Les outils d'évaluation de la diversité génétique et phénotypique

Marianne Lefort-Buson, Yannick Hebert, Catherine Damerval

▶ **To cite this version:**

Marianne Lefort-Buson, Yannick Hebert, Catherine Damerval. Les outils d'évaluation de la diversité génétique et phénotypique. Agronomie, 1988, 8 (3), pp.173-178. hal-00885086

HAL Id: hal-00885086
https://hal.science/hal-00885086

Submitted on 11 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Agronomie, 1988, **8** (3), 173-178.

SYNTHÈSE

Les outils d'évaluation de la diversité génétique et phénotypique

Marianne LEFORT-BUSON, Yannick HEBERT, Catherine DAMERVAL

C.N.R.S.-I.N.R.A.-U.P.S., G.I.S. Le Moulon, Ferme du Moulon, F 91190 Gif-sur-Yvette
I.N.R.A., Station d'Amélioration des Plantes Fourragères, Centre de Recherches de Poitou-Charentes, F 86600 Lusignan

RÉSUMÉ

La description et la mesure de la diversité génétique sont des préalables indispensables à la définition de stratégies d'amélioration ou de gestion des ressources. Le généticien est, à l'heure actuelle, en mesure d'apprécier cette diversité tant au niveau agronomique qu'à plusieurs niveaux moléculaires (isozymes, protéines dénaturées séparées par électrophorèse bidimensionnelle, fragments de restriction de l'ADN), et divers outils mathématiques lui permettent de valoriser les informations recueillies.
Dans cette optique, la spécificité de l'information apportée par chacun des niveaux d'analyse de la variabilité est examinée. Le point est fait sur l'utilisation en sélection de différents outils d'évaluation de la variabilité génétique (indices de distance, indicateurs de divergence, analyses multidimensionnelles). L'accent est mis en particulier sur les possibilités de prédiction de l'hétérosis ou des performances hybrides, ainsi que sur les problèmes de sélection récurrente et de gestion des ressources génétiques.

Mots clés additionnels : *Distance génétique, sélection récurrente, prédiction, ressources génétiques, valeur en croisement, analyse multivariable, parenté, isozyme, protéine dénaturée, fragment de restriction de l'ADN.*

SUMMARY

Tools for assessment of genetic and phenotypic diversity.

Genetic diversity has to be described and measured in order to set breeding strategies and manage genetic resources. Nowadays, several levels of variability (agronomic traits, proteins, isozymes, DNA) can be used and mathematical tools are available to make better use of the genetic information collected. With this aim, the specific contribution for each level of analysis of variability is considered. An assessment is made of the value for plant breeding of genetic tools for determining variability such as distance indices, divergence indicators or multidimensional analyses. Particular attention is paid to the prediction of heterosis and hybrid performance, and to recurrent selection and the management of genetic resources.

Additional key words : *Genetic distance, recurrent selection, prediction, genetic resource, combining ability, multivariate analysis, kinship, isozyme, denatured protein, DNA restriction fragment.*

I. INTRODUCTION

Dans tout schéma d'amélioration génétique, le sélectionneur ne peut agir qu'en présence d'une certaine diversité génétique des caractères. Mais, le plus souvent, ceux-ci se révèlent complexes : variation continue, nombreux gènes en cause, et interactions de l'effet des gènes avec les facteurs du milieu. D'où la nécessité de quantifier la variation observée et de la décrire à l'aide d'un nombre plus réduit de paramètres pour qu'elle devienne plus facilement exploitable en sélection. Jusqu'ici, le problème était surtout abordé avec les caractères agronomiques, qu'ils soient qualitatifs ou quantitatifs. Actuellement, une telle approche se trouve enrichie de l'apport de nouvelles techniques dont l'intérêt réside surtout dans la révélation de niveaux différents d'expression de la variabilité génétique. Mais la maîtrise de ces méthodes doit s'accompagner d'une meilleure connaissance des caractéristiques de la variabilité mise en évidence (importance, impact de la sélection naturelle ou artificielle, relations avec les caractères agronomiques).

Dans ce qui suit, nous présenterons tout d'abord les outils permettant aujourd'hui d'estimer la diversité génotypique. Puis nous évoquerons brièvement, compte tenu des contributions apportées par chacun d'eux, leur intérêt pour divers objectifs de sélection, en nous référant à quelques exemples. Les problèmes liés à la gestion et à l'exploitation des ressources génétiques seront abordés.

