranean Cereal Rust. Conf. (Prague)*: 75-79.
- COMBES D., 1972. Polymorphisme et modes de reproduction dans la section des *Maximae* du genre *Panicum* (Graminées) en Afrique. Thèse, Paris. Mémoire ORSTOM, Paris n° 77: 99 p. (1975).
- COUTURON E., 1979. Maintaining the viability of coffee seeds by checking their water content and storage temperature. *Café, Cacao, Thé* 24 (1) : 27-32.
- D'AMATO F., 1975. Les problèmes de la stabilité génétique dans les cultures de cellules et de tissus végétaux. *In. Crop Genetic Resources for to day and to morrow*. Ed. Frankel and Hawkes: 53-80.
- DAY P.R., 1974. Genetics of host-parasite interaction. *Freeman, San Francisco* 238 p.
- DE VIENNE D., 1977. Variabilité et évolution chez *Medicago sativa* L. Apports de l'étude des isoenzymes du pollen. Thèse 3° cycle, Paris VI.

- DE WET J.M.J., 1971. Reversible tetraploidy as an evolutionary mechanism. *Evolution* 25: 545-548.
- DINOOR A., PELEG N., 1972. The identification of genes for resistance or virulence, without genetic analyses by the aid of the «gene-for-gene» hypothesis. *Proceed. European and Medit. Cereal Rust. Conf. II.*
- DOBZHANSKY Th., 1951. «Genetics and the origin of species». 3ed. Columbia Univ. Press New York.
- ENGELS W.R., 1981. Estimating genetic divergence and genetic variability with restriction endonucleases. *Proc. Natl Acad. Sci. USA*, 78: 6329-6333.
- FARMER R.E. Jr and BARNETT P.E., 1974. Low-temperature storage of black walnut pollen. *Cryobiology*, 11, 4: 366-367.
- FAVRE J.M., 1973. Divers aspects du rôle du bourgeon et des noeuds sur la rhizogenèse de la vigne, cultivée in vitro. *Rev. Cytol. Biol. Végét. Paris*. 37: 393-406.
- FELDMAN M.W. and CHRISTIANSEN F.B., 1975. The effect of population subdivision on two loci without selection. *Gen. Res.* 24: 151-162.
- FIGIER J., 1982. Etude de la variabilité et du déterminisme de la morphologie de l'*Hedysarum coronarium* L en Tunisie... Implications concernant l'amélioration de cette espèce fourragère dans ce pays. Thèse d'Etat Paris XI-Orsay. 234 p.
- FISCHER E. et GÄMANN E., 1929. Biologie der Pflanzen-bewohender parasitischer Pilze. Jena. Fischer 428 p.
- FLAVELL R., BENNETT M., SMITH J. and SMITH D., 1974. Genome size and the proportion of repeated nucleotide sequence DNA in plants. *Biochem. Genet.* 12: 257-279.
- FLAVELL R., O'DELL M., RIMBAU J. and SMITH D.B., 1978. Biochemical detection of Alien DNA incorporated into wheat by chromosome engineering. *Heredity* 40: 439-455.
- FLAVELL R., SMITH D.B. et BEDROVK J.R., 1979. The evolution of plant genome structure. *In: Plant Genome Organisation and Expression.* C.J. Leaver Ed. New York.
- FLOR H.H., 1955. Host-parasite interaction in flax rust, its genetics and other implications. *Phytopathology*, 45: 680-685.
- FLOR H.H., 1956. The complementary genic systems in flax and flax rust. *Advan. Genet.* 8: 29-54.
- FLOR H.H., 1971. Current status of the gene-for-gene concept. *Ann. Rev. Phytopathol.* 9: 275-296.
- FRANKEL O.H., 1974. Genetic conservation: our evolutionary responsibility. *Genetics* 78, 1: 53-65.
- FRANKEL O.H., 1977. Genetic resources as the backbone of plant protection. *In: Induced mutations against plant diseases. Proceeding of a Symposium Vienna 31-1/4-2:* 3-14.
- FRANKEL O.H. et BENNETT E., 1970. «Genetics Resources in Plants». Blackwell Scientific Publ. Oxford.
- GALZY, 1969. Remarques sur la croissance de *Vitis rupestris* Scheele cultivée in vitro sur différents milieux nutritifs. *Vitis*, 8: 191-205.
- GAUTHERET R.J., 1939. Sur la possibilité de réaliser la culture indéfinie des tissus de tubercules de carotte. *C.R. Hebd. Séance. Acad. Sci. Paris* 208: 118-121.
- GHORAID.P., 1976. Viability of rice in storage. *Oryza*, 13, 1: 33-36.

