Study of open pollination in *Hevea* using microsatellites

The international survey organized in 1981 by IRRDB in the Amazon forest of Brazil has made it possible to broaden the base of germplasm available for *Hevea* breeding. Three successive projects to characterize this Amazonian germplasm were funded by the European Union over a cumulated period of twelve years. Different genetic groups were identified. Their improvement requires intense genetic mixing followed by selection. To that end, seed gardens containing samples of Amazonian genotypes from the different groups were set up in Côte d'Ivoire under close collaboration between CIRAD and CNRA. In order to validate this new approach for Hevea, it is worth specifying the conditions of the reproduction scheme resulting from open pollination in seed gardens.

The *Hevea* open pollination reproduction scheme

Hevea varieties are budded clones and seeds are usually only used for the production of rootstocks in nurseries. The seed gardens studied here did not, as for numerous forest species such as eucalyptus, have a seed production function.

The methodological study involved a planting set up in a forest environment at CNRA's Divo station, with total isolation from contamination by pollen from outside rubber trees. The planting comprised 50 Amazonian genotypes from three survey zones: Acre, Rondonia and Mato Grosso, along with the male-sterile clone GT 1, for which any selfing is excluded in principle. The seed garden therefore consisted, in equal parts (25/25), of genotypes belonging to groups I, II and III on the one hand: Acre and Rondonia except Pimenta Bueno (PB) district in Rondonia, corresponding to the western Amazon, and group IV on the other hand: Mato Grosso and Rondonia-PB corresponding to the eastern Amazon. These genetic groups were identified by Seguin *et al.*, based on molecular diversity. Each genotype was represented by 5 to 6 budded trees, randomly distributed in the seed garden, so as to maximize the number of possible recombinations. Each tree in the planting was intended to take part in reproduction as both a male pollen emitter and as a female seed producer, in accordance with its specific flowering and fruiting traits. Only clone GT1 functioned exclusively as a female. The questions raised involved the quality of genetic mixing of the rubber tree population undergoing open pollination within the planting design adopted. For a progeny harvested from a maternal genotype, what were the diversity and balance of the paternal contributions? What was the rate of selfing depending on the genotypes? Were the progenies of the different genotypes similar in their paternal make-up? Depending on the answers to those questions, what harvesting and selection strategy should be adopted to maximize variability in the harvested progenies and to optimize selection between and within the progenies?

Of the genotype characterization criteria (agro-morphological traits, isozymes, molecular markers), microsatellites appeared to be particularly efficient for varietal identification and especially for determining paternity, given the large number of different

alleles found at each locus. The exploratory power of this molecular tool on *Hevea* had already been shown in a project funded by the French genetic resources bureau, involving the constitution of core collections, to which IRD and CIRAD have contributed. Using microsatellites to study seed gardens was made possible through the funding of a thematic research project programmed by CIRAD and involving CNRA in Côte d'Ivoire, for rubber and FOFIFA in Madagascar, for eucalyptus.

For the work in Côte d'Ivoire, and for the 1998 fruiting season, 468 plants derived from 14 genotypes (including GT 1) and 29 different mother-trees were analysed. For the 1999 season, 338 plants derived from 4 genotypes (including GT 1) and 8 different mother-trees were analysed. These large samples enabled a precise evaluation of the paternal participation of each of the 50 genotypes in the seed garden over two consecutive seasons.

Eight microsatellite markers were used, all independent from each other and distributed over six linkage groups of the *Hevea* genome; they comprised from 11 to 21 alleles. For each plant analysed, paternity was sought using Cervus software. Genotyping of all the parental genotypes was carried out first of all, in order to specify the possible origins of alleles observed in each progeny to be analysed. A check of clonal conformity of the parent trees in the seed garden, carried out at CNRA using isozymes, revealed the existence of 15 off-types within the seed garden; microsatellite genotyping of those 15 trees has yet to be carried out.