II. ÉVALUATION DE LA DIVERSITÉ GÉNOTYPIQUE

La diversité génétique ou phénotypique peut être estimée, soit à partir de données recueillies sur les entités considérées ou sur leurs descendances, soit à partir d'informations connues *a priori* sur ces unités (origine géographique, généalogie).

A. Utilisation de données connues *a priori*

La connaissance de l'origine géographique des entités sélectionnées peut rendre partiellement compte de leur diversité génétique. Il est probable que la sélection naturelle ou artificielle pour un caractère dans des milieux assez diversifiés aura conduit à fixer, d'une part des allèles non identiques pour un locus donné, d'autre part des associations de gènes favorables différentes pour l'ensemble des locus gouvernant ce caractère. Ainsi, la diversité génétique est d'autant plus grande que les milieux de sélection d'une part et les idéotypes sélectionnés d'autre part sont différents.

Par ailleurs, la connaissance de la généalogie des entités permet de donner un ordre de grandeur de leur parenté (F ; Malécot, 1948), ou inversement de leur non-apparentement ou distance (1-F). La diversité sera d'autant plus élevée que les distances entre entités 2 à 2 seront proches de 1. Cet indice de distance donne cependant une image biaisée de la diversité potentielle. En effet, il est calculé en supposant l'absence de sélection gamétique et zygotique ce qui, par définition, n'est jamais réalisé dans un schéma de sélection, et en faisant abstraction de toute interaction entre le génotype et le milieu. Enfin, les généalogies des lignées considérées n'étant guère connues au-delà des 30 dernières années, l'hypothèse d'absence de parenté est souvent posée pour les couples de géniteurs à l'origine des premiers schémas de sélection ; or, il est probable que des écotypes géographiquement proches aient eu un ancêtre commun auparavant, et soient donc partiellement apparentés.

B. Données recueillies *a posteriori* sur les individus étudiés

Nous distinguerons les données recueillies au niveau phénotypique sur la plante de celles recueillies à différents niveaux moléculaires (électrophorèses de protéines et/ou d'enzymes, profils de restriction de l'ADN). Quelles que soient les données, cependant, les méthodes statistiques utilisées pour estimer la diversité font souvent appel à des analyses multidimensionnelles (analyse en composantes principales, analyse des correspondances, analyse discriminante, calcul de distances et d'indices de distance multivariables).

1. *Données phénotypiques*

Le sélectionneur s'intéresse classiquement à plusieurs caractères des unités à sélectionner ou de leurs descendances. Une analyse multidimensionnelle sur ces mesures conduit à définir un nouvel espace de représentation de la variation phénotypique. Le référentiel de cet espace peut être interprété en termes de combinaisons linéaires des caractères de départ. Le sélectionneur peut aussi calculer des indicateurs de divergence entre unités (HÉBERT & VINCOURT, 1985 ; MASSON, 1986). La diversité sera d'autant plus grande que les points représentant les phénotypes seront « éloignés » les uns des autres dans l'espace de représentation défini, ou encore que les indicateurs de divergence seront grands.

Ces principes étant posés, il importe de noter que l'estimation de la diversité est fortement conditionnée par 3 éléments :

— les facteurs étudiés, définis à partir des individus eux-mêmes (valeurs propres) ou de leurs descendances (valeurs en croisement, aptitudes générale et spécifique à la combinaison),
— les caractères utilisés,
— la métrique employée.

a) Les facteurs

Le choix des facteurs est relativement délicat et conditionne assez souvent la signification de la variation génétique révélée. Ainsi, utiliser l'aptitude générale à la combinaison, l'aptitude spécifique ou encore des valeurs en croisement avec plusieurs testeurs donne accès à différentes sources de variation.

L'estimation de la diversité phénotypique sera d'autant plus représentative de la diversité génétique que les valeurs phénotypiques seront de bons estimateurs des valeurs génétiques. Pour ce faire, les évaluations plurilocales ou pluriannuelles réalisées sur les descendances des entités étudiées et non sur les entités elles-mêmes donneront une meilleure idée des valeurs génétiques de ces dernières et donc de la diversité génétique qu'elles représentent. Cet intérêt porté à la valeur en croisement s'appuie également sur la relation qui lie l'hétérosis à la distance génétique, relation discutée ultérieurement.

b) Les caractères

Le choix des caractères, leur nombre et leur nature posent un certain nombre de questions auxquelles seul le sélectionneur, confronté à des objectifs spécifiques, peut répondre :

