REVEALING QTLs ASSOCIATED WITH FATTY ACID COMPOSITION AND COMPACTNESS ON AN INTEGRATED LINKAGE MAP OF OIL PALM INTERSPECIFIC BACKCROSS 2 (BC2)

CURRENT STATUS: UNDER REVISION

BMC Plant Biology  BMC Series

Zulkifli Yaakub
Malaysian Palm Oil Board
✉ zulkifliyaakub80@gmail.com Corresponding Author
ORCiD: https://orcid.org/0000-0001-9174-271X

Katialisa Kamaruddin
Malaysian Palm Oil Board

Rajinder Singh
Malaysian Palm Oil Board

Suzana Mustafa
Malaysian Palm Oil Board

Marhalil Marjuni
Malaysian Palm Oil Board

Ngoot-Chin Ting
Malaysian Palm Oil Board

Mohd Din Amiruddin
Malaysian Palm Oil Board

Low Eng-Ti Leslie
Malaysian Palm Oil Board

Ooi Leslie Cheng-Li
Malaysian Palm Oil Board

Kandha Sritharan
United Plantation Berhad
| Authors                        | Institutions                                      |
|-------------------------------|---------------------------------------------------|
| Rajanaidu Nookiah             | Malaysian Palm Oil Board                          |
| Johannes Jansen               | Wageningen University and Research Wageningen Plant Research |
| Meilina Ong Abdullah          | Malaysian Palm Oil Board                          |

**DOI:**
10.21203/rs.2.12850/v2

**SUBJECT AREAS**
*Plant Molecular Biology and Genetics*

**KEYWORDS**
*Oil palm, interspecific hybrids, QTL, fatty acid composition, compactness*
Abstract

Background Molecular breeding has opened new avenues for crop improvement with the potential for faster progress. As oil palm is the major producer of vegetable oil in the world, its improvement, such as developing compact planting materials and altering its oil’s fatty acid composition for wider application, is important.

Results This study sought to identify the QTLs associated with fatty acid composition and vegetative traits for compactness in the crop. It integrated two separate interspecific backcross two (BC2) mapping populations to improve the genetic resolution and evaluate the consistency of the QTLs identified. A total of 1,963 markers (1,814 SNPs and 149 SSRs) spanning a total map length of 1793 cM were integrated into a consensus map. For the first time QTLs associated with vegetative parameters were identified in an interspecific hybrid population, and carotene content apart from these associated with fatty acid composition. The QTL analysis observed 9, 4 and 8 genomic loci associated significantly with fatty acids, carotene content and compactness, respectively.

Conclusions Major genomic region influencing the parameters associated with compactness and fatty acid composition was identified across separate populations using two different methods for QTL detection. Other significant locus influencing compactness, carotene content and FAC were identified either being common to both populations or specific to a particular genetic background. It is hoped that the QTLs identified will be useful tools for marker-assisted selection and accelerate the process of identifying desirable genotypes for breeding.

Background

Global palm oil production now stands at over 65 million tonnes/year, or 34% of the world’s vegetable oil production [1,2]. The oil palm commonly planted commercially is the African origin (*Elaeis guineensis*). It is the most productive vegetable oil crop, with commercial oil yields of ~4 tonnes/ha/yr [3], but up to 13 tonnes/ha/yr have already been achieved in some breeding trials [4]. Although yield is the primary target, there is also need for disease resistance and tailored fatty acid composition (FAC) for the multivariated uses of the oil.

The oil palm, unfortunately has only a single growing point, and continually grows taller making it
more and more difficult to harvest, until well-nigh impossible. Having shorter (dwarf) palms will extend its economic life, with ensuing lower (labour) cost of harvesting. The height increment of current commercial Dura x Pisifera palms is 40 – 75 cm/yr [5], and palms can reach 15-18 meters in height, up to 30 meters in a dense forest [6]. A breeding programme for dwarf palms was initiated by Elmina Estate in Malaysia with selfing of the short and famous Malayan Dumpy dura palm E206 [7]. More recently, MPOB, from its Nigerian prospection, identified Population 12 not only for its dwarfness, but also for its high bunch number, good yield, and desirable fruit characteristics [8]. These palms, when crossed with elite materials, are 5-10% shorter than the standard crosses [9]. In improving the palm, progress can be speeded up by the use of biotechnology, such as using molecular markers to screen for desired traits. Recently, quantitative trait loci (QTL) associated with trunk height and bunch weight identified using a linkage map containing 1,085 single nucleotide polymorphism (SNPs) [10]. A study [11], also constructed a consensus linkage map for a population of oil palm using simple sequence repeats (SSRs) and SNPs, and identified a major QTL for stem height on LG 5. In other study using association mapping, a SNP marker SNPG00006 Fatl was observed to be significantly associated with height increment (P < 0.05) [12].

All the above work was carried out on E. guineensis, the African oil palm. There is, however, a second oil palm, Elaeis oleifera – the American oil palm. Although not much commercially planted and produces very low yields, it has several interesting characteristics, such as shortness, less saturated oil and disease resistance, which may be introgressed to improve E. guineensis. Interspecific hybrids of E. guineensis and E. oleifera have already been made (F1) – they are shorter, but the yield is still very low [13]. There is also a lack of pollen production (as demonstrated by pure E. oleifera), and assisted pollination is even required to produce the low yield. Backcrossing to E. guineensis will quickly improve the yield, but just as quickly lose the desirable E. oleifera characteristics. In other words, the desire to improve E. guineensis by introgressing E. oleifera traits has largely come to nought in a painstakingly slow and costly process.

But that was in the past using conventional breeding. Now, with DNA-based markers there is promise of more efficient crop improvement by introgressing only the specific genes wanted, rather than half
the whole genome of donor palms for some required genes in conventional hybridization. The availability of dense genetic maps for both *E. guineensis* and interspecific hybrid populations [14,15,16,17], as well as markers linked to important quantitative traits such as yield, vegetative measurement and FAC [10,15,18,19,20], have already provided the groundwork for this work. However, and interestingly, no QTL associated with height increment has yet been reported for the interspecific hybrid.

Compact palms with shorter trunks and fronds can be planted at a higher density than the current 148/ha. If the individual palm yields can be maintained, then the yield per unit area will increase [21,22]. In South America, *E. guineensis*-based compact palms [23] have already been developed by multiple rounds of backcrossing OxG hybrids to *E. guineensis* [24], and the outstanding palms cloned for planting [25]. In 2012, it has been reported that the development of an OxG hybrid, known as COMPACT palms, with low height increment (<40 cm/year) and shorter fronds (~6.5 meter), allowing high density planting (180-200 palms/ha) [26]. Backcrossing the COMPACT palms with Deli, Ghana, and Nigeria *E. guineensis* produced fronds of 6.6–6.9 meters which reduced the density to 170 palms/ha.

Interspecific hybrids and their backcrosses also have desirable FACs in their oils. The genomic regions associated with various FACs in an OxG interspecific hybrid [14,20] and interspecific backcross one (BC₁) [16] mapping populations were identified via conventional QTL analysis. A number of these QTLs were validated across interspecific backcross two (BC₂) mapping populations [20]. One of the major restrictions in associating markers to traits in oil palm is the size of the mapping populations employed in such studies. Oil palm breeding trials generally consist of 64 palms per progeny, which is small for effective genetic mapping and QTL analysis. However, a study had shown it possible to develop high quality integrated maps of multi parental populations, which can enhance QTL discovery [15]. This study searched for the QTLs associated with vegetative traits and FAC in two interspecific BC₂ mapping populations - characters important for developing compact palms with higher unsaturated mesocarp oil. This study also in part looked at validating QTLs linked to FAC identified in
previous studies. It integrated two separates interspecific BC$_2$ mapping populations to enhance the
genetic resolution and observe the consistency of the QTLs detected, apart from identifying
population specific genomic regions influencing the traits of interest.

