Theor Appl Genet (2006) 112: 258–268
DOI 10.1007/s00122-005-0123-z

ORIGINAL PAPER

L. Baudouin · P. Lebrun · J. L. Konan · E. Ritter
A. Berger · N. Billotte

QTL analysis of fruit components in the progeny of a Rennell Island Tall coconut (*Cocos nucifera* L.) individual

Received: 22 April 2005 / Accepted: 24 September 2005 / Published online: 24 November 2005

Abstract We investigated the genetic factors controlling fruit components in coconut by performing QTL analyses for fruit component weights and ratios in a segregating progeny of a Rennell Island Tall genotype. The underlying linkage map of this population was already established in a previous study, as well as QTL analyses for fruit production, which were used to complement our results. The addition of 53 new markers (mainly SSRs) led to minor amendments in the map. A total of 52 putative QTLs were identified for the 11 traits under study. Thirty-four of them were grouped in six small clusters, which probably correspond to single pleiotropic genes. Some additional QTLs located apart from these clusters also had relatively large effects on the individual traits. The QTLs for fruit component weight, endosperm humidity and fruit production were found at different locations in the genome, suggesting that efficient marker-assisted selection for yield can be achieved by selecting QTLs for the individual components. The detected QTLs descend from a genotype belonging to the "Pacific" coconut group. Based on the known molecular and phenotypic differences between "Pacific" and "Indo-Atlantic" coconuts, we suggest that a large fraction of coconut genetic diversity is still to be investigated by studying populations derived from crosses between these groups.

Electronic Supplementary Material Supplementary material is available for this article at http://dx.doi.org/10.1007/10.1007/s00122-005-0123-z and is accessible for authorized users.

Communicated by C. Möllers

L. Baudouin (✉) · P. Lebrun · A. Berger · N. Billotte
CIRAD–Centre International de Recherches en Agronomie pour le Développement, Avenue Agropolis - TA 80 / 03,
34398 Montpellier Cedex 5, France
E-mail: luc.baudouin@cirad.fr
Tel.: 334-67-617116
Fax: 334-67-617183

J. L. Konan
CNRA–Centre National de la Recherche agronomique,
Station Marc Delorme 07 B.P. 13, Abidjan 07, Côte d'Ivoire

E. Ritter
NEIKER–Basque Institute for Research and Development in Agriculture, Apartado 46, E-1080 Vitoria, Spain

Introduction

Coconut is widely cultivated in the coastal tropical areas and represented until the 1970's as the most important fat and oil source. Although it is presently facing severe competition from other oil crops such as oil palm and soybean, it still represents an important asset for small farmers in these regions. It can be grown on poor soils where no other crop would survive, does not require an intensive management and offers a large variety of products for domestic consumption as well as for the local and international market (Persley 1992; Harries 2000). The commercially most interesting product is copra, the dried coconut meat that is used for oil extraction (Bourdeix et al. 2001). The solid endosperm (or meat) of the mature fruits is also frequently grated and used for cooking. Mixed with the liquid endosperm (or "coconut water"), it is known as "coconut milk". The liquid endosperm of young fruits provides a refreshing and sweet beverage (Child 1974). Other parts of the fruit are used for non-food purposes: the shell is used to make various household items, but can also provide an excellent activated charcoal, while its husk gives imputrescible fibers used, for example, to make ropes or door mats. Other parts of the plant are also useful. The sap in the inflorescences is used to prepare alcoholic drinks and the terminal bud is known as coconut "cabbage". Coconut leaves are used as thatch and for weaving various articles and coconut wood to make furniture.

Most of the traditional coconut cultivars ("Tall" coconuts) are out-breeding population varieties, but there is also a small group of self-pollinating types ("Dwarf" coconuts), which breed true to type. Improved cultivars are obtained by selecting progenitors among

Tall cultivars using mass selection or—better—by progeny testing. High-yielding varieties can be obtained by intercrossing well-chosen traditional cultivars. Most of these "hybrid" varieties are produced by crossing a Dwarf with a Tall cultivar, but cultivars obtained by Tall × Tall crosses may also be very attractive (Bourdeix et al. 2001).

Some of the characteristics that make coconut an attractive crop for farmers can also be seen as drawbacks from the plant breeder's point of view. Coconut is a tree whose lifespan exceeds 60 years and produces giant seeds, although in relatively small amounts. On the other hand, a single breeding cycle lasts more than 15 years, including controlled pollination and raising seeds in the nursery. Production starts at 3–4 years and is not stabilized before 6 years. Assessing the potential of a new variety requires recording yield and fruit traits of about 72 trees (corresponding to half a hectare) over an additional 6 years (Gascon and Nucé de Lamothe de 1976).

Coconut breeding is, therefore, a time and space consuming activity. The current tendency to promote multi-purpose use of coconut complicates this task even more since it entails multiple breeding criteria and thus requires observing a large number of traits in many genotypes. Other important factors to consider are those related to adaptation to the environment and particularly resistance or tolerance to pests and diseases. Deadly diseases such as lethal yellowing in the Caribbean region and Africa or Cadang-cadang in the Philippines represent serious threats for many farmers. Integrated control measures, including the adoption of suited varieties are necessary to preserve the viability of coconut cultivation.

