

CARACTERISATION VARIETALE Les marqueurs moléculaires : un nouvel outil pour l'inscription et la protection variétale ?

Varietal characterisation Molecular markers: a new tool for the registration and protection of plant varieties?

Oléagineux, Corps Gras, Lipides. Volume 8, Numéro 5, 502-5, Septembre - Octobre 2001, Dossier : Aspects des filières semencières Nord/Sud

Auteur(s) : Claire BARIL, Geves, Unité expérimentale de la Minière, 78285 Guyancourt Cedex, France.

Résumé : Le Geves (Groupe d'étude et de contrôle des variétés et des semences) a pour mission officielle de mener les études nécessaires à l'homologation des variétés végétales nouvelles, la protection juridique du droit des obtenteurs et la certification des semences avant leur commercialisation, dans le cas des espèces soumises à une certification réglementaire. L'inscription d'une variété nouvelle au catalogue officiel est une condition nécessaire à sa commercialisation. Pour être homologuée, une variété doit être distincte, homogène et stable (DHS) et apporter un progrès au niveau de sa valeur agronomique ou technologique (VAT). L'expertise des nouvelles variétés est effectuée pour le compte du CTPS (Comité technique permanent de la sélection), placé sous l'autorité du ministère de l'Agriculture. Le matériel végétal nouveau peut bénéficier d'une protection légale par la délivrance d'un certificat d'obtention végétale (COV) par le CPOV (Comité de protection des obtentions végétales), en application de la convention internationale de l'Upov (Union pour la protection des obtentions végétales). La protection juridique du droit de l'obteneur permet la valorisation de l'innovation variétale et, donc, celle des investissements de recherche. La production de semences certifiées suit des protocoles et respecte des normes de qualité très précises. Avant leur commercialisation, tous les lots de semences d'espèces certifiables sont échantillonnés et analysés. Les contrôles d'identité et de pureté variétale ainsi que les analyses de qualité des semences sont réalisés à la demande du SOC (Service officiel de contrôle et de certification), service technique du GNIS (Groupement national interprofessionnel des semences).

Summary : The use of molecular markers for the registration of new varieties and for the protection of plant breeders has been analysed in many research studies and widely discussed among the international seeds community. Before adopting these powerful varietal characterisation tools, the underlying methods and techniques used remain to be defined.

Keywords : molecular markers, official registration, varietal characterisation, genetic distances, prediction.

ARTICLE

Une évolution inévitable

Les analyses réalisées pour l'homologation et la protection des variétés et pour la certification des semences sont aujourd'hui fondées sur des observations phénotypiques, biochimiques ou physiques. Mais l'augmentation inéluctable du nombre des variétés des collections de référence (ensemble des variétés notoirement connues) et les difficultés croissantes rencontrées par les groupes d'experts du CTPS pour distinguer les variétés nouvelles de certaines espèces rendent l'approche moléculaire nécessaire. La perspective à court terme d'une redistribution des analyses DHS par espèce au niveau européen pousse le Geves à réfléchir davantage sur les moyens d'optimiser la procédure d'inscription actuelle pour la rendre plus compétitive vis-à-vis de ses homologues étrangères.

L'intérêt de la caractérisation moléculaire de la biodiversité apparaît dans tous les domaines de la génétique végétale : taxonomie (pour une meilleure compréhension de l'évolution des espèces), gestion des ressources génétiques (constitution de *cores* collections), amélioration des plantes (identification de groupes d'aptitude à la combinaison, sélection assistée par marqueurs) et, enfin, inscription et protection variétale (traçabilité, harmonisation). Bien que son efficacité ait été largement prouvée en taxonomie, en gestion des ressources génétiques et en amélioration des plantes, la caractérisation moléculaire ne parvient pas à être adoptée dans les procédures d'inscription et de protection des variétés.

Si le bien-fondé de l'utilisation des marqueurs moléculaires dans la procédure d'inscription n'est pas encore totalement admis, il est clair que l'approche moléculaire doit constituer l'un des éléments à la base de l'établissement de la dérivation essentielle d'une variété par rapport à une autre préalablement inscrite [1]. Le droit de l'obteneur et sa protection légale étant reconnus et appliqués par tous les pays signataires de la convention internationale de l'Upov, les modalités d'utilisation de ces caractères alternatifs doivent être définies à l'échelle internationale. Des groupes de travail transversaux ont été constitués à l'Upov pour réfléchir à l'évolution et à l'harmonisation des protocoles en liaison avec les représentants des obtenteurs, notamment au travers de l'Assinsel (Association internationale des sélectionneurs). Enfin, les marqueurs moléculaires semblent être particulièrement bien indiqués pour vérifier l'identité et la pureté variétales nécessaires à la certification des lots de semences.

