Int. Symposium on Oil Palm Genetic Resources and Utilization pp W1-W17

DEVELOPMENT AND CHARACTERISATION OF OIL PALM (*Elaeis Guineensis* Jacq.) MICROSATELLITE MARKERS.

N BILLOTTE[1], AM RISTERUCCI[1], E BARCELOS[2], JL NOYER[1], P AMBLARD[1], FC BAURENS[3].

ABSTRACT

A large Elaeis guineensis *microsatellite development program has been undertaken to generate a wide range of (GA)n, (GT)n and (CCG)n Simple Sequence Repeats (SSR) markers useful for oil palm genetic mapping, genetic diversity study and variety identification. Comparative results in building several microsatellite-enriched libraries demonstrate that an optimised hybridization based capture methodology, using biotin-labelled microsatellite oligoprobes and streptavidin-coated magnetic beads, is the quickest and best adapted procedure to ensure a high enrichment rate with minimal redundancy. A set of 21 SSR loci are evaluated with estimates of allele size range and expected heterozygosity in* E. guineensis *as in the close related* E. oleifera *species where an optimal transportability of the SSR markers was observed. SSR mendelian segregation allows oil palm breeders to use SSR markers for genetic mapping and for the search of agronomic genes based on both intra- or inter-specific progenies. Multivariate data analysis show the ability of our developed SSR markers to reveal the genetic diversity structure of* E. guineensis *and* E. oleifera *natural populations, in accordance with already measured relationships and with geographical origins. Very high levels of allelic variability indicate that SSRs will be a powerful tool for variety identification in the* Elaeis *species.*

Keywords: Elaeis guineensis, *oil palm, microsatellite, DNA markers*

---

[1] CIRAD (Centre de coopération Internationale en Recherche Agronomique pour le Développement) TA 80 / 03 Avenue Agropolis 34398 Montpellier cedex 5 FRANCE.

[2] EMBRAPA Amazonia occidental Rod. AM 010 Km29 AM Manaus BRASIL CEP 69-011-970

[3] VITROPIC SA ZAE Les Avants F-34270 Saint Mathieu de Treviers FRANCE.

e-mail : billotte@cirad.fr Tel : (33) 0467615452 Fax : (33) 0467617183

MANUSCRIPT

N. BILLOTTE
tel : (33) 04 67 61 54 52
fax: (33) 04 67 61 71 83
E-mail: billotte@cirad.fr

Manuscript includes 1 figure and 2 tables

INTRODUCTION

Nowadays, oil palm (*E. guineensis* Jacq.) is grown primarily in Asia, following its introduction in Indonesia in the 19th century, and represents the second largest contribution to human vegetal oil consumption after soybean (Mielke, 1996). Using reciprocal recurrent selection, the purpose of the oil palm breeding schemes undertaken world-wide is to supply improved seeds whose oil productivity is greatest by virtue of the larger amount of mesocarp in their fruits (Meunier and Gascon, 1972). The related American wild species, *E. oleifera*, is the single species belonging to the same *Elaeis* genus. The inter-fertility between *E. guineensis* and *E. oleifera* enabled inter-specific hybridisation programmes for the introgression of interesting *E. oleifera* characters (resistance to the bud rot disease, slow vertical growth, oil fluidity) into *E. guineensis* (Le Guen *et al.*, 1991). Any oil palm genetic trial involves a large number of palms, a vast planting area, and high observation and up-keeping costs for more than a decade. Molecular markers bring a better knowledge of the genetic structure of natural populations (Barcelos *et al.*, 1999; Ghesquière, 1983), the control of the identity and of the heterozygosity of selected progenies, studies of the genetic relationships between selected characters and, moreover, a obvious tool for early selection (Jack and Mayes, 1993; Jones, 1989; Mayes *et al.*, 1996; Shah *et al.*, 1994). The RFLP (Restricted Fragment Length Polymorphism) technique has been chosen for the first genetic mapping in oil palm by Mayes *et al.* (1997), due to the powerfulness and to the reproducibility of its diagnostics (Botstein *et al.*, 1980). But a routine use of the RFLP technique, costly in time and manpower, does not enable to analyse large number of individuals (Jack and Mayes, 1993). Simple Sequence Repeats (SSRs), also called microsatellites, are tandemly arrays of simple nucleotide motifs that are ubiquitous components of eucaryotic genomes (Delseny *et al.*, 1983; Tautz and Renz, 1984). Inherited in a mendelian fashion (Saghai Maroof *et al.*, 1994; Weissenbach et al, 1992), their hypervariable length polymorphism is simply revealed by Polymerase Chain Reaction (PCR) using flanking primers that generate co-dominant markers. Smith *et al.* (1997) resumed the SSR technology presenting the potential advantages of reliability, reproducibility, discrimination, standardisation and cost effectiveness. The main objectives of this study are to: 1/ develop first *E. guineensis* SSR markers; 2/characterize the SSR length variability in *E. guineensis* and in the closely related *E. oleifera* ; 3/evaluate the suitability of SSR markers for genetic diversity and variety identification studies in the *Elaeis* genus.

