Theor Appl Genet (2001) 103:1302–1310

K.A. Rance · S. Mayes · Z. Price · P.L. Jack
R.H.V. Corley

Quantitative trait loci for yield components in oil palm (*Elaeis guineensis* Jacq.)

Received: 26 June 2000 / Accepted: 24 October 2000

Abstract The development of an oil palm RFLP marker map has enabled marker-based QTL mapping studies to be undertaken. Information from 153 RFLP markers was used in combination with phenotypic data from an F_2 population to estimate the position and effects of quantitative trait loci (QTLs) for traits including yield of fruit and its components and measures of vegetative growth. The mapping population consisted of 84 palms segregating for the major gene influencing shell thickness. Marker data were analysed to produce a linkage map consisting of 22 linkage groups. The QTL mapping analysis was carried out by interval mapping and single-marker analysis for the unlinked markers; significance thresholds were generated by permutation. Using both single-marker and interval-mapping analysis significant marker associated QTL effects were found for 11 of the 13 traits analysed. The results of interval-mapping analysis of fruit weight, petiole cross section and rachis length, and ratios of shell:fruit, mesocarp:fruit and kernel:fruit indicated significant (P<0.05) QTLs at the genome-wide threshold. The putative QTLs were associated with between 8.2% and 44.0% of the phenotypic variation, with an average of 27% for the single-marker analysis and 19% for the interval-mapping analysis. The higher percentage of phenotypic variation explained in the single-marker analysis, when compared to the interval-mapping analysis, is likely to be due to the lower stringency associated with the single-marker analysis. Large dominance deviations were associated with a sizeable proportion of the putative QTLs. The ultimate objective of mapping QTLs in commercial populations is to utilise novel breeding strategies such as marker-assisted selection (MAS). The potential impact of MAS in oil palm breeding programmes is discussed.

Keywords QTL · Marker-assisted selection · Oil palm · Economic trait · RFLP map

Communicated by J.W. Snape

K.A. Rance(✉)
Department of Agriculture & Forestry, University of Aberdeen, MacRobert Building, 581 King Street, Aberdeen, AB24 5UA, UK
e-mail: k.rance@abdn.ac.uk
fax: +44(0)1224 273731

K.A. Rance · Z. Price · P.L. Jack
Plant Breeders International Cambridge, Maris Lane, Trumpington, Cambridge, CB2 2LQ, UK

S. Mayes · Z. Price
Department of Genetics, University of Cambridge, Tennis Court Road, Cambridge, CB2 3EF, UK

R.H.V. Corley
Highlands, New Road, Great Barford, Bedford, MK44 4LQ, UK

Introduction

The oil palm (*Elaeis guineensis* Jacq.) is a plantation crop of major economic importance in South East Asia, Africa and South America, giving rise to a diverse range of commercial products ranging from margarine and cooking oils to animal feeds, soaps and detergents. The production of, and demand for, oil palm has risen dramatically in recent years and, as a result, there has been substantial interest in increasing production efficiency by selective breeding. Commercial breeding and selection began in the early 1920s and since then considerable improvements have been made in both yield and quality. Yield has essentially quadrupled in this time (Hartley 1988) with over 50% being attributed to genetic improvement (Corley and Lee 1992). Despite the impressive increases in productivity to date, Corley (1983) estimates that oil palm has the physiological potential to produce 17 tonnes of oil hectare^{-1} year^{-1} compared to the best reported yields of 10.5 tonnes hectare^{-1} year^{-1} achieved in Malaysia.

Of the limited number of simply inherited traits that have been identified in oil palm, only the shell-thickness gene has a clear commercial value, with the heterozygotes having at least a 30% yield advantage over either homozygote. The majority of other traits of economic importance, however, are of a quantitative nature. Identification of the individual genetic factors underlying

quantitative traits or quantitative trait loci (QTLs) will provide the potential for improved oil palm breeding programmes. Establishing new methodologies such as marker-assisted selection (Mohan et al. 1997), where individuals are selected on the basis of genetic marker information, would represent a major step forward in breeding technology. Currently, a typical breeding cycle for mother palms of commercial seed (dura) is 10 years, while that for pollen parents (pisifera) is nearer 16 years. The crossing and nursery stages take two years; seedlings are then field-planted, typically at 143 palms/ha, and start fruiting after $2^1/_2$ years. A further 5 years of recording are then needed to establish reliable phenotypic values for yield components. The ability to use markers to select progeny while they were still in the nursery, allowing only the selected palms to be field-planted, could be of great potential value.

The development of an oil palm RFLP marker map (Mayes et al. 1997; Price, unpublished data) has enabled QTL mapping studies to be carried out. The objective of the study described in this paper was to investigate the underlying genetic basis of production traits in oil palm.

