Available online at www.sciencedirect.com
SCIENCE DIRECT
PHYTOCHEMISTRY
Phytochemistry 67 (2006) 904–915
www.elsevier.com/locate/phytochem

Oil content of Arabidopsis seeds: The influence of seed anatomy, light and plant-to-plant variation

Yonghua Li, Fred Beisson, Mike Pollard, John Ohlrogge *

Department of Plant Biology, Michigan State University, East Lansing, MI 48824-1312, USA

Received 14 December 2005; received in revised form 30 January 2006
Available online 4 April 2006

Abstract

Arabidopsis thaliana is frequently used as a model for the study of oilseed biology and metabolism. However, the very small seeds of Arabidopsis can complicate analysis of their oil content and influence the application of results to larger-seeded plants. Here, we describe how seed anatomy, light, and plant-to-plant variation influence the content and measurement of oil in Arabidopsis seeds. The anatomy of Arabidopsis and *Brassica napus* seeds were compared and the distribution of mass, oil and the fatty acid composition of different seed parts were determined. In Brassica, 90% of the seed oil resides in the cotyledons that contribute 74% of seed mass. By contrast, the values for Arabidopsis are 60% and 45%, respectively, with a higher fraction of the oil deposited in the radicle, hypocotyl, endosperm and seed coat. Growth of Arabidopsis plants with 600 μmol m^{-2} s^{-1} light resulted in a two-fold higher seed yield, a 40% increase in mass per seed and a 60% increase in oil per seed compared to growth at 100 μmol m^{-2} s^{-1}. Factors that influence the analysis of oil content were evaluated. Intact-seed transmethylation followed by gas chromatography (GC) analysis provided reproducible analysis of Arabidopsis seed oil. However, plant-to-plant variation in oil content is large and we analyzed how this influences the ability to detect statistically valid changes in oil between different genotypes. These observations establish a reference data set on the fatty acid composition and distribution of mass and oil between tissues of Arabidopsis seeds that should help to predict the applicability of results obtained with Arabidopsis to other oilseeds.

Keywords: *Arabidopsis thaliana*; *Brassica napus*; Brassicaceae; Oilseeds; Seed coat; Lipid; Fatty acids

1. Introduction

In addition to its role as a general model for plant biology research (Meyerowitz and Somerville, 1994), *Arabidopsis thaliana* also provides an attractive system for the study of oilseeds. Arabidopsis is a close relative of *Brassica napus*, a major oilseed crop. Like *B. napus*, Arabidopsis stores oil in its cotyledons as the major seed carbon reserve. It has similar seed phylogeny and development and on average, genes of these species share ca. 85% nucleotide identity in their coding region (Cavell et al., 1998). Together these factors facilitate the application of information obtained from Arabidopsis research to crops of the Brassicaceae family.

Arabidopsis is also a valuable tool because large numbers of independent transgenic lines can be obtained within weeks or a few months, whereas time and space investments required for generation of *B. napus* transformants are much greater. Thus, Arabidopsis can often serve to provide relatively fast and low-cost "proof of concept" results before larger investments in crop transformation are made (e.g., Murphy, 1996; Voelker et al., 1996; Zou et al., 1997). However, limitations associated with the use of Arabidopsis as a model for crop species are that the plants are small and its seeds are tiny (~20 μg). These factors contribute to at least two related issues. First, a small amount of total seed material is available and this limits the types and number of analytical procedures that can be

* Corresponding author. Tel.: +1 517 353 0611/9399; fax: +1 517 353 1926.
E-mail address: OHLROGGE@msu.edu (J. Ohlrogge).

0031-9422/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.phytochem.2006.02.015

conducted. Second, the proportion of the seed mass that is contributed by the seed coat and endosperm is up to 30–35% for Arabidopsis (this study), which is substantially higher than for larger seeds of *B. napus*. These issues provide challenges for measuring oil content accurately and for extrapolating results from Arabidopsis to Brassica or other species with much larger seeds.

In this study, we focus on several factors that influence the oil content of Arabidopsis seeds. We consider the influence of seed anatomy on seed composition, and we determine the impact of light on seed oil content, size and yield. We also validate and optimize a whole-seed transmethylation method for determination of Arabidopsis seed oil content and fatty acid composition and demonstrate the influence of plant-to-plant variability on the ability to make statistically valid comparisons between genotypes.

2. Results and discussion

2.1. Methods for oil content determination

Because of the very small size of Arabidopsis seeds and the wide range of published results on Arabidopsis seed oil content (Table 1), this study first considered several methods for Arabidopsis seed oil analysis. Gravimetric quantification is often used as a standard method to calibrate other methods and is completely appropriate for samples of Arabidopsis seeds of 100 mg or more. Since a single Arabidopsis plant can produce up to 1 g or more of seeds (Meinke, 1994), replicate analyses can be easily obtained. The major disadvantages of this method are that it requires larger amount of sample and the procedure is more time consuming than the direct methylation method described below. In addition, it does not provide information on fatty acid composition of the oil. Wide line ^{1}H-NMR measurements (Rutar, 1989) and reflectance IR of oil content have been used in several studies of Arabidopsis seed oil (Table 1). These are nondestructive and non-invasive and short scan times ($\leqslant$1 min) lead to high throughput, but the methods generally require more sample than GC-based methods (e.g., 50 mg used in O'Neill et al., 2003) and do not provide fatty acid composition. Complete information on fatty acid composition can be obtained via GC analysis. Although direct methylation has been reported as a method for plant fatty acid analysis (e.g., Browse et al., 1986; Epp and Pollard, 1993; Garces and Mancha, 1993), the almost two-fold variation in results reported in the literature for Arabidopsis seed oil content using this and other methods (Table 1) prompted us to examine the technique and factors that influence analytical results. To foster consistent analysis methods, in Section 4 and below, we present in some detail the procedures we have found most reliable and factors that can influence the accuracy of results.

2.2. Evaluation and validation of direct methylation

In Arabidposis seeds, almost all of the fatty acids are esterified in the form of triacylglycerols (TAG) with the remainder originating largely from diacylglycerols and membrane lipids (this study and Baud et al., 2002). Quantitative extraction of oil from seeds usually requires grinding the seeds to provide access for organic solvents. However, because Arabidopsis seeds are extremely small and have thin seed coats we found that this step is not necessary. Table 2 indicates that no differences in oil content or fatty acid composition were observed between ground and non-ground samples; this demonstrates that whole seed transmethylation provides a simple one-step procedure for extracting oil from Arabidopsis seeds and simultaneous

Table 1
Examples of publications reporting oil content of Arabidopsis wild type seeds

Ecotype	Method	Oil content (mean ± s.e. scale[a])		References
		Oil%	μg/seed	
Col	Whole seed transmethylation–GC	35.88 ± 0.18% (10 mg, $n = 6$)	5.95 ± 0.25 μg oil/seed (50-seed, $n = 6$)	This study
		–	5.5 μg FAs/seed (10-seed, $n = 15$)	Cernac and Benning (2004)
		–	7.45 ± 0.4 μg FAs/seed (100-seed, $n = 5$)	Zou et al. (1997)
		31 ± 0.50%[a]	–	Periappuram et al. (2000)
		26.2 ± 0.8% ($n = 12$)	–	Katavic et al. (1995)
	Oil extraction transmethylation–GC	–	3.2 ± 0.5 μg TAG/seed (10-seed, $n = 2$)	Focks and Benning (1998)
		27 ± 1.0% (100–200 seeds, $n = 5$)		Jako et al. (2001)
		–	6.75 ± 0.25μg FAs/seed (100-seed, $n = 4$)	Marillia et al. (2003)
		31.0 ± 3.1% ($n = 22$)	–	Thelen and Ohlrogge (2002)
		36.4%[a]	–	Zou et al. (1999)
	NMR	36% (50-mg, $n = 10$)	–	O'Neill et al. (2003)
Cvi Ler	Direct methylation/NMR	39.3 ± 1.8% (200 mg, $n = 5$) 43.4 ± 0.9% (200 mg, $n = 5$)	–	Hobbs et al. (2004)
WS	Extraction–methylation–GC	–	6.2 ± 0.5 μg FAs/seed (20-seed, $n = 3$)	Baud et al. (2002)
		23.8 ± 0.5% (20-seed, $n = 3$)	–	Baud et al. (2003)

–: not available.
[a] Scale and sample size indicated where information available.

