

Title: Methods and Kits for Selecting Plants with High Oil Yield or Predicting Oil Yield

Technical Field

- 5 This application relates to methods and kits for selecting plants with high oil yield or predicting oil yield, and more particularly to methods for selecting oil palm plant with high oil yield, methods for predicting oil yield of a test oil palm plant, and kits for increasing oil yield of an oil palm plant or predicting oil yield of a test oil palm plant.

10 **Background Art**

The African oil palm *Elaeis guineensis* Jacq. is an important oil-food crop. Oil palm plants are monoecious, i.e. single plants produce both male and female flowers, and are characterized by alternating series of male and female inflorescences. The male inflorescence is made up of numerous spikelets, and can bear well over 100,000 flowers.

- 15 Oil palm is naturally cross-pollinated by insects and wind. The female inflorescence is a spadix which contains several thousands of flowers borne on thorny spikelets. A bunch carries 500 to 4,000 fruits. The oil palm fruit is a sessile drupe that is spherical to ovoid or elongated in shape and is composed of an exocarp, a mesocarp containing palm oil, and an endocarp surrounding a kernel.

- 20 Oil palm is important both because of its high yield and because of the high quality of its oil. Regarding yield, oil palm is the highest yielding oil-food crop, with a recent average yield of 3.67 tonnes per hectare per year and with best progenies known to produce about 10 tonnes per hectare per year. Oil palm is also the most efficient plant known for harnessing the energy of sunlight for producing oil. Regarding quality, oil
- 25 palm is cultivated for both palm oil, which is produced in the mesocarp, and palm kernel oil, which is produced in the kernel. Palm oil in particular is a balanced oil, having almost equal proportions of saturated fatty acids ($\approx 55\%$ including 45% of palmitic acid) and unsaturated fatty acids ($\approx 45\%$), and it includes beta carotene. The palm kernel oil is more saturated than the mesocarp oil. Both are low in free fatty acids. The current
- 30 combined output of palm oil and palm kernel oil is about 50 million tonnes per year, and demand is expected to increase substantially in the future with increasing global population and per capita consumption of oils and fats.

Although oil palm is the highest yielding oil-food crop, current oil palm crops produce well below their theoretical maximum. Moreover, conventional methods for identifying potential high-yielding palms for use in crosses to generate progeny with higher yields require cultivation of palms and measurement of production of oil thereby over the course of many years, which is both time and labor intensive. In addition, conventional breeding techniques for propagation of oil palm for oil production are also time and labor intensive, particularly because the most productive, and thus commercially relevant, palms exhibit a hybrid phenotype which makes propagation thereof by direct hybrid crosses impractical. Accordingly, a need exists to improve oil palm through improved methods for increasing oil yield of an oil palm plant and predicting oil yield of a test oil palm plant.

Methods for selection based on biomarkers, such as proteins, may offer solutions. For example, difference gel electrophoresis ("DIGE") analysis, corresponding to two dimensional gel electrophoresis employing sensitive fluorescent labeling dyes, as described by Mackintosh *et al.*, 3 Proteomics 2273-88 (2003), has been successfully employed in protein expression analyses in rice and sunflower, as described by Teshima *et al.*, Regulatory Toxicology & Pharmacology (article in press), and Hajduch *et al.*, 6 Journal of Proteome Research 3232-41 (2007), respectively. In rice, this approach was used to differentiate one cultivar from others, and also to compare expression of allergen proteins. In sunflower, several leads in seed oil traits have been identified for further investigation. Genetic modification of particular oil plants to express particular genes has also been reported to result in increased oil yields. For example, Zank *et al.*, US Pub. No. 2009/0083882, discloses methods of increasing the total oil content and/or glycerol 3-phosphate content in transgenic oil crop plants by expressing glycerol-3-phosphate dehydrogenases from yeasts therein and notes that an increase in content of triacylglycerides is achieved by increasing the glycerol-3-phosphate dehydrogenase activity. Also for example, Zou *et al.*, U.S. Pat. No. 6,825,039, discloses a method for increasing oil content in a plant by modulating pyruvate dehydrogenase kinase protein expression, including a step of stably transforming a plant cell with a plant pyruvate dehydrogenase kinase polynucleotide operably linked to a promoter. Also for example, Zou *et al.*, U.S. Pat. No. 7,057,091, discloses a method of changing the oil or biopolymer

content of a plant, plant storage organ, or plant seed, the method including introducing a sense or anti-sense nucleic acid construct into a plant transformation vector to produce a modified plant transformation vector, wherein the sense or anti-sense nucleic acid construct includes an isolated, purified, or recombinant nucleic acid encoding a *Brassica* pyruvate dehydrogenase kinase protein, among other steps. However, given the great diversity of biochemical pathways and regulatory mechanisms across plants, it is not apparent that approaches for increasing oil yields in specific plant species by genetic modification are generalizable to other plant species. It is also not apparent whether selection based on particular combinations of biomarkers, e.g. levels of protein expression, may provide for further increases in oil yields, and if so what those combinations may be.

Disclosure of Invention

In one example embodiment, a method for selecting oil palm plant with high oil yield is disclosed. The method comprises determining the level of triose-phosphate isomerase and the level of glycerol-3-phosphate dehydrogenase in mesocarp tissue of a fruit of a parental oil palm plant 11 to 21 weeks after pollination thereof in comparison to mesocarp tissue of a fruit of a reference oil palm plant 11 to 21 weeks after pollination thereof. The method also comprises selecting progeny of the parental oil palm plant based on the level of triose-phosphate isomerase being lower and the level of glycerol-3-phosphate dehydrogenase being higher for the parental oil palm plant in comparison to the reference oil palm plant, to obtain a high-yielding oil palm plant.

In another example embodiment, a method for predicting oil yield of a test oil palm plant is disclosed. The method comprises determining the level of triose-phosphate isomerase and the level of glycerol-3-phosphate dehydrogenase in mesocarp tissue of a fruit of a test oil palm plant 11 to 21 weeks after pollination thereof in comparison to mesocarp tissue of a fruit of a reference oil palm plant 11 to 21 weeks after pollination thereof. The method also comprises predicting oil yield of the test oil palm plant based on the level of triose-phosphate isomerase being lower and the level of glycerol-3-phosphate dehydrogenase being higher for the test oil palm plant in comparison to the reference oil palm plant.

In another example embodiment, a kit for increasing oil yield of an oil palm plant or predicting oil yield of a test oil palm plant is disclosed. The kit comprises an antibody for detection of triose-phosphate isomerase. The kit also comprises an antibody for detection of glycerol-3-phosphate dehydrogenase. The kit also comprises an extract of a mesocarp tissue of a fruit of a reference oil palm plant.

Brief Description of Drawings

FIG. 1 is a graph of relative amounts of triose-phosphate isomerase protein (density) versus time (12 to 22 weeks after pollination), for mesocarp tissue of fruits of high-yielding oil palm plants (diamonds) versus low-yielding oil palm plants (X symbols). Statistically significant differences are noted (stars).

FIG. 2 is a graph of relative amounts of fructose-1,6-bisphosphate aldolase protein (density) versus time (12 to 22 weeks after pollination), for mesocarp tissue of fruits of high-yielding oil palm plants (diamonds) versus low-yielding oil palm plants (X symbols). Statistically significant differences are noted (stars).

FIG. 3 is a graph of relative amounts of ATP citrate lyase protein (expression intensity) versus time (12 to 22 weeks after pollination), for mesocarp tissue of fruits of high-yielding oil palm plants (squares) versus low-yielding oil palm plants (X symbols). Statistically significant differences are noted (stars).

FIG. 4 is a graph of relative amounts of pyruvate kinase protein (expression intensity) versus time (12 to 22 weeks after pollination), for mesocarp tissue of fruits of high-yielding oil palm plants (squares) versus low-yielding oil palm plants (X symbols). Statistically significant differences are noted (stars).

FIG. 5 is a graph of relative amounts of aconitase protein (expression intensity) versus time (12 to 22 weeks after pollination), for mesocarp tissue of fruits of high-yielding oil palm plants (squares) versus low-yielding oil palm plants (X symbols). Statistically significant differences are noted (stars).

FIG. 6 is a graph of total lipid content (grams per gram of biomass) for yeast strains corresponding to wild-type (1), a strain that over-expresses fructose-1,6-bisphosphate aldolase (2), a strain that over-expresses glycerol-3-phosphate dehydrogenase (3), a strain that over-expresses cytosolic triose-phosphate isomerase (4),

a strain that over-expresses plastidial triose-phosphate isomerase (5), and a strain that over-expresses glyceraldehyde-3-phosphate dehydrogenase (6). Error bars are shown.

FIG. 7 is a graph of relative concentration of 16:0 fatty acids, i.e. saturated 16-carbon fatty acids (expressed as relative abundance), versus time (12 to 22 weeks after
5 pollination), for mesocarp tissue of fruits of high-yielding oil palm plants (triangles) versus low-yielding oil palm plants (circles). Error bars are shown. Statistically significant differences based on a p -value < 0.005 are noted (*).

FIG. 8 is a graph of relative concentration of 18:0 fatty acids, i.e. saturated 18-carbon fatty acids (expressed as relative abundance), versus time (12 to 22 weeks after
10 pollination), for mesocarp tissue of fruits of high-yielding oil palm plants (triangles) versus low-yielding oil palm plants (circles). Error bars are shown. Statistically significant differences based on a p -value < 0.005 are noted (*).

FIG. 9 is a graph of relative concentration of 18:1 fatty acids, i.e. mono-unsaturated 18-carbon fatty acids (expressed as relative abundance), versus time (12 to 22
15 weeks after pollination), for mesocarp tissue of fruits of high-yielding oil palm plants (triangles) versus low-yielding oil palm plants (circles). Error bars are shown. Statistically significant differences based on a p -value < 0.005 are noted (*).

FIG. 10 is a graph of concentration of fructose 1,6-bisphosphate (expressed as nmol/g dry weight), versus time (12 to 22 weeks after pollination), for mesocarp tissue of
20 fruits of high-yielding oil palm plants (triangles) versus low-yielding oil palm plants (circles). Error bars are shown. Statistically significant differences based on a p -value < 0.005 are noted (*).

FIG. 11 is a graph of concentration of glycerol 3-phosphate (expressed as nmol/g dry weight), versus time (12 to 22 weeks after pollination), for mesocarp tissue of fruits
25 of high-yielding oil palm plants (triangles) versus low-yielding oil palm plants (circles). Error bars are shown. Statistically significant differences based on a p -value < 0.005 are noted (*).

FIG. 12 is a graph of concentration of 3-phosphoglyceric acid (expressed as nmol/g dry weight), versus time (12 to 22 weeks after pollination), for mesocarp tissue of
30 fruits of high-yielding oil palm plants (triangles) versus low-yielding oil palm plants

(circles). Error bars are shown. Statistically significant differences based on a p -value < 0.005 are noted (*).

FIG. 13 is a graph of concentration of malic acid (expressed as nmol/g dry weight), versus time (12 to 22 weeks after pollination), for mesocarp tissue of fruits of high-yielding oil palm plants (triangles) versus low-yielding oil palm plants (circles). Error bars are shown. Statistically significant differences based on a p -value < 0.005 are noted (*).

FIG. 14 is a graph of concentration of isocitric acid (expressed as nmol/g dry weight), versus time (12 to 22 weeks after pollination), for mesocarp tissue of fruits of high-yielding oil palm plants (triangles) versus low-yielding oil palm plants (circles). Error bars are shown. Statistically significant differences based on a p -value < 0.005 are noted (*).

FIG. 15 is a graph of concentration of 2-oxoglutaric acid (expressed as nmol/g dry weight), versus time (12 to 22 weeks after pollination), for mesocarp tissue of fruits of high-yielding oil palm plants (triangles) versus low-yielding oil palm plants (circles). Error bars are shown. Statistically significant differences based on a p -value < 0.005 are noted (*).

FIG. 16 is a graph of the ratios of malic acid to citric acid versus time (12 to 22 weeks after pollination), for mesocarp tissue of fruits of high-yielding oil palm plants (triangles) versus low-yielding oil palm plants (circles) (n=8).

