

Claims

1. A method for predicting palm oil yield of a test oil palm plant, the method comprising the steps of:

(i) determining, from a sample of a test oil palm plant of a population of oil palm plants, at least a first single nucleotide polymorphism (SNP) genotype of the test oil palm plant, the first SNP genotype corresponding to a first SNP marker, the first SNP marker (a) being located in a first quantitative trait locus (QTL) for a high-oil-production trait and (b) being associated, after stratification and kinship correction, with the high-oil-production trait with a genome-wide $-\log_{10}(p\text{-value})$ of at least 4.0 in the population or having a linkage disequilibrium r^2 value of at least 0.2 with respect to a first other SNP marker that is linked thereto and associated, after stratification and kinship correction, with the high-oil-production trait with a genome-wide $-\log_{10}(p\text{-value})$ of at least 4.0 in the population;

(ii) comparing the first SNP genotype of the test oil palm plant to a corresponding first reference SNP genotype indicative of the high-oil-production trait in the same genetic background as the population; and

(iii) predicting palm oil yield of the test oil palm plant based on the extent to which the first SNP genotype of the test oil palm plant matches the corresponding first reference SNP genotype,

wherein the first QTL is a region of the oil palm genome corresponding to one of:

- (1) QTL region 1, extending from nucleotide 66542323 to 66776312 of chromosome 1;
- (2) QTL region 2, extending from nucleotide 66807385 to 67299617 of chromosome 1;
- (3) QTL region 3, extending from nucleotide 62277032 to 62355782 of chromosome 2;
- (4) QTL region 4, extending from nucleotide 31132787 to 31173962 of chromosome 4;
- (5) QTL region 5, extending from nucleotide 32863621 to 32964104 of chromosome 5;
- (6) QTL region 6, extending from nucleotide 33355931 to 33509217 of chromosome 5;
- (7) QTL region 7, extending from nucleotide 33658904 to 34233352 of chromosome 5;
- (8) QTL region 8, extending from nucleotide 34358119 to 34997228 of chromosome 5;
- (9) QTL region 9, extending from nucleotide 35004388 to 35125743 of chromosome 5;
- (10) QTL region 10, extending from nucleotide 35191678 to 35193677 of chromosome 5;
- (11) QTL region 11, extending from nucleotide 36108847 to 36272808 of chromosome 5;
- (12) QTL region 12, extending from nucleotide 39210662 to 39225076 of chromosome 5;
- (13) QTL region 13, extending from nucleotide 39518005 to 40469897 of chromosome 5;
- (14) QTL region 14, extending from nucleotide 40535309 to 40690150 of chromosome 5;
- (15) QTL region 15, extending from nucleotide 40789706 to 40983955 of chromosome 5;
- (16) QTL region 16, extending from nucleotide 41001085 to 41302446 of chromosome 5;
- (17) QTL region 17, extending from nucleotide 3050807 to 3241977 of chromosome 8;

- (18) QTL region 18, extending from nucleotide 5354764 to 5445890 of chromosome 8;
 (19) QTL region 19, extending from nucleotide 29488933 to 29602300 of chromosome 9;
 (20) QTL region 20, extending from nucleotide 4797284 to 5717606 of chromosome 11; or
 (21) QTL region 21, extending from nucleotide 8611715 to 8857914 of chromosome 15;

wherein the population of oil palm plants comprises a Nigerian *dura* X AVROS *pisifera* population, and the first SNP marker comprises one or more SNP markers as follows:

