Journal of Plant Physiology 208: 7–16 (2017)
http://dx.doi.org/10.1016/j.jplph.2016.10.011

Identification of target genes and processes involved in erucic acid accumulation during seed development in the biodiesel feedstock Pennycress (*Thlaspi arvense* L.)

Ana Claver[1], Raquel Rey[2], M. Victoria López[3], Rafael Picorel[1], and Miguel Alfonso[1],*.

[1]Department of Plant Nutrition, Estación Experimental de Aula Dei-CSIC, Avda. Montañana 1005, 50059 Zaragoza, Spain.

[2]Laboratorio Agroambiental, Gobierno de Aragón, Avda. Montañana 1005, 50071 Zaragoza, Spain.

[3]Department of Soil and Water, Estación Experimental de Aula Dei-CSIC, Avda. Montañana 1005, 50059 Zaragoza, Spain.

*Corresponding author.

E-mail address: alfonso@eead.csic.es (M. Alfonso). Phone: 00-34-976-716059; FAX: 00-34-976-716145.

ABSTRACT

We studied erucic acid accumulation in the biodiesel feedstock Pennycress (*Thlaspi arvense* L.) as a first step towards the development of a sustainable strategy for biofuel production in the EU territory. To that end, two imbred Pennycress lines of European origin, "NASC" and "French", were cultivated in a controlled chamber and in experimental field plots, and their growth, seed production and seed oil characteristics analyzed. Differences in some agronomical traits like vernalization (winter-French versus spring-NASC), flowering time (delayed in the French line) or seed production (higher in the French line) were detected. Both lines showed a high amount (35-39%) of erucic acid ($22:1^{\Delta 13}$) in their seed oil. Biochemical characterization of the Pennycress seed oil indicated that TAG was the major reservoir of $22:1^{\Delta 13}$. Incorporation of $22:1^{\Delta 13}$ to TAG occurred very early during seed maturation, concomitant with a decrease of desaturase activity. This change in the acyl fluxes towards elongation was controlled by different genes at different levels. *TaFAE1* gene, encoding the fatty acid elongase, seemed to be controlled at the transcriptional level with high expression at the early stages of seed development. On the contrary, the *TaFAD2* gene that encodes the Δ12 fatty acid desaturase or *TaDGAT1* that catalyzes TAG biosynthesis were controlled post-transcriptionally. *TaWRI1*, the master regulator of seed-oil biosynthesis, showed also high expression at the early stages of seed development. Our data identified genes and processes that might improve the biotechnological manipulation of Pennycress seeds for high-quality biodiesel production.

Keywords: *Thlaspi arvense*, erucic acid, seed oil, TAG (triacylglycerol), FAE1 (fatty acid elongase 1), biofuel.

1. Introduction

Field Pennycress (*Thlaspi arvense* L.) is a winter annual that belongs to the Brassicaceae family. As many other members of this family like Miscanthus (Robson et al., 2013), Etiopian mustard (*Brassica carinata*; Bouaid et al., 2009) or Camelina (*Camelina sativa*, Frohlich and Rice, 2005), Pennycress has attracted the attention of researchers as a promising alternative oilseed feedstock for biodiesel production, that accomplishes the land use and sustainability criteria and displaces other plant species like Jatropha (*Jatropha curcas*) or Crambe (*Crambe abyssinica*) that are not well adapted to the temperate climate conditions of Europe and North America (Moser, 2012). Native from Eurasia, Pennycress is widely distributed across temperate regions all around the world and is highly adapted to a wide variety of climatic conditions (Warwick et al., 2002; Vaughn et al., 2005; Moser et al., 2009a). Because of its growth cycle, it can be successfully planted at the end of the summer and germinates in the fall developing a low-growing rosette that protects the plant from low temperatures and cold winds during the winter. Pennycress is an extreme cold tolerant plant (Sharma et al., 2007). The plant resumes its growth in the spring, sets seeds and is harvested at the beginning of the summer. Because of its culture cycle it can be used in rotations, not displacing existing agricultural production (Moser, 2012). It does not need any agricultural inputs like fertilizers or pesticides and has no specific water requirements. Furthermore, Pennycress could be planted in lands not otherwise suited for agricultural production (Moser, 2012). Therefore, it does not compete with food chain cultures. Pennycress has been signalled for its potential to produce biomass for renewable biofuel production (Moser et al., 2009a,b; Moser, 2012), being a prolific seed producer (Fan et al., 2013). Harvested Pennycress seeds contain around 36% oil (w/w), which is twice the amount present in other oil commodities like soybean or sunflower and very similar to that found in camelina