Best Mode for Carrying Out the Invention

The application is drawn to methods for selecting oil palm plant with high oil yield comprising: (i) determining the level of triose-phosphate isomerase and the level of glycerol-3-phosphate dehydrogenase in mesocarp tissue of a fruit of a parental oil palm plant 11 to 21 weeks after pollination thereof in comparison to mesocarp tissue of a fruit of a reference oil palm plant 11 to 21 weeks after pollination thereof; and (ii) selecting progeny of the parental oil palm plant based on the level of triose-phosphate isomerase being lower and the level of glycerol-3-phosphate dehydrogenase being higher for the parental oil palm plant in comparison to the reference oil palm plant, to obtain a high-yielding oil palm plant. The application is also drawn to methods for predicting oil yield

of a test oil palm plant comprising: (i) determining the level of triose-phosphate isomerase and the level of glycerol-3-phosphate dehydrogenase in mesocarp tissue of a fruit of a test oil palm plant 11 to 21 weeks after pollination thereof in comparison to mesocarp tissue of a fruit of a reference oil palm plant 11 to 21 weeks after pollination thereof; and (ii) predicting oil yield of the test oil palm plant based on the level of triose-phosphate isomerase being lower and the level of glycerol-3-phosphate dehydrogenase being higher for the test oil palm plant in comparison to the reference oil palm plant. The application is also drawn to kits for increasing oil yield of an oil palm plant or predicting oil yield of a test oil palm plant comprising: an antibody for detection of triose-phosphate isomerase; an antibody for detection of glycerol-3-phosphate dehydrogenase; and an extract of a mesocarp tissue of a fruit of a reference oil palm plant.

By studying and comparing gene expression, protein levels, and metabolite levels in populations of oil palm that yield high amounts of oil ("high-yielding") versus those that yield low amounts of oil ("low-yielding"), it has been determined that levels of triose-phosphate isomerase and glycerol-3-phosphate dehydrogenase in mesocarp tissue of a fruit of a parental oil palm plant can be used for selecting progeny to obtain a high-yielding oil palm plant. It has also been determined that levels of triose-phosphate isomerase and glycerol-3-phosphate dehydrogenase in mesocarp tissue of a fruit of a test oil palm plant can be used for predicting oil yield of the test oil palm plant. Without wishing to be bound by theory, it is believed that relatively lower levels of triose-phosphate isomerase and relatively higher levels of glycerol-3-phosphate dehydrogenase, e.g. in the parental or test oil palm plant in comparison to a reference oil palm plant, is indicative of increasing carbon flux toward dihydroxyacetone phosphate at a critical branch point in biosynthesis, based on the lower levels of triose-phosphate isomerase, combined with converting dihydroxyacetone phosphate to glycerol 3-phosphate efficiently, based on the higher levels of glycerol-3-phosphate dehydrogenase. It is believed that this combination in particular allows for increased production of glycerol 3-phosphate, which provides the backbone molecule of lipids including triglycerides, without having dihydroxyacetone phosphate accumulate to toxically high levels.

It has also been determined that relatively higher levels of malate dehydrogenase and/or relatively higher levels of ATP citrate lyase can also be used for selecting progeny

to obtain a high-yielding oil palm plant and for predicting oil yield of a test oil palm plant. Without wishing to be bound by theory, it is believed that the higher levels of malate dehydrogenase increase conversion of malic acid to citric acid and that the higher levels of ATP citrate lyase provide for increased conversion of citric acid to acetyl-CoA for fatty acid biosynthesis, with each also contributing to increased oil yield in oil palm.

It has also been determined that relatively lower levels of glyceraldehyde-3-phosphate dehydrogenase, relatively higher levels of fructose-1,6-bisphosphate aldolase, relatively higher levels of pyruvate kinase, and/or relatively higher levels of aconitase can also be used for selecting progeny to obtain a high-yielding oil palm plant and for predicting oil yield of a test oil palm plant. Without wishing to be bound by theory, it is believed that the lower levels of glyceraldehyde-3-phosphate dehydrogenase cause a further shift in the equilibrium toward dihydroxyacetone phosphate by slowing consumption of glyceraldehyde-3-phosphate, that the higher levels of fructose-1,6-bisphosphate aldolase cause increased flux through the glycolytic pathway, that the higher levels of pyruvate kinase cause increased production of pyruvate for carbon supply to the TCA cycle and acetyl-CoA, and that the higher levels of aconitase may benefit equilibria in the TCA cycle, with each further contributing to increased oil yield in oil palm.

The term “parental oil palm plant,” as used herein, means an oil palm plant from which progeny have been generated, are generated, or will be generated during the course of carrying out methods for selecting oil palm plant with high oil yield as disclosed herein or using kits for increasing oil yield of an oil palm plant as disclosed herein.

The term “test oil palm plant,” as used herein, means an oil palm plant which has been subjected, is subjected, or will be subjected to a step of determining the level of triose-phosphate isomerase and the level of glycerol-3-phosphate dehydrogenase in mesocarp tissue of a fruit thereof during the course of carrying out methods for predicting oil yield of a test oil palm plant as disclosed herein.

The term “reference oil palm plant,” as used herein, means an oil palm plant used as a basis for comparison in determining oil palm yield traits. The reference oil palm plant can be, for example, an oil palm plant that produces high, average, or low amounts of palm oil, depending on the context of the particular application. For example, the

reference oil palm plant can be an oil palm plant that produces 10, 9, 8, 7, 6, 5, 4, 3, 2, or 1 tonnes of palm per hectare per year.

The terms "high-yielding," "low-yielding," and "oil yield," as used herein with respect to the methods and kits disclosed herein, refer to yields of palm oil in mesocarp
5 tissue of fruits of palm oil plants.

The singular forms "a," "an," and "the" are intended to include the plural forms as well, unless the context clearly indicates otherwise.

As noted above, a method for selecting oil palm plant with high oil yield is disclosed. The method for selecting oil palm plant with high oil yield can comprise
10 determining the level of triose-phosphate isomerase and the level of glycerol-3-phosphate dehydrogenase in mesocarp tissue of a fruit of a parental oil palm plant 11 to 21 weeks after pollination thereof in comparison to mesocarp tissue of a fruit of a reference oil palm plant 11 to 21 weeks after pollination thereof.

Triose-phosphate isomerase catalyzes the interconversion of glyceraldehyde 3-
15 phosphate and dihydroxyacetone phosphate in plant cell cytosol and plastids and plays an important role in metabolism by replenishing glyceraldehyde-3-phosphate in plant cells as it is consumed during glycolysis. Accordingly, relatively lower levels of triose-phosphate isomerase can result in a relative increase in dihydroxyacetone phosphate in oil palm cells. In some embodiments the triose-phosphate isomerase can comprise a
20 cytosolic isoform, e.g. SEQ ID NO: 1. In some embodiments the triose-phosphate isomerase can comprise a plastidial isoform, e.g. SEQ ID NO: 2.

Glycerol-3-phosphate dehydrogenase catalyzes the reversible redox conversion of dihydroxyacetone phosphate to glycerol 3-phosphate and serves to link carbohydrate metabolism and lipid metabolism. Accordingly, the increase in glycerol-3-phosphate
25 dehydrogenase activity can result in an increase in glycerol 3-phosphate available for lipid production. In some embodiments, the glycerol-3-phosphate dehydrogenase of the parental oil palm plant can comprise a plastidial isoform, e.g. SEQ ID NO: 3.

The level of triose-phosphate isomerase and the level of glycerol-3-phosphate dehydrogenase in mesocarp tissue of a fruit of a parental oil palm plant can be
30 determined, for example, in a preparation of proteins from mesocarp tissue, e.g. a crude preparation, a minimally purified preparation, or a highly purified preparation of

mesocarp proteins. The preparation can include total mesocarp proteins, or a subset of mesocarp proteins, e.g. soluble proteins, insoluble proteins, proteins having an isoelectric point between pH 4 to 7, or proteins having higher or lower isoelectric points. The mesocarp tissue itself may be obtained at a particular developmental stage of the fruit from which it is derived, at any time following pollination. As is known in the art, the onset of oil deposition in oil palm follows pollination according to the following time frame. Oil deposition in the endosperm starts at approximately 12 weeks after pollination and is almost complete by 16 weeks after pollination. Oil deposition in mesocarp starts at approximately 15 weeks after pollination and continues until fruit maturity at about 20 weeks after pollination. More specifically, 12 weeks after pollination marks the start of oil deposition in endosperm but precedes the start of oil deposition in mesocarp, 16 weeks after pollination marks the point of highest transcript expression level in mesocarp, following the initiation of oil biosynthesis after pollination, and 18 weeks after pollination marks the time at which transcript expression would be expected to decrease as the fruit matures. Thus, for example, the mesocarp tissue itself may be obtained 11-21 weeks after pollination, 13-21 weeks after pollination, 15-21 weeks after pollination, 17-21 weeks after pollination, 19-21 weeks after pollination, 11-19 weeks after pollination, 13-19 weeks after pollination, 15-19 weeks after pollination, 17-19 weeks after pollination, 11-17 weeks after pollination, 13-17 weeks after pollination, 15-17 weeks after pollination, 11-15 weeks after pollination, 13-15 weeks after pollination, or 11-13 weeks after pollination, or 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, or 21 weeks after pollination.

The level of the protein can be expressed in terms such as mass protein per mass mesocarp tissue, intensity of signal of protein relative to intensity of signal of reference, or other suitable terms.

At least one of the level of triose-phosphate isomerase and the level of glycerol-3-phosphate dehydrogenase in mesocarp tissue of a fruit of a parental oil palm plant can be determined, for example, by a technique such as an in vitro assay, two-dimensional fluorescence difference gel electrophoresis, antibody-based detection, immunoblot detection, dot-blot detection, liquid chromatography-mass spectrometry, an enzyme activity assay, reverse transcriptase polymerase chain reaction (also termed RT-PCR),

RNA-seq, hybridization, or microarray. As will be appreciated, an in vitro assay encompasses the various techniques noted, as well as others. Two-dimensional fluorescence difference gel electrophoresis can include CyDye labeling of total proteins in a sample, followed by separation and detection of the protein. Antibody-based
5 detection can be carried out, for example, by use of monoclonal antibodies or polyclonal antibodies raised against the protein, for quantitative or qualitative detection, and includes immunoblot detection and dot-blot detection, among others. Liquid chromatography-mass spectrometry can be carried out for quantitation and comparison of protein based, for example, on quantitation of spectral features. An enzyme activity
10 assay can be used to measure levels based on activity. Analyses based on RT-PCR, RNA-seq, hybridization, or microarray can be used to measure levels based on quantitation of corresponding transcripts.

As will also be appreciated, the level of triose-phosphate isomerase and the level of glycerol-3-phosphate dehydrogenase can be determined in parallel or in series, on the
15 same or different samples of mesocarp tissue, and/or by the same or different techniques.

The step of determining the level of triose-phosphate isomerase and the level of glycerol-3-phosphate dehydrogenase in mesocarp tissue of a fruit of a parental oil palm plant can be carried out in comparison to mesocarp tissue of a fruit of a reference oil palm plant. The comparison can be made, for example, by determining the level of
20 triose-phosphate isomerase and the level of glycerol-3-phosphate dehydrogenase in the mesocarp tissue of the fruit of the reference oil palm plant as described above for the parental oil palm plant, e.g. in a preparation of proteins from mesocarp tissue, including total mesocarp proteins or a subset of mesocarp proteins, at a particular developmental stage of the fruit such as 11-21 weeks after pollination, and/or by a technique such as an
25 in vitro assay. Again, the level of triose-phosphate isomerase and the level of glycerol-3-phosphate dehydrogenase can be determined in parallel or in series, on the same or different samples of mesocarp tissue, and/or by the same or different techniques. The comparison can also be made, for example, relative to previously determined levels of triose-phosphate isomerase and glycerol-3-phosphate dehydrogenase in the mesocarp
30 tissue of the fruit of the reference oil palm plant 11 to 21 weeks after pollination thereof, e.g. without need for determining the levels of triose-phosphate isomerase and glycerol-

3-phosphate dehydrogenase in the mesocarp tissue of the fruit of the reference oil palm plant at the time that the levels of triose-phosphate isomerase and glycerol-3-phosphate dehydrogenase in the mesocarp tissue of the fruit of the parental oil palm plant are determined.

5 The comparison can be made, for example, by comparing the respective level of triose-phosphate isomerase and the respective level of glycerol-3-phosphate dehydrogenase, for the parental oil palm plant and the reference oil palm plant, by an in vitro assay as discussed above, and checking for differences between the levels for the parental oil palm plant and the reference oil palm plant. In some embodiments such a
10 comparison is considered to reveal a biologically and/or statistically significant difference based, for example, on the level of triose-phosphate isomerase in the mesocarp tissue of the parental oil palm plant being lower than that of the reference oil palm plant by, for example, more than 1.1 fold, 1.25 fold, 1.5 fold, 2 fold, 4 fold, or more, with p values of, for example, <0.025 , <0.05 , or <0.1 . Also in some embodiments, such a comparison is
15 considered to reveal a biologically and/or statistically significant difference based, for example, on the level of glycerol-3-phosphate dehydrogenase in the mesocarp tissue of the parental oil palm being higher than that of the reference oil palm plant by, for example, more than 1.1 fold, 1.25 fold, 1.5 fold, 2 fold, 4 fold, or more, again with p values of, for example, <0.025 , <0.05 , or <0.1 . The comparison may be facilitated by use
20 of software for determining and comparing signal intensities, for example by use of Image Quant software (version 6.0, Amersham BioSciences), followed by Biological Variation Analysis using DeCyder (TM) 2D software version 6.5 (Amersham BioSciences).