Table 2
Influence of grinding and number of extractions on seed oil content and fatty acid composition for samples processed by direct methylation

Extraction times	1×		3×	1×
Sample preparation	Ground	No grinding	No grinding	No grinding
Oil%	34.2 ± 0.29	35.2 ± 1.05	37.3 ± 0.68	35.7 ± 0.66
Fatty acids (mol%)				
16:0	8.7 ± 0.01	8.8 ± 0.05	8.4 ± 0.01	8.4 ± 0.02
18:0	3.6 ± 0.01	3.6 ± 0.01	3.6 ± 0.01	3.7 ± 0.01
18:1	15.0 ± 0.01	15.1 ± 0.10	16.1 ± 0.05	16.3 ± 0.05
18:2	29.0 ± 0.05	28.8 ± 0.05	28.7 ± 0.02	28.3 ± 6.08
18:3	19.2 ± 0.05	19.3 ± 0.10	17.5 ± 0.09	16.5 ± 0.09
20:0	2.2 ± 0.01	2.2 ± 0.01	2.1 ± 0.01	2.3 ±0.01
20:1	20.2 ± 0.02	20.3 ± 0.07	20.1 ± 0.01	20.9 ± 0.08
20:2	2.0 ± 0.01	1.9 ± 0.01	1.8 ± 0.01	1.8 ± 0.02
22:1	1.7 ± 0.01	1.6 ± 0.01	1.6 ± 0.01	1.7 ± 0.01

Data are the mean ± s.e. ($n = 4$).
Note: Results in the first two columns were obtained from different batches of seeds as that used for analysis in the last two columns.

preparation of fatty acid methyl esters (FAMEs) by acid catalyzed transesterification. In addition, these results show that omitting the grinding step is not likely to be responsible for differences in oil amount as reported in the literature (see Table 1).

Following transmethylation, fatty acid methyl esters are quantified by GC by comparing peak areas of identified FAME to the area of an internal standard. The resulting data provide an analysis of total seed fatty acid quantity while also providing a profile of fatty acid composition. Although free fatty acid or its methyl esters are often used as internal standards for such analysis, it is preferable to add a triacylglycerol internal standard (i.e., triheptadecanoin) before transmethylation because this can minimize errors brought about by variation in the efficiency of methylation.

A time course of the transmethylation reaction indicated that for Arabidopsis intact seeds the reaction was essentially complete after 1 h (data not shown) but to be conservative we routinely transmethylate for 1.5 h. Because an internal triacylglycerol standard is used, this procedure can be further simplified by a single hexane extraction. The calculated recovery of oil and each of its components are the same for single and triplicate extractions (Table 2). With a single extraction the entire procedure will require only approximately 3 h to prepare 100 samples for GC analysis. Thus, the method is attractive for medium to high-throughput screening projects.

2.3. Plant-to-plant variation

We also tested the extent of plant-to-plant variation in oil content. *Arabidopsis* wild-type plants were cultivated under the same conditions, and pots were regularly rotated to different positions in the growth chamber. When mature seeds were harvested from each individual plant, the seed oil content was found to range from 33–40% [36.2 ± 1.8 (mean ± s.d., $n = 18$)] (Fig. 1). This falls in the range (28–40%) reported in the literature for Columbia wild-type seeds (O'Neill et al., 2003). Although most plants produce near the average oil content (36%), a number of "outliers" exist which give much higher (e.g. >40%) or lower (e.g. 30%) oil content. As shown in Fig. 1, the distribution of values can be attributed to biological variation between plants because the analytical variation of the method is considerably lower. Because Arabidopsis is self-fertilizing, the plants are expected to be essentially isogenic and the biological variation is likely due to responses of the individual plants to their environment (e.g., soil or water conditions, despite regular pot rotation).

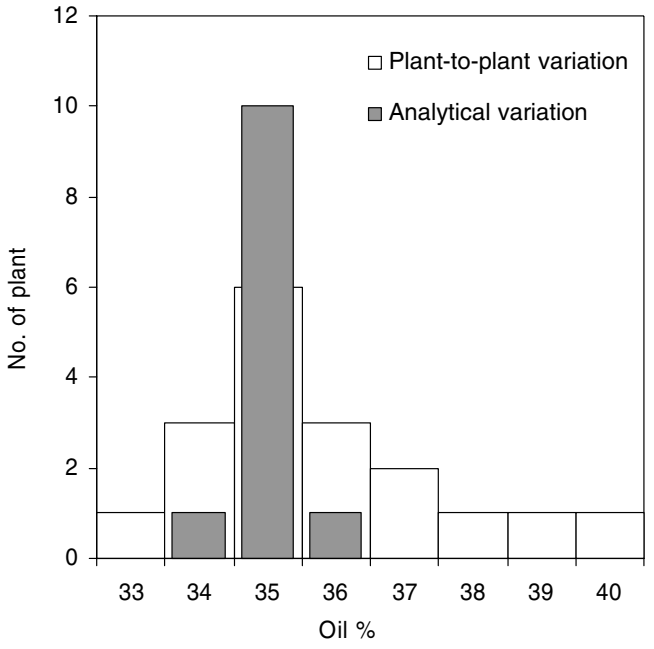


Fig. 1. Distribution of seed oil content in *Arabidopsis* wild-type plants shown together with analytical variation. Seeds from individual plants were collected and oil content determined by direct methylation method (mean = 36.18, s.d. = 1.79, $n = 18$). Analytical variation represents the replicate analysis of seeds from a single plant (mean = 35.88, s.d. = 0.62, $n = 12$).

This variation is important in the context of screening for higher/lower oil content phenotypes because it can easily lead to identification of false positive or false-negative lines. Detection of genes that influence seed oil content is a goal of many laboratories and even small increases (e.g. 5–10%) can provide substantial increased value to oilseed crops. In view of a medium- or high-throughput screening of the seed oil content of transgenic or mutant Arabidopsis populations, it is important to estimate the minimum number of transgenic or mutant plants that can be used to ensure reasonable rates of false positives and also of false negatives. In order to provide an example, Table 3 presents the influence of sample size on the statistical significance of the measurement of a 10% change in oil content. Simple assumptions were made. We considered the case of a transgenic/mutant population with a normal distribution like the WT population (Fig. 1). In this case, it can be calculated that analysis of at least six plants is needed to ensure that a 10% increase in oil content measured in the sample of transgenics/mutants is significant at $p < 0.05$, i.e., that the mean oil content of the WT and transgenic/mutant populations are different. The 95% CI for the actual increase in oil content in the transgenic population lies between 3.6% and 16.4%. Furthermore, if the goal is to be 95% confident that a 10% increase detected in a sample corresponds to at least a 5% increase in the oil content of the population, a higher number of plants is needed, namely 11. Thus, the plant-to-plant variation (Fig. 1) and Table 3 emphasize that reliable detection of small differences in oil content between lines requires analysis of more than a few plants within each line.