Encore au stade prospectif, les bases moléculaires et statistiques de la distinction, de l'identification, de l'homogénéité et de la pureté des variétés passent par le calcul de distances génétiques [2].

La notion de variété est née en 1932, date de la création du catalogue officiel des espèces et des variétés. Il a fallu attendre 1961 pour que la protection de la propriété industrielle soit adaptée au matériel végétal. Enfin, ce n'est qu'en 1991 qu'une définition claire de la variété a vu le jour, de même que le concept de dérivation essentielle. Espérons qu'il ne faudra pas attendre 2011 pour voir entrer en vigueur l'utilisation des marqueurs moléculaires dans la caractérisation officielle des variétés. Si l'on considère qu'il existe un continuum entre l'ADN, l'ARN, les protéines/enzymes et les caractères morphophysiques, l'utilisation des marqueurs moléculaires revient à faire un zoom sur les génotypes en augmentant le pouvoir de discrimination entre individus et groupes d'individus. Il reste à déterminer le seuil de distinction pour chaque espèce, afin d'isoler les différences biologiques significatives et pertinentes des différences dues à des variations aléatoires, non

reproductibles. L'homogénéité étant un corollaire de la distinction, il conviendra de fixer un seuil en fonction de la variabilité génétique présente dans l'espèce et en fonction de son mode de reproduction. La crainte des sélectionneurs à propos d'une exigence accrue en termes d'homogénéité liée à l'introduction des marqueurs moléculaires n'est donc pas justifiée dès lors qu'une distance génétique est calculée.

Relation entre distance phénotypique et distance moléculaire

La *figure 1* montre la relation caractéristique qui lie les distances phénotypiques (Mahalanobis) entre 145 lignées de maïs aux distances génétiques (Rogers) calculées entre ces mêmes lignées. Dans cet exemple, les distances phénotypiques ont été calculées à partir de 10 variables quantitatives. Les distances génétiques ont été calculées à partir de 80 marqueurs moléculaires de type RFLP. Les concepts de distinction (fondée sur les caractères phénotypiques) et de dérivation essentielle (fondée sur les marqueurs moléculaires) sont clairement représentés, respectivement par les seuils horizontaux et verticaux. Différents cas possibles sont envisagés en fonction de la position des couples de lignées vis-à-vis de ces seuils. La dérivation essentielle d'une variété nouvelle par rapport à une variété initiale n'a de sens que si ces deux variétés sont distinctes de l'ensemble de la collection de référence mais également entre elles, donc inscriptibles. Cette relation a depuis été observée sur d'autres espèces [3, 4]. Le Geves développe actuellement un logiciel de prédiction de distances phénotypiques par les données moléculaires, dont l'objectif est de linéariser cette relation.

Plusieurs phénomènes peuvent expliquer cette relation triangulaire :

- la compensation de certains caractères élémentaires dans l'expression de caractères complexes ;
- la faible proportion du génome impliquée dans l'expression de caractères phénotypiques ;
- la multiplicité des origines génétiques engendrant des variations de déséquilibres de liaison (co-sélection de caractères non liés physiquement mais dont la combinaison fournit un avantage adaptatif ou sélectif) ;
- l'absence des conditions environnementales nécessaires à l'expression de tous les caractères.

L'utilisation des marqueurs moléculaires dans la procédure d'inscription variétale peut s'envisager sous plusieurs aspects :

- la gestion des collections de référence ;
- l'établissement de la dérivation essentielle ;
- l'appréciation de l'homogénéité variétale.

Une examen des concepts et méthodes statistiques liés à l'utilisation des marqueurs moléculaires pour l'inscription variétale est présentée par Van Ewijk et Baril [1].