- GILMOUR J.S.L., 1960. The Deme Terminology. Report of the Scottish Plant Breeding Station. Pentlandfield Scotland : 99-105. In Population and Environmental Biology, A.S. Boughey ed. (1967) Dickenson Pub. California.
- GOUJON M., 1971. Considérations à propos de la résistance des plantes. Le cas particulier des caféiers attaqués par les rouilles orangée et farineuse. Café, Cacao, Thé, 15: 308-328.
- GRAHAM K.M., NIEDERHAUSER J.S., et ROMERO S., 1959. Observations on races of *Phytophthora infestans* in Mexico during 1956-1957. Am. Potato J., 36: 196-203.
- GRASSIAS-HUBAUT M., 1980. Etude de la fertilité et du comportement méiotique des hybrides interspécifiques tétraploïdes *Arabusta*. Thèse de 3ème Cycle, Paris XI.
- GROUZIS M., 1979. Sur le *Pennisetum violaceum sensu lato* en Afrique de l'Ouest: Formes, Ecologie et distribution géographique. Bulletin de l'I.F.A.N. 41 A2: 300-316.
- HALPERIN W. et WETHERELL D.T., 1964. Adventice embryony in tissue culture of the wild carrot *Daucus carota*. Am. J. Bot. 51: 274-183.
- HAOUSSOU N., TOURE B., PIQUEPAILLE M., 1981. Etude de la variabilité créée par la nature des organes de multiplication végétative chez *Dioscorea alata* cv. *Braso fuerte*. Coll. INRA Ignames, 28/7-2/8 1980 Guadeloupe.
- HARDY-WEINBERG G.H., 1908. Mendelian proportions in a mixed population. Science 28: 49-50.
- HARLAN J.R., 1970. The evolution of cultivated plants. In: Frankel and Bennett, Ed. « Genetic resources in plants-their exploration and conservation », Blackwell Scientific Publ. Oxford: 19-32.
- HARLAN J.R., 1972. The genetics of disaster. J Environ Quality 1, 3: 212-215.
- HARLAN J.R., 1976. Diseases as a factor in plant evolution. Ann. Rev. Phytopathol. 14: 31-51.
- HARLAN J.R. et DE WET J.M.J., 1971. Toward a rational classification of cultivated plants. Taxon 20 (4): 509-517.
- HARRINGTON J.F., 1970. Seed and pollen storage for conservation of plant genes ressources. In: Genetic Resources in Plants-Their Exploration and conservation. Frankel and Bennett Ed. Blackwell Scientific Publ. Oxford 501-521.
- HARRIS H., HOPKINSON D.A., 1976. Handbook of enzyme electrophoresis in human genetics. Amer. Elsevier Publ. Cy New York.
- HESLOP-HARRISON J. and HESLOP-HARRISON Y., 1970. Evaluation of pollen viability by enzymatically induced fluorescence; intracellular hydrolysis of fluoresceins diacetate. Stain Technology 45: 115-120.
- HILLEL J., FELMAN M.W. et SIMCHEN G. 1973. Mating systems and population structure in two closely related species of the wheat group. I variation between and within population. Heredity, 30: 141-167.
- HOLLY L., 1981. Use of back-garden system and natural reserved for isoclimatic regeneration of germplasm samples in Hungary. Intern. Conf. on Crop Genetic Ressources. Report of FAO, Rome 1981.
- HOOVER A.L., 1969. Widely based resistance to rust in Corn. Field Crops. Spec. rev. Iowa Agric. Home Econ. Exp. Stn. 64: 28-34.

- HUSSEY G., 1978. The application of tissue culture to the vegetative propagation of plants. *Science Progress*, 65, 258: 185-208.
- ITO H., 1972. Organization of the national seed storage laboratory for genetic resources in Japan. *In: Viability of seeds*. 405-416.
- IWANAMI Y., 1972. Retaining the viability of *Camellia japonica* pollen in various organic solvents. *Plant and Cell Physiology*. 13: 1139-1141.
- JAIN S.K., 1975. *In: Plant Genetic resources Today and Tomorrow*. Ed. Frankel and Hawkes: 15-36.
- JAMES E., 1972. Organisation of the United States National Seed Storage Lab. *In: Viability of Seeds*: 397-404.
- JOHRI B.M. et VASIL I.K., 1961. Physiology of pollen. *Bot. Rev.* 27: 326-381.
- KAHLER A.L., CLEGG M.T. and ALLARD R.W., 1975. Evolutionary changes in the mating system of an experimental population of barleys (*Hordeum vulgare* L.). *Proc. Nat. Acad. Sci.* 72, 3: 943-946.
- KING J.R., 1965. The storage of pollen-particularly by the freeze-drying method. *Bull. Torrey Bot. Club* 92: 270-287.
- KING J.L. et JUKES T.H., 1969. Non-Darwinian evolution. Random fixation of selectively neutral mutations. *Science* 164: 788-798.
- KING J.L. et OHTA T., 1975. Polyallelic mutational equilibria. *Genetics* 79: 681-691.
- KONZAK C.F., 1972. Progress toward international standardized documentation of genetic resource collections from plant explorations. *Amer. Sty. Agro.*: 13.
- KURATA N. et OMURA T., 1978. Karyotype analysis on rice. A new method for identifying all chromosome pairs. *Japan J. Genet.* 53: 251-255.
- LANAUD C., 1979. Etudes des problèmes génétiques posés chez le caféier par l'introggression de caractères d'une espèce sauvage (*C. kianjavatensis* *Mascarocoffea*) dans l'espèce cultivée *C. canephora* (*Eucoffea*). *Café, Cacao, Thé* 23: 3-28.
- LEBLANC J.M., 1978. Etude sur le système des alcools déshydrogénases du mil: *Pennisetum typhoideum* (*americanum*). Thèse 3<sup>e</sup> Cycle, Paris XI.
- LEIGH-BROWNE A.J. and LANGLEY C.H., 1979. Reevaluation of genic heterozygosity in naturel population of *Drosophila melanogaster* by two-dimensional electrophoresis. *Proc. Nat. Acad. Sci. US* 76: 2381-2384.
- LEPPIK E.E., 1970. Gene centers of plants as sources of disease resistance. *Ann. Rev. Phytopathol.* 8: 323-344.
- LEROY J.F., 1967. Recherches sur les caféiers. Sur la classification biologiques des caféiers et sur l'origine et l'aire du genre *Coffea*. *C.R. Acad. Sci. Paris* 265: 1043-1045.
- LEROY J.F., 1967. Recherches sur les caféiers. Esquisse d'une théorie sur l'évolution des espèces. *C.R. Acad. Sci. Paris* 265: 1373-1376.
- LETOUZEY R. *In: Flore du Cameroun*.
- LEWONTIN R.C., 1974. The genetic basis of evolutionary change. Columbia Univ. Press. New York.
- LINSKENS H.F., 1964. Pollen physiology. *Ann. Rev. Pl. Physiol.* 15: 255-270.
- LOEGERING W.Q., MacINTOSH R.A. and BURTON C.H., 1971. Computer analysis of disease data to derive hypothetical genotypes for reaction