**Results**

**Traits of Interest**

The vegetative parameters, carotene content and FACs, observed in the 2.6-1 and 2.6-5 mapping
families are summarized in Table 1. All the traits showed wide segregation and followed normal
distribution (P<0.05, Shapiro-Wilk test) in the 2.6-5 population. However, for population 2.6-1, rachis
length (RL), height increment (HI), iodine value (IV) and oleic acid (C18:1) were not normally
distributed. The means for RL, petiole cross section (PCS) and IV were slightly higher in 2.6-1, but HI
and carotene content were higher in 2.6-5. However, both RL and HI were considerably lower than in
the commercial DxP, whereby, they are generally >5 and >0.45 meters, respectively [27]. For FAC,
population 2.6-1 had higher stearic (C18:0), oleic (C18:1) and linoleic (C18:2) acids, whereas palmitic
(C16:0) was slightly higher (35.33%) in the 2.6-5 population.

The relationships between the individual fatty acids were evaluated using Pearson’s correlation
analysis and consistent results were obtained for both mapping families (Table 2). The most abundant
saturated fatty acid, C16:0, was negatively correlated with the unsaturated fatty acids (C18:1 and
C18:2). A negative correlation was also observed between C18:0 and C16:0. As expected, IV, as an
indicator for unsaturation, was positively correlated with C18:1 and C18:2, and negatively with C16:0
and C18:0. In addition, correlations were positive between C18:0 and C18:2 and negative between
C18:1 and C18:2. The results corroborated with those from other studies [14,16,20]. Correlation
trends for vegetative parameters were similar in both populations (Table 3). PCS was positively
correlated with HI and RL, while RL and HI appeared to be negatively correlated, although not
significant at P ≤ 0.05.

| Table 1: Summary of vegetative traits and fatty acid composition (FAC) in the 2.6-1 and 2.6-5 BC$_2$ mapping populations |

| Table 2: Pearson’s correlations between individual fatty acids in the 2.6-1 and 2.6-5 mapping populations |
Table 3: Pearson’s correlations between vegetative parameters in the 2.6-1 and 2.6-5 mapping populations

| Population | 2.6-1 | 2.6-5 |
|------------|-------|-------|
|            | RL    | HI    | RL    | HI    |
| HI         | -0.21 | -0.02 | -0.08 | 0.28* |
| PCS        | 0.45* | 0.10  | 0.55* | 0.48* |

*significant at p ≤ 0.05

**BC₂ Consensus Genetic Map**

In this study a total of 4491 SNP markers were utilized to test for polymorphism in both populations, while 515 and 715 SSR markers were screened for informativeness in population 2.6-1 and 2.6-5, respectively. Only polymorphic markers which met the expected segregation ratios at 5% significance
level (p<0.05) and had nearest-neighbor stress value of > 3cM were considered suitable for utilization in construction of genetic map [see Additional file 1]. The consensus genetic map for both families is shown in Figure 1 while Table 4 summarizes the marker compositions and lengths of the individual linkage groups (LGs). The 2.6-1 and 2.6-5 genetic maps had 1,744 and 1,254 markers covering 1,505 and 1,564 cM, respectively. Both maps were integrated into a consensus map [see Additional file 2] of 16 LGs with 1,963 markers (1,814 SNPs and 149 SSRs), spanning a total map length of 1793 cM. The lengths of the individual LGs on the consensus map ranged from 57 - 195 cM, with a mean of 112 cM. The average distance between markers was 0.91 cM. Initially, LG5 in population 2.6-1 and LGs 1 and 15 in population 2.6-5 had two sub-groups. However, these subgroups were successfully integrated into a consensus linkage group.

**QTLs Associated with Vegetative Parameters, Carotene and FAC**

Results of the QTL analysis obtained using Genstat are summarized in Table 5 and Additional File 3, and distribution of QTLs on the respective LGs is shown in Figure 1. The study revealed eight significant QTLs associated with HI, RL and PCS in the integrated genetic map of BC2 interspecific hybrid populations. The three traits in combination contribute to shorter and more compact palms. Interestingly, a single genomic locus in LG4 was associated with two of the traits, namely HI and PCS, while the QTL for the third trait RL, was located in close proximity in the same LG. An additional QTL for PCS was identified in LG4 for population 2.6-5 (Additional file 3), which also appeared on the integrated LG4. The two QTLs for PCS on LG4 were more than 150 cM apart, so they are likely not linked.

In addition, ten other QTLs were identified for traits associated with fatty acid composition (FAC) namely, IV, C16:0, C18:1 and C18:2 content across six LGs. Single QTLs for IV and C16:0 were, as in a previous study [20], located in LG1. Relatively high LOD levels were detected for some of the compactness and FAC traits, namely HI (7.77), PCS (7.70), C18:2 (8.19) and IV (13.02). Interestingly, this study also for the first time revealed four QTLs associated with carotene content across four different LGs.

Table 4: Distribution of markers on the 16 linkage groups (LG) of the BC$_2$ genetic map
| Population | 2.6-1 |       | 2.6-5 |       | Integrated |
|------------|-------|-------|-------|-------|------------|
|            | Map Length (cM) | No. Markers | Map Length (cM) | No. Markers | Map Length (cM) | No. Markers |
| 1          | 115   | 121   | 88+23* | 95+9* | 138         | 146         |
| 2          | 101   | 87    | 97     | 62    | 113         | 107         |
| 3          | 55    | 48    | 58     | 42    | 57          | 54          |
| 4          | 174   | 178   | 212    | 117   | 195         | 202         |
| 5          | 25.2 + 14.8* | 21+17* | 53     | 22    | 100         | 44          |
| 6          | 106   | 129   | 126    | 82    | 122         | 135         |
| 7          | 78    | 173   | 85     | 117   | 113         | 189         |
| 8          | 122   | 132   | 175    | 97    | 152         | 154         |
| 9          | 82    | 56    | 37     | 15    | 83          | 157         |
| 10         | 88    | 129   | 93     | 93    | 116         | 138         |
| 11         | 128   | 166   | 133    | 125   | 145         | 179         |
| 12         | 108   | 130   | 124    | 91    | 124         | 144         |
| 13         | 77    | 127   | 57     | 96    | 80          | 135         |
| 14         | 98    | 77    | 88     | 69    | 99          | 94          |
| 15         | 69    | 119   | 52 + 46* | 79+10* | 98        | 142         |
| 16         | 50    | 34    | 63     | 33    | 58          | 43          |
| **Total**  | 1505  | 1744  | 1564  | 1254  | 1793        | 1963        |

* Sub-groups
Table 5: QTLs associated with compactness traits and FAC identified via Genstat on the interspecific BC2 integrated map

| Trait | Closest marker to QTL peak | LG | Position (cM) | LOD |
|-------|-----------------------------|----|---------------|-----|
| HI    | SNPM00563                  | 4  | 4.27          | 7.77|
|       | SNPM04928                  | 7  | 110.78        | 3.83|
| RL    | SNPM03201                  | 4  | 11.30         | 3.79|
|       | SNPM03772                  | 8  | 92.56         | 3.17|
|       | SNPM03676                  | 11 | 38.69         | 4.17|
| PCS   | SNPM00563                  | 4  | 4.27          | 7.70|
|       | SNPM03375                  | 4  | 192.47        | 3.25|
|       | sEg00213                   | 8  | 139.86        | 3.81|
| C16:0 | SNPM00796                  | 1  | 133.4         | 3.59|
| C18:1 | SNPM02507                  | 4  | 169.17        | 3.44|
|       | SNPM03249                  | 8  | 44.0          | 3.19|
|       | SNPM00274                  | 12 | 31.38         | 5.97|
| C18:2 | SNPM01602                  | 1  | 124.87        | 4.62|
|       | SNPM00249                  | 4  | 3.43          | 8.19|
|       | SNPM01190                  | 15 | 7.03          | 5.14|
| IV    | SNPM01452                  | 1  | 132.53        | 13.02|
|       | SNPM04197                  | 3  | 15.78         | 3.72|
|       | SNPM03285                  | 15 | 98.11         | 5.98|
| Carotene | SNPM02349             | 3  | 4.35          | 5.12|
|       | SNPM00729                  | 4  | 181.83        | 3.32|
|       | SNPM03960                  | 7  | 108.53        | 3.73|
|       | SNPM03921                  | 10 | 77.39         | 3.63|