(Moser, 2012). Because of its high oil content and fatty acid profile, with high amount of unsaturated fatty acids, particularly erucic acid ($22:1^{\Delta 13}$), Pennycress oil can be used for biodiesel and biojet production with excellent characteristics like high cetane number of 59.8 and low temperature properties (Moser et al., 2009a; Moser, 2012; Fan et al., 2013). These results indicate that Pennycress oil could qualify as a biomass-derived diesel according to the Renewable Fuels Standard (Fan et al., 2013). Despite of these interesting characteristics, Pennycress is still a wild plant that requires much research to evaluate its actual agronomical potential. In the USA, test plots from Illinois reported seed yields that varied from 900 to 2,352 kg ha^{-1} (Evangelista et al., 2012; Fan et al., 2013). In our first experimental campaign in Spain, productions around 1,300 kg ha^{-1} were obtained with European varieties. Several traits, including seed dormancy and size, erucic acid content, pot shatter and flowering time have been identified as potential targets of Pennycress breeding to improve its potential as a dedicated bioenergy crop (Sedbrook et al., 2014). To our knowledge, although it is well adapted to the European climatic conditions, and there is botanical register of its presence all through the EU, very few experiences on Pennycress cultivation have been performed, all of them using Pennycress seeds of American origin, and the agronomic conditions required for its successful cultivation have not been still established (Groeneveld and Klein, 2014, 2015).

In addition to the research efforts to be held at the agronomical level, understanding the biochemical and molecular pathways involved in oil biosynthesis in Pennycress seeds would be necessary to guide future crop improvement efforts directed towards an increase in oil yield and quality as well as other important agronomic traits (dormancy, vernalization, etc). At the molecular level, a transcriptome assembly of Pennycress genes has been reported that might provide tools for the breeding of Pennycress (Dorn et al., 2013). More recent studies

have performed a metabolite profiling of Pennycress seed embryos to determine the biochemical pathways active during oil synthesis (Tsogtbaatar et al., 2015). However, although the enzymes (fatty acid elongases and desaturases) and compounds (TAG, DAG) that participate in seed oil biosynthesis have been characterized in other plant species, mainly *Brassica napus* and *Arabidopsis thaliana* (von Wettstein-Knowles, 1982; James et al., 1995; Millar and Kunst, 1997; Katavic et al., 2001; Katavic et al., 2004), this information is still lacking for Pennycress seeds. This knowledge is necessary to avoid metabolic bottlenecks that usually appear during the manipulation of seed oil biosynthetic pathways (Cahoon et al., 2007). An example of these bottlenecks related with erucic acid manipulation was detected in genetically engineered *Crambe abyssinica* (Guan et al., 2014).

As a first step for the introduction of Pennycress as an alternative feedstock for biofuel production in the Mediterranean regions of the EU, we have studied erucic acid accumulation both at the biochemical and molecular levels in two different *Thlaspi arvense* L. strains of European origin that showed differences in some agronomical interesting traits like vernalization, flowering time, or seed size and production. Our data showed a rapid accumulation of $22{:}1^{\Delta 13}$ into TAG in developing Pennycress seeds. At the molecular level, several genes involved in $22{:}1^{\Delta 13}$ biosynthesis like *TaFAE1*, *TaFAD2* or in the regulation of seed oil biosynthesis like *TaDGAT1* or *TaWRI1*, were also characterized. The *TaFAE1* gene from *T. arvense* showed a high phylogenetic correlation with other FAE1 enzymes from plants that also accumulated high amounts of erucic acid in their seed oil. The expression analysis of the *TaFAE1* gene suggested that the accumulation of $22{:}1^{\Delta 13}$ was controlled at the transcriptional level at the earlier stages of seed development and that a change in the acyl fluxes towards elongation versus desaturation took place to favour $22{:}1^{\Delta 13}$ accumulation

in Pennycress seeds. This data might help to improve its biotechnological manipulation for high-quality biodiesel production.