of host varieties to pathogens. Can. J. Genet. Cytol. 13: 742.

- LOUARN J., 1976. Hybrides interspécifiques entre *Coffea canephora* Pierre et *C. eugenicides* Moore. Café, Cacao, Thé 20: 33-52.
- MALCOLMSON J.F. et BLACK W., 1966. New genes in *Solanum demissum* Lindl. and their complementary races of *Phytophthora infestans* (Mont.) de Bary. Euphytica 15: 199-203.
- MANGNOT G., 1975. La Graine: définition, origine et développement, types structuraux, vie ralentie, germination. In: Les protéines des graines, Ed. J. Miegé. Conservatoire et Jardin Botanique de Genève: 15-29.
- MARGARA K., 1978. Mise au point d'une gamme de milieux minéraux pour les conditions de la culture «in vitro». C.R. Ac. Sci. Agric. France, 64,8: 654-661.
- MARSHALL D.R. and BROWN A.H.D., 1975. Optimum sampling strategies in genetic conservation. In: Crop genetic resources for today and tomorrow. Ed. Frankel and Hawkes: 53-80.
- MARTIN F.W., 1975. La conservation du germplasm des racines et tubercules tropicaux en forme végétative. In: Crop Genetic Resources for today and tomorrow. Ed. Frankel and Hawkes: 369-377.
- MARTIN F.W., 1958. Staining and observing pollen types in the style by means of fluorescence. Stain Technol. 34: 124-128.
- MARTIN T.J. et ELLINGBOE A.H., 1976. Differences between compatible parasite-host genotypes involving the Pm4 locus of wheat and the corresponding genes in *Erysiphe graminis* f. sp. *tritici*. Phytopathology 66: 1435-1438.
- McCLINTOCK B., KATO Y.T.A. et BLUMENSTEIN A., 1981. Chromosome constitution of races of maize. Its significance in the interpretation of relationships between races and varieties in the Americas. Colegio de Postgraduados, Chapingo, Mexico, 518 p.
- METTLER L.E. et GREGG T.G., 1969. Population genetics and evolution. Prentice Hall London.
- MOORE, 1972. Effect of mechanical injuries of viability. In: Viability of Seeds: 207.
- MOREL G. et MARTIN C., 1952. Guérison de dahlias atteints d'une maladie à virus. C.R. Ac. Sci. 235: 1324-1325.
- MOREL G. et MARTIN C., 1955. Guérison de pommes de terre atteintes de maladies à virus. C.R. Ac. Sci. 41: 472-475.
- MORISHIMA H., HINATA K. and OKA H.I., 1963. Comparison of modes of evolution of cultivated forms from two wild rice species, *O. breviligulata* and *O. Perennis*. Evolution 17: 170-181.
- MOTI, 1972. Studies on the morphology and viability of the grape (*Vitis vinifera* L.) pollen. Punjab Horticultural Journal (India), 12, 2/3: 101-110.
- MURASHIGE T. et SKOOG F., 1962. A revised medium for rapid growth and bioassays with tobacco tissue cultures. Physiol. Plantarum. 15: 473-497.
- NAYAR N.M., 1973. Origin and cytogenetics of rice. Adv. in Genet. 17: 153-292.
- NEI M., 1967. Modification of linkage intensity by natural selection. Genetics. 57: 625-641.
- NEI M., 1975. Molecular population genetics and evolution. North-Holland, Amsterdam.

