

Towards Identification and Characterization of Candidate Genes Involved in Coffee Cup Quality

**C. CAMPA, H. CHRESTIN, A. de KOCHKO, C. BERTRAND, T. LEROY,
M. NOIROT**

UR 117, UMR 1097 IRD, BP 5045 34032 Montpellier Cedex 1, France
CIRAD-CP 1097 IRD, BP 5045 34032 Montpellier Cedex 1, France

SUMMARY

A project on the genetics of biochemical compounds involved in the cup quality of coffee was initiated by IRD (ex ORSTOM) in 1994. The major compounds studied were: sucrose, chlorogenic acids (CGA), caffeine and trigonelline. The analysis was done on the offspring of an interspecific cross between *C. pseudozanguebariae* (PSE) and *C. liberica* var. *dewevrei* (DEW). PSE is a wild species with a short fruit maturation period (2.5 months) and differs from DEW for some biochemical compounds related to the cup quality: no caffeine and low content in CGA was evaluated in PSE green beans, whereas high contents of sugar and trigonelline was observed. The role of environment and genetic effects has been shown by traditional quantitative genetics. An AFLP genetics map was obtained using this cross and QTL were located for caffeine, CGA and trigonelline accumulation.

The project currently focuses on CGA and caffeine, both responsible for bitterness. A new cross between PSE and *C. canephora* (CAN), a highly caffeine and CGA producing species, is now being mapped. A biochemical evaluation is also underway. Two cDNA libraries have been obtained from CAN leaves and beans harvested at different maturation stages. Genes involved in the control of caffeine and CGA content are being investigated. The first approach used consists in looking for equivalent heterologous sequences to design specific primers deduced from conserved domains of such genes. The corresponding fragments will be amplified and used as probes to find a possible correspondence with the QTL. If such an equivalence was found, the corresponding genes will be isolated and characterized. An alternate approach is also planned by using differential screening of cDNA libraries established from tissues of different species at different developmental stages.

RÉSUMÉ

Un projet sur la génétique des composés biochimiques impliqués dans la qualité à la tasse du café a été initié à l'IRD (ex ORSTOM) en 1994. Les principaux composés étudiés sont: le saccharose, les acides chlorogéniques (ACG), la caféine et la trigonelline. L'analyse a été faite sur la descendance d'un croisement interspécifique entre *C. pseudozanguebariae* (PSE) et *C. liberica* var. *dewevrei* (DEW). PSE est une espèce sauvage possédant une courte période de maturation des graines (2 mois et demi). Elle diffère de DEW pour la composition biochimique de ces graines en ce qui concerne les composés liés à la qualité du café-boisson. Si l'on ne trouve pas de caféine et si la teneur en ACG y est faible, en revanche, les teneurs en sucres et en trigonelline sont élevées. L'impact relatif du milieu et des gènes a été montré par génétique quantitative. Une carte génétique AFLP a été obtenue en utilisant le croisement (PSExDEW)xDEW et des QTLs ont été localisés pour les teneurs en caféine, en ACG et en trigonelline.

Actuellement, le projet s'intéresse particulièrement aux ACG et à la caféine, tous deux impliqués dans l'amertume du café-boisson. Un nouveau croisement entre PSE et *C. canephora* (CAN), fort producteur de caféine et d'ACG, est en cours de cartographie. Une évaluation biochimique est également commencée. Deux banques d'ADNc ont été obtenues à partir de feuilles de CAN et de graines récoltées à différents stades de maturation. Les gènes impliqués dans le contrôle de la teneur en caféine et ACG sont recherchés. La première approche utilisée consiste à trouver des séquences hétérologues équivalentes pour définir des amorces spécifiques déduites des domaines conservés de ces gènes. Les fragments correspondants seront amplifiés et utilisés comme sondes pour trouver une correspondance possible avec des QTLs. Si une telle co-localisation est trouvée, les gènes correspondants seront isolés et caractérisés. Une autre approche envisagée consiste à utiliser une double comparaison soustractive de banques d'ADNc, établies à partir de tissus de deux espèces différentes à deux stades différents de développement.