Paternity identification using microsatellite markers

Identification of the most likely father was completed for 463 of the 468 progenies harvested in 1998. Of the 338 progenies harvested in 1999, identification of the most likely father was completed for 334 plants. The few cases in which it was not possible to identify the probable father may have been due to the existence in the seed garden of 15 trees that were not true-to-type and not genotyped. Despite a high confidence level in paternal identification, there were sometimes mismatches between parents and progenies at certain loci. The error rate, estimated over the two years of the study, was 0.6%. These errors were probably due to inaccurate allele readings on the microsatellite migration gels. In order to take these errors into account, a probabilistic approach had to be taken to paternity testing by microsatellites. The probability that a genotype was the father of a given progeny was calculated by the Lod score method (Lod = logarithm of the odds ratio).

The sensitivity of the analysis depending on the number of microsatellite markers used was assessed. It declined as the number of markers decreased. When switching successively from 8 markers to 7, 6, 5 and 4 markers, whilst seeking to keep the most effective markers, the percentage of changes in paternal identification, compared to the best situation established with 8 markers, was 5%, 10%, 15% and 30% respectively. The small percentage of changes seen when switching from 7 to 8 markers (5%) afforded a high level of confidence to paternity identifications carried out with 8 markers.

These paternity identifications made it possible to deduce the degree of paternal participation of the different genotypes in the pollinations. A major result was that the distribution proved to be highly unequal. In the sample of seeds harvested in 1998 and 1999, 11 and 18 genotypes respectively, out of the 50 potential pollinators, did not participate in the recombination. When analyses were grouped for the two years, 8 genotypes did not make any paternal contribution. On the other hand, 4 genotypes accounted for 40% of total paternal contribution, 14 genotypes accounted for 80% of total paternal contribution and 25 genotypes accounted for 95% of total paternal contribution.

The genotypes were found to have quite a stable performance for their paternal contribution over the two harvesting years (table). The degree to which a given genotype made a paternal contribution to the seed garden progeny therefore seemed to be a genetic characteristic with high variability, which was relatively little influenced by environmental effects.

Whilst the proportion of mother-trees in the seed garden belonging to genotypes of genetic group IV was 50%, the share of those trees in the paternal contribution was 83%. The major pollinators therefore mainly belonged to group IV (East).

The distributions of paternal contributions for a given mother-tree were examined on five mother-trees for which the analysed progenies exceeded 40 seeds. It was found that the rate of participation in the total paternal contribution per mother-tree of the two predominant genotypes varied quite strongly between 0.34 and 0.71. In addition, these two predominant genotypes were not the same for a given two mother-trees. Although marked by certain important parents, the pollen pools that pollinated these trees were therefore not identical. The figure shows spatial distribution of the paternal contributions for these 5 mother-trees—and for an additional 3 mother-trees.

For the grouping of the 1998 and 1999 harvests (except progenies harvested from GT 1), the average selfing rate of the analysed genotypes (including possible crosses between trees of the same genotype) was estimated at 4.8%. As could be expected for a male-sterile clone, no selfing was seen for GT 1. This selfing rate, which was estimated less precisely for each genotype separately, due to limited numbers of individuals, varied considerably from one

Table. Contributions of paternal genotypes to formation of the progenies analysed for the 1998 and 1999 seed harvests.

Paternal population	Number and proportion of paternal genotypes	Number of fathers per genetic group*		Number and proportion of progenies		
		Groups (I, II, III) West	Group IV East	1998	1999	Cumulated 1998-1999
First 10 fathers°	10 (20 %)	0	10	300 (65%)	237 (71%)	537 (67%)
First 25 fathers°	25 (50 %)	7	18	426 (92%)	323 (97%)	749 (95%)
MT + RO/PB	25 (50 %)	0	25	385 (83%)	274 (82%)	679 (83%)
AC + RO	25 (50 %)	25	0	78 (17%)	60 (18%)	138 (17%)

° classed by decreasing number of progenies for cumulated years 1998 and 1999.

* according to Seguin et al. 1999.

AC = Acre, RO = Rondonia, MT = Mato Grosso, PB = Pimenta Bueno district of Rondonia.

