

L

BΔP BΔ 1776/1

La biologie moléculaire en appui à l'amélioration génétique du caféier Arabica

Anthony F. ¹, Bertrand B. ², Lashermes P. ³, Charrier A. ⁴

¹ ORSTOM/CATIE/PROMECAFE, CATIE, Ap. 59, 7170 Turrialba, Costa Rica

² CIRAD-CP/PROMECAFE, IICA, Ap. 55, 2200 Coronado, San José, Costa Rica

³ ORSTOM, BP 5045, 34032 Montpellier Cedex 1, France

⁴ ENSAM/INRA, 2 Place Viala, 34060 Montpellier Cedex 1, France

L'IICA/PROMECAFE ¹, le CATIE ² et la coopération française développent actuellement un programme régional d'amélioration génétique du caféier Arabica pour enrichir la base génétique des variétés cultivées en Amérique centrale. L'historique de l'introduction du caféier sur le continent américain (Chevalier et Dragon, 1928 ; Carvalho, 1946) montre que le matériel diffusé, ou en cours de diffusion, contient seulement les gènes d'une dizaine de caféiers sauvages. A cause de cette base génétique très étroite, les cultivars présentent un comportement homogène et des

problèmes encore plus sérieux, comme la sensibilité presque généralisée aux parasites et maladies (nématodes, scolyte du fruit, anthracnose du fruit,...) (Anthony *et al.*, 1995).

D'après leur définition la plus courante, les ressources génétiques du caféier Arabica sont composées de toutes les plantes avec lesquelles il peut échanger des gènes, c'est-à-dire les individus sauvages collectés dans le centre d'origine (Ethiopie, Kenya), les variétés et mutants sélectionnés dans divers centres de recherche, et les autres espèces de caféiers (genres *Coffea* et *Psilanthus*). L'espèce la plus connue pour l'amélioration de *C. arabica* est *C. canephora*. Elle a déjà fourni les gènes de résistance à la rouille aux Catimor et Sarchimor, via l'Hybride de Timor (Osorio Garcia, 1990 ; Aguilar, 1995), et possède des gènes de résistance à diverses espèces de nématodes

(1) Instituto Interamericano de Cooperación para la Agricultura/Programa cooperativo regional para el desarrollo tecnológico y modernización de la caficultura, Ciudad Guatemala, Guatemala.

(2) Centro Agronómico Tropical para la Investigación y la Enseñanza, Turrialba, Costa Rica.



du genre *Meloidogyne* (Anzueto *et al.*, 1996).

La biologie moléculaire offre un ensemble d'outils qui permettent d'analyser l'organisation de l'information génétique au niveau de l'ADN et sa traduction en protéines. Les marqueurs moléculaires sont nombreux, presque illimités, phénotypiquement neutres et peuvent révéler plus de polymorphismes que les marqueurs morphologiques. Ils ont permis d'obtenir des résultats spectaculaires sur l'organisation du génome de plusieurs plantes alimentaires, comme le blé, le maïs, la tomate,... (Weining et Langridge, 1991 ; Edwards, 1992 ; Michelmore, 1995).

Sur le caféier, les études moléculaires commencèrent à la fin des années 70 avec la révélation des isozymes (Berthou et Trouslot, 1978). De 1991 à 1994, un projet de la Communauté européenne (contrat CII*CT91-0899) qui associa le CATIE, l'ORSTOM³ et le SCRI⁴ permit le développement de nouvelles techniques pour étudier les ressources génétiques en Amérique centrale. La plupart des résultats présentés dans cet article furent obtenus dans le cadre de ce projet.

Evaluation de la diversité génétique

Deux types de marqueurs moléculaires ont été utilisés pour étudier la diversité génétique des espèces de caféier : les isozymes révélées par électrophorèse et les RAPDs (ADN polymorphe amplifié au hasard).

Les isozymes

Les enzymes sont des protéines qui catalysent les réactions chimiques des cellules. Elles proviennent de la traduction de l'information génétique portée par l'ADN, au cours de la synthèse des protéines. Les enzymes peuvent présenter diverses formes appelées isozymes qui ont la possibilité de se séparer, d'après leur poids et leur charge électrique, dans un gel placé dans un champ électrique (technique d'électrophorèse). La limitation dans l'utilisation des isozymes provient du fait qu'elles représentent une partie très réduite du génome et que leur nombre est limité par la disponibilité des colorants enzymatiques.

La structure de la diversité génétique de l'espèce *C. canephora* fut étudiée en utilisant sept systèmes enzymatiques (Berthaud, 1986). Les 471 individus analysés provenaient de populations naturelles et d'une collection gérée par l'Idefor DCC⁵ en Côte d'Ivoire. Les données obtenues permirent de classer les individus en deux groupes géographiques (figure 1) : l'un avec toutes les populations d'Afrique de l'Ouest (« les Guinéens ») et l'autre avec les populations d'Afrique centrale (« les Congolais »). La collection Idefor-DCC se classa avec le groupe d'Afrique centrale car la plupart des génotypes avaient cette origine. Une étude complémentaire révéla la présence de deux sous-groupes chez « les Congolais » (Montagnon *et al.*, 1992). Ces résultats ont été utilisés en Côte d'Ivoire pour mettre en place un nouveau programme d'amélioration du caféier *Canephora*, basé sur la création d'hybrides entre groupes (Leroy *et al.*, 1993).

Cette étude montre le potentiel de l'électrophorèse de protéines pour détecter le polymorphisme des espèces diploïdes de caféier. Cependant, cette technique n'a pas permis d'étudier la diversité de l'espèce tétraploïde *C. arabica* ($2n = 4x = 44$) : les zymogrammes contiennent de nombreuses isozymes mais présentent peu de variations entre individus (Berthou et Trouslot, 1978).

Les RAPDs

Les RAPDs sont des copies de fragments d'ADN produits au hasard (Welsh et McClelland, 1990 ; Williams *et al.*, 1990). La production de RAPDs se réalise en trois étapes :

- la dénaturation de l'ADN, c'est-à-dire la séparation des deux brins des molécules d'ADN ;

- l'union d'une amorce, constituée généralement par un segment de dix bases, à une séquence complémentaire d'un brin d'ADN (Adénine à Thymine et Cytosine à Guanine) ;
- et la synthèse d'un fragment complémentaire d'ADN à partir de l'une des extrémités de l'amorce, grâce à l'action d'une enzyme, la *Taq* polymérase.

Ces trois étapes sont répétées une quarantaine de fois, ce qui permet une amplification exponentielle des fragments synthétisés. Les RAPDs générés sont séparés par électrophorèse, colorés dans une solution de bromure d'éthidium et observés sous une lampe ultraviolette.

