

## 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 de 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.

## **INTRODUCTION**

Coffee cup quality is based on the characterization of a large number of factors including taste and aroma. These factors are related to the biochemical content of roasted beans. A thousand of compounds, appearing during roasting, are involved in coffee cup quality. These compounds rise from a smaller number of biochemical compounds present in green beans. Their presence could have a favorable effect on the coffee cup quality, as for trigonelline and sugars, or an unfavorable one, as for chlorogenic acids and caffeine (Clifford, 1985; Macrae, 1985).

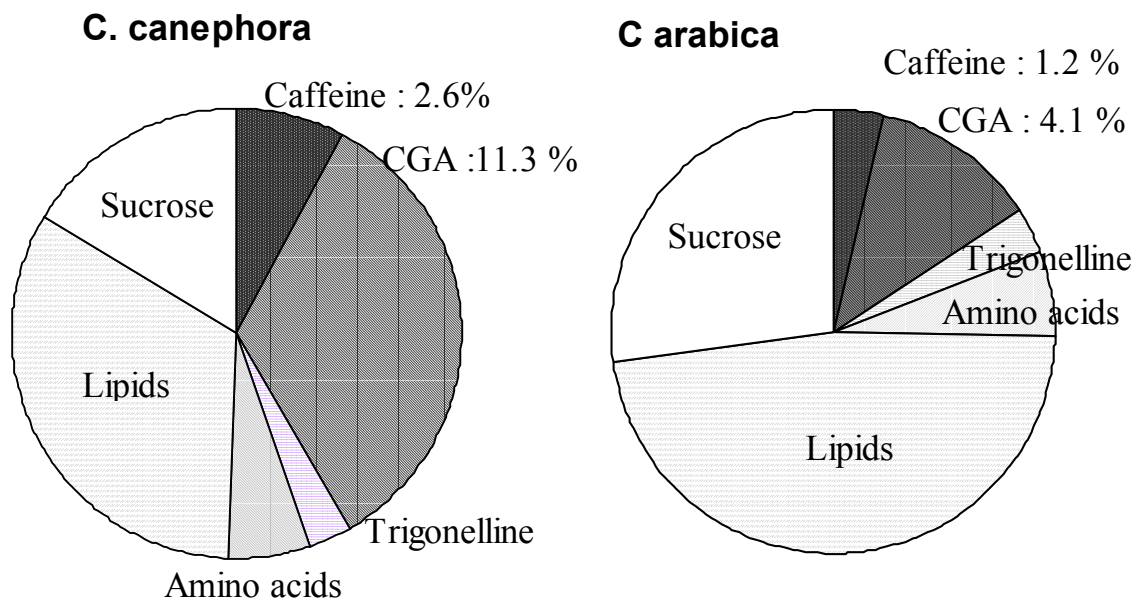
A modification of these compounds in the bean may have an effect, positive or negative, on the coffee cup quality. To improve coffee cup quality, we could increase the synthesis of favorable compounds or decrease the synthesis of negative compounds. In this work, we were only interested in the decrease of negative compounds. Compound level vary according to enzyme activity involved in the biochemical pathways. Corresponding enzymes are themselves regulated by genes for their structure and regulation. Finally, modifying gene expression, particularly for genes implicated in the biochemical compound content, may be useful for the improvement of the coffee cup quality. For doing this, two alternative proceedings are available: reducing gene expression involved in the biosynthesis of unfavorable compounds, which will limit their synthesis, or enhancing the expression of genes involved in their catabolism, which will increase their degradation.

The identification of such genes, candidate genes, requires knowledge of the compound content inheritance. In other words, it firstly requires a biochemical analysis in order to evaluate the biochemical diversity in the genus. Secondly, a genetic analysis of interspecific cross offspring will indicate inheritance. Parental species were those that appeared interesting during the biochemical analysis. Thirdly, genetic mapping and QTL location will be carried out.

## **RESULTS**

### **Biochemical analysis**

Compound content analysis was carried out on green beans harvested on the cultivated species, *C. canephora* and *C. arabica* (Figure 1).