- NEI M. et LI W.H., 1973. Linkage disequilibrium in subdivided populations. *Genetics* 75: 213-219.
- NELSON R.R., 1975. Horizontal resistance in plants: concepts, controversies and applications. *In: Proceedings of the seminar on horizontal resistance to the blast disease of rice*. Ser. CE 0. CIAT Colombia 1-20.
- NELSON R.R., 1978. Genetics of horizontal resistance to plant diseases. *Ann. Rev. Phytopathol.* 16 : 359-378.
- NEVO E., GOLENBERG E., BEILES A., BROWN A.H.D. and ZOHARY D., 1982. Genetic diversity and environmental association of wild wheat: *Triticum dicoccoides* in Israël. *Theor. Appl. Genet.* 62: 241-254.
- NOBECOURT P., 1939. Sur la pérennité et l'augmentation de volume des cultures de tissus végétaux. *C.R. Séance. Soc. Biol.* 130: 1270-1271.
- NOBECOURT P., 1939. Sur les radicules naissant des cultures de tissus végétaux. *C.R. Séance Soc. Biol.* 130: 1271.
- NOZERAN R. et BANCILHON L., 1972. Les cultures in vitro en tant que technique pour l'approche de problèmes posés par l'amélioration des plantes. *Ann. Amélior. Plantes* 22 (2) 167-185.
- O'FARREL P.H., 1957. Asynaptic effect of chromosome V. *Wheat Infor. Serv.* 5: 6.
- PARLEVLIET J.E. et ZADOKS J.C., 1977. The integrated concept of disease resistance; a new view including horizontal and vertical resistance in plants. *Euphytica* 25, 5: 21.
- PERNES J., 1970. Etude du mode de reproduction: apomixie facultative du point de vue de la génétique des populations. *Travaux et Documents ORSTOM, Paris*, 66 p.
- PERNES J., 1975. Organisation évolutive d'un groupe agamique: la section des *Maximae* du genre *Panicum* (Graminées). *Mémoires ORSTOM* 75, ORSTOM Paris.
- PERSON C., 1967. Genetic aspects of parasitism. *Can. J. Bot.* 45: 1193-1204.
- PHILIPPS L.L., 1976. Cotton. *In: Evolution of crops plants*. N.W. Simmonds Ed. Longmann, London : 196-200.
- PIERIK R.L.M., 1975. Vegetative propagation of horticultural crops «in vitro» with special attention to shrubs and trees. *Acta Hortic.* 54: 71-82.
- PURSEGLOVE J.W., 1968. *Tropical Crops*. Longmann, Green and Co.
- DUALSET CO, 1975. Sampling germplasm in a center of diversity: an example of disease resistance in Ethiopian barley. *In Crop Genetic Resources for today and tomorrow*. Ed. Frankel and Hawkes. Cambridge Univ. Press.
- REINERT J. et BAJAJ Y.P.S., 1977. Applied and fundamental aspects of plant cell tissue and organ culture. Springer Verlag.
- ROBERTS E.H., 1975. Problems of long term storage of seed and pollen for genetic resources conservation. *In: Crop genetic resources for today and tomorrow*. Ed. Frankel & Hawkes. 269-296.
- ROBERTS E.H., 1972. *Viability of Seeds*. Ed. ROBERTS. Chapman & Hall. London.
- ROBERTS E.H., ABDALLA F.H. and OWEN R.J., 1967. Nuclear damage and the ageing of seeds, with a model for seed survival curves. *Symp. Soc. Exp. Biol.* 21: 65-100.