Materials and methods

Population and marker data

The QTL mapping study described in this paper was based on the analysis of phenotypic data obtained from a population derived from the self-fertilisation of palm A137/30 to yield an F_2 population previously described by Mayes et al. (1997). The phenotypic and molecular-marker information was available for the population planted at the Univanich Palm Oil Company Ltd., Thailand (cross UV101). The UV101 trial population consists of 84 palms segregating for the major gene influencing shell-thickness. The segregation ratio for the shell thickness gene in the UV101 population was 16 homozygous thick-shell types (dura), 45 heterozygotes (tenera), and 23 homozygous shell-less types (pisifera). Phenotypic data were available for the yield of fruit and its components, bunch number and weight, fruit bunch composition (with the exception of pisifera palms which characteristically show female infertility), and measures of vegetative growth (see Table 1). The genetic basis of the female infertility observed in pisifera palms remains uncertain. No conclusive evidence is available to confirm whether the infertility is due to a pleiotropic effect of the shell-thickness gene itself or is a result of linkage between the shell-thickness gene and an as yet unidentified gene bringing about female infertility (Wonkyi-Appiah 1987).

The phenotypic data available were mean values obtained over a 5-year trial period. The molecular marker information available was generated as part of a study to establish an oil palm RFLP map (Jack et al. 1995; Mayes et al. 1997) together with additional RFLP marker data in the same population (Price, unpublished data). Data for a total of 153 RFLP markers were available. For a detailed explanation of the protocols used for probe development, DNA extraction, Southern blotting, and RFLP probing, see Jack et al. (1995).

Linkage map construction

The molecular-marker information was used to construct a linkage map. All markers were tested for goodness of fit (χ^2 test) to a 1:2:1 (*df*=2) segregation ratio for co-dominant loci or a 1:3 ratio (*df*=1) for markers scored as dominant. The linkage map was constructed using JoinMap 2.0 (Stam 1993). Haldane's mapping function was applied and a LOD 4 threshold and a recombination fraction of 0.49 were used to determine linkage groups. A LOD threshold of 3 was used to determine marker order.

QTL mapping analysis

The data analysis was carried out in two stages; initial exploratory analysis followed by QTL mapping including single-marker analysis and interval mapping. All trait data were explored to determine the significance of the shell-thickness genotype on the measures of production traits, the equality of variance between the different shell-thickness classes, and the normality of phenotypic data (and residual trait data where appropriate). All pisifera individuals were excluded from any further analysis due to the female infertility in the pisifera palms and unequal variances within the shell-thickness genotypic groups; as a result, all relevant analyses were repeated. Unless otherwise stated all statistical analyses were carried out using Genstat (Genstat 5 Committee 1993).

Single-marker analysis was carried out to determine marker-QTL effects associated with unmapped marker loci. A simple ANOVA was employed to determine significant ($P<0.05$) associations between marker information and phenotypic data. Where appropriate, the shell-thickness genotype was also fitted in the ANOVA model.

Interval mapping was carried out using the QTL Cartographer analysis package (Basten et al. 1994, 1999). The model used in the QTL Cartographer package was Model 6, where co-factors (or background markers) are fitted to account for background genetic variance (Jansen 1993, 1996; Zeng 1994). These co-factors were selected by a forward and backwards (FB) stepwise regression, where markers are selected and ranked according to their effect on the quantitative trait. In addition a window-size parameter was used to exclude the use of co-factors from a region of the genome on either side of the markers flanking the test site. This prevents the background markers eliminating variance due to the test site itself [for a more detailed explanation of the selection and use of co-factors in the QTL Cartographer see Basten et al. (1999)]. An upper limit of five co-factors was used in the analysis and the window size was set to 10 cM as suggested by Basten et al. Where appropriate, corrected trait values (fitted mean+residuals) were analysed to account for the fixed effect of shell-type.

The empirical thresholds for QTL detection were calculated by 2,000 replicated permutations, based on the method described by Churchill and Doerge (1994). This method uses a permutation test where phenotypic data are permuted with respect to marker data. Due to the number of traits analysed, 13 in total, a single simulated trait (mean=0, variance=1) was used to generate genome-wide thresholds at $P<0.01$ and $P<0.05$. In addition, empirical chromosomal thresholds were produced by taking the mean of 2,000 replicate permutations for each chromosome.

Results

Linkage analysis

Analysis to investigate the segregation ratio of markers showed that six markers deviated significantly from the expected segregation ratios; these markers were excluded from any further analysis. It should be noted that there was no evidence to support a genuine segregation distortion as the number of markers significantly different from the expected segregation ratios were less than would be expected by chance in a data set of this size. The remaining marker information available was 100 codominant and 47 dominant scored RFLP markers. The ratio of the two alternative dominant marker types was 26: 21. Using a LOD score of 4 and a recombination fraction of 0.49 a total of 22 linkage groups giving a total map length of 852 cM was obtained (Fig. 1). The