MATERIAL AND METHODS

Plant material and DNA extraction

All genomic DNA were extracted from freeze-dried leaf samples (Gawel and Jarret; 1991), and then purified on anion exchange micro-columns. Oil palm materials used for the investigation of the polymorphism detected by SSR markers included 18 accessions representing the *E. guineensis* genetic variability according to agro-morphological characters (Meunier and Gascon, 1972) and molecular data (Ghesquière, 1983; Barcelos, 1998). Twenty-one prospected accessions, representing the *E. oleifera* genetic variability revealed by Barcelos (1998) with RFLP and AFLP markers, were chosen to study the transportability and the polymorphism of oil palm SSR markers on that close related species. The plant material sample included 39 individuals for the 2 *Elaeis* species (table 1).

Di- and tri-nucleotide microsatellites sampling

A single oil palm genomic DNA sample (LM2T genitor), was used to build five *E. guineensis* microsatellite enriched-libraries. Two (GA)n and one (GT)n microsatellite-enriched libraries were constructed according to Billotte *et al.* (1999), following a hybridisation-based capture methodology using biotin-labelled microsatellite oligoprobes and streptavidin-coated magnetic beads. One (GT)n and one (CCG)n microsatellite-enriched libraries were constructed following the membrane-bound oligonucleotide enrichment technique of Edwards *et al.* (1996). Clones were obtained after transformation of the selected DNA fragments into *Escherichia coli* XL1 Blue strain (Stratagene, USA). Overall percentages of positive clones containing a microsatellite locus were detected by hybridisation with microsatellite probes (Billotte *et al.*, 1999). Sequencing was performed on a sample of 166 microsatellite-containing sequences: 50 per (GA)n or (CCG)n library, and 8 per (GT)n library. Sequence data analysis was performed using the Sequence Navigator software (Applied Biosystems). For a given sequence of a library, the microsatellite length was defined as the longer run of uninterrupted repeats. The total number of targeted or alternative repeats was taken into account to sample sequences for designing PCR primers (cut off of at least 8 and 5 repeats respectively for di- or tri -motifs).

The required minimum length of each microsatellite flanking region was 20 bp. A sub-set of 34 non-redundant sequences was chosen for PCR primer design.

PCR primer design

PCR primer pairs for microsatellite amplification were designed using the primer analysis software 'OLIGO' 4.06 (National Biosciences, USA) and synthesized commercially (Eurogentec, Belgium). The major criteria for the primer design procedure were to produce 17-30 nucleotides long well-matched primers characterised by a 3'-end decreasing internal stability with a 3'-end pentamer (G higher than -10 kcal/mol, with an average GC content ranging between 40% and 60%, and PCR ampliflying fragments ranging between 100 and 300 bp. Each sequence was subjected to one single PCR primers design.

PCR amplification of SSR loci

PCR amplifications on the sample were performed using 50 ng of genomic DNA. Each primer pair was firstly tested on the same oil palm DNA extract from LM2T. PCR products were checked by agarose gel electrophoresis. PCR amplification was performed in a final volume of 25 µl with 0.2 µ*M* of each primer, 200 µ*M* of dNTP and 1 unit of Taq DNA polymerase (GibcoBRL) in a final buffer containing 10 m*M* TrisHCl pH 8, 100 m*M* KCl, 0.05% w/v gelatin and 1.5 m*M* MgCl2, and following the PCR program: denaturation at 95°C for 1 min and 35 cycles of [94°C for 30 s, 52°C for 60 s, 72°C for 120 s] and a final elongation step at 72°C for 8 min. A range of annealing temperatures (Ta) was tested for 2°C by 2°C from 46°C to 60°C. Different $MgCl_2$ concentrations of 1.5, 2.0 and 2.5 m*M* were tested when necessary to increase the annealing specificity of the primers and to avoid the appearance of unexpected PCR products.