25 The method for selecting oil palm plant with high oil yield can also comprise selecting progeny of the parental oil palm plant based on the level of triose-phosphate isomerase being lower and the level of glycerol-3-phosphate dehydrogenase being higher for the parental oil palm plant in comparison to the reference oil palm plant, to obtain a high-yielding oil palm plant. The step of selecting progeny of the parental oil palm plant can be carried out, for example, by choosing a parental oil palm plant for propagation
30 based on the comparison, i.e. the level of triose-phosphate isomerase being lower and the level of glycerol-3-phosphate dehydrogenase being higher, and crossing the plant with

another oil palm plant, e.g. another oil palm plant in which the level of triose-phosphate isomerase is lower and the level of glycerol-3-phosphate dehydrogenase is higher, by conventional breeding techniques to obtain progeny corresponding to the high-yielding oil palm plant.

- 5 In some embodiments the selecting is based on the level of triose-phosphate isomerase in the mesocarp tissue of the fruit of the parental oil palm plant 14 to 19 weeks after pollination thereof being lower in comparison to the mesocarp tissue of the fruit of the reference oil palm plant 14 to 19 weeks after pollination thereof. In some
10 embodiments the selecting is based on the level of glycerol-3-phosphate dehydrogenase in the mesocarp tissue of the fruit of the parental oil palm plant 14 to 19 weeks after
pollination thereof being higher in comparison to the mesocarp tissue of the fruit of the reference oil palm plant 14 to 19 weeks after pollination thereof.

As is well known in the art, fruit type is a monogenic trait in oil palm that is important with respect to breeding and commercial production of palm oil. Specifically,
15 oil palms with either of two distinct fruit types are generally used in breeding and seed production through crossing in order to generate palms for commercial production of oil (also termed "commercial planting materials" or "agricultural production plants"). The first fruit type is *dura* (genotype: sh+ sh+), which is characterized by a thick shell corresponding to 28 to 35% of the fruit by weight, with no ring of black fibres around the
20 kernel of the fruit. For *dura* fruits, the mesocarp to fruit ratio varies from 50 to 60%, with extractable oil content in proportion to bunch weight of 18 to 24%. The second fruit type is *pisifera* (genotype: sh- sh-), which is characterized by the absence of a shell, the vestiges of which are represented by a ring of fibres around a small kernel. Accordingly, for *pisifera* fruits, the mesocarp to fruit ratio is 90 to 100%. The mesocarp oil to bunch
25 ratio is comparable to the *dura* at 16 to 28%. *Pisiferas* are however usually female sterile as the majority of bunches abort at an early stage of development.

Crossing *dura* and *pisifera* gives rise to palms with a third fruit type, the *tenera* (genotype: sh+ sh-). *Tenera* fruits have thin shells of 8 to 10 % of the fruit by weight, corresponding to a thickness of 0.5 to 4 mm, around which is a characteristic ring of
30 black fibres. For *tenera* fruits, the ratio of mesocarp to fruit is comparatively high, in the range of 60 to 80%. Commercial *tenera* palms generally produce more fruit bunches than

duras, although mean bunch weight is lower. The extractable oil to bunch ratio is in the range of 20 to 30%, the highest of the three fruit types, and thus *tenera* are typically used as commercial planting materials.

Dura palm breeding populations used in Southeast Asia include Serdang Avenue, Ulu Remis (which incorporated some Serdang Avenue material), Johor Labis, and Elmina estate, including Deli Dumpy, all of which are derived from Deli *dura*. *Pisifera* breeding populations used for seed production are generally grouped as Yangambi, AVROS, Binga and URT. Other *dura* and *pisifera* populations are used in Africa and South America.

Accordingly, in some embodiments the parental oil palm plant is a *dura* palm selected from the group consisting of Deli *dura*, Serdang Avenue *dura*, Ulu Remis *dura*, Johor Labis *dura*, Elmina estate *dura*, and Deli Dumpy *dura*. Alternatively, in some embodiments the parental oil palm plant is a *pisifera* palm selected from the group consisting of Yangambi *pisifera*, AVROS *pisifera*, Binga *pisifera*, and URT *pisifera*.

Oil palm breeding is primarily aimed at selecting for improved parental *dura* and *pisifera* breeding stock palms for production of superior *tenera* commercial planting materials. Such materials are largely in the form of seeds although the use of tissue culture for propagation of clones continues to be developed. Generally, parental *dura* breeding populations are generated by crossing among selected *dura* palms. Based on the monogenic inheritance of fruit type, 100% of the resulting palms will be *duras*. After several years of yield recording and confirmation of bunch and fruit characteristics, *duras* are selected for breeding based on phenotype. In contrast, *pisifera* palms are normally female sterile and thus breeding populations thereof must be generated by crossing among selected *teneras* or by crossing selected *teneras* with selected *pisiferas*. The *tenera* x *tenera* cross will generate 25% *duras*, 50% *teneras* and 25% *pisiferas*. The *tenera* x *pisifera* cross will generate 50% *teneras* and 50% *pisiferas*. The yield potential of *pisiferas* is then determined indirectly by progeny testing with the elite *duras*, i.e. by crossing *duras* and *pisiferas* to generate *teneras*, and then determining yield phenotypes of the fruits of the *teneras* over time. From this, *pisiferas* with good general combining ability are selected based on the performance of their *tenera* progenies. Intercrossing among selected parents is also carried out with progenies being carried forward to the

next breeding cycle. This allows introduction of new genes into the breeding programme to increase genetic variability. Using this general scheme, priority selection objectives include high oil yield per unit area in terms of high fresh fruit bunch yield and high oil to bunch ratio (thin shell, thick mesocarp), high early yield (precocity), and good oil qualities, among other traits.

Accordingly, in some embodiments, the parental oil palm plant is a *dura* breeding stock plant, the progeny comprises an oil palm plant selected from the group consisting of a *dura* breeding stock plant and a *tenera* agricultural production plant, and the high-yielding oil palm plant is selected from the group consisting of a *dura* breeding stock plant and a *tenera* agricultural production plant. For example, in some embodiments the method is carried out with the purpose of generating improved *dura* breeding stock, in which case the parental *dura* breeding stock plant is crossed with another *dura* breeding stock plant to obtain a high-yielding oil palm plant directly among the progeny, which will also be *dura* breeding stock plants. Also for example, in some embodiments the method is carried out with the purpose of generating improved *tenera* agricultural production plants, in which case the parental *dura* breeding stock plant is crossed with a *pisifera* breeding stock plant to obtain a high-yielding oil palm plant directly among the progeny, which will be *tenera* agricultural production plants.

Alternatively, in some embodiments the parental oil palm plant is a *tenera* breeding stock plant, the progeny comprises an oil palm plant selected from the group consisting of a *tenera* breeding stock plant, a *pisifera* breeding stock plant, and a *tenera* agricultural production plant, and the high-yielding oil palm plant is selected from the group consisting of a *tenera* breeding stock plant and a *tenera* agricultural production plant. For example, in some embodiments the method may be carried out with the purpose of generating improved *tenera* breeding stock, in which case the parental *tenera* breeding stock plant is crossed with another *tenera* breeding stock plant, to obtain a *tenera* high-yielding palm plant directly among the progeny, of which 25% will be *dura*, 50% will be *tenera*, and 25% will be *pisifera*. Also for example, in some embodiments the method is carried out with the purpose of generating improved *tenera* agricultural production plants, in which case the parental *tenera* breeding stock plant is crossed with a *pisifera* breeding stock plant, to yield progeny corresponding to 50% *tenera* and 50%

pisifera. The *pisifera* resulting from this cross can in turn be used as *pisifera* breeding stock for generation of *tenera* agricultural production plants.

Progeny plants may be cultivated by conventional approaches, e.g. seedlings may be cultivated in polyethylene bags in pre-nursery and nursery settings, raised for about 12
5 months, and then planted as seedlings, with progeny that are known or predicted to exhibit high yields chosen for further cultivation.

In accordance with the method, the oil yield of the high-yielding oil palm plant can be increased by at least 5% relative to the parental oil palm plant. For example, the oil yield of the high-yielding oil palm plant can be increased by at least 10%, 15%, 20%,
10 or more, relative to the parental oil palm plant. The oil yield can also be increased without causing a deleterious shift in the profile of fatty acids of the oil, e.g. without causing a deleterious shift in the profile of C16:0, C18:0, and/or C18:1 fatty acids.

The method for selecting oil palm plant with high oil yield can further comprise determining at least one of the level of malate dehydrogenase and the level of ATP citrate
15 lyase in the mesocarp tissue of the fruit of the parental oil palm plant 11 to 21 weeks after pollination thereof in comparison to the mesocarp tissue of the fruit of the reference oil palm plant 11 to 21 weeks after pollination thereof. In some embodiments the malate dehydrogenase can comprise SEQ ID NO: 4. In some embodiments the ATP citrate lyase can comprise SEQ ID NO: 5. The at least one level can be determined as described
20 above, e.g. in a preparation of proteins from mesocarp tissue, including total mesocarp proteins or a subset of mesocarp proteins, at a particular developmental stage of the fruit such as 14-19 weeks after pollination, and/or by a technique such as an in vitro assay. The comparison can also be made as described above, e.g. by comparing the at least one respective level, for the parental oil palm plant and the reference oil palm plant, by an in
25 vitro assay as discussed above, and checking for differences between the level for the parental oil palm plant and the reference oil palm plant.

The method for selecting oil palm plant with high oil yield can further comprise selecting progeny of the parental oil palm plant based also on at least one of the level of malate dehydrogenase being higher and the level of ATP citrate lyase being higher for the
30 parental oil palm plant in comparison to the reference oil palm plant, to obtain the high-yielding oil palm plant. The selecting can also be carried out as described above, e.g. by

choosing a parental oil palm plant for propagation based on the comparison, i.e. at least one of the level of malate dehydrogenase being higher and the level of ATP citrate lyase being higher, and crossing the plant with another oil palm plant, e.g. another oil palm plant in which at least one of the level of malate dehydrogenase is higher and the level of ATP citrate lyase is higher, by conventional breeding techniques to obtain progeny corresponding to the high-yielding oil palm plant.

In some embodiments, the step of selecting can be based on both the level of malate dehydrogenase being higher and the level of ATP citrate lyase being higher.

The method for selecting oil palm plant with high oil yield can further comprise determining at least one of the level of glyceraldehyde-3-phosphate dehydrogenase, the level of fructose-1,6-bisphosphate aldolase, the level of pyruvate kinase, and the level of aconitase in the mesocarp tissue of the fruit of the parental oil palm plant 11 to 21 weeks after pollination thereof in comparison to the mesocarp tissue of the fruit of the reference oil palm plant 11 to 21 weeks after pollination thereof. In some embodiments the glyceraldehyde-3-phosphate dehydrogenase can comprise SEQ ID NO: 6. In some embodiments the fructose-1,6-bisphosphate aldolase can comprise SEQ ID NO: 7. In some embodiments the pyruvate kinase can comprise SEQ ID NO: 8. In some embodiments the aconitase can comprise SEQ ID NO: 9. Again, the at least one level can be determined as described above, and the comparison can also be made as described above.

The method for selecting oil palm plant with high oil yield can further comprise selecting progeny of the parental oil palm plant based also on at least one of the level of glyceraldehyde-3-phosphate dehydrogenase being lower, the level of fructose-1,6-bisphosphate aldolase being higher, the level of pyruvate kinase being higher, and the level of aconitase being higher for the parental oil palm plant in comparison to the reference oil palm plant, to obtain the high-yielding oil palm plant. Again, the selecting can be carried out as described above.

In some embodiments, the step of selecting can be based on at least two or at least three of the level of glyceraldehyde-3-phosphate dehydrogenase being lower, the level of fructose-1,6-bisphosphate aldolase being higher, the level of pyruvate kinase being higher, and the level of aconitase being higher, e.g. the level of glyceraldehyde-3-

phosphate dehydrogenase being lower and the level of fructose-1,6-bisphosphate aldolase being higher. In some embodiments, the step of selecting can be based on all four of the level of glyceraldehyde-3-phosphate dehydrogenase being lower, the level of fructose-1,6-bisphosphate aldolase being higher, the level of pyruvate kinase being higher, and the
5 level of aconitase being higher.

A method for obtaining palm oil from a high-yielding oil palm plant is also disclosed. The method comprises obtaining the high-yielding oil palm plant as described above. The method also comprises isolating palm oil from a fruit of the high-yielding oil palm plant. The step of isolating palm oil may be carried out by conventional
10 approaches, e.g. harvesting of fruit bunches followed by extraction of oil, within 24 hours, from the fresh and non-wounded fruits thereof.