SNP No.	QTL region	Chromosome	Position	Major allele	Minor allele
2	2	1	66972538	A	G
4	2	1	67248054	A	C
5	3	2	62287970	G	A
6	4	4	31149521	G	A
8	6	5	33457617	G	A
10	7	5	34027909	A	G
15	8	5	34568717	G	A
17	8	5	34612126	G	A
19	8	5	34828628	G	A
20	8	5	34856258	C	A
21	8	5	34863191	A	G
22	8	5	34980654	A	G
24	9	5	35080572	G	A
26	9	5	35121695	A	G
27	10	5	35192678	A	G
29	11	5	36218554	A	G
30	11	5	36234729	G	A
31	12	5	39210662	A	G
34	13	5	39613906	G	A
35	13	5	39620126	A	G
37	13	5	39653455	G	A
38	13	5	39763460	A	C
43	13	5	39860983	G	A
44	13	5	39966907	A	G
45	13	5	40066844	G	A
46	13	5	40112108	G	A
47	13	5	40129189	C	A
48	13	5	40145585	A	G
49	13	5	40158838	A	G
50	13	5	40249343	G	A
52	13	5	40300709	A	G
53	13	5	40396733	A	C
54	13	5	40405137	A	G
55	14	5	40577880	G	A
56	14	5	40587726	A	G
57	14	5	40597291	A	G
58	15	5	40902353	A	G
60	16	5	41036282	A	G
61	16	5	41189141	A	G
62	16	5	41200140	G	A

63	16	5	41203008	A	G
64	16	5	41204629	G	A
65	17	8	3154572	G	A
66	18	8	5393101	C	A
68	20	11	4828172	A	G
76	20	11	5412986	A	G
77	20	11	5420279	A	G
78	20	11	5422706	G	A
79	20	11	5436885	G	A
81	20	11	5442401	C	A
82	21	15	8740489	A	G

and wherein the population of oil palm plants comprises a Deli *dura* X AVROS *pisifera* population, and the first SNP marker comprises one or more SNP markers as follows:

SNP No.	QTL region	Chromosome	Position	Major allele	Minor allele
1	1	1	66639699	A	G
3	2	1	67033874	A	G
6	4	4	31149521	G	A
7	5	5	32866989	A	G
8	6	5	33457617	G	A
9	7	5	33975929	G	A
10	7	5	34027909	A	G
11	7	5	34142044	C	A
12	8	5	34454937	G	A
13	8	5	34521370	G	A
14	8	5	34525454	G	A
15	8	5	34568717	G	A
16	8	5	34603394	G	A
17	8	5	34612126	G	A
18	8	5	34773282	A	G
19	8	5	34828628	G	A
20	8	5	34856258	C	A
21	8	5	34863191	A	G
22	8	5	34980654	A	G
23	9	5	35070248	A	C
24	9	5	35080572	G	A
25	9	5	35100527	A	G
26	9	5	35121695	A	G
27	10	5	35192678	A	G
28	11	5	36158880	G	A
29	11	5	36218554	A	G
30	11	5	36234729	G	A
31	12	5	39210662	A	G
32	13	5	39607208	A	G
33	13	5	39610847	G	A
34	13	5	39613906	G	A
35	13	5	39620126	A	G
36	13	5	39642505	C	A
37	13	5	39653455	G	A
38	13	5	39763460	A	C

39	13	5	39799450	A	G
40	13	5	39804720	G	A
41	13	5	39805514	G	A
42	13	5	39806797	A	G
43	13	5	39860983	G	A
44	13	5	39966907	A	G
45	13	5	40066844	G	A
46	13	5	40112108	G	A
47	13	5	40129189	C	A
48	13	5	40145585	A	G
49	13	5	40158838	A	G
50	13	5	40249343	G	A
51	13	5	40269392	G	A
52	13	5	40300709	A	G
53	13	5	40396733	A	C
54	13	5	40405137	A	G
58	15	5	40902353	A	G
59	15	5	40916454	G	A
60	16	5	41036282	A	G
61	16	5	41189141	A	G
62	16	5	41200140	G	A
63	16	5	41203008	A	G
64	16	5	41204629	G	A
67	19	9	29537541	0	G
68	20	11	4828172	A	G
69	20	11	4831923	G	A
70	20	11	4838915	G	A
71	20	11	5165520	0	G
72	20	11	5204949	0	G
73	20	11	5255710	C	A
74	20	11	5369351	A	C
75	20	11	5372237	G	A
76	20	11	5412986	A	G
77	20	11	5420279	A	G
78	20	11	5422706	G	A
79	20	11	5436885	G	A
80	20	11	5439423	G	A
81	20	11	5442401	C	A
82	21	15	8740489	A	G

2. The method of claim 1, wherein the high-oil-production trait comprises increased oil-to-dry mesocarp.

3. The method of claims 1 or 2, wherein the population of oil palm plants comprises a Nigerian *dura* x AVROS *pisifera* population, a Deli *dura* x AVROS *pisifera* population, or a combination thereof.