The polymorphism study of selected primer pairs was based on PCR amplifications performed as previously described, but with 0.2 µ*M* of each 5'-end (^{33}P radio-labelled primer using the best annealing temperature. A volume of 25 µl of 80% formamide containing 0.5% (w/v) bromophenol blue and xylene cyanol was mixed to the PCR product. The mix was denatured at 94°C for 5 min, kept at 70°C during gel loading, and then separated on a 5% denaturing

polyacrylamid gel containing in 8*M* urea and 0.5XTBE buffer at 55W for 1 h to 2.5 h. The gel was dried for 20 min in a Bio Rad gel dryer (model 583) and then exposed to KODAK Omat-XLS X-ray film for 4 days. Patterns were sized by comparison to a 10-bp repeats standard DNA ladder (GibcoBRL).

Data analysis

The number of alleles and the polymorphism level of SSR markers were estimated in the two *Elaeis* species, separately, then on combined data. The polymorphism level of each SSR locus was evaluated by the calculation of the expected heterozygosity (He) according to Nei (1973):

$$He = 1 - \sum_{i=1}^{i=n} f_i^2$$ where f_i is the frequency of the allele i.

Probability for identical genotypes (PI value) was estimated according to Paetkou et al. (1995):

$$PI = \sum_{i=1} f_i^4 + \sum_{i=1}^{i=n-1} \sum_{j=i+1}^{n} (2(f_i f_j)^2)$$ where f_i and f_j are the frequencies of the alleles i and j.

Factorial Analyses of Correspondences (Benzecri, 1973) were performed, using the 'ADDAD' software (Lebeaux, 1985). Based on a χ^2 distance, FAC identifies several independent axes or eigenvectors that are linear combinations of the studied characters (SSRs here) and that accounted for the largest part of the variation (SSR length polymorphism). Each band level was scored as '1' in case of heterozygosity and '2' in case of homozygosity for the *E. guineensis* or *E. oleifera* individuals. A null allele was added to the data set as absence of any PCR product was considered as due to a genetic cause. In all other situations, data were considered as missing. Individuals showing missing data were considered as non-active objects for the FAC calculation.

RESULTS

E. guineensis *microsatellite-enriched libraries*

A total of 1880 clones were obtained after transformation of the selected DNA fragments. The analysis of 1341 randomly picked clones showed an insert size ranging from 300 to 1200 bp. The overall percentage of positive clones was high with a mean value of 72% and rather limited variations depending on the libraries or on enrichment methodologies using magnetic beads (72%) or membranes (68%). About 60% of the positive clones were unique according to sequencing data of a subset from each library. That percentage seemed equivalent for clones captured with the magnetic beads procedure (73%) or selected by membranes (58%). The (GA)n *Pst*I digested enriched-library gave the highest percentage of unique positive sequences, with 80% against 40% and 54% for the two other best characterised (GA)n sonicated and (CCG)n *Rsa*I digested enriched-libraries. The 166 sequenced clones revealed a single false positive belonging to the (CCG)n library. All (CCG)n microsatellites were perfect SSRs as defined by Weber (1990). Predominant perfect repeats (62%) were observed in the (GT)n libraries, as in the (GA)n libraries (83%). Other SSRs were compound one.

PCR primer pairs testing on E. guineensis *and on* E. oleifera

Ninety-nine positive non-redundant sequences that were analysed contained 52% of potential SSRs with an adequate size and well located inside the clone insert. The (GA)n *Pst*I digested library gave a percentage of 72%, compared to 34% and 26% respectively for the (GA)n sonicated and (CCG)n libraries. A total of 28 primer pairs were designed and tested on the *E. guineensis* LM2T genitor. Expected PCR products were obtained for 22 of them. On the other hand, the 22 primer pairs generating amplification in *E. guineensis* also exhibited clear PCR products in *E. oleifera*. Patterns on polyacrylamid gels were all similar for the two species. One primer pair, designed from a clone containing a (GT)n microsatellite sequence, was totally monomorphic in both *Elaeis* species (data not shown). We finally measured that about 62% of the potential (GA)n, (GT)n or (CCG)n non-redundant sequences would yield a polymorphic SSR locus in one or both *Elaeis* species.

SSR length polymorphism in E. guineensis *and in* E. oleifera.