Gestion des collections de référence

Les marqueurs moléculaires pourraient être utilisés pour prédire les distances phénotypiques entre variétés. Ainsi, avant la première expérimentation au champ, les marqueurs moléculaires renseigneraient sur la proximité de la variété candidate vis-à-vis des variétés de la collection de référence et des autres candidates. Il suffirait alors d'implanter côte à côte les variétés candidates et les variétés de la collection de référence jugées proches par l'équation de prédiction. Toutefois, si le pouvoir explicatif de tels modèles peut être élevé, il est fondamental de vérifier leur pouvoir prédictif. La bonne capacité de prédiction de ces modèles ne sera conservée que si les déséquilibres de liaison mis en évidence dans les ensembles de variétés ayant permis d'estimer les paramètres des modèles sont conservés au sein des nouvelles variétés candidates à l'inscription. La *figure 2* représente l'intervalle de confiance (à 95 %) autour de la distance observée en fonction de la distance prédite par ces modèles. Le seuil horizontal correspond au seuil de distinction et les deux seuils verticaux correspondent aux bornes inférieure et supérieure de la distance prédite. L'utilisation des marqueurs moléculaires pour prédire les distances phénotypiques est fondée sur le raisonnement suivant :

- pour une distance prédite inférieure à la borne inférieure ou supérieure à la borne supérieure, il n'est pas nécessaire d'implanter au champ les lignées de référence ;
- pour une distance prédite comprise entre ces deux bornes, les lignées de références concernées doivent être mises au champ à côté de la lignée candidate en question.

Les erreurs commises peuvent être de deux types : fausse similarité ou fausse distinction. L'application d'un modèle linéaire liant chaque variable phénotypique quantitative à des marqueurs moléculaires a permis d'économiser la mise au champ d'un tiers de la collection de référence, sans risque d'erreur significatif, dans une étude réalisée sur les 145 lignées de maïs présentées précédemment [5]. Pour plus de sécurité, dans les simulations réalisées par Nuel *et al.*, seules les fausses distinctions ont été autorisées (dès qu'une lignée de référence avait une distance prédite avec la lignée candidate inférieure à la borne supérieure, elle était implantée à ses côtés). Ces premiers résultats sont encourageants et ces modèles vont être appliqués à d'autres espèces. Par ailleurs, ils sont en cours d'extension à la prédiction de variables phénotypiques discrètes. Enfin, il semble évident que la prédiction des distances morphologiques par les données moléculaires sera améliorée lorsque des QTL de caractères de distinction seront localisés.

Établissement de la dérivation essentielle

La *figure 3* synthétise les trois dimensions à prendre en compte avant d'utiliser des marqueurs moléculaires dans le cadre de la distinction variétale (DHS) ou de la dérivation essentielle (EDV). Les méthodes de calcul de distances entre variétés dépendent en effet du type de caractère (sélectionné ou neutre), du type de variété (homogène ou hétérogène) et du type de marqueur (dominant ou co-dominant). Dans le cas de la dérivation essentielle, le *tableau* synthétise les distances génétiques les plus adaptées aux différentes situations [6]. Pour des variétés homogènes (83 variétés de colza) caractérisées par des marqueurs dominants (154 marqueurs AFLP cartographiés), une étude comparative a été menée pour juger de l'efficacité de la distance BLUE (*best linear unbiased estimator*) par rapport à la distance de Rogers.

Distance de Rogers :

$$\hat{D}_R^{ij} = \frac{1}{2L} \sum_{l=1}^L \sum_{a=1}^{A_l} (\hat{P}_{al}^i - \hat{P}_{al}^j)^2$$

Distance BLUE :

$$\tilde{D}_B^{ij} = (1' V^{-1} 1)^{-1} (1' V^{-1}) D_R^{ij}$$

L est le nombre de loci, A_l est le nombre d'allèles au locus l , \hat{P}_{al}^i est la fréquence de l'allèle a au locus l pour la variété i , D_R^{ij} est le vecteur de distances de Rogers entre les variétés i et j à chaque locus, 1 est le vecteur identité, V est la matrice de variance-covariance fondée sur les distances cartographiques.

La distance BLUE est fondée sur l'intégration de l'information fournie par la carte génétique (corrélations entre les marqueurs) dans la matrice de pondération utilisée pour calculer la distance génétique. La *figure 4* montre, d'une part, que l'estimateur de la distance BLUE présente une variance inférieure à celle de l'estimateur de la distance de Rogers et, d'autre part, que le gain de précision apporté par cet estimateur est d'autant plus important que la densité des marqueurs sur la carte génétique est élevée. L'intérêt de la distance BLUE par rapport à la distance de Rogers tombe en cas d'indépendance des marqueurs utilisés et/ou en cas d'absence de consanguinité entre les variétés étudiées. Si l'indépendance peut être un des critères de choix des marqueurs moléculaires, l'hypothèse d'absence de consanguinité entre deux variétés soupçonnées d'être essentiellement dérivées l'une de l'autre a peu de chance d'être vérifiée.