**Common and Population Specific QTL**

Figure 2A shows a major genomic region in LG4 influencing vegetative parameters namely HI, PCS and RL in the interspecific hybrid populations evaluated in the study. The closest marker to the QTL peak for HI and PCS was the same, namely SNPM00563, both in the two independent maps and the integrated map. Genstat revealed that the QTL related to RL for population 2.6-1 and the integrated map was also detected in the same region, about 7cM away from HI and PCS. Interestingly the QTL for C18:2, one of the most abundant unsaturated fatty acids in oleifera and interspecific hybrids also mapped around the same region for the two independent populations and the integrated maps. The
SNP markers corresponding to the QTLs for HI, PCS, C18:2 and RL were also physically positioned on the genome build, spanning about 3600 kb (Figure 2B). This confirmed that the QTLs influencing the traits were located in close proximity, thus suggesting a major genomic region influencing compactness and unsaturation level in interspecific hybrids. Another QTL for PCS, specific to population 2.6-5, was also located at the other end of LG4.

It was also clear that there were other population specific QTLs [see Additional file 3], where for example, a significant QTL for RL was located in LG8 for population 2.6-1 as well as the integrated map. Similarly, there were population specific QTLs for HI, PCS and fatty acids compositions (C18:1, C18:2 and IV) reflecting the diversity of the 2 populations used in the study. The QTLs linked to carotene content were also specific to population 2-6-5 (also in the integrated map), likely because the variation for the trait was higher in population 2-6-5 (40.27%) than that observed in 2-6-1 (33.27%) (Table 1).

The integrated map also proved useful in detecting QTL not detected in an individual population, especially for FAC. QTL was revealed for C18:2 in the integrated map (confirmed as minor QTL in MapQTL analysis below in Table 6), and correspond to the same linkage group previously linked to FAC in other studies [17]. Although a QTL was associated with RL in LG13 in the population 2.6-5 (Additional File 3), this was not reproducible in the integrated map. Interestingly, the integrated map while maintaining the population specific QTL for RL in LG4, revealed a new QTL for RL in LG11, which was not detected in the individual populations. The QTL in LG11 was also confirmed as a minor QTL based on analysis carried out using MapQTL (Table 6).

**Further Confirmation of QTLs via Interval Mapping (IM)**

The interval mapping approach implemented via MapQTL was used to further validate the QTLs identified earlier. The analysis was restricted to the integrated map and the results are presented in Table 6 and Additional file 4. Generally, the QTLs detected earlier (with one exception), were also revealed as influencing the specific traits and could be further divided into major QTLs (significant at genome-wide) and suggestive minor QTLs (significant at chromosome-wide). The QTLs linked to HI in LG4 and LG7 as well those linked to PCS in LG4 and LG8 were major QTLs, significant at genome-wide
level. The genomic region associated with RL in LG8 was also a major QTL. However, the QTLs linked to RL via Genstat in LG04 and LG11, were considered as minor QTLs (significant at chromosome wide) in the MapQTL analysis. In terms of FAC, QTLs linked to C18:1 (LG04 and LG08), C18:2 (LG4), C16:0 (LG1) and IV (LG1 and LG15) were also significant at genome-wide, as were the QTLs linked to carotene content in LG03, LG04, LG07 and LG10. However, the QTLs linked to C18:2 in LG1 and IV (LG03) were only significant at chromosome-wide, indicating they were minor (suggestive) QTLs pointing to a region that may be regulating the traits concerned. The only exception was the QTL linked to C18:1 in LG12, which was not significant in the MapQTL analysis, and as such, not considered as representing a genomic region associated with the trait. A point to emphasize is that the QTLs revealed in integrated maps, but not in the individual maps, particularly C18:2 (LG1) and RL (LG11), were actually revealed as minor QTLs. This clearly demonstrates that the detection power for minor QTLs is significantly enhanced in the integrated map.

The phenotypic variation observed is also indicated in Table 6. The phenotypic variation described by the major QTL for PCS, HI and Carotene content were above 20%, indicating their major influence on the traits concerned, while that for minor QTLs linked to RL was below 20%. Similarly, the phenotypic variation observed for the major QTLs linked to FAC were mostly between 20 – 35 %, similar to that reported by [39], but slightly higher than reported by [16], where the range was 11 – 20% for similar FAC. The highest phenotypic variation observed was that for IV (measure of level unsaturation) consistent with [16], who also observed highest phenotypic variation for IV.

**Segregation profiles of markers associated with QTLs**

Relationships between the genotype profile of the closest marker (to the QTL peak) and the traits of interest was determined. Palms were differentiated by their genotype (“aa”, “ab” or “bb”), and the phenotype means compared between the different genotypes (Figure 3 and Additional file 5). All the selected markers associated with FAC and vegetative traits were significantly different (P<0.05) in the traits for the different genotypes. Interestingly, the markers SNPM01452 and SNPM00796 in LG01 linked to IV and C16:0 content respectively are located in close proximity (<1cM apart) and show significant difference for the traits (P<0.05) between the genotypes. The “aa” genotype for the
respective markers had higher C16:0 content (saturated fatty acid) and lower IV content (lower levels of unsaturation), indicating the segregation profile of the markers in close proximity are in agreement with the negative correlation between both traits. Similarly, PCS and HI are positively correlated and have the same marker, SNPM00563 linked to both traits in LG04. The “aa” genotype results in smaller PCS and lower HI than the “ab” genotype, also in agreement with their positive correlation.

Table 6. QTLs detected using both the Genestat and MapQTL methods

| QTL  | Marker   | Genestat | MapQTL |
|------|----------|----------|---------|
|      | Marker   | LG       | Position (cM) | LOD Score for Genestat | LOD threshold for MapQTL | LOD score |
| HI   | SNPM00563 | 4        | 4.27        | 7.77                   | 3.4                     | 5.69      |
|      | SNPM04928 | 7        | 110.78      | 3.83                   |                        | 4.33      |
| RL   | SNPM03201 | 4        | 11.30       | 3.79                   | 3.3                     | 3.0*      |
|      | SNPM03772 | 8        | 92.56       | 3.17                   |                        | 4.28      |
|      | SNPM03676 | 11       | 38.69       | 4.17                   |                        | 3.05*     |
| PCS  | SNPM00563 | 4        | 4.27        | 7.70                   | 3.5                     | 4.61      |
|      | SNPM03375 | 4        | 192.47      | 3.25                   |                        | 3.87      |
|      | sEg00213  | 8        | 139.86      | 3.81                   |                        | 3.94      |
| C16:0| SNPM00796 | 1        | 133.4       | 3.59                   | 3.6                     | 4.20      |
| C18:1| SNPM02507 | 4        | 169.2       | 3.44                   | 3.3                     | 4.33      |
|      | SNPM03249 | 8        | 44.03       | 3.19                   |                        | 4.25      |
|      | SNPM00274 | 12       | 31.38       | 5.97                   |                        | -         |
| C18:2| SNPM01602 | 1        | 124.87      | 4.62                   | 3.6                     | 3.31*     |
|      | SNPM00249 | 4        | 3.43        | 8.19                   |                        | 5.55      |
|      | SNPM01190 | 15       | 70.32       | 5.14                   |                        | 3.80      |
| IV   | SNPM01452 | 1        | 132.53      | 13.02                  | 3.5                     | 8.93      |
|      | SNPM04197 | 3        | 15.78       | 3.72                   |                        | 3.4*      |
|      | SNPM03285 | 15       | 98.11       | 5.98                   |                        | 8.51      |
| Carotene | SNPM02349 | 3        | 4.35        | 5.12                   | 3.7                     | 4.24      |
|      | SNPM00729 | 4        | 181.83      | 3.32                   |                        | 4.36      |
|      | SNPM03960 | 7        | 108.53      | 3.73                   |                        | 4.07      |
|      | SNPM03921 | 10       | 77.39       | 3.63                   |                        | 4.83      |