Eighteen primer pairs generated polymorphic PCR products in *E. guineensis* and twenty in *E. oleifera* (table 2). Based on amplification pattern analysis, all polymorphic SSRs seemed to be single-locus, with no more than two displayed bands, except one showing a putative double-loci pattern (mEgCIR0134). Their concerned ten (GA)n SSRs (17±3 repeats), seven (GT)n SSRs (10±6 repeats) and four (CCG)n SSRs (6±1 repeats). The total number of alleles per SSR single-locus reached an average value of 8 with about 5 alleles per locus in each *Elaeis* species. About 70% of these alleles were specific to one or to the other species (about 3 alleles per single-locus). Another information was that putative null alleles were detected for two loci in *E. oleifera* individuals. Expected heterozygosities (He values) informed upon the high average polymorphism rates of our oil palm SSR markers: 0.68 in *E. guineensis* and 0.63 in *E. oleifera*, considering SSR single-loci. Combined estimates for the two species gave a 0.69 average He value, with ten loci equalling or above 0.80. Low probabilities of identities (PI) were in accordance with negatively correlated expected heterozygosities. The 6 most informative loci enabled to discriminate unrelated individuals with very low probabilities for identical genotypes: 7.2×10^{-10} in *E. guineensis* and 3.2×10^{-10} in *E. oleifera*.

Genetic diversity structure of the Elaeis genus

The FAC was performed from 20 single-locus SSR markers totalling 167 alleles. Distinct sub-groups were discriminated on the plan 1-2, which explained 22.3 % of the total molecular variation (fig. 1). One is exclusively formed by all close-spotted *E. guineensis* individuals that are opposed, along the first axis, to 4 distinguishable *E. oleifera* groups issued from 4 different geographic areas: Brazilian, Central American, Peruvian, and particularly isolated on the representation, French Guiana or Surinam.

DISCUSSION

Our study showed that the two enrichment procedures we used gave equivalent results in terms of percentage of positive or unique clones. However, for building up enriched-microsatellite libraries, major arguments favour a methodology using biotin-labelled oligoprobes and

streptavidin-coated magnetic beads. The enrichment step itself took just 2 hours using that methodology, compared to 2-3 weeks with the membrane-bound oligonucleotide enrichment technique, and 5 days only from the genomic DNA shearing to the first results of the hybridisation step with microsatellite probes. A second point is that the method should be also suitable for the enrichment of all types of SSR by simply switching microsatellite oligoprobes. Finally, the redundancy of clones, a major problem generally not reported in the literature, can be theoretically reduced to single sequences in the library by decreasing the number of PCR cycles (Billotte *et al.*, 1999). Such optimised protocol gives a technical basis for an easy microsatellite marker development not only in oil palm but also in any tropical crops.

All PCR primer pairs generating amplification in *E. guineensis* exhibited clear PCR products in *E. oleifera*. High levels of length polymorphism in that closely related species confirmed the optimal transportability of *E. guineensis* SSR markers to *E. oleifera*. The average SSR polymorphism level in the *Elaeis* genus is in the 0.40-0.80 range commonly reported in plants by various authors (Bryan *et al.*, 1997; Diwan and Cregan, 1997; Grapin *et al.*, 1998; Ridout and Donini, 1999; Smith *et al.*, 1997). The use of SSR co-dominant markers presenting a multi-allelic richness will greatly facilitate the building up of both intra- or inter-specific genetic maps, and increase genes and Quantitative Trait Loci (QTL) detection power (Muranti, 1997). SSR multi-loci could increase their informative power especially if they can reveal independent loci, a point that will have to be checked and for mEgCIR0134.

Barcelos *et al.* (1999) observed the high molecular differentiation between *E. guineensis* and *E. oleifera* as well as the narrower genetic basis of *E. guineensis*, at the nuclear and mitochondrial genomes with AFLP and RFLP markers. Despite a limited number of palms, our FAC on both *Elaeis* species revealed a genetic diversity that was in accordance with these results. Both data revealed the low genetic diversity of *E. guineensis* that Barcelos *et al.* (1999) explained by a continuous distribution in the African continent due to an absence of geographical gene flow barrier and to the material dispersion by human for this spécies. The same authors evidenced a discrimination of *E. oleifera* accessions from French Guiana, Surinam and Peru due to specific alleles supporting a hypothesis of drift and recent bottleneck events experienced by these groups despite the continuous *E. oleifera* distribution in Amazonia from Peru to Brazil. The projection of the SSRs alleles on the 1-2 plane of our FAC indicated similar results with 10

specific alleles (from 9 SSRs) mostly responsible of the particular location of the *E. oleifera* French Guiana or Surinam individuals (data not shown).

