Plant Methods

METHODOLOGY Open Access

Rapid quantification of whole seed fatty acid amount, composition, and shape phenotypes from diverse oilseed species with large differences in seed size

Matthew G. Garneau[1], Prasad Parchuri[1], Nora Zander[1] and Philip D. Bates[1*]

Abstract

Background Seed oils are widely used in the food, biofuel, and industrial feedstock industries, with their utility and value determined by total oil content and fatty acid composition. Current high throughput seed oil analysis methods either lack accuracy in total fatty acid profiling or require extensive labor for lipid extraction prior to derivatization to fatty acid methyl esters (FAME) and quantification by gas chromatography (GC). Alternatively, direct whole seed FAME production methods have been developed for the very small seeds in the model species *Arabidopsis thaliana* but these have generally not been adapted to larger seeds of most oilseed crops.

Results High-throughput direct whole seed FAME production methods were optimized for seeds up to 5 mg each utilizing acid-catalyzed esterification. For the oilseed species *Camelina sativa*, *Thlaspi avernse* (pennycress), *Cuphea viscosissima*, and *Brassica napus* (var. Canola), the total seed fatty acid content and composition from direct seed esterification to FAME matched that of lipid extract derivatization demonstrating the accuracy of the methods. In combination with seed phenotyping using GridFree, this approach enabled the development of a rapid pipeline for simultaneous seed weight, count, size/shape phenotyping, and oil analysis. For the larger and tougher seeds produced by *Limnanthes alba* (Meadowfoam) and *Cannabis sativa* L. (hemp) the whole seed acid-based method proved insufficient, and prior laborious homogenization of seeds was required. Therefore, a rapid one-tube bead homogenization and base catalyzed-esterification method was developed. Base-derived fatty acid esterification cannot derivatize free fatty acids leading to slightly lower total seed fatty acid than acid-catalyzed methods, however the seed oil content and fatty acid composition that is valuable for screening large numbers of samples in research populations was accurately measured.

Conclusions New rapid whole seed fatty acid esterification and phenotyping protocols were developed to accurately assess oilseed lipid content. These methods are particularly valuable in oilseed research, breeding, and engineering applications where efficient analysis of large numbers of samples and accurate oil fatty acid profiling is essential. While having been developed for current and emerging oilseed crops, these methods also provide a foundation from which protocols might be established for new and emerging crop species.

*Correspondence:
Philip D. Bates
phil_bates@wsu.edu

Full list of author information is available at the end of the article

Keywords Seed oil, Oil analysis, Fatty acid, Fatty acid composition, FAME, Camelina, Hemp, Pennycress

Background

Seed oils are a critical agricultural resource, essential for producing a wide range of products, including cooking oils, biofuels, and various industrial compounds such as lubricants, industrial precursors, cosmetics and bioplastics [1, 2]. The suitability of a specific seed's oil for any application is dependent on the fatty acid composition which determines its physicochemical properties and potential uses [3, 4]. In response, plant breeding/engineering of oilseed crops focuses on both lipid accumulation and fatty acid composition to improve both the harvestable yield and oil value [5, 6]. A range of seed oil analysis techniques exist, each with their strengths and weaknesses including non-destructive whole seed imaging/spectroscopy, lipid extraction followed by gravimetric quantification or subsequent derivatization and fatty acid quantification, or in situ lipid derivatization and subsequent fatty acid quantification. Non-destructive spectroscopy techniques (that can also be high throughput) such as Near Infrared spectroscopy [7] or Nuclear magnetic resonance [8, 9] quantify oil content, whereas Magnetic Resonance Imaging [10] is proficient at quantification of oil distribution. However, these techniques are limited by the accuracy of measured fatty acid composition especially for minor components [7–9] which can be especially important for breeding or engineering plant oils with unique/tailored fatty acid content. In direct opposition are the labor-intensive total lipid extraction-based methods, often regarded as the 'gold standard' for quantifying both oil content and composition when coupled with gas chromatography (GC, for review see; [4, 11]). Prior to GC analysis of fatty acid content, seed lipids must be derivatized to a more volatile form, most commonly fatty acid methyl esters (FAME). Many FAME production methods exist and typically involve methanol as the methyl donor combined with a strong acid or strong base as the catalyst [12]. Acid-based methods perform transesterification of ester bonds and direct esterification of carboxylic acids (e.g. produce FAME from both glycerolipids and free fatty acids) whereas base-catalyzed methods only perform transesterification and thus will not derivatize free fatty acids. Accurate measurement of the total fatty acid composition of seed oils requires first total lipid extraction; second, isolation of the oil (triacylglycerol, TAG) fraction from membrane lipids or other non-TAG neutral lipids; third, derivatization to FAME; and finally, quantification by GC coupled with either mass spectrometry (GC-MS) or flame ionization detection (GC-FID) [13]. However, prior lipid extraction and TAG isolation require significant investment of expertise, time, materials, and labor. Alternatively, because seed oil fatty acid content is typically > 90% esterified within the TAG fraction, direct seed tissue derivatization to FAME can be utilized to more rapidly determine seed oil amount and fatty acid composition [14, 15]. While direct transesterification of plant tissue is more efficient than prior lipid extraction and more sensitive than non-destructive techniques, it can be limited by the efficiency of the whole tissue derivatization conditions that simultaneously extract and derivatize the lipids within a complex biological tissue. In particular seed size and strength of seed coat can limit extraction/derivatization.

Seed-specific methodologies have largely focused on hydrochloric acid or sulfuric acid as a catalyst to produce FAME from either small seeds or ground plant tissue to ensure total tissue derivatization. Additionally, co-solvents such as toluene have aided in the concomitant lipid extraction/solubilization and derivatization of seed tissue [14, 16]. Most direct whole seed derivatization to FAME protocols have focused on small seeds such as *Arabidopsis thaliana* where the very small size allows for efficient simultaneous tissue disruption, lipid derivatization, and extraction. Development of whole seed direct FAME production protocols for larger seeds has been limited. The goal of the current study was to develop direct whole seed derivatization to FAME protocols for seeds from current and emerging oilseed crops and experimental systems by optimizing in situ lipid derivatization to FAME reaction conditions for larger/tougher seeds. Methods were developed for high throughput seed phenotyping and direct esterification to FAME in oilseeds with an ~ 900-fold difference in single seed weight ranging from very small (20 µg/seed), to relatively large (18 mg/seed), including: *Arabidopsis thaliana* and a variety of current and upcoming research crops including *Camelina sativa* (Camelina), *Thlaspi arvense* (pennycress), *Brassica napus* (Canola), *Cuphea viscosissima* (Cuphea), *Limnanthes alba* (meadowfoam), and *Cannabis sativa* L. (hemp), that represent a variety of seed shapes and sizes (Fig. 1).

The results herein demonstrate optimized direct transmethylation methods that account for: acid or base catalysis, co-solvent concentration, incubation time, and suitability of different seeds for whole seed FAME production or the requirement of seed homogenization. The high throughput methodology utilizes image-based seed counting and size measurements, and the output includes total seed fatty acid composition, total seed fatty acid amount per seed and by seed weight, seed size, and weight per seed. For all species tested to assess the accuracy of the direct tissue esterification to FAME, the optimized methods were compared to FAME production from a total lipid extract from those seed tissues. These