Table 3
Significance levels and confidence intervals (CIs) for a 10% increase in oil content measured in a sample of transgenics of variable size[a]

Transgenics sample size ($=n_2$)	Significance level α ($\beta = 0.2$)	Increase in oil in the population (%) (95% CI)
4	0.086	2.4–17.6
5	0.054	3.1–16.9
6	0.034	3.6–16.4
7	0.022	4.1–16.0
8	0.014	4.4–15.6
9	0.009	4.6–15.4
10	0.006	4.9–15.1
11	0.004	5.1–15.0
18	2.9×10^{-4}	5.9–14.1

[a] The WT sample described in Fig. 1 (mean $x_1 = 36.18$, size $n_1 = 18$, standard deviation $s_1 = 1.79$) was compared to a theoretical sample of transgenics normally distributed (mean $x_2 > x_1$ size n_2, standard deviation s_2) using an upper single-tail t-test. Given in the table is the type I error rate (α) for a type II error rate (β) equal to 0.2 at various sample sizes n_2. The type I error rate is the rate of false positives (transgenic populations falsely considered to show an increase in oil%) and the type II error rate is the rate of false negatives (transgenic populations with a real increase in oil% that are missed by the test). The sample means for the transgenics is $x_2 = 39.8$ for a 10% increase, and s_2 is supposed to be 50% higher than s_1. A 95% CI, within which the % increase in the mean oil content of the transgenic population is expected to lie, was also calculated for each case.

2.4. Oil content determination at two scales

In this study, we compared the standard error (s.e.) of results both on a 10-mg scale and on a 50-seed basis (~1 mg seeds). We considered that 10 mg is the lower limit for accurate weight determination on commonly available analytical balances. It is also often of interest to determine the weight of oil per seed. Furthermore, in many projects such as mutant screening where small quantities of seeds are available, or functional genomics where many tests of the same seeds are conducted, a smaller scale procedure is advantageous. Therefore, we also analyzed a 50-seed scale.

Seeds were first dried to constant moisture content in a desiccator, and then 10 mg samples were weighed on an analytical balance. An oil content of $35.9 \pm 0.2\%$ (mean $\pm$ s.e., $n = 6$) was determined from FAME analysis by GC (Table 4). On the 50-seed scale a value of 6.0 ± 0.3 μg FAME/seed ($n = 6$) was obtained. Converting this value using the determined batch seed weight as 17.8 ± 0.5 μg ($n = 6$) within the range of 17–21 μg reported for Arabidopsis seeds (e.g., Alonso-Blanco et al., 1999; Katavic et al., 1995) gave an oil content of $33.7 \pm 1.7\%$. The standard error of the 50-seed analysis is 4.2% of the mean compared to 0.5% for the weighed sample. It is likely that this larger variation represents heterogeneity of seed sizes in the samples.

Gravimetric analysis for oil content of the seeds from the same seed lot resulted in a value of $41.1 \pm 2.1\%$ (mean $\pm$ s.e., $n = 6$, at 10 mg scale). This value is slightly higher than that determined via direct methylation due mainly to inclusion of non-saponifiable lipids or other materials in the weighed lipid extract. Although the result is comparable to that obtained from direct methylation, the standard deviation of gravimetric analysis was 5.1% of the mean, which is substantially higher than direct methylation (0.5% of mean). Thus, better reproducibility is obtained from analyses based on FAME analysis of 10 mg samples than by analysis per seed or by gravimetric analysis.

2.5. Carbon and nitrogen analysis of seeds

The major seed storage components in Arabidopsis are TAGs (~35% by weight) and storage proteins (~30% by

Table 4
Comparison of precision of methods tested in this study

Methods	Oil content (mean)	s.e.	n	s.e. as % of mean
Direct methylation (10-mg) (%)	35.88	0.18	6	0.5%
Direct methylation (50-seed) (μg oil/seed)	5.95	0.25	6	4.2%
Gravimetric (%)	41.11	2.10	6	5.1%
	C:N ratio			
Elemental analysis	11.99	0.04	6	0.3%

weight) (Baud et al., 2002). Based on this composition and the molecular formulas for seed storage proteins, TAGs and carbohydrates, it can be estimated that more than 90% of seed nitrogen is in protein and over 50% of seed carbon resides in oil (see Section 4). Therefore, we hypothesized that alterations in the accumulation of these major storage components would be reflected in the C and N composition of seeds as determined by elemental analysis. In typical analysis of wild-type mature seeds, the elemental composition of C was 58.0 ± 3.06% (mean ± s.e., $n = 6$), N was 4.84 ± 0.24% (mean ± s.e., $n = 6$) and the C:N ratio was 11.99 ± 0.11 (mean ± s.e., $n = 6$). With an increasing proportion of seed oil, a positive correlation between oil content and %C is expected, as well as a decrease in %N in seeds. Indeed, when C:N ratio was plotted against the oil % for seeds that ranged from 35% to 45% oil, a linear relationship was obtained ($R = 0.89$). This analysis demonstrates that C and N analysis of seeds is an alternative method for estimating relative oil content. It has several advantages including providing information on relations between oil and protein levels, small sample size (2–4 mg), low cost per sample and high reproducibility (Table 4, s.e. is 0.3% of mean). All these make it an attractive alternative as a first-stage screen for mutants with higher or lower oil or protein content in seeds. Nonetheless, two limitations are that C and N contents do not provide data on fatty acid compositions and are not direct measures of oil and protein, and therefore could also reflect altered proportions of other seed components such as seed coat (see below).

2.6. Light conditions during plant growth strongly influence seed size, seed yield and oil content of seeds

Developing seeds of Arabidopsis and most members of the Brassicaceae family are green and several studies