— s'il veut avoir une idée de la diversité globale de son matériel, il utilisera de nombreux caractères. L'intégration d'un grand nombre de variables dans les analyses de données multidimensionnelles, si elle permet d'analyser globalement la diversité, rend délicate l'interprétation des composantes principales définies par l'analyse. Une autre méthode consiste alors à répartir les caractères en groupes ayant subi des pressions de sélection à peu près équivalentes, et à travailler ensuite indépendamment pour chaque groupe : différentes représentations de la diversité sont alors obtenues. Le choix des voies d'exploitation du matériel génétique résultera de la confrontation de ces représentations ;

— s'il souhaite conserver une variabilité maximale pour tous les caractères quelle que soit la valeur des géniteurs pour les caractères économiquement importants tels que le rendement (par exemple, pour pouvoir répondre, dans 5 ou 10 ans, à un nouveau problème parasitaire), la diversité ne sera appréciée que pour les caractères secondaires à conserver, non ou peu corrélés à la productivité.

— s'il veut disposer d'une variabilité importante pour différents caractères, en imposant une productivité moyenne minimum à l'ensemble de la population (constitution de populations de départ dans un schéma de sélection récurrente), il recherchera alors des génotypes relativement productifs, choisis aussi pour assurer une grande diversité pour les autres caractères étudiés.

Quel que soit l'objectif, si l'on est amené à introduire un grand nombre de caractères, on s'efforcera de réduire autant que possible, compte tenu des résultats fournis par l'analyse, la dimension de l'espace de représentation (celle-ci dépend de la part de variation expliquée par chacun des axes). Les interprétations seront dans tous les cas facilitées par une dimension réduite de l'espace retenu (2, 3 ou 4), c'est-à-dire par un petit nombre de directions de variation.

c) La métrique

Les conclusions dépendront fortement de la métrique choisie. La métrique euclidienne et la métrique du Chi^2 sont couramment utilisées dans l'analyse en composantes principales et l'analyse des correspondances, respectivement. Dans certains cas, il peut s'avérer plus judicieux de tenir compte de la structure de variance-covariance du groupe d'individus étudiés. C'est le principe de l'analyse factorielle discriminante qui permet de prendre en compte la variation relative des caractères dans les groupes homogènes d'individus ainsi que des corrélations entre les mesures phénotypiques. Dans d'autres cas, il est justifié de tenir compte de la structure de la population F_2 issue des individus étudiés, afin que l'écart phénotypique entre les individus soit ramené à celui qui peut être observé dans une ségrégation mettant en jeu une large variabilité génétique, ou également afin qu'un individu présentant une association rare de caractères dans la population F_2 soit distingué de l'ensemble des autres (GOODMAN, 1968).

L'estimation de la diversité est donc fortement conditionnée par le choix des caractères mesurés et la définition du ou des espaces de représentation (choix de la dimension et de la métrique). HÉBERT & VINCOURT (1985) en ont montré, par un exemple, les conséquences sur l'interprétation des distances. Par ailleurs, les indicateurs de divergence, qui synthétisent en une valeur unique l'ensemble des informations recueillies, sont souvent restrictifs pour l'étude de la diversité. Ils seront par contre plus fréquemment utilisés dans les études de prédiction de l'hétérosis.

2. *Données moléculaires*

Trois types de techniques donnent actuellement accès à l'étude de la variabilité au niveau moléculaire : l'électrophorèse unidimensionnelle, qui permet de révéler la diversité isozymatique ou de certaines protéines particulières (de réserve, par exemple), les profils de restriction caractéristiques de segments d'ADN (PLFR) et l'électrophorèse bidimensionnelle des protéines dénaturées.

a) Les isozymes

Les isozymes sont les premiers marqueurs moléculaires à avoir été étudiés et la découverte de leur important polymorphisme, il y a plus de 20 ans, a suscité et suscite encore des débats passionnés.

Des associations entre variation morphologique et variation isozymatique ont parfois été mises en évidence (MARSHALL & ALLARD, 1970 ; STUBER *et al.*, 1982 ; FREY *et al.*, 1986) ; il ne semble pourtant pas exister de règle générale dans ce domaine (EL-KASSABY, 1982 ; KAHLER, 1983 et pour revue DAMERVAL & DE VIENNE, 1985). Quoi qu'il en soit, ces relations, lorsqu'elles existent, semblent rarement le fait d'une part prépondérante des allozymes dans les processus métaboliques sous-jacents aux caractères phénotypiques analysés. La neutralité des allozymes vis-à-vis de la sélection naturelle et leur absence d'effet phénotypique visible peuvent être admises d'une façon générale. Ils se comportent comme des marqueurs de segments chromosomiques, à preuve la possibilité de déséquilibre de liaison associant différents caractères quantitatifs avec les mêmes locus d'enzymes (POLLACK *et al.*, 1984 ; KAHLER & WHERHAHN, 1986).

