Supplement to:

Quantitative profiling and pattern analysis of triacylglycerol species in Arabidopsis seeds by electrospray ionization mass spectrometry

Maoyin Li, Ethan Baughman, Mary R. Roth, Xianlin Han, Ruth Welti, and Xuemin Wang

Method S1. Derivation of the regression curves for deduction of the adjustment factors at 46 seed oil TAG *m/z*s.

Adjustment factor is defined as the NL signal intensity of internal standard divided by the NL signal intensity for each TAG, when the TAG and the internal standard are equimolar. Thus, the adjustment factor for the internal standard C51:3 (tri17:1) is 1. The adjustment factors of 15 TAGs were measured (Figure S3a; Table S1). Comparing the measured adjustment factors of the ammoniated TAG ions of the C54 group (C54:0, C54:3, C54:6), C57 group (C57:0, C57:3, C57:6), and C60 group (C60:0, C60:3, C60:6) as a function of *m/z* (Figure S3a), the observed trend was different from that observed for lithiated TAG ions (Han and Gross, 2001) with more dependence of the ammoniated TAG ion response on acyl unsaturation. Due to the complexity of the adjustment factor variation for ammoniated TAG ions and the lack of a full collection of TAG standards, we applied a stepwise strategy of fitting regression curves as described below and illustrated in Figures S3-S6. Online tools, including an exponential regression tool and a polynomial regression tool, were utilized to fit the data with the regression curves (www.xuru.org).

The adjustment factors for saturated TAGs and TAGs with three fatty acyl double bonds were first deduced from regression curves in which the measured adjustment factors were plotted as a function of TAG *m/z*, as shown in Figure S3. The adjustment factors were measured for C42:0, C45:0, C48:0, C51:0, C54:0, C57:0, C60:0 (Figure S3a). These 7 data points were fitted with an exponential regression curve, from which the adjustment factors of additional saturated TAGs were deduced, including C50:0, C52:0, C56:0, and C58:0 (Figure S3b). Adjustment factors were measured for C48:3, C51:3, C54:3, C57:3, C60:3 (Figure S3a), and these data were also fitted with an exponential regression curve, from which the

adjustment factors of TAGs with 3 double bonds were deduced, including C50:3, C52:3, C56:3, and C58:3 (Figure S3c).

The adjustment factors for C48:6, C50:6, C52:6, C56:6, and C58:6 were deduced with two approaches, TAG:0→TAG:6 and TAG:3→TAG:6, in which the quantitative relationships among measured adjustment factors were used to create a curve from which additional adjustment factors can be approximated (Figure S4). The measured adjustment factor pairs (x, y) of the TAGs (C54:0, C54:6), (C57:0, C57:6), and (C60:0, C60:6), were fitted with a quadratic regression curve (Figure S4a), so that additional values of adjustments factors for TAGs with 6 double bonds could be calculated from adjustment factors for TAGs with 0 double bonds. Using the adjustment factors for C48:0, C50:0, C52:0, C56:0, C58:0, determined as shown in Figure S3b, the adjustment factors for C48:6, C50:6, C52:6, C56:6, and C58:6 were calculated with the equation shown in Figure S4a. Similarly, the measured adjustment factor pairs (x, y) of the TAGs (C54:3, C54:6), (C57:3, C57:6), and (C60:3, C60:6), were fitted with a quadratic regression curve (Figure S4b). From the adjustment factors of C48:3, C50:3, C52:3, C56:3, C58:3, calculated from the curve fitting shown in Figure S3c, the adjustment factors for C48:6, C50:6, C52:6, C56:6, and C58:6 were determined (Figure S4b). The adjustment factors from the separate approaches of fitting TAG:0→TAG:6 and fitting TAG:3→TAG:6 were: C48:6 (3.03 from TAG:0→TAG:6, 3.29 from TAG:3→TAG:6), C50:6 (3.04, 3.21), C52:6 (3.11, 3.18), C56:6 (4.00, 3.99), and C58:6 (5.72, 5.75). A log likelihood test was performed to compare the difference between the data from these two approaches and the test value is 0.32 (p value > 0.05), indicating that there is no significant difference between these two groups of calculated values. The adjustment factors for C48:6, C50:6, C52:6, C56:6, and C58:6 obtained from the two approaches were averaged and utilized for the steps described below (Figure S4c).

The adjustment factors for species with varied double bond numbers in the C48, C50, C52, C54, C56, C58, and C60 groups were deduced from regression curves derived for each TAG group and from the measured or calculated values for adjustment factors for TAGs with 0, 3, and 6 double bonds (Figure S3, S4c). The adjustment factors for C48:0, C48:3, and C48:6 were taken from Figure S3 and S4c, and a

quadratic regression curve was fitted for the deduction of adjustment factors in the C48 group (Figure S5a). Similarly, using the :0, :3, and :6 values from Figures S3 and S4c, quadratic regression curves were derived to obtain adjustment factors for TAGs with C50, C52, C54, C56, and C58 (Figure S5b-f).