**Figure 1. Biochemical content of green beans of *Coffea arabica* and *Coffea canephora***

The major studied compounds were sucrose, lipids, chlorogenic acids (CGA), caffeine and trigonelline. More than 80% of the *C. arabica* bean content (on dry matter basis: dmb) consisted in compounds such as sucrose, lipids, amino acids and trigonelline showing positive effect on coffee cup quality. In *C. canephora* beans, these compounds represented only 58% (dmb) of the biochemical content. At the opposite, the negative compounds content, as caffeine and chlorogenic acids, was more than two-fold in *C. canephora* beans.

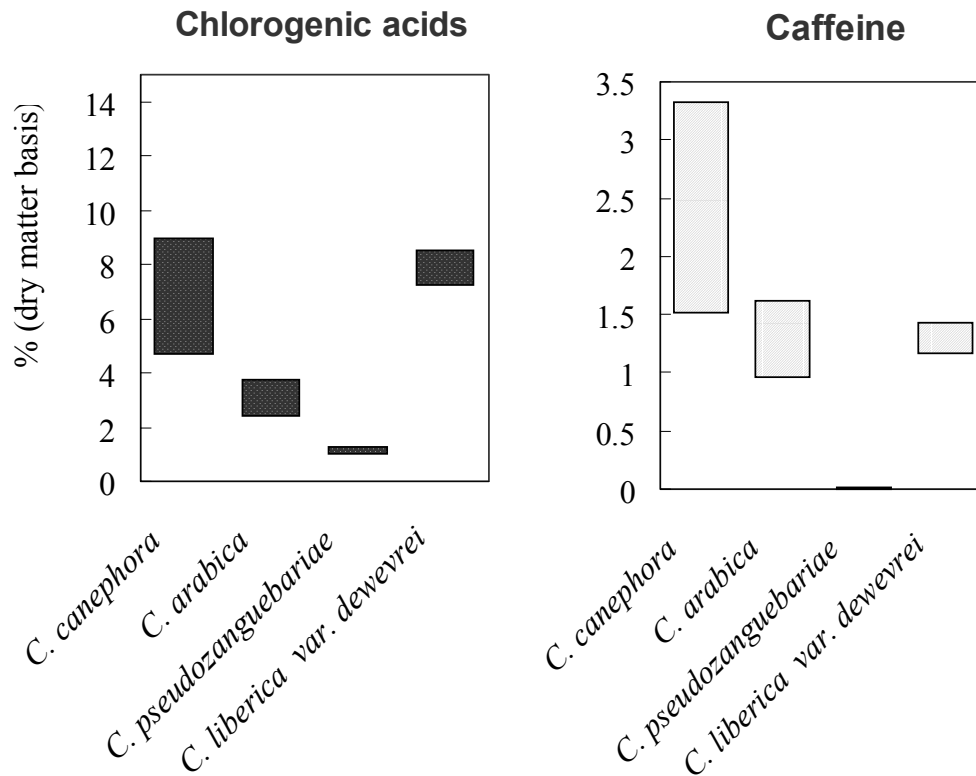
The biochemical diversity in the genus *Coffea* was studied for chlorogenic acids and caffeine in green beans (Anthony et al., 1993). Compared to the cultivated species, several wild species exhibited large differences in caffeine and chlorogenic acid contents. Two species were interesting (Figure 2): *C. pseudozanguebariae* (PSE) (low levels in chlorogenic acids and caffeine) and *C. liberica* var. *dewevrei* (DEW) (high chlorogenic acid content and a caffeine level close to *C. arabica*). To develop the genetics of corresponding biochemical compounds, the interspecific offspring (PSE x DEW or PSE x CAN) were used.

### Genetic analysis

Inheritance of compound content has been evaluated using the cross between PSE and DEW. Only the caffeine and chlorogenic acid content studies are reported here. Caffeine content was under a polygenic control with a recessive Mendelian gene coding for the absence. Barre et al. (1998) reported a weak effect of environment (6%) on the bean caffeine content. The polygenic control and additivity were also observed for chlorogenic acid content. Ky et al. (1999) showed that inheritance was polygenic and influenced by environment (20-50%). The same type of analysis was done in the offspring for sugar and trigonelline (Ky et al., 2000a; Ky et al., 2001). An AFLP genetic map has been obtained and quantitative trait loci (QTL) have been located for caffeine, CGA and trigonelline accumulation (Ky et al., 2000b).