L'analyse de 20 individus de l'espèce *C. arabica* aboutit à la définition de trois groupes, formés par les individus sauvages d'Ethiopie et les variétés des deux groupes botaniques Bourbon et Typica (figure 2). D'après Lashermes *et al.* (1996c), les individus sauvages d'Ethiopie se séparèrent nettement des variétés cultivées et générèrent la majeure partie du polymorphisme détecté. Cependant, le nombre réduit de marqueurs utilisés n'a pas permis la caractérisation des variétés au sein de chacun des deux groupes botaniques. Le classement des deux individus sauvages du Kenya (Marsabit 3058 et 3099) (Anthony *et al.*, 1987) dans le groupe Bourbon devra être confirmé en augmentant le nombre de marqueurs.

Cette structuration de la diversité en trois groupes génétiques avait été observée dans une étude moléculaire préliminaire, avec d'autres origines de *C. arabica* (Orozco-Castillo *et al.*, 1994). Elle est confirmée par la vigueur des hybrides entre groupes, que ce soit chez les lignées Mundo Novo qui proviennent de croisements entre Typica et Bourbon ou chez les hybrides F1 entre indi-

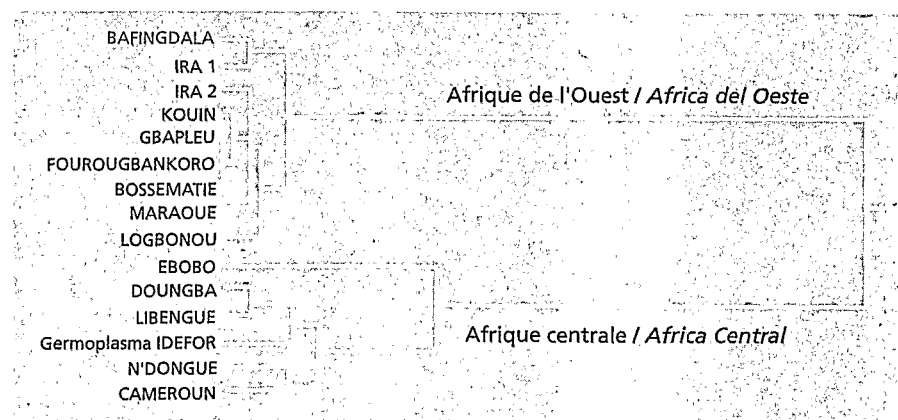


Figure 1. Structure de la diversité de l'espèce *C. canephora*, révélée par sept systèmes enzymatiques (Berthaud, 1986).

Estructura de la diversidad genética dentro de la especie *C. canephora* revelada por siete sistemas enzimáticos (Berthaud, 1986).

(3) Institut français de recherche scientifique pour le développement en coopération, Paris, France.

(4) Scottish Crop Research Institute, Dundee, Ecosse.

(5) Institut des Forêts - Département du café et du cacao, Abidjan, Côte-d'Ivoire.

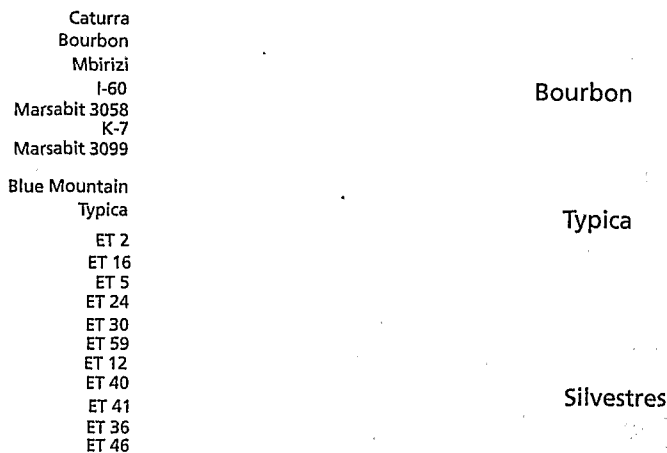


Figure 2. Classification de 20 individus de *C. arabica*, obtenue avec les données des marqueurs moléculaires RAPD (Lashermes *et al.*, 1996c).
Clasificación de 20 individuos de C. arabica, utilizando datos de marcadores moleculares RAPD (Lashermes *et al.*, 1996c).

vidus sauvages d'Ethiopie et variétés cultivées (Bouharmont, 1995 ; Bertrand *et al.*, 1997).

Les relations phylogénétiques

Les relations phylogénétiques s'étudient au niveau moléculaire en utilisant principalement les analyses de l'ADN chloroplastique (ADNcp) et de l'ADN ribosomique (ADNr) en raison du niveau d'informations que possèdent ces molécules. L'ADNcp constitue un matériel privilégié car il est abondant dans les chloroplastes, possède une très faible fréquence de changements structuraux et présente des séquences conservées, c'est-à-dire qui ont eu une évolution lente. Chez le caféier, l'ADNcp a une hérédité strictement maternelle (Lashermes *et al.*, 1996b). C'est pourquoi il peut être utilisé pour révéler la filiation maternelle de polyploïdes ou d'hybrides. L'ADNr est également intéressant parce qu'il permet d'étendre les résultats de l'ADNcp à d'autres génomes et parce qu'il produit des informations supplémentaires. Il se rencontre dans les génomes nucléaires, mitochondriaux et chloroplastiques de tous les organismes. En général, les séquences des régions non codantes sont variables alors que les séquences des régions codantes apparaissent plus conservées.

Les relations phylogénétiques entre espèces de caféiers ont été étudiées au moyen de l'analyse des sites de restriction, par les marqueurs RFLP (polymorphismes de longueur des fragments de restriction), et au moyen du séquençage de région d'ADN. La révélation des RFLPs (Botstein

et al., 1980) peut être décomposée en cinq étapes :

- la digestion de l'ADN par une enzyme de restriction ;
- la séparation des fragments d'ADN par électrophorèse ;
- le transfert des fragments séparés à une membrane de nitrocellulose ou de nylon (technique « Southern ») ;
- l'hybridation des fragments avec une sonde marquée qui contient une séquence spécifique de bases ;
- la visualisation des hybrides sonde / fragments d'ADN.

Le séquençage s'effectue au moyen des opérations suivantes (technique de Maxam-Gilbert) : isolement des deux brins d'un fragment déterminé d'ADN ; marquage d'une extrémité ; séparation de chaque échantillon en quatre lots et élimination d'une base (Adénine, Thymine, Cytosine ou Guanine) dans chacun des lots, ce qui provoque la cassure de la chaîne à l'endroit où la base a été éliminée ; comparaison du résultat du traitement de chaque brin et détermination de la séquence. Comme les séquences des bases sont complémentaires entre les deux brins, il est possible de vérifier la séquence de l'un avec celle de l'autre.