A method for predicting oil yield of a test oil palm plant is also disclosed. The method for predicting oil yield of a test oil palm plant can comprise determining the level of triose-phosphate isomerase and the level of glycerol-3-phosphate dehydrogenase in
15 mesocarp tissue of a fruit of a test oil palm plant 11 to 21 weeks after pollination thereof in comparison to mesocarp tissue of a fruit of a reference oil palm plant 11 to 21 weeks after pollination thereof. In some embodiments, the triose-phosphate isomerase of the test oil palm plant can comprise SEQ ID NO: 1 or SEQ ID NO: 2. In some embodiments, the glycerol-3-phosphate dehydrogenase of the test oil palm plant can comprise SEQ ID
20 NO: 3.

The level of triose-phosphate isomerase and the level of glycerol-3-phosphate dehydrogenase can be determined as described above, e.g. in a preparation of proteins from mesocarp tissue, including total mesocarp proteins or a subset of mesocarp proteins, and/or at a particular developmental stage of the fruit, e.g. 11-21 weeks after pollination,
25 13-21 weeks after pollination, 15-21 weeks after pollination, 17-21 weeks after pollination, 19-21 weeks after pollination, 11-19 weeks after pollination, 13-19 weeks after pollination, 15-19 weeks after pollination, 17-19 weeks after pollination, 11-17 weeks after pollination, 13-17 weeks after pollination, 15-17 weeks after pollination, 11-15 weeks after pollination, 13-15 weeks after pollination, or 11-13 weeks after
30 pollination, or 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, or 21 weeks after pollination. Moreover, at least one of the level of triose-phosphate isomerase and the level of

glycerol-3-phosphate dehydrogenase in mesocarp tissue of a fruit of a parental oil palm plant can be determined, for example, by a technique such as an in vitro assay, two-dimensional fluorescence difference gel electrophoresis, antibody-based detection, immunoblot detection, dot-blot detection, liquid chromatography-mass spectrometry, an
5 enzyme activity assay, RT-PCR, RNA-seq, hybridization, or microarray.

The comparison can also be made as described above, e.g. by comparing the respective level of triose-phosphate isomerase and the respective level of glycerol-3-phosphate dehydrogenase, for the test oil palm plant and the reference oil palm plant, by an in vitro assay as discussed above, and checking for differences between the levels for
10 the test oil palm plant and the reference oil palm plant. In some embodiments such a comparison is considered to reveal a biologically and/or statistically significant difference based, for example, on the level of triose-phosphate isomerase in the mesocarp tissue of the test oil palm plant being lower than that of the reference oil palm plant by, for example, more than 1.1 fold, 1.25 fold, 1.5 fold, 2 fold, 4 fold, or more, with p values of,
15 for example, <0.025, <0.05, or <0.1. Also in some embodiments, such a comparison is considered to reveal a biologically and/or statistically significant difference based, for example, on the level of glycerol-3-phosphate dehydrogenase in the mesocarp tissue of the test oil palm being higher than that of the reference oil palm plant by, for example, more than 1.1 fold, 1.25 fold, 1.5 fold, 2 fold, 4 fold, or more, again with p values of, for
20 example, <0.025, <0.05, or <0.1.

The method for predicting oil yield of a test oil palm plant can also comprise predicting oil yield of the test oil palm plant based on the level of triose-phosphate isomerase being lower and the level of glycerol-3-phosphate dehydrogenase being higher for the test oil palm plant in comparison to the reference oil palm plant. The predicting
25 may be carried out, for example, based on the extent to which the level of triose-phosphate isomerase is lower and/or the extent to which the level of glycerol-3-phosphate dehydrogenase is higher, and/or based on correlations between the levels and yield.

In some embodiments the predicting is based on the level of triose-phosphate isomerase in the mesocarp tissue of the fruit of the test oil palm plant 14 to 19 weeks after
30 pollination thereof being lower in comparison to the mesocarp tissue of the fruit of the reference oil palm plant 14 to 19 weeks after pollination thereof. In some embodiments

the predicting is based on the level of glycerol-3-phosphate dehydrogenase in the mesocarp tissue of the fruit of the test oil palm plant 14 to 19 weeks after pollination thereof being higher in comparison to the mesocarp tissue of the fruit of the reference oil palm plant 14 to 19 weeks after pollination thereof.

5 The method for predicting oil yield of a test oil palm plant can further comprise determining at least one of the level of malate dehydrogenase and the level of ATP citrate lyase in the mesocarp tissue of the fruit of the test oil palm plant 11 to 21 weeks after pollination thereof in comparison to the mesocarp tissue of the fruit of the reference oil palm plant 11 to 21 weeks after pollination thereof. Again, in some embodiments the
10 malate dehydrogenase can comprise SEQ ID NO: 4, and/or the ATP citrate lyase can comprise SEQ ID NO: 5. Again, the at least one level can be determined as described above, and the comparison can also be made as described above.

 The method for predicting oil yield of a test oil palm plant can further comprise predicting oil yield of the test oil palm plant based also on at least one of the level of
15 malate dehydrogenase being higher and the level of ATP citrate lyase being higher for the test oil palm plant in comparison to the reference oil palm plant. The predicting can also be carried out as described above.

 In some embodiments, the predicting can be based both the level of malate dehydrogenase being higher and the level of ATP citrate lyase being higher.

20 The method for selecting a test oil palm plant with high oil yield can further comprise determining at least one of the level of glyceraldehyde-3-phosphate dehydrogenase, the level of fructose-1,6-bisphosphate aldolase, the level of pyruvate kinase, and the level of aconitase in the mesocarp tissue of the fruit of the test oil palm plant 11 to 21 weeks after pollination thereof in comparison to the mesocarp tissue of the
25 fruit of the reference oil palm plant 11 to 21 weeks after pollination thereof. Again, in some embodiments the glyceraldehyde-3-phosphate dehydrogenase can comprise SEQ ID NO: 6, the fructose-1,6-bisphosphate aldolase can comprise SEQ ID NO: 7, the pyruvate kinase can comprise SEQ ID NO: 8, and/or the aconitase can comprise SEQ ID NO: 9. Again, the at least one level can be determined as described above, and the
30 comparison can also be made as described above.

The method for predicting oil yield of a test oil palm plant can further comprise predicting oil yield of the test oil palm plant based also on at least one of the level of glyceraldehyde-3-phosphate dehydrogenase being lower, the level of fructose-1,6-bisphosphate aldolase being higher, the level of pyruvate kinase being higher, and the
5 level of aconitase being higher for the test oil palm plant in comparison to the reference oil palm plant.

In some embodiments, the predicting can be based on at least two or at least three of the level of glyceraldehyde-3-phosphate dehydrogenase being lower, the level of fructose-1,6-bisphosphate aldolase being higher, the level of pyruvate kinase being
10 higher, and the level of aconitase being higher, e.g. the level of glyceraldehyde-3-phosphate dehydrogenase being lower and the level of fructose-1,6-bisphosphate aldolase being higher. In some embodiments, the predicting can be based on all four of the level of glyceraldehyde-3-phosphate dehydrogenase being lower, the level of fructose-1,6-bisphosphate aldolase being higher, the level of pyruvate kinase being higher, and the
15 level of aconitase being higher.

A kit for increasing oil yield of an oil palm plant or predicting oil yield of a test oil palm plant is also disclosed. The kit can comprise: (i) an antibody for detection of triose-phosphate isomerase; (ii) an antibody for detection of glycerol-3-phosphate dehydrogenase; and (iii) an extract of a mesocarp tissue of a fruit of a reference oil palm
20 plant. The triose-phosphate isomerase, as detected, can comprise SEQ ID NO: 1 or SEQ ID NO: 2. The glycerol-3-phosphate dehydrogenase, as detected, can comprise SEQ ID NO: 3. The antibodies can be prepared by methods that are well known in the art. As will be appreciated from the foregoing, the kit can be used to carry out the method for increasing oil yield of an oil palm plant and/or the method for predicting oil yield of a test
25 oil palm plant as discussed above.

The kit can further comprise instructions indicating use of the antibodies for: (i) determining the level of triose-phosphate isomerase and the level of glycerol-3-phosphate dehydrogenase in mesocarp tissue of a fruit of a parental or test oil palm plant 11 to 21
30 weeks after pollination thereof in comparison to the extract; and (ii) selecting progeny of the parental oil palm plant or predicting oil yield of the test oil palm plant based on the level of triose-phosphate isomerase being lower and the level of glycerol-3-phosphate

dehydrogenase being higher for the parental or test oil palm plant in comparison to the reference oil palm plant.

The following examples are for purposes of illustration and are not intended to limit the scope of the claims.

5

Example 1

Defining high-yielding oil palms and low-yielding oil palms. Two screening populations of oil palm plants, a high-yielding screening population (also termed HY) and a low-yielding screening population (also termed LY), were used. The screening
10 populations were derived from crosses of Serdang Avenue *dura* (at least 75% of Serdang Avenue *dura*) and AVROS *pisifera* (at least 75% of AVROS *pisifera*) to yield *tenera* progeny. The high-yielding and low-yielding palms are defined by the quantity of oil produced by them in tonnes per hectare per year. In this study, an oil palm was considered high-yielding if it produced more than 10 tonnes of oil per hectare per year,
15 and an oil palm was considered low-yielding if it produced less than 6 tonnes of oil per hectare per year. Both screening populations were derived from a Carey Island oil palm plantation. The yield determinations were based on historical oil yield data for each sample.

For the examples that follow, samples of mesocarp tissue of individual plants
20 selected from the high-yielding screening population and the low-yielding screening population were collected at specific time points of the fruit development. The mesocarp samples were sliced, pulverized in liquid nitrogen, and stored at -80 °C. The pulverized mesocarp samples were then extracted and profiled using high-throughput omics platforms including transcriptomics, proteomics, and metabolomics to identify potential
25 candidates/biomarkers (e.g. genes) related to yield trait, as follows.

Example 2

Arabidopsis Microarray Experiment

Arabidopsis one-color microarray-based gene expression analysis. Samples of
30 oil palm transcripts obtained from high-yielding palms and low yielding palms at 16 weeks after pollination were hybridized on 4 x 44K microarray formats. Procedures for

the preparation, labeling of complex biological targets, and hybridization, washing, scanning, and feature extraction of Agilent's 60-mer oligonucleotide microarrays for gene expression analysis were adapted from "One-color microarray-based gene expression analysis version 6.0" by Agilent Technologies.

- 5 **Results.** The Arabidopsis microarray data showed triose-phosphate isomerase was down-regulated (> 1.7 fold) in high-yielding palms compared to low-yielding palms at week 16 after pollination, with a p value of 0.02.

Example 3

10 **2-D DIGE Experiment**

- Preparation of samples.** A modified protein extraction method published by He *et al.* (1) was used to isolate total mesocarp protein from oil palm fruitlets. Subsequently, the extracted protein samples were resuspended in 2-D cell lysis buffer (30 mM Tris-HCl, pH 8.8, containing 7 M urea, 2 M thiourea and 4% CHAPS). The mixture was sonicated
15 at 4 °C, followed by shaking for 30 minutes at room temperature. The samples were then centrifuged for 30 minutes at 14,000 rpm and the resulting supernatant was collected. Protein concentration of the supernatant fraction was measured using Bio-Rad protein assay method (Bradford, 1976). To aid downstream analysis, an internal standard (IS) was prepared, by mixing equal amount of protein from each sample, and included in the
20 2D-DIGE experiment.

- CyDye labeling.** For each sample, 30 μ g of protein was mixed with 1.0 μ l of diluted CyDye, and kept in dark on ice for 30 minute. Samples from each pair of high-yielding and low-yielding palms were labeled with Cy3 and Cy5, respectively, while the internal standard was labeled with Cy2. The labeling reaction was stopped by adding 1.0
25 μ l of 10 mM Lysine to the sample, and incubating in dark on ice for additional 15 minutes. The 3 labeled samples were then mixed together. The 2X 2-D Sample buffer (8 M urea, 4% CHAPS, 20 mg/ml DTT, 2% pharmalytes and trace amount of bromophenol blue), 100 μ l Destreak solution and rehydration buffer (7 M urea, 2 M thiourea, 4% CHAPS, 20 mg/ml DTT, 1% pharmalytes and trace amount of bromophenol blue) were
30 added to the labeling mix to make the total volume of 250 μ l. The resulting mixture was

thoroughly mixed and spun down prior to loading the labeled samples onto immobilized pH gradient gel (IPG) strips housed in a strip holder.

Set up of 2D-DIGE analytical gels. DIGE gels were designed to contain the appropriate sample pairings in order to facilitate gel analysis in the later part of the experiment. A total of 9 DIGE gels were produced with the sample pairing.

Internal Standard. The internal standard (also termed IS) is used to match and normalize protein patterns across different DIGE gels, therefore negating the problem of inter-gel variation. It allows accurate quantification of differences between samples with an associated statistical significance. Quantitative comparisons of protein between samples are made on the relative change of each protein spot to its own in-gel internal standard.

