Franco-Thai Project 2005-2008:
"Towards the improvement of rubber tree productivity"

Molecular genetic markers and rubber breeding in Thailand 1 – Genetic mapping of the family RRIM600 x PB217 by use of microsatellite markers

K. Prapan[1], N. Lekawipat[1], C. Weber[2], M. Rodier-Goud[2], A. Clément-Demange[2], M. Seguin[2]

[1]RRIT-DOA, Chachoengsao Rubber Research Center, Chatuchak, Bangkok 10900 - Thailand (rrit@doa.go.th)

[2]CIRAD Tree crops department –TA 80/01 – Avenue Agropolis, 34398 Montpellier Cedex 5 - France (marc.seguin @cirad.fr, andre.clement-demange @cirad.fr)

Abstract

Rubber breeding has long been based on conventional breeding and quantitative genetics based on phenotypic observations. Since recently, molecular genetic markers have brought new possibilities for characterizing genotypes, with view to identify cultivars, or their parents, to analyse genetic diversity, to establish relationships between agricultural traits and genetic factors (QTLs), and to identify genes of interest. In order to analyse the genetic bases of latex production, the cross RRIM600 x PB217 was created in Thailand for genotyping the progenies and building a genetic linkage map. The two parents were chosen for their contrasted physiological behaviour, as shown by the metabolic typology of cultivated rubber clones. Based on the double pseudo testcross strategy, two parental maps and a consensus map were built, with 334 progenies, 247 microsatellite markers, and 198 AFLP markers. The 18 linkage groups were assigned to the corresponding groups, equivalent to the 18 chromosomes of the *Hevea* haploid genome, of a Cirad *Hevea* genetic map used as reference. This new consensus map will be used together with field data for QTLs detection on heritable traits. Efforts will be priorily focussed on studying the physiological parameters of the latex diagnostic, in association with latex production.

Keywords:
Hevea brasiliensis ; breeding ; latex production ; metabolic typology of rubber clones ; latex diagnostic ; molecular genetic markers ; microsatellites ; genetic linkage mapping ; heritability ; QTLs ; expressed genes.

This research was funded by Agropolis (Montpellier, France), and the Programme of French-Thai Cooperation for Education and Research.

1. Introduction

Since the introduction of rubber tree to Thailand in 1900, rubber cropping was developed by smallholders in the form of seedlings. From the beginning of the 1960s, the involvement of the Governement and ORRAF for replanting old seedlings by new monoclonal plots, mainly based on clone RRIM600, was responsible for a spectacular increase in yields and planted areas. However there is a need for clonal diversification and increase in genetic performances. RRIT-DOA develops a conventional rubber breeding programme for the joint production of latex and rubberwood, with a classical selection scheme including : a) preliminary screening of seedlings issued from hand pollination (recombination, creation of genetic variability), b) cloning of the best trees, c) early selection in small scale clonal trial, d) final selection in a multilocal large scale clonal trials network, e) clonal recommendations for the improvement of the performances and for the diversification of the planted clones, and f) distribution of budwood to the development. The current recommendations in Thailand are as following:

Latex clones:
Class 1: RRIT251, RRIT226, BPM24, RRIM600
Class 2: RRIT209, RRIT214, RRIT218, RRIT225, RRIT250, RRIT319, RRIT405, RRIT406, RRIC100, RRIC101, PR302, PR305, Haiken2

Latex-timber clones :
Class 1: PB235, PB255, PB260, RRIC110
Class 2: RRIT312, RRIT325, RRIT404, RRIT407, RRIT409, RRIC121

Timber clones:
Class 1: Chachoengsao50, AVROS2037, BPM1
Class 2: RRIT401, RRIT403, RRII118, RRII203.

For improving rubber breeding efficiency (genetic progress per time unit), Thai universities and research institutions develop biotechnologies in cooperation with other countries such as France.

Living populations are determined by a combination of genetics and environment. For one population, heritability of one trait is the share of the genetic variation over the total phenotypic variation (genetics + environment). Most of the traits studied in rubber are “quantitative”: they display a continuous variation among individuals due to genetics and to the environment. Quantitative genetics, which is used in conventional rubber breeding, assumes that these traits are genetically determined by a large number of genes that cumulate many small effects. Measurements on different traits (phenotypic values) are used for estimating the genotypic value of each trait in each rubber genotype (seedling or clone), thereby allowing the selection of the best genotypes.