Appréciation de l'homogénéité variétale

La *figure 5* présente les premiers résultats obtenus sur trois variétés de colza (une variété issue d'haplodiploïdisation très homogène, une variété hétérogène présentant six individus hors-type et une variété très hétérogène, non fixée) concernant l'utilisation des marqueurs moléculaires (173 marqueurs AFLP) pour apprécier l'homogénéité variétale et la présence d'individus hors-type (HT) [6]. L'indice de diversité de Nei, calculé pour chaque variété en excluant les individus hors-type, est en accord avec l'appréciation phénotypique de l'homogénéité des variétés. L'ACP (analyse en composantes principales) réalisée sur les données moléculaires de l'ensemble des individus des trois variétés étudiées montre une correspondance parfaite entre homogénéité phénotypique et homogénéité génétique. Tous les individus hors-type (HT) au champ apparaissent hors-type (HT) sur le plan moléculaire. Cela tendrait à prouver que, sur du matériel végétal sélectionné et suffisamment fixé, la variabilité neutre est très liée à la variabilité phénotypique. Toutefois, ces résultats doivent être corroborés par des expérimentations à plus grande échelle.

CONCLUSION

À la lumière de ces quelques illustrations, il apparaît que les travaux doivent être poursuivis sur les liens entre la variabilité neutre et la variabilité des caractères agronomiques.

Pour conclure, il est clair que la prise en compte du marquage moléculaire dans les procédures d'inscription des variétés d'espèces végétales relèvent de nombreuses disciplines (génétique des populations, génétique quantitative, biologie moléculaire, agronomie, biostatistique) et qu'une intégration de ces approches s'avère aujourd'hui nécessaire.

REFERENCES

1. VAN EEUWIJK FA, BARIL CP (2001). Conceptual and statistical issues related to the use of molecular markers for distinctness and essential derivation. Molecular markers for characterizing genotypes and identifying cultivars in horticulture. International symposium under the aegis of the commission of biotechnology of ISHS, Montpellier, France, March 6-8. *Acta Horticulturae*, 546 : 35-54.
2. LOMBARD V, DUBREUIL P, DILLMANN C, BARIL CP (2001). Genetic distance estimators based on molecular data for plant registration and protection: a review. Molecular markers for characterizing genotypes and identifying cultivars in horticulture. International symposium under the aegis of the commission of biotechnology of ISHS, Montpellier, France, March 6-8. *Acta Horticulturae*, 546 : 55-64.
3. MAHALANOBIS PC (1936). On the generalized distances in statistics. *Proceedings of the national institute for science*, India, 2 : 49-55.
4. ROGERS JS (1972). Measures of similarity and genetic distance. *Studies in genetics*, VII, 7213 : 145-53.
5. LOMBARD V, BARIL CP, DUBREUIL P, BLOUET F, ZHANG D (2000). Genetic relationships and fingerprinting of rapeseed cultivars using AFLP: consequences for varietal registration. *Crop Science*, 40 : 1417-25.
6. ROLDAN-RUIZ I, VAN EEUWIJK FA, GILLILAND TJ, *et al.* (2001). A comparative study of molecular and morphological methods describing relationships between perennial ryegrass (*Lolium perenne* L.) varieties. *Theor Appl Genet* (in press).
7. NUEL G, ROBIN S, BARIL CP (2001). Predicting distances using a linear model: the case of varietal distinctness. *J Applied Stat*, 28 : 607-21.
8. LOMBARD V, TIREAU B, BLOUET F, ZHANG D, BARIL CP (2001). Usefulness of AFLP markers to estimate varietal homogeneity. Assessment of the uniformity of rapeseed inbred line varieties. *Euphytica* (in press).
9. DILLMANN CH, CHARCOSSET A, GOFFINET B, SMITH JSC, DATTEE Y (1997). Best linear unbiased estimator of the molecular genetic distance between inbred lines. *Advances in Biometrical Genetics*. In : KRAJEWSKI P, KACZMAREK Z, eds. *Proceedings of the Tenth meeting of EUCARPIA Section Biometrics in Plant Breeding*, Posnan, 14-16 May : 105-10.