Le polymorphisme des allozymes intègre une partie des variations de l'ADN signifiantes au niveau protéique. L'une des plus importantes limitations de cette technique de marquage est le faible nombre de locus accessibles à l'analyse (30 chez le blé, 36 chez le maïs, 28 chez la tomate).

b) Le polymorphisme de la longueur des fragments de restriction de l'ADN (PLFR)

L'analyse du polymorphisme de la longueur des fragments de restriction de l'ADN est une technique de marquage génétique récente. La signification des variations au niveau de l'ADN non traduites au niveau des protéines reste très controversée (DOOLITTLE, 1982 ; NINIO, 1983). Le polymorphisme révélé par cette technique est important et variable selon les espèces (HELENTJARIS *et al.*, 1985).

Tout comme les allozymes, le PLFR fournit des marqueurs de segments chromosomiques. Mais alors que dans le premier cas un petit nombre de locus seulement est accessible à l'analyse, le PLFR permet un marquage beaucoup plus extensif du génome ; en effet, toute séquence d'ADN, codante ou non, peut théoriquement être utilisée comme sonde (BECKMANN & SOLLER, 1983). De plus, ces 2 types de marquage n'apportent pas nécessairement une information de même nature : les allozymes sont le produit d'une certaine catégorie de gènes de structure, non répartis au hasard sur le génome, alors que les sondes utilisées pour le PLFR représentent un échantillon aléatoire de séquences, indépendant de leur rôle fonctionnel.

L'utilité du PLFR pour l'étude des caractères d'intérêt agronomique est à l'heure actuelle discutée (SOLLER & BECKMAN, 1983 ; ELLIS, 1986). Récemment, OSBORN *et al.* (1987) ont donné un exemple de liaison entre un « locus » de PLFR et des gènes déterminant la teneur en certains composés du fruit chez la tomate. L'efficacité de ce type de marquage paraît dépendre de 3 facteurs au moins : un petit nombre de gènes à effets majeurs déterminant le caractère que l'on cherche à prédire ; l'absence d'interaction entre les locus contrôlant le caractère ; le fait de travailler chez des autogames (ou des lignées) où le déséquilibre de linkage serait maximisé par rapport à des allogames.

c) Données issues de l'électrophorèse bidimensionnelle des protéines dénaturées

L'électrophorèse bidimensionnelle des protéines dénaturées permet de visualiser conjointement plusieurs centaines de produits de gènes de structure, et de suivre leur variation entre génotypes au cours du développement et/ou selon l'organe chez un même individu. La variation quantitative des protéines entre génotypes reflète le polymorphisme d'éléments génétiques qui contrôlent l'expression de gènes de structure (régulation aux niveaux transcriptionnels, traductionnels...). Ceci revêt un intérêt tout particulier dans l'optique de la recherche de marqueurs moléculaires de la variation morphologique, le rôle majeur des facteurs de régulation ayant maintes fois été suggéré dans ce domaine (WILSON *et al.*, 1977, MCINTYRE, 1982).

Ce type de marquage génétique devrait donc apporter une information de nature différente de celles fournies par le PLFR et le polymorphisme enzymatique. D'ores et déjà, des phénomènes de dominance au niveau des quantités de protéines ont été mis en évidence (LEONARDI *et al.*, 1987) ; ni le PLFR, ni la variation allozymique telle qu'elle est couramment envisagée ne permettent de révéler des écarts à l'additivité dans les descendances hybrides. D'autre part, il a été montré, pour 5 lignées de maïs, que la distance entre lignées, estimée au niveau des quantités de protéines, était corrélée à la distance calculée à partir de plusieurs caractères agronomiques (DAMERVAL *et al.*, 1987). La prise en compte de plusieurs centaines de spots protéiques pour évaluer la diversité génétique à l'intérieur d'un lot d'une centaine d'individus (ou moins) pose question. Une première étape consiste donc à sélectionner les spots les plus intéressants selon différents critères : existence d'un polymorphisme, absence de redondance de l'information apportée, nature de la variation qualitative et/ou quantitative. On peut ensuite pratiquer différentes analyses sur les spots sélectionnés et calculer des indices de distance (HOFMANN *et al.*, 1986). Ultérieurement, le déterminisme génétique des variations observées pourra être étudié.