It is practically infeasible to take all these factors into account in a single breeding experiment. A practical way to consider multiple selection criteria in a breeding scheme is to combine conventional breeding with marker-assisted selection using information collected from distinct experiments. Various kinds of molecular markers have been developed for genetic analysis in coconut. These include dominant markers such as RAPD (Ashburner et al. 1997; Wadt et al. 1999), AFLP (Perera et al. 1998) or ISTR (Duran et al. 1997; Rohde et al. 1999) as well as codominant markers such as RFLP (Lebrun et al. 1998) and microsatellite markers (Lebrun et al. 1999; Perera et al. 2000; Rivera et al. 1999; Meerow et al. 2003)). These markers have been used in population diversity studies by different research teams. Independent of the marker type, all results agreed on the partition of coconut populations into two large groups (Lebrun et al. 1998; Teulat et al. 2000). The "Pacific" group originates from an area between South-East Asia and the South Pacific and is the largest and most variable. This group also includes the Dwarf coconuts. The "Indo-Atlantic" group was differentiated in the Indian sub-continent and germplasm was subsequently disseminated by humans to West Africa and to the Atlantic coast of America (Harries 1977). Although this classification was based on molecular and geographical basis, it confirms the broad lines of the phenotypic classification proposed by Harries (1978): the molecular groups correspond broadly to the "wild " (or *Niu Kafa*) and to the "domesticated" (or *Niu Vai*) coconut types, respectively. Cultivars with *Niu Vai* traits in the Indian Ocean are the result of introgression from SouthEast Asia. However, cultivars with *Niu Kafa* traits in the Pacific do not seem to have connections with the Indo-Atlantic group.

Several segregating populations have been used to construct linkage maps and to identify putative QTL markers: an initial coconut linkage map was presented by Rohde et al. (1999). Herrán et al. (2000) used a cross between a Laguna Tall genotype from the Philippines and a Malayan Yellow Dwarf to construct a linkage map and to identify QTLs for germination precocity. The first linkage map involving an adult population was presented by Lebrun et al. (2001) and allowed the investigation of QTLs for yield components in a Cameroon Red Dwarf × Rennell Island Tall cross. The present paper is an extension of this study, which involves a larger number of markers and examines the distribution of QTLs involved in fruit composition and yield in this progeny.

Material and methods

Plant material

The mapping population was the same as used by Lebrun et al. (2001). This progeny represents a half sib family but may be considered as the progeny of a single heterozygous Rennell Island Tall (RIT) pollen donor (P02664) from the Solomon Islands, crossed with a pure line, represented by 12 Cameroon Red Dwarf (CRD) mother palms (P04383 to P04394) which are almost completely homozygous. Homozygosity and uniformity of this Dwarf cultivar is warranted by its highly self-pollinating habit and by the distinctive pale orange colour of leaf petioles, which is a recessive trait and allows elimination of off-type plants at the nursery stage. In addition, the 12 CRD mother palms were genotyped and only a minor degree of non-segregating polymorphisms appeared. Only two of the 69 progeny genotypes had to be excluded from the mapping population as off-type. In the 67 remaining palms, genetic variation is attributed to disjunction between alleles from the RIT parent. This progeny is one of the 16 treatments of progeny trial PBGC25, planted in 1981 in field M51 at the Marc Delorme Research Station (Côte d'Ivoire). According to the design of the genetic trial, the palms were planted in six different blocks.

Trait recording

Fruit component analyses were performed according to Santos et al. (1996) from September 2000 to July 2003 every 2 months corresponding to 18 harvesting rounds.

Due to fruit availability at the moment of harvest, the number of collected samples varied from 8 to 17 per palm and the number of fruits per sample varied from 1 to 4. In total 3,623 fruits were analyzed (on average 56 fruits/palm). The following fruit components were weighted:

(a) Whole fruit
(b) Husk = fibrous mesocarp
(c) Nut = fruit without husk = (a)-(b)
(d) Shell = ligneous endocarp
(e) Meat = solid endosperm
(f) "Coconut water" = liquid endosperm in the nut = (c)−(d)−(e)
(g) Weight of a sample of fresh endosperm
(h) Weight of the same sample after oven drying at 105°C for 10 hDerived traits were also calculated:

(i) Nut/fruit ratio = (c)/(a) × 100
(j) Shell/nut ratio = (d)/(c) × 100
(k) Meat/nut ratio = (e)/(c) × 100
(l) Water/nut ratio = (e)/(c) × 100
(m) Endosperm humidity ((g)−(h))/(g) ×100 to calculate copra yield.

In order to reduce the influence of variations of soil fertility across the field, adjusted values were calculated by subtracting from the individual values the deviations of the block means from the grand mean.

Statistical methods

A one-way analysis of variance was performed with each observed trait, according to the random model:

$$Y_{ij} - m = P_i + R_{ij}$$

where Y_{ij} represents the result of an observation, m the grand mean, P_i the effect associated to individual palms and R_{ij} the random deviation between observations made on, the same palm (Dagnelie 1970). The variance of R_{ij} (σ_r^2) was estimated as the residual mean square MS_r, while the variance of the "palm" effect (σ_r^2) was estimated as

$$\sigma_r^2 = (MS_p - MS_r)(p-1)/n' \text{ with } n' = n + \left(\sum n_i^2\right)\Big/ n$$

where MS_p is the "palm" mean square, p the number of palms, while n and n_i are, respectively, the total number of observations and the number of observations for palm number i.

Repeatability was calculated as $\sigma_p^2/(\sigma_p^2 + \sigma_r^2)$.

For correlation analyses Pearson coefficients were calculated using SAS Software (SAS Institute Inc. 1989).

Molecular markers

A total of 230 AFLP, SSR or RFLP markers were already available for linkage mapping from a previous study (see details in Lebrun et al. 2001). For constructing the present map we included one additional AFLP primer combination (E-GAA/P-AA; No. 97), generating four segregating fragments, eight heterologous SSR markers from oil palm (Billotte et al. 2005) and 48 new coconut SSR markers. These markers are part of 235 functional microsatellites in a coconut SSR library recently developed within the framework of a European project (Contract ICA-CT-2001−10066), described in the EMBL database (see electronic supplementary material for details). Among the 48 segregated in our progeny, 42 could be assigned to one of the 16 linkage groups.