Molecular biology now holds great promises. Many methodologies make possible the identification of more and more genes. These candidate genes are studied for identifying their function and better understanding the most important metabolic pathways (molecular physiology). Most efforts are focussed on model plants such as *Arabidopsis* or rice. As the benefit for rubber breeding may be only for the long term, integrating molecular biology with breeding is a real challenge.

Molecular genetic markers (identified fragments of the DNA genome) can be observed independently from the environment. As their heritability is of 100 %, characterizing one rubber clone for one marker requires only the DNA of one plant, and can be done at a very early stage, which can be very useful for early selection (reduction of the time of selection of new clones). Some of these molecular genetic markers are full copies or partial fragments of expressed genes, but most of them are not expressed: they are said “neutral” or “anonymous” as they do not contribute to the expression of the genotype. Now one question is how can such markers be useful for breeding? We hereafter refer to “markers” as “non expressed anonymous molecular genetic markers”.

Coming after the use of phenotypic morphological markers observed on the living plants, and after isozymes, proteins that can be used for the identification of genotypes (rubber clones), the development of a huge number of molecular genetic markers has made possible the detailed identification of the two alleles in each known locus of any diploid genotype (genotyping). At the level of the research, this was developed in rubber in the framework of the Thai-French cooperation for analysing the genetic diversity of the rubber tree germplasm, with a comparison of the efficiency of anonymous markers (microsatellites) and of expressed genes (Lekawipat et al, 2003a, 2003b, 2004). At the level of the development, microsatellites are now also used in Thailand for controlling the clonal conformity of the budwood gardens and budded plants issued from nurseries (quality control applied to the process of vegetative multiplication in rubber).

Microsatellite markers (or Simple Sequence Repeats, SSRs) were used in these studies. They are made of a very short DNA sequence with a varying number of repetitions from one genotype to another; differences in the length of the fragments can be observed by use of labelling, PCR amplification, and electrophoresis. They have many advantages, which explains that they are more and more widely used. For microsatellite loci, most alleles are codominant (no dominant or recessive alleles), which makes possible the direct observation of the 2 alleles in each locus of any genotype. In a population, a high number of different alleles (about 10 to 20) can be found in microsatellite loci (multiallelic markers with a high polymorphism), which provides a high resolution for distinguishing the different genotypes.

Genetic mapping for QTL identification is a general methodology that can be developed due to the availability of increasing numbers of molecular genetic markers. This methodology was chosen and developed in rubber by the French-Thai project : “Towards the improvement of rubber tree productivity” because it puts the bases for integrating laboratory molecular

genetics and field breeding at the level of latex production and of other traits. As a matter of fact, the general assumption of quantitative genetics (many minor genes determining one trait) is not always true. For the most heritable traits, it is often possible to identify major genes, or individual genes able to determine a significant share of the variation of the trait in a population. The principle, for a chosen population, is to associate, genetic mapping of the markers with field characterization of some targeted agricultural traits, in order to find relationships between these traits and some markers, and to localize the concerned areas on the genome (Quantitative Trait Loci, QTLs). Markers related with these QTLs might be used, under some specific conditions, for breeding in a Markers Assisted Selection (MAS). Expressed genes for which probes are available can also be mapped and associated to neighbouring markers. The QTL approach is a primary exploration, not able to precisely localize genes of interest but to inform about the existence of such genes and to provide roughly their localization with indication of the neighbouring markers. In rubber tree, QTL mapping had been previously successfully applied to South American Leaf Blight resistance traits (Lespinasse et al 2000b, Le Guen et al 2003).