10. NEI M, LI WH (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc Natl Acad Sci USA*, 76 : 3269-73.
11. JACCARD P (1900). Contribution au problème de l'immigration post-glaciaire de la flore alpine. *Bulletin de la Société Vaudoise des Sciences Naturelles*, 37 : 547-79.
12. JACCARD P (1908). Nouvelles recherches sur la distribution florale. *Bulletin de la Société Vaudoise des Sciences Naturelles*, 44 : 223-70.
13. GHERARDI M, MANGIN B, GOFFINET B, BONNET D, HUGUET T (1998). A method to measure genetic distance between allogamous populations of alfalfa (*Medicago sativa*) using RAPD markers. *Theor Appl Genet*, 96 : 406-12.
14. FOULLEY JL, HILL WG (1999). On the precision of estimation of genetic distance. *Genet Sel Evol*, 31 : 457-64.

Illustrations

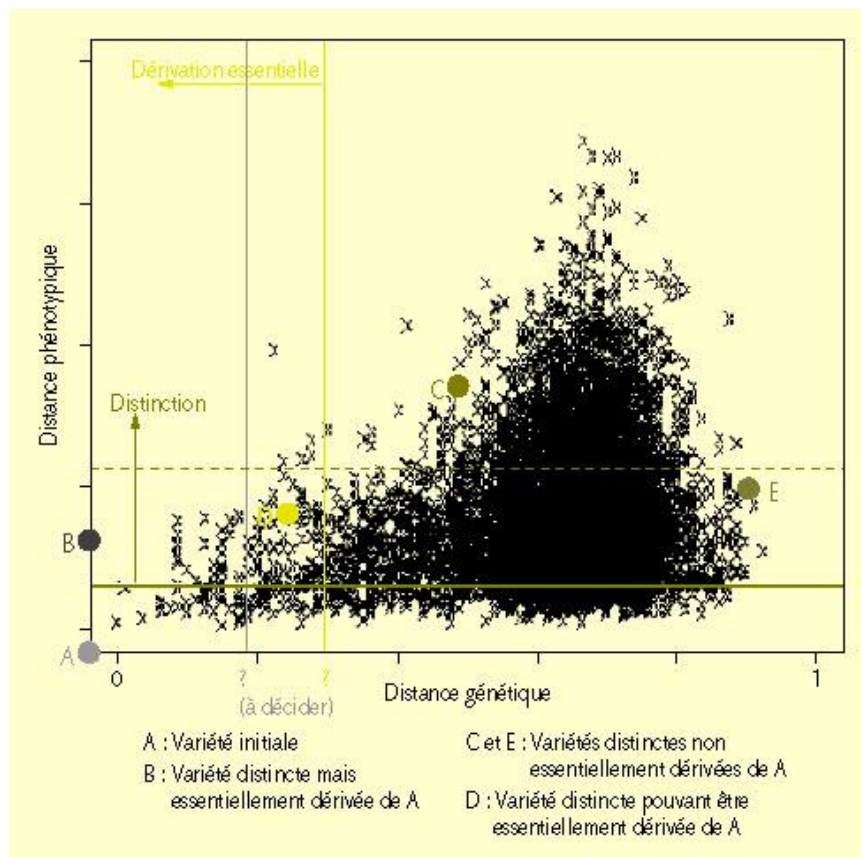


Figure 1. Distance phénotypique en fonction de la distance génétique fondée sur des données RFLP de 145 lignées de maïs : représentation graphique des concepts de distinction et de dérivation essentielle.