IEF and SDS-PAGE. After loading the labeled samples onto pH 4-7 IPG strips, isoelectric focusing (IEF) was run following the protocol provided by Amersham BioSciences (GE Healthcare, 2004). Upon finishing the IEF, the IPG strips were incubated in the freshly made equilibration buffer-1 (50 mM Tris-HCl, pH 8.8, containing 6 M urea, 30% glycerol, 2% SDS, trace amount of bromophenol blue and 10 mg/ml DTT) for 15 minutes with gentle shaking. The strips were then rinsed in the freshly made equilibration buffer-2 (50 mM Tris-HCl, pH 8.8, containing 6 M urea, 30% glycerol, 2% SDS, trace amount of bromophenol blue and 45 mg/ml DTT) for 10 minutes with gentle shaking. Following that, the IPG strips were rinsed in the SDS-polyacrylamide gel electrophoresis (SDS-PAGE) gel running buffer before transferring into 12% SDS-gels. The SDS-gels were run at 15 °C until running of the dye front out of the gels.

Image scan and data analysis. Gel images were scanned immediately following the SDS-PAGE using Typhoon TRIO (Amersham BioSciences) (data not shown). The scanned images were then analyzed by Image Quant software (version 6.0, Amersham BioSciences), followed by Biological Variation Analysis (BVA) using DeCyder (TM) 2D software version 6.5 (Amersham BioSciences). The protein spots that were found significantly being expressed differentially in at least 1.5 fold change were selected for mass spectrometry identification work using MALDI-ToF/ToF.

Results. Based on 2D-DIGE analysis, three proteins that were involved in glycolysis pathway were determined to be differentially expressed between high-yielding and low-yielding palms across time points. These proteins are triose-phosphate isomerase, which is expressed at lower levels in high-yielding versus low-yielding palms, fructose biphosphate aldolase, which is expressed at higher levels in high-yielding versus low-yielding palms, and glyceraldehyde-3-phosphate dehydrogenase, which is expressed at lower levels in high-yielding versus low-yielding palms.

Example 4

10 Dot-Blot Immunoassay

Protein extraction for dot-blot immunoassay validation. TCA extraction buffer (pre-cooled at -20 °C, 0.2 g + 0.5 mL) was added to fine powder and further ground using mini plastic grinder. The mixture was mixed and mashed well before adding another 1 mL buffer and samples were incubated at -20 °C for 1 hour. The samples were then centrifuged at max speed (13.2 g) at 4 °C for 15 min. Sample tubes were kept on ice and the supernatant was removed using a pipette. Approximately 1.8 ml wash buffer was added to the sample and pellets were re-suspended and/or crushed using pipette tips. Samples were incubated at -20 °C for 1 hour and centrifuged at max speed at 4 °C for 15 min. Supernatant was removed and the washing step was repeated for a total of three times. Sample powder was air-dried on ice for 30 min and re-suspended in 500 µl of lysis/USB buffer. The solution was incubated at 37 °C for 1 hour with continuous shaking. Samples were centrifuged at max speed at room temperature for 15 min and the supernatant was transferred to clean tubes before storing at -80 °C (pellet was stored at this temperature as back up for further use). To further elute protein from pellet, an additional 500 µl of lysis/USB buffer was added to samples, which were then shaken at room temperature for 1 hour. The back-up supernatant was transferred to clean microcentrifuge tubes and stored at -80 °C.

Protein quantification (Bradford assay). Protein stocks (1.4 µg/µl) that were diluted five-fold were used to construct the standard curve for quantification. Sample concentrations ranged from 0.244 µg/µl (lowest) to 2.934 µg/µl (highest). Upon obtaining the protein stock concentration, a working stock (330 µl) was prepared at a final

concentration of 0.2 µg/µl using PBS buffer (+10% glycerol) for dot-blotting onto membranes.

Dot-blot array using 386 pin replicator. Samples were prepared at concentrations of 0.20 µg/µl and 0.02 µg/µl (10x dilution), respectively. Using a replicator, the proteins were applied onto the membrane and fan-dried after each application. The procedure was repeated for a total of 5 rounds (equivalent to stamping 0.20 µg or 0.02 µg of protein on each spot since replicator pins deliver 0.2 µl of sample). Membranes were allowed to air dry overnight. Membranes were then removed from plates and cut down to size. Membranes were kept in between the original protective paper and stored in air-tight containers in dry environment until further use.

Preparation for screening. Individual membranes were clipped onto glass slides, two on each slide, back facing inwards. The clipped membranes were dipped into a container filled with cold 0.1% PBS-T (pH 7.4) and stirred on a magnetic stirrer for 40 min. The 0.1% PBS-T was replaced with cold 0.05% PBS-T and washed continuously for 15 min. The solution was replaced with fresh cold 0.05% PBS-T and stirred continuously for 7 min and the step above was repeated.

Antibody incubations. Approximately 1 ml of PBS-T 0.05% was pipetted onto blank membranes and 1 ml diluted antibody was pipetted onto corresponding membranes in incubation containers. The membranes were incubated overnight on a Belly Dancer brand laboratory shaker (IBI Scientific) at 4 °C or room temperature for 2 to 3 hours depending on optimized conditions for individual antibodies. The used sera were retained for further experiments or discarded into a bottle to be autoclaved. The membranes were clipped onto glass slides and washed in cold 0.05% PBS-T for 15 min, followed by 7 min twice using fresh 0.05% PBS-T. The membranes were laid back into clean incubation containers in a manner to ensure that no bubbles were trapped underneath the membranes. Diluted secondary antibody (1 mL) was added to each membrane. For secondary antibodies with background signals, pre-adsorption was performed with 1% BSA, followed by shaking at 37 °C for 2 hours. The containers were covered and incubated for 2.5 hours on a Belly Dancer laboratory shaker at 4 °C or room temperature. The secondary antibody was discarded and washing steps were repeated as in steps above for 15 min, 7 min, and 7 min, each round with fresh, cold 0.05% PBS-T.

Development and documentation. The membranes were laid back in incubation containers, again in a manner to ensure that no bubbles were trapped underneath the membranes, and any remaining 0.05% PBS-T was removed. Fresh NBT/BCIP was prepared according to manufacturer guidelines using alkaline phosphatase (AP) buffer (100 mM Tris [pH 9.0], 150 mM NaCl, 1 mM MgCl₂). NBT/BCIP solution (1.5 mL) was added per membrane, followed by incubation on a Belly Dancer laboratory shaker until purple color is developed (approximately 30-45 min). The reaction was then stopped and membranes were soaked in water. The developed membranes were scanned and processed using Adobe Photoshop C84 Extended and Olympus Micro software to automatically capture and transform spot densities into excel sheets.

Results. The dot blot immunoassay indicated that triose-phosphate isomerase expression was lower in high-yielding palms than in low-yielding palms at week 16 after pollination (FIG. 1) and that fructose biphosphate aldolase expression was higher in high-yielding palms than in low-yielding palms at week 18 after pollination (FIG. 2).

Example 5

Transcriptional Expression Analyses

Experimental setup and sampling. Eight biological samples were selected respectively from high and low yielding populations based on the average yearly performance of oil content in the mesocarp tissue of the oil palm. Inflorescences of the selected palms were open pollinated in the field. Each pollinated inflorescence was kept until a specific time point according to the pollination date thereof, specifically 12, 14, 16, 18, 20, and 22 weeks after pollination. The fresh fruit bunches of selected palms were harvested at the specific time points and mesocarp tissues were obtained from harvested fruit bunches.

Oil palm mesocarp RNA extraction. Total RNA was extracted from oil palm mesocarp tissue of six different time points as described above by using a modified RNA extraction method by Asemota *et al.* (2). The modification includes use of a composition of extraction buffer to which no phenol was added. The modified protocol also included an extra step of chloroform cleanup before LiCl precipitation. The concentrations and purity of total RNA were determined by quantification with a NanoDrop brand

spectrophotometer (Thermo Scientific). The AU 260/280 and AU 260/230 were measured and samples with ratios of 1.8-2.0 were accepted in quality. Gel electrophoresis was also performed on 1 µg of total RNA by using 1% agarose gel in TAE buffer to further determine the RNA quality. Gels were imaged by use of AlphaImager 2200
5 brand imaging system (Alpha Innotech).

Custom Design Oil Palm Mesocarp Agilent Array. The transcriptional expression of high and low yielding oil palm populations throughout the six different time points were determined by hybridizing the samples on custom oil palm mesocarp gene expression arrays. Customized gene expression oil palm mesocarp arrays were
10 designed based on an Agilent microarray platform in 2 x 105K format. The probes were designed based on 31794 sequences from mesocarp transcriptome sequencing whereby the annotations were obtained by comparing to the UniProt database (unpublished data). Probes were designed using Agilent internal designed program through Agilent's eArray website. The transcriptome sequences were represented by three different probes which
15 covered different parts of the transcriptome sequences.

Synthesis of cRNA, Microarray Hybridization, and Scanning. The total RNA samples from the mesocarp were treated and labeled with one-color Cy3 dye according to Agilent's Low Input Quick Amp Labeling protocol version 6.0 (3) as follows. A total of 100 ng of total RNA was used to synthesize cRNA, which was labeled with Cy3 dye.
20 Labeled cRNAs were used to hybridize on the mesocarp array at 65° C for 16 hours. After hybridization, the mesocarp array was subjected to two steps of washing with wash buffer 1 and 2 for 1 minute for both washing steps. The array was air-dried for a few seconds before the image was scanned by using an Agilent B scanner.

Data Extraction, Analysis, Selection of Candidates, and Classification. Raw
25 microarray data were extracted from scanned images by using Feature Extraction software version 10.7.31 from Agilent. After the extraction, raw microarray data from each time point were further analyzed using R software. In the analysis of R software, the raw microarray data were normalized using a Quantile normalization algorithm. After normalization, each probe signal value from high yielding population samples was
30 compared to probe signal value of samples from low yielding population, to obtain the relative gene expression ratios of each probe between high and low yielding populations.

Probes with expression ratios of greater than 1.5 fold were selected. The differentially expressed candidates (probes) were further filtered by using the corresponding p-value obtained. The p-value cut-off for each time point was based on the distribution of general p-value for all probes in that particular time point where non-differentially expressed candidates were equally distributed, in accordance with Fodor *et al.* (4). In this experiment, the cut-off p-values of weeks 12, 14, 16, 18, and 20 were 0.05 and the cut-off p-value for week 22 was 0.01. The differentially expressed candidates from each time point were classified into different categories according to their gene functions using WEGO Gene Ontology (GO) system.

Results and Discussion. A total of 2116 candidate differentially expressed probes were obtained from six different time points by applying the cut-off above. The candidate differentially expressed probes were distributed as follows: week 12 (100 probes); week 14 (270 probes); week 16 (134 probes); week 18 (171 probes); week 20 (588 probes); and week 22 (853 probes). All 2116 candidates probes represent specific sequences, termed isotigs, which can be used to design probes on arrays.

Annotation of the candidate probes was carried out based on the Swiss-Prot database. Annotations were distributed as follows: week 12 (56 annotations from the 100 probes); week 14 (180 annotations from the 270 probes); week 16 (93 annotations from the 134 probes); week 18 (108 annotations from the 171 probes); week 20 (370 annotations from the 588 probes); and week 22 (524 annotations from the 853 probes). In total, 1331 annotated candidate probes were obtained from the six time points. The Blast2GO program was used to obtain Gene Ontology (GO) annotations for all annotated candidate probes. The annotated candidate probes were further classified according to GO using a WEGO web tool of Ye *et al.* (5) and cluster of orthologous group as mentioned in Wang *et al.* (6).

In accordance with the annotations, the transcriptional expression data indicated that ATP citrate lyase expression was higher in high-yielding palms than in low-yielding palms at week 20 after pollination (FIG. 3), that pyruvate kinase expression was higher in high-yielding palms than in low-yielding palms at week 14 after pollination (FIG. 4), and that aconitase expression was higher in the high-yielding palms than in low-yielding palms at 14 weeks after pollination (FIG. 5).

Example 6**Mass Spectrometry Profiling**

Capillary Electrophoresis–Mass Spectrometry method. Mesocarp samples (55 to 70 mg) were added to 500 µl of methanol including 50 µM of internal standard (Methionine sulfone and D-camphor-10-sulfonic acid) in the tubes and were frozen with liquid nitrogen. They were homogenized using a cell breakage machine with beads (TOMY, MS-100R). Chloroform (500 µl) and Milli-Q water (200 µl) were added to the homogenates, followed by thorough mixing and then centrifugation ($2,300 \times g$, 4 °C, 5 min). After centrifugation, the resulting water layer (400 µl) was harvested and filtered with a 5-kDa cut-off filter (MILLIPORE, Ultrafree MC UFC3LCC). The filtrates were desiccated and then dissolved with 50 µl of Milli-Q water. Judging from peak shapes and intensities, they were diluted by 50% and 20% for CE-TOFMS analysis in cation and anion modes, respectively.