CONCLUSION AND PROSPECTS

The relatively simple laboratory techniques based on PCR and the non-radioactive detection of their polymorphism (Grapin *et al.*, 1998) makes SSRs a tool of choice for small-sized laboratory units in oversee oil palm research stations. Our *E. guineensis* microsatellite markers are especially bound to genetic mapping, diversity studies and variety identification. We observed the high polymorphism and discrimination capability of the SSRs in *E. guineensis*, but also in the closely related *E.oleifera* to which all markers were transferable. Developed SSRs are actually used for *E. guineensis* variety checking. A large number of high informative *E. guineensis* SSRs are being developed for the building up of a microsatellite-based reference linkage map that will be a stone towards oil palm QTL search and marker-assisted breeding. Both common and specific SSR alleles of *E. guineensis* and *E. oleifera* enable a fast building up of inter-specific genetic maps supporting inter-specific hybridisation programmes. Consequences of SSRs transportability will be of great importance to help the introgression of favourable agronomic characters from one species to the other. In a short term, the use of SSRs will drastically contribute to valorise prospective genetic collections of the wild *E. oleifera* species and inter-specific back-crosses already planted and studied in oil palm research stations.

ACKNOWLEDGEMENTS

We gratefully acknowledge our partners for providing *Elaeis* samples i.e. the CNRA oil palm research station in La Mé Côte d'Ivoire and the EMBRAPA research station in Rio Urubu Brazil. We are very grateful to the Genoscope Centre in Evry France and to its general director, Dr J. Weissenbach, for the mass sequencing of CIRAD's oil palm SSR-containing clones.

REFERENCES

Barcelos, E. 1998. Etude de la diversité génétique du genre *Elaeis* (*E. oleifera* (Kunth) Cortés et *E. guineensis* Jacq.) par marqueurs moléculaires (RFLP et AFLP). Thèse. Université Montpellier II France. pp 137.

Barcelos, E., Second, S., Kahn, F., Amblard, P., Lebrun, P. and Seguin, M. 1999. *In* Evolution, variation and classification of palms. Memoirs 83 of the New York Botanical Garden. Editors : A. Henderson and F. Borchsenius. The New York Botanical Garden Press. New York. pp. 191-201.

Benzecri, J.P. 1973. L'analyse des correspondances. Vol. 2. Editor: Dunod. Paris. 1973.

Billotte, N., Lagoda, P.J.L., Risterucci, A.M., and Baurens, F.C. 1999. Microsatellite-enriched libraries: applied methodology for the development of SSR markers in tropical crops. Fruits. **54**: 277-288.

Botstein, O., White, R.L., Skolnick, M., Davis, R.V. 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphism. Am. J. Hum. Gene. 32: 314-331.

Bryan, G.J., Collins, A.J., Stephenson, P., Orry A., Smith J.B., and Gale, M.D. 1997. Isolation and characterisation of microsatellites from hexaploid bread wheat. Theor. Appl. Genet. **94**: 557-563.

Delseny, M., Laroche, M., Penon, P., 1983. Detection of sequences with Z-DNA forming potential in higher plants. Biochem. Biophys. Res. Commun. **116**: 113:120.

Diwan, N., and Cregan, P.B. 1997. Automated sizing of fluorescent-labeled simple sequence repeat (SSR) markers to assay genetic variation in soybean. Theor. Appl. Genet. **95**: 723-733.

Edwards, K.J., Barker, J.H.A., Daly, A., Jones, C., and Karp, A. 1996. Microsatellite libraries enriched for several microsatellite sequences in plants. BioTechniques, **20**: 758-760.

Gawel, N.J., and Jarret, R.L. 1991. A modified CTAB DNA extraction procedure for *Musa* and *Ipomoea*. Molecular Biology Reporter. **9**: 262-266.

Grapin, A., Noyer, J.L., Carreel, F., Dambier, D., Baurens, F.C., Lanaud, C., Lagoda, P.J.L. 1998. Diploid *Musa Accuminata* genetic diversity assayed with sequence-tagged microsatellite sites. Electrophoresis. **19**: 1374-1380.

Guesquière, M. 1983. Contribution à l'étude de la variabilité génétique du palmier à huile (*Elaeis guineensis* Jacq.). Le polymorphisme enzymatique. Thèse. Université Paris-Sud. Centre d'Orsay. pp 116.