4. The method of claims 1, 2, or 3, wherein:
the population of oil palm plants comprises a Nigerian *dura* x AVROS *pisifera* population;
the first QTL corresponds to one of QTL regions 2, 3, 8, 10, 13, 14, 16, 17, or 18; and
step (iii) further comprises applying a genotype model, thereby predicting the palm oil yield of the test oil palm plant.
5. The method of claims 1, 2, or 3, wherein:
the population of oil palm plants comprises a Nigerian *dura* x AVROS *pisifera* population;
the first QTL corresponds to one of QTL regions 3, 8, 10, 13, 15, 16, 17, or 18; and
step (iii) further comprises applying a dominant model, thereby predicting the palm oil yield of the test oil palm plant.
6. The method of claims 1, 2, or 3, wherein:
the population of oil palm plants comprises a Nigerian *dura* x AVROS *pisifera* population;
the first QTL corresponds to one of QTL regions 3, 4, 6, 7, 8, 9, 10, 11, 12, 13, 14, 16, 20, or 21; and
step (iii) further comprises applying a recessive model, thereby predicting the palm oil yield of the test oil palm plant.
7. The method of claims 1, 2, or 3, wherein:
the population of oil palm plants comprises a Deli *dura* x AVROS *pisifera* population;
the first QTL corresponds to one of QTL regions 1, 2, 4, 5, 6, 7, 8, 9, 11, 12, 13, 15, 16, 19, 20, or 21; and
step (iii) further comprises applying a genotype model, thereby predicting the palm oil yield of the test oil palm plant.
8. The method of claims 1, 2, or 3, wherein:
the population of oil palm plants comprises a Deli *dura* x AVROS *pisifera* population;
the first QTL corresponds to one of QTL regions 8, 10, or 13; and
step (iii) further comprises applying a dominant model, thereby predicting the palm oil yield of the test oil palm plant.
9. The method of claims 1, 2, or 3, wherein:
the population of oil palm plants comprises a Deli *dura* x AVROS *pisifera* population;

the first QTL corresponds to one of QTL regions 1, 2, 4, 5, 6, 7, 8, 9, 11, 12, 13, 15, 16, 19, 20, or 21; and

step (iii) further comprises applying a recessive model, thereby predicting the palm oil yield of the test oil palm plant.

10. The method of any one of claims 1-9, wherein the test oil palm plant is a *tenera* candidate agricultural production plant.

11. The method of claims 1 or 2, wherein the population of oil palm plants comprises a Nigerian *dura* x Nigerian *dura* population, a Nigerian *dura* x Deli *dura* population, a Deli *dura* x Deli *dura* population, an AVROS *pisifera* x AVROS *tenera* population, an AVROS *tenera* x AVROS *tenera* population, or a combination thereof.

12. The method of claims 1, 2, or 11, wherein the test oil palm plant is a plant for mother palm selection and propagation, a plant for introgressed mother palm selection and propagation, or a plant for pollen donor selection and propagation.

13. The method of any one of claims 1-12, wherein the test oil palm plant is a seed, a seedling, a nursery phase plant, an immature phase plant, a cell culture plant, a zygotic embryo culture plant, or a somatic tissue culture plant.

14. The method of any one of claims 1-12, wherein the test oil palm plant is a production phase plant, a mature palm, a mature mother palm, or a mature pollen donor.