*Significant at chromosome wide
# SNPM00274 was considered not significant, as association with C18:1 was not reproducible in the MapQTL analysis
**Candidate genes identified within the QTL intervals**

Candidate genes residing within the QTL interval were identified using the existing oil palm genome assembly [28]. Blast results to the genome build identified 21 candidate genes within the QTL confidence intervals affecting the vegetative traits and FAC. The *ERECTA* gene [GenBank: XM_010910431.1] was found in the QTL interval linked to PCS, RL and HI on LG4. In addition, the QTL region for HI in LG7 also revealed an interesting gene with high similarity to the auxin transport protein *BIG* [GenBank: XM_010943964.2] in oil palm. Similarly, we identified *BAM1* [GenBank: XM_010914345.2] which co-localized with markers in the QTL region associated with RL on LG11. For FAC, two 3-ketoacyl-CoA synthase genes in Arabidopsis, *CUT1* [GenBank: XM_010917870.2] and *KCS11* [GenBank: XM_010916640.2] flanked the QTLs for IV, C16:0 and C18:2 on LG1. Details of all 21 genes identified are provided [see Additional file 6].

**Discussion**

The traits analyzed in this study - HI, RL, PCS, C16:0, C18:0, C18:1, C18:2, IV and carotene content - are generally highly heritable in oil palm [14,29,30], indicating that they are amenable to selection either via conventional or molecular breeding. This makes the traits attractive for QTL analysis. Four of them (RL, HI, IV and C18:1) in the 2.6-1 mapping family did not follow a normal distribution, but, as reported by [31], deviation from normal distribution did not appear to greatly affect QTL detection. As expected, all the trait values were intermediate between the means for *E. oleifera* and *E. guineensis*, similar to the observations by other studies [16,32]. The wide distribution for all traits measured, suggests that both BC2 populations are ideal for QTL mapping and also for selection and improvement in oil palm. RL and HI in both populations are considerably lower than in commercial DxP as reported by [27], suggesting that these populations can be used for the development of compact palms.

The genetic linkage maps for the two BC2 populations was successfully integrated, improving the resolution of the combined map. The number of palms in the individual populations used in this study
was relatively small compared to in other crops. However, such small family sizes are common in oil palm trials with >64 palms being rare. Although the study focused on highly heritable traits, the small population size could have led to an underestimation of the QTL numbers and restricted the QTL analysis to only the most prominent effects. [33] found that the number of QTLs detected increased with population size. To obviate this limitation, the map resolution and hence QTL detection power, were improved by integrating the two BC2 maps. Other factors, such as the phenotypic measurement accuracy and marker density, also contributed to the QTL detection and localization [34].

Development of an ordered set of markers along the oil palm chromosome allows the genome to be screened systematically for linkages to complex traits. A total of 149 SSR and 1814 SNP markers were found to meet the expected segregation ratios and had a near-neighbour stress value at < 3cM, indicating the suitability for map construction. The number of markers that meet the criteria for map construction was similar to that reported previously [14,16]. SNP markers employed in this study resulted in better genome coverage, increasing the potential for realization of effective marker-assisted selection (MAS). A total of 1,963 polymorphic loci generated 16 LGs, which is consistent with the 16 chromosome-pairs in oil palm [35]. The genome length observed (1,793 cM) was close to that reported by [14,17] of 1,815 – 1,867 cM for E. guineensis. The average length of the LGs is 112 cM, which is in the range of most agricultural crops [36]. Resolution of the two individual genetic maps was good with an average gap of 0.86 cM (2.6-1) and 1.32 cM (2.6-5). An average interval of 0.91 cM was observed on the BC2 integrated map. This gap was much smaller than those previously reported on oil palm interspecific hybrids of 1.2 - 7.2 cM [14,16,17]. The density of both BC2 genetic maps allowed the genomic segments associated with compactness and FAC traits to be identified. Consistency of some of the QTLs detected in the independent and integrated maps [Figure 1 and Additional File 3] adds confidence to their detection.

The interspecific hybrid breeding programme in Malaysia aims to develop palms with higher unsaturated oil and compact characteristics, without sacrificing yield. Applying MAS can accelerate the programme as markers can be linked to selected traits to enable early detection. A high logarithm of the odds ratio (LOD) score will provide confidence for integrating the markers in breeding lines, at
least in palms of similar genetic backgrounds. To ensure the robustness of the genomic region linked to the traits of interest, two independent QTL analysis (Genstat and MapQTL) were employed in this study. Only QTL detected consistently using both methods were considered as significant. Generally, majority of the QTLs detected were common in both methods, with the added advantage of defining them as major or minor QTLs. In the QTL analysis of vegetative traits, interestingly, the QTLs associated with RL, PCS and HI were located at the same genomic region on LG4, likely representing a major locus influencing compactness in oil palm. The markers found in the QTL interval be useful to identify favorable alleles for developing compact interspecific hybrid palms. As PCS is positively correlated with HI and RL, the strategy would be to enrich alleles which favor smaller PCS, lower HI and shorter RL in subsequent screening to develop compact palms.

In the study, population specific QTLs were also identified. Only population specific QTLs that were reproducible in the integrated map using both QTL detection methods were considered as significant. A case in point is the QTL for PCS and RL which was found overlapping around the same region in LG8 for population 2.6-1 as well as the integrated map. The population specific QTLs were most likely due to the female parent, where in 2.6-1 it was a result of the cross between advanced breeding line (Serdang pisifera) and a germplasm from Nigeria (tenera palm, T128). The female parent in 2.6-5 was a result of selfing of the Nigerian palm T128. This likely resulted in the phenotypic variability for the two traits in both populations which contributed to the difference in the QTL results observed in both populations. Similar explanation is also likely for other population specific QTLs such as carotene content and FAC (C18:1, C18:2). Nevertheless, this information could be used to accumulate favorable alleles for unsaturated oil and compactness in future crosses involving the specific populations. Similarly, population specific QTLs were also reported for pod-dehiscence in two separate families of soybean linked by a common parent [37]. Population specific QTLs are generally considered common in QTL analysis involving bi-parental populations [38]. This suggests that QTLs that are detected consistently across two separate populations, as in this study for PCS, HI, C18:2 and IV, have potential for utilization with high confidence in a marker-assisted selection programme. The traits, PCS and HI are positively correlated and the same marker namely SNPM00563, is significantly linked to both
traits in LG04. From Figure 3 and Additional file 5 is clear that selecting for allele “aa” for markers SNPM00563 will resulting in smaller PCS and lower HI, which is further corroborated by the positive correlation observed for both traits (Table 3). The SNP marker SNPM00563 as such, is a good candidate to assist in selecting for palms with smaller PCS and lower HI which will result in more compact palms, that will facilitate harvesting, (due to shorter palms) and planting of more palms per Ha, which will raise productivity [22].