2. Materials and methods

2.1. Plant growth and characterization of plant lines

Pennycress seeds (*Thlaspi arvense* L.) were obtained from different sources. "NASC" seeds were obtained from the *Nothingham Arabidopsis Stock Centre,* UK (NASC). "French" seeds were obtained from the *B&T World Seeds* company (France). Seeds were germinated in plates on wet Whatman paper without addition of any other supplement. Once germinated, seeds were transferred to pots containing a 75:25 mixture of substrate (peat moss, Kekkilä white 420W): vermiculite and grown in a bioclimatic chamber under a light intensity of 120-150 μmol m^{-2} s^{-1}, with a 16h/8h light/dark photoperiod at 22 ºC and a relative humidity of 45%. When grown in the culture chamber, French plants required a vernalization treatment to fully develop into flowers and seeds. To that end, upon emergence of the first pair of true leaves, the plants were transferred to a cold chamber at 6 ºC for 5 weeks with a light intensity of 120 μmol m^{-2} s^{-1} and 10/14 h day/night cycle, respectively. The plants were then placed back into the growth chamber under normal growth conditions and allowed to fully develop.

Pennycress was also grown under field conditions in small experimental plots (10 m^2) located at the research farm of the Estación Experimental de Aula Dei (Consejo Superior de Investigaciones Científicas), in the Aragon region (NE Spain) (41º44'N, 0º46'W, 259 m alt.). Sowing was done at the end of September with a density of 6-8 kg ha^{-1}. The area is characterized by a semiarid Mediterranean climate with an average annual rainfall of 355 mm and an average annual air temperature of 14.5 ºC. The soil of the experimental plots is representative of the soils in semiarid Aragon with a medium texture, alkaline and generally with low organic carbon content. Specifically, for the 0-40 cm depth, this soil is a loam soil (23% sand, 53% silt and 24% clay) with pH=8.3, 13.6 g kg^{-1} of organic matter content and 345 g kg^{-1} of $CaCO_3$. Mean soil nutrient contents (0-40 cm depth) were 11, 130, 28 and 193

mg kg^{-1} for P, K, N, and Mg, respectively. No herbicide or biocide treatments were applied during the crop growth. Seeds were manually harvested at the end of May. The weight and size of seeds, number of seeds per pod, number of pods per plant, and number of seeds per plant were obtained from at least 10 different plants from each Pennycress lines. The amount of oil per seed dry weight basis was obtained gravimetrically by the method of Li et al. (2006) from 1 g of seeds.

2.2. Lipid and fatty acid composition analysis

Total lipids were extracted from *Thlaspi arvense* seeds (0.5 g) in four successive maturation phases (green, green-yellow, yellow-green and dry phase) with chloroform/methanol (2:1, v/v) as described by Bligh and Dyer (1959). Fatty acid methyl esters of total lipids or individual lipid classes were produced by acid-catalyzed transmethylation (Garcés and Mancha, 1993) and analyzed by gas chromatography (GC), using a 7890A (Agilent, Santa Clara, CA USA) fitted with a capillary column (60-m length; 0.25-mm inner diameter; 0.2-µm film thickness) of fused silica (Supelco, Bellafonte, PA, USA) and a FID detector. Helium was used as a carrier gas with a linear rate of 1.2 ml min^{-1} and split ratio of 1/100. The injector temperature was 250 ºC and the detector temperature was 260 °C. The oven temperature was modified as follows: 170 °C for 30 min, then raising the temperature by 5 ºC/min to 200 ºC. 17:0 was used as an internal standard. Data from fatty acid analysis were obtained from three biological experiments in the case of plants grown in culture chamber and two independent biological experiments for plants grown in the field plots. Analysis of variance (ANOVA) was applied to compare treatments. Statistical analyses were carried out with the program Statgraphics Plus for Windows 2.1, using a level of significance of 0.05.

2.3. TAG and DAG analysis by TLC

TLC plates (Silica Gel 60, Merck) were activated by heating to 110 ºC for at least one hour prior to the analysis in order to drive off any moisture. Then, total lipids extracted from the seeds corresponding to the different stages of Pennycress seed development were loaded onto the TLC plates. TAG and DAG fractions were separated in TLC plates developed with a mixture of heptane:diethylether:acetic acid (70:30:1, v/v/v), following the method described by Li-Beisson et al (2013). Detection of lipids was performed by short exposure to iodine vapour. Bands were marked, scraped and extracted with a Methanol:chloroform:water (100:50:40, v/v/v) solution followed by an additional separation in a chloroform:water (50:50, v/v) mixture. Commercial standards of TAG and DAG (Sigma) were used as a reference for band identification. 17:0 was used as an internal standard for quantification purposes. Fatty acid methyl esters of the different TAG and DAG fractions obtained were performed as described above.