La diversité génétique estimée à partir de ces 3 types de marquage moléculaire recouvre différents aspects du polymorphisme. Si leur spécificité est confirmée dans le futur (trop peu de résultats expérimentaux sont encore disponibles pour les méthodes de profils de restriction et d'électrophorèse bidimensionnelle), on voit clairement, là encore, que l'évaluation de la diversité génétique pour un matériel donné sera fortement conditionnée par les objectifs du sélectionneur. Nous rappelons, dans la partie suivante, les objectifs les plus classiques et discutons de l'intérêt de différents prédicteurs, compte tenu de leur coût.

III. DIVERSITÉ GÉNÉTIQUE ET AMÉLIORATION DES PLANTES

L'estimation de la variabilité génétique est un problème fondamental en amélioration des plantes. On le rencontre à tous les niveaux d'un schéma de sélection :

— au départ d'un programme, lorsqu'on recense la variabilité disponible, soit pour une utilisation directe, soit pour la constitution de populations de base ;

— au cours d'un schéma de sélection récurrente, lorsqu'on veut maintenir la variabilité initiale ;

— en sortie variétale, lorsqu'on veut prédire la valeur potentielle des familles sélectionnées.

L'utilisation des marqueurs phénotypiques pour l'appréciation de la diversité génétique, en particulier dans l'étude de la relation entre l'hétérosis et la distance, a déjà été largement débattue (cf. LEFORT-BUSON, 1985*c*, pour revue). Nous n'évoquerons donc que les résultats les plus marquants et les plus récents.

L'apport des marqueurs moléculaires pour la sélection pourrait se concevoir à 2 niveaux :

— accélérer les étapes de la sélection grâce à des liaisons marqueurs — caractères sélectionnés. Ceci serait particulièrement intéressant pour la sélection de lignées performantes ou la prédiction de la vigueur hybride. Les marqueurs auxquels on s'intéressera alors devront être soit à la fois neutres vis-à-vis de la sélection naturelle et en déséquilibre de linkage avec des locus contrôlant les caractères recherchés, soit non neutres et impliqués dans la réalisation de ces caractères ;

— apprécier la ressemblance génétique d'entités destinées à constituer des collections (gestion des ressources génétiques) ou des populations à améliorer. Dans ce cas, des marqueurs neutres comme les isozymes ou les PLFR seront précieux.

A. Sélection et prédiction

1. *Prédiction de la valeur hybride*

D'un point de vue théorique, plusieurs auteurs ont montré que l'hétérosis était d'autant plus élevé que les entités croisées étaient génétiquement différentes (LEFORT-BUSON, 1985*a*). Le sélectionneur a donc intérêt à rechercher une certaine distance génétique pour favoriser la complémentarité en croisement et, par là même, un hétérosis plus important. Toutefois, différents auteurs ont montré que des croisements entre des parents trop divergents pouvaient conduire à une régression de l'hétérosis, ceci en raison d'un fonctionnement disharmonieux des allèles en présence à certains locus (LEFORT-BUSON, 1985*b*).

Sur le plan pratique, les résultats sont extrêmement variables (LEFORT-BUSON, 1985*c*) :

— les études monolocales réalisées avec les mesures directes des individus (utilisant souvant la distance de Mahalanobis) montrent que l'hétérosis est, dans certains cas, une fonction croissante de la distance et, dans d'autres cas, non relié à celle-ci. La diversité des résultats obtenus par plusieurs auteurs sur la même espèce témoigne bien, entre autres, des difficultés d'appréciation des génotypes au travers des phénotypes ;

— les études réalisées sur les descendances des entités vont presque toujours dans le sens d'une relation positive entre l'hétérosis et la distance ; ces résultats reflètent probablement une meilleure appréciation des valeurs génétiques des individus à partir de leurs descendances, relativement à celle obtenue à partir de leur valeur propre.

Dans le cas des isozymes, peu de résultats favorables ont été rapportés pour la prédiction de l'hétérosis (BRUNEL, 1985) probablement à cause du petit nombre de locus étudiés et surtout, du grand découplage existant entre la variation allozymique et la variation phénotypique d'un caractère complexe. Cependant, il a été récemment montré que la valeur prédictive des isozymes était meilleure dans le cas où les génotypes parentaux dérivaient d'un même pool ancêtre que dans le cas où ces génotypes dérivaient de pools différents (FREY *et al.*, 1986 ; CHARCOSSET, 1986 ; STRAUSS, 1986). Ceci renforce l'idée d'un marquage par les isozymes via un déséquilibre gamétique avec des locus impliqués dans les caractères quantitatifs sélectionnés.