To date, there are no QTL analysis of vegetative traits in interspecific hybrids. The QTLs detected in this study were compared to those described previously [15], for a segregating E. guineensis population. However, most of the QTLs detected in the current BC2 populations were located on different chromosomes compared with those reported by [15], with the exception of those associated with RL which was on LG11. This suggests that separate genomic regions influence compactness in E. guineensis and the interspecific hybrids. A previous study [18] reported two QTLs related to RL and PCS in E. guineensis. However, a comparison between similarity of the linkage groups could not be made as the sequence information for restriction fragment length polymorphism (RFLP) markers linked to the traits in the study were not publicly available for further analysis. More specifically on the HI, recent reports revealed QTLs and candidate genes influencing it in E. guineensis [10,11]. However, the genomic region linked to HI in this study was different from these two reports.

A total of nine significant QTLs were detected for IV, C16:0, C18:1 and C18:2 in five LGs. The result is lower when compared to previous reports on interspecific hybrids [16,20], with 19 and 12 QTLs found, respectively. Five of the identified QTLs in this study were similar to those in both reports. The QTL for IV on LG15 is in agreement with that by [16]. The major QTLs for IV and C16:0 on LG1 were previously reported for an interspecific hybrid family [20] which shows their potential to be used for making informed decisions in breeding. The results support a previous postulation that the same genomic region has a major influence on the unsaturation (IV) and saturation (C16:0) of palm oil. A major QTL for C18:1 and C18:2, were also located on LG4, revealing other chromosomal regions influencing fatty acid composition. Interestingly in this study, the QTL for C18:2 was located around the same region in
LG4 for both populations and in the integrated map. Generally, higher unsaturation level of *E. oleifera* is due to high C18:2 content, which is almost twice that of *E. guineensis* [39]. A strong and consistent QTL for C18:2 content suggests its potential in accumulating favorable allele for unsaturation in interspecific hybrids. The detection of strong QTLs linked C16:0 content, the most abundant saturated fatty acid and IV (measure of unsaturation), similar to previous studies, adds confidence and provides the information necessary to implement MAS selection for higher unsaturated FAC in interspecific hybrid breeding, at least in the genetic background utilized in this study. Interestingly this study also revealed QTLs associated with carotene content in interspecific hybrids. Since similar work for carotene content in oil palm has not been reported, comparison with another research could not be made. Breeding for higher carotene content is desirable due to its pro-vitamin A and antioxidant activities which will help enhance the nutritional attributes of palm oil [40]. The current commercial material, namely *E. guinnensis* has carotene content ranging from 500 to 700 ppm, while oleifera palms can have carotene content as high as 3000ppm [41]. It is clear from this study that individual palms in the interspecific hybrid backcross two populations utilized in this research have carotene content as high as 2700ppm (Table 1). As such, the QTLs identified in this study can help select favorable alleles to enrich future hybrid breeding populations of similar genetic background for higher carotene content, apart from compactness and more desirable FAC.

Availability of the oil palm genome sequence [28] has allowed the underlying QTL interval to be positioned on the EG5 physical map to identify potential candidate genes influencing the traits of interest. The auxin transport protein, BIG, located within the QTL region for HI, is required for auxin efflux and polar auxin transport (PAT) and could influence auxin-mediated developmental responses (e.g. cell elongation, apical dominance, lateral root production, inflorescence architecture, general growth and development) [42]. Generally, BIG controls elongation of the pedicel and stem internodes through auxin action, which support it potential role in regulating HI in oil palm interspecific hybrids. In addition, BIG also plays a role in Arabidopsis with respect to regulation of responses to phytohormones, such as auxin, cytokinins, ethylene and gibberellic acid (GA), particularly during light-mediated stimuli (e.g. shade avoidance and etiolation) [43,44]. BAM1, found in the QTL region for RL,
encodes a leucine-rich repeat receptor-like serine/threonine-protein kinase is known to regulate cell division and differentiation, such as in the formation of shape, size and symmetry of leaves [45], suggesting its possible influence on rachis length of oil palm. *ERECTA*, linked to PCS on LG4, regulates aerial architecture (including inflorescence), e.g. shoot apical meristem-originating organ shape, elongation of internodes and pedicels, and adaxial-abaxial polarity, and stomatal patterning, probably by tuning cell division and expansion [46], which explains how it may control size of petioles (PCS) in oil palm.

In terms of FAC, *KCS11* gene is associated with the QTLs for IV and C16:0 on LG1. The *KCS11* gene is involved in fatty acid biosynthesis on both saturated and mono-unsaturated acyl chains C16 to C20 [47]. As such, it is interesting that the gene is located within the QTL interval that regulating both the C16:0, most abundant saturated fatty acid and IV, measure of unsaturation, mostly related to C18:1 and C18:2 content. Although the candidate genes within the QTL interval are interesting, it is however important to keep in mind that their involvement and influences in controlling compactness and FAC of oil palm are still speculative. Further studies are necessary to characterize these genes to determine their functions in regulating the traits in oil palm.

Conclusions
The oil palm planted area in Malaysia has expanded rapidly in the last few decades largely by assuming the land for other crops, such as rubber and coconut, and secondary forest [48]. However, there is now limited land for further expansion. Thus, improving productivity is the only way. Developing compact palms is a step in that direction, as it can prolong the economic lifespan of the palms, and allow higher planting density to increase the yield per area. Lower HI, shorter RL and smaller PCS are preferred since more nutrients can be channeled into FFB production instead of vegetative growth [27]. In fact, compact palms at a density of 180/ha are being touted [49] for a possible 20% increase in yield. Breeding for compact palms can be accelerated using MAS, as the QTLs identified in this study can facilitate identification of desirable genotypes for the traits of interest.

In addition, reducing saturated and increasing unsaturated FAs will open up the prospects for oil palm
to compete more effectively with other oil crops, such as soybean, rapeseed and sunflower, in the liquid oil sector [50]. There are already efforts to alter the FAC of palm oil both through traditional breeding [51] and genetic engineering [52], but these approaches are still very much in their infancy. Identifying the QTLs associated with FAC and the resulting candidate genes, can prove useful for selection or genetic manipulation in the quest.

Methods

Mapping Populations

The two independent mapping families used in this study were interspecific hybrid backcross two (BC2) populations referred to as ‘2.6-1’ and ‘2.6-5’, consisting of 74 and 80 palms, respectively. The first BC2 population (‘2.6-1’) was derived from a cross involving *E. guineensis* (1084/TP51/22.32) (palm T128, Nigerian germplasm x a Serdang *pisifera*), with an interspecific backcross one (BC1) palm (335/5.2–5/23.96). The second BC2 population (2.6-5’) was derived from a cross between *E. guineensis* (320/TT113/22.32) (self-pollination of palm T128), with the same BC1 palm (335/5.2–5/23.96). The BC1 palm itself involves a cross between GxO F1 hybrid (983/2.4–43/15.90) with palm T128. The GxO F1 hybrid was derived from an interspecific cross of La Mé *E. guineensis* (L2T) and a Colombian *E. oleifera* (79/4.4–12/6.61). The BC2 mapping populations were planted in United Plantations (UP) Berhad in the year 2000. The breeding scheme is illustrated in Figure 4.