15. The method of any one of claims 1-14, wherein:

step (i) further comprises determining, from the sample of the test oil palm plant, at least a second SNP genotype of the test oil palm plant, the second SNP genotype corresponding to a second SNP marker, the second SNP marker (a) being located in a second QTL for the high-oil-production trait and (b) being associated, after stratification and kinship correction, with the high-oil-production trait with a genome-wide $-\log_{10}(p\text{-value})$ of at least 4.0 in the population or having a linkage disequilibrium r^2 value of at least 0.2 with respect to a second other SNP marker that is linked thereto and associated, after stratification and kinship correction, with the high-oil-production trait with a genome-wide $-\log_{10}(p\text{-value})$ of at least 4.0 in the population; and

step (ii) further comprises comparing the second SNP genotype of the test oil palm plant to a corresponding second reference SNP genotype indicative of the high-oil-production trait in the same genetic background as the population,

wherein the second QTL corresponds to one of QTL regions 1 to 21, with the proviso that the first QTL and the second QTL correspond to different QTL regions.

16. The method of claim 15, wherein step (iii) further comprises predicting palm oil yield of the test oil palm plant based on the extent to which the second SNP genotype of the test oil palm plant matches the corresponding second reference SNP genotype.

17. The method of claims 15 or 16, wherein:

step (i) further comprises determining, from the sample of the test oil palm plant, at least a third SNP genotype to a twenty-first SNP genotype of the test oil palm plant, the third SNP genotype to the twenty-first SNP genotype corresponding to a third SNP marker to a twenty-first SNP marker, respectively, the third SNP marker to the twenty-first SNP marker (a) being located in a third QTL to a twenty-first QTL, respectively, for the high-oil-production trait and (b) being associated, after stratification and kinship correction, with the high-oil-production trait with a genome-wide $-\log_{10}(p\text{-value})$ of at least 4.0 in the population or having linkage disequilibrium r^2 values of at least 0.2 with respect to a third other SNP marker to a twenty-first other SNP marker, respectively, that are linked thereto and associated, after stratification and kinship correction, with the high-oil-production trait with a genome-wide $-\log_{10}(p\text{-value})$ of at least 4.0 in the population; and

step (ii) further comprises comparing the third SNP genotype to the twenty-first SNP genotype of the test oil palm plant to a corresponding third reference SNP genotype to a corresponding twenty-first reference SNP genotype, respectively, indicative of the high-oil-production trait in the same genetic background as the population,

wherein the third QTL to the twenty-first QTL each correspond to one of QTL regions 1 to 21, with the proviso that the first QTL to the twenty-first QTL each correspond to different QTL regions.

18. The method of claim 17, wherein step (iii) further comprises predicting palm oil yield of the test oil palm plant based on the extent to which the third SNP genotype to the twenty-first SNP genotype of the test oil palm plant match the corresponding third reference SNP genotype to the corresponding twenty-first reference SNP genotype, respectively.

19. A method of selecting a high-palm-oil-yielding oil palm plant for agricultural production of palm oil, the method comprising the steps of:

(a) predicting palm oil yield of a test oil palm plant according the method of any one of claims 1-18; and

(b) field planting the test oil palm plant for agricultural production of palm oil if the palm oil yield of the test oil palm plant is predicted to be higher than average for the population based on step (a).

20. A method of selecting a high-palm-oil-yielding oil palm plant for cultivation in cell culture, the method comprising the steps of:

(a) predicting palm oil yield of a test oil palm plant according the method of any one of claims 1-18; and

(b) subjecting at least one cell of the test oil palm plant to cultivation in cell culture if the palm oil yield of the test oil palm plant is predicted to be higher than average for the population based on step (a).

21. A method of selecting a parental oil palm plant for use in breeding to obtain agricultural production plants or improved parental oil palm plants, the method comprising the steps of:

(a) predicting palm oil yield of a test oil palm plant according the method of any one of claims 1-18; and

(b) selecting the test oil palm plant for use in breeding if the palm oil yield of *tenera* progeny of the test oil palm plant is predicted to be higher than average for the population based on step (a).