2.4. RNA isolation and cDNA synthesis

Total RNA was isolated from 0.1 g of *Thlaspi arvense* seeds from the four maturation stages analyzed (green, green-yellow, yellow-green and dry) using the CTAB-LiCl extraction method of Gasic et al. (2004). RNA concentration and integrity was measured in Nanodrop 2000 UV-Vis Spectrophotometer (Thermo Scientific). cDNAs were synthesized from 3 µg of total RNA using Thermoscript Reverse Transcriptase (Life Technologies) and oligo dT primer, according to the manufacturer's instructions.

2.5. Sequence analyses and manipulation

All sequences were obtained from GENBANK. Protein alignment was performed using the CLUSTALW2 multiple alignment tool. Phylogenetic trees were generated using the PHYML software (Guindon and Gascuel, 2003; www.atgc-montpellier.fr/PHYML), with bootstraps 500.

2.6. RT-PCR expression analysis

The expression patterns of *TaFAE1*, *TaFAD2*, *TaDGAT1*, *TaWRI1* and *ACTIN* genes were examined by reverse-transcription (RT)-PCR assay. Three primer pairs were designed with 5' and 3' flanking sequences of the genes, according to the gene sequences in the GenBank data (KF664164 for *TaFAE1*, DQ518324 for *TaFAD2* and JQ435880 for *ACTIN*, respectively). For the *TaDGAT1* and *TaWRI1* genes, degenerated primers were designed in conserved regions of the protein from other plant species and used for the amplification of a partial fragment of the gene. The fragments were cloned and sequenced to confirm its correspondence to the *TaDGAT1* and *TaWRI1* genes. The primers sequences were for the *TaFAE1* gene, F-FAE1: 5'- ATGACGTCCGTTAACGTTAAGCT-3'; R-FAE1: 5'- TTAGGACCGACCGTTTTGGACACGA-3'; for the *TaFAD2* gene, F-FAD2: 5'- TTCACGCTCGGAGAACTGAAG; R-FAD2: 5'- AACTCCGTAGACGCAGATCA-3'; for the *TaDGAT1* gene, F-DGAT1: 5'- CGTATCGACCGTCGGTTCCA- 3'; R-DGAT1: 5'- GAAGCAGTAGAACATGCAGAGCCA- 3'; for the *WRI1* gene, F-WRI1: 5'- GGAGTCACTAGACATAGATGGA- 3'; R-WRI1: 5'- CGAAATTGGTAACCGCGTT-3'; for the *ACTIN*, F-ACTIN: 5'- CATGTATGTCGCCATCCAAGCTGTTCTC; R-ACTIN: 5'- CAGCAGCTTCCATTCCCACGAAC-3'. The PCR reactions were performed using a GeneAmp PCR system 9700 (Perkin Elmer, Applied Biosystems) and Phusion High-Fidelity DNA Polymerase (New England Biolabs). The PCR products were examined electrophoretically in 1% (w/v) agarose gels. The *TaFAE1*, *TaFAD2*, *TaDGAT1*, *TaWRI1* and *ACTIN* purified PCR products were sequenced after cloning into the pGEM-T-easy Vector (Promega), according to the manufacturer's instructions. The obtained sequences of *TaFAE1* from the two Pennycress lines used in this study have been deposited in the NCBI database with the accession numbers KT223024 and KT223025 for *TaFAE1* gene from NASC and French lines, respectively.