In rubber, whereas growth clearly is a very complex trait, there are some indications that latex production could be determined by some significant genes, and the biochemical parameters of the latex diagnostic showed a good heritability, especially inorganic phosphorus content and sucrose content (Gnagne et al., 1997). For latex production, rubber clones are structured within a metabolic typology which was established from the data of the latex diagnostic. Considering inorganic phosphorus content (Pi) and sucrose content (Suc) in the latex of tapped trees, clones can be classified depending on their metabolic activity and on their ability to maintain a high availability of sucrose in the laticifer (source of energy and basis for isoprenic synthesis). Clones such as PB260 or RRIM600 have a high metabolic activity and a fast increase in production after the initiation of tapping but a medium productivity in the long term. By contrast, clones such as PB217 express a slow increase in production but a very high potential after 5 tapping years. Can genetic recombination combine these two properties in a complementary way for the selection of new clones with higher performances ?

The main objective of this research was to apply the methodology of genetic mapping and QTL identification to the analysis of latex production, in relation with the metabolic typology of the clones and the physiological parameters of the latex diagnostic. This first communication presents the currently achieved genetic mapping of the family RRIM600 x PB217.

2. Materials and methods

2.1. Plant material

Following the most common method in vegetatively-propagated crops, and due to their heterozygous nature, a full-sib F1 family of a large size, made of the progenies from two clonal parents, was created by RRIT-DOA at Crrc station (Chachoengsao) in 2000.

The female parent RRIM600 was chosen because of its high metabolic activity and relatively low level of sucrose in the laticifer. The male parent PB217 was chosen for its low metabolic activity and high level of sucrose in the laticifer. Crossing these two clones is assumed to generate recombination of the genetic factors of the metabolic typology and to display a wide variability among progenies for the concerned parameters of the latex diagnostic. Moreover, segregation of alleles during meiosis and recombination in the progenies make possible the analysis of linkage between alleles of neighbouring loci, which is at the base of genetic mapping.

2.2. Molecular resources

Identification of microsatellite sequences was developed at Cirad based on the building of microsatellite-enriched libraries and sequencing under a grant from Genoscope/National Sequencing Center (Evry, France) in 1999 and 2000 (Seguin et al 2001). The microsatellite sequences belong to the public domain and are registered in an international DNA sequence database (EMBL/Genbank).

For achieving the mapping with a higher density of markers, another type of markers, although less informative (AFLP), were used in addition to the available polymorphic microsatellites.

2.3. Genotyping and genetic mapping

DNA from the progenies was extracted in Thailand. Genotyping (identification of the two alleles of each diploid microsatellite marker in each progeny) and genetic mapping were developed by two Thai researchers at Cirad-Montpellier with assistance of the French team until May 2005, with two intensive periods in 2002 and at the beginning of 2005. The full working time of one researcher necessary for achieving this genetic mapping was estimated at about 18 months.

Genetic linkage maps of each chromosome are made by determining how frequently the alleles of two markers are passed together from parents to offspring. Because genetic material is frequently exchanged during the production of gametes in meiosis (crossing over), groups of alleles, originally together on one chromosome, may not be inherited together. Closely linked markers are less likely to be separated by crossing over. The recombination rate “r” between two loci that are located on the same chromosome can be estimated ; it is considered as a “distance” between the two linked loci. The precision of these estimations is improved by a large number of progenies, by codominant and multiallelic markers such as microsatellites, and by the recombination between two very different parents, which increases the occurence

of loci with 4 or 3 different alleles (Lespinasse, 1999). Integration of the informations from every couple of loci in an iterative way allows the building of linkage groups with marker loci arranged in real order.

Genetic mapping was carried out by software JoinMap 3.0. Although not absolutely necessary, the strategy of double pseudo-testcross applied to the F1 cross between two heterozygous parents was chosen: it requires the building of two distinct parental maps prior to the building of a synthetic map. This strategy is more reliable as it makes possible to check that each marker is really "monolocus" and can be mapped at the same location in the two parental maps. The orders of the markers can be compared in the two maps. Then a synthetic "consensus" map is built by use of the common "bridge" markers between the two parental maps, and it can cumulate the informations from each parental map. A good genetic linkage map will comprise marker loci that are evenly spaced and span the genome. Average intermarker distances of 5 to 10 centimorgans would be optimal.

3. Results

The segregating family RRIM600 x PB217 was created with more than 600 progenies.

The parentage of 365 progenies supposed to be issued from the family RRIM600 x PB217 was checked in October 2001 by use of 10 microsatellite markers, and 334 were validated as actually belonging to the family. Genetic mapping was applied to these 334 progenies.