- ROBERTS E.H. and ELLIS R.H., 1977. Prediction of seed activity longevity at subzero temperature and genetic resources conservation. *Nature* (London) 268: 431-432.
- ROBERTS E.H. and ROBERTS D.L., 1972. Viability of seeds. Ed. ROBERTS. Chapman & Hall. London.
- ROBINSON R.A., 1973. Horizontal resistance. *Rev. Plant Pathol.* 52 (8): 483-501.
- ROCHE L.R., 1976. Méthodologie de la conservation des ressources génétiques forestières. FAO (Rome) 134 p.
- ROUX J., 1968. Sur le comportement des axes aériens chez quelques plantes à rameaux végétatifs polymorphes: le concept de rameau plagiotrope. *Ann. Sci. Nat. Bot. Ser.* 12 (9): 111-244.
- SAKAI et NOSHIRO, 1975. Quelques facteurs contribuant à la survie des semences de plantes cultivées refroidies à la température de l'azote liquide. *In: Genetic Resources in Plants.* Ed. O.H. Frankel & Bennett: 317-326.
- SCHICK R., SCHICK E. et HAUSSDORFER M., 1958; Ein Beitrag zur physiologischen Spezialisierung von *Phytophthora infestans*. *Phytopathol.* 2, 31: 225-236.
- SECKINGER G.R. and Mc COWN B.H., 1978. Microculture and woody plants. *Fundamental methods, problems and potentials.* *Amer. Nurseryman*, 148, 8, 11: 112-119.
- SECOND G., 1982. Origin of the genic diversity of cultivated rice (*Oryza* spp.). Study of the polymorphism scored at 40 isozyme loci. *Jap. J. Genetics*, 57: 25-57.
- SECOND G. and TROUSLOT P., 1980. Polymorphisme de treize zymogrammes observés parmi diverses espèces sauvages et cultivées du genre *Oryza*. *In Electrophorèse d'enzymes de riz (Oryza Spp.)*. Travaux et documents 120, 50-88 ORSTOM PARIS.
- SHAHI B.B., MORISHIMA H. and OKA H.I., 1969. A survey of variations in peroxidase, acid phosphatase and esterase isozymes of wild and cultivated *Oryza* species. *Japan. J. Genet.* 44: 303-319.
- SIDHU G.S., 1975. Gene-for-gene relationships in plant parasites systems. *Sci. Prog.* 62, 467-485.
- SIMMONDS, N.W., 1962. The Evolution of the Bananas.
- SOLDATOV O.P., 1976. Spontaneous mutation and the germination rate of seeds after prolonged storage. *Byulleten' Vsesoyuznogo Ordena Lenina i Ordena Druzhby Narodov Institute Rastenievodstva Inreni N.I. Vavilova (Leningrad)* 60: 20-22. (*In Plant Breeding Abstracts*, 1977, 47, 6 = 427).
- STANLEY R.G. and LINSKENS H.F., 1974. Pollen biology biochemistry management. Ed. Springer Verlag.
- STOKES M.J., 1974. Plant propagation by means of aseptic techniques. *Proceed. Internat. Plant Propagator's Society*, 24: 196-206.
- THOMPSON P.H., 1974. The use of seed banks for the conservation of populations of species and ecotypes. *Biological conservation* 6: 15-19.
- TOXOPEUS H.J., 1956. Reflexions on the origin of new physiologic races of *Phytophthora infestans* and the breeding for resistance in potatoes. *Euphytica* 5: 221-237.

- TSÁI-YING CHENG, 1978. Propagating woody plants through tissue culture. *Amer. Nurseryman*, 147 (10) 7-8: 94-102.
- ULLRICH J., 1976. Epidemiologische Aspekte bei der Krankheitsresistenz von Kulturpflanzen. In: *Advances in plant breeding*; Supplement to J.P. Breed. 6, Verlag Paul Parey, Berlin-Hamburg. 88 p.
- VAN DER PLANK J.E., 1963. *Plant diseases: Epidemics and control*. Academic Press New York, London, 349 p.
- VAN DER PLANK J.E., 1968. *Disease resistance in plants*. Academic Press, New York, London 206 p.
- VAN DER VOSSEN, 1978. Methods of preserving the viability of coffee seed in storage. *J. Seed Science and Technology* : 6.
- VAVILOV N.I., 1951. The origin, variation, immunity and breeding of cultivated plants. In «Selected writings of N.I. Vavilov». Transl. by K. Starr Chester. *Chronica bot.* 13, 1-364.
- VEDEL F., QUETIER F. et BAYEN M., 1976. Specific cleavage of chloroplast DNA from higher plant by *ECO RI* restriction nuclease. *Nature*, 263: 440-442.
- VEDEL F., QUETIER F., 1978. Hydrolyse spécifique de l'ADN chloroplastique et de l'ADN mitochondrial des végétaux supérieurs par les enzymes de restriction. *Physiol. Vég.* 16 (3) : 441-425.
- VEYRET Y., 1965. L'apomixie chez les angiospermes. Thèse Univ. Paris.
- VILLIERS T.A., 1975. Genetic maintenance of seeds in imbibed storage. In: *Crop genetic Resources*. Ed. Frankel and Bennet : 297-316.
- VILLIERS T.A. and EDGCUMBE D.J., 1975. On the cause of seed deterioration in dry storage. *Seed Science and Technology* 3: 761-764.
- VISSER T., 1955. Germination and storage of pollen. *Meded. van de Land Hoogeschool Wageningen* 55: 1-68.
- VISSER T., VRIES D.P. de, WELLES G.M.H. and SCHEURINK J.A.M. & c<sup>è</sup>. Hybrid tea-rose pollen. 1. Germination and storage. *Euphytica*, 26 (3): 721-728.
- VOGEL M., 1975. Croissance rythmique du cacaoyer. Thèse 3<sup>è</sup> cycle, Univ. Paris XI.
- WALYARO D.J. et VAN DER VOSSEN H.A.M., 1977. Pollen longevity and artificial cross pollination in *Coffea arabica*. *Euphytica* 26 (1): 225-231.
- WATSON I.A., 1970. The utilization of wild species in the breeding of cultivated crops resistant to plant pathogens. In: *Genetic resources in plants. Their exploration and conservation*. Ed. Frankel and Bennett, Oxford. Blackwell : 554 p.
- WEIR B.S., ALLARD R.W. and KAHLER A.L., 1975. Further analysis of complex allozyme polymorphism in a barley population. *Genetics* 78: 911-919.
- WEIR B.S. and COCKERHAM, CLARCK C., 1973. Mixed self and random mating at two loci. *Genet. Res.* 21: 247-262.
- WELLMAN F.L. et BLAISDELL D.J., 1940. Differences in growth characters and pathogenicity of *Fusarium* with isolations tested on three tomato varieties. *U.S. Dept. Agr. Tech. Bull.* 705: 28 p.
- WHITE P.R., 1939. *Bull. Torrey Bot. Club*, 66: 505-513.
- WHITE P.R., 1963. *The cultivation of animal and plant cells*. New York. Ronald Press 2e ed.
- WILKES H.G., 1972. Maize and its wild relatives. *Science* 177: 1071-1077.