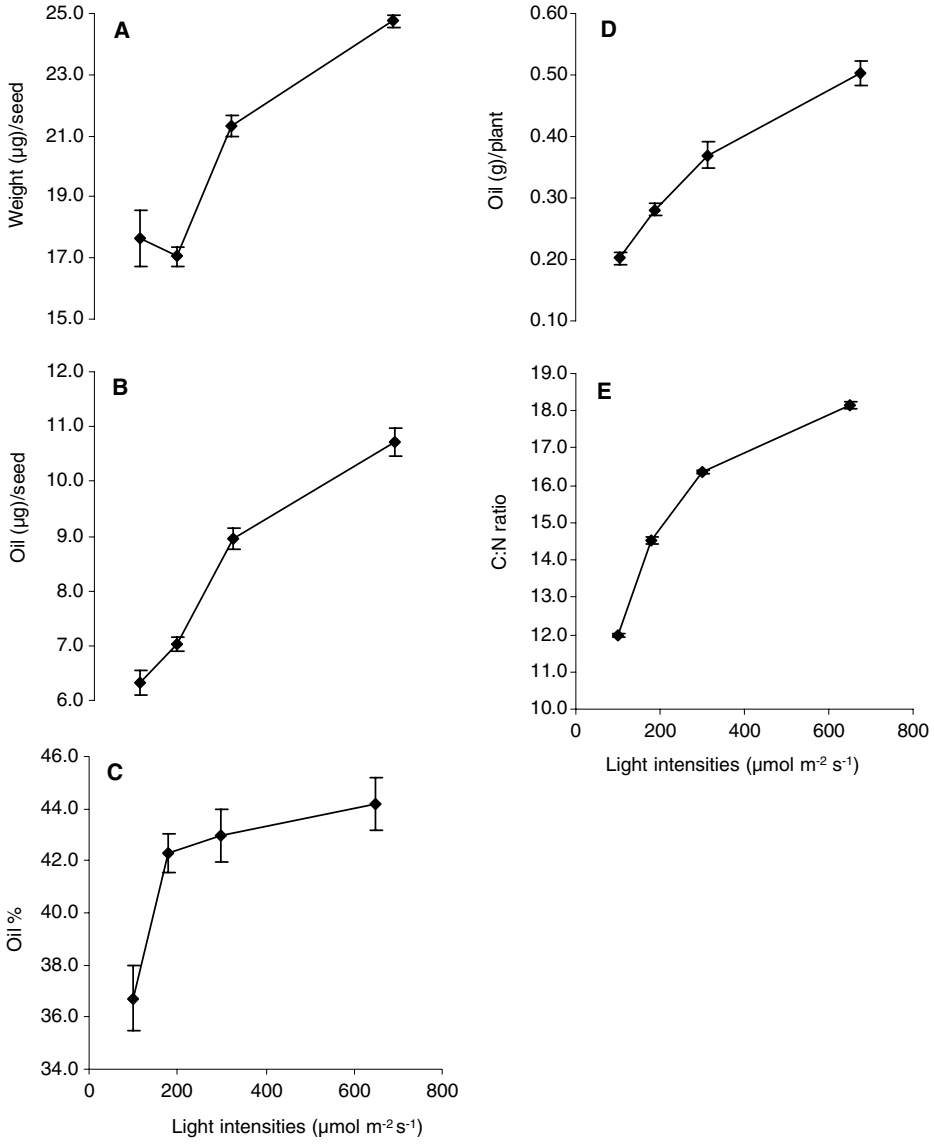
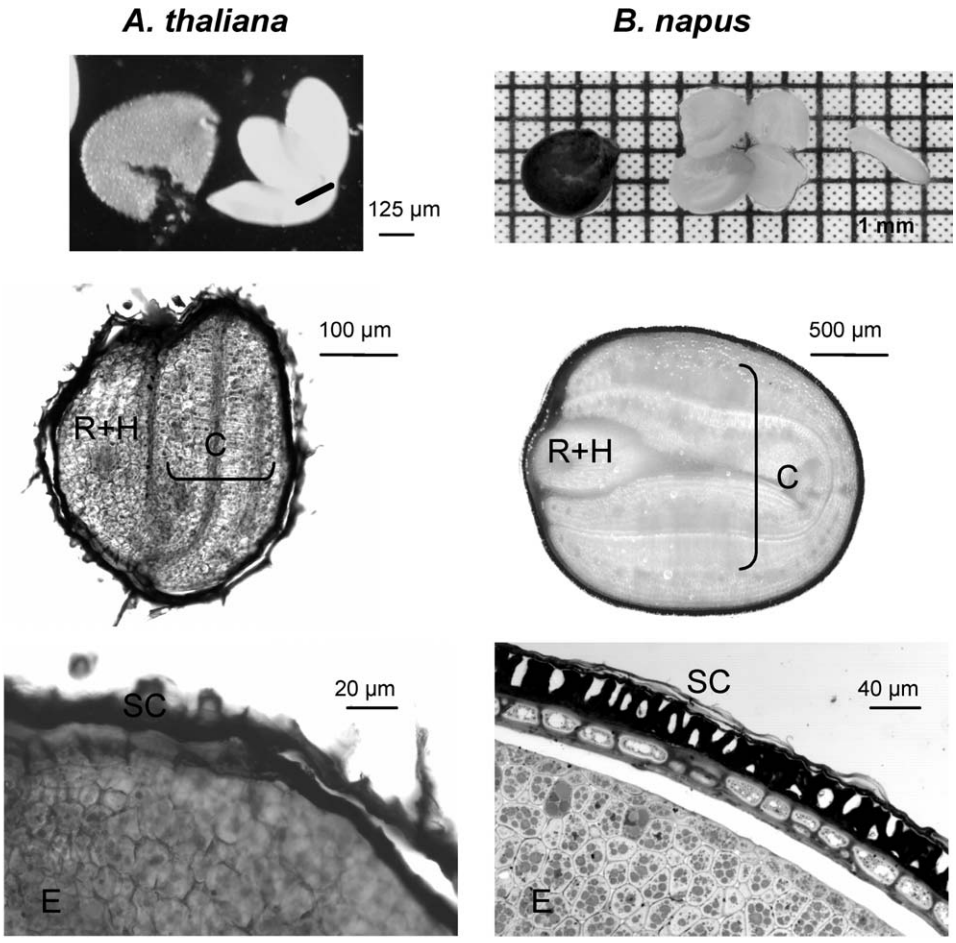


Fig. 2. Influence of light intensities (100, 180, 300 and 650 µmol m^{-2} s^{-1}) on Arabidopsis seed weight (A), oil per seed (B), oil% (C), total oil produced per plant (D) and seed C:N ratio (E). Error bars are s.e. ($n = 4$).



	Weight (μg)	Length (μm)	Width (μm)
A. thaliana	17.8 ± 0.5	485.2 ± 12.6	270.4 ± 8.7
B. napus	4160.1 ± 80.2	2083.6 ± 47.2	1907.1 ± 37.2

Fig. 3. Size comparison between *B. napus* and Arabidopsis mature seeds. Left panel shows dissection of Arabidopsis seed parts (top), cross section of Arabidopsis whole seed (middle) and seed coat (bottom); and right panel shows dissection of *B. napus* seed parts (top), cross section of *B. napus* whole seed (middle) and seed coat (bottom), R+H: radicle + hypocotyl, C: cotyledon, SC: seed coat, E: embryo. The table indicates measured dimensions of Arabidopsis and *B. napus* mature seeds (±s.e., $n = 50$).

(e.g., Eastmond et al., 1996; Eastmond and Rawsthorne, 1998) have demonstrated the presence of photosynthetic activity in these green seeds. The expression of genes that are involved in photosynthesis has also been found to be maximal during the period of oil synthesis in Arabidopsis (Ruuska et al., 2004). Recent studies from our lab (Goffman et al., 2005; Ruuska et al., 2004; Schwender et al., 2004) have confirmed the importance of light passing through the silique wall on green oilseed metabolism. Therefore, one potential factor influencing seed oil content might be light intensities used to grow Arabidopsis plants in different laboratories. Furthermore, when we monitored the light intensities at various areas within a growth chamber, it varied up to 25% (data not shown). To test the extent to which light intensity alters seed oil content, Arabidopsis wild-type plants were grown in a high light chamber and different light intensities (100, 180, 300 and 650 μmol m^{-2} s^{-1}) were produced by shading the plants with up to 15 layers of cheese cloth. These values compare to 1000–2000 μmol m^{-2} s^{-1} that occur outdoors on a sunny day.

At all light conditions plants grew normally, except slightly purple rosette leaves were observed with plants grown with 650 μmol m^{-2} s^{-1} light. After reaching maturation, seed yield, size, and oil content were analyzed (Fig. 2A–E). A positive correlation was found with all the parameters studied. We found a 40% increase in weight per seed when the light intensity was increased from 100 to 650 μmol m^{-2} s^{-1} (Fig. 2A). Visual inspection and measurements of the dimensions of seeds grown under different light levels indicated that seeds from higher light tend to be rounder, whereas seeds from lower light have a more

oblong shape. From the seed dimensions an increase in seed volume of 50% occurs at 650 μmol m^{-2} s^{-1} compared to 100 μmol m^{-2} s^{-1} light, which is a greater change than the increase in seed weight (40%: from 18 to 25 μg). This is expected because the increase in seed volume is largely contributed by an increase in seed oil (Fig. 2B), which has a lower density than other components (e.g., starch and protein). This again supports our finding that seed oil content increases with higher light.