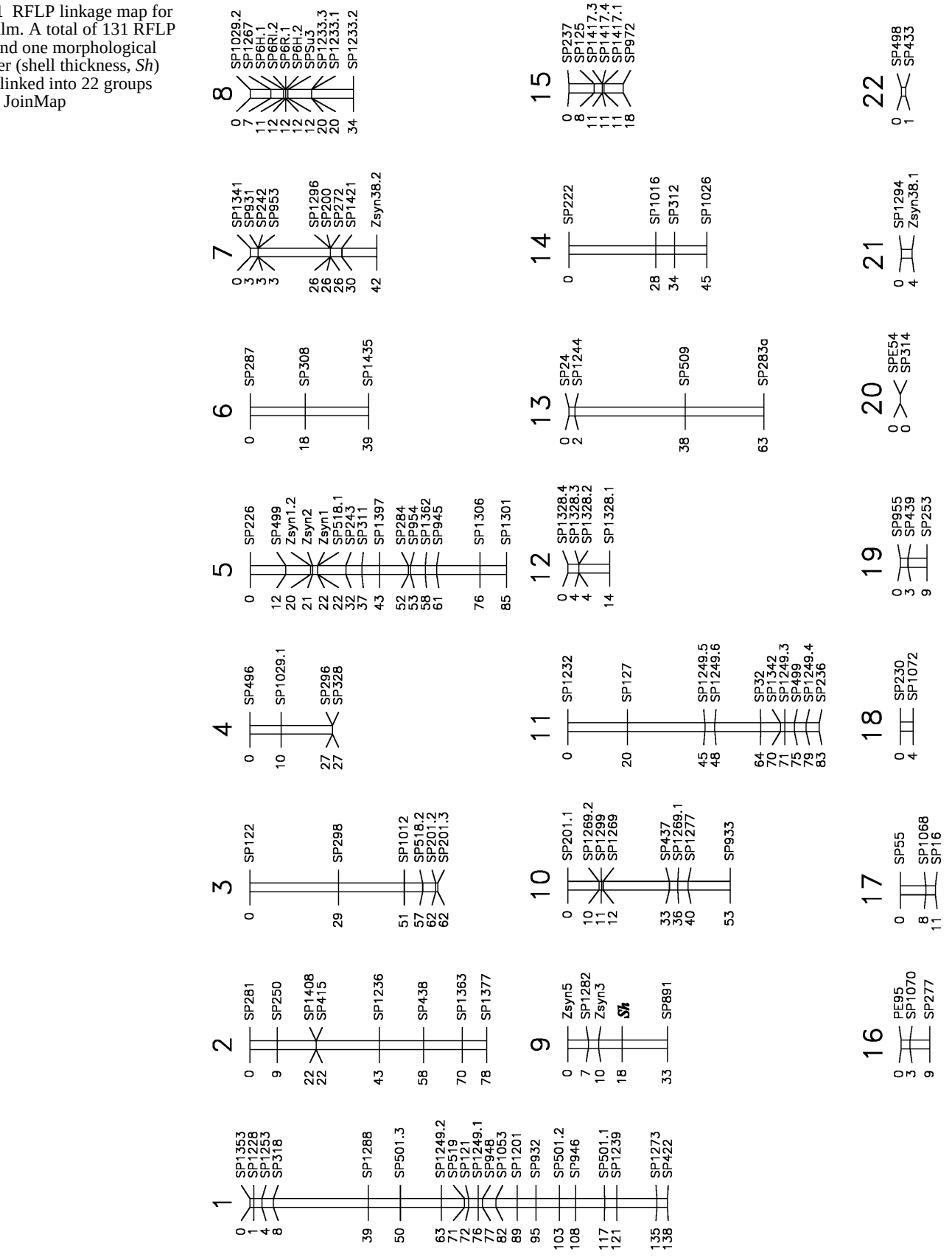

Fig. 1 RFLP linkage map for oil palm. A total of 131 RFLP loci and one morphological marker (shell thickness, *Sh*) were linked into 22 groups using JoinMap

Table 1 Summary statistics for production data used in QTL mapping analysis. The estimates shown are for raw means or where appropriate, least square means

Trait	Abbreviation	Mean	SD[a]	CV %[b]
Bunch number palm^{-1} year^{-1}	BN	37.1	13.9	31.6
Bunch weight (kg)	BWt	3.74	1.35	36.2
Fruit per bunch (%)	FB	53.7	8.60 (7.65)	16.1 (14.2)
Fruit weight (g)	FWt	8.31	3.24	31.2
Fresh Fruit bunch yield (kg)	FFB	151.6	77.8	45.2
Kernel : fruit (%)	KF	9.21	2.53	27.4
Shell : fruit (%)	SF	12.8	9.74 (2.24)	76.8 (17.5)
Mesocarp : fruit (%)	MF	71.7	11.2 (5.54)	15.7 (7.75)
Oil : bunch (%)	OB	18.7	3.87 (3.71)	20.7 (19.8)
Height (m)	Ht	0.60	0.21	28.8
Leaf area (m^2)	LeafA	3.80	1.05	27.6
Rachis length (m)	Rach	3.73	0.50	13.4
Petiole cross section (cm^2)	PCS	12.6	2.65	21.0

[a] Standard deviation
[b] Percentage coefficients of variation; residual standard deviations and the resultant coefficients of variation are shown in parenthesis where appropriate

Table 2 Marker genotype associated effects estimated for unmapped marker loci; standard errors shown in parethesis

Trait	Marker	Marker genotype[a]										% Var[b]	pr(F)[c]
		AA		Aa		aa		aa + Aa		AA + Aa			
BN	Z962					28.8	(3.73)			44.36	(2.97)	36.0	0.005
BN	ZSy38.3	35.6	(4.14)					47.5	(3.04)			18.9	0.032
FFB	Z962					95.6	(23.7)			199.8	(18.9)	38.8	0.003
FFB	ZSy38.3	125.5	(22.7)					218.8	(16.6)			34.4	0.004
KF	Z316	9.00	(0.71)	8.66	(0.62)	13.6	(1.82)					13.1	0.050
KF	SP963	10.6	(0.64)	8.99	(0.53)	8.25	(0.68)					8.8	0.043
OB	Z316	20.1	(1.00)	18.4	(0.88)	13.3	(2.56)					13.0	0.050
OB	SP963	17.7	(0.92)	18.2	(0.76)	20.9	(0.99)					8.9	0.042
LeafA	Z989.1	5.62	(0.85)	4.13	(0.19)	3.29	(0.38)					18.4	0.037
Rach	Z1233	4.51	(0.23)	4.00	(0.18)	3.71	(0.13)					43.2	0.043