3. Results and Discussion

3.1. Growth characteristics of two varieties of Pennycress

Germination percentages near to 100% were obtained for Pennycress seeds from either NASC or French lines in Petri dishes on a wet Whatman paper (Fig. 1D). It is worth mentioning that these germination rates were obtained with seeds that were stored from around 10-12 months since harvest. Germination of freshly harvested Pennycress seeds resulted in very low germination efficiencies (less than 10%). It has been reported that freshly harvested Pennycress seeds from winter varieties have relatively low germination rates unless treated with the plant hormone gibberellic acid (GA) (Sedbrook et al., 2014). Our results are consistent with this observation, indicating that seed dormancy is very high in these European lines as it seems to occur in the American ones (Sedbrook et al., 2014). Usually, 7-10 days were necessary from the sowing in the Petri dishes to the transfer to soil pots. Plants grew rapidly in the growth chamber producing flowers and generating pods with seeds. Several differences were encountered between both varieties at this stage. NASC line always produced flowers that further developed to seed pods. On the contrary, approximately one third of each batch of germinated French plants transferred to the growth chamber, did not develop flowers maintaining a big rosette for months (Fig. 1A). When French plants were subjected to vernalization at the earlier stages of germination (4-5 leaf stage), in a growth chamber at 6 ºC for 5 weeks, all French line plants produced flowers and seeds. These results might suggest that the French line is “winter or winter facultative” since vernalization was required for fully development. On the contrary, the fact that NASC plants did not require vernalization for flowering and seed development, might suggest that NASC plants are representative of “spring” lines. This observation might be relevant from the agronomic point of view since NASC seeds could be used in short cycle cultivation (sowing in January/February and harvest in May) while those from the French line might require

winter cultivation in a long cycle (sowing in September and harvest in May/June). In addition to this, other differences between lines were also observed. Thus, the NASC plants were taller and showed earlier flowering when compared to the French ones (Figs. 1A-C). It is important to note that these differences in flowering phenotype were observed in plants from both varieties cultivated either in the culture chamber or under field conditions. Seeds from both lines were harvested and their size and seed oil characteristics were compared. To that end, imbred lines by single seed descent were obtained. As shown in Table 1, Pennycress seed weight was higher (22%) in the NASC line than in the French line (1.23 vs. 1.01 mg). These differences in seed weight were statistically significant ($p<0.05$). These seed weights for both lines are within the range of previous data from USA Pennycress accessions that ranged between 0.4-1.3 mg (Sedbrook et al., 2014) and, even, closer to the top of the range. Differences in seed colour were also detected, being the seeds from the French plants of a light red-brown colour while those from the NASC plants were dark-brown (Fig. 1E). Both Pennycress lines showed a similar average number of seeds per pod (9.32 and 9.21 for NASC and French lines, respectively), but again, statistically significant differences ($p<0.05$) were found when the number of pods per plant was counted, resulting in a higher average value for the French line (490 vs 349). This higher number of pods per plant resulted in a higher average number of seeds per plant from the French line when compared with the NASC plants (4,513 vs 3,253). It is worth mentioning that the lower weight of one seed in the French line was compensated with a higher number of pods per plant and, therefore, a higher number of seeds per plant. No problems of pod shattering were detected during plant manipulation. Seed oil characteristics were also analyzed. As shown in Table 1, total seed oil content ranged from 39.08 to 41.18% for NASC and French lines, respectively. No statistically significant differences were detected in this parameter between both Pennycress lines, suggesting that the differences in seed weight and seed production

compensated for a similar oil production. These results are similar to those obtained from the high oil content Pennycress strains from USA (Sedbrook et al., 2014), indicating that both European lines accumulated high oil content in their seeds. The fatty acid composition of the seed oil from both Pennycress NASC and French lines cultivated in the growth chamber was also compared. Erucic acid ($22:1^{\Delta 13}$) was the major fatty acid species in the seed oil, with an average percentage of 35.12 and 36.34 for NASC and French lines, respectively. This high percentage of $22:1^{\Delta 13}$ is within the range of that previously published from USA Pennycress accessions (Sedbrook et al., 2014). No statistically significant differences were detected in this parameter between both Pennycress lines.

At this point, we found interesting to compare the seed oil characteristics from Pennycress plants grown in the controlled chamber with those cultivated in experimental plots under field conditions. Once completed the growing season (September-May) and harvested the Pennycress seeds, their fatty acid content was analyzed. As shown in Table 2, the fatty acid composition of seeds from the NASC line cultivated in field plots was essentially similar to that obtained in the controlled chamber. Erucic acid was also the major fatty acid species in their seed total lipids with an average percentage of 37.23% and 39.68% for the NASC and French plants, respectively (Table 2). Although these values seemed to be slightly higher than those obtained for the same Pennycress lines cultivated under controlled conditions in the chamber (35.12% and 36.34%, for NASC and French lines, respectively), no statistically significant differences were found between both cultivation conditions. This observation might be very useful for the extrapolation of the results obtained from plants grown under controlled conditions to field conditions.