Vegetative Measurements and Fatty Acid Composition

The vegetative traits were measured non-destructively [53]. Among the measurements were rachis length (RL) and petiole cross-section (PCS) area, both on Frond 17 following the standard procedure. Height increment (HI) was measured from ground to the base of Frond 41. The difference in height between two years of measurement is the increment, to be divided by the number of years to obtain the annual increment. However, a total of 13 palms from both populations died before any leaf could be sampled. FAC of the mesocarp oil was determined using the MPOB Test Method [54] on 112 palms from both populations. The remaining palms could not be sampled due to sterility issues, abortive bunches and palms having died before sampling.

DNA Extraction and PCR Programme
Spear leaves were harvested from each palm including the parental palms for DNA extraction. DNA extraction was carried out using the modified CTAB method [55]. SSR analysis was carried out using the respective primers as described previously [56,57], using the following PCR parameters: pre-denaturation at 95°C for 1 min, denaturation at 95°C for 30 s, annealing (temperature depends on primer) for 30 s, and extension at 72°C for 30 s. This programme was repeated for 35 cycles, followed by a final extension at 72°C for 5 min. SNP analysis was carried out as previously described [56,57].

**Development of linkage map and QTL analysis**

SNP and SSR data were analysed as described previously [20,56,57]. A total of 515 and 715 SSR markers were genotyped on mapping populations 2.6-1 and 2.6-5, respectively. Meanwhile, a total of 4,451 SNPs and an additional 40 candidate SNPs flanking various fatty acid and oil biosynthesis related genes were genotyped on both populations using the Illumina Infinium assay and iPLEX, respectively. A genetic linkage map was first constructed separately for each population using JoinMap 4.1 [58]. Regression mapping algorithm with the default parameters (minimum LOD score threshold of 1.0, recombination fraction threshold of 0.4, ripple value of 1.0 and jump threshold of 5.0) was used to assign markers to individual linkage groups. Haldane’s mapping function was then used to determine map order. Only markers with < 5% of missing data and following Mendelian inheritance ratio (p <0.05) were used in map construction. Markers showing nearest-neighbour stress of more than 3 cm were also omitted. Subsequently, the two maps were integrated, and QTL analysis carried out using Genstat 18th edition [59]. QTL threshold value was determined as LOD 3, using default parameters, with the effective number of independent test set as defined by [60].

QTL analysis for the integrated map was also carried out using MapQTL 5 [61]. Interval mapping was used for QTL detection using maximum of 5 neighboring markers and 1 cM mapping size. QTL threshold values (genome-wide and chromosome-wide) were determined with 1000 permutation test. Confidence interval of the QTL was determined using the standard method of LOD-1, while the phenotypic variation observed was corrected as described by [39], by multiplying the explained variance with the following formulae: (see Formula 1 in the Supplementary Files)

Correction was necessary to overcome the overestimation of phenotypic variance in small populations.
Phenotyping and Marker Sequence Similarity Search

Using SAS version 9.3, the t-test and Duncan analyses were carried out to compare the phenotypic values of the different genotypes. For comparison, palms were grouped according to their genotypes and phenotypic values averaged for each genotype. The sequences of markers flanking all QTLs identified were extracted from the oil palm genome build (EG5) [28] and searched for sequence similarity (BLASTN and BLASTX) against the NCBI databases. Sequences with significant similarity (BLASTN e-value of < 1e-25 and 90 % identity over total sequence length) to genes of interest were shortlisted for further analysis. Putative functions of the selected genes were derived from UniProt, a freely accessible database of protein sequence and functional information, and literature.

Abbreviations

1. Interspecific backcross two (BC₂)
2. Fatty acid composition (FAC)
3. Quantitative trait loci (QTL)
4. Interspecific backcross one (BC₁)
5. Rachis length (RL)
6. Height increment (HI)
7. Iodine value (IV)
8. Petiole cross section (PCS)
9. Oleic acid (C₁₈:₁)
10. Stearic acid (C₁₈:₀)
11. Linoleic acid (C₁₈:₂)
12. Palmitic acid (C₁₆:₀)
13. Linkage group (LG)
14. Single nucleotide polymorphism (SNP)
15. Simple sequence repeat (SSR)
16. Marker-assisted selection (MAS)
17. Logarithm of the odds ratio (LOD)
18. Oil palm genome build (EG5)
19. Restriction fragment length polymorphism (RFLP)
20. Polar auxin transport (PAT)
21. Gibberellic acid (GA)
22. Very-long-chain fatty acids (VLCFAs).

Declarations

Ethics approval and consent to participate
Not Applicable

Consent for publication
Not Applicable

Availability of data and material
All data generated or analysed during this study are included in this published article [and its supplementary information files]

Competing interests
The authors declare that they have no competing interests

Funding
The funding of this study including data collection, analysis, interpretation of data and this manuscript publication was funded by Malaysian Palm Oil Board (MPOB)

Authors’ contributions
ZY, RS, KS and RN conceived and designed the experiments. ZY, KK, and JJ performed and analyzed the experiments. ZY and RS thoroughly interpreted the data and revised the manuscript for intellectual content. TNC, MM, MDA, LETL, OLCL and MOA coordinated the project and participated in the direction of the study. ZY, KK, SM and RS wrote the manuscript. All authors discussed the results and commented on the manuscript.

Acknowledgements
The authors would like to thank the Director-General of MPOB for permission to publish this paper. The second backcross (BC2) oil palm population used in this study was obtained from United Plantations Berhad, Teluk Intan, Perak. Part of the work was carried out at Biometris, Wageningen University and Research Centre, the Netherlands. The authors also would like to thank Andy Chang Kwong Choong for valuable comments on this manuscript.

References
1. Oil World. http://www.oilworld.biz/app.php (2017). Accessed on 11 February 2019.
2. Kushairi A, Soh KH, Azman I, Elina H, Meilina OA, Zanal Bidin MNI, Razmah G, Shamala S, Ahmad Parveez GK. Oil Palm Economic Performance in Malaysia and R&D Progress in 2017. Journal of Oil Palm Research. 2018; 30(2): 163–195.
3. USDA. Malaysia: Stagnating palm oil yields impede growth. Commodity Intelligence Report. http://www.pecad.fas.usda.gov/highlights/2012/12/Malaysia (2012). Accessed on 28 February 2019.
4. Sharma M, Tan YP. Oil palm breeding programme and the performance of DXP planting materials at United Plantation Bhd. Proceedings of Seminar on Sourcing of Oil Palm Planting Materials for Local and Overseas Join Venture: PORIM, Malaysia. 1999. p. 118-135.
5. Kushairi A, Rajanaidu N, Jalani BS, Mohd Rafii Y, Mohd Din A. PORIM oil palm planting materials. PORIM Bull. 1999; 38: 1–13.
6. Barcelos E, Rios SA, Cunha RNV, Lopes R, Motoike SY, Babiychuk E, Skirycz A, Kushnir S. Oil palm natural diversity and the potential for yield improvement. Frontiers Plant Science. 2015; 6: 190.
7. Jagoe RB. Deli oil palms and early introductions of *Elaeis guineensis* to Malaya. Malay Agric J. 1952; 3: 4–11.
8. Rajanaidu N, Jalani BS. Oil Palm genetic resources – collection, evaluation, utilization and conservation. PORIM Colloquium on Oil Palm Genetic Resources. PORIM, Bangi;
9. Bilal M, Veerappan P, Nazeeb M. Evaluation of Deli Nigerian DxT and Avros-Nigerian TxT crosses in Sime Darby Plantations. In: Rajanaidu N, Jalani BS, editors. Proc. Seminar on PS1 and PS2 in oil palm planting materials. Palm Oil Research Institute of Malaysia, Kuala Lumpur; 1999. p. 54-64.

