Euphytica **101:** 103–108, 1998.
© 1998 *Kluwer Academic Publishers. Printed in the Netherlands.*

Genetic diversity in coconut (*Cocos nucifera* L.) revealed by restriction fragment length polymorphism (RFLP) markers

P. Lebrun[1], Y.P. N'cho[2], M. Seguin[1], L. Grivet[1] & L. Baudouin[1]
[1] *CIRAD/CP, UR Amélioration des plantes, BP 5035, 34032, Montpellier Cedex, France;* [2] *IDEFOR/DPO, Station Marc Delorme, 07 BP 13, Abidjan 07, Ivory Coast*

Received 17 December 1996; accepted 30 September 1997

Key words: Cocos nucifera, RFLP, genetic diversity

Summary

A study of the genetic diversity in coconut by RFLP analysis was performed in 100 individuals representing 10 Tall and seven Dwarf local populations or 'ecotypes' from various geographical origins. Nine cDNA clones from rice, one mitochondrial DNA clone (*CoxI*) and one genomic clone (rDNA) from wheat were used as probe for southern hybridization. The distribution of the 40 polymorphic bands revealed by rice cDNA clones was studied using a multivariate analysis and allowed to identify two main genetical groups. The first one includes the ecotypes from the Far East and from the South Pacific, whereas the other one comprises the ecotypes from India, Sri Lanka and Western Africa. The *rDNA* and the *CoxI* probes confirm this distinction. The Far East and the Pacific regions which were the most likely center of origin also exhibit the widest polymorphism. The associations between the Panama Tall and the Pacific group and between the West African Tall and the Indian Ocean group reflect their likely origin. The Comoro Tall appears to be intermediate between the two main groups, and could reflect the old migration route between Indonesia and Madagascar. All Dwarf varieties belong to the first group, even those collected in West Africa. Those were probably introduced from Asia and Pacific at the beginning of the century. The cross-pollinating Tall ecotypes were generally more polymorphic than the self pollinating Dwarf ecotypes. The legitimacy of two hybrids between ecotypes was confirmed and maternal inheritance of mitochondrial genome was observed.

Introduction

The coconut palm is a perennial oilseed plant found throughout the intertropical zone. Its probable origin is in the Southeast Asia/West Pacific region if the considerable morphological variability, number of local names and many uses made of the plant in that region are anything to go by (Persley, 1992). The species would seem to have been disseminated with human migration and seaborne nuts floating eastward in the Pacific Ocean and westward in the Indian Ocean. Its presence on the Atlantic coasts of Africa and America seems to be a recent introduction by European navigators.

Traditional coconut cultivars are population varieties known as ecotypes. There are two major types, Talls and Dwarfs, which differ in numerous characters. The Talls have strong vertical growth, high productivity, low precocity and are preferentially cross-fertilizing, whereas the Dwarfs are shorter, less productive, more precocious and generally self-fertilizing (Taffin, 1993). Breeding programmes are currently geared towards the creation of Tall $\times$ Tall and Tall $\times$ Dwarf hybrid varieties (hybrids between ecotypes) to combine precocity and productivity (Bourdeix, 1989).

A study of the genetic diversity for ecotype is of particular importance for breeding, especially in the context of hybrid variety creation. Indeed, it should make it possible to structure ecotypes into genetic groups, thus identifying redundant ecotypes in collections and rationalize the choice of crosses to be tested.

Beside the Tall-Dwarf duality, there is considerable morphological variability between ecotypes including the characteristics of the fruit and vegetative organs. This variability is expressed in the size, shape and colour of the fruit and has been used to propose a