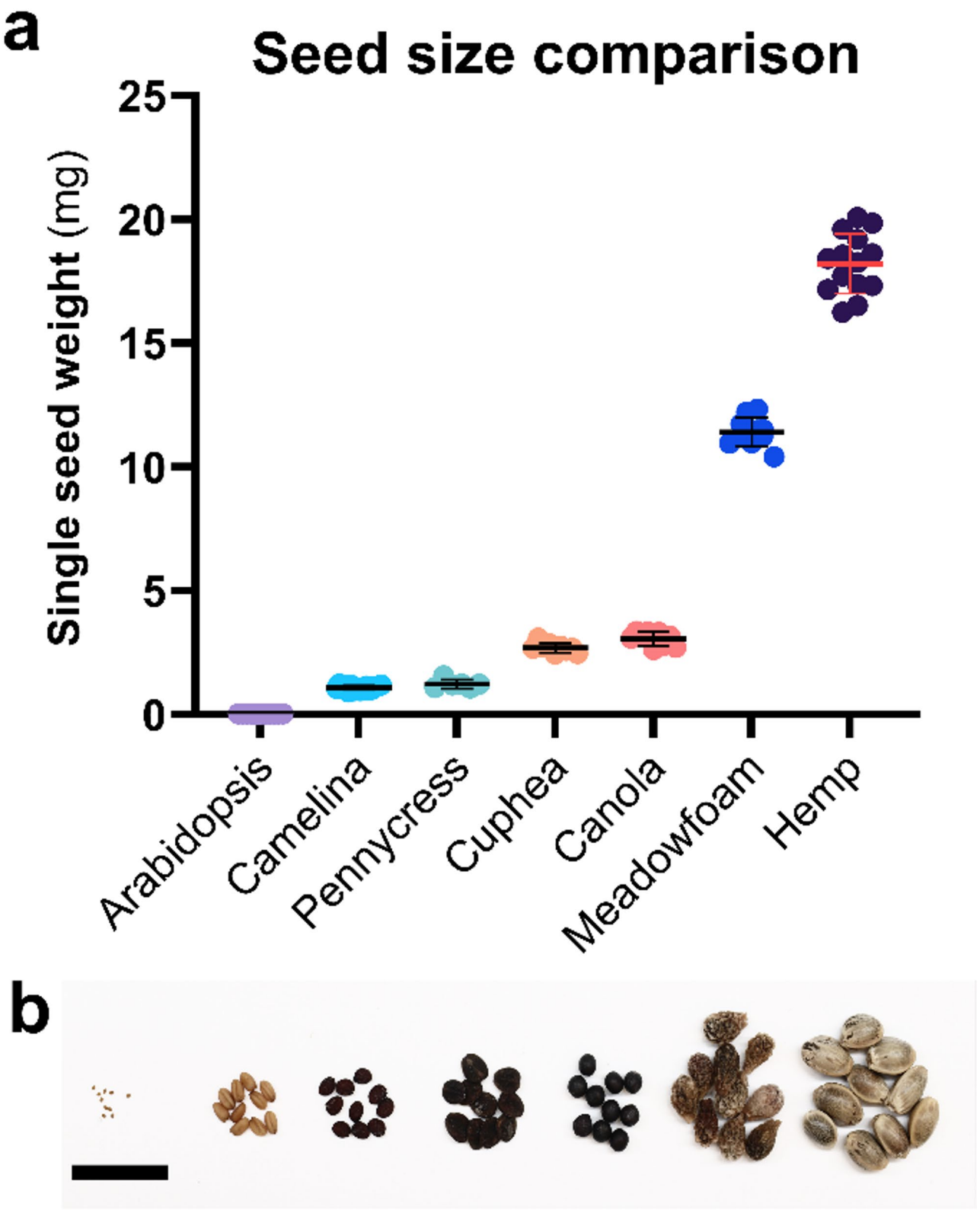

Fig. 1 Comparison of agronomically important oil seeds. The average single seed weight (**a**) and morphology (**b**) of the model species *Arabidopsis thaliana* (Arabidopsis) and oil crops with current or upcoming agronomic importance: *Camelina sativa* (Camelina), *Thlaspi avernse* (Pennycress), *Brassica napus* (Canola), *Cannabis sativa* (Hemp), and plant species with unique lipid compositions *Cuphea viscosissima* (Cuphea) and *Limnanthes alba* (Meadowfoam). Data is presented as mean ± standard deviation, $n = 6+$; scale bar = 0.5 cm

species-specific optimized methods may also provide a basis for the adaptation of tissue direct fatty acid analysis for a wide variety of species with seeds of similar sizes.

Methods

Total seed lipid extraction and acid-based esterification to FAME as control

Total lipids were extracted from *Arabidopsis thaliana* (eco. Columbia-0), *Camelina sativa* (cv. Suneson), *Thlaspi avernse* (pennycress, cv. Spring 32), *Cuphea viscosissima*, and *Brassica napus* (Canola, cv. Griffin) by first quenching 50–100 mg of seeds in 1 mL of 85 °C isopropanol for 15 min in an 8 mL glass tube with PTFE-lined cap. For *Cannabis sativa* L. (hemp, cv. Hlesia) and *Limnanthes alba* (meadowfoam, cv. Foamore), 15 (~250–330 mg) and 4 seeds (~40-50 mg) were taken for lipid extraction, respectively, and quenched as above. Lipids were extracted from seeds using a method modified from Folch et al. [17]. First, quenched samples were ground in a polytron (PT 10–35 GT, Kinematica, NY, USA), followed by the addition of chloroform, methanol, and water to produce a single-phase solvent mixture of chloroform: methanol/isopropanol: water (v/v/v) at a ratio of 2:1:0.1. After mixing well the samples were centrifuged at 2500 x *g* for 5 min. The supernatant was transferred to a new 13 mL glass tube and the solvent extraction above was repeated with an additional 6 mL 2:1 chloroform: methanol and centrifuged again. The supernatants were combined and 0.8% KCl was added to form a phase separation with chloroform: methanol: aqueous/KCl ratio of 2:1:0.8 (v/v/v), samples were shaken well, and the lower organic phase separated from the upper aqueous phase by centrifugation at 2500 x *g* for 5 min. The organic phase was collected with a Pasteur pipet and the total extraction was dried down under N_2 and resuspended in toluene containing + 0.005% butylated hydroxytoluene (BHT).

To quantify the total fatty acid content and composition of the lipid extracts, an aliquot of the lipid extract, 40 µg of a Tripentadecanoin (15:0) internal standard dissolved in 50 µL toluene, and 1 mL of 2.5% (v/v) sulfuric acid in methanol were added to 8 mL glass vials with PTFE lined caps. The samples were derivatized to FAME by heating to 85 °C for 50 min. The resulting FAMEs were collected in hexane by the addition of 1 mL 0.8% (w/v) KCl, mixing well, and 0.5 mL hexane and centrifugation for 5 min at 2500 x *g*.

Seed phenotyping prior to fatty acid quantification using computational analysis

Seed length, width, area, and number were determined using the image-based seed size and counting software GridFree [18]. A photography workflow setup was built to image small seeds at low magnification with high resolution. First, a digital single-lens reflex camera (DSLR; Rebel T7i, Canon, NY, USA) was fitted with a macro-photography Lens (EF-S 35 mm F/2.8 Macro IS STM, Canon) and placed on an overhead tabletop camera mounting stand with built-in photography lighting (Fig. S1). Next, the camera and lens optics were adjusted to take a clear 5 × 3 cm image while still able to move a disposable plastic weigh dish under the lens. The camera and lens were both set to manual, image size set to 1920 × 1280 pixel resolution, and settings/lighting were adjusted once to produce consistent in focus high contrast seed images for image analysis. Large (330 mL) white square plastic weigh dishes were then marked with a 5 × 3 cm rectangle using white tape to provide a size reference and background for seed imaging. During high throughput sampling, a scoop of Arabidopsis seeds (~5 mg) or 5–15 seeds for lager oilseeds (see Table 1) were spread into the weigh boat within the 5 × 3 cm tape rectangle, a photo was then taken, and the photograph number recorded for that sample. Seeds were then weighed and placed in an 8 mL glass tube for oil analysis. Following sampling, seed sample images were processed using GridFree software program following the authors instructions [18]. If seed size analysis was used in the GridFree program the 5 × 3 cm rectangle was used to calculate pixel length and determine seed size. For seed counting of large numbers of samples irrespective of seed size batch counting mode was used.

General direct seed transesterification using sulfuric acid/methanol

Direct seed transesterification was developed based off a protocol described in [14] with optimization for larger sample sizes. Arabidopsis seed aliquots of 5 mg were first transferred to 8 mL glass vials with PTFE lined caps. An internal standard of 50 µg of Tripentadecanoin (15:0) dissolved in 50 µL toluene and 1 mL of freshly made 5% sulfuric acid (H_2SO_4) in methanol were added to each sample. Next, additional toluene was added as a co-solvent to improve reaction efficiency. To determine the optimal addition of toluene ratios of 10:1 to 2:1 (5% sulfuric acid/methanol: toluene) were tested. Samples were then incubated at 85 °C for times ranging from 30 min to 3 h to identify the optimal incubation times. Seed FAMEs were separated by the addition of 1.5 mL 0.8% (w/v) KCl and 1 mL hexane followed by centrifugation for 5 min at 2500 x *g*.