L'intérêt d'utiliser simultanément plusieurs prédicteurs de l'hétérosis ou des performances hybrides pour améliorer l'efficacité de la prédiction a aussi été abordé. *A priori,* les résultats seront d'autant meilleurs que les prédicteurs apporteront des informations de nature différente. Chez le colza, dans le cas du rendement en grains d'hybrides F_1, LEFORT-BUSON *et al.* (1986) ont souligné l'intérêt de combiner des indices multivariables de distance entre lignées évalués d'une part sur les valeurs propres des lignées et d'autre part sur leurs valeurs en croisement avec diverses autres lignées. Dans le cas du rendement d'hybrides F_1 de maïs fourrage, CHARCOSSET (1986) a montré que les indices de distance (estimés à partir de données enzymatiques d'une part et à partir de données agronomiques d'autre part) apportaient un complément d'information net à la prédiction réalisée par la valeur propre des lignées ou par leur valeur en croisement avec 2 testeurs complémentaires.

2. *Prédiction de la valeur en lignée*

Les sélectionneurs de plantes autogames ont depuis longtemps cherché à croiser des lignées présentant des caractéristiques différentes pour obtenir, dans la descendance du croisement, des lignées porteuses des caractéristiques favorables des 2 lignées parentales. On peut alors se demander si des croisements réalisés entre lignées très « distantes » et de bonne valeur propre, ne fourniraient pas en moyenne des lignées meilleures que des croisements réalisés entre lignées « voisines », de bonne valeur propre également. La question peut alors être posée en termes de prédiction : y a-t-il un indicateur de divergence qui, parmi un ensemble de croisements potentiels, permettrait de trier ceux qui donneraient les meilleures lignées ? En d'autres termes, peut-on prédire la moyenne et la variance des lignées dérivables d'un croisement compte tenu de la « distance » entre les 2 parents du croisement ?

Malgré l'importance de la question, aucun résultat bibliographique n'a été trouvé sur ce sujet où tout reste encore à faire. Des résultats préliminaires favorables ont cependant été obtenus en comparant des descendances F_3 de 3 hybrides F_1 de colza entre lignées plus et moins apparentées (LEFORT-BUSON, non publié) : la moyenne et la variance des lignées dérivables du croisement entre lignées les plus divergentes sont supérieures à celle des 2 autres croisements.

B. Sélection récurrente

Au cours de différentes étapes, la méthodologie de sélection récurrente peut faire appel à des indices de distance pour mieux gérer et exploiter la variabilité.

Lors de la constitution des populations, le choix des géniteurs peut se faire en utilisant, conjointement aux performances agronomiques qui donneront une idée de la valeur moyenne de la population, des critères de distances estimés soit à partir de données moléculaires, soit à partir de caractères quantitatifs qui donneront des informations sur la diversité à l'intérieur de ces populations. Il est probable que les distances « moléculaires » seront surtout efficaces pour des gammes de divergence limitées, dans lesquelles le sélectionneur recherchera une diversité maximale. Les indices de distances « moléculaires » et « agronomiques » pourront aussi être utilisés pour constituer des populations homogènes qui seront menées en parallèle dans le cadre de programmes de sélection récurrente réciproque.

Au cours des cycles de sélection récurrente, les mesures de distances entre les géniteurs dans une même population pourront être utilisées afin de maintenir la variabilité au sein de celle-ci. De plus, les marqueurs moléculaires permettront de contrôler la dérive génétique, par exemple tous les 3 ou 4 cycles.

L'introduction de nouvelles sources de variabilité génétique peut aussi s'appuyer sur l'estimation des distances entre les entrées d'une part et entre les entrées et les populations en cours de sélection d'autre part.

C. Conservation des ressources

Dans la gestion des ressources génétiques, bien qu'encore peu utilisées, les distances génétiques pourraient être des outils précieux. Comme dans le cas de la sélection récurrente, elles permettraient de constituer des pools homogènes, en veillant à ne pas trop restreindre la diversité à l'intérieur de ceux-ci. Dans tous les cas, le choix des caractères à considérer implique une réflexion approfondie du sélectionneur sur l'importance présente et future de ceux-ci. Par la suite, le choix de nouvelles introductions à maintenir sera raisonné en tenant compte de leurs distances par rapport aux représentants des groupes déjà maintenus, de l'originalité du caractère qu'ils sont censés apporter et de l'évolution des objectifs de sélection.