Samples were analyzed using an Agilent CE-TOFMS system (Agilent Technologies) equipped with a fused silica capillary (i.d. 50 µm × 80 cm). For cation mode, the run buffer was Cation Buffer Solution (p/n: H3301-1001), and the rinse buffer was Cation Buffer Solution (p/n: H3301-1001). Sample injection pressure was 50 mbar (10 sec). The MS parameters were as follows: CE voltage : Positive, 27 kV; MS ionization : ESI Positive; MS capillary voltage : 4,000 V; MS scan range : m/z 50-1,000; Sheath liquid : HMT Sheath Liquid (p/n : H3301-1020). For anion mode, the run buffer was Anion Buffer Solution (p/n: H3302-1021), and the rinse buffer was Anion Buffer Solution (p/n: H3302-1022). Sample injection pressure was 50 mbar (25 sec). The MS parameters were as follows: CE voltage : Positive, 30 kV; MS ionization : ESI Negative; MS capillary voltage : 3,500 V; MS scan range : m/z 50-1,000; Sheath liquid : HMT Sheath Liquid (p/n: H3301-1020).

Data Acquisition and Analysis. Peaks detected in CE-TOFMS analysis were extracted using the automatic integration software (MasterHands ver.2.1.0.1). Peak information including m/z , migration time (MT) and area was obtained. Peak area was converted into relative peak area according to the following equation: Relative Peak Area = (Metabolite Peak Area)/(Internal Standard Peak Area x Sample Amount). Principle

component analysis (PCA) and Orthogonal Partial Least Square-Discrimination Analysis (OPLS-DA) were used to identify metabolites that reflect the differences between high and low yielding palm trees. PCA and OPLS-DA were performed with SPSS ver. 18 and SIMCA-P ver. 12. Quantitative estimation was performed using 107 metabolites, including intermediates in glycolysis, TCA cycle, amino acids, and nucleic acids. Concentrations of the metabolites were calculated by normalization of the peak area of an internal standard. Standard curves for each of the metabolites were obtained by single-point at 100 μ M of standard metabolites.

Results. Metabolites such as 2-phosphoglyceric acid, 3-phosphoglyceric acid, and fructose-1,6-bisphosphate were found to be in low concentration in high-yielding palms. These metabolites are highly consumed or have been utilized for oil biosynthesis in the mesocarps of high-yielding palms. It was found that the concentration of glycerol-3-phosphate was higher in high-yielding palms from week 16 to matured stage. This indicates that a reduced triose-phosphate isomerase activity and an increased glycerol-3-phosphate dehydrogenase activity enhanced production of dihydroxyacetone phosphate, which was converted into glycerol-3-phosphate, providing building blocks for oil biosynthesis.

Example 7

Functional Characterization of Glycolytic Genes

Model systems. Functional characterization of oil palm genes of glycolysis and lipid biosynthesis can be carried out using transformation-based approaches in model plant and microbial systems such as *Arabidopsis* sp. and yeast. Changes in levels of metabolites, lipid content, protein expression, or other phenotypic changes can be compared between control (i.e. wild-type) plants or microbes, mutants that over-express a particular gene, and mutants that are complemented for a particular gene. SNP analysis and function prediction of genes related to the oil palm glycolytic pathway can be carried out and the SNPs effects on these genes can be studied in yeast.

Experimental design for oil palm gene over-expression study in yeast. The following approach was used to characterize the effect of over-expression of four oil palm genes, specifically triose-phosphate isomerase, fructose-1,6-bisphosphate aldolase,

glycerol-3-phosphate dehydrogenase, and glyceraldehyde-3-phosphate dehydrogenase, in yeast. Each of the four genes was isolated from DNA of oil palm mesocarp. The sequence of each gene was verified. Each gene of interest was sub-cloned into pYES2.1 TOPO vector. Sequence verification was carried out to check the orientation of the cloned gene in the vector. A yeast strain was transformed with the vector. Yeast transformants were verified using a PCR approach. Metabolite profiling and analysis was then carried out as follows.

Yeast strain and growth conditions. *Saccharomyces cerevisiae* cells were grown in synthetic complete (SC) minimal agar medium supplemented with appropriate auxotrophic supplements at 30 °C for 2 days. The yeast strain was grown overnight in SC minimal broth medium (without Uracil in the case of cultures of over-expressed mutants) at 30 °C, 200 rpm. The overnight culture was centrifuged at 7500 rpm at ambient temperature for 2 min. The supernatant was discarded and replenished with fresh broth medium. Initial OD₆₀₀ was diluted to 0.5 and inoculated to fresh liquid medium. The yeast cultures (3 replicates) were harvested at 6 hours after incubation and OD₆₀₀ was adjusted to 1.0 for extraction.

Samples for the intracellular metabolites (quenching). Samples of 5 mL (OD₆₀₀ fixed to 1.0) were harvested at 6 hours and centrifuged for 2 min at 7500 rpm, 4 °C. Each supernatant was discarded and the corresponding cell pellet was re-suspended with 5 mL fresh medium. A quenching method was adapted from Castrillo *et al.* (7). In accordance with the method, 10 mL of 60% (v/v) buffered MeOH:tricine (10 mM, pH 7.4) at -40 °C was added to each sample, the mixture quickly vortexed (about 1 second) and incubated at -50 °C for about 5 min. Supernatants were separated by centrifugation (7500 rpm, 5 min, -9 °C, adaptors pre-cooled at -40 °C). The quenching was performed in quadruplicate samples.

Boiling ethanol extraction. An extraction method using buffered ethanol:ammonium acetate (10 mM, pH 7.5) was adapted from Ewald *et al.* (8). Extraction was performed in 50 mL Falcon tubes (2 replicates). Tubes containing 5 mL of 75% (v/v) buffered ethanol:ammonium acetate (10 mM, pH 7.5) and 0.4 mg/mL ribitol were pre-heated in a water bath at 80 °C. Samples were added to the boiling buffered ethanol, immediately mixed, and placed in water at 80 °C for 3 min. Samples were

centrifuged for 2 min at 7500 rpm to remove biomass debris. Samples were then filtered through 0.2 μ m filter, freeze-dried and stored at -80 °C prior to LC-MS analysis.

Ultra Performance Liquid Chromatography. Solvents used for liquid chromatography were of Optima LC-MS grade (Fisher Scientific, Fair Lawn, NJ, USA).

5 Millipore purified water was used in the preparation of standards and sample solutions. The standards of sugar phosphates/glycolytic-related compounds were of analytical grade and obtained from Sigma (St. Louis, MO, USA). An Acquity UPLC system coupled to a detector Xevo TQs was used for analytical determination of the targeted metabolites in the mesocarp sample. Chromatography was conducted with an Amide (2.1 x 100 mm, 1.7
10 μ m) column. Column and sample were maintained at 35 °C and 4 °C, respectively. The injection volume was 3 μ L with partial loop injection. The mobile phase A consisted of 10 mM ammonium formate, pH 9.0, while mobile phase B was 90% (v/v) acetonitrile in 10 mM ammonium formate, pH 9.0. A linear gradient was used as follows: 20% to 50% B from initial time to 5 min; 50% B from 5 min to 8 min; 50% to 20% B from 8 min to
15 10 min; 20% B from 10 min to 15 min. The total run time was 15 min, the flow was constant at 0.3 mL/min, and the curves for gradients were each 6.

Mass Spectrometry. The mass spectrometry was operated in both positive and negative mode with multiple reaction monitoring using ESI. The capillary voltage was set at 3.5 kV, desolvation gas set at 800 L/hr at temperature of 300 °C. The collision gas flow
20 was set at 0.15 mL/min. The MRM settings in the MS/MS function with corresponding cone voltage and collision energy were optimized for each standard compound. Auto dwell times were set for positive mode and negative mode, respectively. Total acquisition durations for both UPLC and MS were set at 15 min and 5 min, respectively. Data were acquired and processed using MassLynx V4.1 and TargetLynx, respectively. The
25 glycolytic metabolites determined were fructose-1,6-bisphosphate; dihydroxyacetone phosphate; glyceraldehyde-3-phosphate; glycerol-3-phosphate; 3-phosphoglyceric acid; phosphoenolpyruvate; and ribitol (as internal standard).

Total Lipid Extraction. Freeze-dried samples of 50 mg were extracted with chloroform:methanol:water (2 mL:2 mL:2 mL), vortexed for 30 seconds, and centrifuged
30 at 4000 rpm, 4 °C for 10 min. The chloroform layer was collected in a new pre-weighed tube. The samples were re-extracted twice with 2 mL of chloroform, vortexed, and

centrifuged. The chloroform extracts (second and third extractions) were combined with the first chloroform extract and dried. The weight of dried extracts (lipid content) were measured.

Results: Differential levels of metabolites in overexpressed transformants.

5 Overexpression of oil palm fructose biphosphate aldolase in yeast resulted in intracellular fructose biphosphate levels that were much lower than for the wild-type yeast. Fructose biphosphate aldolase stimulated the metabolism of fructose biphosphate to dihydroxyacetone phosphate (9 fold) and glyceraldehyde 3-phosphate (11 fold), which greatly decreased the fructose biphosphate pool. High fructose biphosphate aldolase
10 activity also results in accumulation of other metabolites such as glycerol 3-phosphate, 3-phosphoglyceric acid, and phosphoenolpyruvate. Thus, the oil palm fructose biphosphate aldolase showed functionality in directing flux towards dihydroxyacetone phosphate and glyceraldehyde 3-phosphate production.

Overexpression of oil-palm cytosolic triose-phosphate isomerase, which catalyzes
15 interconversion of dihydroxyacetone phosphate into glyceraldehyde 3-phosphate reversibly in yeast resulted in an increase of glyceraldehyde 3-phosphate and dihydroxyacetone phosphate. Accumulation of 3-phosphoglyceric acid was observed as glyceraldehyde 3-phosphate was rapidly converted to 3-phosphoglyceric acid by glyceraldehyde-3-phosphate dehydrogenase in the next step of glycolysis. Thus, the oil
20 palm cytosolic triose-phosphate isomerase showed functionality in directing flux towards glyceraldehyde 3-phosphate and 3-phosphoglyceric acid production.

Overexpression of glyceraldehyde-3-phosphate dehydrogenase resulted in a decreased level of glyceraldehyde 3-phosphate, while an accumulation of 3-phosphoglyceric acid (> 15 fold) was observed in the transformants in comparison to
25 wild-type. Glyceraldehyde-3-phosphate dehydrogenase is essential for the maintenance of cellular ATP levels and carbohydrate metabolism (Rius *et al.* (9)). Thus, the oil palm glyceraldehyde-3-phosphate dehydrogenase showed functionality in directing flux towards 3-phosphoglyceric acid production.

Overexpression of glycerol-3-phosphate dehydrogenase in yeast resulted in
30 differential levels of glycolytic metabolites in comparison to wild-type yeast. Glycerol-3-phosphate dehydrogenase catalyzes the reduction of dihydroxyacetone phosphate to

glycerol 3-phosphate. An increase in conversion of glycerol 3-phosphate from dihydroxyacetone phosphate was observed in the overexpression of glycerol-3-phosphate dehydrogenase transformants. Glycerol 3-phosphate and dihydroxyacetone phosphate can serve as precursors to synthesize phospholipids and glycerolipids (Wang *et al.* (10)).

- 5 Thus, the oil palm glycerol-3-phosphate dehydrogenase showed functionality in directing flux towards glycerol 3-phosphate production. Overexpression of glycerol-3-phosphate dehydrogenase in yeast diverts the carbon flux towards glycerol, but results in poor growth rate due to cytotoxic effect of acetaldehyde accumulation (Remize *et al.* (11)). Results on the biomass growth showed that there was a decrease in dry biomass weight of
10 about 30% in yeast over-expressing glycerol-3-phosphate dehydrogenase compared to wild-type yeast.

- Total lipid contents of the overexpression glycolytic genes in yeast.** Total lipid contents of yeast transformants and wild-type yeast are shown (FIG. 6). Overexpression of fructose biphosphate aldolase and glycerol-3-phosphate dehydrogenase leads to an
15 increase of lipid contents of 10% and 33%, respectively. Moreover, overexpression of triose-phosphate isomerase and glyceraldehyde-3-phosphate dehydrogenase resulted in a decrease of lipid contents of 59% and 21%, respectively. In addition, it is predicted that reduced expression and/or activity of triose-phosphate isomerase and glyceraldehyde-3-phosphate dehydrogenase would result in an increase in lipid contents.

20

Example 8

Additional metabolite studies in high-yielding and low-yielding oil palms.