For larger seeds the Arabidopsis protocol above was adapted through increasing volume of 5% sulfuric acid in methanol at a rate of 0.1 mL per each mg of tissue. Aliquots of 15 seeds (≈15 mg) were tested for Camelina, pennycress, hemp while in larger seeds (Cuphea and Brassica) 10 seeds aliquots were used (30–45 mg). Toluene concentration was tested again with Camelina seeds

Table 1 Optimized methods for whole seed direct acid-catalyzed derivatization to FAME

Species	Seed weight/ count	Reagent Volumes (mL) Mix the following reagents and standard (0.8 g/L 15:0-TAG in toluene)			Incubation time at 85 °C (h)	Separation centrifuge at 2500 x g for 5 min following addition of:	
		5% H_2SO_4 in MeOH	Standard in toluene	Toluene		Hexane (mL)	0.8% KCl (mL)
Arabidopsis	5 mg seed (≈ 200)	1	0.1	0.4	1.5	1	1.5
Camelina	15 seeds (≈ 15 mg)	1.5	0.2	0.4	2 (mix/0.5 h)	2	1.5
Pennycress	15 seeds (≈ 15 mg)	1.5	0.2	0.4	2–3 (mix/0.5 h)	2	1.5
Cuphea	10 seeds (≈ 25 mg)	2 mL	0.2	0.8	2 (mix/0.5 h)	2	1.5
Canola	10 seeds (≈ 30 mg)	3 mL	0.2	0.8	1.5-2 (mix/0.5 h)	2	2

(~ 50 fold larger than Arabidopsis, Fig. 1) and subsequently the ratio of 5% sulfuric acid in methanol: toluene (2–3:1) was used for all other seeds to efficiently derivatize seed lipids. For all seed samples except Arabidopsis, sample tubes were inverted at half hour intervals during the incubation and/or pre-homogenization of the tissue was also tested. After derivatization, FAMEs were separated by the addition of both 0.8% (w/v) KCl and hexane (for volumes specific to different oilseeds see Table 1) followed by centrifugation for 5 min at 2500 x *g*. The optimized reaction conditions for each species can be found in Table 1.

Base transmethylation of hemp and meadowfoam seed lipids

Bead mill homogenization and base transmethylation of hemp and meadowfoam seed lipids was performed following the method of Vick et al. [19], with modifications to facilitate high-throughput screening of hemp seed samples. Briefly, approximately 15 hemp seeds (~ 250–300 mg) or 4 meadowfoam seeds (~ 40–45 mg) were transferred into VWR® Reinforced 2 mL Bead Mill Tubes (Cat. No: 10158-556). An internal standard of 150 µg tripentadecanoin (15:0), dissolved in toluene containing 0.005% butylated hydroxytoluene (BHT), was added. To homogenize the seeds, five 2 mm metal beads (VWR® Metal Beads size 2.4 mm; Cat. No: 10158-550) were added to each tube, and the samples were homogenized using a VWR bead mill (Cat. No. 75840-022) at the following settings: speed (S): 6:00; cycle time (T): 2:00; number of cycles (C): 1; pause dwell (D): 0:00. Following homogenization, the tubes were centrifuged at 12,000 rpm for 1 min to pellet the crushed seed material.

For transmethylation, 1.5 mL of 0.5 N sodium methoxide in methanol (Sigma Cat. No: 403067) was added to the tubes containing the crushed seed residue and extracted lipids. The tubes were vortexed thoroughly and incubated at 45 °C with vigorous shaking (1,500–2,000 rpm) in a thermomixer for 2 h to prevent sedimentation. After incubation, the tubes were centrifuged at 11,000 × *g* for 4 min to pellet the seed debris. The supernatant was transferred to an 8 mL Teflon-lined screw-capped glass tube. FAMEs were separated by adding 1.5 mL of 0.88% KCl and 3 mL of hexane, followed by vortexing and centrifugation at 2,000 × *g* for 3 min. The upper hexane layer, containing FAMEs, was transferred to gas chromatography (GC) vials for further analysis. Additionally, the effect of co-solvents, such as toluene (30% of the total reaction volume), was evaluated to enhance transesterification efficiency. Acid-based transesterification method of whole seeds (as above for Camelina) and crushed seeds were also carried out for comparison.

Fatty acid methyl ester quantification by GC-FID

FAMEs generated from seed lipids were separated and quantified against the 15:0 internal standard using an Agilent model 7890 GC-FID utilizing a DB HEAVYWAX UI column (Agilent, Santa Clara, CA, USA; 30 m length, 0.25 mm inner diameter, 0.25 µm film thickness). The GC-FID was run in split mode at 1:20 for lipid extracts, 1:40 seed samples with the exception of hemp which was run at a 1:90 split ratio. Both the GC injector and flame ionization detector were held at 255 °C with a helium flow of 1.05 mL/min. GC oven temperature started at 140 °C and increased at 20 °C/min to 200 °C, then at 5 °C/min to 260 °C, followed by a 3 min hold.

Statistical analysis

All data in this publication is expressed mean ± stand deviation with statistical significance analyzed using GraphPad Prism Version 10.4.0 using one-way ANOVA followed by a Holm-Šidák multiple comparison test. All pairwise comparisons were analyzed by Welches t-test.

Results

Developing high throughput Arabidopsis seed analysis

For Arabidopsis, techniques for seed oil analysis by direct seed lipid transmethylation to FAME have been developed [14, 15]. Our enhanced methodology focuses on developing a pipeline for image based rapid seed phenotyping to both count seeds and quantify seed size prior to oil quantification. The optimized seed phenotyping pipeline was designed around widely available commercial lenses for macrophotography and a DLSR camera to take consistent low magnification, high contrast seed images (Fig. 2a, S1). First, ~5 mg of seed tissue (roughly one micro-sampling spoon) was spread in a 3 × 5 cm frame outlined by tape in a weigh boat (Fig. 2a), a picture was taken, and the exact seed weight measured on an analytical balance. Image quality and seed counting/phenotyping were enhanced through consistent lighting and focal distance between sample and camera (Fig. S1). Clear seed images enabled accurate calculation of both seed number and seed size using the seed counting program GridFree [18] (Fig. 2a). From GridFree software output data the seed length, width, and area were quantified (Fig. 2b) in tandem with whole seed fatty acid analysis (Fig. 2c). Due to the very small size of Arabidopsis seeds, obtaining accurate oil analysis requires a large number of seeds, making manual counting of 100 or more seeds across numerous samples highly laborious. The image-based counting and seed phenotyping allows for rapid and accurate counting of Arabidopsis seeds, and determining seed shape phenotypes (length, width, area) are valuable when characterizing mutant or transgenic lines that may alter seed morphology. For larger studies batch analysis can be used for seed counting but will not output seed area data. Using the same setup as Arabidopsis, high throughput phenotyping is also possible for larger seeds such as Camelina (Fig. S2).

Oil quantification was then optimized for 5 mg samples of seed samples post phenotyping. Seed lipids were derivatized to FAMEs using 1 mL 5% H_2SO_4 in methanol with increasing concentrations of toluene as a co-solvent, incubated for 2 h, and compared to extracted lipids from the same seed stock as a control (Fig. 3). The most efficient derivatization was found with the addition of 500 µL toluene (33% of reaction volume) (Fig. 3a). Using the optimized reaction conditions (1.5 mL 5% H_2SO_4 in methanol, 500 µL toluene) incubation time was then tested. Total seed lipids were completely esterified to FAMEs after heating for 1 h at 85 °C (Fig. 3b). The fatty acid composition of direct whole seed esterification of lipids after 1 h was consistent with lipids extracted from the same tissue (Fig. 3c), which confirms the accuracy of the method.

Method optimization for multiple *Brassicaceae* oilseed crops

The direct seed quantification method developed from Arabidopsis was adapted to the emerging oil crop *Camelina sativa*. First, 10 crushed and 10, 15, 20 or 25 whole Camelina seeds were derivatized using the direct seed FAME protocol established for Arabidopsis. No significant difference was observed between measurements from 10 crushed and whole seeds (Fig. S3). To minimize the effects of seed-to-seed variability in individual sample measurements, subsequent analyses used 15 seeds per sample. Additionally, the method was further optimized by comparing to extracted lipids to establish conditions for total lipid fatty acid derivatization. Total fatty acid content of whole Camelina seeds was also lower than lipid extract, therefore, a whole seed Camelina method was further optimized. To ensure efficient and total transesterification of lipids in 15 seed samples, 5% H_2SO_4 in methanol was increased to 1.5 mL and the concentration of toluene and incubation time were tested. The addition of 600 µL toluene as a co-solvent (Fig. 4a) and a 2 h incubation time including mixing by inversion of samples every 0.5 h (Fig. 4b) produced the same total fatty acid content as the lipid extract control. It is important to note that prolonged incubation times of 3 h and longer reduced total fatty acid content, possibly due to sample degradation enhanced by components inherent in Camelina seeds. This result further emphasizes the importance of species-specific methodology for whole seed fatty acid derivatization methods. The Camelina method was then adapted to pennycress which have a similar size and shape (Fig. 1b). The incubation time for 15 whole pennycress seeds was tested with the reaction mix from Camelina (1.5 mL 5% H_2SO_4, 600 µL toluene)