IV. CONCLUSION

Le sélectionneur dispose donc de méthodes lui permettant d'apprécier d'un point de vue nouveau la structure de la variabilité génétique qu'il manipule. Ces outils l'aident à mieux valoriser la diversité génétique dans sa conservation et son utilisation. Malgré la rareté des expériences de prédiction réalisées jusqu'à ce jour, les prédicteurs biométriques et enzymatiques paraissent prometteurs.

Une bonne gestion des origines des entités sélectionnées (généalogie, origine géographique) devrait être réalisée quelle que soit l'espèce. C'est une information de base peu coûteuse, accessible à tous les sélection-

neurs et utile dans tout programme de sélection où un minimum de diversité est requis.

Par ailleurs, une meilleure exploitation de l'ensemble des données agronomiques recueillies au cours des expérimentations permettrait au sélectionneur d'obtenir des renseignements complémentaires sur la diversité potentielle de son matériel. La mise en œuvre systématique d'analyses multidimensionnelles simples, accompagnées de sorties graphiques claires, serait aussi un investissement peu coûteux conduisant à faciliter grandement la tâche du sélectionneur.

En ce qui concerne le marquage moléculaire, par contre, un certain nombre de questions restent encore sans réponse claire. En effet, à l'heure actuelle, seule peut être envisagée dans un programme de sélection l'utilisation de l'électrophorèse unidimensionnelle d'enzymes, son coût étant tout à fait raisonnable. L'analyse des profils de restriction de l'ADN est encore une technique trop récente et trop lourde pour que ses avantages puissent être précisés ; quant à l'électrophorèse bidimensionnelle des protéines dénaturées, elle reste encore aujourd'hui associée à des thèmes de recherche développés en laboratoire. Le coût d'un gel et celui de son exploitation automatique la maintiennent encore éloignée de l'analyse de routine.

Reçu le 22 mai 1987.
Accepté le 20 décembre 1987.

RÉFÉRENCES BIBLIOGRAPHIQUES

Beckman J. S., Soller M., 1983. Restriction fragment length polymorphism in genetic improvement : methodology, mapping and costs. *Theor. Appl. Gen.*, **67**, 35-43.

Brunel D., 1985. Distances génétiques et hétérosis. 5. Utilisation des marqueurs moléculaires, pp. 159-168. *In :* M. Lefort-Buson & D. de Vienne. *Les distances génétiques. Estimations et applications.* I.N.R.A., Paris, 181 p.

Charcosset A., 1986. *Prédiction de la performance d'hybrides F_1 chez le maïs (*Zea mays *L.).* Mémoire de fin d'études (INA-PG), 71 p.

Damerval C., de Vienne D., 1985. Divergence morphologique et divergence moléculaire. 1. Apport des marqueurs protéiques, pp. 61-80. *In :* M. Lefort-Buson & D. de Vienne. *Les distances génétiques. Estimations et applications I.N.R.A.* Paris, 181 p.

Damerval C., Hebert Y., de Vienne D., 1987. Is the polymorphism of protein amounts related to phenotypic variability ? A comparison of two-dimensional electrophoresis data with morphological traits in maize. *Theor. Appl. Gen.*, **74**, 194-202.

Doolittle W. F., 1982. Selfish DNA after fourteen months, pp. 3-28. *In :* G. A. Dover & R. B. Flavell. *Genome evolution.* Academic Press, N.Y., 382 p.

El-Kassaby Y., 1982. Associations between genotypes and quantitative traits in Douglas fir (*Pseudotsuga menziesii* (Mirb.) Franco). *Genetics*, **101**, 103-115.

Ellis T. H. N., 1986. Restriction fragment length polymorphism markers in relation to quantitative characters. *Theor. Appl. Genet*, **72**, 1-2.

Frey O. M., Stuber C. W., Goodman M. M., 1986. Use of allozymes as genetic markers for predicting performance in maize single cross hybrids. *Crop Sci.*, **26**, 37-42.

Goodman M. M., 1968. The races of maize. II. Use of multivariate analysis of variance to measure morphological similarity. *Crop. Sci.*, **8**, 693-698.

Hebert Y., Vincourt P., 1985. Mesures de la divergence génétique. 2. Distances calculées sur des critères biométriques, pp. 23-38. *In :* M. Lefort-Buson & D. de Vienne. *Les distances génétiques. Estimations et applications.* I.N.R.A., Paris, 181 p.