In the same samples, oil per seed increased more than 60% with increasing light, from 6.5 to 10.96 μg per seed. We also observed a 20% increase in oil content (% by weight). Oil yield (Fig. 2D), the product of seed yield and oil content for each Arabidopsis plant, is at least two-fold higher when grown at 650 μmol m^{-2} s^{-1} than under lower light (100 μmol m^{-2} s^{-1}). The positive correlation between oil content and light intensities was also confirmed by elemental analysis which demonstrated an increased C:N ratio with increasing light intensities (Fig. 2E). Furthermore, both the increase in % oil by weight (Fig. 2C) and the increase in C:N ratio indicate that light preferentially enhances oil over protein storage, in agreement with recent results observed with *B. napus* (Goffman et al., 2005). Fatty acid composition, however, remained the same under the different light conditions (data not shown). Of course the influence of light on the seed parameters presented in Fig. 2 can be attributed to both the light effect on maternal plant growth as well as light stimulation of metabolism of the seed within the silique (Goffman et al., 2005).

2.7. Comparison of Arabidopsis and B. napus seed tissues

As mentioned in Section 1, Arabidopsis has been widely used as a model for oilseed crop species. As described below, differences in seed anatomy, oil content distribution and fatty acid composition in each tissue of Arabidopsis and *B. napus* seeds can influence the comparison of results between these two species.

As shown in Fig. 3, the Arabidopsis seed is ~6 times smaller in diameter compared to *B. napus* seeds, but in terms of mass Arabidopsis seeds (15–20 μg) are approximately 200 times lighter than *B. napus* seeds (4.2 mg). Also shown in Fig. 3, the different seed sizes are accompanied by different proportions of the tissues within each seed. The seed coat + endosperm of Arabidopsis is ~20 μm thick compared to ~50 μm for *B. napus* seeds. If we model Arabidopsis seed as an ellipsoid, with a length of 485 μm and a radius of 188 μm (this study and Debeaujon et al., 2000), the seed coat (plus endosperm) will comprise 30% of total seed volume. If we apply the same model to *B. napus* seeds (length = 2 mm; radius = 1 mm; seed coat = 50 μm), the seed coat + endosperm would represent less than 15% of total volume. Therefore, the percentage of seed volume represented by the seed coat + endosperm is approximately two-fold larger in Arabidopsis. This ratio was confirmed experimentally on a weight basis. As indicated in Fig. 4A for Arabidopsis the seed coat + endosperm comprise 34% of total seed weight, yet only half of this amount (17%) for *B. napus* seed. Likewise, the proportion of radicle and hypocotyl of Arabidopsis is substantially larger than that of *B. napus* seeds. On a weight basis, the proportions of cotyledons, (radicle + hypocotyls), and (seed coat + endosperm) for Arabidopsis seed were determined to be 45:21:34 and for *B. napus* seed were 74:9:17 (Fig. 4A).

Whole mature seeds of *B. napus* and Arabidopsis were dissected and analyzed for oil content and FA composition. As shown in Fig. 4B, in *B. napus* seeds, the cotyledons contain 90% of total seed fatty acids, with 6% in radi-

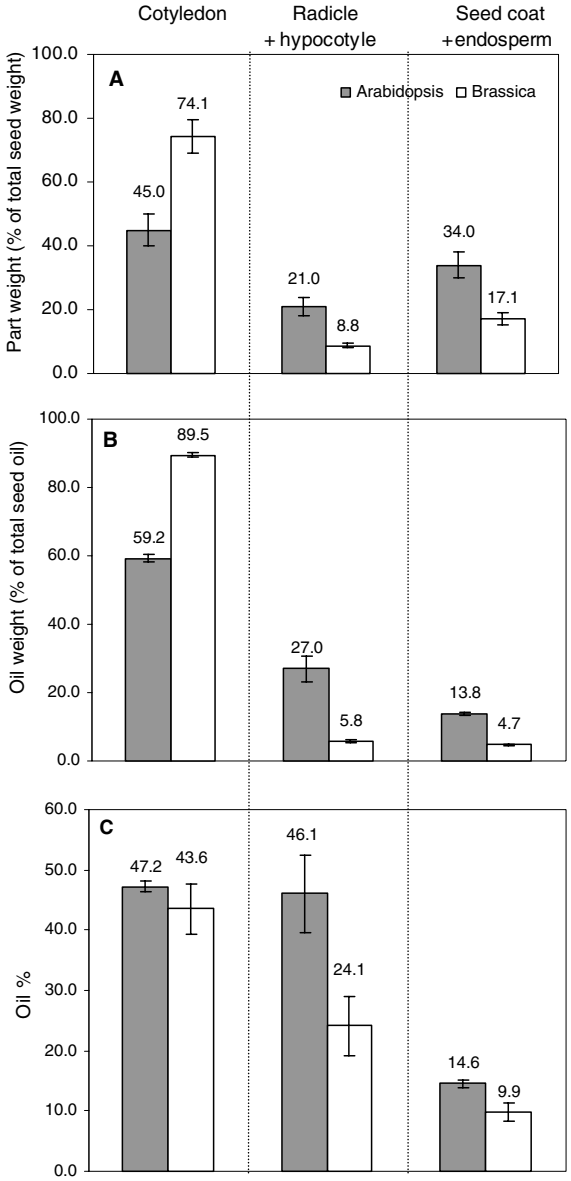


Fig. 4. Comparison of Arabidopsis and *B. napus* mature seeds in terms of: (A) proportion of total seed weight contributed by seed parts; (B) proportion of total seed oil contributed by seed parts; (C) oil% by weight in each seed part. Error bars are s.e. ($n = 3$).

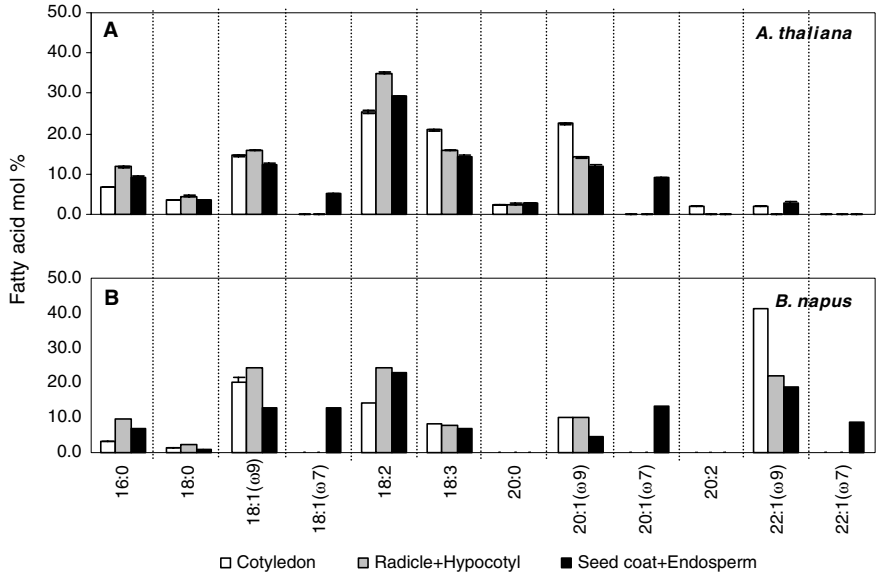


Fig. 5. Fatty acid composition in each seed part of Arabidopsis (A) and *B. napus* (B). Bars = s.e. ($n = 3$).

cle + hypocotyl and 4% in seed coat + endosperm. This contrasts substantially to that of Arabidopsis seeds, in which the cotyledons contain 59% of total fatty acids, with 27% in radicle + hypocotyl and 14% in seed coat + endosperm. When calculated on a % of weight basis (Fig. 4C), Arabidopsis and *B. napus* have a similar content of oil in cotyledons (47% and 44%, respectively) but Arabidopsis has nearly two-fold higher oil content (46%) in the hypocotyl + radical than its counterpart in *B. napus* seeds (24%). There also is a higher amount of oil in Arabidopsis seed coat + endosperm (15%) than for *B. napus* (10%). These differences in distribution indicate that some results obtained from Arabidopsis may not apply quantitatively to *B. napus*. For example, a transgene that reduces seed coat thickness, or increases embryo size would be expected to have a larger impact on oil content in Arabidopsis than in *B. napus*.