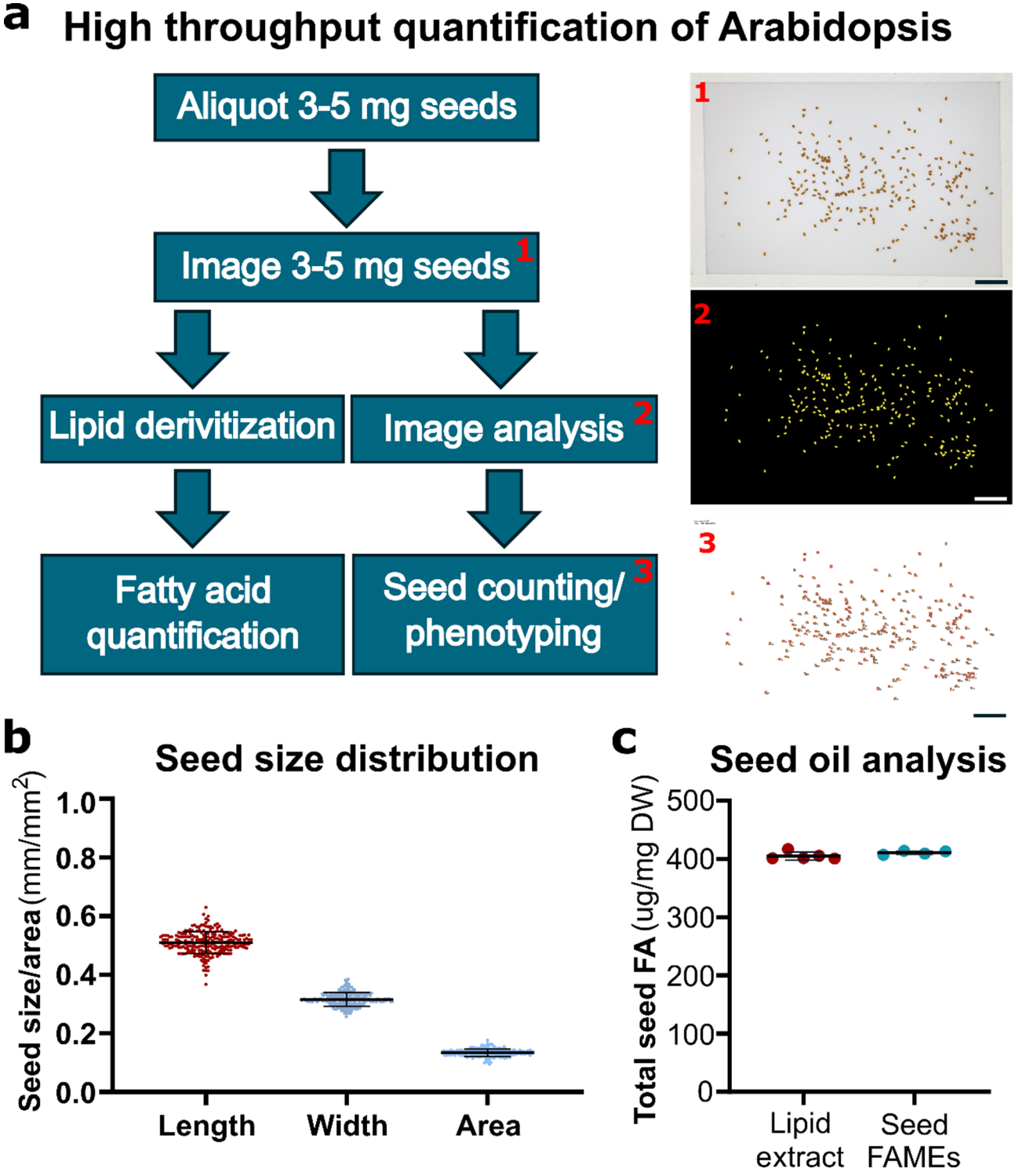

Fig. 2 Workflow of Arabidopsis seed phenotype and oil characterization. An optimized workflow for Arabidopsis for seed phenotyping and analysis (**a**), demonstrates how seeds are weighed and photographed for image analysis of seed number as well as seed size/area (**b**, $n = 230$, scale bars = 0.5 mm). Oil content was then analyzed through direct derivatization of whole seed tissue and compared to oil extracts from the same tissue (**c**, $n = 5$). Data is presented as mean ± standard deviation

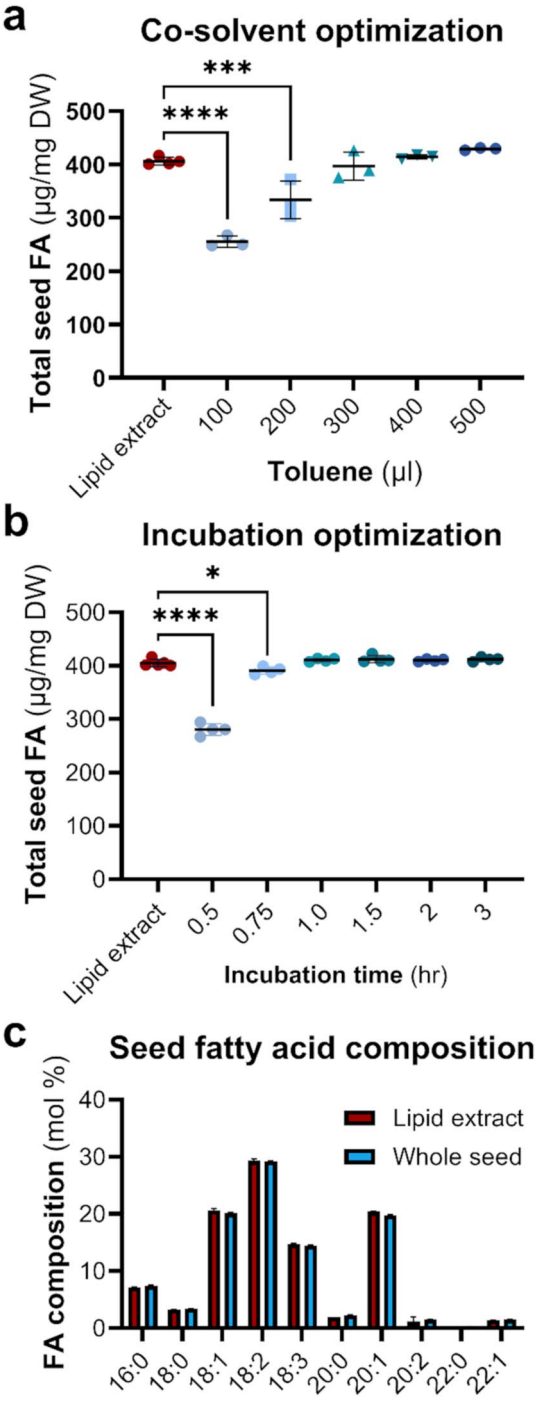

Fig. 3 Optimization of Arabidopsis oil quantification through direct seed transmethylation. The methodology for converting total seed fatty acids to fatty acid methyl esters (FAMEs) was optimized with respect to co-solvent volume (**a**) and incubation time (**b**). The optimized method of direct seed transmethylation (1mL 5% H_2SO_4 in methanol, 500 µL toluene, 1 h incubation) was compared to that of transmethylated whole seed lipid extracts (**c**). All samples contained 5 mg seed. Data is presented as mean ± standard deviation $n = 3+$, with significant differences measured by ANOVA followed by a multiple comparison test to lipid extract using the Holm-Šidák method; *, $p < 0.05$; ***, $p < 0.001$; ****, $p < 0.0001$

at 85 °C. Similar to Camelina, whole pennycress seed transmethylation produced the same quantity of FAME as the lipid extract control with a 2 h incubation, and slightly more with incubation extended to 3 h (Fig. 4c). Additionally, the fatty acid composition was consistent with lipid extracts for both Camelina (Fig. 4d) and pennycress (Fig. 4e) which demonstrates that direct seed esterification is accurate for larger seeds.

The Camelina protocol was adapted to the oilseed crop Canola. The volume of both 5% sulfuric acid in methanol and toluene co-solvent were increased to 3 and 1 mL, respectively, to account for larger seed size. Additionally, only 10 seeds were used for each reaction. Whole seeds and seeds ground with a polytron homogenizer were compared to lipid extracts (Fig. 4f) and no significant change was found between methods. Derivatization incubation time for the whole seed method was tested between 1 and 3 h and the optimal incubation time was 90 min (Fig. 4g). Lastly, fatty acid composition of lipid extracts was also compared to whole and ground seeds and no change was seen (Fig. 4h) supporting that direct seed transmethylation of 10 Canola seeds is consistent with lipid extraction methods.