- WILLIAMS J.T. and HANSON J., 1974. The potential of vigour testing for long term seed storage. *J. Horticult. Sci. (Birmingham)* 49, 4: 395-401.
- WOLD A., 1975. The national seed storage laboratory for genetic resources in Japan. *ISTA News Bulletin*, 49: 2-3.
- WRIGHT S., 1969. Evolution and the genetics of Populations. Vol. 2: The Theory of Gene Frequencies. Chicago U.P.
- YARCHUK T.A., 1974. Duration of seed viability in maize in relation to length and method of storage. *Byulleten' Vsesoyuznogo Ordenan Lenina Instituta Rastemevodstva Imeni N.I. Vavilova (Leningrad)* 43: 16-20. (In *Plant Breeding Abstracts*, 1975, 45, 11: 712).
- ZADOKS J.C., 1966. Problem in race identification of wheat rusts. In: 5th Yugoslav Symposium on Research in Wheat (Novi Sad). *Contemp. Agric.* 11: 299-305.
- ZOHARY D. and FELDMAM M., 1962. Hybridization between amphidiploids and the evolution of polyploids in the wheat (*Aegilops-Triticum*) groups. *Evolution* 16: 44-61.
- ZIELINSKI Q.B., 1968. Techniques for collecting, handling, germinating and storing of pollen of the FILbert (*Corylus spp.*). *Euphytica*, 17: 121-125.

## INDEX BIBLIOGRAPHIQUE

Abbott	223, 224
Abdalla et Roberts	211, 212
Agnew	113, 119
Allard	200, 202
Anderson	13
Antonovics	20
Aquadro et Avise	186
Aubertie et al.	227
Autran et Bourdet	175
Bajaj	225
Bancilhon	220, 225
Barnabas et Rajki	216
Bass	210
Behnke	227
Benard	217
Bendich et Mc Carthy	188, 189
Bennett	65, 197
Benzecri et al	271
Berthou	187
Bigot	222, 223
Black	94
Bogyo et al.	128
Bonga	224
Bouzid et Lasram	224
Brewbaker	216
Brown	124
Brunken et al.	197
Burton	201
Button et Botha	224
Carlson	224
Cauderon	157, 160
Carvalho	225
Cavalli-Sforza et Bodmer	53
Champagnat	224
Charrier	115, 164, 225
Chevalier	116
Chiang	218
Chu et Oka	163
Clifford	100
Cochran	104, 152
Combes	128, 157, 163, 164
Couturon	209
Couturon et al	219
Cox	152
D'Amato	226
Darwin	7

Day	104
De Vienne	179
De Wet	9
Diday	141
Dinoor et Peleg	154
Dobzhansky	11
Earle et Langhans	225
Engels	187
Farmer et Barnett	217
Favre	224
Federer	152
Feldman et Christiansen	45
Figier	220
Fischer et Gaudmann	92
Flavell et al.	188, 189
Flor	90, 92, 93, 94, 104
Frankel	91, 194
Frankel et al	197, 216, 217
Galzy	224, 225, 227
Gautheret	222, 223
Ghorrai	210
Gilmer et Gregor	11
Goujon	92
Graham et al	98
Grassias	166
Grouzis	113
Halperin et Wetherell	222
Haoussou et Toure	220
Hardy-Weinberg	39
Harlan	8, 9, 13, 68, 69, 90, 91, 194, 195, 197
Harrington	205, 215, 216
Harris et Hopkinson	177
Heller	223
Henshaw	225
Heslop-Harrison	218
Holly	320
Holman et Brubaker	215
Hooker	100
Hu	159
Hussey	223, 224
Ishii et Mitsukuri	159
Ito	213, 214, 215
Iwanami	218
Jain et al	194
James	213
Johr et Vasil	216

Kahler et al	46
King	217
King et al	174, 177, 184
Konzag	249
Kuckuck	198
Kurata et Omura	158, 159
Lanaud	166
Leblanc	53, 62
Leigh-Brown et. Langley	185
Leppik	90
Leroy	114
Letouzey	117
Lewis	211
Lewontin	183
Linskens	216
Linsmaier et Skoog	227
Loegering et al	154
Louarn	161
Malcolmson et Black	94
Mangenot	19, 205, 206
Marchoux et al.	227
Margara	223
Marshall et Brown	127, 176, 184, 300
Martin	204, 218, 227
Martin et Ellingboe	105
McClintock et al	158
Mettler et Gregg	11
Moore	210
Morel	22.2, 223, 228
Morel et Martin	204, 222, 227
Morishima et al	162
Morrison et Majhatly	163
Moti	216
Murashige et al.	223, 225
Nayar	159
Nei et al	40, 46
Nelson et al	104, 105, 154
Nevo et al.	127
Nobecourt	222
Nozeran	109, 203
Nozeran et Bancilhon	222, 224, 225
O'Farrell	185
Okamoto	163
Parlevliet et Zadoks	104, 105
Pernès	14, 47, 57, 58, 195, 197, 200, 220
Person	99