10. Pootakham W, Jomchai N, Ruang-Areerate P, Shearman JR, Sonthirod C, Sangsrakru D, Tragoonrungrung S, Thangphatsornruang S. Genome-wide SNP discovery and identification of QTL associated with agronomic traits in oil palm using genotyping-by-sequencing (GBS). Genomics. 2015; 105: 288–295.

11. Lee M, Xia JH, Zou Z, Ye J, Rahmadsyah, Alfiko Y, Jin J, Lieando JV, Purnamasari MI, Lim CH, Suwanto A, Wong L, Chua NH, Yue GH. A consensus linkage map of oil palm and a major QTL for stem height. Scientific Reports. 2015; 5: 8232.

12. Ong PW, Maizura I, Marhalil M, Rajanaidu N, Abdullah NAP, Rafii, MY, Ooi LCL, Low ETL, Singh R. Association of SNP Markers with Height Increment In MPOB-Angolan Natural Oil Palm Populations. Journal of Oil Palm Research. 2018; 30 (1): 61–70.

13. Kushairi A, Mohd Din A, Rajanaidu N. Oil palm breeding and seed production. In: Mohd Basri W, Choo YM, Chan KW, editors. Further advances in oil palm research (2000–2010) MPOB, Bangi; 2011. p. 47-101.

14. Singh R, Tan SG, Panandam JM, Rahimah AR, Ooi LCL. Mapping quantitative trait loci (QTLs) for fatty acid composition in an interspecific cross of oil palm. BMC Plant Biology. 2009; 9: 114.

15. Billotte N, Jourjon MF, Marseillac N, Berger A, Flori A, Asmady H, Adon B, Singh R, Nouy B, Potier F, Cheah SC, Rohde W, Ritter E, Courtois B, Charrier A, Mangin B. QTL detection by multi-parent linkage mapping in oil palm (Elaeis guineensis Jacq.). Theor Appl Genet. 2010; 120:1673-87.
16. Montoya C, Lopes R, Flori A, Cros D, Cuellar T, Summo M, Espeout S, Rivallan R, Risterucci AM, Bittencourt D, Zambrano JR, Alarcon GWH, Villeneuve P, Pina M, Nouy B, Amblard P, Ritter E, Leroy T, Billote N. Quantitative trait loci (QTLs) analysis of oil palm fatty acid composition in an interspecific pseudo-backcross from *Elaeis oleifera* (H.B.K.) Cortes and oil palm (*Elaeis guineensis* Jacq.). Tree Genetics & Genomes. 2013; 9: 1207-1225.

17. Ting NC, Jansen J, Mayes S, Massawe F, Sambanthamurthi R, Ooi RCL, Chin CW, Arulandoo X, Seng TY, Syed-Alwee SSR, Ithnin M, Singh R. High density SNP and SSR-based genetic maps of two independent oil palm hybrids. BMC Genomics. 2014; 15: 309.

18. Rance KA, Mayes S, Price Z, Jack PL, Corley RHV. Quantitative trait loci for yield components in oil palm (*Elaeis guineensis* Jacq.). Theor. Appl. Genet. 2001;103: 1302-1310.

19. Jeennor S, Volkaert H. Mapping of quantitative trait loci (QTLs) for oil yield using SSRs and gene-based markers in African oil palm (*Elaeis guineensis* Jacq.). Tree Genet Genomics. 2014; 10: 1-14.

20. Ting NC, Zulkifli Y, Katialisa K, Mayes S, Massawe F, Ravigadevi S, Jansen J, Maizura I, Kushairi A, Xaviar A, Rozana R, Chan KL, Nadzirah A, Kandha S, Lim CC, Rajanaidu N, Mohd Din A, Rajinder S. Fine-mapping and cross-validation of QTLs linked to fatty acid composition in multiple independent interspecific crosses of oil palm. BMC Genomics. 2016; 17: 289.

21. Escobar R, Alvarado A. Strategies in production of oil palm seed varieties and clones for high density planting. ASD Oil Palm Papers (Costa Rica). 2004; 27:1-12.

22. Zulkifli Y, Norziha A, Naqiuddin MH, Fadila AM, Nor Azwani AB, Suzana M, Samsul KR, Ong-Abdullah M, Singh R, Ahmad Parveez GK, Kushairi A. Designing the Oil Palm of
23. Escobar R, Alvarado A, Chinchilla C, Guzman N, Peralta F. Performance of ASD de Costa Rica oil palm planting materials: seeds and compact clones. In: Asgard Seminar, Sourcing of oil palm planting materials for localand overseas joint ventures. Kuala Lumpur. 2006.

24. Sterling F, Richardson DL, Chavez C. Some phenotypic characteristics of the descendants of 'QB49: 238', and exceptional hybrid of oil palm. International Oil Palm/Palm Oil Conference-Agriculture Proceedings: PORIM, Malaysia. 1988. p. 94-102.

25. Alvarado A, Escobar R, Peralta R, Chinchilla C. Compact seeds and clones and their potential for high density planting. Proceedings of International ISOPB Seminar on Yield Potential in The Oil Palm. Phuket, Thailand; 2006. p.10.

26. Cadena T, Prada F, Perea A, Romero H. Lipase activity, mesocarp oil content, and iodine value in oil palm fruits of Elaeis guineensis, Elaeis oleifera, and the interspecific hybrid OxG (E. oleifera x E. guineensis). J Sci Food Agr. 2012; 93(3): 674-80.

27. Noh A, Rafii MY, Saleh G, Kushairi A, Latif MA. Genetic performance and general combining ability of oil palm Deli dura x AVROS pisifera tested on inland soils. Scientific World Journal. 2012; 792601.

28. Singh R, Ong-Abdullah M, Low ET, Manaf MA, Rosli R, Nookiah R, Ooi LC, Ooi SE, Chan KL, Halim MA, Azizi N, Nagappan J, Bacher B, Lakey N, Smith SW, He D, Hogan M, Budiman MA, Lee EK, DeSalle R, Kudrna D, Goicoechea JL, Wing RA, Wilson RK, Fulton RS, Ordway JM, Martienssen RA, Sambanthamurthi R. Oil palm genome sequence reveals divergence of interfertile species in Old and New worlds. Nature. 2013; 500: 335-339.

29. Hardon JJ, Rao V, Rajanaidu N. A review of oil palm breeding. In: Progress in Plant
Breeding, Rusell GE, editors. Butterworths: United Kingdom; 1985. p. 139- 163.

30. Marhalil M, Rafii MY, Afizi MMA, Arolu W, Noh A, Mohd Din A, Kushairi A, Norziha A, Rajanaidu N, Latif MA, Malek MA. Genetic variability in yield and vegetative traits in elite germplasm of MPOB-Nigerian *dura* × AVROS *pisifera* progenies. Food, Agriculture and Enviroment. 2013; 11: 515-519.

31. Septiningsih EM, Prasetiyono J, Lubis E, Tai TH, Tjubaryat T, Moeljopawiro S, McCouch SR. Identification of quantitative trait loci for yield and yield components in an advancedbackcross population derived from the Oryza sativa varietyIR64 and the wild relative O. rufipogon. Theor. Appl. Genet. 2003; 8: 1419-1432.

32. Meunier J, Boulin D. *Elaeis melanococca* and *E. melanococca* × *E. guineensis* hybrids. First results. Oleagineux. 1975; 30:5-8.

33. Vales MI, Schön CC, Capettini F, Chen XM, Corey AE, Mather DE, Mundt CC, Richardson KL, Sandoval-Islas JS, Utz HF, Hayes PM. Effect of population size on the estimation of QTL: a test using resistance to barley stripe rust. Theor. Appl. Genet. 2005; 111: 1260-1270.