[a] aa, Aa and AA represent co-dominant marker-associated effects; dominant marker-associated effects are shown by aa and (AA+Aa) or AA and (aa+Aa)
[b] The proportion of total variance explained by the marker
[c] The probability associated with the marker-QTL effect

linkage analysis produced a total of 18 linkage groups of three or more loci and four groups of two markers; 15 markers remained unmapped.

Phenotypic data

The phenotypic data were initially analysed to determine the normality of data; the results indicated that no measurements deviated significantly from a normal distribution (P>0.05). The influence of the shell thickness genotype (dura or tenera) was determined using ANOVA and was found to have a significant effect (P<0.05) on the phenotypic traits of FB, SF, MF and OB. As a result, corrected values for these traits were used in interval mapping analysis. The mean trait values, standard deviations, coefficients of variation and least squares means, where appropriate, are shown in Table 1. The traits PCS and Rach, and FB, SF, MF and OB, corrected for the shell-thickness genotypes, were associated with coefficients of variation between 8 and 21%. The remaining traits were associated with CVs within the range of 27 to 45%, indicating a high degree of variation in the population for these traits.

Single-marker analysis of unmapped loci

Single-marker analysis was carried out to allow an estimation of the QTL effects associated with the unmapped marker loci. Significant (P<0.05) marker-associated effects were estimated with the traits BN, FFB, KF, OB, Rach and LeafA (Table 2). The single-marker results showed the phenotypic data for BN and FFB, and KF and OB, to be associated with the same markers; however, this is not unexpected as the correlations between these pairs of traits are r=0.85 (P<0.001) and r=–0.44 (P<0.001) respectively. The percentage total trait variance explained by these markers ranges between approximately 8 and 19% for the traits KF, OB, LeafA and BN (marker Zsyn38.3 only). The remaining marker-associated QTL effects account for between 34 and 43% of the total variance. Highly significant (P<0.005) marker associated effects were detected for the traits FFB and BN (marker SP962 only). However, the results obtained from the single-marker analysis should be interpreted with caution. An appropriate threshold for single-marker analysis becomes problematic when multiple tests are performed and, therefore, only the most highly significant (P<0.01) tests should be regarded as indicative of marker-associated QTL effects. The interpretation of the results of the single-marker analysis will be discussed in greater depth later.

Table 3 QTLs for production traits found to be significant at the empirical genome-wide mapping threshold; LOD 4.2 in bold and LOD 3.4 in normal-type face

Trait	Group	Closest marker	Position[a] (cM)	LOD score	Additive	Dominance	% Var [b]
FWt	16	PE95	2	3.9	0.58	–2.64	21.5
KF	**11**	**SP1342**	**70**	**5.7**	**–1.22**	**–1.46**	**8.5**
SF	3	SP298	47	3.4	–2.39	–0.01	44.0
MF	11	SP1342	71	3.5	1.28	3.05	12.3
Rach	14	SP1016	28	3.5	0.37	–0.43	23.8
PCS	5	SP284	47	3.9	–0.96	–1.64	8.2

[a] Estimated position from first marker on linkage group
[b] The proportion of total variance explained by the QTL