For the C60 group, four adjustment factors were measured, including values of C60:0, C60:3, C60:6 and C60:9 (Figure S3a). Therefore, a cubic regression curve was fitted for C60 group (Figure S5g). A quadratic regression curve, $y = 1.369011274 \cdot 10^{-1} x^2 - 270.0585003 x + 133188.8296$, can also be fitted for C60 group from the measured adjustment factors of C60:0, C60:3, and C60:6. The adjustment factors of C60 group deduced from the quadratic regression model and the cubic regression model, respectively, were C60:0 (11.88; 12.02), C60:1 (8.78; 8.31), C60:2 (6.79; 6.52), C60:3 (5.91; 6.22), C60:4 (6.14; 6.99), C60:5 (7.49; 8.39), and C60:6 (9.95; 10.00). A log likelihood test was performed to compare the difference of these two sets of adjustment factors and the test p-value is 0.43, which is above 0.05 and indicates that there is no significant difference between these two sets of calculated values. The quadratic model and cubic model are equally sufficient to deduce the adjustment factors for C60 group. The cubic regression curve was utilized for the deduction of the adjustment factors for C60:0, C60:1, C60:2, C60:3, C60:4, C60:5, and C60:6 (Figure S5g).

The seven regression curves for the corresponding seven Arabidopsis seed TAG mass groups are summarized in Figure S6 and the deduced adjustment factors are shown in Table S1. The adjustment factor for C52:3 was determined to be 1.39 from the regression curve for C52 group (Figure S5c). The synthetic TAG 16:0-18:1-18:3 (C52:3) was measured to be 1.37 ± 0.05 (SE, n = 5), which is very close to the deduced adjustment factor of 1.39 for C52:3. This implies that the multiple regression approach employed here is appropriate to derive adjustment factors that can be used for analysis of seed TAGs.

Method S2. Derivation of the formulae for calculation of the fatty acyl combinations at each TAG *m/z*.

The calculation formulae for the fatty acyl combination at each TAG mass are listed in Table S3. The derivation of these formulae is illustrated by calculation of the levels of fatty acyl combinations at C52:4 (*m/z* 872.8), C56:6 (*m/z* 924.8), C56:5 (*m/z* 926.8), C58:2 (*m/z* 960.9), and C60:4 (*m/z* 984.9)

(Figure 3, S7). The input data into each calculation are the molar acyl chain abundances (nmol/mg dry mass) calculated from intensities from the NL scans at the indicated *m/z* by isotopic deconvolution and application of the adjustment factor appropriate for that *m/z* (i.e., for that TAG group as defined by total acyl carbons: total acyl double bonds); data were normalized to the internal standard as the adjustment factor is applied. The formulae were determined by logical deduction from the possible acyl combinations in a TAG group and the molar abundance of the fatty acyl species making up TAGs in the group. In some cases, practical considerations, such as very low abundance of some fatty acyl species at some *m/z*, were taken into account in determining the formulae. In all cases, values were set to zero if the calculation results in a value equal or below zero when there was subtraction in the formula.

Derivation of the calculation formulae for C52:4 (*m/z* 872.8):

Two fatty acyl chain combinations, 16:0-18:2-18:2 and 16:0-18:1-18:3, can be deduced for *m/z* 872.8 (C52:4) based on total fatty acyl carbon number and total fatty acyl double bonds.

1. 16:0-18:2-18:2 is the only species that contains 18:2. The 18:2 molar amount is designated $A_{18:2}$ as determined by NL (Figure 3a; Table S2). Since 16:0-18:2-18:2 contains two 18:2, the molar amount of 16:0-18:2-18:2 is equivalent to $A_{18:2}/2$.

2. 16:0-18:1-18:3 contains 18:1 and 18:3 and their total amount is ($A_{18:1}$ + $A_{18:3}$) as determined by NL (Table S2). Therefore, the level of 16:0-18:1-18:3 can be deduced to be ($A_{18:1}$ + $A_{18:3}$)/2.

Derivation of the calculation formulae for C56:6 (*m/z* 924.8):

Four fatty acyl chain combinations, 18:3-18:3-20:0, 18:2-18:3-20:1, 18:2-18:2-20:2, and 18:1-18:3-20:2, can be deduced for *m/z* 924.8 (C56:6) based on total fatty acyl carbon number and total fatty acyl double bonds.

1. 18:3-18:3-20:0 is the only species that contains 20:0. The 20:0 molar amount is designated $A_{20:0}$ as determined by NL (Figure 3a; Table S2). Since 18:3-18:3-20:0 contains one 20:0, the molar amount of 18:3-18:3-20:0 is equivalent to $A_{20:0}$.