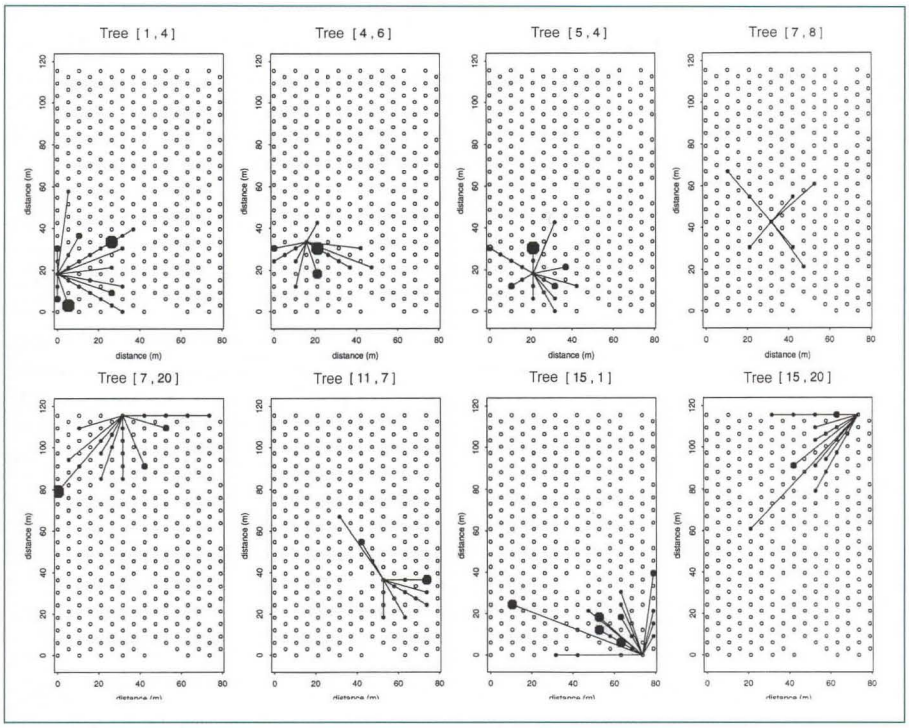

Figure. Microsatellite analysis of gene flows in a *Hevea* seed garden: paternal contributions in the progeny of 8 mother-trees in 1998. Each point represents a living tree in the seed garden. The lines link the mother-tree to the nearest trees of the parental genotypes that contributed to the progeny of the mother-tree in question. The size of the black circles is proportional to the number of progenies derived from that paternal genotype.

mother-tree to the next. The following individual results were obtained for five mother-trees: 0%, 1.8%, 6.8%, 8.6%, along with 15.9% for a mother-tree located on the edge of the seed garden and with a very large canopy. Observations on two mother-trees of the same genotype, with 56 and 44 progenies respectively, indicated selfing rates of 1.8 and 6.8%. It seemed that these variations could be partly explained by the genetic selfing ability, but also by the architecture of the mother-trees—canopy volume—the flowering and fruiting ability, along with the environment and canopy access to light—position of the tree on the edge of the garden or next to a gap.

An analysis of seeds from the same fruit revealed the possibility that two different fathers may have contributed to seed formation; this was seen in 4 of the 9 fruits studied. It is therefore not possible to envisage analysing just one seed per fruit to determine the paternal origin of the other seeds.

A further study, currently being conducted at CNRA, is intended to compare these results with architectural and phenological characteristics, along with the flowering periods and the abundance of flowering in the genotypes and individual trees in the seed garden. Indeed, considerable architectural differences are known to exist between clones from Acre and Rondonia—often marked by strong apical dominance, limited branching and flowering—and clones from Mato Grosso—often marked by strong branching and flowering. This research will provide a classification of genotypes according to their fruiting ability, and will elucidate the effect of staggered flowering on any formation of preferential crosses. The study takes into account climatic differences over the two years, and the effect of tree canopy growth from one year to the next. It could provide indications on distance effects associated with pollination.

Application to genetic variability management, and selection

The power of microsatellites, and the suitability of these markers for studying the reproduction scheme in seed gardens, are clearly shown through the abundance and precision of the information gathered.

