Accelerating gene function discovery by rapid phenotyping of fatty acid composition and oil content of single transgenic T<sub>1</sub> Arabidopsis and camelina seeds

Shijie Ma<sup>1</sup> | Chang Du<sup>1</sup> | John Ohlrogge<sup>2</sup> | Meng Zhang<sup>1</sup> 

<sup>1</sup>College of Agronomy, Northwest A&F University, Yangling, Shaanxi, China
<sup>2</sup>Department of Plant Biology, Michigan State University, East Lansing, MI, USA

Abstract
Arabidopsis is wildly used as a model plant and camelina is increasingly used for oilseed research and applications. Although the Arabidopsis genome has been sequenced for two decades, the functions of many lipid-related genes and their regulators have not been well characterized. Improvements in the efficiency and accuracy of gene investigations are key to effective discovery of gene function and downstream bioengineering of plant oil quantity and quality. In this study, a visible marker was used to quickly identify transgenic T<sub>1</sub> seeds and a method has been developed to phenotype fatty acid compositions and oil content of single T<sub>1</sub> seeds. A whole seed direct transmethylation method was first optimized with multiple seeds and incubation at 85°C for 2 hours in a transmethylation solvent (5% H<sub>2</sub>SO<sub>4</sub> in methanol with 30% toluene cosolvent) is recommended. Based on this method, a single Arabidopsis seed mini-transmethylation (SAST) method has been established in a 1.5 ml GC sample vial with 200 μl transmethylation solvent. Characteristics of the method were evaluated and it was used to phenotype transgenic T<sub>1</sub> seeds expressing AtFAD2 or RcWRI1. Our results indicate that fatty acid composition of T<sub>1</sub> individual seeds are consistent with those of pools of multiple seeds from higher generations. However, oil content per individual seed varied substantially and therefore pooling five seeds is recommended for phenotyping oil content of T<sub>1</sub> seeds. Additionally, a whole camelina single-seed direct transmethylation was evaluated and results confirm its feasibility. The suitability of partial seed analysis of camelina was investigated but variation in composition of different seed tissues limits this approach.

Keywords
Arabidopsis, camelina, fatty acid composition, oil content, single seed, transgenic T<sub>1</sub> seed
Arabidopsis is widely used as a model plant and its small genome was the first of plants to be sequenced 20 years ago. More than 800 candidate genes were predicted to be involved in lipid metabolism (Beisson et al., 2003, Li-Beisson et al., 2013) (ARALIP: http://arabidopsisacyl lipids.plantbiology.msu.edu/), but among these less than 300 have been experimentally characterized (Mcgrew et al., 2015). Moreover, there are over one thousand transcription factors in Arabidopsis, but only a few of them has been revealed to be involved in regulation of lipid metabolism. Therefore, there is still a long road ahead to reveal functions of lipid genes and their regulators. Thus, investigation of gene function is now a major task and a main factor limiting our understanding of plants in the post genomic era. As more and more gene sequences and their expression patterns are available, many gene functions have been revealed through reverse genetics (Meissner et al., 1999, Østergaard and Yanofsky, 2004, Matus et al., 2014). Overexpression and knock-out/knockdown of gene expression are major strategies of reverse genetics for identification of gene functions. Generally speaking, overexpression vectors or gene silencing vectors need to be constructed and transferred into host species, and then gene functions can often be revealed by phenotyping the changes of transgenic lines. Due to the segregation of traits in heterozygous transgenic plants, analysis of homozygous transgenic lines is optimal, but is time-consuming and requires at least three generations. Furthermore, due to position effects or other unknown reasons, multiple independent transgenic lines are usually needed for phenotyping, which further increases the burden of investigators. In downstream applications of genetic engineering, it is also important to phenotype a substantial number of independent engineered lines to identify optimal traits. Therefore, increasing the efficiency of identification and phenotyping of transgenic lines can provide a major boost to gene function studies and bioengineering research.

Arabidopsis can be easily transformed by the flower dipping method (Clough and Bent, 1998), which further enhances its application for reverse genetic investigation and proof-of-concept tests for bioengineering. Because of its 30%–35% oil content in seeds, Arabidopsis is often used as a model plant for oil crop research (Katavic et al., 1995, Li et al., 2006). Fatty acid composition and oil content of seeds can be determined by different methods. Oil content of seeds can be determined by the gravimetric method where total lipid is extracted (e.g., by the Folch method [Folch et al., 1957] or the Soxhlet method) and weighed. Fatty acid composition can be determined by transmethylation of lipid extracts followed by GC analysis (Focks and Benning, 1998, Marillia et al., 2003). Direct transmethylation of Arabidopsis leaves increases the efficiency of phenotyping fatty acid composition (Browse et al., 1986) and combination with the internal standard triheptadecanoin was used for quantitative measurement of fatty acid contents (Epp and Pollard, 1993). Single transgenic T₅ seed was homogenized and transmethylated for detecting unusual fatty acid (Cahoon and Shanklin, 2000). Later, a whole intact Arabidopsis seed direct transmethylation method to determine fatty acid composition as well as oil contents was evaluated (Li et al., 2006). Although seed oil content and fatty acid composition may also be determined by non-destructive methods, such as nuclear magnetic resonance (NMR) (Alexander et al., 1967, O’Neill et al., 2003) and near infrared reflectance spectroscopy (NIRS) (Jiang et al., 2007, Rudolphi et al., 2012, Kaur et al., 2017), they are not often used in Arabidopsis gene function studies because they need larger quantities of seeds.