2. 18:2-18:3-20:1 is the only species that contains 20:1. The 20:1 molar amount is designated $A_{20:1}$ as determined by NL (Figure 3a; Table S2). Since 18:2-18:3-20:1 contains one 20:1, the molar amount of 18:2-18:3-20:1 is equivalent to $A_{20:1}$.

3. 18:1-18:3-20:2 is the only species that contains 18:1. The 18:1 molar amount is designated $A_{18:1}$ as determined by NL (Figure 3a; Table S2). Since 18:1-18:3-20:2 contains one 18:1, the molar amount of 18:1-18:3-20:2 is equivalent to $A_{18:1}$.

4. 18:2-18:3-20:1 and 18:2-18:2-20:2 contain 18:2. The molar amount of 18:2 in 18:2-18:3-20:1 is $A_{20:1}$ (step 2). Therefore the molar amount of 18:2 in 18:2-18:2-20:2 is ($A_{18:2}$ - $A_{20:1}$). Since 18:2-18:2-20:2 contains two 18:2, the molar amount of 18:2-18:2-20:2 can be deduced to be ($A_{18:2}$ - $A_{20:1}$)/2.

Derivation of the calculation formulae for C56:5 (*m/z* 926.8):

Five fatty acyl chain combinations, 18:2-18:3-20:0, 18:2-18:2-20:1, 18:1-18:3-20:1, 18:1-18:2-20:2, and 18:0-18:3-20:2, can be deduced for *m/z* 926.8 (C56:5) based on total fatty acyl carbon number and total fatty acyl double bonds.

1. 18:2-18:3-20:0 is the only species that contains 20:0. The 20:0 molar amount is designated $A_{20:0}$ as determined by NL (Figure 3a; Table S2). Since 18:2-18:3-20:0 contains one 20:0, the molar amount of 18:3-18:3-20:0 is equivalent to $A_{20:0}$.

2. 18:0-18:3-20:2 is the only species that contains 18:0. The 18:0 molar amount is designated $A_{18:0}$ as determined by NL (Figure 3a; Table S2). Since 18:0-18:3-20:2 contains one 18:0, the molar amount of 18:0-18:3-20:2 is equivalent to $A_{18:0}$.

3. 18:1-18:2-20:2 and 18:0-18:3-20:2 are the species that contain 20:2. The 20:2 molar amount is designated $A_{20:2}$ as determined by NL (Figure 3a; Table S2). The molar amount of 20:2 in 18:0-18:3-20:2 is equivalent to $A_{18:0}$ (step 2). Therefore the molar amount of 20:2 in 18:1-18:2-20:2 is equivalent to ($A_{20:2}$ - $A_{18:0}$). The level of 18:1-18:2-20:2 is deduced to be ($A_{20:2}$ - $A_{18:0}$), which is also designated as B.

4. 18:1-18:3-20:1 and 18:1-18:2-20:2 contain 18:1. The molar amount of 18:1 in 18:1-18:2-20:2 is B (step 3). Therefore the molar amount of 18:1 in 18:1-18:3-20:1 is ($A_{18:1}$ - B). The level of 18:1-18:3-20:1 is deduced to be ($A_{18:1}$ - B), which is also designated as C.

5. 18:2-18:2-20:1 and 18:1-18:3-20:1 contain 20:1. The molar amount of 20:1 in 18:1-18:3-20:1 is C (step 4). Therefore the molar amount of 20:1 in 18:2-18:2-20:1 is ($A_{20:1}$ - C). The level of 18:2-18:2-20:1 is deduced to be ($A_{20:1}$ - C).

Derivation of the calculation formulae for C58:2 (*m/z* 960.9):

The abundance of each fatty acyl chain at TAG *m/z* 960.9 was calculated from molar acyl chain abundances (nmol/mg dry mass) from the spectra acquired in each of 10 NL scan modes and corrected as indicated (Figure S7a; Table S2). Six fatty acyl chain combinations, with appropriate total fatty acyl carbon number and total fatty acyl double bonds, can be deduced for *m/z* 960.9 (C58:2) (Figure S7a). The logical steps used in deduction of the calculation formulae follow.

1. The level of 18:0-18:1-22:1 is equivalent to the fatty acyl abundance of 22:1 ($A_{22:1}$), because 22:1 is uniquely found in this species at this *m/z*.

2. Similarly, because 20:1 is found only in 18:1-20:0-20:1 at this *m/z*, the level of 18:1-20:0-20:1 is equivalent to the level of $A_{20:1}$.