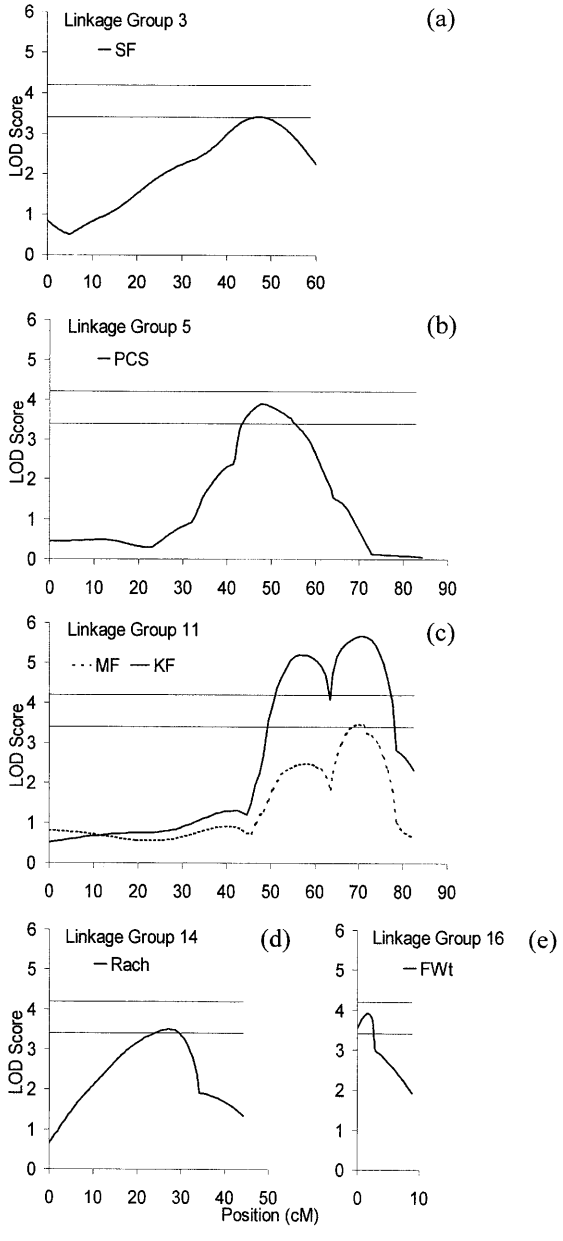

Fig. 2a–e Likelihood profiles for QTLs found to be significant at the genome-wide empirical significant thresholds; $P<0.01$ LOD 4.2 and $P<0.05$ LOD 3.4. Parts *a–e* show the likelihood profiles for linkage groups 3, 5, 11, 14 and 16 respectively

Interval mapping

Using the estimated genome-wide empirical LOD score thresholds of 4.2 for $P<0.01$, and 3.4 for $P<0.05$, a total of six QTLs were detected. One highly significant ($P<0.01$) association was found on linkage group 11 and five QTLs at the $P<0.05$ level (Fig. 2 and Table 3). No QTLs were found to be significant at the suggestive linkage threshold level ($P<0.10$). The QTLs for MF and KF mapped on linkage group 11 both show similar-shaped likelihood profiles suggesting that the same QTLs may be influencing both traits. This hypothesis is supported by the highly significant ($P<0.005$) phenotypic correlation between the traits $r=–0.417$. Despite the double-peaked likelihood profile obtained for both traits it is not possible to conclude that there are two QTLs present. The low probability of recombination events between the two peaks means that it is not possible to distinguish between two linked QTLs and a single QTL. The shape of the likelihood profile therefore suggests an estimated QTL position encompassing the two peaks. The single-QTL hypothesis is supported by the 95% confidence interval estimated using the one-LOD drop-off method (Lander and Botstein 1989) (data not shown). The estimated QTL position was taken at the higher of the two peaks (Table 3). The QTLs mapped on linkage group 11 explain 12.3 and 17.9% of the total variance of the traits MF and KF respectively (Table 3). The direction of QTL effects indicates the influence of each of the two possible allelic types coming from the selfed parent. Unfortunately, the origins of the two alternative allelic types are not known, as marker information from the parents of the selfed palm (A137/30) was not available. As a result, the positive or negative additive effects can only be used to make comparisons of the direction of QTL effects between different loci and traits. The additive QTL effects are in different directions in KF and MF, as predicted by the negative phenotypic correlation between these traits. The dominance deviations associated with the QTLs are greater than the additive effects and in the same direction as the additive effect.

Quantitative trait loci were detected at the $P<0.05$ significance level for PCS, Rach and FWt on linkage groups 5, 14 and 16 respectively. The PCS QTL explains 8.2% of the total phenotypic variance and the estimates of additive effects and dominance deviations are both

Table 4 Putative QTLs found to be significant at the empirical chromosome-wide significance threshold; LOD 2.6 for $P<0.05$

Trait	Group	Closest Marker	Position [a] (cM)	LOD score	Additive	Dominance	% Var[b]
BN	4	SP496	0	2.7	6.52	–12.64	19.7
BWt	14	SP1016	32	2.7	0.86	–1.23	19.1
FFB	4	SP1029.1	26	2.6	–25.69	–25.72	5.5
KF	5	SP243	64	2.7	1.35	–0.71	6.7
MF	15	SP237	4	3.0	–4.89	1.81	13.8
OB	7	Z953	2	2.6	3.08	–0.98	21.5
LeafA	5	SP1397	53	2.7	0.08	–0.76	13.9

[a] Estimated position from first marker on linkage group
[b] The proportion of total variance explained by the QTL

negative. The QTLs detected for Rachis length and fruit weight explain a large proportion of the total phenotypic variance, being 23.8% and 21.5% respectively. Analysis of the shell-to-fruit ratio indicated a QTL on linkage group 3. It should be noted that the data for the shell-to-fruit ratio was corrected for the shell-thickness gene on linkage group 4. The QTL for the corrected shell on linkage group 3 explains a large proportion (44%) of the total corrected trait variance.

In addition to generating empirical genome-wide thresholds for QTL mapping, empirical chromosome-wide thresholds were calculated. Using an empirical chromosomal threshold does not indicate a significant QTL on a genome wide basis but points to a certain level of association between markers and traits. The QTLs detected at the $P<0.05$ chromosomal threshold (LOD=2.6; SD=0.31) are presented in Table 4. The analysis of FFB indicated a putative QTL on linkage group 13 explaining 5.5% of the total variance. QTLs were also detected for the traits BN, BWt, MF, OB and LeafA. These QTLs explained a moderate proportion of the total phenotypic variation between 14 and 22%.