Camelina (Camelina sativa) is also a member of Brassicaceae, and was cultured as an oil crop 2000 years ago. Due to high content of ω-3 fatty acid, camellinas has regained attention of vegetable oil consumers. To improve the seed oil quality, the three FAE1s were knocked out in camellina, which leads to greatly reduced levels of very long chain fatty acid in its seed oil (Ozseyhan et al., 2018). Additionally, because camellina is adapted to various soil and climatic conditions and has strong resistance to diseases and pests, it is gradually used as an eco-friendly oil crop for biodiesel (Zubr, 1997) and other products. Due to the successful application of the flower dipping method (Lu and Kang, 2008), camellina is also increasingly used as a platform oil crop for gene function investigation and bioengineering (Bansal and Durrett, 2016).

Although life cycles of Arabidopsis and camellina are as short as 3 months, in transgenic studies, it will take nearly 1 year to identify phenotypes of fatty acid composition and oil content if homozygous lines are used. However, if a visible marker is used to screen transgenic seeds, we asked if it is possible to accurately phenotype seed lipids at an early generation of transgenic lines. To improve the efficiencies of reverse genetic studies and bioengineering on seed lipids, in this study, a quick method was developed after first optimizing the whole seed transmethylation conditions. Reliability and applicability of this method for phenotyping fatty acid composition and oil content were evaluated by comparison with the multiple seed method. By combining the SAST method with application of the DsRed visible marker, T₅ Arabidopsis seeds were identified and phenotyped. Furthermore, the validity of whole camellina seed transmethylation was compared with other methods and a single camelina seed transmethylation method was also evaluated. Additionally, the feasibility of a camellina partial-seed analysis method was examined.
2 | RESULTS

2.1 | Optimization of the whole seed transmethylation-GC method for oil content and fatty acid composition of Arabidopsis seeds

The whole seed transmethylation-GC method (WST-GC) is very efficient and widely used to determine oil content as well as fatty acid composition of pooled Arabidopsis seeds (Li et al., 2006). Based on the WST-GC protocol, a method for single Arabidopsis seed was explored and developed. As a control to evaluate reliability, multiple seeds (50 seeds) of Arabidopsis wild type from the same batch were randomly pooled and measured by the WST-GC method. However, variation in results, especially oil contents, were frequently found for the 50-seed controls. Some reasons have been identified later, such as the uncalibrated temperatures of incubators. Furthermore, oil contents of wild type, determined by different laboratories or different members of our laboratory varied to a certain extent (discussed below). Therefore, the WST-GC method was optimized first by using the same batch of seeds to establish a standard protocol.

The effectiveness of toluene as a cosolvent of methanol and triacylglycerol was evaluated first. Results showed that oil content measured in 30% toluene was about two times higher than without toluene, while the determined fatty acid compositions did not differ significantly (Figure 1a,b). This result emphasizes the indispensability of toluene (or other cosolvent) in the transmethylation solution, even if the internal standard 17:0 is added. Second, oil contents and fatty acid compositions were determined under two different concentrations of sulfuric acid, the catalyst for the transmethylation. The fatty acid compositions measured under 2.5% and 5% sulfuric acid were nearly identical, while oil content under 5% sulfuric acid was slightly higher than that under 2.5% sulfuric acid (Figure 1c,d). Therefore, 5% sulfuric acid was used in the following experiments.

As mentioned above, incubation temperature and duration time of transmethylation vary in previous studies. Therefore, oil contents and fatty acid compositions were determined and compared under different combinations of incubation temperatures (80, 85, 90 and 95°C) and duration time (1.5, 2, 2.5 and 3 hr). Results (Figure 2a) indicated that the measured oil contents varied substantially (from 5.5 to 7.4 μg/seed) under different treatments of temperature and incubation time. In general, the measured oil contents were relatively low at 80 and 95°C, and the contents were also relatively low under 1.5 hr duration time. It is worth noting that the combination of lower temperature (80°C) and shorter duration time (1.5 hr) resulted in a ~25% lower oil content measurement. The combination of long duration time (3 hr) and high temperature (95°C) also resulted in lower oil content results. Fatty acid compositions also varied. High temperature and long duration time often led to higher relative contents of saturated fatty acids, especially 16:0, and lower polyunsaturated fatty acids, especially 18:2 (Figure 2b and Table S1), which may be due to preferential loss of PUFA. Taken together, we conclude an incubation temperature of 85°C and 2 hours of duration as the optimum conditions.