As analyses stand at the moment, the study revealed substantial variability in the performance of parental genotypes as regards their maternal contribution to fruiting and their paternal contribution to male flowering, pollination and fertilization, and their selfing rate. The reproduction scheme in the seed garden was therefore far from a panmictic situation where all the combinations between parents would be equally probable in the progenies.

The relatively small number of genotypes (50), combined with the imbalance in paternal contributions, limited the recombinations. Since the abundance of flowering in the genotypes appeared to be the main factor limiting the diversity of paternal contributions, the propagation of each genotype by budding and randomization of trees in the field did not improve the situation much. However, it is advisable to harvest small samples of seeds from the largest possible number of mother-trees. The "East" group accounted for 83% of pollination: in order to increase the genetic diversity of the overall progeny, seeds could preferably be harvested from genotypes with a low paternal contribution, in this case clones from the "West" group, which would give a high proportion of hybrid progenies between the two groups. On the other hand, harvesting seeds from genotypes of the "East" group will give a large proportion of progenies derived from the internal recombination of that group.

The composition of the seed garden does not seem suitable for obtaining progenies of half-sib families derived from the same diversified and balanced pollen pool, which is prejudicial to effective selection.

Different arrangements need to be considered to try and diversify and improve the balance of paternal contributions in the seed garden. However, it is not sure they will be effective and would have to be systematically assessed.

Towards microsatellite marker assisted recombination

The power of paternity determination by microsatellites, used to assess the genetic recombination of the seed garden, could also be used as a tool for identifying the paternal origin of each seed harvested, after the event. The cost and time required for analysis do not necessarily rule out this possibility, notably because these analyses are set to become increasingly automated.

Sorting could be carried out in the large quantity of harvestable seeds, to make up batches according to the objectives sought:

- an overall progeny maximizing genetic diversity,
- a specialized progeny for a type of recombination, between- or within-group,
- progenies derived from the same pollen pool,
- progenies of full-sib families possibly structured in mating designs.

With this approach the particular conditions for the genetic recombination of a seed garden have very little influence—apart from isolation conditions designed to limit pollution from outside pollen—and the introduction of various constraints can be avoided when setting up seed gardens. The availability of a recombination seed garden with a large enough number of parents having undergone microsatellite genotyping could thus constitute an abundant source of naturally renewable breeding material, which is easy to dispatch in "seed" form for use by distant partners, and which is preservable in living form in the nursery at high density, sortable and usable over several years, in line with objectives which may evolve over time. This approach, which is envisaged for managing and utilizing wild germplasm, could be extended to selection within the advanced Wickham population, notably to carry out further genetic mixing involving little used parents.

Microsatellite genotyping of seed gardens is an example of marker-assisted selection applied to Hevea. The control of gene flows, which has been made possible by molecular markers, opens up new methods for two key aspects of rubber tree selection: dynamic management of genetic resources, and the acquisition of legitimate, half-sib and full-sib progenies.

In addition to seed gardens, molecular marking of gene flows in *Hevea* will also be very important when planting out genetically modified rubber trees. Microsatellite markers will make it possible to answer the question of pollen propagation distances, and the risk of GMO dispersion in this species. ■

List of publications

CLÉMENT-DEMANGE A., SEGUIN M., VERHAEGEN D., LIDAH Y. J., RODIER-GOUD M., BLANC G., RAZAFIARIVELO S., RAZAFIMAHORO V., 2001. Utilisation de marqueurs génétiques pour l'étude de systèmes à pollinisation naturelle ou artificielle chez trois espèces tropicales pérennes. Compte-rendu final de l'Atp Cirad 98/20.

LIDAH Y. J., 1998. Observations dans les jardins de pollinisation libre d'hévéa. Brève présentation des résultats de la campagne 1997. Abidjan, Côte d'Ivoire, IDEFOR-DPL note (internal document).

SEGUIN M., FLORI A., LEGNATÉ H., CLÉMENT-DEMANGE A., 1999. L'hévéa. *In*: Diversité génétique des plantes tropicales cultivées, P. Hamon, M. Seguin, X. Perrier, J.C. Glaszmann éd., Montpellier, France, Cirad, coll. Repères, p. 241-269.