Le polymorphisme de l'ADNcp

La région entre les gènes *trnL* - *trnF* de l'ADNcp fut séquencée chez 38 individus qui représentaient 25 taxa de caféier (Cros *et al.*, 1997). Ces gènes codent les ARN de transfert de deux acides aminés (la leucine et la phénylalanine). Seulement 20 séquences distinctes furent observées. Pour obtenir une classification

relativement fiable, les données du séquençage furent complétées par celles du polymorphisme de l'ADNcp, détecté par les marqueurs RFLP (Lashermes *et al.*, 1996b). Les principales branches de la classification apparaissent étroitement liées à l'origine des espèces (figure 3). Les deux individus de l'espèce tétraploïde *C. arabica* (sauvage ET-12 et Caturra) se classèrent avec *C. eugenioides* d'Afrique orientale et avec une nouvelle espèce d'Afrique centrale, *C. sp.* « Moloundou » (Anthony, 1992). *C. arabica* pourrait donc avoir le même ascendant maternel que ces deux espèces.

Le polymorphisme de l'ADNr

La région ITS 2 (espaceur transcrit interne) de l'ADNr nucléaire qui code les sous-unités 18S et 26S de l'ARNr fut séquencée chez 37 individus qui représentaient 32 taxa de caféier (Lashermes *et al.*, 1997). Comme pour la classification basée sur le polymorphisme de l'ADNcp, les principales branches furent étroitement liées à l'origine des espèces (figure 4). Cependant, les deux individus *C. arabica* se classèrent avec des espèces d'Afrique centrale et occidentale : *C. brevipes*, *C. canephora* et *C. congensis*. Par ailleurs, les deux espèces repérées par l'étude de la filiation maternelle, *C. eugenioides* et *C. sp.* « Moloundou », se classèrent ensemble mais séparées de *C. arabica*.

En conclusion, l'étude phylogénétique a permis de préciser les espèces génétiquement les plus proches de *C. arabica* et de confirmer son origine allotétraploïde. Elle conforte l'hypothèse émise par Lashermes *et al.* (1996c) que *C. arabica* pourrait provenir d'une hybridation entre deux espèces de *Coffea*, proches de *C. eugenioides* et de *C. canephora*. Les affinités entre *C. arabica*, *C. congensis* et *C. eugenioides* avaient déjà été observées par l'analyse de l'ADN mitochondrial (Berthou *et al.*, 1983).

Le marquage de caractères

Les résultats se rapportent à la détection d'introgessions naturelles et à la construction d'une carte génétique.

La détection d'introgessions naturelles

Deux fragments d'ADN résultant d'introgessions naturelles furent détectés par les RAPDs et vérifiés par hybridation moléculaire en les utilisant comme sonde :

- un fragment de 1 500 paires de bases, présent chez *C. canephora* et l'Hybride

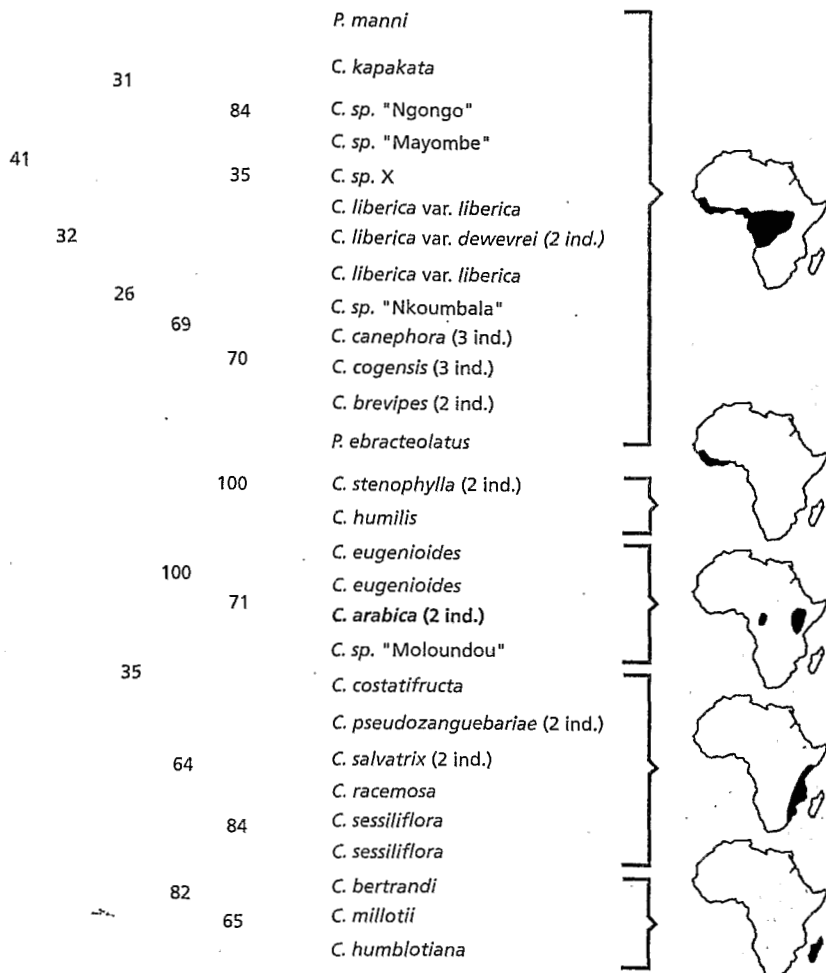


Figure 3. Classification phylogénétique de 25 taxa de caféier, sur la base du polymorphisme de l'ADNcp, et distribution géographique des principales branches de la classification (Cros et al., 1997). Clasificación filogenética de 25 taxa de cafeto, a través del polimorfismo del ADNcp, y distribución geográfica de los ramos principales de la clasificación (Cros et al., 1997).

de Timor (832/1) mais absent chez *C. arabica* (Lashermes et al., 1993) ;

- un fragment de 200 paires de bases, présent chez *C. canephora*, Rume Sudan (RS-510) et un Catimor (T.5175), mais absent chez *C. arabica* (Orozco-Castillo et al., 1994).

La construction d'une carte génétique

Une carte génétique se présente comme une carte routière d'un pays, où les chromosomes sont les routes et les loci qu'ils portent sont les villes et villages (Phillips-Mora et al., 1993). Les marqueurs moléculaires permettent de localiser les gènes qui

codent les caractères, le long des chromosomes.