The SSR analyses were performed on an automatic sequencer Li-Cor IR2 (Lincoln, Nebraska). For each SSR locus, one of the primers was designed with a 5′-end M13 extension (Steffens et al. 1993). For the PCR amplification, 25 ng of DNA was used in a 10 µl final volume, containing 0.08 µM of the M13 labeled primer, 0.1 µM of the other primer and 0.06 µM of M13 primer-fluorescent dye IR700 or IR800 (Biolego, The Netherlands). The PCR mix contained 1X Buffer (10 mM Tris–HCl pH 8, 50 mM KCl and 2 mM $MgCl_2$), 200 µM DNTP and 1 U Taq DNA polymerase. The PCR program started with an initial denaturation at 94°C for 5 min, then 35 cycles of 94°C for 30 s, 51°C for 1 min 15 s and 72°C for 1 min 30 s, and stopped after a final elongation at 72°C for 5 min. Each mix of the PCR products contained one or two IR700 and IR800 labeled M13 reverse complement extensions, diluted to one-fourth with formamid blue; 0.8 µl of the final mix was loaded on a 6.5% polyacrylamide gel and then detected by the IR fluorescence scanning system of the sequencer.

Data analysis and linkage mapping

Polymorphic DNA fragments were scored for presence or absence in parents and progeny genotypes. Linkage analysis between marker fragments, estimation of recombination frequencies, and determination of linear order between linked loci including multipoint linkage analysis and the EM algorithm for handling missing data were performed as described in Ritter et al. (1990) and Ritter and Salamini (1996). The MAPRF program (Ritter and Salamini 1996) was applied for the computational methods. The half-sib family was analyzed in the same way as a backcross progeny, considering only segregating fragments which were present in the RIT parent and absent in all CRD parents. Initially more closely linked fragments were arranged into many linkage sub-groups using a minimum LOD threshold of 5.0 between consecutive markers. Fragments composing each linkage subgroup were ordered by minimizing the sum of LOD scores for alternative orders. Finally, appropriate subgroups were connected based on maximum LOD values/minimum recombination frequencies between lateral markers of different linkage groups and

arranged into 16 linkage groups using always LOD values greater than 3.

QTL analyses

The complete linkage map was used to identify genomic regions controlling the traits mentioned above. QTLs were mapped using the least square interval mapping method developed for backcross progenies according to Knapp et al. (1990) and Knapp and Bridges (1990) and applied to all intervals composed of individual markers from the RIT parent. SAS software (SAS Institute Inc. 1989) and, in particular, the procedure PROC NLIN was used for computational analysis. The percentage of total variance explained by the sum of the individual QTLs was calculated by performing multiple regression analysis on the corresponding intervals.

Results

Variation in fruit components

Table 1 presents some basic statistics of the observed traits. All traits varied significantly between palms. Repeatabilities were relatively low, reflecting the natural variation of fruit size between successive bunches. Component weights and endosperm humidity were generally more repeatable (from 0.10 to 0.18) than the ratios between components (0.03–0.13). Despite low repeatabilities, the number of repetitions was sufficient to find significant differences between trees and to allow the detection of QTLs for all traits under study. The coefficient of variation (CV) among trees and the maximal and minimal values of each trait are presented in Table 1, reflecting the extent of these variations. The CVs of fruit component weights were around 7–10%, except the CV of water weight which was 15.3%. Variations of fruit component ratios were smaller and ranged between 2.2 and 6.7%.

Table 2 presents correlations between the studied traits. Fruit component weights were all positively correlated (Table 2a), which is a consequence of the interdependence between the different parts of the fruit, indicating that the weights of all components of the fruit tend to vary in the same direction. However, they do not vary in the same proportion: shell/nut and meat/nut ratios are negatively correlated with water/nut ratio (Table 2b) as well as with all weights (Table 2c), showing that water/nut ratio tends to be highest in the largest fruits, at the expense of the other two components of the nut. In addition, the moderate but significant correlations of endosperm humidity with fruit component traits suggest that larger fruits also tend to have a higher endosperm humidity.

Correlations with the yield traits studied in the previous article were also calculated, although we have to consider that these measurements were performed at different periods. According to our results (Table 2d), no significant correlations between fruit components and bunch numbers were found, as well as with nut numbers in younger trees (FN1). However, most fruit components were negatively and moderately correlated with fruit numbers recorded when fruit yield was stabilized (FN2). These significant correlations witness differences between palms in their allocation of resources within bunches (Bourdeix 1989). Some palms tend to produce fewer and larger fruits, while others tend to produce a larger number of smaller fruits. Lack of significant correlation with the early stage of production (FN1) suggests that this trait is more influenced by the age at which palms reach flowering maturity than by its yield potential. Considering fruit component ratios, positive correlations with FN2 were found only for meat/nut and shell/nut ratios, which

Table 1 Trait characteristics and results of QTL analyses for fruit and yield components

Trait	Abbreviation	Mean	σ_p	σ_r	F test	Repeatability	C.V. [%]	Max	Min	QTL code	Number of QTLs	Total R^2
Fruit weight (g)	FRUIT	1242.7	86	264	< 0.001	0.10	8.8	1495.8	999.4	QF	6	56.9
Husk weight (g)	HUSK	356.8	47	124	< 0.001	0.12	9.5	434.7	292.2	QH	3	40.7
Nut weight (g)	NUT	885.8	85	179	< 0.001	0.23	9.5	1061.1	701.7	QN	7	65.1
Shell weight (g)	SHELL	196.7	17	38	< 0.001	0.16	7.9	229.7	160.6	QS	7	57.3
Meat weight (g)	MEAT	415.5	29	66	< 0.001	0.17	7.5	473.0	346.4	QM	6	52.6
Water weight (g)	WATER	273.6	35	94	< 0.001	0.12	15.3	367.8	189.2	QW	4	46.9
Nut/fruit ratio (%)	NUT%F	71.2	1.1	5.8	0.004	0.03	2.2	74.1	67.1	Q%N	2	12.3
Shell/nut ratio (%)	SHELL%N	22.3	0.8	2.8	< 0.001	0.07	3.8	23.8	20.3	Q%S	5	39.4
Meat/nut ratio (%)	MEAT%N	47.0	1.4	4.9	< 0.001	0.08	3.6	50.6	44.0	Q%M	4	46.3
Water/nut ratio (%)	WATER%N	30.7	2.1	5.4	< 0.001	0.13	6.7	35.1	26.0	Q%W	4	43.0
Endosperm humidity (%)	HUMID	47.7	1.2	3.2	0.004	0.12	4.0	53.4	43.7	Q%U	4	44.7
Bunch number (years 3–6)[a]	BN1	9.3	0.2	–	–	–	–	–	–	Q1	3	19.4
Bunch number (years 7–14)[a]	BN2	13.5	0.1	–	–	–	–	–	–	Q2	3	40.3
Fruit number (years 3–5)[a]	FN1	84.1	1.7	–	–	–	–	–	–	Q3	1	12.8
Fruit number (years 3–5)[a]	FN2	121.8	6.7	–	–	–	–	–	–	Q4	3	48.3
Total											62	