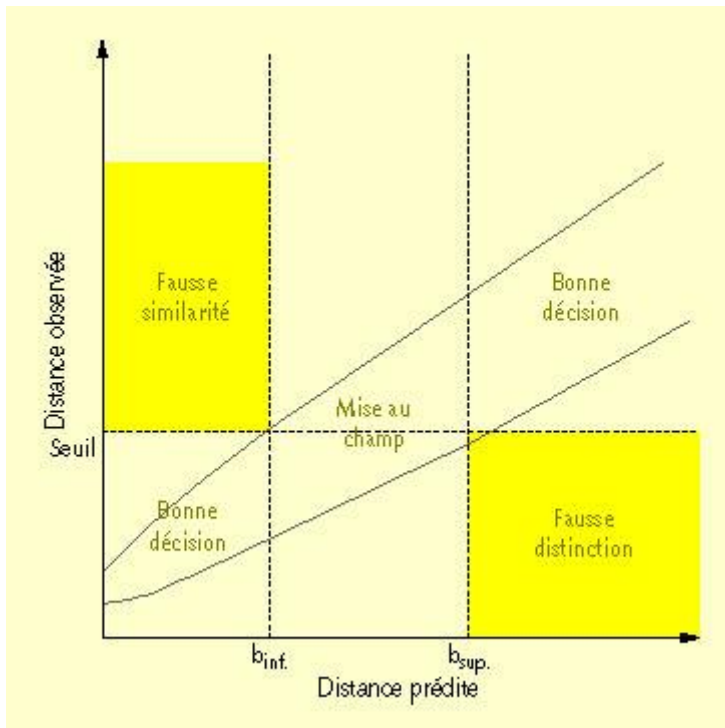


Figure 2. Marqueurs moléculaires et gestion des collections de référence : distance phénotypique observée en fonction de la distance phénotypique prédite par les marqueurs moléculaires.

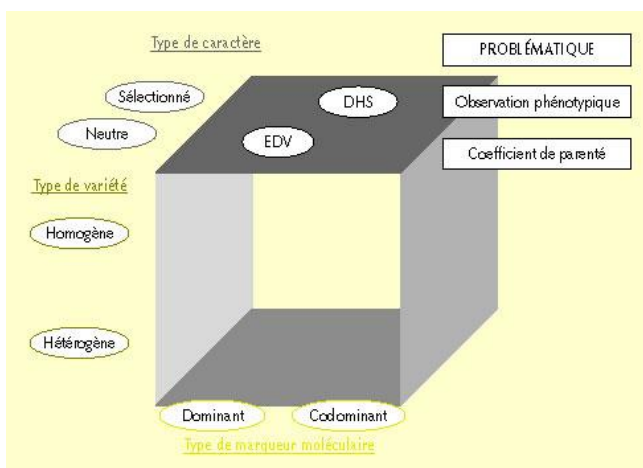


Figure 3. Trois sources de variabilité à prendre en compte pour l'inscription et la protection variétale : les types de caractère, de variété et de marqueur moléculaire.

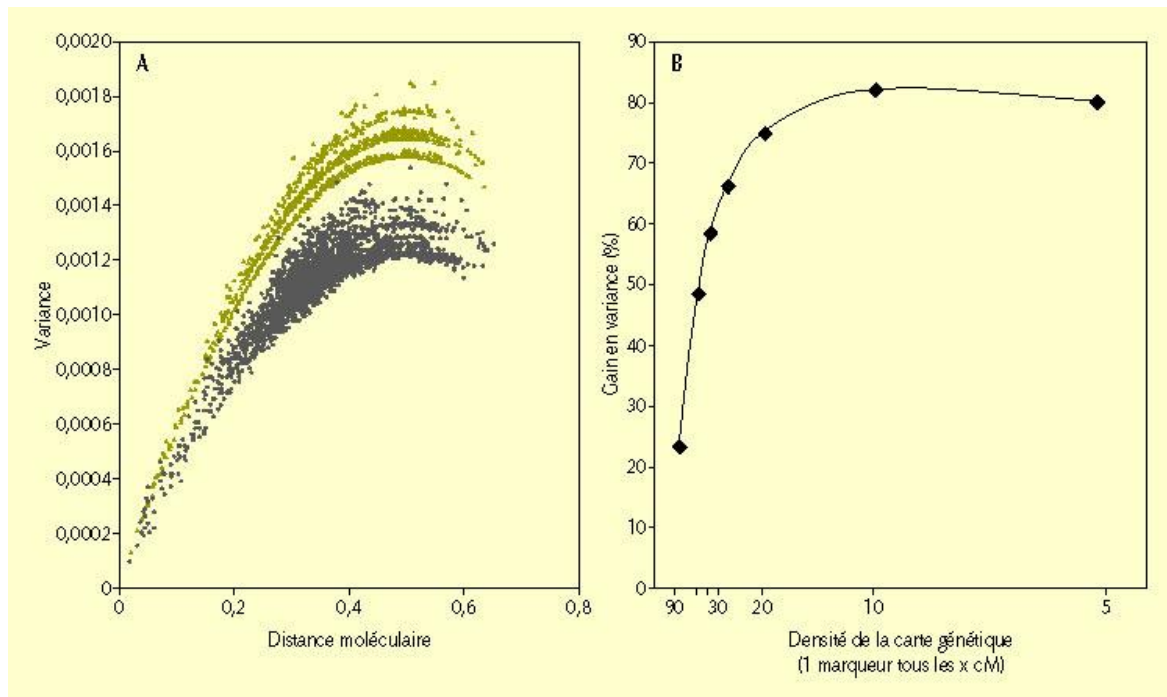


Figure 4. Comparaison entre les distances génétiques BLUE et Rogers. A. Variance de l'estimateur en fonction de la distance moléculaire. B. Gain de précision apporté par l'estimateur BLUE en fonction de la densité de la carte génétique.