Developing methods for the derivatization of oil in hemp and Meadowfoam seeds

Seeds of hemp and meadowfoam plants differ greatly from those of the *Brassicaceae* family. Chiefly, these seeds are much larger (Fig. 1) and include a tough outer hull as well as specialized metabolites such as terpenes, phytosterols, tocopherols, phenolic compounds and cannabinoids (hemp) [20, 21] each of which could affect direct seed transesterification. The direct seed quantification method developed for Brassicaceae crops was attempted for hemp (cv. Hlesia) and meadowfoam seed. First, 15 whole hemp seeds or 4 meadowfoam seeds were derivatized using the direct seed esterification protocol developed from Camelina and compared to that of lipid extract (Fig. 5a, c). For both sets of seeds, direct whole seed acid-catalyzed derivatization to FAME only produced ~ 16–25% as much FAME as the extracted lipid controls, however homogenization of the hemp and meadowfoams seeds prior to derivatization recovered the seed oil content (Fig. 5a, c). These results suggest the much larger seed size and/or tough outer hulls limit the efficacy of direct whole seed derivatization methods.

For rapid screening of seed lipid amount and fatty acid composition of a breeding population the homogenization of individual samples by hand or a polytron was found to be time consuming, and acid-catalyzed transmethylation is not compatible with plastic tubes utilized with bead mill homogenizers because it leads to plasticizers contaminating the FAME extract. Alternatively, a base-derived method of lipid esterification was attempted

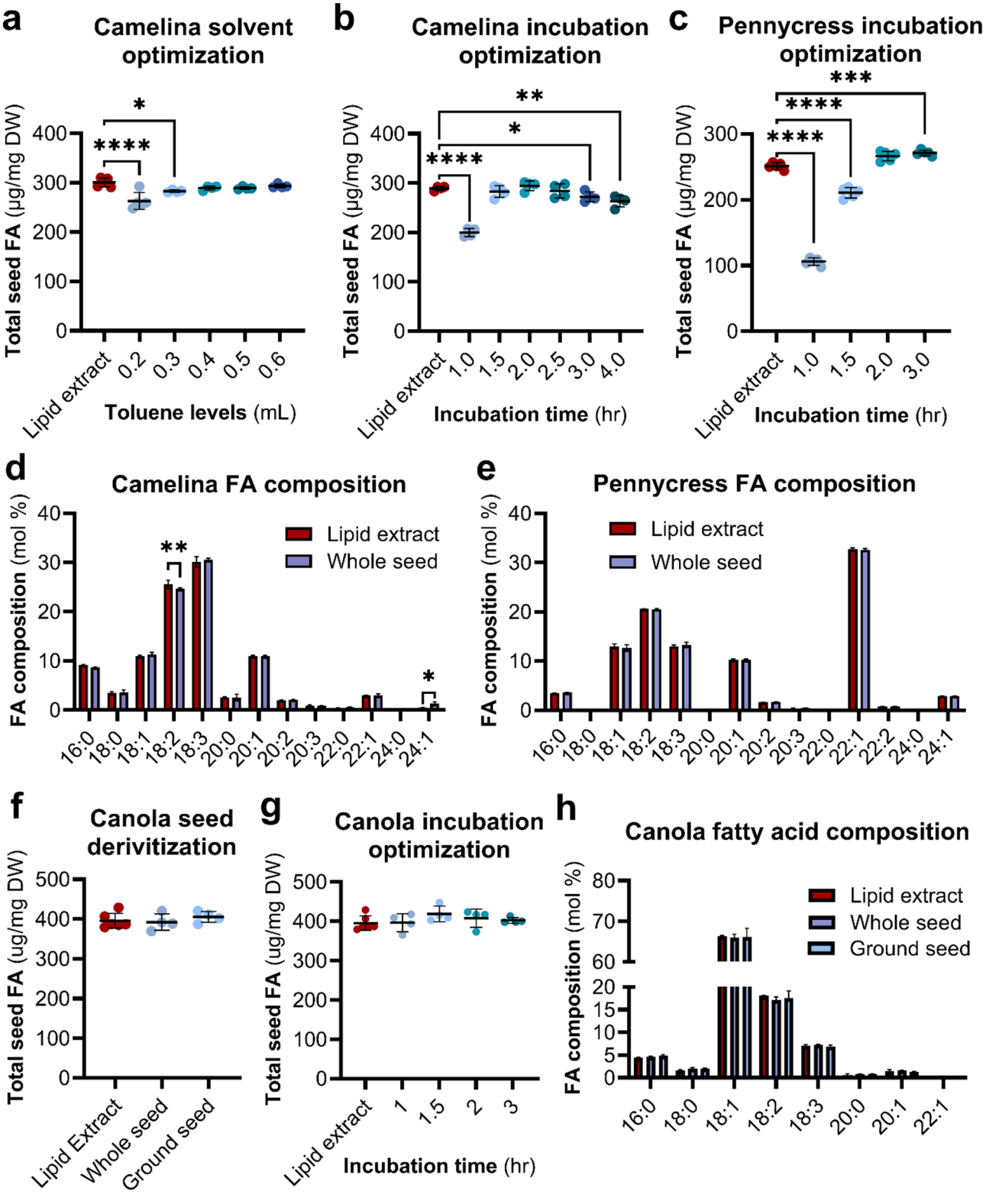

Fig. 4 (See legend on next page.)

(See figure on previous page.)

Fig. 4 Optimization of seed oil quantification in Brassicaceae species. Camelina fatty acid content was measured in response to increasing concentrations of toluene co-solvent (**a**; 1.5 mL 5% H_2SO_4, 2 h incubation). Fatty acid incubation times were then optimized for Camelina seeds (**b**) in a reaction containing 600 µL toluene and 1.5 mL 5% H_2SO_4. As pennycress is similar in size to that of Camelina, its incubation time was optimized under camelina reaction conditions (**c**). Fatty acid composition from the direct seed transmethylation under optimized conditions (600 µL toluene, 2 h incubation) was compared to that of total extracts for both Camelina (**d**) and pennycress (**e**). Methods of lipid quantification for larger Canola seed including whole seed and ground seed acid transmethylation (3 mL H_2SO_4, 1 mL toluene) were compared to lipid extract (**f**). Incubation time for whole Canola seeds was evaluated in comparison to lipid extract (**g**). The fatty acid composition of Canola was compared between directly transmethylated whole seeds, ground seeds, and lipid extracts (**h**). Data is presented as mean ± standard deviation $n = 4+$, with significant differences from lipid extract measured by ANOVA with follow-up multiple comparison test to WT using the Holm-Šidák method; *, $p < 0.05$; **, $p < 0.01$; **** $p < 0.0001$

using sodium methoxide (CH_3ONa) as it allowed for samples to be ground in a plastic tube with a commercial bead mill to facilitate high-throughput sample homogenization. The resulting base-catalyzed FAMEs were compared to extracted lipids using the same method as well as acid catalysis (Fig. 5a, c). Measured fatty acid content (Fig. 5a, c) and composition (Fig. 5b, d) were consistent between the base-esterified extracted lipid control and homogenized seeds. However, total fatty acid levels for base transmethylated hemp were ~10% lower than the acid-catalyzed FAME production method likely due to the inability of base-catalyzed transmethylation to esterify free fatty acids. The base transmethylated meadowfoam was not significantly different than the acid method, likely due to less free fatty acids in meadowfoam seeds than hemp. The use of co-solvent toluene was also tested with the sodium methoxide-based transmethylation for hemp, but it did not improve FAME yield (Fig. S4) and thus was not included in the optimized methods. Overall, while acid-catalyzed derivatization lipid extracts measured higher total oil content for hemp (but not meadowfoam), the rapid sample processing of bead mill homogenization combined with base esterification of seed tissue allows for efficient analysis of seed oil amount and fatty composition that is congruent with lipid extracts (Fig. 5). Therefore, for rapid screening of breeding or other research populations of plants that produce larger/tougher seeds such as hemp or meadowfoam, the higher throughput bead mill homogenization in plastic tubes combined with base catalyzed-transmethylation is suitable to determine the total fatty acid amount of larger/tougher seeds with ~90–100% accuracy, and a fatty acid composition that comparable to acid-catalyzed methods. Additionally, our results indicate that confirmation of total seed fatty acid content of large/tough seeds by acid-catalyzed derivatization to FAME methods can be done more quickly with direct derivatization of crushed/homogenized seeds than the more laborious total lipid extraction prior to derivatization methods.