CV coefficient of variation *Total* R^2: Percentage of variance explained by all QTLs for the trait σ_b Standard deviation of "between palm" variance component σ_r Standard deviation of residual variance component

[a]Previous results obtained by Lebrun et al. (2001)

Table 2 Results of correlation analyses for fruit component traits and yield components

A Between fruit component weights

	Husk	Nut	Shell	Meat	Water
FRUIT	82.8***	97.4***	91.0***	89.3***	95.3***
HUSK		68.0***	66.8***	56.8***	69.3***
NUT			92.1***	93.9***	96.6***
SHELL				82.8***	86.1***
MEAT					83.3***

B Between fruit component weights and ratios

	Fruit	Husk	Nut	Shell	Meat	Water
NUT (%F)	18.0ns	−39.9***	39.7***	30.4*	46.2***	34.0**
SHELL (%N)	−53.3***	−28.8*	−58.1*	−22.2ns	−62.8***	−61.7***
MEAT (%N)	−71.8***	−61.4***	−69.2***	−71.5***	−40.4***	−82.1***
WATER (%N)	80.9***	62.2***	80.7***	67.8***	59.1***	92.8***
HUMID	51.3***	49.2***	47.5***	43.6***	40.6***	49.2***

C Between fruit component ratios

	Shell (%N)	Meat (%N)	Water (%N)	Humid
NUT (%F)	−39.1***	−9.7ns	24.1*	−2.8ns
SHELL (%N)		23.3ns	−60.5***	−30.4*
MEAT (%N)			−91.5***	−39.8***
WATER (%N)				44.5***

D Between fruit production and fruit component weights and ratios

	BN1	BN2	FN1	FN2
FRUIT	3.9ns	−16.2ns	−14.3ns	−36.4*
HUSK	10.7ns	−4.0ns	−10.2ns	−24.6ns
NUT	0.8ns	−19.6ns	−14.6ns	−37.9**
SHELL	5.8ns	−13.3ns	−15.0ns	−31.2*
MEAT	−4.0ns	−22.2ns	−17.7n	−37.4**
WATER	2.3ns	−18.2ns	−10.9ns	−37.2**
NUT (%F)	−13.2ns	−20.2ns	−6.0ns	−16.5ns
SHELL (%N)	10.0ns	22.8ns	6.3ns	30.5*
MEAT (%N)	−12.3ns	3.5ns	−0.1ns	21.7ns
WATER (%N)	5.8ns	−12.2ns	−2.5ns	−29.9*
HUMID	0.0ns	−14.9ns	−12.1ns	−3.7*

ns not significant at the 5% error level
see Table 1 for abbreviations

simply reflect the already noted relations between fruit size and fruit composition.

Linkage mapping

In total, 290 segregating markers (80 SSR, including 8 *E. guineensis* markers, six RFLP and 204 AFLP markers) were available for linkage mapping. A linkage map of 16 linkage groups was established for the RIT parent of the mapping population as described above. The total length of the map was 1,849.8 cM and linkage groups varied in lengths from 51.9 to 181.8 cM. A total of 274 markers were located on the updated linkage map (47 more than in the previous map by Lebrun et al. 2001). The number of markers per linkage group varied between 6 and 28 and the average density between 3.3 and 13.8 cM/marker. A total of 16 markers (eight AFLP, one oil palm and seven coconut SSR) revealed ambiguous map locations and were discarded.

QTL analysis

Forty-eight putative QTLs were identified for fruit component traits at a significance level of 5% and additional four QTLs for endosperm humidity (Table 3 and Fig. 1). QTLs were identified for all studied traits and their number varied from two for nut/fruit ratio to seven for nut and shell weight (Table 1). However, 34 of these QTLs were located in six clusters, identified as C1 to C6 in Table 3. These clusters grouped 4 to 9 QTLs in small regions of the genome (less than 20 cM). Within the same cluster, the effect of all QTLs associated with weight components had the same sign. The importance of these clusters is not only underlined by the number of QTLs, but also by the magnitudes of their effects, which tend to be higher in clusters than for ''scattered'' QTLs. The percentages of variance explained by individual QTLs (R^2) varied from 9 to 28% within clusters, but only from 5 to 15% for the 14 remaining QTLs. The four QTLs identified for endosperm humidity explained between 11 and 22% of the genetic variance and were not located within these clusters. A large part of genetic variance for fruit components was explained by the sum of the effects of all individual QTLs for each trait (Total R^2; Table 1). As usual total R^2 was smaller than the sum of the individual QTL contributions due to the correlative effects, which exist between QT alleles from different QTLs. Total R^2 was above 40% for all traits except for nut/fruit and shell/nut ratios and was even above 50% for most fruit component weights indicating that marker-assisted selection on the various traits is likely to be effective.