3.2. Fatty acid composition during seed maturation in two varieties of Pennycress

To further characterize the biosynthesis of $22:1^{\Delta 13}$ in Pennycress seeds, we monitored the

fatty acid composition during seed maturation in both NASC and French plants grown in the culture chamber. Seeds from four developmental stages corresponding to green (G), green-yellow (GY), yellow-green (YG), and dry pods (D) were chosen for the analysis. The results are shown in Figure 2. Both lines showed an increase of total lipid accumulation on a seed fresh weight basis in a nearly linear fashion (Fig. 2A). At the initial stages of seed maturation (green), Pennycress seeds were characterized by high levels of erucic acid ($22:1^{\Delta 13}$) that was the major fatty acid present in the analysis, with a content of 27-28% in this stage in both Pennycress lines (Fig. 2B). Seeds also contained relatively high levels of 18:2 (22-23%) and 18:3 (18%), (Figs. 2C and D). With seed maturation, levels of $22:1^{\Delta 13}$ increased from 27-28 to 35-37% in seeds from GY and YG pods. In the French seeds, $22:1^{\Delta 13}$ peaked at the YG developmental stage while in the NASC plants $22:1^{\Delta 13}$ levels were very similar in both GY and YG stages (Fig. 2B). However, no statistically significant differences in the relative $22:1^{\Delta 13}$ values were detected when fatty acid composition from NASC and French lines was compared. In both cases, the increase in $22:1^{\Delta 13}$ was accompanied by a concomitant decrease of 18:2 and also 18:3 (Figs. 2C and D). $22:1^{\Delta 13}$ is synthesized outside the plastid by a membrane bound fatty acid elongation complex (elongase) that uses 18:1 CoA substrates. This increase of $22:1^{\Delta 13}$, concomitant with the decrease of 18:2 and 18:3 in Pennycress seeds during maturation, suggested a change of the acyl flux towards elongation that might result in less 18:1 CoA available for its incorporation into PC at the sn-2 position where it could be desaturated by the microsomal Δ12 desaturase FAD2.

In developing seeds, oil is accumulated in the form of triacylglycerol (TAG) which is the major fraction in total seed-oil lipids (Bates et al., 2013). The direct lipid precursor for TAG assembly is diacylglycerol (DAG) which represents an important intermediate step in the

synthesis between TAG and membrane lipids (Li-Beisson et al., 2013). To further characterize oil biosynthesis in Pennycress seeds, we found interesting to determine the fatty acid composition of TAG and DAG fractions from developing Pennycress seeds. To that end, seeds from at least five plants from both lines grown in the culture chamber were pooled to extract total lipids and subjected to TLC to separate TAG, and DAG. These fractions were subsequently analysed for their fatty acid composition. $22:1^{\Delta 13}$ was the major fatty acid detected in TAG in Pennycress seeds. In the first stages of seed development, corresponding to the G stage, levels of $22:1^{\Delta 13}$ fatty acid represented a mole percent of 34 and 35 % of total fatty acids in TAG for both French and NASC lines, respectively (Figs. 3A and B). In both lines, $22:1^{\Delta 13}$ levels increased with seed development. Thus, in the French seeds, $22:1^{\Delta 13}$ increased from a 34 % in the G stage to the highest level of 42% in the YG stage (Fig. 3A). In the NASC plants, $22:1^{\Delta 13}$ increased from 35% in the G stage to peak at 42 % in the YG stage as occurred in the French line seeds (Fig. 3B). The increase of $22:1^{\Delta 13}$ was concomitant with a decrease in 18:2 and 18:3 during seed development that also occurred in both Pennycress lines (Fig. 3). The $22:1^{\Delta 13}$ content in the TAG fractions from both Pennycress lines was consistent with the 22:1 levels obtained from their total lipids (Fig. 2B and Tables 1 and 2). The high levels of 22:1 detected in TAG at the earlier stages of seed development might suggest that 22:1-CoA pools were rapidly available for their incorporation to TAG in Pennycress plants. In that sense, during the analysis of 22:1 incorporation to TAG in *Crambe abyssinica*, another high-erucic containing species, it was shown that the incorporation of 22:1-CoA to TAG was rapidly accelerated upon seed development (Furmanek et al., 2014).