La construction d'une carte génétique a commencé chez l'espèce *C. canephora*, avec une population en ségrégation d'haploïdes doublés (Paillard et al., 1996). Les avantages de ce matériel sont sa structure diploïde ($2n = 2x = 22$) et le niveau de polymorphisme élevé des individus qui permet d'obtenir la ségrégation de nombreux caractères. La carte génétique contient actuellement 47 loci RFLP et 100 loci RAPD, situés sur 15 groupes de liaison. C'est une carte de densité moyenne (10 cM en moyenne entre marqueurs) qui peut déjà être utilisée pour localiser des gènes,

comme celui de l'auto-incompatibilité sur le groupe de liaison n° 9 (Lashermes et al., 1996a).

Conséquences pour l'amélioration génétique

En raison du nombre réduit de marqueurs étudiés, les résultats qui ont été présentés doivent être interprétés comme des résultats préliminaires. Malgré cela, il est possible de dégager quelques conclusions importantes pour l'amélioration génétique du caféier Arabica :

- le classement du matériel sauvage, séparé des variétés qui proviennent du Typica et du Bourbon, confirme son intérêt pour élargir la base génétique du matériel cultivé. Cependant, le nombre réduit de marqueurs n'a pas encore permis d'estimer la distance génétique entre le matériel sauvage et les variétés cultivées ;
- l'analyse de la diversité génétique chez le matériel sauvage permettra de définir des groupes génétiques. La réalisation de croisements entre ces groupes fournira de nouveaux géniteurs pour la création variétale, qui pourront présenter une recombinaison de caractères intéressants ;
- l'introgession de gènes dans les variétés cultivées peut être tentée à partir de toutes les espèces phylogénétiquement proches de *C. arabica*, c'est-à-dire *C. brevipes*, *C. canephora*, *C. congensis*, *C. eugenioides* et *C. sp. « Moloundou »* ;
- la création de variétés porte-greffes de *C. canephora*, résistantes aux principaux nématodes, doit être basée sur la réalisation de croisements entre individus d'Afrique centrale et d'Afrique occidentale. Les nouvelles variétés présenteront une vigueur hybride (hétérosis) qui pourra aider la croissance des caféiers Arabica greffés ;
- l'adaptation de la carte génétique à *C. arabica* permettra l'analyse génétique des caractères d'intérêt agronomique et la localisation de loci de caractères quantitatifs (*Quantitative Trait Loci*). L'identification de marqueurs moléculaires liés aux caractères intéressants constituera un outil très puissant pour le programme d'amélioration génétique parce qu'il sera possible d'effectuer une sélection précoce (en pépinière) des individus qui posséderont ces caractères et de ne planter en essai que les plantes ainsi présélectionnées. En développant une telle méthodologie, on peut espérer diminuer le coût du programme de sélection et accroître son efficacité. ■

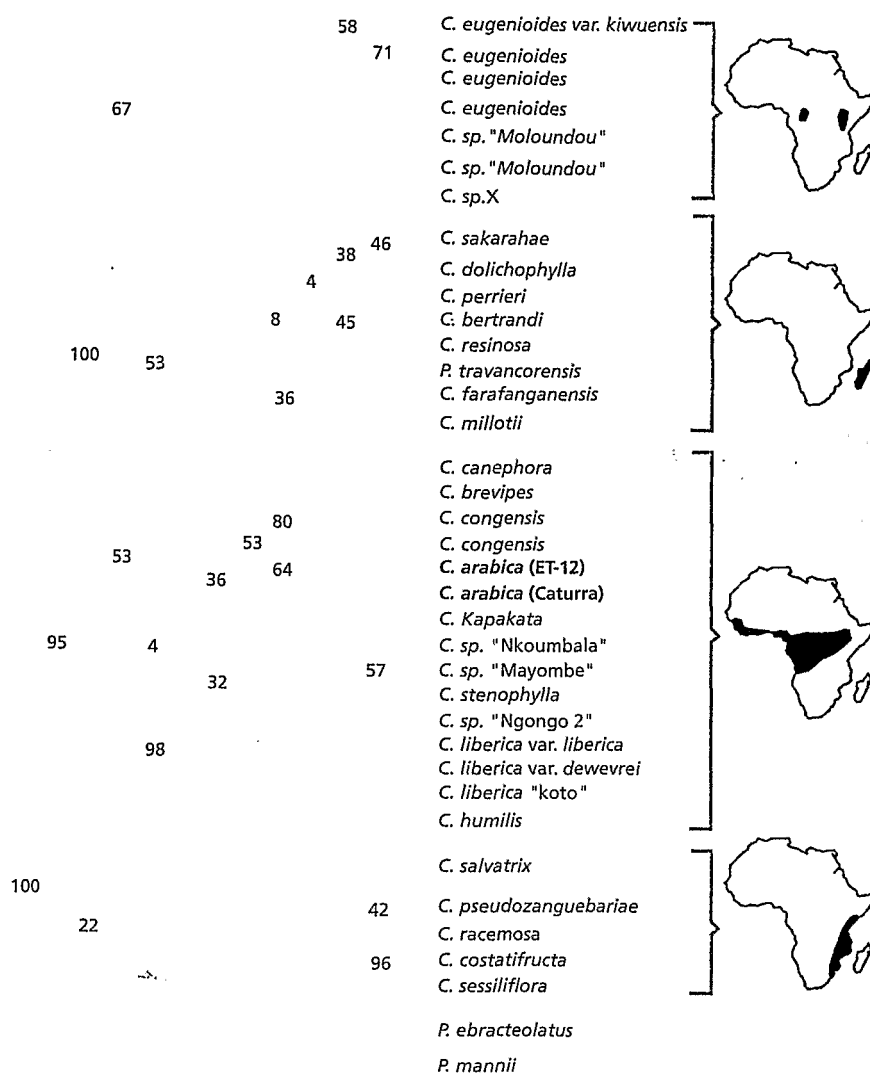


Figure 4. Classification phylogénétique de 32 taxa de caféier, sur la base du polymorphisme de l'ADNr, et distribution géographique des principales branches de la classification (Lashermes et al., 1997).

Clasificación filogenética de 32 taxa de cafeto, a través del polimorfismo del ADNr, y distribución geográfica de los ramos principales de la clasificación (Lashermes et al., 1997).

Remerciements

Les auteurs remercient M^{me} Irma Hernández de Phillips et M. Wilberth Phillips-Mora pour la relecture du manuscrit original, en espagnol, et le Dr Albertús Eskes pour ses précieux commentaires.

Bibliographie / Bibliografía

AGUILAR G., 1995. Variedad Costa Rica. Convenio del Instituto del Café de Costa Rica y del Ministerio de Agricultura y Ganadería, San José, Costa Rica, ICAFE-MAG, 30 p. (document interne).