Table 1. Coconut germplasm under study

Ecotype code	Type	Place of origin	Geographic group	Number of individuals studied
PNT	Tall	Panama	America	2
CMT	Tall	Comoro Isl.	Indian Ocean	5
LMT	Tall	Laccadives Isl. India	Indian Ocean	5
SLT	Tall	Sri Lanka	Indian Ocean	5
MLT	Tall	Malaysia	Far East	7
TAGT	Tall	Tagnanan, Philippines	Far East	2
WAT4	Tall	Côte d'Ivoire	Africa	15
WAT6	Tall	Benin	Africa	10
RIT	Tall	Rennell Isl. (Solomon Isl.)	Pacific	7
SIT	Tall	Solomons Isl.	Pacific	7
MYD	Yellow Dwarf	Malaysia	Far East	15
MRD	Red Dwarf	Malaysia	Far East	2
MGD	Green Dwarf	Malaysia	Far East	2
CATD	Green Dwarf	Tacunan, Philippines	Far East	2
GYD	Yellow Dwarf	Ghana	Africa	7
CRD	Red Dwarf	Cameroun	Africa	5
NLAD	Dwarf	Niu Leka, Fidji	Pacific	2
PB111	Hybrid	CRD × WAT4		7
PB121	Hybrid	MYD × WAT4		7

diversification model for the coconut palm based on a comparison between a wild type (Niu kafa) and a selected type (Niu vai) (Harries, 1978). Nevertheless, the structuring obtained with these types of characteristics is difficult to relate to geographical origin and to the assumed species dissemination scenario (N'Cho et al., 1993). Studying diversity by enzyme electrophoresis came up against technical problems, as numerous systems proved to be monomorphic or not very active. Four systems revealed polymorphism and only two alleles per locus could be detected for each system (Benoit & Ghesquière, 1984). On the other hand, analysis of leaf polyphenol polymorphism provided a picture of variability that matched geographical origins (Jay et al., 1989). This technique seems to be a good tool for studying relations between ecotypes, but the sensitivity of the polyphenol banding patterns to ecological conditions limits its application, particularly for comparing origins from different stations. An overview of the work was presented by Meunier (1992).

In this report we present the first results on coconut germplasm diversity using Restriction Fragment Length Polymorphism.

Materials and methods

Plant material. All the material studied has been collected at the Marc Delorme Station (IDEFOR, Côte d'Ivoire), a breeding station with a major collection comprising some sixty ecotypes from throughout the intertropical zone. In this collection, each ecotype is represented by numerous individuals (forty-four are represented by over a hundred individuals), enabling the analysis of both between- and within-ecotype diversity. One hundred trees samples from 10 Tall and seven Dwarf ecotypes of various geographic origins (Table 1), along with two hybrid populations of seven individuals each, have been studied. The hybrid populations were PB 111 from the cross WAT 4 (female) × CRD and PB121 from the cross WAT 4 (female) × MYD.

DNA extraction. DNA was extracted from the leaflets of frond number one (first open frond) by the CTAB method (Hoisington, 1992). The frehs material sent from Côte d'Ivoire to Montpellier, was freeze-dried and ground to a fine powder, which can be conserved several years when stored at − 20 °C.

RFLP analysis. Taking advantage of the fact that the existence of phylogenetic links between different

Table 2. Probe/enzyme pairs used to reveal polymorphism among coconut accessions

Probe	Type	Restriction enzyme	Number of polymorphic bands
c848	Rice cDNA	*Eco* RI	6
c356	Rice cDNA	*Eco* RI	3
		Bgl II	2
		Sst I	5
c746	Rice cDNA	*Sst* I	4
c496	Rice cDNA	*Eco* RI	5
		Bgl II	3
c975	Rice cDNA	*Bgl* II	3
c147	Rice cDNA	*Eco* RV	2
c285	Rice cDNA	*Sst* I	2
c131	Rice cDNA	*Sst* I	2
c74	Rice cDNA	*Bgl* II	3
Pta71	Wheat genomic rDNA	*Sst* I	2
CoxI	Wheat cytoplasmic	*Bgl* II	2

species is reflected in the conservation of certain genic sequences, rice cDNA clones were used as probe for hybridization on coconut DNA (both plants are monocotyledons). The probes used were chosen dispersed throughout the rice chromosomes. They were kindly provided by the Rice Genome Research Program, NIAR/STAFF, Japan (Kurata et al., 1994). Of the 51 probes tested, six did not hybridize, 32 gave very low intensity bands or a smear, two gave sufficiently intense bands, but revealed no polymorphism, and 12 gave intense bands and revealed polymorphism. Of these 12, nine were used for hybridization on the DNA of the 114 individuals, restricted by enzymes *Bgl*II, *Eco*RI, *Eco*RV or *Sst*I. All of the nine probes revealed complex hybridization patterns, probably corresponding to gene families. Twelve enzyme/probe pairs were studied in all (Table 2). A cytoplasmic DNA clone from wheat, *CoxI* (B. Lejeune et al., 1988) and a genomic clone *Pta71* (Gerlach & Bedbrook, 1979) corresponding to genes *18S-5.8S-25S* of wheat ribosomal RNA were also used as probe.