3. 18:1-20:0-20:1 (step 2), 18:0-20:0-20:2, and 18:2-20:0-20:0 are potential species containing 20:0. The fatty acyl abundance of 20:2 at this *m/z* is very low, indicating that the 20:2-containing 18:0-20:0-20:2 is

very low in abundance, and the amount of 20:0 in 18:0-20:0-20:2 is negligible compared with the total NL abundance of 20:0 (Figure S7a). The amount of 20:0 in 18:1-20:0-20:1 (deduced in step 2) is equivalent to $A_{20:1}$. Therefore the amount of 20:0 in 18:2-20:0-20:0 can be approximated by ($A_{20:0}$ - $A_{20:1}$). Since 18:2-20:0-20:0 contains two 20:0, the level of 18:2-20:0-20:0 is ($A_{20:0}$ - $A_{20:1}$)/2.

4. 18:0-18:1-22:1 (step 1), 18:1-20:0-20:1 (step 2), and 18:1-18:1-22:0 are the three possible species that contain 18:1. The amount of 18:1 in 18:0-18:1-22:1 is equivalent to $A_{22:1}$ (step 1); the amount of 18:1 in 18:1-20:0-20:1 is equivalent to $A_{20:1}$ (step 2); therefore the amount of 18:1 in 18:1-18:1-22:0 is ($A_{18:1}$ - $A_{22:1}$ - $A_{20:1}$). Since 18:1-18:1-22:0 contains two 18:1, the level of 18:1-18:1-22:0 is ($A_{18:1}$ - $A_{22:1}$ - $A_{20:1}$)/2. The abundance of 18:1-18:1-22:0 is here designated as B.

5. 18:1-18:1-22:0 (step 4) and 18:0-18:2-22:0 contain 22:0. The amount of 22:0 in 18:1-18:1-22:0 is B. Therefore the amount of 22:0 in 18:0-18:2-22:0 is equivalent to ($A_{22:0}$ - B). The level of 18:0-18:2-22:0 is ($A_{22:0}$ - B), which is here designated as C.

6. The level of 18:0-20:0-20:2 can be calculated based on the amount of 18:0. The 18:0-containing combinations are 18:0-18:1-22:1 (step 1), 18:0-18:2-22:0 (step 5), and 18:0-20:0-20:2. Therefore, the amount of 18:0 in 18:0-20:0-20:2 is equivalent to ($A_{18:0}$ - $A_{22:1}$ - C). The level of 18:0-20:0-20:2 is ($A_{18:0}$ - $A_{22:1}$ - C).

Derivation of the calculation formulae for C60:4 (*m/z* 984.9):

The abundance of each fatty acyl chain at TAG *m/z* 984.9 was calculated from the molar acyl chain abundances (nmol/mg dry mass) calculated from correction of the intensities in the spectra acquired in each of 10 NL scan modes (Figure S7b; Table S2). Seven fatty acyl chain combinations can be deduced for *m/z* 984.9 (C60:4) based on total fatty acyl carbon number and total fatty acyl double bonds (Figure S7b). The logic is described below for determining the levels of TAG with each of the fatty acyl combinations.

1. 18:2-20:2-22:0 and 18:3-20:1-22:0 each contain one fatty acyl 22:0. The combined total of 18:2-20:2-22:0 and 18:3-20:1-22:0 is equivalent to the amount of 22:0, or $A_{22:0}$, and these species are calculated in combination.

2. 18:1-20:2-22:1, 18:2-20:1-22:1, and 18:3-20:0-22:1 each contain one fatty acyl 22:1. The total level of 18:1-20:2-22:1, 18:2-20:1-22:1, and 18:3-20:0-22:1 is equivalent to the amount of 22:1, or $A_{22:1}$, and these species are calculated in combination.

3. 20:0-20:2-20:2 and 20:1-20:1-20:2 contain 20:0, 20:1, and 20:2. In the fatty acyl combinations of 18:2-20:2-22:0 (step 1), 18:3-20:1-22:0 (step 1), 18:1-20:2-22:1 (step 2), 18:2-20:1-22:1 (step 2), and 18:3-20:0-22:1 (step 2), each contains one fatty acyl 20:0, 20:1, or 20:2 plus one 22:0 or 22:1. Thus, the amount of 20:0 + 20:1 + 20:2 in these five species is equivalent to $A_{22:0} + A_{22:1}$. The total amount of 20:0, 20:1, and 20:2 in these plus the other possible species, 20:0-20:2-20:2 and 20:1-20:1-20:2, is equivalent to ($A_{20:0} + A_{20:1} + A_{20:2}$), and the total amount of 20:0, 20:1, and 20:2 in 20:0-20:2-20:2 and 20:1-20:1-20:2 is equivalent to $[A_{20:0} + A_{20:1} + A_{20:2} - (A_{22:0}+A_{22:1})]$. Since each combination has 3 fatty acyl chains, the total level of 20:0-20:2-20:2 and 20:1-20:1-20:2 is equivalent to $[A_{20:0} + A_{20:1} + A_{20:2} - (A_{22:0} + A_{22:1})]/3$.