Discussion

Comparison of the oil palm marker map generated by JoinMap 2.0 (Stam 1993) analysis is in close agreement with the previously reported oil palm map (Mayes et al. 1997) where MAPMAKER 2.0 was used (Lander et al. 1987). A total of 49 additional marker loci have been added to the map resulting in an improvement of map resolution from 24 linkage groups (Mayes et al. 1997) to 22 linkage groups (n=16). The addition of extra markers, and re-checking of the previous marker data, has resulted in a reduction of total map length from 860 cM estimated of Mayes et al. (1997) to 852 cM in the present study. There are a number of minor discrepancies between the RFLP map by Mayes et al. (1997) and the map presented in this study. The top of linkage group 1 from Mayes et al. (1997) has joined with linkage group 15 (Mayes et al. 1997) to form linkage group 13 in the current study. This fragmentation alone has resulted in a reduction of map length by 50 cM (the distance from marker SP283 to SP1053; Mayes et al. 1997) and, therefore, may account substantially for the reduction in map length. In addition to the fragmentation of the top of linkage group 1 there have been a number of groups joining together. For example, groups 1 and 2 from Mayes et al. (1997) joined to form linkage group 1 in the current study, and groups 11 and 19, and groups 12 and 14 from Mayes et al. (1997), have joined to form linkage groups 10 and 11 respectively in the current study. Clearly, additional markers are required to resolve the oil palm marker map such that each chromosome is represented by a single linkage group. A multiplexed marker system, for example AFLPs, would offer a rapid generation of marker data around the RFLP framework map.

Exploration of the phenotypic data suggested that all pisifera (homozygous shell-less type) should be discarded from any further QTL analysis. The pisifera palms are typically female-infertile and consequently more energy is available for vegetative growth. Analysis of the vegetative traits indicated unequal variances between the three shell-thickness genotype classes, with pisifera individuals being associated with lower levels of phenotypic variance. As a result of the lack of bunch data and the unequal variances between the shell-thickness classes, data from the pisifera palms were not included in the QTL analysis. Additional QTL mapping projects in oil palm would have to consider the problems associated with the homozygous shell-less genotype. If future QTL mapping projects were to be based on a selfed (tenera) F_2 population, the population size would have to be large to allow for the exclusion of pisifera individuals. Alternatively, a design based on a population derived from dura×tenera or dura×dura crosses would not give rise to pisifera individuals.

The influence of shell-thickness genotype on the dura and tenera palms was found to be significant for FB, SF, MF and OB. The use of corrected-trait values enables variance due to the shell-thickness genotype to be accounted for in the model. However, it should be noted that the use of a two-stage analysis (stage 1 correcting trait values for shell thickness genotype, and stage 2 estimating QTL effects) could result in the erroneous allocation of phenotypic variance.

Single-marker analysis was carried out for all unmapped marker-trait combinations yielding a total of ten significant ($P<0.05$) marker-QTL effects. As stated above, the formulation of an appropriate threshold for single-marker analysis becomes problematic when multiple tests are performed. In the single-marker analysis a

total of 13 traits and 15 markers were analysed, giving a total of 195 tests; therefore, by chance ten tests might be significant at $P<0.05$ and two significant at $P<0.01$. As a result, only the most-significant ($P<0.01$) marker-associated QTL effects should be considered to indicate a putative QTL. Alternatively, it is possible to apply the permutation method, as described by Churchill and Doerge (1994), to single-marker analysis to determine a genome-wide empirical threshold for single-marker analyses. By permuting the marker information with respect to phenotypic data over a large number of replications, the distribution of the *F* values (obtained using ANOVA), under the null hypothesis that there is no association between the marker information and phenotypic data, can be determined. By combining the resulting F values from all markers, the empirical genome-wide thresholds for single-marker analyses can be determined. This approach would be particularly useful if QTL mapping analysis was based solely on single-marker analysis.

The single-marker analysis yielded a total of three highly significant ($P\leq0.005$) marker-associated effects explaining between 34 and 39% of the total phenotypic variation for traits BN and FFB. The unmapped marker SP962 was associated with highly significant effects on both BN and FFB. The common marker-associated QTL effect observed for BN and FFB is likely to be explained by FFB being a product of the total bunch number and bunch weight, and this association is reflected in the strong correlation between these two traits ($r=0.85$). As a result of the relationship between BN and FFB, any QTL effects observed in BN are likely to be detected in FFB unless there is a compensating effect associated with BWt. Taking account of the predicted number of ”chance” marker-associated QTL effects, it is difficult to draw any conclusions from the single-marker results to-date. However, it should be noted that population UV101 is comparatively small, and the analysis of further populations, derived from the same cross and planted in multiple environments, may yield more conclusive results.

The QTL-mapping package QTL Cartographer (Basten et al. 1994, 1999) enabled co-factors to be included in the QTL-mapping model. The inclusion of co-factors is used to eliminate background genetic noise (QTLs elsewhere on the genome) and neutralise the effects of linked QTLs (from outside the window parameter) resulting in an increase in the power and reduction of interference due to linked QTLs (Jansen 1993, 1996; Zeng 1994). The interval mapping of production traits yielded a total of six significant ($P<0.05$) associations at the genome-wide significance threshold. The QTLs mapped to linkage group 11 influence the negatively correlated traits MF and KF explaining 12.3 and 8.5% of the total phenotypic variance, respectively. The likelihood profiles for these two traits suggest a single QTL having an effect on both traits, rather than two different QTLs. Of the remaining QTLs found to be significant at the genome-wide threshold, the FWt QTL on linkage group 16 influencing 21.5% of the total phenotypic variance would be of most-commercial interest.