- Additional comparisons of differences in metabolite concentrations in high-yielding (HY) palms versus low-yielding (LY) palms were carried out. Specifically, eight palms were
25 selected from each of the high- and low-yielding screening populations essentially as described in Example 1. The high-yielding population was previously determined to yield 10 to 12 tonnes of palm oil per hectare per year, and a low-yielding population was previously determined to yield 4 to 7 tonnes of palm oil per hectare per year. Fruits were harvested from plants at 12, 14, 16, 18, 20, and 22 weeks after pollination to represent
30 fruit development before, during, and after lipid biosynthesis and ripening occurs. Profiling of the two populations was carried out using CE-MS with the same extraction

and acquisition parameters described in Example 6. Secondary profiling was carried out using LC-MS (LTQ-Orbitrap and LC-triple quadrupole MS) as well as GC-MS in order to investigate the instrument-specificity of the results, as follows.

LC-MS (LTQ-Orbitrap) method: Extraction for LC-MS (LTQ-Orbitrap)

5 profiling was carried out as follows. Mesocarp tissue from each fruit was lyophilized. Approximately 100 mg of each lyophilized mesocarp tissue was weighed into a 15 mL FALCON tube. Then 2 mL of 75% (v/v) isopropanol with 0.01% BHT was added to the tissue. The mixture was heated and shaken in a thermo-mixer for 15 minutes at 750 rpm. A mixture of (i) 0.2 mg/mL phenanthrene dissolved in 2 mL chloroform, (ii) 0.2 mg/mL
10 ribitol dissolved in methanol, and (iii) 1.6 mL water was used for internal standards for lipids and polar metabolites. The mixture was shaken for 30 seconds using a vortex mixer. Incubation was carried out for 30 minutes at 250 rpm and 4° C in a thermo-mixer. Then 4 mL of chloroform and water at a ratio 1:1 was added to the mixture. The mixture was shaken 30 seconds using a vortex mixer. The mixture was then centrifuged at 4° C
15 and 4000 rpm for 10 minutes to separate the lipid layer (in chloroform, at the bottom) and the polar layer (in methanol and water, at the top).

LC-MS (LTQ-Orbitrap) profiling was conducted on the polar (top) layer as follows. LC-MS data were acquired using Accela-LTQ Orbitrap brand instrument (Thermo Fisher, Germany). Sample analysis was carried out in positive and negative ion
20 modes of detection. The mass scanning range was 100 to 2000 m/z, while capillary temperature was 300° C and sheath gas auxiliary gas flow rates were 35 and 15 arbitrary units ("arb"), respectively. The sweep gas flow rate was set at 1 arb I-spray voltage at 4.5 kb. The resolution was 30,000 at 1 microscan and maximum injection time at 500 ms. The capillary voltage and tube lens were set at 40 V and 80 V, respectively for positive
25 ion modes. Both respective parameters were set at -2.00 V and -47.44 V for negative ion mode. The MS/MS spectra of metabolites were obtained by collision energy ramp at 35 V. Autosampler temperature was set at 10° C with 3.0 µL injection volume. The LC/MS system (controlled by Xcalibur brand software version 2.0, Thermo Fisher Corporation) was run in binary gradient mode. Solvent A was 0.1 % v/v formic acid/water and solvent
30 B was acetonitrile containing 0.1% v/v formic acid. The flow rate was 0.2 mL/min. An Acquity UPLC HSS T3 chromatography column (1.8 µm, 2.1 x 100 mm; Waters,

Malaysia) set at 45° C was used for analyses. The gradient was as follows: 1% B (0 - 1.8 min), 10% B (3 min) to 40% B at 20 min and hold for 3 min, 90% B at 26-28 min and 1% B at 29-35 min. The raw data were processed and compared using Sieve version 1.2 (Thermo Fisher, Alpha Analytical, Malaysia) with the frame time and m/z width set at 1.5 min and 0.002, respectively.

LC-MS (Xevo Triple Quad) method: Mesocarp sample extraction was performed in the same manner as used for LC-MS (LTQ-Orbitrap) analysis. An Acquity UPLC system coupled to a detector Xevo TQs was used for analytical determination of the targeted metabolites in the mesocarp samples. Chromatography was conducted with an HSS T3 (2.1 x 100 mm, 1.7 µm) column. The column and samples were maintained at 45° C and 4° C, respectively. The LC mobile phases used were 0.1% formic acid in water (solvent A) and 0.1% formic acid in acetonitrile (solvent B). The flow rate was 0.3 mL/min. The elution gradient was as follows: initial hold at 95% solvent A; 0 to 3 min linear gradient to 60% solvent A; 3 to 5 min 5% solvent A; 5 to 5.1 min linear gradient to 95% solvent A and hold on to 7 min. Injection volume was 3 µL.

The mass spectrometry was operated in both positive and negative mode with multiple reaction monitoring using ESI. The capillary voltage was set at 2.9 kV, and desolvation gas was set at 800 L/hr at temperature of 350° C. The collision gas flow was set at 0.15 mL/min. The MRM settings in the MS/MS function with corresponding cone voltage and collision energy were optimized for each standard compound. Auto dwell times were set for positive mode and negative mode, respectively. Total acquisition duration for both UPLC and MS was set at 15 min. Data were acquired and processed using MassLynx V4.1 and TargetLynx, respectively.

Standard compounds were weighed and dissolved with 5% (v/v) acetonitrile in MilliQ water to make a stock solution with final concentration of 1 mg/mL (1000 ppm). The stock solutions were diluted to 1 ppm as working stock solutions. A mix standard at concentration of 1 ppm was prepared and injected into LC-MS daily to check system sensitivity and reproducibility.

GC-MS method: Mesocarp sample extraction was performed in the same manner as used for LC-MS (LTQ-Orbitrap) analysis. Derivatization of the polar metabolites was then conducted to enable detection using GC-MS. Samples were taken

out from storage and placed in a SpeedVac concentrator for 30 minutes prior to derivatization. This was to ensure that all water content is dried off as moisture will hinder the derivatization process. Then 120 μ L of 20 mg/mL methoxyamine hydrochloride in pyridine was added and samples incubated at 60° C for 4 hours.

- 5 Following this, 120 μ L of N-methyl-N-(trimethylsilyl) trifluoroacetamide (MSTFA) with 1% trimethylchlorosilane was added. This was followed with further incubation at 60° C for 1 hour. Derivatized samples were then analyzed using an Agilent 6890N Gas Chromatograph coupled with an Agilent 5973i Mass Detector and 6890 series autosampler. Volatiles were separated on a DB-5ms column, 30 m x 0.25 mm i.d. x 0.25
- 10 μ m film thickness. The temperature programme was set at 80° C, held for 3 minutes, followed by a 5° C per minute ramp to 315° C, and holding at this temperature for 5 minutes. The injection port was splitless at 280° C while column flow was maintained constant at 1 mL/min of Helium gas. The MS source was set at 230° C with a scanning range of m/z 50 to m/z 650. Sample of 1 μ L was injected into a programmable injector.
- 15 The total run time is 55 minutes.

Results: Data from primary and secondary profiling were analyzed as described in Example 6. Results indicating relative metabolite concentrations versus time (12 to 22 weeks after pollination) for mesocarp tissue of fruits of high-yielding oil palm plants (triangles) versus low-yielding oil palm plants (circles) are shown for 16:0 fatty acids

20 (FIG. 7), 18:0 fatty acids (FIG. 8), 18:1 fatty acids (FIG. 9), fructose 1,6-bisphosphate (FIG. 10), glycerol 3-phosphate (FIG. 11), 3-phosphoglyceric acid (FIG. 12), malic acid (FIG. 13), isocitric acid (FIG. 14), and 2-oxoglutaric acid (FIG. 15). Ratios of malic acid to citric acid versus time (12 to 22 weeks after pollination), for mesocarp tissue of fruits of high-yielding oil palm plants (triangles) versus low-yielding oil palm plants (circles)

25 are also shown (FIG. 16).

As can be seen, citric acid (FIG. 16), isocitric acid (FIG. 14) and, more markedly, 2-oxoglutaric acid (FIG. 15) all appeared at lower concentrations in the HY group of palms just preceding and during the early stages of lipid biosynthesis (14-18 weeks after pollination). In contrast, malic acid (FIGS. 13 and 16) exhibited a very different

30 concentration profile between the HY and LY groups, maintaining a significantly higher concentration in HY palms from 12 weeks after pollination through to the mid-point of

lipid biosynthesis (18 weeks after pollination) when it dropped to a level similar to that of the LY group. All of the organic acids in the TCA cycle were at their lowest concentrations during the final stage of fruit maturation. The differing trends of malic acid and citric acid concentration in HY and LY palms can be seen (FIG. 16), where the ratio of malic acid to citric acid is higher in HY palms from 12 weeks after pollination until when lipid biosynthesis reaches a maximum at 16-18 weeks after pollination, at which time the ratio decreases steadily to the same as LY palms. The higher consumption rate of malic acid during peak lipid production in HY palms indicates higher malic acid dehydrogenase activity. The relatively low concentrations of isocitric acid and 2-oxoglutaric acid observed in HY palms during lipid biosynthesis is likely to result from higher utilization of acetyl-CoA for fatty acid production and possibly of other major biosynthetic precursors leading to amino acid/protein production.

As can also be seen, the results show that there is not a deleterious shift in the profile of fatty acids C16:0 (FIG. 7), C18:0 (FIG. 8), or C18:1 (FIG. 9) of the oil of the HY palms in comparison to the oil of the LY palms.

The results indicate a divergence in carbon flux away from pyruvate and towards glycerol-3-phosphate as well as significant increases in the malate to citrate ratio preceding and during the lipid biosynthesis periods of fruit development. Simultaneous increases in production of glycerol-3-phosphate through glycolysis and diversion of carbon utilization for acetyl-CoA from the TCA cycle could be significant drivers of increased lipid production in oil palm. The results suggest more broadly that concerted changes in the glycolytic pathway enzyme activities, as discussed earlier, combined with changes to expression or activities of certain TCA cycle enzymes, can simultaneously direct flux towards the two key building blocks of lipid, i.e. glycerol and fatty acids, as an approach for increasing oil yields of oil palm plants.

Example 9

Obtaining High-Yielding Oil Palm Plants. A high-yielding oil palm plant can be obtained from a *dura* x *pisifera* cross, e.g. a *dura* x *pisifera* cross that is likely to segregate in terms of palm-oil yield performance or for commercial production of palm-oil producing plants, as described in the following prophetic example.

A *dura* x *pisifera* cross is carried out by standard methods, and *tenera* progeny are obtained. The progeny are field planted and grown, until fruiting, for at least 18 months. Fruit bunch specimens are taken from each progeny plant for testing at one or more time points from 11 to 23 weeks post-pollination. Testing at multiple time points is preferable to testing at a single time point, as the former provides a higher number of comparative points. Fruitlets are separated immediately from the bunch and ten fruits samples are randomly selected therefrom. Mesocarp tissue is obtained from the ten fruit samples and combined into one mesocarp tissue sample per progeny plant. The resulting mesocarp tissue samples are frozen under liquid nitrogen until further processing.

10 The level of triose-phosphate isomerase and the level of glycerol-3-phosphate dehydrogenase in the mesocarp tissue samples from the progeny plants are then determined, in comparison to mesocarp tissue of a fruit of a reference oil palm plant. For example, upon resuming processing, the mesocarp tissue samples are subjected to extraction as described in Example 5 above. Then gene expression is measured with
15 respect to the samples by use of RT-PCR, RNA-seq, hybridization, or microarray techniques. Alternatively or additionally, the mesocarp tissue samples are subjected to protein extraction as described in Example 3 or Example 4 above. Then expression of the corresponding proteins is measured with respect to the samples by use of gel
20 electrophoresis, mass spectrometry, or antibody binding methods. In either case, the level of triose-phosphate isomerase being lower and the level of glycerol-3-phosphate dehydrogenase being higher for the progeny oil palm plant in comparison to the reference oil palm plant is indicative of a high-yielding phenotype for the progeny. Moreover, for increased selectivity, at least one of the level of malate dehydrogenase being higher and the level of ATP citrate lyase being higher for the progeny oil palm plant in comparison
25 to the reference oil palm plant is further indicative of a high-yielding phenotype for the progeny. In addition, for further increased selectivity, at least one of the level of glyceraldehyde-3-phosphate dehydrogenase being lower, the level of fructose-1,6-bisphosphate aldolase being higher, the level of pyruvate kinase being higher, and the level of aconitase being higher for the progeny oil palm plant in comparison to the
30 reference oil palm plant is further indicative of a high-yielding phenotype for the progeny.

Based on these comparisons, the progeny plants are classified in terms of palm oil yields, resulting in a distribution of progeny spanning a spectrum of high-yielding palm plants to low-yielding palm plants. The set of high-yielding progeny are then selected for further breeding stock development. Thus, for example, progeny for which the level of
5 triose-phosphate isomerase is lower and the level of glycerol-3-phosphate dehydrogenase is higher in comparison to the reference oil palm plant are classified as having a high oil palm yield. Moreover, such progeny can then be selected for further breeding stock development.