Table 3 QTL locations and characteristics for fruit component traits detected in linkage groups of a RIT × CRD progeny

Linkage group	QTL Code	Marker Interval	Absolute and relative effects					
			R	*R*1			Pr1	R^2
Lg 1	QMa	mCnCir202-74/2	9.8	0	22.2	5%	1.2	11.5
C1	QFa	36/2-94/1	3.2	0	-74.4	-6%	2.9	10.6
	QNa			0	-56.2	-6%	2.8	10.7
	QWa			0	-30.9	-11%	1.3	12.8
	%Wa			0	-1.51	-5%	1.6	11.8
C2	QFb	37/2-mCnCir73	5.5	0	81.5	7%	1.0	14.8
	QNb			0	60.5	7%	1.4	13.9
	QSa			0	10.5	5%	2.4	12.5
	Q%Ma			4.4	-1.11	-2%	4.4	11.1
Lg 2	Q3a	38/3-mCnCir206	3.4	0	1.42	2%	3.0	12.8
C3	QMb			0	-25.0	-6%	0.5	15.9
	QFc			0.3	-74.8	-6%	2.1	12.8
	QNc			1.3	-62.3	-7%	1.2	14.8
	QSb			3.4	-13.1	-7%	0.3	18.4
	Q2a	mCnCir215-90/2	5.6	5.6	0.11	1%	0.1	18.9
	Q4a			5.6	6.54	5%	0.1	22.6
	Q%Ua	CN11E10-CNZ33	6.8	3.9	1.37	3%	2.9	11.8
Lg 3	QNd	53/1-53/5	6.2	0	-55.8	-6%	3.2	9.4
	QWb			0	-27.8	-10%	3.9	8.7
	QSc	53/5-64/3	11.5	1.8	11.1	6%	2.9	10.9
	QNe	43/2-37/3	1.6	1.6	-54.7	-6%	3.3	9.4
	QSd			0	-11.5	-6%	1.2	12.8
Lg 4	Q%Sa	CnCirE1-55/4	6.3	6.3	0.64	3%	0.9	11.7
	Q%Wb	33/2-CnCirD1	4.5	4.5	1.33	4%	3.2	6.0
Lg 5	Q%Ub	72/1-Eg5/EcoR1	1.7	0	2.01	4%	0.0	24.6
C4	Q%Wc	29/1-CnCirH11	6.1	0.2	1.41	5%	2.1	10.9
	QWc			0.4	26.1	10%	4.2	9.1
	Q%Mb			1.8	-1.4	-3%	0.4	15.9
	QFd			2.5	70.1	6%	4.1	9.6
	QSe			3.3	10.2	5%	3.7	10.0
	QHa			4.9	24.4	7%	1.5	12.2
Lg 7	Q%Mc	81/4-33/3	27.6	3.8	1.23	3%	3.9	9.2
	Q%Uc	33/3-71/3	13	3.6	1.8	4%	0.1	22.2
	QMc	76/1-58/1	9.8	9.8	21.6	5%	2.1	8.4
Lg 8	Q4b	CnCirD8-mCnCir121	29.5	8.6	6.55	5%	1.5	17.2
Lg 9	Q1b	31/5-54/7	9.3	0	0.12	1%	3.5	9.0

Table 3 (Contd.)

Linkage group	QTL Code	Marker Interval	*R*	*R*1	Absolute and relative effects		Pr1	R^2
Lg 10	QHb	CnCirB11-CNZ26	7.8	3.4	29.5	8%	0.3	17.6
	QNf	CNZ26-60/2	16.9	3.1	65.5	7%	1.7	12.5
	QMd			3.2	26.6	6%	0.5	15.8
C5	Qfe			4.5	96.1	8%	0.7	15.5
	QSf	60/2-63/2	10.5	0	11.1	6%	1.5	12.3
	Q2c			0	-0.07	-1%	3.1	9.5
	Q%Na	35/6-CNZ23	32.8	21.4	1.29	2%	2.5	11.0
	Qme			32.8	22.6	5%	1.5	6.3
	Q%Sb	CNZ23-37/1	11.5	6.8	0.71	3%	1.0	14.5
	Q2b	37/1-79/1	1.9	1.9	0.09	1%	1.9	13.7
	Q4c	79/1-76/3	3.8	2.2	5.7	5%	1.2	16.6
Lg 11	Q%Sc	87/1-73/1	3.3	3.3	0.56	3%	3.3	9.6
Lg 12	Q%Ud	110/1-54/1	11.3	11.3	1.63	3%	2.6	10.9
	Q1c	147/1-57/2	23.2	12.3	0.13	1%	2.0	13.8
Lg 13	QSg	mCnCir179-74/1	8.1	0	-14	-7%	0.1	18.1
	QMf			0.6	-25.7	-6%	0.5	15.5
	QNg			1.5	-85.9	-10%	0.0	23.4
	Q%Sd			7.7	0.61	3%	2.3	11.2
C6	QFf	74/1-43/3	1.7	0	-110.1	-9%	0.0	23.1
	QHc			0	-23.5	-7%	0.0	26.3
	QWd			0	-47.5	-17%	0.0	28.7
	Q%Md			0	1.8	4%	0.0	28.6
	Q%Wd			0	-2.35	-8%	2.3	11.9
Lg 14	Q%Nb	CnCirA9-31/2	27.3	0	1.01	1%	3.3	5.1
	Q%Se				0.53	2%	3.7	5.7
	Q1a	31/2-43/1	16.4	16.4	0.11	1%	4.6	6.7
Lg 15	Q%Sf	91/5-Eg5/EcoR1	1.6	1.6	0.67	3%	0.0	14.3

Ci Clusters of QTLs for different traits indicated by brackets
R Length of the marker interval where the QTL is located
*R*1 position in the interval, measured form the left marker of the given interval
Effects Absolute and relative [%] effects measured as deviations between marker classes (relative to the presence of the left marker in the interval) caused by the different alleles of the QTL
Prl Probability for the null hypothesis of no QTL
R^2 Portion of the total variance explained by the QTL

Discussion

Our work represents the first attempt to identify QTLs for fruit component weights in coconut. Over 3,000 fruits had to be analyzed on a tree-by-tree basis specially for this purpose, since this type of analyses is usually not done in a breeding experiment such as PBGC25. Although variation among fruits of the same tree was