The QTLs mapped during the single-marker analysis and interval-mapping analysis explained between 8.2% and 44.0% of the total phenotypic variation, with an average of 27% and 19% for the single-marker and interval-mapping analysis respectively. It is likely that the differences in total phenotypic variation explained are due to the lower stringency associated with the single-marker analysis (as discussed above) rather than the greater power associated with the single-marker analysis (where power is the probability of detecting QTLs where QTLs are present).

In addition to the genome-wide empirical thresholds, chromosome-wide thresholds were generated. Because the accepted experimental error rate is 5%, the term ”suggestive linkage” might be used for a chromosome-wide significance level of 5% (Van Ooijen 1999). The empirical chromosome-wide LOD threshold estimated in this experiment (LOD=2.6) is in close agreement with the proposed comparative-mapping suggestive-linkage threshold of LOD=2.7 for F_2 populations (Van Ooijen 1999). The objective of the suggestive-linkage threshold is to give an indication of the putative QTLs which failed to reach the stringent empirical genome-wide thresholds. One of the most effective means of verifying the presence and location of QTLs is for the QTLs to be isolated in replicate experiments. The publication of these ”suggestive” marker QTL associations makes these data available to other workers in oil palm genetics. Additionally, the population size available in this study was comparatively small, which has a dramatic effect on the power of the QTL mapping study (Darvasi et al. 1993). The putative QTLs detected at the suggestive linkage threshold include a number of QTLs for traits of extreme economic importance, most notably BN, BWt and OB, all influencing approximately 20% of the total phenotypic variance.

The ultimate objective of mapping QTLs in commercial populations, such as oil palm, is to enable new selection strategies such as marker-assisted selection (MAS) to be established. Marker-assisted selection can be used in two distinct ways: to increase speed or improve efficiency for objectives which could be achieved eventually by conventional breeding, or for objectives which could not be approached in any other way. The latter includes such things as selection for non-expressed characters (e.g. disease resistance in a disease-free area) and ‘pyramiding’ of resistance genes. The QTLs considered in our work all fit into the first category. In this case, the relative time-scales and the costs of conventional and marker-assisted selection, are important in judging the value of MAS.

Xie and Xu (1998a) considered the advantages of two-stage selection, using markers at the first stage to reduce the scale and costs of subsequent phenotypic screening. In terms of selection gain, the benefit of including markers is greater the lower the heritability of the character, and also the greater the proportion (*p*) of

variation associated with the markers. Of all the traits of economic importance in oil palm, MAS could be most effective for FFB yield, which generally has a fairly low heritability (Breure and Corley 1983). Marker-only selection may be useful if p is high, h^2 is low, and the cost of phenotypic recording is high. The cost of recording FFB yield for 5 years is high, so a search for further markers associated with FFB yield, to increase p, could be worthwhile. Xie and Xu (1998a) estimated that in most situations the selection gain per unit cost will argue against MAS, until progress is made in reducing the cost of molecular-marker assays. However, when the cost of phenotypic selection exceeds that of the marker assay, MAS could be effective. For FFB yield, the cost of recording is roughly four times the current cost of assaying for a single RFLP marker.

Marker-assisted selection may also be used to shorten the generation time. Hospital et al. (1997) confirmed other observations that the value of MAS selection diminishes over successive generations of selection, as linkages break down. However, they found that a strategy of one generation of phenotypic scoring, to establish linkages with QTLs, followed by two generations of marker selection without phenotypic scoring, is more efficient, in terms of genetic gain, than phenotypic selection alone, and also has the potential to be faster. This would be particularly useful with a long-generation crop such as oil palm. After one generation of normal field recording, with marker work to find QTLs, one might then go through two generations using MAS in the nursery, and making crosses as soon as flowering starts, at about 3 years after germination. This would reduce the mean generation time from 10 years to about 5.

Hospital and Charcosset (1997) discuss the introgression of one or several QTLs into a different background. For a single QTL, quite small minimum population sizes at each generation (often less than ten, depending on the position of the markers relative to the QTL) are sufficient to ensure that at least one individual has the required genotype; with more than one QTL, numbers are larger. Such a programme would be quite feasible with oil palm, and the approach might be useful to incorporate novel characteristics from the American oil palm *Elaeis oleifera* (Hardon 1969) into a high yielding *E. guineensis* background. However, a number of factors can influence the effectiveness of marker assisted introgression. The confidence interval (CI) around the QTL will influence the probability of introgressing unfavourable genes due to linkage, whereas a large CI will increase the probability of introducing unfavourable loci from the donor parent (*E. oleifera*). In addition, the effect of linkage drag will influence the proportion of donor parent genetic material in the recurrent backcross (Stam and Zeven 1981). Using additional markers flanking the QTL to accelerate recurrent parent (*E. guineensis*) genome recovery (Visscher 1999) can reduce the effect of linkage drag, but may influence the cost of the marker-assisted introgression.