10 **Example 10**

Identifying High Yielding Variants in a Population, and Subsequent Selection for Breeding Stock Development. High yielding oil palm plants including useful variants of the genes triose-phosphate isomerase and glycerol-3-phosphate dehydrogenase, and optionally also containing useful variants of the genes malate
15 dehydrogenase, ATP citrate lyase, glyceraldehyde-3-phosphate dehydrogenase, fructose-1,6-bisphosphate aldolase, pyruvate kinase, and/or aconitase can be identified and used for selection of breeding stock development, as described in the following prophetic example.

A population of *tenera* oil palms is obtained from a *dura* x *pisifera* cross of
20 commercial oil palms, breeding parental palms, or wild germplasm material. The *tenera* population is analyzed in order to identify sequence variants of the genes coding for triose-phosphate isomerase and glycerol-3-phosphate dehydrogenase. Each gene variant so detected is then cloned and over-expressed separately in a suitable model system such as yeast or another oleaginous microbe using an appropriate promoter. The level of
25 activity of the corresponding enzyme and/or the level of corresponding lipid production are then measured in the resulting yeast or other oleaginous microbe strains.

For triose-phosphate isomerase, and optionally glyceraldehyde-3-phosphate dehydrogenase, i.e. enzymes for which decreased activity leads to higher lipid production, variants are selected based on producing the least decrease in lipid production
30 versus wild-type and/or based on exhibiting the lowest enzymatic activity in a corresponding enzyme activity assay.

For glycerol-3-phosphate dehydrogenase, and optionally malate dehydrogenase, ATP citrate lyase, fructose-1,6-bisphosphate aldolase, pyruvate kinase, and/or aconitase, i.e. enzymes for which increased activity leads to higher lipid production, variants are selected based on producing the greatest increase in lipid production versus wild-type
5 and/or based on exhibiting the highest enzymatic activity in a corresponding enzyme activity assay.

The *tenera* palms that possess the greatest numbers of these beneficial variants are then identified and selected as breeding stock. A selection of palms with beneficial variants of different enzymes can be intercrossed to bring together multiple desirable
10 genes in subsequent progeny plants based on this method.

References

1. He CY, Zhan JG, Duan AG, Yin JY and Zhou DS. Comparison of Methods for Protein Extraction from Pine Needles. *Forestry Studies in China*, 7: 20-23 (2005).
- 15 2. Asemota O and Shah F H. Detection of Mesocarp Oleoyl-Thioesterase Gene of the South American Oil Palm *Elaeis oleifera* by Reverse Transcriptase Polymerase Chain Reaction. *African Journal of Biotechnology*, 3: 595-598 (2004).
3. Agilent. One-Color Microarray-Based Gene Expression Analysis, Low Input Quick Amp Labeling Protocol. Agilent Technologies, Version 6.0 (December 2009).
- 20 4. Fodor A, Tickle T, and Richardson C. Towards the Uniform Distribution of Null P Values on Affymetrix Microarrays. *Genome Biology*, 8: R69 (2007).
5. Ye J, Fang L, Zheng H, Zhang Y, Chen J, Zhang Z, Wang J, Li S, Li R, Bolund L and Wang J. WEGO: A Web Tool for Plotting GO Annotations. *Nucleic Acids Research*, 34 (Web Server issue): W293–W297 (2006).
- 25 6. Wang ZY, Fang BP, Chen JY, Zhang XJ, Luo ZX, Huang LF, Chen XL and Li YJ. De Novo Assembly and Characterization of Root Transcriptome Using Illumina Paired-End Sequencing and Development of cSSR Markers in Sweet Potato (*Ipomoea batatas*). *BMC Genomics*, 11:726 (2010).
- 30 7. Castrillo JI, Hayes A, Mohammed S, Gaskell SJ and Oliver SG. An Optimized Protocol for Metabolome Analysis in Yeast Using Direct Infusion Electrospray Mass Spectrometry. *Phytochemistry*, 62 : 929-937 (2003).

8. Ewald JC, Heux S and Zamboni N. High-Throughput Quantitative Metabolomics: Workflow for Cultivation, Quenching, and Analysis of Yeast in a Multiwell Format. *Anal. Chem.*, 81: 3623-3629 (2009).
9. Rius SP, Casati P, Iglesias AA and Gomez-Casati DF. Characterization of Arabidopsis Lines Deficient in GAPC-1, a Cytosolic NAD-Dependent Glyceraldehyde-3-Phosphate Dehydrogenase. *Plant Physiology*, 148: 1655-1667 (2008).
10. Wang ZX, Zhugea J, Fang HY and Prior PA. Glycerol Production by Microbial Fermentation: A Review. *Biotechnol. Advances*, 19: 201-223 (2001).
11. Remize F, Barnavon L, and Dequin S. Glycerol Export and Glycerol-3-Phosphate Dehydrogenase, but not Glycerol Phosphatase, Are Rate Limiting for Glycerol Production in *Saccharomyces cerevisiae*. *Metabolic Eng.*, 3: 301-312 (2001).

Industrial Applicability

- The methods and kits disclosed herein are useful for increasing oil yield of an oil palm plant and for predicting oil yield of a test oil palm plant, and thus for improving commercial production of palm oil.

Claims

1. A method for selecting oil palm plant with high oil yield comprising:
 - (i) determining the level of triose-phosphate isomerase and the level of glycerol-3-phosphate dehydrogenase in mesocarp tissue of a fruit of a parental oil palm plant 14 to 19 weeks after pollination thereof in comparison to mesocarp tissue of a fruit of a reference oil palm plant 14 to 19 weeks after pollination thereof; and
 - (ii) selecting progeny of the parental oil palm plant based on the level of triose-phosphate isomerase being lower and the level of glycerol-3-phosphate dehydrogenase being higher for the parental oil palm plant in comparison to the reference oil palm plant, to obtain an oil palm plant with high oil yield.
2. The method of claim 1, further comprising:
 - (i) determining at least one of the level of malate dehydrogenase and the level of ATP citrate lyase in the mesocarp tissue of the fruit of the parental oil palm plant 11 to 21 weeks after pollination thereof in comparison to the mesocarp tissue of the fruit of the reference oil palm plant 11 to 21 weeks after pollination thereof; and
 - (ii) selecting progeny of the parental oil palm plant based also on at least one of the level of malate dehydrogenase being higher and the level of ATP citrate lyase being higher for the parental oil palm plant in comparison to the reference oil palm plant, to obtain the high-yielding oil palm plant.
3. The method of claim 2, wherein the selecting is based on both the level of malate dehydrogenase being higher and the level of ATP citrate lyase being higher.
4. The method of claim 1, 2, or 3, further comprising:
 - (i) determining at least one of the level of glyceraldehyde-3-phosphate dehydrogenase, the level of fructose-1,6-bisphosphate aldolase, the level of pyruvate kinase, and the level of aconitase in the mesocarp tissue of the fruit of the parental oil palm plant 11 to 21 weeks after pollination thereof in comparison to the mesocarp tissue of the fruit of the reference oil palm plant 11 to 21 weeks after pollination thereof; and

(ii) selecting progeny of the parental oil palm plant based also on at least one of the level of glyceraldehyde-3-phosphate dehydrogenase being lower, the level of fructose-1,6-bisphosphate aldolase being higher, the level of pyruvate kinase being higher, and the level of aconitase being higher for the parental oil palm plant in
 5 comparison to the reference oil palm plant, to obtain the high-yielding oil palm plant.

5. The method of claim 4, wherein the selecting is based on all four of the level of glyceraldehyde-3-phosphate dehydrogenase being lower, the level of fructose-1,6-bisphosphate aldolase being higher, the level of pyruvate kinase being higher, and the
 10 level of aconitase being higher.

6. The method of any of claims 1 to 5, wherein the triose-phosphate isomerase of the parental oil palm plant comprises a cytosolic isoform comprising SEQ ID NO: 1 or a plastidial isoform comprising SEQ ID NO: 2, and the glycerol-3-phosphate
 15 dehydrogenase of the parental oil palm plant comprises a plastidial isoform comprising SEQ ID NO: 3.

7. The method of any of claims 1 to 6, wherein at least one of the levels is determined by a technique selected from the group consisting of an in vitro assay, two-
 20 dimensional fluorescence difference gel electrophoresis, antibody-based detection, immunoblot detection, dot-blot detection, liquid chromatography-mass spectrometry, an enzyme activity assay, reverse transcriptase polymerase chain reaction, RNA-seq, hybridization, and microarray.

25 8. The method of any of claims 1 to 7, wherein:
 the parental oil palm plant is a *dura* breeding stock plant;
 the progeny comprises an oil palm plant selected from the group consisting of a *dura* breeding stock plant and a *tenera* agricultural production plant; and
 the high-yielding oil palm plant is selected from the group consisting of a *dura*
 30 breeding stock plant and a *tenera* agricultural production plant.

9. The method of any of claims 1 to 7, wherein:
the parental oil palm plant is a *tenera* breeding stock plant;
the progeny comprises an oil palm plant selected from the group consisting of a
tenera breeding stock plant, a *pisifera* breeding stock plant, and a *tenera* agricultural
5 production plant; and
the high-yielding oil palm plant is selected from the group consisting of a *tenera*
breeding stock plant and a *tenera* agricultural production plant.
10. A method for obtaining palm oil from an oil palm plant with high oil yield
10 comprising:
(i) obtaining the oil palm plant with high oil yield by the method of any of claims
1 to 9; and
(ii) isolating palm oil from a fruit of the oil palm plant with high oil yield.
- 15 11. A method for predicting oil yield of a test oil palm plant comprising:
(i) determining the level of triose-phosphate isomerase and the level of glycerol-
3-phosphate dehydrogenase in mesocarp tissue of a fruit of a test oil palm plant 14 to 19
weeks after pollination thereof in comparison to mesocarp tissue of a fruit of a reference
oil palm plant 14 to 19 weeks after pollination thereof; and
20 (ii) predicting oil yield of the test oil palm plant based on the level of triose-
phosphate isomerase being lower and the level of glycerol-3-phosphate dehydrogenase
being higher for the test oil palm plant in comparison to the reference oil palm plant.
12. The method of claim 11, further comprising:
25 (i) determining at least one of the level of malate dehydrogenase and the level of
ATP citrate lyase in the mesocarp tissue of the fruit of the test oil palm plant 11 to 21
weeks after pollination thereof in comparison to the mesocarp tissue of the fruit of the
reference oil palm plant 11 to 21 weeks after pollination thereof; and
30 (ii) predicting oil yield of the test oil palm plant based also on at least one of the
level of malate dehydrogenase being higher and the level of ATP citrate lyase being
higher for the test oil palm plant in comparison to the reference oil palm plant.

13. The method of claim 12, wherein the predicting is based both the level of malate dehydrogenase being higher and the level of ATP citrate lyase being higher.

5 14. The method of claim 11, 12, or 13, further comprising:

(i) determining at least one of the level of glyceraldehyde-3-phosphate dehydrogenase, the level of fructose-1,6-bisphosphate aldolase, the level of pyruvate kinase, and the level of aconitase in the mesocarp tissue of the fruit of the test oil palm plant 11 to 21 weeks after pollination thereof in comparison to the mesocarp tissue of the
10 fruit of the reference oil palm plant 11 to 21 weeks after pollination thereof; and

(ii) predicting oil yield of the test oil palm plant based also on at least one of the level of glyceraldehyde-3-phosphate dehydrogenase being lower, the level of fructose-1,6-bisphosphate aldolase being higher, the level of pyruvate kinase being higher, and the level of aconitase being higher for the test oil palm plant in comparison to the reference
15 oil palm plant.

15. The method of claim 14, wherein the predicting is based on all four of the level of glyceraldehyde-3-phosphate dehydrogenase being lower, the level of fructose-1,6-bisphosphate aldolase being higher, the level of pyruvate kinase being higher, and the
20 level of aconitase being higher.

16. The method of any of claims 11 to 15, wherein the triose-phosphate isomerase of the test oil palm plant comprises a cytosolic isoform comprising SEQ ID NO: 1 or a plastidial isoform comprising SEQ ID NO: 2, and the glycerol-3-phosphate
25 dehydrogenase of the test oil palm plant comprises a plastidial isoform comprising SEQ ID NO: 3.

17. The method of any of claims 11 to 16, wherein at least one of the levels is determined by a technique selected from the group consisting of an in vitro assay, two-
30 dimensional fluorescence difference gel electrophoresis, antibody-based detection, immunoblot detection, dot-blot detection, liquid chromatography-mass spectrometry, an

enzyme activity assay, reverse transcriptase polymerase chain reaction, RNA-seq, hybridization, and microarray.