Optimizing FAME methodology for *Cuphea viscosissima*

Direct fatty acid transmethylation was also adapted to seeds of *Cuphea viscosissima* which produces unusual medium chain fatty acids (MCFA). Recent research suggests that engineering MCFA into high yielding oilseed crops can be a valuable carbon source for sustainable aviation fuel [22–24], and thus understanding lipid metabolism in species such as Cuphea is gaining interest. Due to the unique shape and size of Cuphea seeds (Fig. 1), both Camelina (acid) and hemp (base) methodologies were tested and compared to lipid extracts (Fig. 6a). Direct seed transmethylation using sulfuric acid-based methylation was superior to the ground seed sodium methoxide method which demonstrated an ~10% decrease in measurable fatty acid content (Fig. 6a). Further analysis of fatty acid composition showed small albeit significant shifts in fatty acid composition of the sodium methoxide-catalyzed seed samples (Fig. 6b), that together with the reduced total lipid content suggests that the base method does not capture a free fatty acid fraction that contains a significant amount of 8:0. Overall, acid-catalyzed direct transmethylation of *Cuphea* seeds provides a nearly accurate measurement of oil content and fatty acid composition, and demonstrates the method works well in non-Brassicaceae seeds of similar size to that of Brassicaceae crops.

The optimized methodology for direct whole seed lipid derivatization to FAME for each species is summarized in Table 1. Table 2 summarizes the methodology for seeds that require prior crushing before derivatization by either acid or base catalyzed methods.

Discussion

Rapid direct seed transmethylation/esterification is a viable strategy for oil and fatty acid quantification and is adaptable to a range of oilseed species with diverse seed sizes

To date, methods of detailed seed oil analysis for seeds larger than Arabidopsis have been mostly dependent on labor intensive lipid extraction prior to derivatization and GC or less accurate non-destructive methods. Our results demonstrate that a large variety of plant seeds are amenable to acid-based direct seed lipid transmethylation including camelina, pennycress, canola, and cuphea (Figs. 3, 4 and 6). Efficient derivatization of lipids in seeds up to 5 mg in size is dependent on increasing reagent volumes, utilizing toluene as a co-solvent, and increasing incubation time with sample mixing during derivatization (Figs. 3, 4, 5 and 6). The reagent volume of 5% sulfuric acid in methanol was scalable at a rate of 0.1 mL per

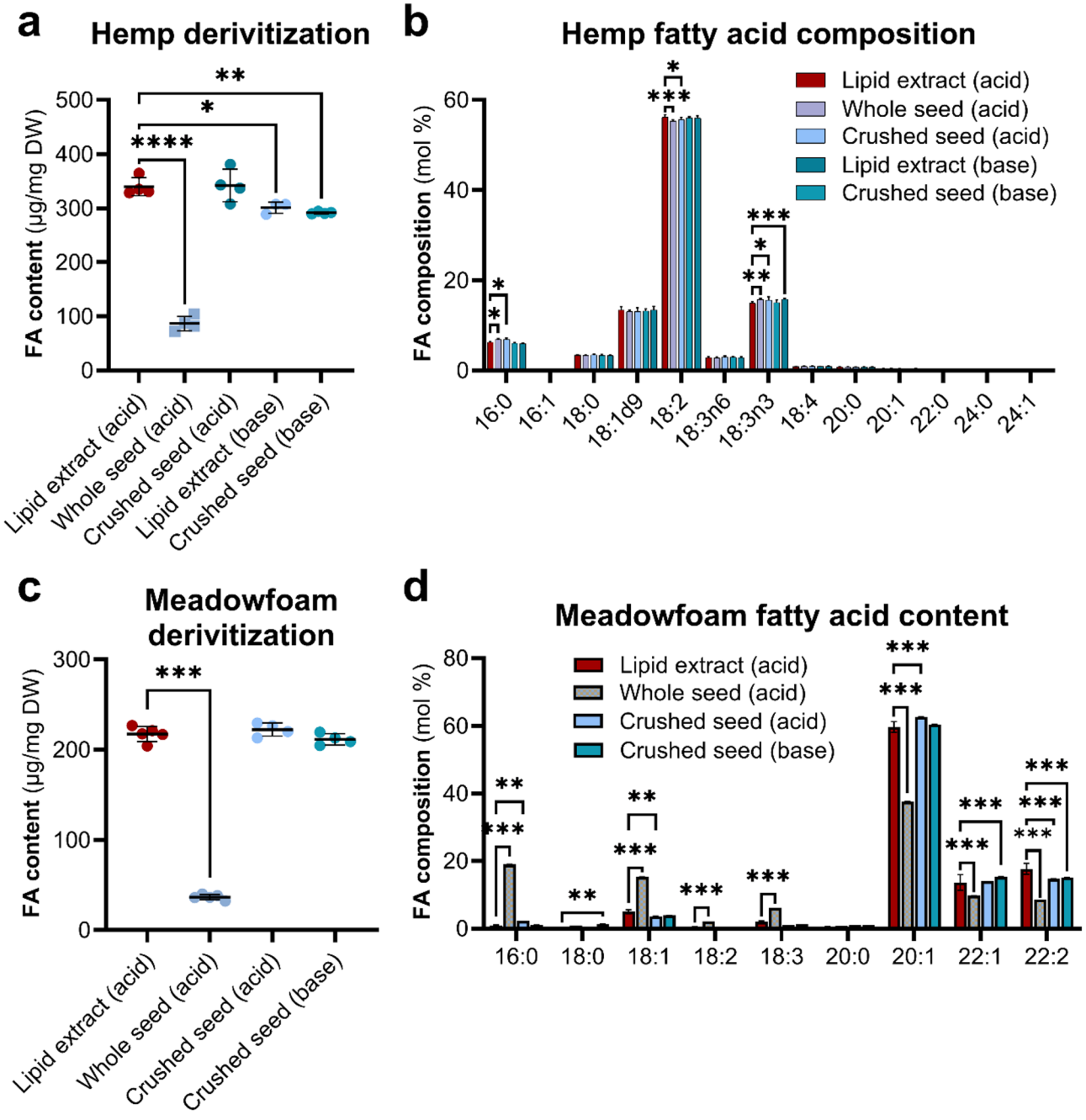

Fig. 5 Optimization of high throughput analysis of larger hemp and meadowfoam seeds. Total Seed lipids were quantified from hemp using 15 whole and crushed seed transmethylation by acid (1.5 mL H_2SO_4 in methanol, 0.45 mL toluene, 2 h incubation) and base (1.5 mL 0.5 M CH_3ONa in methanol, 50 µL toluene, 2 h incubation) and compared to lipid extracts derivatized by each method, respectively. The corresponding measurements are shown as total lipid content (**a**) and fatty acid composition (**b**). The developed hemp method was then adapted to 4 seed meadowfoam samples to determine oil content (**c**) and fatty acid composition (**d**). Data is presented as mean ± standard deviation $n = 4$, significant differences from lipid extract measured by ANOVA with follow-up multiple comparison test to WT using the Holm-Šidák method; *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$; ****, $p < 0.0001$

mg sample in seed tissue (Fig. 3b and c 3f, 6a) and the volume of added toluene was also increased so it accounted for 25–30% of total reagent volume. Aliquots of at least 10 + seeds were used, incubation times lengthened, and samples were inverted every 30 min to improve reaction efficiency and reduce variation in measured lipids. For all seeds other than Arabidopsis, a single seed contains enough lipid to measure directly. However, due to seed-to-seed variability from a single plant, use of 10 + seeds gives a better representation of the average seed lipid content from the individual plant. When compared to lipid extract controls direct whole-seed acid-catalyzed