Philipps	164
Pierik	224
Porceddy	128
Portères	110, 116
Purseglove	115
Qualset	91
Reinert et Bajaj	223
Roberts	206, 207, 208, 209, 210, 211, 216
Roberts et al	206, 208, 210, 211, 213
Robinson	102
Roche	203
Roux	220
Sacristan	226
Sakai et al	210, 214
Seckinger et al	224
Second	195
Second et al	179
Sen	159
Sidhu	93
Simmonds	115, 199
Soldatov	211
Stanley et Linskens	216, 217, 218
Stokes	223, 224
Street	223
Strigemura	225
Thompson	206, 213
Toxopeus	96
Tsai-Ying Cheng	224
Turner	223
Ullrich	100
Van Der Plank	90, 93, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106
Van Der Vossen	209
Vasil et al.	221
Vavilov	7, 68, 69, 90. 194
Vedel et al.	171, 186
Vessereau	152
Veyret	205
Villiers	208, 211
Villiers et al	211, 212
Visser	216
Visser et al.	216
Vogel	220
Walyaro et al.	217
Watson	91

Weir et al.	43, 44
Wellman et Blaisdell	100
White	222, 223
Wilkes	197, 198
Williams et al.	210
Withers et Street	225
Wold	213
Wright	11, 48
Yarchuck	210
Yasui	159
Zadoks	104
Zielenski	217

## INDEX DES TERMES SCIENTIFIQUES

adaptation	49
ADN	167, 171, 186
ARN	168
agriculture traditionnelle	197
aire d'origine	V centre d'origine
agressivité parasitaire	100
analyse	
des correspondances	279
des nuées dynamiques de Diday	275
discriminante	280
de variance hiérarchisée	143
en composantes principales	272
apomictique	195
apomixie	58
autogamie	34, 43, 59
base de données	235
gestion	239, 247
logiciels	268
programmation	257
ressources génétiques	256
vocabulaire	250
bivalent	166
bouturage	221
caryotype	157
centres de ressources génétiques	305
Bari	306
Taiwan	309
Tapioszele	317
Tsukuba	313

centre	
de diversité	67
d'origine	7, 67
de conservation	296, 301
chromosomes	
accidents chromosomiques	211
classification	140, 270
collection	
évolution théorique	56, 65
de conservation	197
complexe d'espèces	9-16, 130, 268
organisation géographique	64
compartiment	11, 66
conservation	
des graines	205
du -pollen	215 - 219
in vitro	221
des ressources génétiques	87, 193 - 233
culture in vitro	106, 221
cytogénétique	156
dendrogramme	140, 274
dérive	57, 64
descripteur	244
déséquilibre gamétique	31 - 49
coefficients	35
évolution	41
global	42
distance génétique	19, 24
de Nei	24 - 31
domestication	8, 30, 64
échantillonnage	121
écosystème	194
électrophorèse	175 - 186
bidimensionnelle	185
enzyme	
de restriction	186
alcool déshydrogénase	62
glumate déshydrogénase	178
peroxydase	181
évaluation	135 - 189
agronomique	137
en collection	139 - 142
génétique	142 - 189
cytogénétique	156
biologie moléculaire	167 - 189
fichier de données	264
flore	115
flux de gènes	12 - 16
fréquence allélique	20
gène de résistance	92

génétiq <u>ue</u> quantitative	142, 268
germination (taux de)	214
gestion de bases de données	238, 247
herbier	115
hétérozygotie	143
hybridation	159
de l'ADN	187
hybride interspécifique	7
interaction hôte-pathogène	92
introgression	56, 86
isolat	84
linkage	34
marqueur	33
mildiou	91, 94
mode de reproduction	33, 46, 129
modèles	
informatiques	251
en génétique des populations	47
multiplication	
sexuée	199
asexuée	202
végétative	219
mutation	212
nématode	93
nuées dynamiques	276
ordinateur	254
parasite	
diversité	88
obligatoire	96
phytopathologie	153
ploidie	157
pollen	215
polymorphisme	143
enzymatique	28, 123, 174
polyploidisation	163
prospection 107 - 121	
bilan	121
itinéraire	117
stratégie	107, 118
protéines	175
quarantaine	228 - 233, 309
recombinaison génétique	33

réservoir massai	201
résistance	
aux maladies	90 - 106
horizontale	99, 101
test	101
verticale	94, 96, 99, 104
ressources génétiques	
formation des personnels	319
rouille	93
sélection stabilisatrice	96
spéciation	16
stabilité génétique	225
stérilité mâle	59
taxonomie	268
numérique	270
unité taxonomique	7
variabilité	139
enzymatique	183
virulence	89
zone centrale	68
marginale	68