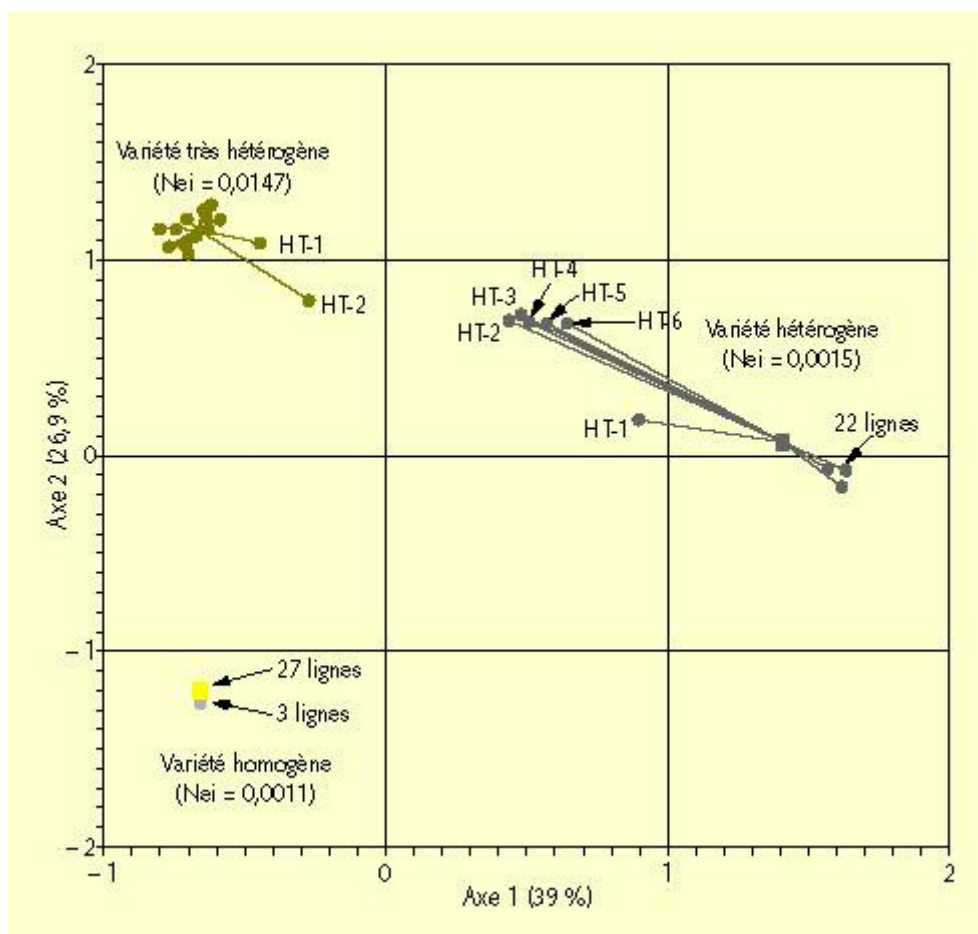


Figure 5. Marqueurs moléculaires et homogénéité variétale : ACP des représentants de trois variétés de colza fondée sur des données AFLP.

Tableau. Marqueurs moléculaires et dérivation essentielle : distances génétiques en fonction des types de variétés et de données moléculaires.

Type de variétés	Marqueurs codominants (Données locus -- allèle)	Marqueurs dominants (Données bandes)
Variétés homogènes (lignées, clones, hybrides...)	- distance de Rogers [8] - distance BLUE (cartographie) [9]	- distance de Nei & Li [10] - distance de Jaccard [11, 12]
Variétés hétérogènes (synthétiques, populations...)	- distance de Rogers débiaisée [13] - distance de Sanghvi [14]	- distance de Rogers débiaisée sous l'hypothèse de diallélisme [13]