Penfield et al. (2004) recently reported that lipids deposited in endosperm of Arabidopsis seeds have a different fatty acid composition from that of cotyledons. As indicated in Fig. 5A, we have confirmed and extended these observations. The most striking characteristic of the seed coat + endosperm lipids is the presence of 18:1 and 20:1 ω-7 fatty acids which together constitute 15 mol% of total fatty acids but are very minor in the other seed structures. A similar composition occurs in *B. napus* seed coat + endosperm (Fig. 5B), where ω-7 fatty acids (18:1, 20:1 and 22:1) represent 35 mol% of the total fatty acid. We also observed that the oil content and fatty acid profile are different between cotyledons and radicle + hypocotyl of both Arabidopsis and *B. napus*. In Arabidopsis seeds, the radicle + hypocotyl has a similar amount of oil compared to cotyledons, but has a higher proportion of 16:0, 18:2 and less 18:3 and 20:1. In *B. napus* seeds, the radicle + hypocotyl has a lower oil content than cotyledons, but, similar to Arabidopsis, a higher proportion of 16:0, 18:1, 18:2 and less 22:1. Analogous to the comments above about seed coat + endosperm, the much larger contribution of the radicle + hypocotyl to Arabidopsis seed composition means that selecting for Arabidopsis mutants/transgenics having different fatty acid composition or quantity may result in identifying traits that will be quantitatively smaller in *B. napus*.

3. Conclusions

In light of the frequent use of Arabidopsis as a model oilseed and also the substantial variation in published data on Arabidopsis seed oil content, this paper evaluated the accuracy of quantifying oil content of Arabidopsis seeds and factors that influence the results. The direct methylation procedure described is rapid and the amount of seed material required is minimal. Variability in seed oil content between plants was found to be substantial and is a factor that can limit the detection of transgenic or mutant lines with altered oil content. We also show that elemental analysis provides a simple and highly reproducible alternative for estimating seed oil content and may be of particular value for labs without convenient access to GC equipment.

The procedures described were applied to plants grown under different light levels and strong positive correlations between Arabidopsis seed size and oil content and light intensities were observed. Finally, the observations provided in this study also establish a reference data set on the distribution of mass, oil and fatty acid composition between tissues of Arabidopsis and *B. napus* seeds that should be useful to help understand or predict the applicability of results obtained with Arabidopsis to other oilseeds.

4. Experimental

4.1. Plant material

Seeds of *A. thaliana* ecotype Columbia (Col 0) were sown onto planting soil (1:1:1 mixture of peat moss-enriched soil:vermiculite:perlite) and cold treated (at 4 °C) for 3–4 days before being transferred to a controlled growth chamber. Plants were watered once a week with nutrient solution (5 μM KNO_3, 2.5 μM KPO_4 (pH 5.5), 2.0 μM $MgSO_4$, 2.0 μM $Ca(NO_3)_2$, 0.05 μM FeEDTA, 35 nM boric acid, 7 nM $MnCl_2$, 25 nM $CuSO_4$, 0.5 nM $ZnSO_4$) and unless otherwise stated grown for 7–8 weeks at 22 °C (16 h photoperiod) with a light intensity of 100 μmol m^{-2} s^{-1} provided by fluorescent bulbs. After all siliques were formed, watering was stopped and plants were transferred to constant light to dry the seeds. Mature, dry seeds were harvested and cleaned by passing through a 425 μm sieve (Test sieve, W.S. Tyler, USA) several times to remove as much debris as possible. Because seed lipid content decreases ~10% in the last stages of maturation (Baud et al., 2002; Chia et al., 2005) it is important to harvest seeds in a consistent manner from fully mature siliques.

For growth under different light intensities, plants were grown in a high light chamber under 1–15 layers of cheesecloth to produce different light intensities (600, 300, 180 and 100 μmol m^{-2} s^{-1}), which were monitored using a Quantum Meter Model QMSS (Apogee Instruments Inc., USA).

Brassica napus v. Reston were grown in an air-conditioned green house under natural light supplemented with lamps to provide l6/8-h photoperiod and an approximate light intensity of 800 μmol m^{-2} s^{-1}.

4.2. Seed weight determination

Arabidopsis mature seeds (300–500 seeds per replicate) were counted under a dissecting microscope. The samples were then weighed carefully on a Sartorius CP225D analytical balance to ±0.02 mg.

Because Arabidopsis seeds are tiny it can be difficult to cleanly separate the seeds from dried plant debris, and any such contamination will introduce error in the final oil content determination. We therefore used a hand sieve with a mesh size of 425 μm to remove debris with seeds passing through the mesh and collected on clean paper. After sieving, the seeds are still frequently mixed with small amounts of residue. A combination of additional sieving, gentle blowing and visual inspection are employed to clean the seeds completely.

4.3. Stabilizing seed water content

Typically, mature Arabidopsis seeds at the time of harvest have water content of 8–10% (Baud et al., 2002), but this content will vary depending on the atmospheric humidity and length of drying time. It is important to minimize and stabilize seed water content prior to methylation for two reasons: (1) if oil content is to be determined on a weight basis, it is important that all the seed samples analyzed have essentially the same water content and; (2) water can impair methylation and it is important to minimize total water content in the methylation reaction to less than 10% of reagent volume (Lepage and Roy, 1984).

Although oven drying (e.g., at 100–110 °C) can be used, it proved less suitable because it destroys the seeds tested and the seeds can gain water more rapidly when dried to very low water content. In this study, we found that by storing the seeds in a desiccator (with desiccant) seeds reach constant water content ($2.72 \pm 0.23\%$; $n = 5$) within 24–48 h (data not shown). Within 1 h after being removed from the dessicator, dried seeds start to absorb moisture from the atmosphere (at a humidity of 30% and 20 °C). At higher humidity it is expected that water content would increase more rapidly. Therefore, it is important that samples are weighed immediately after removal from the desiccator. This procedure ensures that all samples analyzed contain equal amount of moisture content, and accurate comparisons can be made.

4.4. Determining seed water content

Pre- weighed Arabidopsis seeds (ca.100 mg per replicate) were heated for 18 h at 110 °C. The samples were weighed again after cooling down in a desiccator with anhydrous calcium sulfate as desiccant. Final water content was determined as the difference between the two measurements divided by the initial sample weight.