22. A SNP detection kit for predicting palm oil yield of a test oil palm plant, the kit comprising:

(i) a set of at least 21 nucleotide molecules suitable for determining, from a sample of a test oil palm plant of a population of oil palm plants, a first SNP genotype to a twenty-first SNP genotype, respectively, of the test oil palm plant, the first SNP genotype to the twenty-first SNP genotype corresponding to a first SNP marker to a twenty-first SNP marker, respectively, the first SNP marker to the twenty-first SNP marker (a) being located in a first QTL to a twenty-first QTL, respectively, for a high-oil-production trait in the population and (b) being associated, after stratification and kinship correction, with the high-oil-production trait with a genome-wide $-\log_{10}(p\text{-value})$ of a least 4.0 in the population or having linkage disequilibrium r^2 values of at least 0.2 with respect to a first other SNP marker to a twenty-first other SNP marker, respectively, that are linked thereto and associated, after stratification and kinship correction, with the high-oil-production trait with a genome-wide $-\log_{10}(p\text{-value})$ of at least 4.0 in the population; and

(ii) a reference sample of a reference high-oil-yielding oil palm plant of the population,

wherein the first QTL to the twenty-first QTL are regions of the oil palm genome corresponding, respectively, to:

- (1) QTL region 1, extending from nucleotide 66542323 to 66776312 of chromosome 1;
- (2) QTL region 2, extending from nucleotide 66807385 to 67299617 of chromosome 1;
- (3) QTL region 3, extending from nucleotide 62277032 to 62355782 of chromosome 2;
- (4) QTL region 4, extending from nucleotide 31132787 to 31173962 of chromosome 4;
- (5) QTL region 5, extending from nucleotide 32863621 to 32964104 of chromosome 5;
- (6) QTL region 6, extending from nucleotide 33355931 to 33509217 of chromosome 5;
- (7) QTL region 7, extending from nucleotide 33658904 to 34233352 of chromosome 5;
- (8) QTL region 8, extending from nucleotide 34358119 to 34997228 of chromosome 5;
- (9) QTL region 9, extending from nucleotide 35004388 to 35125743 of chromosome 5;
- (10) QTL region 10, extending from nucleotide 35191678 to 35193677 of chromosome 5;
- (11) QTL region 11, extending from nucleotide 36108847 to 36272808 of chromosome 5;
- (12) QTL region 12, extending from nucleotide 39210662 to 39225076 of chromosome 5;
- (13) QTL region 13, extending from nucleotide 39518005 to 40469897 of chromosome 5;
- (14) QTL region 14, extending from nucleotide 40535309 to 40690150 of chromosome 5;
- (15) QTL region 15, extending from nucleotide 40789706 to 40983955 of chromosome 5;
- (16) QTL region 16, extending from nucleotide 41001085 to 41302446 of chromosome 5;
- (17) QTL region 17, extending from nucleotide 3050807 to 3241977 of chromosome 8;
- (18) QTL region 18, extending from nucleotide 5354764 to 5445890 of chromosome 8;
- (19) QTL region 19, extending from nucleotide 29488933 to 29602300 of chromosome 9;
- (20) QTL region 20, extending from nucleotide 4797284 to 5717606 of chromosome 11; and
- (21) QTL region 21, extending from nucleotide 8611715 to 8857914 of chromosome 15;

wherein the population of oil palm plants comprises a Nigerian *dura* X AVROS *pisifera* population, and the first SNP marker comprises one or more SNP markers as follows:

SNP No.	QTL region	Chromosome	Position	Major allele	Minor allele
2	2	1	66972538	A	G
4	2	1	67248054	A	C
5	3	2	62287970	G	A
6	4	4	31149521	G	A
8	6	5	33457617	G	A
10	7	5	34027909	A	G
15	8	5	34568717	G	A
17	8	5	34612126	G	A
19	8	5	34828628	G	A
20	8	5	34856258	C	A
21	8	5	34863191	A	G
22	8	5	34980654	A	G
24	9	5	35080572	G	A