## Identification of candidate genes

Our project focuses on CGA and caffeine, both involved in bitterness and the cross between *C. pseudozanguebariae* (PSE) and *C. canephora* (CAN), very different for bean compound contents (Figure 2). This cross is also being mapped and a biochemical evaluation is also underway.



**Figure 2. Variations in chlorogenic acid and caffeine content in green beans of different *Coffee* species**

A first approach was developed to investigate genes involved in the control of caffeine and CGA content. It consists into the search of heterologous sequences on data libraries. Consensus primers were obtained by comparing gene sequences available from other plants and by identifying their conserved domains. These primers will be used to amplify genomic DNA with different PCR techniques and the corresponding fragments will be used as probes. Gene was then mapped in order to find a possible co-location with the QTL yet described. In case of co-location, the corresponding genes will be considered as candidate gene, then isolated and characterized. The corresponding enzymes will be studied and the modification of gene expression will be checked by genetic transformation.

CGA and caffeine have negative effect on coffee cup quality, but they have a positive role in resistance to diseases. Then, modification must be done only in seeds. For this, the best tool seems to plan a genetic transformation *via Agrobacterium tumefaciens* carrying a plasmid containing a cassette with a fruit-specific promoter gene, the gene of interest and a terminator. In conclusion, preliminary results for five enzymes of the biosynthesis pathway, the amplification of genomic DNA from *C. canephora* with consensus primers conducted to the isolation of homologous probes. Analysis of genetic diversity and mapping are now initiated.

A double differential screening of cDNA libraries (or protein profiles obtained by two dimensional electrophoresis) was also developed. Two cDNA libraries were obtained from CAN, one from leaves and the other from beans at different maturation stages. The same procedure will be done for PSE.

## REFERENCES

- Anthony F., Clifford M.N., Noiro M., 1993 Biochemical diversity in the genus *Coffea* L.: chlorogenic acids, caffeine and mozambioside contents. *Genetic Resources and Crop Evolution*, **40**: 61-70.
- Barre P., Akaffou S., Louarn J., Charrier A., Hamon S., Noiro M., 1998. Inheritance of caffeine and heteroside contents in an interspecific cross between a cultivated coffee species *Coffea liberica* var “*dewevrei*” and a wild species caffeine free *C. pseudozanguebariae*. *Theor. Appl. Genet.*, **96**: 306-311.
- Clifford M.N., 1985. Chlorogenic acids. In : R.J. Clarke and R. Macrae (eds), *Coffee, Volume 1: Chemistry*, pp. 153-202. Elsevier Applied Science, London.
- Ky C.L., Barre P., Lorieux M., Trouslot P., Akaffou S., Louarn J., Charrier A., Hamon S., Noiro M., 2000b. Interspecific genetic linkage map, segregation distortion and genetic conversion in coffee (*Coffea* sp.). *Theor. Appl. Genet.*,...
- Ky C.L., Doulbeau S., Guyot B., Charrier A., Hamon S., Louarn J., Noiro M., 2000a. Inheritance of sucrose content in the interspecific cross: *Coffea pseudozanguebariae* x *Coffea liberica* var “*dewevrei*”. *Plant Breeding*, **119**: 165-168.
- Ky C.L., Guyot B., Louarn J., Hamon S., Noiro M., 2001. Trigonelline inheritance in the interspecific cross between *Coffea pseudozanguebariae* x *Coffea liberica* var “*dewevrei*”. *Theor. Appl. Genet.*,.
- Ky C.L., Louarn J., Guyot B., Charrier A., Hamon S., Noiro M., 1999. Relations between and inheritance of chlorogenic acid content in an interspecific cross between *Coffea pseudozanguebariae* and *Coffea liberica* var “*dewevrei*”. *Theor. Appl. Genet.*, **98**: 628-637.
- Macrae R., 1985. Nitrogenous compounds. In : R. J. Clarke and R. Macrae (eds), *Coffee, Volume 1: Chemistry*, pp. 115-152. Elsevier Applied Science, London.