All the available microsatellite primer pairs were screened for PCR amplification efficiency and for genetic polymorphism on the two parents RRIM600 and PB217 as well as 3 available grand-parents (PB49, PB86, and TJIR1), and 276 polymorphic microsatellite markers were identified for genotyping the 334 progenies.

For achieving the mapping with a higher density of markers, AFLP markers were added to microsatellites for genotyping 196 out of the 334 progenies (the progenies which are phenotyped in the field). Genetic linkage groups were built and then assigned to the 18 chromosomes of the rubber haploid genome, by reference to the *Hevea* genetic map already built at Cirad (Lespinasse et al., 2000a) and with common "bridge" markers.

The PB217 parental map encompasses 189 microsatellite markers clustered in 21 linkage groups. The RRIM600 parental map encompasses 199 microsatellite markers clustered in 18 linkage groups. The comparison of marker orders for common markers in the two maps confirmed that each microsatellite marker was assigned to the same locus in the two parental maps; microsatellites were so confirmed as monolocus markers. The consensus map RRIM600 x PB217 was built with 247 microsatellite and 198 AFLP markers (a total of 445 markers) clustered in 18 linkage groups or chromosomes and covering a total length of 2100 cM.

4. Discussion

This mapping was facilitated by the reference to the first published map of *Hevea* genome based on the cross PB260 x RO38 (Lespinasse et al. 2000a) that was developed for studying resistance to the South American Leaf Blight (SALB). Sixty common "bridge" microsatellite markers were used for complete assignment of the linkage groups of the new map to their corresponding groups in the reference map (equivalent to the 18 chromosomes of the haploid *Hevea* genome), and for marker order comparison, control and correction. This was achieved by using microsatatellite markers previously mapped by Lespinasse et al. (2000a), and by mapping several markers from the new RRIM600 x PB217 map on the formerly built reference map.

Microsatellite markers appear well distributed over all the *Hevea* genome, and rarely form clusters of markers, thereby confirming their efficiency for genome mapping in *Hevea*. More than half of these markers were common to the two parental maps, thereby allowing to achieve a reliable synthetic map. However, microsatellite markers by themselves were insufficient in number to saturate the map. In contrast, AFLPs appear frequently associated in clusters of tightly linked markers. In addition, if AFLPs reveal dominant markers and will be less informative than microsatellites for the future QTL analysis, they were nevertheless, very useful to rapidly achieve map saturation. Finally, our results confirm that the two techniques, AFLPs and microsatellites genotyping, are efficient complementary tools for efficient genome mapping in rubber tree.

The total length which could be deduced from the mapping of these 445 markers (247 microsatellites and 198 AFLP) was found equivalent to the 2144 cM-length initially obtained in the reference map (Lespinasse et al., 2000a) that was built with 717 markers including 285 RFLP and 359 AFLP markers. The average length between two neighbouring markers is a little less than 5 cM, and this new map, in its current stage, appears dense enough for QTL detection in relation with the factors of the metabolic typology of the clones and with latex production. However, there are some areas which can still be densified; the number of bridge markers between the two parental maps can be increased for improving the coverage of the parental maps and the precision of the estimated genetic distances between the markers in the synthetic map. This new map will be enriched with already known expressed genes in order to increase the possible number of links between genes, markers, and agricultural traits.

References

1 - Gnagne M., Clément-Demange A., Legnaté H., Chapuset T., and Nicolas D. - 1997
Results of the rubber breeding programme in Ivory Coast. Proc. International Rubber Research and Development Board Symp. on Natural Rubber in Vietnam, 13-15 October 1997, Vol. 1. In : M.E. Cronin (ed), p. 101-113.

2 - Le Guen V., Lespinasse D., Oliver G., Rodier-Goud M., Pinard F., and Seguin M. - 2003
Molecular mapping of genes conferring field resistance to South American Leaf Blight (Microcyclus ulei) in rubber tree. Theoretical and Applied Genetics 108: 160-167.

3 - Lekawipat N. - 2005
Development of genetic map of RRIM600 x PB217 based on microsatellite markers. Cirad-Biotrop, Montpellier, January-May 2005. Doras-Rubber : Towards the improvement of the productivity of the rubber tree. Genmap component : "Variability analysis and genetic determinism of some physiological characteristics of the productivity in Thailand".