ANTHONY F., 1992. Les ressources génétiques des caféiers : collecte, gestion d'un conservatoire et évaluation de la diversité génétique. Paris, France, ORSTOM, coll. Travaux & Documents Microédités 81, 320 p.

ANTHONY F., BERTHAUD J., GUILLAUMET J. L., LOURD M., 1987. Collecting wild *Coffea* species in Kenya and Tanzania. *Ressour. Gén. Vég. Bull.* 69 : 23-29.

ANTHONY F., BERTRAND B., DUFOUR M., ESCALANT J.V., 1995. Evaluación y caracterización de los recursos genéticos de café conservados en el germoplasma del CATIE. In : XVI simposio sobre la caficultura latinoamericana, Managua, Nicaragua, 25-29 oct. 1993. Tegucigalpa, Honduras, IICA/PROMECAFE, 6 p.

ANZUETO F., BERTRAND B., PEÑA M., MARBAN-MENDOZA N., VILLAIN L., 1996. Desarrollo de una variedad porta-injerto resistente a los principales nematodos de América Central. In : XVII simposio sobre la caficultura latinoamericana, San Salvador, Salvador, 23-27 octubre 1995. Tegucigalpa, Honduras, IICA/PROMECAFE, 7 p.

BERTHAUD J., 1986. Les ressources génétiques pour l'amélioration des caféiers africains diploïdes. Evaluation de la richesse génétique des populations sylvestres et de ses mécanismes organisateurs. Conséquences pour l'application. Paris, France, ORSTOM, coll. Travaux et Documents de l'ORSTOM 188, 379 p.

BERTHOU F., TROUSLOT P., 1978. L'analyse du polymorphisme enzymatique dans le genre *Coffea* : adaptation d'une méthode d'électrophorèse en série ; premiers résultats. In : VIII colloque scientifique international sur le café, Abidjan, Côte d'Ivoire, 28 nov.-3 déc. 1977. Paris, France, ASIC, p. 373-383.

- BERTHOUE F., MATHIEU C., VEDEL F., 1983. Chloroplast and mitochondrial DNA variation as indicator of phylogenetic relationships in the genus *Coffea* L. *Theor. Appl. Genet.* 65 : 77-84.
- BERTRAND B., SANTACREO R., AGUILAR G., ANZUETO F., ESKES A.B., 1997. El mejoramiento genético en América Central: ¿ la solución a todos los problemas ? *In* : Desafíos de la caficultura centroamericana, San José, Costa Rica, IICA (sous presse).
- BOTSTEIN D., WHITE R. L., SKOLNICK M., DAVIS R.W., 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.* 32 : 314-331.
- BOUHARMONT P., 1995. La sélection du caféier Arabica au Cameroun, 1964-1991. Montpellier, France, CIRAD, coll. Documents de travail du CIRAD-CP 1-95, 131 p.
- CARVALHO A., 1946. Distribuição geográfica e classificação botânica do gênero *Coffea* com referência especial à espécie Arabica. *Boletim da Superintendência dos Serviços do café* 21 (230) : 174-180.
- CARVALHO A., 1988. Principles and practice of coffee plant breeding for productivity and quality factors: *Coffea arabica*. *In* : Coffee, 4 : Agronomy, R. J. Clarke et R. Macrae éd., Londres, Royaume-Uni, Elsevier Applied Science, p. 129-165.
- CHEVALIER A., DAGRON M., 1928. Recherches historiques sur les débuts de la culture du caféier en Amérique. *Communications et Procès Verbaux de l'Académie des Sciences Coloniales de Paris*, 38 p.
- CROS J., COMBES M. C., TROUSLOT P., ANTHONY F., HAMON S., CHARRIER A., LASHERMES P., 1997. Phylogenetic analysis of chloroplast DNA variation in *Coffea* L. *Mol. Phylogenet. Evol.* (sous presse).
- EDWARDS M., 1992. Use of molecular markers in the evaluation and introgression of genetic diversity for quantitative traits. *Field Crops Res.* 29 : 241-260.
- LASHERMES P., CROS J., MARMERY P., CHARRIER A., 1993. Use of random amplified DNA markers to analyse genetic variability and relationships of *Coffea* species. *Genet. Resour. Crop Evol.* 40 : 91-99.
- LASHERMES P., COUTURON E., MOREAU N., PAILLARD M., LOUARN J., 1996a. Inheritance and genetic mapping of self-incompatibility in *Coffea canephora* Pierre. *Theor. Appl. Genet.* 93 : 458-462.
- LASHERMES P., CROS J., COMBES M. C., TROUSLOT P., ANTHONY F., HAMON S., CHARRIER A., 1996b. Inheritance and restriction fragment length polymorphism of chloroplast DNA in the genus *Coffea* L. *Theor. Appl. Genet.* 93 : 626-632.
- LASHERMES P., TROUSLOT P., ANTHONY F., COMBES M. C., CHARRIER A., 1996c. Genetic diversity for RAPD markers between cultivated and wild accessions of *Coffea arabica*. *Euphytica* 87 : 59-64.
- LASHERMES P., COMBES M. C., TROUSLOT P., CHARRIER A., 1997. Phylogenetic relationships of coffee-tree species, *Coffea* L. as inferred from ITS sequences of nuclear ribosomal DNA. *Theor. Appl. Genet.* 94 : 947-955.
- LEROY T., MONTAGNON C., CHARRIER A., ESKES A. B., 1993. Reciprocal recurrent selection applied to *Coffea canephora* Pierre I. Characterization and evaluation of breeding populations and value of intergroup hybrids. *Euphytica* 67 : 113-125.
- MICHELMORE R., 1995. Molecular approaches to manipulation of disease resistance genes. *Ann. Rev. Phytopathol.* 15 : 393-427.
- MONTAGNON C., LEROY T., YAPO A., 1992. Etude complémentaire de la diversité génotypique et phénotypique des caféiers de l'espèce *Coffea canephora* en collection en Côte d'Ivoire. *In* : XIVe colloque scientifique international sur le café, San Francisco, Etats-Unis, 14-19 juillet 1991. Paris, France, ASIC, p. 444-450.
- OROZCO-CASTILLO C., CHALMERS K. J., WAUGH R., POWELL W., 1994. Detection of genetic diversity and selective gene introgression in coffee using RAPD markers. *Theor. Appl. Genet.* 87 : 934-940.
- OSORIO GARCIA F. O., 1990. Evaluación de la estabilidad del rendimiento de introducciones en un ensayo regional de café. IX reunión regional de mejoramiento de café de PROMECAFE, Managua, Guatemala, IICA/PROMECAFE, 14 p. (document interne).
- PAILLARD M., LASHERMES P., PÉTIARD V., 1996. Construction of a molecular linkage map in coffee. *Theor. Appl. Genet.* 93 : 41-47.
- PHILLIPS-MORA W., RODRIGUEZ H., FRITZ P. J., 1993. Marcadores de ADN: teoría, aplicaciones y protocolos de trabajo, con ejemplos de investigaciones en cacao, *Theobroma cacao*. San José, Costa Rica, CATIE, Serie Técnica, Informe Técnico 252, 183 p.
- WEINING S., LANGRIDGE P., 1991. Identification and mapping of polymorphisms in cereals based on the polymerase chain reaction. *Theor. Appl. Genet.* 82 : 209-216.
- WELSH J., MCCLELLAND M., 1990. Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Res.* 18 : 7213-7218.
- WILLIAMS J.G.K., KUBELIK A.R., LIVAK K.J., RAFALSKI J.A., TINGEY S.V., 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res.* 18 : 6531-6535.