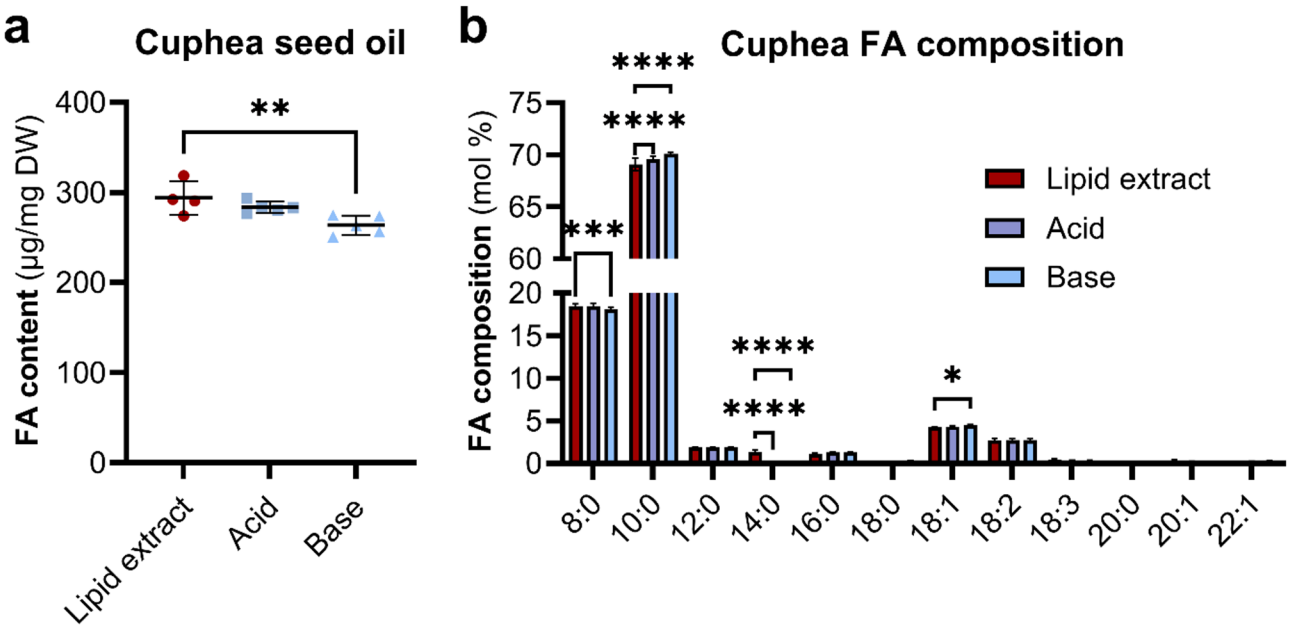

Fig. 6 Optimized methods for *Cuphea viscosissima*. Total Seed lipids were quantified from cuphea using both direct whole seed transmethylation by acid (2 mL of 5% H_2SO_4 in methanol, 1 mL toluene, 2 h incubation) and ground base (1.5 mL CH_3ONa, 50 µL toluene, 2 h incubation) method and compared to lipid extracts by total lipid content (**a**) and fatty acid composition (**b**). Data is presented as mean ± standard deviation $n = 5$, with significant differences to lipid extract measured by ANOVA with follow-up multiple comparison test to WT using the Holm-Šidák method; *, $p < 0.05$; **, $p < 0.01$; ****, $p < 0.0001$

Table 2 Optimized acid- and base-catalyzed derivatization methods for ground seed tissue

Species; Seed weight/ sample prep	Reagent Volumes (mL) Mix the following regents and standard (1.6 g/L 15:0-TAG in toluene)			Incubation time (h)	Separation centrifuge at 2500 x g for 5 min following addition of:	
Acid method						
Reagent	5% H_2SO_4 in MeOH	Standard in toluene	Toluene	incubation at 85 °C	Hexane (mL)	0.8% KCl (mL)
Meadowfoam; ≈100 mg tissue/ 4 seeds ground by polytron	3	0.2	0.8	2 (mix/0.5 h)	2	1.5
Hemp; ≈350 mg tissue/ 10 seeds ground by bead mill	3	0.2	0.8	2 (mix/0.5 h)	2	1.5
Base method						
Reagent	0.5 M CH_3ONa	Standard		incubation at 45 °C	Hexane (mL)	0.8% KCl (mL)
Meadowfoam; ≈100 mg tissue/ 4 seeds ground by polytron	1.5	50 µL		2 (continuous mixing)	3	1.5
Hemp; ≈350 mg tissue/ 10 seeds ground by bead mill	1.5	50 µL		2 (continuous mixing)	3	1.5

derivatization to FAME was able to accurately quantify the total seed lipid content for all plant seeds tested except hemp and meadowfoam. The fatty acid composition of these seed derivatization methods was also consistent with the lipid extracts supporting that our optimized methods (Table 1; Figs. 4d, e and h and 6b) are an accurate measurement of total seed lipid amount and fatty acid composition suitable for rapidly screening seed lipid content within breeding or bioengineering research.

Interestingly, hemp and meadowfoam seeds were not efficiently esterified using whole seed methods (Fig. 5a). This is likely the result of the large seed size and thick protective hull relative to the other tested seeds [20]. Prior seed disruption (crushing/homogenization) before derivatization by acid- or based-catalyzed methods was required to accurately quantify total seed lipid content compared to the extracted lipid controls (Fig. 5a and c). As an alternative, a one-tube method utilizing a bead mill for high throughput homogenization and subsequent transmethylation was developed to improve speed and reduce data variability associated with grinding seed tissue and transfer of ground seed meal to reaction tubes. To avoid concerns of polypropylene plastic tube degradation from solvents (such as toluene and hexane; [25] and the resultant degradation products affecting both equipment and sample analysis, a base-catalyzed derivatization

method was developed that is plastic tube compatible and accurately measures both oil content and fatty acid composition (Table 2; Fig. 5). A drawback of base-catalyzed transesterification is that free fatty acids are not esterified. Therefore, the total fatty acid content of a tissue will only reflect the esterified fatty acids (e.g. oils and membrane lipids) and the variability of measured fatty acid content will be dependent of the free fatty acid content of the seed. For hemp this accounted for an ~10% reduction in measured fatty acid content with the base-catalyzed method, however meadowfoam fatty acid content was not statistically different utilizing the base-catalyzed method likely indicating less free fatty acids in meadowfoam seeds than in hemp seeds. Alternatively for hemp, the base-catalyzed method gives a better estimate of the actual oil fraction of the seed that is likely preferred in breeding programs. Together, both acid- and base-catalyzed methods are high-throughput seed procedures for seed lipid derivatization to FAME that are scalable to various oilseeds. These methods are valuable for plant breeding and engineering research where accurate seed oil amount and fatty acid composition of many individual plants are needed.

Quality of high throughput seed phenotyping data is dependent on both image quality and repeatability

In addition to oil content, seed size and shape are often key attributes investigated within oilseed basic research or breeding and engineering efforts. Determining these features on large numbers of seeds can be especially difficult with smaller seeds, whose size make physically measuring individual seed exceedingly difficult. Additionally, it is useful to be able to correlate the seed size, shape, weight, and lipid content for the same sample. Our optimized workflow (Fig. 2a) has greatly increased the speed of producing these combined data sets. We took advantage of the seed counting software GridFree which was originally designed for larger seeds such as black bean, corn, chickpea, and lentil [18] and adapted it to count a large number of small seeds captured under relatively low magnification using a macrophotography lens (Fig. 2). The GridFree software gives a rapid analysis of seed length, width, and area (although volume cannot be determined from a 2D image). Using this methodology, we were able to design a high throughput counting workflow that eliminates the need to either physically count seeds or use more labor-intensive image analysis tools (such as ImageJ [26]), to manually measure seed area.

However, for accurate seed counting and phenotyping consideration needs to be taken into account when designing a camera for consistent seed images. Rather than a high-resolution camera sensor, investment into high quality macrophotography lens, lighting and stationary camera mount are especially important. The above study used stationary lighting and camera mount to control image quality and improve the repeatability of seed image lighting and staging to ease computer processing. For example, during batch analysis the camera settings (e.g. image ISO, F-stop settings) and the lens focus were controlled manually for consistent high contrast images. During image processing, high resolution images were found to increase the load on the analysis software with no benefit to seed analysis. Therefore, image resolution was reduced from camera maximum to 1920×1280 to lower the computational requirements and increase the throughput speed. Using these common settings for a large sample batch allowed for accurate automated seed counting and reduced the error rate of the GridFree software batch analysis system. It should be noted that these settings may be unique to each individual setup depending on the amount of directed and incidental light and a test seed batch may be helpful when validating the seed counting functions and optimizing any seed counting system.

Conclusions

Breeding and engineering research on oilseeds requires accurate analysis of both oil yield and fatty acid composition. Since these analyses often involve examining large sample sets, a reliable and high-throughput methodology is essential. We have developed improved in situ oil analysis methods for multiple oilseed crops that can provide in-depth total lipid content and fatty acid composition analysis without the time consuming and costly resource expenditure for prior lipid extraction. When coupled with the seed counting and phenotyping software GridFree, this analysis workflow can serve as a valuable tool for characterizing breeding and engineered populations of seeds. The methodology is optimized for small numbers of seeds from various crop plants (Tables 1 and 2). We demonstrated the precision of the method with just 4 replicates from a single seed stock. Under field conditions the environment can affect the plant-to-plant variability of seed oil content, and thus the number of plants analyzed should be considered when designing field experiments. The value of the methods optimized here allows for a rapid analysis of only a few seeds from each plant to facilitate processing of the large numbers of samples from many individual plants in a breeding population. Here we optimized direct seed tissue derivatization to FAME methods for seven oilseed species, establishing a framework for optimization across many other oilseed species. Finally, the adoption of consistent accurate and high throughput methodology between research groups will provide enhanced reliability when comparing results between different oilseed research studies including breeding, oilseed engineering, or the responses of oil content various biotic and abiotic stresses.