Helentjaris T., King G., Slocum M., Siedenstrang C., Wegman S., 1985. Restriction fragment polymorphisms as probes for plant diversity and their development as tools for applied plant breeding. *Plant Molecular Biology*, **5**, 109-119.

Hofmann J. P., Taylor J., Pascoe S., Bessières P., 1986. Evaluation de la divergence à l'aide des marqueurs protéiques révélés en électrophorèse bidimensionnelle. *Séminaire du Département de Génétique et d'Amélioration des Plantes de l'I.N.R.A.* Méribel, 11-15 mars 1986, 33-40.

Kahler A. L., 1983. Effect of half-sib and S_1 recurrent selection for increasing grain yield on allozyme polymorphisms in maize. *Crop Sci.*, **23**, 572-576.

Kahler A. L., Wehrhahn C. F., 1986. Associations between quantitative traits and enzyme loci in the F_2 population of a maize hybrid. *Theor. Appl. Genet.*, **72**, 15-26.

Lefort-Buson M., 1985*a*. Distances génétiques et hétérosis. 1. Mise en évidence d'une relation entre hétérosis et divergence génétique, pp. 111-118. *In :* M. Lefort-Buson & D. de Vienne. *Les distances génétiques. Estimations et applications.* I.N.R.A., Paris, 181 p.

Lefort-Buson M., 1985*b*. Distances génétiques et hétérosis. 2. Aspects théoriques, pp. 119-140. *In :* M. Lefort-Buson & D. de Vienne. *Les distances génétiques. Estimations et applications.* I.N.R.A., Paris, 181 p.

Lefort-Buson M., 1985*c*. Distances génétiques et hétérosis. 4. Utilisation des critères biométriques, pp. 143-158. *In :* M. Lefort-Buson & D. de Vienne. *Les distances génétiques. Estimations et applications.* I.N.R.A., Paris, 181 p.

Lefort-Buson M., Guillot-Lemoine B., Dattee Y., 1986. Heterosis and genetic distance in rapeseed (*Brassica napus* L.). Use of different indicators of genetic divergence in a 7 × 7 diallel. *Agronomie*, **6**, 893-844.

Leonardi A., Damerval C., de Vienne D., 1987. Inheritance of protein amounts : comparison of two-dimensional electrophoresis patterns of leaf sheaths of two maize lines (*Zea mays* L.) and their hybrids. *Genet. Res. Camb.*, **48** *(sous presse).*

McIntyre R. J., 1982. Regulatory genes and adaptation. Past, present and future. *Evol. Biol.*, **15**, 247-285.

Malecot G., 1948. *Les mathématiques de l'hérédité.* Masson et Cie, Paris.

Marshall D. R., Allard R. W., 1970. Enzyme polymorphisms in natural populations of *Avena fatua* and *A. barbata*. *Heredity*, **25**, 273-382.

Masson J. P., 1986. Notions de distance et d'indices de distance. Comparaison des matrices de dispersion. *Séminaire du Département de Génétique et d'Amélioration des Plantes de l'I.N.R.A.* Méribel, 11-15 mars 1986, 6-22.

Ninio J., 1983. L'explosion des séquences : les années folles 1980-1990. *Biochem. Syst. Ecol.*, **11**, 305-313.

Osborn T. C., Alexander D. C., Fobes J. F., 1987. Identification of restriction fragment length polymorphisms linked to genes controlling soluble solids content in tomato fruit. *Theor. Appl. Genet.*, **73**, 350-356.

Pollack L. M., Gardner C. O., Parkhurst A. M., 1984. Relationships between enzyme marker loci and morphological traits in two mass-selected maize populations. *Crop Sci.*, **24**, 1174-1179.

Soller M., Beckman J. S., 1983. Genetic polymorphism in varietal identification and genetic improvement. *Theor. Appl. Genet.*, **67**, 25-33.

Strauss S. H., 1986. Heterosis at allozyme loci under inbreeding and crossbreeding in *Pinus attenuata*. Genetics, **113**, 114-134.

Stuber C. W., Goodman M. M., Moll R. H., 1982. Improvement of yield and ear number resulting from selection at allozyme loci in a maize population. *Crop Sci.*, **22**, 737-740.

Wilson A. C., Carlson S. S., White T. J., 1977. Biochemical evolution. *Ann. Rev. Biochem.*, **46**, 573-639.