Jack, P.L. and Mayes, S. 1993. Use of molecular markers for oil palm breeding. II. Use of DNA markers (RFLPs). Oléagineux. **48**: 1-8.

Jones, L.H. 1989. Prospects for biotechnology in oil palm (*Elaeis guineensis*) and coconut (*Cocos nucifera*) improvement. Biotechnology and Genetic Engineering Reviews. **7**: 281-296.

Lebeaux, M.O. 1985. ADDAD, *Manuel de référence Version 89.1.*, Association pour le Développement et la Diffusion de l'Analyse des Données, Paris.

Le Guen, V., Amblard, P., Omore, A., Koutou, A., and Meunier, J. 1991. Le programme hybride interspécifique *Elaeis oleifera* X *Elaeis guineensis* de l'IRHO. Oléagineux **46**: 479-487.

Mayes, S., Jack, P.L., Marshall, D.F., and Corley, R.H.V. 1997. Construction of a RFLP genetic linkage map for oil palm (*Elaeis guineensis* Jacq.). Genome. **40**: 116-122.

Mayes, S., James, C., Price, Z., Groves, L.LM., Jack, P.L., and Corley R.H.V. 1996. Integration of DNA markers into oil palm breeding programmes. Proceedings 1996 International Palm Oil Conference. Malaysia. pp. 49-54.

Meunier, J., and Gascon, J.P. 1972. Le schéma général d'amélioration du palmier à huile à l'I.R.H.O.. Oléagineux. **27**: 1-12.

Mielke, T. 1996. Outlook for oils and fats from 1996 on wards with special emphasis on oil palm and kernel oil. PAC Seminars of PORIM Malaysia. 28th March 1996. pp. 25.

Muranti, H. 1997. Power of tests for quantitative trait loci detection using full-sib families in different schemes. Heredity. **76**: 156-165.

Nei, M. 1973. Analysis of gene diversity in subdivided populations. P. Natl. Acad. Sci. USA. 70:3321-3323.

Paetkou, D., Calvert, W., Stirling, I., and Strobeck, C. 1995. Microsatellite analysis of population structure in Canadian polar bears. Mol. Ecol. **4**: 347-354.

Ridout, C.J., and Domini, P. 1999. Use of AFLP in cereals research. Trends in Plant Sci., **4**: 76-79.

Saghai Maroof, M.A., Biyashev R.M., Yang G.P., Zhang Q., and Allard, R.W. 1994. Extraordinarily polymorphic microsatellite DNA in barley: Species diversity, chromosomal locations, and population dynamics. Proc. Natl. Acad. Sci. USA. **91**: 5466-5470.

Shah, F.H., Rashid, O., Simons, A.J., and Dunsdon, A., 1994. The utility of RAPD markers for the determination of genetic variation in the oil palm (*Elaeis guineensis*) . Theor. Apppl. Genet. **89**: 713-718.

Smith, J.S.C., Chin, E.C.L., Shu, H., Smith, O.S., Wall, S.J., Senior, M.L., Mitchell, S.E., Kresovich, S., and Ziegle, J. 1997. An evaluation of the utility of SSR loci as molecular markers in maize (*Zea mays* L.), comparisons with data from RFLPs and pedigree. Theor. Appl. Genet. **95**: 163-173.

Tautz, D., and Renz, M. 1984. Simple sequences are ubiquitous repetitive components of eucaryotic genomes. Nucleic Acids Res. **12**: 4127-4138.

Weber, J.L. 1990. Informativeness of human (dC-dA)n (dG-dT)n polymorphisms. Genomics. 7: 524-530.

Weissenbach, J., Gyapay, G., Dib, C., Vignal, A., Morissette, J., Millaseau, P., Vaysseix, G. and Lathrop, M. 1992. A second generation map of the human genome. Nature. **359**: 794-801.

Table 1. List of accessions chosen for the study of oil palm SSRs in *Elaeis guineensis* and in *Elaeis oleifera*.