Abbreviations

FAME	Fatty acid methyl esters
GC	Gas chromatography
MS	Mass spectrometry
FID	Flame ionization detection

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13007-025-01388-3.

Supplementary Material 1

Acknowledgements

The authors would like to thank Abraham Baker for his assistance in the analysis of ***Brassica napus*** seed material.

Author contributions

MGG, PP, and PDB designed experiments, MGG, PP, and NZ performed experiments, MGG, PP, and PDB analyzed data and wrote the manuscript, all authors have read the manuscript and approve it for publication.

Funding

This work is supported by the Agriculture and Food Research Initiative Competitive Grants Program Foundational and Applied Science Program, project award no. 2020-67013-30899 and 2023-67013-39022, from the U.S. Department of Agriculture's National Institute of Food and Agriculture. In addition, this work was supported by the Hatch Project #1015621 and Multi-State Project #1013013 from the U.S. Department of Agriculture. This material is based upon work supported by the National Science Foundation under Award No. NSF-MCB 2242822 and NSF-PGRP 1829365; and the U. S. Department of Energy, Office of Science, Office of Biological and Environmental Research, under award number DE-SC0023142.

Data availability

No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

[1]Institute of Biological Chemistry, Washington State University, Pullman, Washington 99164, USA

Received: 1 April 2025 / Accepted: 9 May 2025
Published online: 22 May 2025

References

1. Biermann U, Bornscheuer UT, Feussner I, Meier MAR, Metzger JO. Fatty acids and their derivatives as renewable platform molecules for the chemical industry. Angew Chem Int Ed. 2021;60(37):20144–65. https://doi.org/10.1002/anie.202100778.
2. Dyer JM, Stymne S, Green AG, Carlsson AS. High-value oils from plants. Plant J. 2008;54(4):640–55. https://doi.org/10.1111/j.1365-313X.2008.03430.x.
3. Baud S. Seeds as oil factories. Plant Reprod. 2018;31(3):213–35.
4. Mumtaz F, Zubair M, Khan F, Niaz K. Chap. 22 - Analysis of plants lipids. In: *Recent Advances in Natural Products Analysis.* Edited by Sanches Silva A, Nabavi SF, Saeedi M, Nabavi SM: Elsevier; 2020: 677–705. https://doi.org/10.1016/B978-0-12-816455-6.00022-6
5. Correa SM, Fernie AR, Nikoloski Z, Brotman Y. Towards model-driven characterization and manipulation of plant lipid metabolism. Prog Lipid Res. 2020;80:101051. https://doi.org/10.1016/j.plipres.2020.101051.
6. Bates PD, Shockey J. Towards rational control of seed oil composition: dissecting cellular organization and flux control of lipid metabolism. Plant Phys. 2025;197(2):kiae658. https://doi.org/10.1093/plphys/kiae658.
7. Li X, Zhang L, Zhang Y, Wang D, Wang X, Yu L, Zhang W, Li P. Review of NIR spectroscopy methods for nondestructive quality analysis of oilseeds and edible oils. Trends Food Sci Technol. 2020;101:172–81. https://doi.org/10.1016/j.tifs.2020.05.002.
8. Castejón D, Mateos-Aparicio I, Molero MD, Cambero MI, Herrera A. Evaluation and optimization of the analysis of fatty acid types in edible oils by 1H-NMR. Food Anal Methods. 2014;7(6):1285–97. https://doi.org/10.1007/s12161-013-9747-9.
9. Castejón D, Fricke P, Cambero MI, Herrera A. Automatic 1H-NMR screening of fatty acid composition in edible oils. Nutr. 2016;8(2):93. https://doi.org/10.3390/nu8020093.
10. Borisjuk L, Rolletschek H, Fuchs J, Melkus G, Neuberger T. Low and high field magnetic resonance for in vivo analysis of seeds. Mater. 2011;4(8):1426–39. https://doi.org/10.3390/ma4081426.
11. Saini RK, Prasad P, Shang X, Keum Y-S. Advances in lipid extraction methods—a review. Nt J Mol Sci. 2021;22(24):13643. https://doi.org/10.3390/ijms222413643.
12. Christie WW, Han X. Chap. 7 - Preparation of derivatives of fatty acids. In: *Lipid Analysis (Fourth Edition).* Edited by Christie WW, Han X: Woodhead Publishing; 2012: 145–158. https://doi.org/10.1533/9780857097866.145
13. Wu Z, Zhang Q, Li N, Pu Y, Wang B, Zhang T. Comparison of critical methods developed for fatty acid analysis: A review. J Sep Sci. 2017;40(1):288–98. https://doi.org/10.1002/jssc.201600707.
14. Li Y, Beisson F, Pollard M, Ohlrogge J. Oil content of Arabidopsis seeds: the influence of seed anatomy, light and plant-to-plant variation. Phytochem. 2006;67(9):904–15. https://doi.org/10.1016/j.phytochem.2006.02.015.
15. McDonald K, Xu Y, Chen G. A simple and cost-effective direct transmethylation procedure for plant lipid analysis. Am Oil Chem Soc. 2023;100(7):521–8. https://doi.org/10.1002/aocs.12709.
16. Garces R, Mancha M. One-Step lipid extraction and fatty acid Methyl esters Preparation from fresh plant tissues. Anal Biochem. 1993;211(1):139–43. https://doi.org/10.1006/abio.1993.1244.
17. Folch J, Lees M, Stanley GHS. A simple method for the isolation and purification of total lipides from animal tissues. J Biol Chem. 1957;226(1):497–509. https://doi.org/10.1016/S0021-9258(18)64849-5.
18. Hu Y, Zhang Z. GridFree: a python package of imageanalysis for interactive grain counting and measuring. Plant Physiol. 2021;186(4):2239–52. https://doi.org/10.1093/plphys/kiab226.
19. Vick BA, Jan C-C, Miller JF. Two-year study on the inheritance of reduced saturated fatty acid content in sunflower seed. Helia. 2004;27(41):25–40. https://doi.org/10.2298/HEL0441025B.
20. Farinon B, Molinari R, Costantini L, Merendino N. The seed of industrial hemp (Cannabis sativa L.): nutritional quality and potential functionality for human health and nutrition. Nutr. 2020;12(7). https://doi.org/10.3390/nu12071935.
21. Stevens JF, Reed RL, Morré JT. Characterization of phytoecdysteroid glycosides in Meadowfoam (Limnanthes alba) seed meal by positive and negative ion LC-MS/MS. J Agri Food Chem. 2008;56(11):3945–52. https://doi.org/10.1021/jf800211k.
22. Kim HJ, Silva JE, Vu HS, Mockaitis K, Nam J-W, Cahoon EB. Toward production of jet fuel functionality in oilseeds: identification of FatB acyl-acyl carrier protein thioesterases and evaluation of combinatorial expression strategies in Camelina seeds. J Exp Bot. 2015;66(14):4251–65. https://doi.org/10.1093/jxb/erv225.
23. Iskandarov U, Silva JE, Kim HJ, Andersson M, Cahoon RE, Mockaitis K, Cahoon EB. A specialized Diacylglycerol acyltransferase contributes to the extreme Medium-Chain fatty acid content of Cuphea seed oil. Plant Physiol. 2017;174(1):97–109. https://doi.org/10.1104/pp.16.01894.
24. Esfahanian M, Nazarenus TJ, Freund MM, McIntosh G, Phippen WB, Phippen ME, Durrett TP, Cahoon EB, Sedbrook JC. Generating Pennycress (Thlaspi arvense) seed triacylglycerols and Acetyl-Triacylglycerols containing Medium-Chain fatty acids. Front Ener Res. 2021;9. https://doi.org/10.3389/fenrg.2021.620118.
25. Blackadder DA, Le Poidevin GJ. Dissolution of polypropylene in organic solvents: 1. Partial dissolution. Polym. 1976;17(5):387–94. https://doi.org/10.1016/0032-3861(76)90233-0.

26. Schneider CA, Rasband WS, Eliceiri KW. NIH image to imageJ: 25 years of image analysis. Nat Methods. 2012;9(7):671–5. https://doi.org/10.1038/nmeth.2089.

Publisher's note