From Traditional Genetics Towards Genomics: A New Approach for an Old Problem

M. NOIROT, A. de KOCHKO, T. LEROY, C. CAMPA, P. HAMON, M. DUFOUR, S. DUSSERT, J. JOUARN, H. CHRESTIN

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

SUMMARY

Since about ten years, advances in genomics increased exponentially, as well as their application in breeding. Now, I would like to present the contribution genomics in the coffee research project of an IRD-CIRAD team. Other talks will be developed today by this team.

Traditionally, coffee breeding, as in other crops, depends on the mode of reproduction. In autogamous crops, as *C. arabica*, we use the genealogic selection. In contrast, in allogamous plants as *C. canephora*, the reciprocal recurrent selection is applied. The second factor acting on breeding way is the range of the intraspecific diversity in regards to the objectives and the access possibilities to the interspecific diversity. In practice, introgression of wild traits was attempted for the two cultivated species. Another breeding way consisted to obtain interspecific F1 hybrids, as in the *Arabusta* program.

The new program, we proposed, is based on genomics and presents three objectives. The first objective is to define new tools for sustainable breeding. This concerns the marker assisted selection strategy and the genetic transformation. The second objective is to improve cup quality with two steps: firstly, to locate and identify genes and, secondly to analyse their regulation of expression. The third objective concerns the cryopreservation of genetic resources.

NEW TOOLS FOR SUSTAINABLE BREEDING

In order to define tools for sustainable breeding, we have to understand relationships between gene location and genome structure an one hand, and between gene regulation and genome structure on the other hand. By this point of view, coffee genus is a model. Indeed, there are 22 chromosomes in all diploid species, whereas the genome size of diploid species varies from 0.9 pg to 1.7 pg. Secondly, in interspecific crosses, sterility of F1 hybrids rises from presence of univalents at the meiosis metaphasis and there is a relationship between F1 hybrid sterility and parental genome size difference. Since genomes are homologous, coding gene number should not vary strongly. Consequently, genome size differences could be due to repeated sequences. But, what is the role of uncoding repeated sequences in terms of evolution? Indeed, a difference of genome size should have a selective advantage to invade the population. And this advantage could concern the gene transfer through recombination limitation and distortion increasing. This leads to the notion of qualitative reproductive barriers. Some genes can be transferred by introgression, others not.

Concerning the repeated sequences, we look for their identification their location and their quantification on one hand, and an analysis of their roles on distortion of segregation, recombination, chromosome pairing and gene expression.

A second way to study the structure of genome is the within-genus synteny (synteny is the resemblance between genetic map). To do that, we will compare four interspecific maps. All maps include *C. canephora*. Other parents were:

- *C. pseudozanguebariae* selected because it is a caffeine-free species and its genome size is low (1.1 pg);
- *C. eugenioides* , selected for its affinity with *C. arabica* genome;
- *C. liberica,* chosen for its affinity with *C. canephora* but also for its seed weight and its fructification synchronisation;
- lastly, *C. heterocalyx* was selected for its high genome size (1.7 pg) and its autogamy.

We also compare these maps with an intraspecific map between Guinean and Congolese trees.

We expect to relate the segregation distortion variation and the recombination rate differences to the parental genome size differences and the gene function and fitness. Indeed, what are genes present on distorted segment and what are their role in the fitness?

COFFEE CUP QUALITY GENES

The second objective of our project is to improve cup quality. So, we have to look for implied genes. This is a five steps approach.

- the genetic variance analysis give results on the relative influence of genotype and environment;
- the QTL analysis allows to define gene number and their location;
- this will be followed by the gene identification;
- their co-location with QTL should allow to define candidate-gene;
- the last step will consist to study their expression regulation.

First results for caffeine

Studies were carried out on an interspecific cross between C. *pseudozanguebariae* (PSE) and *C. liberica dewevrei* (DEW). Parental species presents differences for caffeine: 0% in PSE, and 1% in DEW, and for fructification time: 10 weeks in PSE, and 10 months for DEW.

We decomposed caffeine content in theses components. Indeed, caffeine content resulting from an accumulation, it can be decomposed in caffeine flow by day and day number of accumulation. In addition, we were interested by the CAF/CQA ratio.

We obtained 4 independent QTL:

- one for the fructification time;
- one for the caffeine flow by day;
- one for the CAF/CQA ratio;
- and one explaining the presence/absence of caffeine in parental species.

Further studies

To identify candidate genes, we have four possible ways knowing biosynthesis pathway and using bio-informatics, we expect obtaining of gene sequences from other plants. Proteome comparison between PSE and CAN fruits using two-dimensional electrophoresis will be carry out if the first way gives no expected results. Transcriptome comparison between PSE and

CAN fruits using cDNA libraries will also be carried out. Simultaneously, EST mapping is underway.

In addition and to summarise, all these points emphasise the importance of genetic resources as gene do nor for breeding.

GENETIC RESOURCES CRYOPRESERVATION

Genetic resources cryopreservation constitutes the third part of our project. Coffee genus is again a model. Why? Because there is a large diversity for tolerance to dehydration and low temperatures. Our first objective is to analyse the physiological basis of tolerance through scavenging effects and lipid behaviour at cold temperature. Today, Stephane Dussert showed that sucrose content does not explain tolerance diversity. He developed also a model to estimate seed dessication sensitivity in various *Coffea* L. species. This led to obtain the first *C. arabica* cryopreserved gene bank at CATIE.

The second objective is to establish the genetic basis of tolerance. With two approaches:

- the quantitative inheritance approach in order to analyse additivity;
- the location of QTL for tolerance, but also for lipid composition.

To do that, the PSE DEW cross is again useed. Indeed, PSE is tolerant to cryopreservation, whereas DEW is totally sensitive to dehydration and low temperature.

CONCLUSIONS

The three objectives of our project release on a team IRD-CIRAD. The team includes ten researches, three ingeeners and five technicians.

Partnership concerns traditional partners including CNRA, BRG, IPGRI and CATIE. It also concerns new partners in India, Uganda and Spain, but the list is open. We propose to partners collaboration on genomics projects, training masters and PhD and technicians too. This will concern molecular biology and data analysis. We can also propose expertise.