4.5. Oil content determination via direct methylation

Dried seeds (~10 mg) were weighed on an analytical balance, or 50 seeds were counted carefully under a dissecting microscope. A glass tube (1 cm × 10 cm) with Teflon-lined screw cap was pre-rinsed thoroughly with chloroform and dried to remove any contaminating lipid residues. To this tube was added 1 ml of 5% (v/v) conc. sulfuric acid in MeOH (freshly prepared for each use), 25 μl of BHT solution (0.2% butylated hydroxy toluene in MeOH), 10–100 μg of triheptadecanoin (as a triacylglycerol internal standard to generate methyl heptadecanoate) and 300 μl of toluene as co-solvent. The mixture was vortexed for 30 s then heated at 90–95 °C for 1.5 h. After cooling to room temperature, 1.5 ml of 0.9% NaCl (w/v) was added and FAMEs were extracted with 3 × 2 ml hexane. Pooled extracts were evaporated under nitrogen and then dissolved in 400 μl of the hexane. The FAME extracts were analyzed by GC with a flame ionization detector (FID) on a DB23 column (30 m by 0.25 mm i.d., 0.25 μm film; J&W Scientific, Folsom, CA). The GC conditions were: split mode injection (1:40), injector and flame ionization detector temperature, 260 °C; oven temperature program 150 °C for 3 min, then increasing at 10 °C/min to 240 °C and holding this temperature for

5 min. In preliminary experiments, a time course of the transmethylation reaction was generated using squalane and triheptadecanoin as internal standards.

4.6. Calculation of oil content from GC analysis

The area of each peak in the GC chromatogram was first corrected for the theoretical response factor of the FID in which peak area (pA.sec) is a function of mass of C atoms with at least one bound H atom, and therefore which differs for each fatty acid methyl ester (Christie, 1991). The corrected areas were used to calculate the mass of each FAME in the sample by comparison to the internal standard mass. Because oil is stored in seeds primarily in the form of triacylglycerol, an additional correction to convert FAME weight to TAG weight is needed. This correction will depend on the fatty acid composition of the oil. To determine the moles of TAG the moles of each FAME were calculated from the weight and the molecular weight. The oil content was then calculated from the formula: Percent oil by weight $= 100 \times ((4 \times \text{total mol FAME}/3) + \text{total g FAME})/\text{g tissue}$, where 4 is the M_r difference between TAG and three moles of FAME.

4.7. Seed lipid extraction and gravimetric lipid determination

The lipid extraction method of Hara and Radin (1978) was followed. Briefly, isopropanol (2 ml) was added to a known and exact weight of seeds and the mixture was heated at 85 °C for 10 min. Samples were then ground with a polytron. Hexane (3 ml) was added to the mixture,and after vortexing, was allowed to stand for 5 min at room temperature. To the mixture, aqueous sodium sulphate (2.5 ml of 15% wt/vol) was added to provide phase separation. The upper phase was removed and the lower aqueous phase re-extracted with 7/2 hexane/isopropanol. The combined lipid extracts were evaporated to constant weight under nitrogen. The oil content was calculated by dividing the initial seed weight by the weight of lipid extracted.

4.8. Elemental analysis

The composition of carbon (C) and nitrogen (N) in Arabidopsis seeds were determined at Duke Environmental Stable Isotope Laboratory, Department of Biology, Duke University using a CE Instruments NC 2100 elemental analyzer (ThermoQuest Italia, Milan). 2–4 mg of seeds were accurately weighed and combusted at 1200 °C in an elemental analyzer in the presence of chemical catalysts to produce CO_2 and N_2. Data from Baud et al. (2002) and typical molecular formulas for protein, i.e., the two most abundant storage proteins in Arabidopsis seeds (12S globulins $C_{2318}H_{3628}N_{686}O_{698}S_{10}$ and 2S albumins $C_{827}H_{1300}N_{236}O_{249}S_{15}$ although different in their molecular weight, however, these two molecules have very similar C% and N%), for lipid (e.g., $C_{57}H_{104}O_6$ for triolein) and for carbohydrate (e.g., $C_6H_{12}O_6$, for glucose) allowed us to estimate that more than 90% of Arabidopsis seed nitrogen is in protein and over 50% of seed carbon resides in oil.

4.9. Determination of oil and weight distribution pattern in mature seed tissue

Mature Arabidopsis seeds were imbibed overnight in distilled water at 4 °C in the dark. Imbibed seeds were then pressed gently between two glass plates. Each part (cotyledons, radicle + hypocotyl and seed coat + endosperm) was dissected under a light microscope and it was picked up very carefully to avoid contamination from the other tissues. Twenty cotyledons, 20 radicles + hypocotyls and 200 seed coats + endosperms per replicate were collected for FAME analysis (whole seed transmethylation). For weight determination, 100 seed coats + endosperms, 50 embryos and 50 radicles + hypocotyls per replicate were carefully collected under a dissecting microscope, dried and weighed. Seeds analyzed before and after 16 h imbibition at 4 °C indicated no significant TAG breakdown or mobilization during this period.

Brassica napus v. Reston mature seeds were imbibed for 1 h in distilled water in the dark at 4 °C before being dissected. Each part (cotyledons, radicle + hypocotyl and seed coat + endosperm) was separated carefully by pressing the seeds gently. Seed coat and the rudimentary endosperm tissue were collected together since it is impractical to separate one from the other. Five cotyledon pairs, five radicle + hypocotyls, and 20 seed coats + endosperms per replicate were ground and subjected to FAME analysis. For weight determination, 10 mature dry seeds were dissected and the weight of each part was determined on an analytical balance.

4.10. Statistical analysis

The mean and the standard deviation of the seed oil% measured in 18 individual *Arabidopsis* wild-type plants (reported in Fig. 1) was 36.18 and 1.79, respectively. The entire data set was first tested for normality using PROC UNIVARIATE procedure in SAS (SAS Institute). The normality was well held in all tests. In addition, the distribution had no significant skewness or kurtosis problem.

The statistical significance of an increase in oil content for a sample of the transgenics was estimated assuming normality of the transgenic population. The null hypothesis is H_0: $\mu_2 \leqslant \mu_1$, and the alternative hypothesis is H_1: $\mu_2 > \mu_1$ where μ_2 and μ_1 are the mean of the transgenic and WT populations, respectively. An upper one-tail t statistics was calculated for various values of n_2 with a t-test for unequal variances (the probability of equality of variances was estimated by an f-test in each case). The power retained in the t-test was set to 0.8. In addition to this hypothesis testing approach, a Wald's confidence interval (CI) for the difference between the means was also calculated in each case. CIs were expressed as a % increase of the mean oil% of the transgenic population. All the

statistics, powers and confidence intervals were computed using Simple Interactive Statistical Analysis (http://home.clara.net/sisa/) and Power Calculator (http://calculators.stat.ucla.edu/powercalc/).

4.11. Measurement of seed size

Mature seeds of Arabidopsis and *B. napus* were imbibed in water overnight at 4 °C, fixed for 4 days in 3.7% formaldehyde in phosphate buffer, and then washed in phosphate buffer to remove excess fixative. After the washed seeds were dehydrated through a graded series of ethanol and infiltrated in LR-White low viscosity resin, the seeds were embedded in gelatin capsules and polymerized for 2 days at 60 °C. Sections of 1 μm thick were cut with an ultramicrotome (Power Tome XL, RMC, Boeckeler Instruments, Tucson, AZ). Sections were stained with Toluidine blue (1%) for 2 min at 50 °C, and pictures taken. If an Arabidopsis seed is modeled as as ellipsoid, the volume (V) is calculated based on formula $V = 1/6 \times \pi \times a \times b \times c$ where a, b and c refer to length, width and height, whereas Brassica seeds are round, therefore V is calculated based on formula $V = 4/3 \times \pi \times r^3$ where r is the radius.