La biología molecular en apoyo al mejoramiento genético del café Arabica

Anthony F.¹, Bertrand B.², Lashermes P.³, Charrier A.⁴

¹ ORSTOM/CATIE/PROMECAFE, CATIE, Ap. 59, 7170 Turrialba, Costa Rica

² CIRAD-CP/PROMECAFE, IICA, Ap. 55, 2200 Coronado, San José, Costa Rica

³ ORSTOM, BP 5045, 34032 Montpellier Cedex 1, Francia

⁴ ENSAM/INRA, 2 Place Viala, 34060 Montpellier Cedex 1, Francia

La biología molecular ofrece un conjunto de herramientas que permiten analizar la organización de la información genética al nivel del ADN. Este artículo estudia la aplicación de estas herramientas para la mejora genética del café.

El IICA/PROMECAFE¹, el CATIE² y la Cooperación Francesa están desarrollando ahora un programa regional de mejoramiento genético del café Arabica para enriquecer la base genética de las variedades cultivadas en América Central. La reseña histórica de la introducción del café en el continente americano (Chevalier y Dragon, 1928; Carvalho, 1946), muestra que el material divulgado, o pendiente de serlo, contiene únicamente los genes de unos diez cafetos silvestres. A causa de esta base genética muy limitada, los cultivares presentan un comportamiento homogéneo y problemas aún más serios, como son, una susceptibilidad casi generalizada a plagas y enfermedades (nematodos, broca, antracnosis del fruto, ...) (Anthony *et al.*, 1995).

Según su definición más común, los recursos genéticos del café Arabica están compuestos por todas las plantas con las cuales puede intercambiarse genes, ya sea, individuos silvestres recolectados en el centro de origen (Etiopía, Kenia), las variedades y los mutantes seleccionados en varios centros de investigación, y las demás especies de cafetos (género *Coffea* y *Psilanthus*). La especie más conocida para el mejoramiento de *C. arabica* es *C. canephora* que proporcionó ya los genes de resistencia a la roya hacia los Catimores y Sarchimores, a través del Híbrido de Timor (Osorio Garcia, 1990; Aguilar, 1995) y posee genes de resistencia a varias especies de nematodos del género *Meloidogyne* (Anzueto *et al.*, 1996).

La biología molecular brinda un conjunto de herramientas que permiten analizar la organización de la información genética a nivel del ADN y su traducción en proteínas. Los marcadores moleculares son numerosos, casi

ilimitados, fenotípicamente neutros, y pueden revelar más polimorfismos que los marcadores morfológicos. Permitieron obtener resultados espectaculares sobre la organización del genoma de varias plantas alimenticias, como trigo, maíz, tomate, ... (Weining y Langridge, 1991; Edwards, 1992; Michelmores, 1995).

Los estudios moleculares sobre el café comenzaron a fines de los años 70 con el revelado de las isoenzimas (Berthou y Trouslot, 1978). De 1991 a 1994, un proyecto de la Comunidad Europea (contrato GI1*CT91-0899) que asoció al CATIE, al ORSTOM³ y al SCRI⁴, permitió desarrollar nuevas técnicas para estudiar los recursos genéticos en América Central. La mayoría de los resultados en este artículo, fueron obtenidos en el marco de este proyecto.

La evaluación de la diversidad genética

Para estudiar la diversidad genética de las especies de café, se utilizaron dos tipos de marcadores moleculares: las enzimas reveladas por electroforesis y los RAPDs (ADN polimórfico amplificado al azar).

Las isoenzimas

Las enzimas son proteínas que catalizan las reacciones químicas de las células. Resultan de la traducción de la información genética que se encuentra en el ADN, durante la síntesis de las proteínas. Las enzimas pueden presentar varias formas llamadas isoenzimas, que pueden separarse acorde a su peso y su carga eléctrica, en un gel colocado en un campo eléctrico (técnica de electroforesis). La limitante en el uso de las isoenzimas resulta del hecho de que representan una parte muy reducida del genoma

y que su número se halla limitado por la disponibilidad de los colorantes enzimáticos.

La estructura de la diversidad genética de la especie *C. canephora* fue estudiada utilizando siete sistemas enzimáticos (Berthoud, 1986). Los 471 individuos analizados provenían de poblaciones espontáneas y de una colección manejada por el IDEFOR-DCC⁵ en Côte d'Ivoire. Los datos obtenidos permitieron clasificar los individuos en dos grupos geográficos (figure 1): uno que incluía todas las poblaciones de África del Oeste ("Los Guineanos") y el otro con las poblaciones de África Central ("Los Congoleños"). La colección del IDEFOR-DCC se clasificó en el grupo de África Central, puesto que éste era el origen de la mayor parte de los genotipos. Un estudio complementario reveló la presencia de dos subgrupos en los "Congoleños" (Montagnon *et al.*, 1992). Estos resultados han sido utilizados en Côte d'Ivoire, para desarrollar un nuevo programa de mejoramiento genético del café *Canephora*, basado en la creación de híbridos entre los grupos (Leroy *et al.*, 1993).

Este estudio muestra el potencial que tiene la electroforesis de proteínas para detectar el polimorfismo de las especies diploides de café. Con todo, esta técnica no permitió estudiar la diversidad de la especie tetraploide *C. arabica* ($2n = 4x = 44$): los zimogramas contienen muchas isoenzimas, pero presentan pocas variaciones entre los individuos (Berthou y Trouslot, 1978).

Los RAPDs

Los RAPDs son copias de fragmentos de ADN producidos al azar (Welsh y McClelland, 1990; Williams *et al.*, 1990). La producción de RAPDs se realiza en tres pasos:

- la desnaturalización del ADN, o sea la separación de las dos hebras de las moléculas de ADN;

(1) Instituto Interamericano de Cooperación para la Agricultura/Programa cooperativo regional para el desarrollo tecnológico y modernización de la caficultura, Ciudad Guatemala, Guatemala.