2.2 | Method of SAST

When the above WST-GC method was applied directly to single seeds, some problems were observed in pilot experiments. The first issue was weak signals of GC-FID detector when the extracted FAMEs from single seeds were applied, making it very difficult to identify some minor GC peaks, such as the peak of 18:1Δ11. More importantly, unstable fatty acid profiles and lower oil contents were found compared to that from the same batch of seeds determined by the WST-GC method on multiple seeds. In order to concentrate the FAMEs and increase the signal, one-tenth volume of solvent was first tested in a regular 15 ml screw-cap tube. In this protocol, a single seed was transmethylated in 100 μl solvent (5% sulfuric acid methanol with 30% toluene) in a 15 ml tube. Peak signals of FAMEs did increase significantly, however, the fatty acid profiles were not stable. Relative contents of saturated fatty acids were frequently higher than those determined by the multiple seed WST-GC method (Figure 3a). It was speculated that this increase may be due to increased oxidation or easier evaporation of polyunsaturated FAMEs in a lower volume solvent system with more oxygen in the headspace. Some evaporation of solvent in a relatively large space or any minor gas leakage could be significant to the small volume of solvent system. To deal with this problem, the 100 μl volume of transmethylation solvent was applied in a 1.5 ml screw-cap GC-vial and incubated in a water bath for transmethylation. FAME extracts were transferred into an insert of GC-vial. Then the insert was placed into the same vial without emptying the residue solvent, which can be used as an effective coolant for the FAME sample during queuing for GC sample loading. Results showed that fatty acid profiles determined in the small vial system were much more stable and repeatable. For determining optimal solvent volume, 100 μl and 200 μl solvent systems were tested and compared with the multiple seed method under the same condition. Results indicated that fatty acid profiles determined in the small GC-vial under both solvent systems were nearly identical to that of the multiple seed method. Smaller error bars (±SD) for the 200-μl system indicate that the 200-μl volume is more reproducible than the 100-μl system (Figure 3b,c). Additionally, the 200-μl system is easier than 100 μl for the subsequent FAME extraction and the peak signals can be improved by increasing the loading volume to GC. Therefore, the 200-μl solvent system was used in the following single seed measurements. This method of single Arabidopsis seed mini-transmethylation is referred to as SAST (Figure S1).

After setting up the SAST method, fatty acid composition and oil contents of single seeds (from the same batch of seeds as above) were determined to address whether the SAST is suitable for phenotyping both fatty acid composition and oil contents of a genotype. Results showed that relative fatty acid composition of individual seeds were very similar to each other and also to values determined by the multiple seed method. However, oil contents of these seeds varied over a large range from 4.60 μg/seed to 8.15 μg/seed (Figure 4a), which may be due to the variation in individual seed size and weight. Taken together, these results
suggest that the SAST method is suitable to phenotype fatty acid composition, but may not provide a representative oil content per seed of a genotype. However, as discussed later, by comparing results from a number of single seeds, in some cases, a useful correlation between fatty acid composition and oil content may be discovered.

Considering the relatively easy transformation of Arabidopsis and the oil content variation of single seed analysis, a few seeds (5 and 10 seeds, respectively) were pooled instead of single seed and the oil content was determined by the SAST method. Very similar oil contents with small error bars were determined for the 5-seed pool, 10-seed pool, and 50-seed pool (Figure 4b,c). This result suggests that pooling five seeds (or 10 seeds if abundant T₁ seeds are available) (a modified SAST method), instead of a single seed, may be a suitable choice for phenotyping oil content (per seed) of limited Arabidopsis seeds.
2.3 | Application of the SAST or modified SAST to phenotype fatty acid and oil contents of transgenic T₁ seeds

To evaluate SAST as an early identification method for phenotyping fatty acid and oil contents of transgenic lines, the DsRed visible marker was used to select T₁ seeds. In order to test the validity of the SAST or modified SAST for this early identification, we characterized the phenotypes of transgenic lines in their T₁ seeds. In our previous studies, overexpression of sense FAD2 triggers a strong co-suppression and the substrate of FAD2 (18:1) is accumulated at a high frequency in seeds of flax, carinata, camelina, and Arabidopsis (Du et al., 2018). To characterize this strong co-suppression, we previously determined fatty acid profiles of pooled transgenic Arabidopsis seeds in dozens of independent T₂ and higher generation lines (Du et al., 2018). This same phenotype was used in the present study to test the changes of fatty acid composition on individual T₁ seeds using the SAST method. The construct used by Du et al. was transferred to Arabidopsis and T₁ seeds were screened by DsRed fluorescence and analyzed together with untransformed negative control seeds from the same plants. Fatty acid composition of single T₁ and untransformed seeds were determined by the SAST method. Net changes of 18:1 and 18:2 + 18:3 of T₁ seeds compared to untransformed seeds were also used to compare with the previous T₂ results. Results showed that 18:1 increased in 23 out 30 T₁ seeds compared with untransformed seeds and most lines showed an absolute 20%–40% net increase of 18:1 compared with untransformed seeds (e.g., 18:1 increased from 15% to 35% and to 55%). This nearly 80% frequency of co-suppression and the intensity of co-suppression closely matched the previous results for T₂ seeds (Figure 5b in this study and Figure 2b of Du et al., 2018). These results indicate that determining fatty acid composition of T₁ seeds by the SAST method can phenotype transgenic lines early and reliably.