Each polymorphic band revealed by a rice cDNA was scored 1 for presence and 0 for absence. A multivariate analysis termed Factorial Analysis of Correspondences (FAC) (Benzecri, 1973) was performed on the binary matrix after disjunction of the variables. This enabled us to give the same weight to all the individuals. When two bands revealed by the same probe had the same distribution pattern over all the individuals, they were considered as a single allele and only one was scored. Computation was performed using the ADDAD software (ADDAD, 1983).

Results

The 12 enzyme/probe pairs used to reveal polymorphism with the nine rice cDNA probes led to 40 polymorphic bands being identified. This gives a mean number of 4.4 polymorphic bands per probe. For a given probe, polymorphic bands correspond to a serie of distinct alleles revealed at either one locus or several loci of a same gene family. The cytoplasmic probe *CoxI* and the ribosomal DNA (rDNA) probe *Pta71* each revealed two distinct alleles.

FAC was performed on the 40 polymorphic bands revealed by rice cDNA probes. It led to the identification of two groups along axis one (36.3% of the variability), one (group 1) containing Tall ecotypes originating from the Far East and the Pacific and all the Dwarf ecotypes, and the other (group 2) containing ecotypes from the Indian Ocean and Africa (Figure 1). The ecotype from Comoro Islands (CMT), fell midway between the two groups.

Computation of band frequencies when excluding the individuals from intermediate CMT ecotype showed that among the 40 polymorphic bands, 10 were specific of group 1 and 3 were specific of group 2. Among the 10 specific bands of group 1, two were present in CMT and among the three specific bands of group 2, two were also present in CMT. No band was specific of CMT ecotype.

The two alleles revealed by the rDNA probe corresponded to the two major groups, respectively (Figure 2). Both alleles were present in the individuals of the CMT ecotype.

For the cytoplasmic *CoxI* probe, allele 1 was present in all individuals of group 1 except the two individuals of Dwarf ecotype NLAD which had allele 2. Allele 2 was most frequent in group 2 (Figure 2). The individuals from the CMT ecotype all had the allele 1.

FAC also permitted to reveal a differentiation between the ecotypes from the Far East and those from the Pacific along axis two (9.5% of the variability), which overlapped slightly (Figure 1).

Discussion

With the genetic diversity structuring obtained by RFLP analysis, it was possible to clarify and complete