Acknowledgements

This research was supported by funding from The Dow Chemical Company and Dow AgroSciences.

References

Alonso-Blanco, C., Blankestijn-de Vries, H., Hanhart, C.J., Koornneef, M., 1999. Natural allelic variation at seed size loci in relation to other life history traits of *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. USA 96, 4710–4717.

Baud, S., Boutin, J.P., Miquel, M., Lepiniec, L., Rochat, C., 2002. An integrated overview of seed development in *Arabidopsis thaliana* ecotype WS. Plant Physiol. Biochem. 40, 151–160.

Baud, S., Guyon, V., Kronenberger, J., Wuilleme, S., Miquel, M., Caboche, M., Lepiniec, L., Rochat, C., 2003. Multifunctional acetyl-CoA carboxylase 1 is essential for very long chain fatty acid elongation and embryo development in Arabidopsis. Plant J. 33, 75–86.

Browse, J., McCourt, P.J., Somerville, C.R., 1986. Fatty acid composition of leaf lipids determined after combined digestion and fatty acid methyl ester formation from fresh tissue. Anal. Biochem. 152, 141–145.

Cavell, A.C., Lydiate, D.J., Parkin, I.A., Dean, C., Trick, M., 1998. Collinearity between a 30-centimorgan segment of *Arabidopsis thaliana* chromosome 4 and duplicated regions within the *Brassica napus* genome. Genome 41, 62–69.

Cernac, A., Benning, C., 2004. WRINKLED1 encodes an AP2/EREB domain protein involved in the control of storage compound biosynthesis in Arabidopsis. Plant J. 40, 575–585.

Chia, T.Y., Pike, M.J., Rawsthorne, S., 2005. Storage oil breakdown, during embryo development of *Brassica napus* (L.). J. Exp. Bot. 56, 1285–1296.

Christie, W.W., 1991. Gas chromatographic analysis of fatty acid methyl esters with high precision. Lipid Technol. 3, 97–98.

Debeaujon, L., Leon-Kloosterziel, K.M., Koornneef, M., 2000. Influence of the testa on seed dormancy, germination, and longevity in Arabidopsis. Plant Physiol. 122, 403–414.

Eastmond, P.J., Rawsthorne, S., 1998. Comparison of the metabolic properties of plastids isolated from developing leaves or embryos of *Brcasica napus* L.. J. Exp. Bot. 49, 1105–1111.

Eastmond, P., Kolacna, L., Rawsthorne, S., 1996. Photosynthesis by developing embryos of oilseed rape (*Brasscia napus* L.). J. Exp. Bot. 47, 1763–1769.

Epp, M.D., Pollard, M.R., 1993. Fatty acid content and composition of *Oenothera hookeri* seeds containing mutant plastids. Lipids 28, 1005–1009.

Focks, N., Benning, C., 1998. Wrinkledl: a novel, low-seed-oil mutant of Arabidopsis with a deficiency in the seed-specific regulation of carbohydrate metabolism. Plant Physiol. 118, 91–101.

Garces, R., Mancha, M., 1993. One-step lipid extraction, and fatty acid methyl esters preparation from fresh plant tissues. Anal. Biochem. 211, 139–143.

Goffman, F.D., Alonso, A.P., Schwender, J., Shachar-Hill, Y., Ohlrogge, J., 2005. Light enables a very high efficiency of carbon storage in developing embryos of rapeseed. Plant Physiol. 138, 2269–2279.

Hara, A., Radin, N.S., 1978. Lipid extraction of tissues with a low-toxicity solvent. Anal. Biochem. 90, 420–426.

Hobbs, D.H., Flintham, J.E., Hills, M.J., 2004. Genetic control of storage oil synthesis in seeds of Arabidopsis. Plant Physiol. 136, 3341–3349.

Jako, C., Kumar, A., Wei, Y., Zou, J., Barton, D.L., Giblin, E.M., Covello, P.S., Taylor, D.C., 2001. Seed-specific over-expression of an Arabidopsis cDNA encoding a diacylglycerol acyltransferase enhances seed oil content and seed weight. Plant Physiol. 126, 861–874.

Katavic, V., Reed, D.W., Taylor, D.C., Giblin, E.M., Barton, D.L., Zou, J., Mackenzie, S.L., Covello, P.S., Kunst, L., 1995. Alteration of seed fatty acid composition by an ethyl methanesulfonate-induced mutation in *Arabidopsis thaliana* affecting diacylglycerol acyltransferease activity. Plant Physiol. 108, 399–409.

Lepage, G., Roy, C.C., 1984. Improved recovery of fatty acid through direct transesterification without prior extraction or purification. J. Lipid Res. 25, 1391–1396.

Marillia, E.F., Micallef, B.J., Micallef, M., Weninger, A., Pedersen, K.K., Zou, J., Taylor, D.C., 2003. Biochemical and physiological studies of *Arabidopsis thaliana* transgenic lines with repressed expression of the mitochondrial pyruvate dehydrogenase kinase. J. Exp. Bot. 54, 259–270.

Meinke, D.W., 1994. Seed development in *Arabidopsis thaliana*. In: Meyerowitz, E.M., Somerville, C.R (Eds.), Arabidopsis. CSHL Press, New York, pp. 253–295.

Meyerowitz, E.M., Somerville, C.R., 1994. Arabidopsis. CSHL Press, New York.

Murphy, D.J., 1996. Engineering oil production in rapeseed and other oil crops. Trends Biotechnol. 14, 206–213.

O'Neill, C.M., Gill, S., Hobbs, D., Morgan, C., Bancroft, I., 2003. Natural variation for seed oil composition in *Arabidopsis thaliana*. Phytochemistry 64, 1077–1090.

Penfield, S., Rylott, E.L., Gilday, A.D., Graham, S., Larson, T.R., Graham, I.A., 2004. Reserve mobilization in the Arabidopsis endosperm fuels hypocotyl elongation in the dark, is independent of abscisic acid, and requires PHOSPHOENOLPYRUVATE CARBOXYKINASE1. Plant Cell 16, 2705–2718.

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 Physiol. 122, 1193–1199.

Rutar, V., 1989. Magic angle sample spinning NMR-spectroscopy of liquids as a nondestructive method for studies of plant seeds. J. Agric. Food Chem. 37, 67–70.

Ruuska, S.A., Schwender, J., Ohlrogge, J.B., 2004. The capacity of green oilseeds to utilize photosynthesis to drive biosynthetic processes. Plant Physiol. 136, 2700–2709.

Schwender, J., Goffman, F., Ohlrogge, J.B., Shachar-Hill, Y., 2004. Rubisco without the Calvin cycle improves the carbon efficiency of developing green seeds. Nature 432, 779–782.

Thelen, J.J., Ohlrogge, J.B., 2002. Both antisense and sense expression of biotin carboxyl carrier protein isoform 2 inactivates the plastid acetyl-coenzyme A carboxylase in *Arabidopsis thaliana*. Plant J. 32, 419–431.

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. Plant J. 9, 229–241.

Zou, J., Katavic, V., Giblin, E.M., Barton, D.L., Mackenzie, S.L., Keller, W.A., Hu, X., Taylor, D.C., 1997. Modification of seed oil content and acyl composition in the Brassicaceae by expression of a yeast sn-2 acyltransferase gene. Plant Cell 9, 909–923.

Zou, J., Wei, Y., Jako, C., Kumar, A., Selvaraj, G., Taylor, D.C., 1999. The *Arabidopsis thaliana* TAG1 mutant has a mutation in a diacylglycerol acyltransferase gene. Plant J. 19, 645–653.