To evaluate the validity of phenotyping oil contents of T₁ seeds with SAST, a homolog of WRINKLED1 was tested. In a recent study, RcWRI1 was cloned from castor bean and seed-specific overexpression of RcWRI1 in Arabidopsis increased oil contents of T₃ seeds (Yang et al., 2019). In this study, the same construct was transferred into Arabidopsis and T₁ seeds were identified by their DsRed fluorescence. As suggested above, five T₁ seeds and untransformed seeds were pooled, respectively. Their oil content and fatty acid composition were determined by the modified SAST method. Results showed that oil content of T₁ seeds increased from 5 μg/seed to nearly 8 μg/seed, which is consistent with the improvement in the transgenic T₃ seeds in the previous study (Figure 5a in this study and Figure 5a of Yang et al., 2019). These results indicate that quantifying oil content of T₁ seeds by the five-seed modified SAST method is valid to predict the variation of oil contents in transgene experiments.

2.4 | Evaluation of whole camelina seed transmethylation method

Arabidopsis seed is small enough for successful whole seed direct transmethylation (Li et al., 2006), whereas in large seeds the reagents fail to access all lipid pools. Camelina seed weights are about 60- to 80-fold higher than Arabidopsis seed. Whole seed transmethylation has been reported in camelina lipid studies (Kang et al., 2011). However, to our knowledge, there has been no evaluation of this method to verify its validity on these larger seeds. To better characterize this method, camelina seeds from the same batch were used for determining fatty acid composition and oil content by different methods. The gravimetric method in which
total lipid was extracted and weighed was also used and compared, and the extracted lipid was also transmethylated and quantified by GC. At the same time, ground seed and whole seeds were transmethylated directly. Oil contents are presented as the total fatty content per seed by using GC to quantity FAMEs. The results indicated that oil contents determined by three GC methods were about 480 μg per seed, which was about 60 μg lower than that by the gravimetric method, suggesting that there are about 10% unsaponified lipids in the extracts. More importantly, oil contents and fatty acid contents determined by the three transmethylation methods were nearly identical (Figure 6), indicating that whole camelina seed can be directly and fully transmethylated and that grinding of seeds or preparation of lipid extracts is not needed before transmethylation.
2.5 | Fatty acid compositions and oil contents of individual camelina seeds

Similar to Arabidopsis, camelina can be transformed by the flower dipping method (Lu and Kang, 2008). Transgenic T1 seeds can be easily identified by using a visible marker, such as the DsRed fluorescence marker used in this study (Figure 7a). In principle, it should also be possible to investigate gene function or performance of bioengineered seeds early by phenotyping transgenic T1 camelina seed. Before using T1 seeds for phenotyping, it is important to investigate the seed-to-seed variation in oil contents and fatty acid composition of individual camelina seeds. Therefore,
oil contents and fatty acid compositions of individual seeds were determined and compared with the multiple seed method. Due to the larger size of camelina seeds, it is easy to measure the oil content and fatty acid composition of single camelina seed using a regular scale of transmethylation solution. In order to compare relative oil contents of individual seeds, each camelina seed was

**FIGURE 5** Oil content and fatty acid composition in transgenic lines of Arabidopsis. (a) Oil contents of T1 seeds overexpressing RcWRI1 under wild-type background. Five independent T1 seeds were pooled and transmethylated by using the modified SAST method. Error bars represent SD of three replicates; (b) Net change in proportions of 18-carbon fatty acids in T1 seeds overexpressing AtFAD2 under wild-type background. To compare with the results of T2 seeds in Du et al. (2018), the same format was used to present the desaturation changes of fatty acids. Net changes of 18:1 and 18:2 + 18:3 are indicated by purple and pink columns, respectively

**FIGURE 6** Comparison of different methods for determining oil content and fatty acid compositions. (a) Oil content and (b) Fatty acid composition of camelina seeds determined by different methods. EE refers to esterification of extracted lipid, EG refers to esterification of ground seeds, ED refers to direct esterification of whole seeds and GM refers to the gravimetric method. All samples were transmethylated in a 15 ml tube. Error bars represent SD of three replicates
weighed before transmethylation. Results showed that similarly to Arabidopsis, fatty acid compositions of individual seeds closely resemble each other and are similar to that determined by multiple seed analysis. However, absolute oil contents varied over a large range from about 250 \( \mu \)g/seed to nearly 600 \( \mu \)g/seed. Nevertheless, the relative oil contents (weight %) of individual camelina seeds varied over a much smaller range (Figure 7b). These results indicate that individual camelina seeds are suitable for phenotyping fatty acid compositions and relative oil contents, but absolute oil contents per seed will be strongly influenced by variations in seed size. By referring to the above Arabidopsis methods, a five-seed pool may be valid for preliminary phenotyping of absolute oil contents.