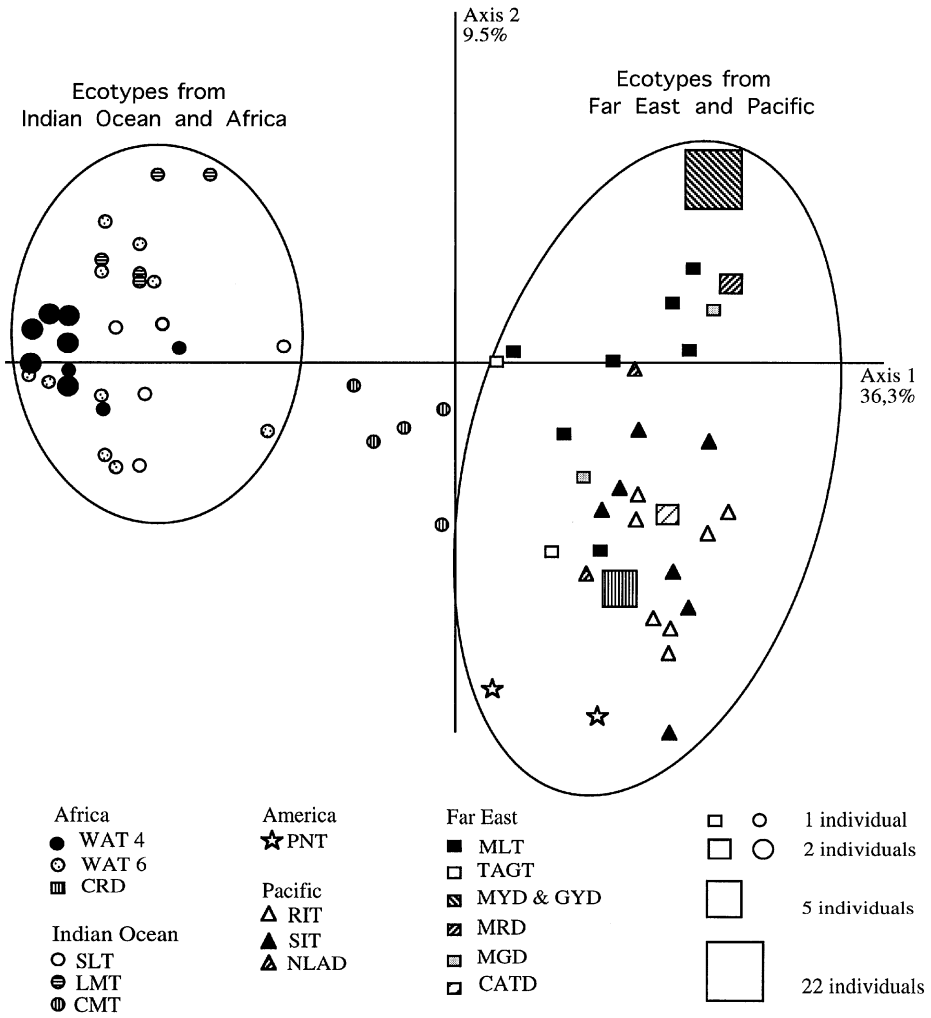

Figure 1. First plane of the FAC performed on 100 individuals of 17 coconut ecotypes characterized by 40 polymorphic bands revealed by nine nuclear rice cDNA probes.

the definition of genetic groups already outlined using polyphenolic markers. The results tallied well with what is already known about the plant's biology. The Tall ecotypes, which are all cross-pollinating, exhibited large amount of polymorphism. Dwarf ecotypes NLAD (cross-pollinating) and MGD (partly cross-pollinating) were represented but only two individuals each, which in both cases had different banding patterns. For Dwarf ecotypes MYD, GYD and CRD, which are completely self-pollinating, the banding pattern for individuals were completely identical inside each ecotype.

The results also tallied with what is known about the historical dispersion of the species. There was substantial between- and within-ecotype variability for the ecotypes collected from the putative area of origin of the coconut palm (Far East and Pacific). The ecotype from the west coast of Panama proved similar to the Pacific ecotypes, which tallies with dissemination of species from the area of origin eastward as far as America (Harries, 1978). The West African ecotypes were related to the Indian and Sri Lankan ecotypes, adding weight to the hypothesis of recent extension of the species along the Atlantic coasts of Africa through nuts originating from the Indian Ocean. The individuals from ecotype CMT (Comoro Islands) had a midway position between groups 1 and 2 on the FAC based on nuclear cDNAs. Two specific bands of group 1 and two specific bands of group 2, were also present in individuals from this ecotype. The two alleles detected for