DNA number	Species Name	Origin Country	Population	Fruit type	Number of individuals	Collection
543, 566	*Elaeis guineensis*	Zaïre	Sibiti	*tenera, dura*	2	Brazil/RIO URUBU
562	*Elaeis guineensis*	Zaïre	Yangambi	*tenera*	1	Côte d'Ivoire/LAME
560, 561	*Elaeis guineensis*	Côte d'Ivoire	La Mé	*tenera, tenera*	2	Côte d'Ivoire/LAME
531	*Elaeis guineensis*	Côte d'Ivoire	Yocoboué	*dura*	1	Côte d'Ivoire/LAME
555	*Elaeis guineensis*	Angola	Salazar	*dura*	1	Côte d'Ivoire/LAME
533	*Elaeis guineensis*	Nigeria	Aba/Calabar	*dura*	1	Brazil/RIO URUBU
549	*Elaeis guineensis*	Nigeria	Abak	*tenera*	1	Côte d'Ivoire/LAME
551	*Elaeis guineensis*	Nigeria	Uli	*tenera*	1	Côte d'Ivoire/LAME
552	*Elaeis guineensis*	Nigeria	Ahoada	*dura*	1	Côte d'Ivoire/LAME
553	*Elaeis guineensis*	Nigeria	Ayangba	*tenera*	1	Côte d'Ivoire/LAME
547	*Elaeis guineensis*	- unknown -	Brazil/Bahia	*dura*	1	Brazil/RIO URUBU
535	*Elaeis guineensis*	Cameroon	Widikoun	*dura*	1	Brazil/RIO URUBU
537	*Elaeis guineensis*	Cameroon	Lobé	*tenera*	1	Brazil/RIO URUBU
542	*Elaeis guineensis*	Cameroon	ASD	*pisifera*	1	Brazil/RIO URUBU
530	*Elaeis guineensis*	Benin	Pobé	*dura*	1	Brazil/RIO URUBU
541	*Elaeis guineensis*	Deli	Deli Serdang	*dura*	1	Brazil/RIO URUBU
Sub-total					18	
286	*Elaeis oleifera*	Brazil	Manaus-Careiro/Ig Japones	*dura*	1	Brazil/RIO URUBU
300	*Elaeis oleifera*	Brazil	Manaus-Careiro/Olaria	*dura*	1	Brazil/RIO URUBU
311	*Elaeis oleifera*	Brazil	Rio Madeira/Manicoré	*dura*	1	Brazil/RIO URUBU
314	*Elaeis oleifera*	Brazil	Rio Madeira/Libertade	*dura*	1	Brazil/RIO URUBU
396, 402	*Elaeis oleifera*	Brazil	BR 174	*dura, dura*	2	Brazil/RIO URUBU
427	*Elaeis oleifera*	Brazil	Rio Negro/Ig Arraia	*dura*	1	Brazil/RIO URUBU
434	*Elaeis oleifera*	Brazil	Rio Negro/Acajatuba	*dura*	1	Brazil/RIO URUBU
439	*Elaeis oleifera*	Brazil	Rio Solimoes/Ig Vendaval	*dura*	1	Brazil/RIO URUBU
448	*Elaeis oleifera*	Brazil	Rio Solimoes/Catua	*dura*	1	Brazil/RIO URUBU
455	*Elaeis oleifera*	French Guiana	Mana	*dura*	1	-
474	*Elaeis oleifera*	Colombia	Sinu	*dura*	1	Côte d'Ivoire/LAME
484	*Elaeis oleifera*	Colombia	Aguas Vivas	*dura*	1	Côte d'Ivoire/LAME
472	*Elaeis oleifera*	Colombia	San Alberto	*dura*	1	Côte d'Ivoire/LAME
488	*Elaeis oleifera*	Costa Rica	Golfito	*dura*	1	Côte d'Ivoire/LAME
498	*Elaeis oleifera*	Nicaragua	El Castillo	*dura*	1	Côte d'Ivoire/LAME
501	*Elaeis oleifera*	Panama	Playa Honda	*dura*	1	Côte d'Ivoire/LAME
515, 517	*Elaeis oleifera*	Peru	Genaro Herrera	*dura, dura*	2	Côte d'Ivoire/LAME
521	*Elaeis oleifera*	Surinam	Colakkreek	*dura*	1	Côte d'Ivoire/LAME
524	*Elaeis oleifera*	Surinam	Marihat	*dura*	1	Indonesia/MARIHAT
Sub-total					21	
Grand Total					**39**	

Table 2. Type of repeats, allele size range, allele numbers, expected heterozygosity and probability of identity of 21 *E. guineensis* SSR loci in *E. guineensis* and *E. oleifera*.

a single-locus SSR b putative double-loci SSR * including one putative null allele with an absence of band. £ = calculated on 19 individuals instead of 21 due to missing data for the 2 palms n°448 and 314. h = homozygous locus