### 2.6 Exploration of partial camelina seed direct transmethylation methods

The ability to analyze fatty acid composition from part of a seed, while preserving the remaining seed for germination, provides a substantial advantage for rapidly identifying and propagating transgenic lines with desired traits. We therefore explored whether a partial Arabidopsis seed mini-transmethylation method could be used but we had very low success in germinating the half seeds. Due to the much larger size of camelina seeds, we tested partial camelina seed direct transmethylation methods for screening the phenotype of T\( _1 \) seeds. First, different cutting methods were tested for the germination of partial seeds. Results showed that cutting of hypocotyl and...
radicle resulted in cutting off seedlings or damage to root, respectively. However, excising part of the cotyledons did not affect the germination and produced normal seedlings (Figure 8a). Therefore, part of the cotyledon was used for determining fatty acid compositions by the SAST method. However, fatty acid compositions varied substantially from different cuttings in pilot experiments. Due to the fact that fatty acid composition of different seeds from the same batch vary only slightly, it was speculated that this variation might be from the different depth of cutting, which results in different ratios of seed coat to cotyledon in the excised tissue. Correspondingly, oil contents and fatty acid composition of seed coat, hypocotyl and cotyledon were determined and compared with those of whole seeds. Results showed that the oil content of hypocotyls was similar to that of whole seed, the oil content of cotyledons was higher than that of whole seeds, while oil contents of seed coats was much lower than that of whole seeds. Fatty acid compositions of different tissues were also quite different from each other, and only that of cotyledons was relatively similar to whole seeds. It is worth noting that fatty acid compositions of seed coats showed large differences in 18:2 and 18:3 from those of other seed tissues (Figure 8b,c). Therefore, different cutting depth showed a different fatty acid composition (Figure 8d). These results suggest that due to the problem of sampling identical structures of partial seeds, it is difficult to accurately phenotype fatty acid compositions or oil contents by using partial camelina seeds. Nevertheless, this method should be useful for rapid preliminary analysis to identify traits such as expression of novel fatty acid structures.

3 | DISCUSSION

It often takes nearly 1 year to create multiple independent homozygous transgenic lines and then to phenotype them for gene function investigations or bioengineering exploration. In this study, an improved strategy was developed to increase the efficiency of these time-consuming operations through the phenotyping of fatty acid composition and oil content of transgenic T₁ seeds. A direct transmethylation method was developed decades ago to rapidly

**Figure 8** Germination after partial cutting of camelina seed and oil analyses of seed tissues. (a) Camelina seeds cut by different incisions and the corresponding seedlings after germination. The positions of incisions are indicated by the red dotted lines. Control is an intact seed; (b) Relative oil content and (c) fatty acid composition of different tissues of camelina seeds and whole seeds; (d) Fatty acid composition of partial cotyledons from different incisions and whole seed. The thin-cut and the deep cut are indicated by the blue and red line in (a). Whole seed was used as a control (CT). Error bars represent SD of three replicates.
determine fatty acid composition of Arabidopsis leaves (Browse et al., 1986). Due to the tiny seed size, this simple and efficient direct transmethylation method has also been widely used to determine the fatty acid composition and oil content of pooled intact Arabidopsis seeds (Li et al., 2006). However, it is worth noting that the composition of transmethylation solution and parameters of incubation varies among previous studies. For examples, concentration of catalyst H$_2$SO$_4$ in transmethylation solution range from 2.5% to 5% (Epp and Pollard, 1993, Lee et al., 2017, Yang et al., 2019). Toluene is a cosolvent of triacylglycerol and methanol and 30% toluene is added in transmethylation solution or it may not be included (Zhang et al., 2009, Lee et al., 2017). Incubation temperature varies from 80°C to 95°C (Katavic et al., 1995, Zou et al., 1997, Li et al., 2006), while reported incubation times differ from 1 to 3 hr (Periappuram et al., 2000, Cernac and Benning, 2004, Li et al., 2017). Furthermore, oil contents of Arabidopsis wild-type seeds show a large range of variation in previous reports. For examples, oil content of ecotype Col 0 was as low as 3.2 μg/seed (Focks and Benning, 1998) and up to 8.0 μg/seed (Zhang et al., 2009). However, fatty acid compositions of Arabidopsis wild-type seeds show a relatively small range of variation (Katavic et al., 1995, Zou et al., 1997, Lee et al., 2017). Oil content may be affected by plant culture conditions (Canvin, 1965, Goffman et al., 2005, Karki and Bates, 2018), and differences of oil content were even found between individual plants (Li et al., 2006). However, the more than twofold differences of reported oil content may not arise only from different plant culture conditions, and therefore the effects of different measurement protocols cannot be excluded. To confirm this speculation, oil content and fatty acid composition of the same batch of Arabidopsis wild-type seeds were determined under different transmethylation solvents and different conditions of incubation. Our results indicate that utilization of cosolvent in transmethylation solvent does not obviously affect fatty acid composition, however, it strongly affects the quantification of oil content (Figure 1a). Two often-used concentrations of catalyst H$_2$SO$_4$ do not affect both the quantification of oil content and fatty acid composition, although the oil content is slightly higher under 5% H$_2$SO$_4$ than 2.5% (Figure 1c). Investigation of temperature and time of incubation revealed that low temperature or short incubation time clearly leads to a measurement of lower oil content, and their combination further worsens the quantification of oil contents. Theoretically speaking, higher temperature and longer time incubation result in more complete transmethylation of lipids. This was true when moderate incubation time and temperature were combined. However, lower oil content occurred under conditions with very high temperature and very long incubation time (Figure 2a). It is deduced that this decrease may be from increased oxidation of unsaturated fatty acids and/or greater evaporation of their methyl ester than of the saturated internal standard. This speculation is supported by the higher relative content of saturated fatty acid under very high temperature and very long incubation time (Figure 2b). Therefore, moderate temperature and incubation time are more suitable and a treatment at 85°C and 2 hours is recommended. Additionally, it is important to calibrate the temperature of incubators before the transmethylation. Standardization of the method is strongly recommended, which will provide more reliable comparison of results from different researchers.