## INDEX DES VÉGÉTAUX (EN LATIN)

<i>Acer pseudoplatanus</i>	225
<i>Adiantum</i>	222
<i>Aegilops</i> (genre)	15, 65, 67, 126, 151
<i>Aegilops longissimum</i>	151
<i>Aegilops speltoides</i>	9, 12, 151
<i>Allium cepa</i>	206
<i>Betula verrucosa</i>	217
<i>Citrus</i>	206
<i>Coffea</i> (genre)	116, 162
<i>C. arabica</i>	92, 110, 119, 164, 168, 195, 203, 206, 225
<i>C. canephora</i>	114, 161, 162, 203
<i>C. congensis</i>	121, 161, 203
<i>C. eugenioides</i>	118, 161, 203
<i>C. humilis</i>	116
<i>C. liberica</i>	161
<i>Coix lachryma jobi</i>	157
<i>Cola nitida</i>	206
<i>Corylus</i>	217

<i>Crepis capillares</i>	226
<i>Cucumis melo</i>	210
<i>Daucus carota</i>	222
<i>Elaeis guineensis</i>	206
<i>Euchlena mexicana</i>	20, 29, 30, 66, 196, 197
<i>Eucoffea</i>	91
<i>Gossypium</i>	205
<i>Hedysarum</i>	220
<i>Hevea brasiliensis</i>	205, 206
<i>Hieracium</i>	210
<i>Hordeum distichum</i>	206, 208, 210, 211
<i>H. spontaneum</i>	124, 126
<i>H. vulgare</i>	206
<i>Lactuca sativa</i>	206, 208, 211
<i>Leersia</i>	114
<i>Lepidium</i>	210
<i>Linum</i>	92
<i>Lycopersicon esculentum</i>	206
<i>Maximae (complexe)</i>	195
<i>Medicago sativa</i>	210, 217
<i>Mimosa glomerata</i>	205
<i>Nelumbo</i>	205
<i>Nicotiana</i> (genre)	86
<i>N. glauca</i>	222
<i>N. langsdorffii</i>	10, 222
<i>N. bonariensis</i>	10
<i>N. tabacum</i>	226, 227
<i>Oryza</i> (complexe)	86
<i>O. breviligulata</i>	14, 15, 114, 196, 270, 282
<i>O. glaberrima</i>	14, 15, 86, 114, 115, 195, 196, 197, 211, 270, 282
<i>O. indica</i>	86, 195, 196, 282
<i>O. japonica</i>	86, 195, 270, 282
<i>O. longistaminata</i>	86, 114, 196, 282
<i>O. perennis</i>	67, 181
<i>O. sativa</i>	86, 114, 115, 158, 159, 162, 195, 196, 197, 206, 207, 208, 209, 210, 211, 270, 282
<i>O. sativa javanica</i>	86, 270
<i>O. stapfii</i>	7, 86, 196
<i>Panicum</i> (genre)	59, 86, 113, 115, 119, 157, 164, 195, 205, 214

<i>P. infestum</i>	114
<i>P. maximum</i>	11, 12, 14, 95, 109, 110, 114, 121, 130, 164, 169, 195, 220
<i>P. trichocladum</i>	114
<i>Paspalum</i> (genre)	165
<i>Pennisetum</i> (genre)	16, 114, 157, 325
<i>P. mollissimum</i>	66, 67
<i>P. purpureum</i>	114
<i>P. typhoides</i> ( <i>americanum</i> )	12, 66, 114, 201
<i>P. violaceum</i>	12, 114
<i>Phyllanthus</i> (genre)	18, 220, 225
<i>Pisum sativum</i>	206, 211
<i>Prunus cerasus</i>	217
<i>P. nigra</i>	217
<i>Psophocarpus tetragonolobus</i> gombo	113
<i>Pyrus communis</i>	216, 217
<i>P. malus</i>	217
<i>Quercus</i>	206
<i>Raphanus brassica</i>	210
<i>Saccharum</i>	226
<i>Setaria</i> (genre)	67, 325
<i>S. italica</i>	66
<i>S. viridis</i>	66
<i>Solanum</i> (genre)	102
<i>S. demissum</i>	91, 94, 97, 98, 103, 104, 105
<i>S. tuberosum</i>	91, 94, 97, 98, 105, 227
<i>Sorghum album</i>	10
<i>S. bicolor</i>	10
<i>S. halepense</i>	10
<i>Taxus</i>	206
<i>Thalictrum</i>	91
<i>Theobroma cacao</i>	205, 206
<i>Triticum</i> (genre)	15, 65, 67, 126, 151
<i>T. aestivum</i>	206, 210
<i>T. beoticum</i>	65
<i>T. durum</i>	9, 12
<i>T. monococcum</i>	65
<i>Tripsacum dactyloides</i>	157
<i>Vicia faba</i>	206, 211
<i>Vitis vinifera</i>	216, 217
<i>Zea mays</i>	20, 29, 30, 154, 196, 197, 210

Imprimé en **Belgique**





agen  
cultu

coopération  
et technique

13, quai An é Citroën - 75015 Paris

ISBN t . me 2 92-9028-043-3