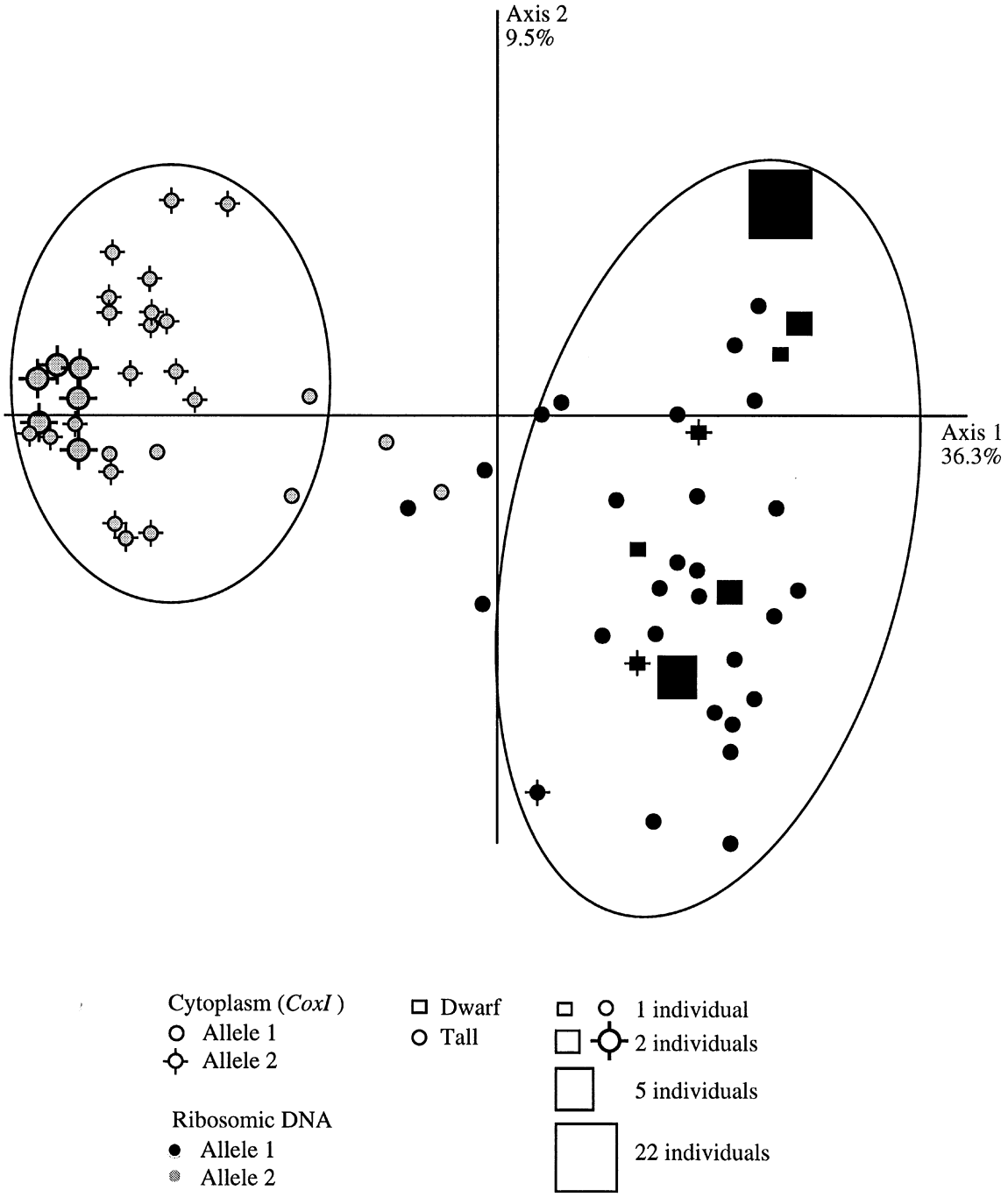

Figure 2. Distribution of *CoxI* and rDNA alleles among coconut individuals represented on the FAC performed with polymorphic bands revealed by rice cDNA clones (see Figure 1).

rDNA were present in this group and all individuals had the cytoplasmic *CoxI* allele 1. This ecotype could thus derive from hybridization between representative individuals of genetic groups 1 and 2. This could be explained by the geographical position of Comoro, at the parting of the ways between the ancient Arabic trading road along East African, Middle East and Indi-

an coasts and the ancient polynesian migration road from Indonesia to Madagascar.

The Dwarf ecotypes collected in West Africa (GYD and CRD) were distributed in genetic group 1 (Far East and Pacific) like all other Dwarfs. Based on historical data, it is highly probable that these ecotypes were introduced at the turn of the century from Southeast Asia (GYD), and from the Pacific (CRD). For GYD, this is backed up by the fact that the RFLP markers for all the individuals were strictly the same as those of the individuals from another Dwarf ecotype, MYD, from Malaysia.

Analysis of the two hybrid cultivars PB 121 and PB 111 confirmed that the alleles of the nuclear genes indeed came from the two assumed parents. This permitted to establish that the mitochondrial DNA was inherited from the mother parent which is usually the case in plants, although there are exceptions (Fauré et al., 1993).