Based on the optimized conditions, the SAST method has been established successfully in 1.5 ml GC-vials with 200 μl transmethylation solution, which reduces the possible oxidation of unsaturated fatty acids and/or evaporation of their methyl ester. Fatty acid composition determined by the SAST method is repeatable and reliable when compared with the pooled-seed WST-GC method (Figure 4a,b). It is important to note that the oil content per Arabidopsis seed varies substantially, which may be due to the varied seed size as was demonstrated by camelina single seed phenotyping (Figure 7b). Variation of seed size occurs naturally within siliques, or can be due to other maternal factors such as position of a silique on the plant. Therefore, oil contents can best be determined representatively by pooling a few (5–10) Arabidopsis seeds together. However, in some cases oil contents may also vary in transgenic seeds due to changes in fatty acid composition. For example, production of ricinoleic acid in transgenic Arabidopsis seeds results in substantially lower oil content (Bates et al., 2014). If a sufficient number of single seeds are analyzed, the SAST method can provide an early discovery of this phenomenon if data for ricinoleic (or other unusual lipid) are plotted versus oil content.

To determine whether the larger intact camelina seeds can be fully transmethylated, our results showed that identical oil content and fatty acid composition were determined by the whole seed transmethylation, ground seed transmethylation, and extracted lipid transmethylation (Figure 6), which demonstrate the reliability and feasibility of the camelina WST-GC method (Kang et al., 2011). Although camelina seeds are large enough for half seed phenotyping and half seed germination strategy (Figure 8a), our results indicate that measurements of oil content and fatty acid composition of partial camelina seeds is influenced by the uneven distribution of oil and fatty acids in seed tissues (Figure 8b,c). Therefore this approach is most suitable to provide a preliminary screen of traits such as expression of a novel fatty acid structure.

Advantages of SAST method

T$_1$ seeds can be easily identified by visible markers, such as DsRed (Stuitje et al., 2003), which can then be analyzed (together with untransformed seeds from the same plant) for fatty acid composition and oil content by the SAST method. This strategy provides several key advantages: (a) The SAST analysis of T$_1$ seeds can substantially improve the efficiency of gene function studies. It will take only about 3 months if the construction of vector is conducted when Arabidopsis is growing. Considering there are often ready-to-use Arabidopsis seedlings in many laboratories, this system may be accomplished in as short as 1 month (1 week for vector preparation, 3 weeks from transformation to collect seeds, and 1 day for phenotyping); (b) Results of T$_1$ seed lipid analysis determined by SAST are very reliable and accurate, as indicated by comparing T$_1$ results of AtFAD2 and RcWRI1 with their higher generations (Figure 5).
4 | EXPERIMENTAL PROCEDURES

4.1 | Plant materials and growth conditions

Arabidopsis thaliana ecotype Col-0 was used as the wild type for all Arabidopsis experiments. The wild-type seeds were surface-sterilized with 75% ethanol (v/v) for 1 min and rinsed with distilled water, and then germinated at 22°C on 1/2 MS medium supplemented with 1.5% sucrose after stratification at 4°C for 48 hr. Seedlings were transferred into a soil mixture (peat moss-enriched soil: vermiculite: perlite = 3:1:1) and cultured under conditions as previously described (Zhang et al., 2016). The Suneson cultivar, released by Montana State University, was used as the wild type of camelina. Seeds were sowed in October under an open field condition and harvested in May at the Crop Experimental Station of Northwest A&F University, Yangling, Shaanxi, China. Growth conditions for indoor camelina seedlings were as described by Lu et al (Lu and Kang, 2008). All mature seeds were collected and naturally dried for more than 2 weeks. A single batch of Arabidopsis and camelina seeds were used in all the experiments.

4.2 | Optimization of the whole Arabidopsis seed direct transmethylation method

Fifty WT seeds were pooled and the intact seeds were placed into a 15 ml screw-cap glass tube with 2 ml transmethylation solvent including 20 μg triheptadecanoin (17:0) and 50 μl BHT (0.2% butylated hydroxyl toluene). For evaluating the effect of cosolvent, 600 μl toluene was or was not included in 5% (v/v) sulfuric acid in methanol. For evaluating the effect of catalyst concentration, 2.5% or 5% (v/v) sulfuric acid was used in methanol with toluene. The transmethylation solvent with seeds was fully blended and heated at 80°C for 2 hr. After transmethylation reaction, 2 ml of 0.9% NaCl and 1 ml hexane were added, and the FAMEs in hexane were extracted and transferred into GC vials. FAMEs were quantified by GC-FID with a 30 m DB21 column (Agilent) and GC parameters given as follows: 50°C for 1 min and ramped to 175°C at 35°C/min with a 1-min temperature hold followed by a ramp to 230°C at 4°C/min with a final 5-min temperature hold. Injection volume was 2 μl. Three biological replicates were performed. The same methods of FAME extraction and GC analysis were used in all pooled seed WST-GC experiments.