(2) Centro Agronómico Tropical para la Investigación y la Enseñanza, Turrialba, Costa Rica.

(3) Institut français de recherche scientifique pour le développement en coopération, Paris, Francia.

(4) Scottish Crop Research Institute, Dundee, Escocia.

(5) Institut des Forêts - Département du café et du cacao, Abidjan, Côte d'Ivoire.

- la unión de un cebo, constituido normalmente por un segmento de diez bases, y de una secuencia complementaria en la hebra de ADN (Adenina a Timina y Citosina a Guanina);
- y la síntesis de un fragmento complementario de ADN a partir de una de las extremidades del cebo, mediante la acción de la enzima *Taq* polimerasa.

Se repiten unas cuarenta veces estos tres pasos, lo que permite una amplificación exponencial de los fragmentos sintetizados. Los RAPDs producidos se separan por electroforesis, se colorean en una solución de bromuro de etidio y se observan bajo una lámpara de rayos ultravioleta.

El análisis de 20 individuos de la especie *C. arabica* condujo a la definición de tres grupos, constituidos por los individuos silvestres de Etiopía y las variedades de los dos grupos botánicos Borbón y Típica (figure 2). Según Lashermes *et al.* (1996c), los individuos silvestres de Etiopía se separaron claramente de las variedades cultivadas y generaron la mayor parte del polimorfismo detectado. Sin embargo, el número reducido de marcadores utilizados no permitió caracterizar las variedades dentro de cada uno de los dos grupos botánicos. Se tendrá que confirmar la clasificación de los dos individuos silvestres de Kenia (Marsabit 3058 y 3099) (Anthony *et al.*, 1987) en el grupo Borbón ampliando el número de marcadores.

Esta estructuración de la diversidad en tres grupos genéticos había sido observada en un estudio molecular preliminar, con otras procedencias de *C. arabica* (Orozco-Castillo *et al.*, 1994). Se corrobora con el vigor de los híbridos entre los grupos, que pertenezcan a las descendencias Mundo Novo, procediendo de cruzamientos entre Típica y Borbón, o en los híbridos F1 entre individuos silvestres de Etiopía y variedades cultivadas (Bouharmont, 1995; Bertrand *et al.*, 1997).

Las relaciones filogenéticas

Las relaciones filogenéticas se estudian a nivel molecular, usando principalmente los análisis de ADN cloroplástico (ADNcp) y de ADN ribosómico (ADNr), debido al nivel de informaciones que tienen estas moléculas. El ADNcp constituye un material privilegiado, dado que es abundante en los cloroplastos, posee una frecuencia muy baja de cambios estructurales y presenta secuencias conservadas, a saber, que han tenido una evolución lenta. En el café, el ADNcp tiene un carácter de herencia estrictamente materna (Lashermes *et al.*, 1996b). Por eso, se puede utilizar para revelar la filiación materna de poliploides o de híbridos. El ADNr es también interesante, porque permite extender los resultados del ADNcp a otros genomas y porque

proporciona informaciones suplementarias. Se encuentra en los genomas nucleares, mitocondriales o cloroplásticos de todos los organismos. Por lo general, las secuencias de las regiones no codificantes son variables mientras que las secuencias de las regiones codificantes aparecen más conservadas.

Las relaciones filogenéticas entre las especies de cafetos se estudiaron mediante el análisis de los sitios de restricción, por los marcadores RFLP (Polimorfismos de Longitud de Fragmentos de Restricción), y por medio de la secuenciación de regiones del ADN. El revelado de los RFLPs (Botstein *et al.*, 1980) puede dividirse en cinco etapas:

- la digestión del ADN con una enzima de restricción;
- la separación de los fragmentos de ADN por electroforesis;
- la transferencia de los fragmentos separados a una membrana de nitrocelulosa o de nilón (técnica "Southern");
- la hibridación de los fragmentos con una sonda marcada que tiene una secuencia específica de bases, y
- la visualización de los híbridos sonda / fragmentos de ADN.

La secuenciación se realiza a través de los siguientes pasos (técnica de Maxam-Gilbert): aislamiento de las dos hebras de un fragmento determinado de ADN, marcación de una extremidad, separación de cada muestra en cuatro partes y eliminación de una base (Adenina, Timina, Citosina o Guanina) en cada una de ellas, lo que provoca una quebradura de la cadena en el lugar donde se ha eliminado la base, comparación del resultado del tratamiento de cada hebra y determinación de la secuencia. Dado que las secuencias básicas son complementarias entre las dos hebras, resulta posible verificar la secuencia de una con la de la otra.

El polimorfismo del ADNcp

Se realizó la secuenciación de la región entre los genes *trnL* - *trnF* del ADNcp, en 38 individuos que representaban 25 taxa de café (Cros *et al.*, 1997). Estos genes codifican a los ARNs de transferencia de dos aminoácidos (la leucina y la fenilalanina). Se pudo distinguir solamente 20 secuencias distintas. Para obtener una clasificación bastante fiable, se completaron los datos de secuenciación por los del polimorfismo del ADNcp, detectado por los marcadores RFLP (Lashermes *et al.*, 1996b). Los principales ramos de la clasificación se revelaron estrechamente relacionados con el origen de las especies (figure 3). Se clasificaron los dos individuos de la especie tetraploide *C. arabica* (silvestre ET-12 y Caturra) con *C. eugenioides* de África del Este y con una nueva especie de África Central,

C. sp. "Moloundou" (Anthony, 1992). Por lo tanto, *C. arabica* podría tener el mismo ascendente materno que estas dos especies.

Le polimorfismo del ADNr

La región ITS 2 (espaciador transcrito interno) del ADNr nuclear que codifica a las subunidades 18S y 26S del ARNr fue secuenciada en 37 individuos, que representaban 32 taxa de café (Lashermes *et al.*, 1997). Igual que para la clasificación basada en el polimorfismo del ADNcp, los principales ramos se relacionaron estrechamente con el origen de las especies (figure 4). Sin embargo, los dos individuos *C. arabica* se clasificaron con especies de África Central y occidental: *C. brevipes*, *C. canephora* y *C. congensis*. Por otro lado, las dos especies identificadas por el estudio de la filiación materna, *C. eugenioides* y *C. sp.* "Moloundou", se clasificaron juntas, pero separadas de *C. arabica*.