4 - Lekawipat N. - 2004
Comparison of gene and non-gene specific molecular markers for evaluating genetic diversity in rubber (*Hevea brasiliensis* Muell. Arg). A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy (Tropical Agriculture), Graduate School, Kasetsart University, April 2004. ISBN 974-273-560-3.

5 - Lekawipat N., Teerawatanasuk K., Rodier-Goud M., Seguin M., Vanavichit A., Toojinda T., and Tragoonrung S. - 2003a
Genetic diversity analysis of wild germplasm and cultivated clones of *Hevea brasiliensis* Muell. Arg. by using microsatellite markers. *Journal of Rubber Research* 6, 36-47.

6 - Lekawipat N., Teerawatanasuk K., Vanavichit A., Toojinda T., and Tragoonrung S. - 2003b
Evaluating the genetic relatedness of wild and cultivated *Hevea brasiliensis* accessions with SSCP markers. *SABRAO Journal of Breeding and Genetics* 35, 123-134.

7 - Lespinasse D. - 1999
Cartographie génétique chez des plantes hétérozygotes non génétiquement fixées. Document de travail de la délégation Mathématiques et informatique appliquées n° 5/99. Cirad.

8 - Lespinasse D., Rodier-Goud M., Grivet L., Leconte A., Legnate H., and Seguin M. – 2000a
A saturated genetic linkage map of rubber tree (Hevea spp.) based on RFLP, AFLP, microsatellite, and isozyme markers. *Theoretical and Applied Genetics* 100, 127-138.

9 - Lespinasse D., Grivet L., Troispoux V., Rodier-Goud M., Pinard F., and Seguin M.- 2000b
Identification of QTLs involved in the resistance to South American leaf blight (Microcyclus ulei) in the rubber tree. Theoretical and Applied Genetics 100: 975-984.

10 - Prapan K. - 2002
Genetic determinism and localization of physiological components of the productivity of the heveaculture by QTL approach in Thailand : Genetic mapping of *Hevea brasiliensis* using microsatellite markers. Report. Rrit-Doa and Cirad, Montpellier, December 2001 - November 2002.

11 - Prapan K., Clément-Demange A., Teerawatanasuk K., Rodier-Goud M., and Seguin M. - 2004
Genetic mapping and field study of a full-sib family (RRIM600 x PB217) in *Hevea brasiliensis* (Genmap project). First results. " Towards the improvement of the productivity of the rubber tree ". Kasetsart University - Rrit-Doa - Inra-Piaf - Cirad Seminar, Bangkok, 27-28 May, 2004.

12 - Prapan K., and Lekawipat N. - 2001
Progeny legitimacy control to identify a sample of 365 progenies (RRIM600 x PB217) by using microsatellite markers. September 20 - November 30, 2001, Cirad Montpellier.

13 - Seguin M. - 2005
Second update of the mapping work of Genmap in Montpellier between September 2003 and December 2004 - April 2005. Development of a genetic map of RRIM600 x PB217 based on microsatellite markers. Doras-Rubber, Genmap component. Variability analysis and genetic determinism of some physiological characteristics of the productivity in Thailand.

14 - Seguin M. - 2003
Update of the mapping work of Ms Kanlaya Prapan In Montpellier (Agropolis funding), July 2003. Development of a genetic map of RRIM600 x PB217 based on microsatellite markers. Doras-Rubber, Genmap component. Variability analysis and genetic determinism of some physiological characteristics of the productivity in Thailand.

15 - Seguin M., Gay C., Xiong T.-C., and Rodier-Goud M. - 2001
Microsatellite markers for genome analysis of rubber tree (*Hevea* spp.). *In "Proceedings of the IRRDB Symposium 2001 – Biotechnology & Rubber Tree"*, Cirad, Montpellier, France, 25-28 September 2001, Montpellier, France, 6p.

Figure 1 : Genmap genetic linkage map issued from RRIM600 x PB217 family

Genetic map built from the 2 parents and 334 progenies of the cross RRIM600 x PB217, with 247 microsatellite and 198 AFLP markers localised (software JoinMap 3.0).