Based on the results of above experiments, 5% (v/v) sulfuric acid in methanol (freshly prepared with 12 μg triheptadecanoin (17:0) and 50 μl BHT, and extra 600 μl toluene) was used as the optimized transmethylation solvent in following experiments. To determine the effects of incubation temperature and time, double factorial experiment was carried out. Four temperature treatments are 80, 85, 90, and 95°C in four calibrated water bath incubators, and four durations of incubation are 1.5, 2, 2.5, and 3 hr. The incubation durations were controlled with an error less than 2 minutes.
4.3 | Development of a SAST method

The 2 ml 5% (v/v) sulfuric acid in methanol with 12 μg triheptadecanoin (17:0) and 50 μl BHT plus 600 μl toluene is designated as "one volume" transmethylation solvent. Transmethylation at 85°C for 2 hr was used as a standard condition in all the following experiments. To evaluate the incubation systems, single seeds were esterified in four different reaction systems: one volume of standard transmethylation solvent in a 15 ml tube, 1/10 volume in a 15 ml tube, 1/10 volume in a 1.5 ml GC sample vial and 1/20 volume in a 1.5 ml GC vial. FAMEs were extracted with corresponding volume of 0.9% NaCl and hexane. The same GC analysis method was as described above, while injection volume was increased to 4 μl instead of 2 μl. Experiments included three replicates. To evaluate the feasibility of the SAST method, 20 Arabidopsis wild-type seeds were transmethylated under above optimized condition (85°C for 2 hr) in a 1.5 ml GC sample vial with 200 μl 5% (v/v) sulfuric acid in methanol with 1.2 μg triheptadecanoin (17:0) and 5 μl BHT plus 60 μl toluene. To modify the SAST method for oil content measurement, 5 or 10 wild-type seeds were randomly pooled and their FAMEs were prepared with the SAST method. Meanwhile, 50 seeds were transmethylated in a 15 ml tube as a control.

4.4 | Plant transformation, identification of transgenic T1 seeds, and determination of their fatty acid compositions and oil contents

Construction of binary vector AtFAD2 and RcWRI1 were previously described by Du et al. (2018) and Yang et al. (2019), respectively. These two vectors were used for Arabidopsis transformation by the flower dipping method (Clough and Bent, 1998). A green light was used as the exciting light and red optical filter was used to identify the transgenic T1 seeds in which the visible marker DsRed was co-expressed with the target gene. The same methods were also used for camelina transformation and identification of transgenic T1 seeds.

For determining fatty acid compositions, FAMEs of Arabidopsis single transgenic T1 seeds were prepared with the SAST method. The fatty acid compositions were determination by the same GC method. To determine the oil contents, five transgenic T1 seeds were pooled and FAMEs were prepared with the SAST method. The oil content (μg/seed) was calculated with the formula: oil content = (total FAMEs − 17:0 FAME)/17:0 FAME x weight of 17:0 (μg).

4.5 | Lipid extraction and determination of oil content and fatty acid composition of camelina seeds

Three biological samples of camelina seeds were weighed carefully on an electronic analytical balance to ±0.01 mg. Fifty camelina dry seeds (about 100 mg) with 200 μg 17:0 were ground and 3 ml Folch solution (2 ml chloroform:1 ml methanol) were added. After vortexing, 2 ml 0.9% NaCl was added to provide phase separation. The lower phase was kept and the seed residues were re-extracted with 2 ml chloroform. The combined lipid extracts were evaporated by N2 and weighed. The oil content was calculated by the weight of extracted lipid/total seeds weight.

To evaluate the whole seed direct transmethylation, 50 ground seeds and 50 intact seeds with 200 μg 17:0 were transmethylated under the standard condition in a 15-ml screw-cap tube. At the same time, about 40 mg extracted lipid was also transmethylated. FAMEs were determined with the standard GC method as described above.

To determine the variation of fatty acid compositions and oil contents of individual seeds, 20 camelina seeds were randomly collected and each seed was weighed. Intact single seeds with 20 μg 17:0 were transmethylated under standard condition and FAMEs were analyzed by GC. Absolute and relative oil content of individual seed were calculated.

4.6 | Histologic oil content and separation of camelina seeds

Camelina seed coats and embryos were separated after imbibition at 4°C for 24 hr in distilled water, by squeezing the soaked seeds gently with two glass slides. Seed coat, hypocotyl, and cotyledon were isolated and dried in a lyophilizer for 48 hr. These tissues were weighed and directly transmethylated and the FAMEs were determined with the standard method above.

Camelina mature seeds were observed and cut under a dissecting microscope. While holding with tweezers, a surgical blade was used to carefully cut the seeds; the thickness of the cut part is less than 1 mm for thin cutting and is more than 2 mm for deep cutting. The remaining seed was used for germination experiments.