En conclusión, el estudio filogenético permitió precisar las especies genéticamente cercanas de *C. arabica* y confirmar su origen alotetraploide. Corroboró la hipótesis emitida por Lashermes *et al.* (1996c) que *C. arabica* podría provenir de una hibridación entre dos especies de *Coffea*, cercanas de *C. eugenioides* y de *C. canephora*. Las afinidades entre *C. arabica*, *C. congensis* y *C. eugenioides* ya habían sido observadas por el análisis del ADN mitocondrial (Berthou *et al.*, 1983).

El marcado de caracteres

Los resultados se refieren a la detección de introgresiones naturales y a la construcción de un mapa genético.

La detección de introgresiones naturales

Dos fragmentos de ADN resultantes de introgresiones naturales fueron detectados por los RAPDs y verificados por hibridación molecular, utilizándolos como sonda:

- un fragmento de 1500 pares de bases, presente en *C. canephora* y en el Híbrido de Timor (832/1), pero ausente en *C. arabica* (Lashermes *et al.*, 1993),
- un fragmento de 200 pares de bases, presente en *C. canephora*, Rume Sudan (RS-510) y un Catimor (T.5175), pero ausente en *C. arabica* (Orozco-Castillo *et al.*, 1994).

La construcción de un mapa genético

Un mapa genético se presenta como un mapa de carreteras de un país, donde los cromosomas son las carreteras y los loci contenidos en ellos, ciudades y pueblos (Phillips-Mora *et al.*, 1993). Los marcadores moleculares permiten situar los genes que codifican a los caracteres, a lo largo de los cromosomas.

Se inició la construcción de un mapa genético en la especie *C. canephora*, con una población en segregación de haploides duplicados (Paillard *et al.*, 1996). Las ventajas de este material son su estructura diploide ($2n = 2x = 22$) y el alto nivel de polimorfismo de los individuos lo que permite obtener la segregación de numerosos caracteres. El mapa genético contiene actualmente 47 loci RFLP y 100 loci RAPD, localizados en 15 grupos de enlace. Es un mapa de mediana densidad (10 cM como promedio entre marcadores) que ya puede utilizarse para ubicar genes, como el de la autoincompatibilidad en el grupo de enlace n° 9 (Lashermes *et al.*, 1996a).

Consecuencias para el mejoramiento genético

Debido al número reducido de marcadores estudiados, los resultados que se han presentado deben interpretarse como resultados preliminares. A pesar de esto, se pueden sacar algunas conclusiones importantes para el mejoramiento genético del café Arabica:

- la clasificación del material silvestre, separando de las variedades que provienen del Típica y del Borbón, confirma su interés para ampliar la base genética del material cultivado. Sin

embargo, el número reducido de marcadores no ha permitido estimar todavía la distancia genética entre el material silvestre y las variedades cultivadas;

- el análisis de la diversidad genética dentro del material silvestre permitirá definir grupos genéticos. La realización de cruzamientos entre estos grupos proporcionará nuevos genitores para la creación varietal, que podrán presentar una recombinación de los caracteres interesantes;
- la introgresión de genes en las variedades cultivadas puede intentarse a partir de todas las especies filogenéticamente cercanas de *C. arabica*, o sea *C. brevipes*, *C. canephora*, *C. congensis*, *C. eugenioides* y *C. sp.* "Moloundou";
- la creación de variedades portainjertos de *C. canephora*, resistentes a los principales nematodos, debe fundamentarse en la realización de cruzamientos entre los individuos de África Central y de África del Oeste. Las nuevas variedades presentarán un fuerte vigor híbrido (heterosis) que podrá ayudar al crecimiento de los cafetos Arabica injertados;
- la adaptación del mapa genético a *C. arabica* permitirá el análisis genético de los carac-

teres de interés agronómico y la localización de loci de caracteres cuantitativos (Quantitative Trait Loci). La identificación de marcadores moleculares ligados a los caracteres interesantes constituirá una herramienta muy valiosa para el programa de mejoramiento genético porque permitirá hacer una selección precoz (en vivero) de los individuos que presentarán estos caracteres, y no sembrar en ensayo más que las plantas así preseleccionadas. Al desarrollar semejante metodología, se puede esperar una disminución del costo del programa de selección y un incremento de su eficiencia. □

Agradecimientos

Los autores agradecen a los Msc. Irma Hernández de Phillips y Wilberth Phillips-Mora por la revisión del manuscrito original en español, y al Dr. Albertus Eskes por sus valiosos comentarios.

Résumé

En Amérique centrale, l'IICA / PROMECAFE, le CATIE et la coopération française développent un programme régional d'amélioration génétique du caféier dont le but est d'enrichir la base génétique des variétés cultivées. Ce programme repose sur l'utilisation des ressources génétiques disponibles afin d'introduire de nouvelles caractéristiques dans le matériel cultivé. Cependant, une utilisation rationnelle des ressources génétiques requiert des informations sur l'organisation de l'information génétique au niveau moléculaire.

Cet article présente les résultats des recherches entreprises par diverses équipes qui ont utilisées les outils de la biologie moléculaire pour étudier les ressources génétiques caféières. Les résultats concernent la structure de la diversité génétique des espèces, les relations phylogénétiques entre espèces et le marquage de caractères. Les techniques utilisées sont discutées avec une vision critique dans le but de pouvoir appuyer le programme régional d'amélioration à court et moyen termes.

Resumen

En América Central, el IICA/PROMECAFE, el CATIE y la Cooperación francesa están desarrollando un programa regional de mejoramiento genético del café, cuya meta es enriquecer la base genética de las variedades cultivadas. Este programa se fundamenta en el uso de los recursos genéticos disponibles, a fin de introducir nuevas características en el material cultivado. No obstante, el uso racional de los recursos genéticos, requiere de conocimientos con respecto a la organización de la información genética a nivel molecular.

Este artículo presenta los resultados de las investigaciones emprendidas por varios equipos que han usado las herramientas de la biología molecular para el estudio de los recursos genéticos del café. Los resultados se refieren a la estructura de la diversidad genética de las especies, a las relaciones filogenéticas entre especies y al marcado de caracteres. Se discuten las técnicas utilizadas bajo una visión crítica, con miras a poder apoyar al programa regional de mejoramiento genético a corto y mediano plazo.

Abstract

IICA/PROMECAFE, CATIE and French cooperation are conducting a regional coffee breeding programme with a view to broadening the genetic basis of varieties cultivated in Central America. The programme is based on the use of genetic resources to introduce new characteristics into cultivated material. However, rational use of this germplasm requires knowledge of how genetic information is organized at molecular level.

This article presents coffee germplasm studies by several teams, using molecular biology tools. The results concern the structure of the genetic diversity within species, the phylogenetic relations between species and character marking. The merits of the technologies used in supporting the regional breeding programme in the short and medium term are discussed.