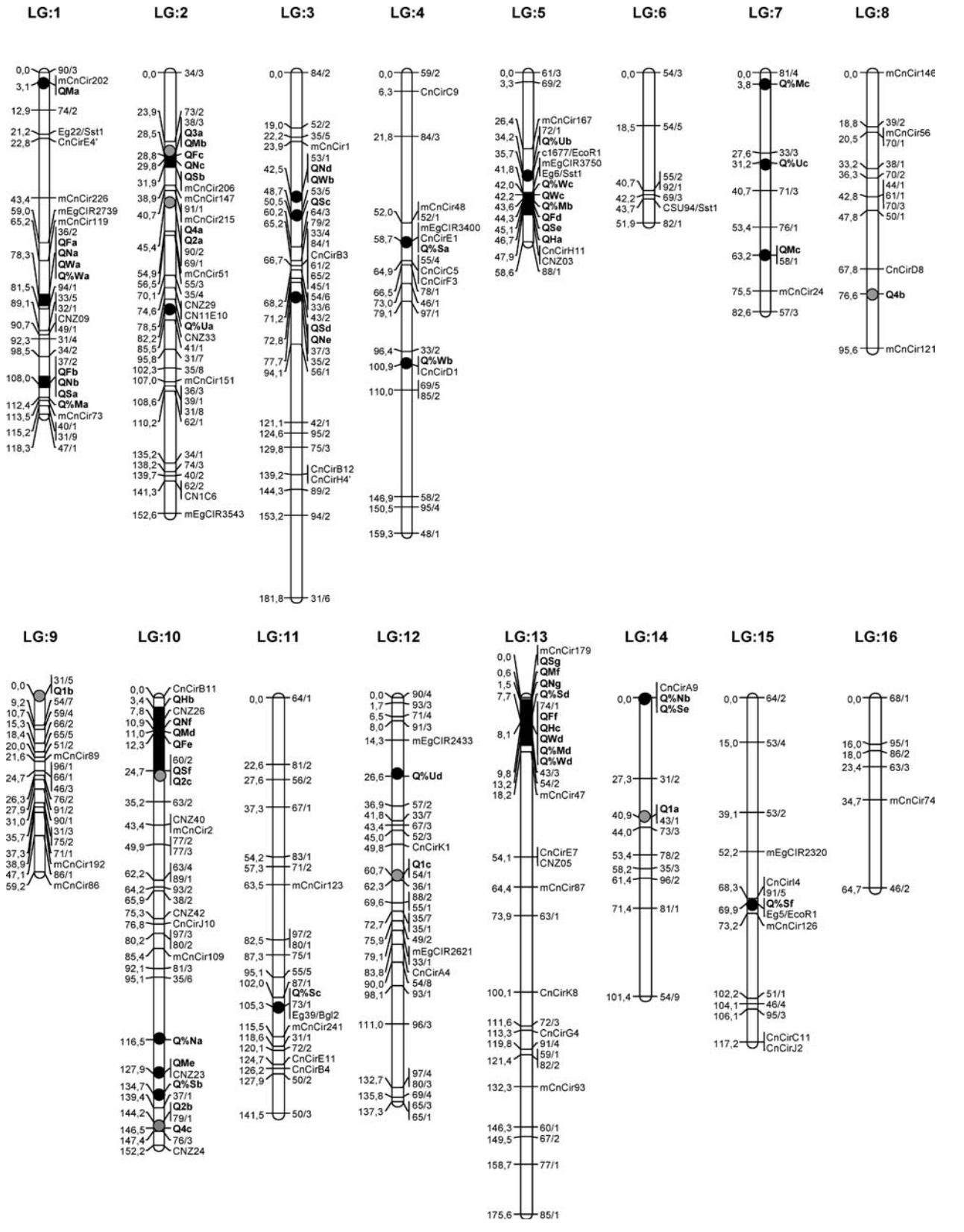

Fig. 1 Coconut linkage map based on the Rennell Island Tall genotype P02664 and location of the QTLs (in *bold* and with prefix "Q"). See Table 1 for QTL abbreviations. AFLP and RFLP markers are labeled according to Lebrun et al. (2001). Marker labels which are not numbers correspond to SSR or RFLP markers. SSR markers derived from oil palm are indicated with the prefix "mEgCIR". *Black Rectangles* correspond to the six major QTL clusters mentioned in the text. *Black circles* indicate scattered QTLs for fruit components and *grey circles* for fruit and bunch numbers

high, trait differences between trees were significant and QTLs were detected for all traits under study.

Linkage mapping

Compared to the previous work presented by Lebrun et al. (2001), the updated map presented here incorporates 47 additional markers (20%), among which 44 are SSRs. There is now, at least, one codominant marker (SSR or RFLP) on each of the linkage groups. The presence of a large number of SSRs in a linkage map is important since they are highly polymorphic, codominant markers and amplify mainly single loci. They are particularly useful for comparing linkage maps from different genetic backgrounds, since they usually map to identical chromosomal locations. In this way it is possible to align maps, which were constructed in different segregating populations. Moreover, the sequence information and map locations of SSR markers provided in electronic supplementary material is useful for constructing framework maps in coconut.

Of particular interest is also the mapping of seven of the eight heterologous SSR markers from oil palm to our linkage map. Six of them have also been mapped in an oil palm population (Billotte et al. 2005) and markers mEgCIR2433 and mEgCIR2621 are located on the same linkage group in both species (see LG12 in Fig. 1 and electronic supplementary material). With an increased number of such markers, it will be possible to compare linkage maps from both the species and perform syntheny studies between these related species.

The general structure of the new map was the same as in the previous study, but some fragments, already considered as weakly linked before, were associated in a different way. This concerns distal marker subgroups located on the linkage groups 1, 3, 12 and 13 of the new map. Such adjustments were not unexpected, since in the first map 15 intervals were longer than 30 cM. This value represents the upper limit of linkage detection with 67 segregating genotypes. However, for the available data the previous map represented the "best" arrangement of the markers among various possible solutions, based on the smallest recombination frequencies or highest LOD scores. In the present map the number of intervals longer than 30 cM dropped to six (located on LG 3,4, 13, 14 and 16). Marginal rearrangements still remain possible in the future, as more markers are incorporated into the map.