ACKNOWLEDGMENT

We thank Dr. Chaofu Lu (Montana State University) who kindly provided the camelina cultivar Suneson. This work was supported by the National Natural Science Foundation of China (31972964), the Key International Cooperation Project of Shaanxi Province (2020KWZ-012) and the Fundamental Research Funds for the Central Universities (2452019168).

CONFLICT OF INTEREST

The authors have no conflict of interest to declare.

AUTHOR CONTRIBUTIONS

SJM and MZ conceived the original screening and research plans. MZ supervised the experiments. SJM and CD performed the experiments. SJM analyzed the data. SJM, JO and MZ conceived the project and wrote the manuscript with contributions of all the authors.

ORCID

Meng Zhang https://orcid.org/0000-0002-0738-7899
Periappuram, C., Steinhauer, L., Barton, D. L., Taylor, D. C., Chatson, B., & Zou, J. (2000). The plastidic phosphoglucomutase from Arabidopsis. A reversible enzyme reaction with an important role in metabolic control. *Plant Physiology, 122*, 1193. https://doi.org/10.1104/pp.122.4.1193

Rudolfi, S., Becker, H. C., Schierholt, A., & von Witzke-Ehbrecht, S. (2012). Improved estimation of oil, linoleic and oleic acid and seed hull fractions in safflower by NIRS. *Journal of the American Oil Chemists’ Society, 89*(3), 363–369. https://doi.org/10.1007/s11746-011-1920-y

Ruuska, S. A., Schwender, J., & Ohlrogge, J. B. (2004). The capacity of green oilseeds to utilize photosynthesis to drive biosynthetic processes. *Plant Physiology, 136*, 301–309. https://doi.org/10.1104/pp.104.047977

Stuitje, A. R., Verbree, E. C., Van Der Linden, K. H., Mietkiewska, E. M., Nap, J.-P., & Knappe, T. J. A. (2003). Seed-expressed fluorescent proteins as versatile tools for easy (co)transformation and high-throughput functional genomics in Arabidopsis. *Plant Biotechnology Journal, 1*, 301–309. https://doi.org/10.1046/j.1467-7652.2003.00028.x

Voelker, T. A., Hayes, T. R., Cranmer, A. M., Turner, J. C., & Davies, H. M. (1996). Genetic engineering of a quantitative trait: Metabolic and genetic parameters influencing the accumulation of laurate in rapeseed. *The Plant Journal, 9*, 229–241. https://doi.org/10.1046/j.1365-313x.1996.09020229.x

Wang, T., Xing, J., Liu, X., Liu, Z., Yao, Y., Hu, Z., ... Zhang, Y. (2016). Histone acetyltransferase general control non-repressed protein 5 (GCN5) affects the fatty acid composition of arabidopsis thaliana seeds by acetylating fatty acid desaturase3 (FAD3). *Plant Journal for Cell & Molecular Biology, 88*, 794–808.

Wei, H., Li, R. J., Zhan, G. M., Jing, L., & Wang, H. Z. (2011). Maternal control of seed oil content in Brassica napus: The role of silique wall photosynthesis. *The Plant Journal, 69*, 432–444.

Yang, Z., Liu, X., Li, N., Du, C., Wang, K., Zhao, C., ... Zhang, M. (2019). WRINKLED1 homologs highly and functionally express in oil-rich endosperms of oat and castor. *Plant Science, 287*, 110193. https://doi.org/10.1016/j.plantsci.2019.110193

Zhang, M., Cao, X., Jia, Q., & Ohlrogge, J. (2016). FUSCA3 activates triacylglycerol accumulation in Arabidopsis seedlings and tobacco BY2 cells. *The Plant Journal, 88*, 95–107.

Zhang, M., Fan, J., Taylor, D. C., & Ohlrogge, J. B. (2009). DGAT1 and PDAT1 acyltransferases have overlapping functions in arabidopsis triacylglycerol biosynthesis and are essential for normal pollen and seed development. *Plant Cell, 21*, 3885–3901.

Zhang, X., Henriques, R., Lin, S.-S., Niu, Q.-W., & Chua, N.-H. (2006). Agrobacterium-mediated transformation of Arabidopsis thaliana using the floral dip method. *Nature Protocols, 1*, 641–646. https://doi.org/10.1038/nprot.2006.97

Zou, J., Katavic, V., Giblin, E. M., Barton, D. L., MacKenzie, S. L., Keller, W. A., ... Taylor, D. C. (1997). Modification of seed oil content and acyl composition in the brassicaceae by expression of a yeast sn-2 acyltransferase gene. *The Plant Cell, 9*, 909. https://doi.org/10.1105/tpc.9.6.909

Zubr, J. (1997). Oilseed crop: Camelina sativa. *Industrial Crops & Products, 6*, 113–119. https://doi.org/10.1016/S0926-6690(96)00203-8

**SUPPORTING INFORMATION**

Additional supporting information may be found online in the Supporting Information section.

**How to cite this article:** Ma S, Du C, Ohlrogge J, Zhang M. Accelerating gene function discovery by rapid phenotyping of fatty acid composition and oil content of single transgenic T1 Arabidopsis and camelina seeds. *Plant Direct*. 2020;00:1–15. https://doi.org/10.1002/pld3.253