From the viewpoint of commercial seed production, which involves making crosses between phenotypically selected duras and progeny-tested pisiferas to produce thin-shelled teneras, MAS might be used in several ways. Dura selection would be more efficient if markers were used to identify individuals carrying as many favourable alleles as possible. The individuals with the best genotypes could be intercrossed or selfed to create parental populations homozygous for all favourable alleles. Selection for desired marker genotypes could be carried out as soon as flowering started while still in the nursery, therefore reducing the generation interval and enabling the field planting of only selected palms. Pisiferas, which are typically female-sterile, are normally selected after progeny testing; this adds 6 or 7 years to the cycle time. Individuals for testing are chosen at random, from segregating families in which the duras and teneras are superior. MAS is unlikely to be useful for family selection except when family size is very small and p is large (Xie and Xu 1998b), but could be used after phenotypic selection of the better families to identify pisiferas carrying favourable alleles, based on QTL analysis of their dura and tenera sibs. This would certainly be more efficient than random selection, and might even eliminate the need for progeny testing altogether.

Arguably, the most important results to come from this study are that initial QTL analysis has yielded a number of significant and suggestive QTLs, some of which influence a large proportion of the total phenotypic variance (in excess of 10%). The ultimate objective of QTL mapping within oil palm populations is to provide a tool, in the form of marker-assisted selection (MAS), to improve the efficiency of selection. This study is a confirmation that QTLs can be detected even in comparatively small populations.

Acknowledgements We acknowledge Dr. Palat of the Univanich Palm Oil Company for providing the phenotypic data and thank the two anonymous referees for their comments.

References

Basten C, Weir B, Zeng Z-B (1994) Zmap-a QTL cartographer. Proceed 5th World Congr Genet Applied to Livestock Prod, Guelph, Canada 22:65–66

Basten CJ, Weir BS, Zeng Z-B (1999) QTL Cartographer. Version 1.13. Department of Statistics, North Carolina State University, Raleigh, North Carolina

Breure CJ, Corley RHV (1983) Selection of oil palms for high density planting. Euphytica 32:177–186

Churchill GA, Doerge RW (1994) Empirical threshold values for quantitative trait mapping. Genetics 138:963–971

Corley RHV, Lee CH (1992) The physiological basis for genetic improvement of oil palm in Malaysia. Euphytica 60:179–184

Corley RHV (1983) Potential productivity of tropical perennial crops. Exp Agric 19:217–237

Darvasi A, Weinreb A, Minke V, Weller JI, Soller M (1993) Detecting marker-QTL linkage and estimating QTL gene effects and map locations using a saturated genetic map. Genetics 134:943–951

Genstat 5 Committee (1993) Genstat 5 release 3 reference manual. Clarendon Press, Oxford

Hardon JJ (1969) Interspecific hybrids in the genus *Elaeis*. II. Vegetative growth and yield of F_1 hybrids *E. guineensis*×*E. oleifera*. Euphytica 18:380–388

Hartley C (1988) The oil palm, 3rd edn. Longmans, London

Hospital F, Charcosset A (1997) Marker-assisted introgression of quantitative trait loci. Genetics 147:1469–1485

Hospital F, Moreau L, Lacoudre F, Charcosset A, Gallais A (1997) More on the efficiency of marker-assisted selection. Theor Appl Genet 95:1181–1189

Jack PL, Dimitrijevic TAF, Mayes S (1995) Assessment of nuclear, mitochondrial and chloroplast RFLP markers in oil palm (*Elaeis guineensis* Jacq). Theor Appl Genet 90:643–649

Jansen RC (1993) Interval mapping of multiple quantitative trait loci. Genetics 135:205–211

Jansen RC (1996) A general Monte Carlo method for mapping multiple quantitative trait loci. Genetics 142:305–311

Lander ES, Botstein D (1989) Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121:185–199

Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Linclon SE, Newburg L (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1:174–181

Mayes S, Jack PL, Marshall DF, Corley RHV (1997) Construction of a RFLP genetic linkage map for oil palm (*Elaeis guineensis* Jacq). Genome 40:116–122

Mohan M, Nair S, Bhagwat A, Krishna T, Yano M, Bhatia C, Sasaki T (1997) Genome mapping, molecular markers and marker-assisted selection in crop plants. Molecular Breeding 3:87–103

Stam P (1993) Construction of integrated genetic linkage maps by means of a new computer package: JOINMAP. Plant J 3:739–744

Stam P, Zeven AC (1981) The theoretical proportion of the donor genome in near-isogenic lines of self-fertilizers bred by back-crossing. Euphytica 30:227–238

Van Ooijen JW (1999) LOD significance thresholds for QTL analysis in experimental populations of diploid species. Heredity 83:613–624

Visscher, PM (1999) Speed Congenics: accelerated genome recovery using genetic markers. Genet Res 74:81–85

Wonkyi-Appiah JB (1987) Genetic control of fertility in oil palm (*Elaeis guineensis* Jacq). Euphytica 36:505–511

Xie CQ, Xu SZ (1998a) Efficiency of multistage marker-assisted selection in the improvement of multiple quantitative traits. Heredity 80:489–498

Xie CQ, Xu SZ (1998b) Strategies of marker-aided recurrent selection. Crop Sci 38:1526–1535

Zeng Z-B (1994) Precision mapping of Quantitative trait loci. Genetics 136:1457–1468