QTL identification

From three to seven QTLs were detected for the various fruit component traits under study. Most of them and also the most important ones are located in six clusters. These correspond probably to pleiotropic loci rather than to closely linked genes for individual traits. This assumption is supported by the fact that the QTL effects of various weighted fruit components were in the same direction in the QTLs of all clusters. Moreover, pleiotropy is consistent with the relationships between the different components described above. In contrary, the QTL effects associated with meat/nut and shell/nut ratios were of the opposite direction in clusters C2, C4 and C6, confirming that genetic factors that tend to increase fruit size are particularly effective for increasing its water contents. In this context it is worth to notice that high pressures (5 atm) are observed in the liquid endosperm during fruit growth (Frémond et al. 1966). It is likely that the accumulation of liquid endosperm plays an important role in this process, both, through the mechanical force for enlarging the fruits and by providing nutrients for the formation of the solid endosperm. The six QTL clusters appear to act on fruit development as a whole, while the remaining 14 QTLs seem to act on specific, individual fruit components. Among them, three can be highlighted due to their notable effects on meat and shell amount in the fruit: QMa on LG1, QSd on LG3 and Q%Sa on LG4. Such QTLs may be useful for breeders, who want to alter the fruit composition, rather than its size.

Although, some correlations were observed between fruit components and endosperm humidity, this trait seems to be governed by different genetic factors, since its QTLs are not co-located with those of the fruit components. The same conclusion applies also for fruit and bunch numbers, since only the QTL Q2c is located within cluster C5. Also QTLs associated with "adult" nut productions (Q4a, Q4b and Q4c) are located apart on linkage groups 2, 8 and 10, respectively. These findings suggest that the negative physiological correlations between fruit size and fruit numbers are not strict. In other words, it is possible to select for copra yield by selecting for the QTLs of its individual components. The trait "Copra yield" was not included per *se* in our study. However, identified QTLs for meat, endosperm humidity and fruit production are highly indicative, since regression coefficients larger than $R^2 = 99\%$ between copra yield and its components were determined (data not shown). These findings leave little space for possible QTLs associated with interactions between copra yield and these traits.

Fruit and fruit component weights are together with fruit production the essential yield components in coconut. Moreover, the desirable fruit composition may differ according to the intended use of coconut products. For these reasons, knowledge about the location of the determining factors of these traits in the genome and the magnitude of their effects is an asset for plant breeders, particularly in a species with a long life cycle like coconut. It is, for example, possible to apply a marker-assisted preselection in nursery and to reduce the time and space needed for the final selection, which must take place in the field. Marker-assisted breeding would also allow additional selections for fruit and yield traits in disease resistance trials, which normally are planted in environments that are not well-suited for phenotypic selection.

However, the implementation of marker-assisted selection in coconut breeding requires some additional experiments. For example, analyses of larger progenies are necessary, which would allow the application of regressions on multiple intervals in the same model and in this way avoid overestimation of individual QTL effects or even the appearance of minor "false positive" QTLs. On the other hand, the identification of a QTL requires alleles with different, measurable effects at the QTL. The probability that a given QTL will segregate also in a new progeny of a different progenitor depends on the relative frequencies of the corresponding alleles in a population and ranges between 0 and 0.5. The reproducibility of identified QTLs can be assessed by genotyping the progenitors together with a small number of progeny individuals with low- and high-trait expressions using known QTL markers. Such progenies crossed with various testers exist, for example, in other genetic trials in Ivory Coast. They offer an excellent opportunity to assess the general validity of the identified QTLs, to determine allelic QTL configurations in progenitors and their potential breeding value.

To date, all QTL studies on coconut, including the present one, were based on the first generation of a controlled cross and segregating genes came from the same population. Hence, there is a large part of the coconut diversity left unexploited from the molecular breeding point of view: the differences between coconuts belonging to the mentioned "Pacific" and "Indo-Atlantic" groups. In addition to the molecular differences, phenotypic characteristics are contrasted (Lebrun et al. 2003). Despite some exceptions, most "Pacific" coconuts have round fruits with a low husk/fruit ratio and a high water/nut ratio, while most "Indo-Atlantic" coconuts have elongated nuts with a high husk/fruit ratio and a low water/nut ratio. This distinction corresponds to the *Niu Kafa* and *Niu Vai* types described by Harries (1978). The genetic origin of these differences suggests at least the presence of different QT alleles in the corresponding populations and perhaps additional QT loci. F1 progenies derived from parents of each group would allow combining QT alleles from both groups in a highly heterozygous condition. The genetic and phenotypic differences between the two groups would probably also provide a large number of segregating loci and an elevated phenotypic variance. However, in order to contrast directly the particular effects of the QT alleles, a further cross is required. The use of a dwarf genotype for this purpose has several advantages, such as the ease of seed production and the simplicity of segregation analysis. One advantage of such a crossing plan is that it makes it possible to identify QTLs that are fixed in each parent population, which is impossible with conventional mapping populations.

Such three-way crosses exist to our knowledge in Côte d'Ivoire in the form of a (Malayan Yellow Dwarf × [West African Tall × Rennell Island Tall]) population and in the Philippines as a (Catigan Green Dwarf × [West African Tall × Laguna Tall]) progeny. These populations represent interesting study objects to validate these hypotheses. The generated knowledge could be applied in breeding programmes, to introduce, for example, specific QT alleles modulating desirable fruit traits into locally adapted populations through marker-assisted selection.

Acknowledgement The research was supported, in part, by the European Community under contract ICA4–CT-2001-10066 of the INCO-DEV programme.