34. Li XM, Quigg RJ, Zhou J, Xu SZ, Masinde G, Mohan S, Baylink DJ. A critical evaluation of the effect of population size and phenotypic measurement on QTL detection and localization using a large F2 murine mapping population. Genetics and Molecular Biology. 2006; 29:166-173.

35. Maria M, Clyde MM, Cheah SC. Cytological analysis of *Elaeis guineensis* (tenera) chromosomes. Elaeis. 1995; 7:122-134.

36. Maliepaard C, Alston FH, van Arkel G, Brown LM, Chevreau E, Dunemann F, Evans KM, Gardiner S, Guilford P, van Heusden AW, Janse J, Laurens F, Lynn, JR, Manganaris AG, den Nijs APM, Periam N, Rikkerink E, Roche P, Ryder C, Sansavini S, Schmidt H, Tartarini S, Verhaegh JJ, Vrieling-van Ginkel M, King CJ. Aligning male and female
linkage maps of apple (*Malus pumilla* Mill.) using multi-allelic markers. Theoretical and Applied Genetics. 1998; 97: 60-73.

37. Kang ST, Kwak M, Kim HK, Choung MG, Han WY, Baek IY, Kim MY, Van K, Lee SH. Population-specific QTLs and their different epistatic interactions for pod dehiscence in soybean [*Glycine max* (L.) Merr.] Euphytica. 2009; 166:15-24.

38. Würschum T. Mapping QTL for agronomic traits in breeding populations. Theoretical and Applied Genetics. 2012; 125: 201-210.

39. Montoya C, Cochard B, Flori A, Cros D, Lopes R, Cuellar T, Espeout S, Syaputra I, Villeneuve P, Pina M, Ritter E, Leroy T, Billotte N. Genetic architecture of palm oil fatty acid composition in cultivated oil palm (*Elaeis guineensis* Jacq.) compared to its wild relative *E. oleifera* (H.B.K) Cortes. PLoS ONE. 2014; 9: e95412.

40. Mba OI, Dumont MJ, Ngadi M. Palm oil: Processing, characterization and utilization in the food industry-a review. Food Bioscience. 2015; 10: 26-41.

41. Mohd Din A, Rajanaidu N, Kushairi A, Mohd Rafii Y, Mohd Isa ZA, Noh A. PS4: High carotene E. oleifera planting materials. 2012. MPOB Information Series No. 137. MPOB, Bangi.

42. Gil P, Dewey E, Friml J, Zhao Y, Snowden KC, Putterill J, Palme K, Estelle M, Chory J. BIG: a callosin-like protein required for polar auxin transport in Arabidopsis. Genes Dev. 2000; 15:1985-1997.

43. Li HM, Altschmied L, Chory J. *Arabidopsis mutants define downstream branches in the phototransduction pathway*. Genes Dev. 1994; 8:339-349.

44. Sponsel VM, Schmidt FW, Porter SG, Nakayama M, Kohlstruk S, Estelle M. *Characterization of new gibberellin-responsive semidwarf mutants of arabidopsis*. Plant Physiol. 1997; 115:1009-1020.

45. Deyoung BJ, Clark SE. *BAM receptors regulate stem cell specification and organ
development through complex interactions with CLAVATA signaling. Genetics. 2008; 180:895-904.

46. Torii KU, Mitsukawa N, Oosumi T, Matsuura Y, Yokoyama R, Whittier RF, Komeda Y. The Arabidopsis ERECTA gene encodes a putative receptor protein kinase with extracellular leucine-rich repeats. Plant Cell. 1996; 8:735-746.

47. Blacklock BJ, Jaworski JG. Substrate specificity of Arabidopsis 3-ketoacyl-CoA synthases. Biochem Biophys Res Commun. 2006; 346:583-90.

48. May CY. Malaysia: Economic transformation advances oil palm industry. American Oil Chemists’ Society. 2012.

49. Alvarado A, Henry J. Evolution Blue: A new oil palm variety with reduced growth and high oil content. ASD Oil Palm Papers; 2015. p. 45.

50. Murphy DJ. Plant Breeding and Biotechnology: Societal Context and the Future of Agriculture. Cambridge University Press, Cambridge; 2007.

51. Isa ZA, Mohd Din A, Maizura I, Noh A, Kushairi A, Rajanaidu N. (2006). PS12: Breeding population for high oleic acid palm oil. MPOB Information Series No. 313. MPOB, Bangi; 2006.

52. Parveez GKA, Abrizah O, Tarmizi HA, Zamzuri I, Kushairi DA, Salmah Y, Bahariah B, Sabariah K. (2003). Transformation of oil palm with antisense palmitoyl-ACP thioesterase gene for increasing oleic acid content. Proc. of the PIPOC 2003 International Palm Oil Congress – Agriculture Conference. 2003. p. 869-878.

53. Breure CJ, Powell MS. The one-shot method of establishing growth parameters in oil palm. Proceedings of 1987 Int. Oil Palm Conference Agriculture: PORIM, Malaysia; 1988.

54. PORIM. PORIM Test Method. In: Methods of test for palm oil and palm oil products. Bandar Baru Bangi: Palm Oil Research Institute Malaysia (PORIM); 1995.
55. Dellaporta SL, Wood J, Hicks JB. A plant DNA minipreparation: Version II. Plant Mol. Biol. Rep. 1983; 1: 19-21.

56. Katialisa K, Ithnin M, Ooi LCL, Ting NC, Musa B, Singh R. Construction of SNP- and SSR-based genetic linkage map in the oil palm backcross two (BC2) mapping population. In: Proceedings of the 2013 PIPOC International Palm oil Congress. Kuala Lumpur, Malaysia; 2013. p.158.

57. Zulkifli Y, Rajinder S, Mohd Din A, Ting NC, Rajanaidu N, Kushairi A, Musa B, Mohamad O, Ismanizan I. Inheritance of SSR and SNP loci in an oil palm interspecific hybrid backcross (BC2) population. Journal of Oil Palm Research. 2014; 26: 203-213.

58. Van Ooijen JW. JoinMap® 4.1. Software for calculation of genetic linkage maps in experimental populations. 2011. Kyazma B.V., Wageningen, the Netherlands.

59. VSN International. Genstat for Windows 18th Edition. 2015. VSN International, Hemel Hempstead, UK.

60. Li J, Ji L. Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. Heredity.2005; 95:221–227.

61. Van Ooijen JW. MapQTL® 6. Software for the mapping of quantitative trait loci in experimental populations of diploid species. 2009. Kyazma B.V., Wageningen, the Netherlands.

62. Raghavan C, Collard BCY. Effect of small mapping population sizes on reliability of quantitative trait locus (QTL) mapping. African J Biotechnol. 2012; 11: 10661-10674.

Figures
Figure 1

The integrated genetic map of the backcross 2 populations. QTLs detected in the two individual populations, as well as in the integrated map, is indicated on the individual linkage groups. Overlapping QTLs for the different traits can be observed (e.g., LG4, LG8).

Figure 2

Corresponding regions of LG4 and pseudochromosome 2, showing a major genomic region influencing compactness and fatty acid composition in LG4.
Figure 3

Boxplot showed means of phenotypes ((A) C16:0, (B) C18:1, (C) C18:2, (D) IV, (E) Carotene, (F) RL, (G) HI, (H) PCS were compared using the independent t-test for a marker with two genotypes, and Duncan’s test for markers with three genotypes (SAS 9.3 statistical package). Means of the different genotypes were significantly different at P<0.05.
Breeding scheme of the two interspecific backcross two (BC2) mapping populations

Supplementary Files
This is a list of supplementary files associated with this preprint. Click to download.

Additional File 2.doc
Formula 1.jpg
Additional File 4.docx
Additional File 6.docx
Additional File 5.docx
Additional File 3.docx
Additional File 1.docx