26	9	5	35121695	A	G
27	10	5	35192678	A	G
29	11	5	36218554	A	G
30	11	5	36234729	G	A
31	12	5	39210662	A	G
34	13	5	39613906	G	A
35	13	5	39620126	A	G
37	13	5	39653455	G	A
38	13	5	39763460	A	C
43	13	5	39860983	G	A
44	13	5	39966907	A	G
45	13	5	40066844	G	A
46	13	5	40112108	G	A
47	13	5	40129189	C	A
48	13	5	40145585	A	G
49	13	5	40158838	A	G
50	13	5	40249343	G	A
52	13	5	40300709	A	G
53	13	5	40396733	A	C
54	13	5	40405137	A	G
55	14	5	40577880	G	A
56	14	5	40587726	A	G
57	14	5	40597291	A	G
58	15	5	40902353	A	G
60	16	5	41036282	A	G
61	16	5	41189141	A	G
62	16	5	41200140	G	A
63	16	5	41203008	A	G
64	16	5	41204629	G	A
65	17	8	3154572	G	A
66	18	8	5393101	C	A
68	20	11	4828172	A	G
76	20	11	5412986	A	G
77	20	11	5420279	A	G
78	20	11	5422706	G	A
79	20	11	5436885	G	A
81	20	11	5442401	C	A
82	21	15	8740489	A	G

and wherein the population of oil palm plants comprises a Deli *dura* X AVROS *pisifera* population, and the first SNP marker comprises one or more SNP markers as follows:

SNP No.	QTL region	Chromosome	Position	Major allele	Minor allele
1	1	1	66639699	A	G
3	2	1	67033874	A	G
6	4	4	31149521	G	A
7	5	5	32866989	A	G
8	6	5	33457617	G	A
9	7	5	33975929	G	A
10	7	5	34027909	A	G
11	7	5	34142044	C	A

12	8	5	34454937	G	A
13	8	5	34521370	G	A
14	8	5	34525454	G	A
15	8	5	34568717	G	A
16	8	5	34603394	G	A
17	8	5	34612126	G	A
18	8	5	34773282	A	G
19	8	5	34828628	G	A
20	8	5	34856258	C	A
21	8	5	34863191	A	G
22	8	5	34980654	A	G
23	9	5	35070248	A	C
24	9	5	35080572	G	A
25	9	5	35100527	A	G
26	9	5	35121695	A	G
27	10	5	35192678	A	G
28	11	5	36158880	G	A
29	11	5	36218554	A	G
30	11	5	36234729	G	A
31	12	5	39210662	A	G
32	13	5	39607208	A	G
33	13	5	39610847	G	A
34	13	5	39613906	G	A
35	13	5	39620126	A	G
36	13	5	39642505	C	A
37	13	5	39653455	G	A
38	13	5	39763460	A	C
39	13	5	39799450	A	G
40	13	5	39804720	G	A
41	13	5	39805514	G	A
42	13	5	39806797	A	G
43	13	5	39860983	G	A
44	13	5	39966907	A	G
45	13	5	40066844	G	A
46	13	5	40112108	G	A
47	13	5	40129189	C	A
48	13	5	40145585	A	G
49	13	5	40158838	A	G
50	13	5	40249343	G	A
51	13	5	40269392	G	A
52	13	5	40300709	A	G
53	13	5	40396733	A	C
54	13	5	40405137	A	G
58	15	5	40902353	A	G
59	15	5	40916454	G	A
60	16	5	41036282	A	G
61	16	5	41189141	A	G
62	16	5	41200140	G	A
63	16	5	41203008	A	G
64	16	5	41204629	G	A
67	19	9	29537541	0	G
68	20	11	4828172	A	G
69	20	11	4831923	G	A

70	20	11	4838915	G	A
71	20	11	5165520	0	G
72	20	11	5204949	0	G
73	20	11	5255710	C	A
74	20	11	5369351	A	C
75	20	11	5372237	G	A
76	20	11	5412986	A	G
77	20	11	5420279	A	G
78	20	11	5422706	G	A
79	20	11	5436885	G	A
80	20	11	5439423	G	A
81	20	11	5442401	C	A
82	21	15	8740489	A	G

23. The SNP detection kit of claim 22, further comprising a solid substrate, the nucleotide molecules being attached to the solid substrate.

24. The SNP detection kit of claims 22 or 23, wherein the nucleotide molecules are oligonucleotide or polynucleotides.