References

Ashburner GR, Thompson WK, Halloran GM, (1997) RAPD Analysis of South Pacific coconut palm populations. Crop Sci 37: 992–997

Billotte N, Marseillac N, Risterucci A-M, Adon B, Brottier P, et al (2005) Microsatellite-based high density linkage map in oil palm. TAG 110: 754–765

Bourdeix R (1989) La sélection du cocotier, Cocos nucifera L.: Etude théorique et pratique, optimisation des stratégies d'amélioration génétique, pp196. Université Paris XI, Orsay

Bourdeix R, Baudouin L, Billotte N, Labouisse JP, Noiret J-M (2001) Coconut. In: Charrier A, Jacquot M, Hamon S, Nicolas D (eds) Tropical Plant Breeding. Science, pp 106–127

Child R (1974) Coconuts. Longman, London

Dagnelie P (1970) Théorie et méthodes statistiques. Presses Agronomiques de Gembloux, Gembloux

Duran Y, Rohde W, Kullaya A, Goikoetxea P, Ritter E (1997) Molecular analysis of East African Tall coconut genotypes by DNA marker technology. J Genet Breed 51: 279–288

Frémond Y, Ziller R, Nucé de Lamothe de M (1966) Le Cocotier. Maissonneuve et Larose

Gascon JP, Nucé de Lamothe de M (1976) Amélioration du cocotier: méthode et suggestions pour une coopération internationale. Oléagineux 31: 479–482

Harries HC (1977) The Cape Verde region (1499 to 1549); the key to coconut in the Western hemisphere? Turrialba 27: 227–231

Harries HC (1978) The evolution, dissemination and classification of Cocos nucifera L. The botanical review 44: 266–319

Harries HC (2000) The coconut palm Elsevier: 321–338

Herran A, Estioko L, Becker D, Rodriguez MJB, Rohde W, et al (2000) Linkage mapping and QTL analysis in coconut (cocos nucifera L). Theor Appl Genet 101: 292–300

Knapp SJ, Bridges WC (1990) Programs for mapping quantitative trait loci using flanking mo-lecular markers and nonlinear models. J Hered 81: 234–235

Knapp SJ, Bridges WC, Birkes D (1990) Mapping quantitative trait loci using molecular marker maps. Theor Appl Genet 79: 583–592

Lebrun P, Baudouin L, Bourdeix R, Konan JL, Barker JHA, et al (2001) Construction of a linkage map of the Rennell Island Tall coconut type (cocos nucifera L) and QTL analysis for yield characters. Genome 44: 962–970

Lebrun P, Grivet L, Baudouin L (1999) Use of RFLP markers to study the diversity of the coconut palm. In: Oropeza C, Verdeil JL, Ashburner GR, Rolando C, Santamaria JM (eds) Current advances in coconut biotechnology. Kluwer Academic Publishers, Mérida, pp 73–87

Lebrun P, N'Cho YP, Bourdeix R, Baudouin L (2003) Coconut. In: Hamon P, Seguin M, Perrier X, Glaszmann J-C (eds) Genetic diversity of cultivated tropical plants. SPI and Cirad, Montpellier, pp 219–238

Lebrun P, N'cho YP, Seguin M, Grivet L, Baudouin L (1998) Genetic diversity in coconut (Cocos nucifera L) revealed by restriction fragment length polymorphism (RFLP) markers. Euphytica 101:103–108

Meerow AW, Wisser RJ, Brown JS, Kuhn DN, Schnell RJ, et al (2003) Analysis of genetic diversity and population structure within Florida coconut (cocos nucifera L) germplasm using

microsatellite DNA, with special emphasis on the Fiji Dwarf cultivar. Theor Appl Genet 106:715–726
Perera L, Russel JR, Provan J, McNicol JW, Powel W (1998) Evaluating genetic relationships between indigenous coconut (*Cocos nucifera* L) accessions from Sri Lanka by means of AFLP profiling. Theor Appl Genet 96:545–550
Perera L, Russell JR, Provan J, Powell W (2000) Use of microsatellite DNA markers to investigate the level of genetic diversity and population genetic structure of coconut (*Cocos nucifera* L). Genome 43:15–21
Persley GJ (1992) Replanting the Tree of Life. Commonwealth Agricultural Bureau International, Wallingford UK
Rivera R, Edwards KJ, Arnold GM, Ayad G, Hodgkin T, et al (1999) Isolation and characterisation of polymorphic microsatellites in Cocos nucifera L. Genome 42:668–675
Ritter E, Gebhardt C, Salamini F (1990) Estimation of recombination frequencies and construction of RFLP linkage maps in plants from crosses between heterozygous parents. Genetics 224:645–654
Ritter E, Salamini F (1996) The calculation of recombination frequencies in crosses of allogamous plant species with application to linkage mapping. Genet Res 67:55–65
Rohde W, Becker D, Kullaya A, Rodriguez J, Herran A, et al (1999) Analysis of coconut germplasm biodiversity by DNA marker technologies and construction of a genetic linkage map. In: Oropeza C, Verdeil JL, Ashburner GR, Cardena R, Santamaria JM (eds) Current advances in coconut biotechnology. Kluwer Academic Publishers, Dordrecht, pp 99–120
Santos GA, Batugal PA, Othman A, Baudouin L, Labouisse JP (1996) Manual on standardized research techniques in coconut breeding. IPGRI, Cogent
SAS Institute Inc. (1989) SAS/STAT Users Guide, Version 6, Fourth Edn, vol 2, SAS Institute Inc., Cary, NC
Steffens DL, Sutter SL, Roemer SC (1993) An alternate universal forward primer for improved automated sequencing of M13. Biotechniques 15:63–68
Teulat B, Aldam C, Trehin R, Lebrun P, Barker JHA, et al (2000) An analysis of genetic diversity in coconut (*Cocos nucifera*) populations from across the geographic range using sequence-tagged microsatellites (SSRs) and AFLPs. Theor Appl Genet 100:764–771
Wadt LHO, Sakiyama NS, Pereira MG, Tupinamba EA, Ribeiro FE, et al (1999) RAPD markers in the genetic diversity study of the coconut palm. In: Oropeza C, Verdeil JL, Ashburner GR, Cardena R, Santamaria JM (eds) Current advances in coconut biotechnology. Kluwer Academic Publishers, Dordrecht, pp 89–97