Although the observed structuring was partly detectable with polyphenolic markers, their use remains tricky and RFLP markers appear as a much more convenient tools to study the genetic diversity of coconut. The present study was conducted with only nine RFLP probes but polymorphism was revealed at a larger number of loci since most probes gave complex banding pattern, probably corresponding to gene families. This number of loci was enough to identify the main tendencies in the structuration of the material but a larger number will probably be needed for finer purposes. Studies are now continuing with a view toward improving our knowledge of diversity in the areas where phenotypic variability seems to be at a maximum, and determining the geographical origins of populations located in the most outlying zones.

Acknowledgements

The plant material samples were kindly supplied by IDEFOR-DPO (Côte d'Ivoire), to whom the authors are most grateful. We also thank Dr Y. Nagamura for providing the rice cDNA and Dr C. Kaye and anonymous reviewers for improving our english.

References

ADDAD, 1983. Manuel de référence. Association pour le développement et la diffusion de l'analyse des données, Paris.

Benoit, H. & M. Ghesquière, 1984. Electrophorèse, compte-rendu cocotier. IV. Déterminisme génétique. Internal Report IRHO-CIRAD, France.

Benzecri, J.P., 1973. L'analyse des données. Tome II. 616 p. Dunod, Paris.

Bourdeix, R., 1989. La sélection du cocotier, *Cocos nucifera* L. Etude théorique et pratique, optimisation des stratégies d'amélioration génétiques. PhD Thesis. Université Paris Sud, France.

Fauré, S., J.L. Noyer, F. Carreel, J.P. Horry, F. Bakry & C. Lanaud, 1993. Maternal inheritance of chloroplast genome and paternal inheritance of mitochondrial genome in bananas (*Musa acuminata*). Curr Genet 25: 265–269.

Gerlach, W.R. & J.R. Bedbrook, 1979. Cloning and characterization of ribosomal RNA genes from wheat and barley. Nucl Acids Res 7: 1869–1885.

Harries, H.C., 1978. The evolution, dissemination and classification of *Cocos nucifera* L. Bot Rev 44: 265–320.

Hoisington, D., 1992. Laboratory Protocols. CIMMYT Applied Molecular Genetics Laboratory. D.F. CIMMYT, Mexico.

Jay, M., R. Bourdeix, F. Potier & C. Sanlaville, 1989. Premiers résultats de l'étude des polyphénols foliaires du cocotier. Oléagineux 44: 151–161.

Kurata, N., Y. Nagamura, K. Yamamoto, Y. Harushima, N. Sue, J. Wu, B.A. Antonio, A. Shomura, T. Shimuzu, S-Y. Lin, T. Inoue, A. Fukuda, T. Shimano, Y. Kuboki, T. Toyama, Y. Miyamoto, T. Kirihara, K. Hayasaka, A. Miyao, L. Monna, H.S. Zhong, Y. Tamura, Z.X. Wang, T. Momma, Y. Umehara, M. Yano, T. Sasaki & Y. Minobe, 1994. A 300 kilobase interval genetic map of rice including 883 expressed sequences. Nature Genetics 8: 365–372.

Lejeune, B., F. Quetier, M.F. Jubier, D. Falconet, A. Rode & C. Hartmann, 1988. Le génome mitochondrial des plantes supérieures: organisation moléculaire et expression. Bull Soc Fr 135: 49–55.

Meunier, J., 1992. Genetic diversity in Coconut, a brief survey of IRHO's work. In: Coconut Genetic Resources, pp. 59–62. Paper of an IBPGR Workshop, Cipanas, Indonesia, 8–11 October, 1991. International Crop Network Series n° 8. IBPGR, Rome.

N'Cho, Y.P., A. Sangaré, R. Bourdeix, F. Bonnot & L. Baudouin, 1993. Evaluation de quelques écotypes de cocotier par une approche biométrique. I. Etude des populations de Grands. Oléagineux 48: 3: 121–132.

Persley, G.J., 1992. Replanting the Tree of Life: Toward an International Agenda for Coconut Palm Research. Redwood Press Ltd., Melksham.

Taffin de, G., 1993. Le cocotier. Collection Le technicien d'agriculture tropicale. Maisonneuve et Larose, Paris.