SRR locus	EMBL Accession N°	Motif	Ta (°C)	MgCl2 (mM)	Allele size range (bp)	# Alleles *E. guineensis*	# Alleles *E. oleifera*	# Shared alleles	Total # alleles	Expected heterozygosity (He) *E. guineensis*	*E. oleifera*	Combined	Probability of identity (PI) *E. guineensis*	*E. oleifera*
a														
mEgCIR0008	AJ271625	(GA)18	52	1.5	195-220	6	11	6	11	0.73	0.87	0.84	0.07	0.02
mEgCIR0009	AJ271633	(GA)20	52	1.5	162-204	4	7	2	9	0.64	0.82	0.80	0.13	0.03
mEgCIR0018	AJ271634	(GA)18	52	2.5	158-177	11	3*	2	12	0.76	0.56	0.81	0.06	0.19
mEgCIR0046	AJ271635	(GA)19	52	1.5	198-262	5	9	1	13	0,79 h	0.86	0.90	0,04 h	0.02
mEgCIR0067	AJ271636	(GA)17	52	1.5	135-187	8	9	4	13	0.78	0,82 £	0.82	0.05	0,03 £
mEgCIR0219	AJ271637	(GA)17	56	1.5	205-233	7	6	4	9	0.71	0.71	0.83	0.08	0.08
mEgCIR0221	AJ271638	(GA)11	52	1.5	195-213	7	3*	1	9	0.61	0,38 h	0.74	0.15	0,38 h
mEgCIR0230	AJ271639	(TA)6GAG(GA)19	52	1.5	326-354	6	8	3	11	0.76	0.82	0.86	0.06	0.03
mEgCIR0254	AJ271640	(GA)18	52	1.5	148-179	10	10	4	16	0.86	0.75	0.83	0.02	0.06
mEgCIR0304	AJ271931	(GT)4(GC)2(GT)2	52	1.5	106-114	1	3	1	3	0.00 h	0.22	0.12	1,00 h	0.61
mEgCIR0326	AJ271932	(GT)9	52	1.5	275-281	1	3	0	4	0.00 h	0.57	0.41	1,00 h	0.19
mEgCIR0332	AJ271933	(GT)8	52	1.5	131-147	4	1	1	4	0.59	0.00 h	0.59	0.17	1,00 h
mEgCIR0350	AJ271934	(GT)8	52	1.5	269-281	4	3	2	5	0.45	0,25 h	0.63	0.30	0,56 h
mEgCIR0353	AJ271935	(GT)11(GA)15	52	1.5	80-102	7	6	1	12	0.83	0.70	0.86	0.03	0.09
mEgCIR0377	AJ271936	(GT)6(GC)4	52	1.5	145-149	1	3	1	3	0.00 h	0,57 h	0.38	1,00 h	0,18 h
mEgCIR1772	AJ271937	(GT)22	52	1.5	166-198	12	8	5	15	0.87	0.84	0.89	0.02	0.03
mEgCIR0391	AJ271938	(CCG)5	52	1.5	293-297	1	4	1	4	0.00 h	0.65	0.44	1,00 h	0.12
mEgCIR0437	AJ271939	(CCG)6	58	1.5	196-206	3	3	2	4	0.48	0.55	0.64	0.28	0.20
mEgCIR0465	AJ271940	(CCG)6	52	1.5	125-137	4	3	2	5	0.67	0.54	0.74	0.11	0.21
mEgCIR0476	AJ271941	(CCG)7	56	1.5	165-177	3	4	2	5	0.36	0.48	0.67	0.40	0.23
Average						**5.25**	**5.35**	**2.25**	**8.35**	**0.68**	**0.63**	**0.69**	**0.30**	**0.21**
Standard-error						3.22	2.87	1.54	4.21	0.15	0.19	0.20	0.36	0.25
b														
mEgCIR0134	AJ271714	(GA)15	52	1.5	128-162	10	5	2	13	0.82	0.62	0.84	0.03	0.15
					173-182	5	5	5	5	0.10	0.50	0.55	0.80	0.25
Polymorphic markers (%)										82	95	100		

Fig.1. Factorial Analysis of Correspondances performed on 20 single-locus microsatellite markers over 18 accessions of *E. guineensis* (Africa) and 21accessions of *E. oleifera* (Brazil, Central America French Guiana, Peru and Surinam). Axis 1 and 2 represent 22.3 % of the total molecular variability.