The fatty acid composition of the DAG fraction during seed development was also analyzed. Differently to what happened in TAG, $22:1^{\Delta 13}$ levels were only slightly higher

with respect to other fatty acids in DAG. $22:1^{\Delta 13}$ levels reached a 17% or 22% at the G developmental stage in both French and NASC lines, respectively (Fig. 4). 18:2 levels were 14 and 19 % for the same developmental stage in both French and NASC lines, respectively (Fig. 4). Interestingly, 18:2 levels increased in DAG during seed development between the G and GY stages in both Pennycress lines. This increase was concomitant with a decrease in 22:1 (Figs. 4A and B), further indicating that the change in acyl fluxes to favour elongation, that occurred during seed development in Pennycress seeds, acts conversely in TAG and DAG fractions. This observation could be related with the fact that a portion of the DAG pool is channelled for PC production instead of TAG biosynthesis, through the action of phosphatidylcholine:diacylglycerol choline phosphotransferase (PDCT; Lu et al., 2009). This PC is incorporated into membrane lipids that require polyunsaturated fatty acids like 18:2.

3.3. Characterization of the TaFAE1 gene from NASC and French Pennycress plant lines

Biosynthesis of very long chain fatty acids (VLCFAs), like $22:1^{\Delta 13}$, is performed in plants by a microsomal fatty acid elongation complex that catalyzes the sequential addition of malonyl-CoA moieties to C18 fatty acids. Each condensation step involves four different activities: (i) condensation of malonyl-CoA with a long chain acyl-CoA; (ii) reduction of 13-keto-acyl-CoA; (iii) dehydration to an enoyl-CoA; and, (iv) reduction of the enoyl-CoA resulting in an elongated acyl-CoA (James et al., 1995). Despite the fact that these four activities are required for the elongation (von Wettstein-Knowles, 1982), ectopic expression experiments of the *FAE1* gene in Arabidopsis, together with heterologous expression experiments in yeast, confirmed that the FAE1 activity is the rate-limiting step of the elongase reaction (Millar and Kunst, 1997; Katavic et al., 2001). Our lipid analysis indicated

that 22:1 was rapidly synthesized and incorporated to seed oil in Pennycress. According to this, as a first step to characterize the biosynthesis of $22:1^{\Delta 13}$ in Pennycress seeds at the molecular level, we focused our attention in the *FAE1* gene. Based on the sequence similarity among plant fatty acid elongases, we designed specific primers from *FAE1* sequences retrieved from available databases. RT-PCR screening of a cDNA obtained from Pennycress seeds from either NASC or French lines resulted in the amplification of a 1,518 bp band. This RT-PCR band was cloned and sequenced, confirming that it corresponded to the full-length *TaFAE1* gene sequence. The same primers were used for amplification of the *TaFAE1* gene from genomic DNA from both Pennycress lines. Sequencing of the genomic fragments from both lines confirmed the sequence data obtained with the cDNA. The RT-PCR and genomic analysis suggested that a single *TaFAE1* gene is present in the *T. arvense* genome. It is worth mentioning that no introns were detected in the genomic amplification of the *TaFAE1* gene. This feature was shared with other *FAE1* genes from other *Brassicaceae* like *Arabidopsis thaliana* (Katavic et al., 2002; Mietkiewska et al., 2004). Figure 5 shows the deduced protein sequence alignment of the *Ta*FAE1 enzyme from either NASC (Genbank accession number KT223024) or French (Genbank accession number KT223025) Pennycress lines, as well as that from Arabidopsis (that does not accumulate erucic acid in seeds) and that from *Brassica napus* (a high erucic acid producer). The *Ta*FAE1 complete ORF encoded a protein of 506 amino acid that showed an 86% identity (93.67% homology) with respect to that from Arabidopsis and a 91.3% of identity (95.45% homology) with respect to that from *B. napus*. It should be noted that no differences were observed between the *Ta*FAE1 protein sequences from both Pennycress lines (Fig. 5). Several residues have been identified as crucial for FAE1 activity (Ghanevati and Jaworski, 2001, 2002). The cysteine (Cys85, Cys223, Cys312, Cys389, Cys460) and histidine (His302, His387, His391, His420) residues identified as part of the active site of the FAE1

enzyme (Ghanevati and Jaworski, 2002) were all present in both Pennycress lines. Particularly, Cys223, His391, and Asn424 (Ghanevati and Jaworski, 2001, 2002) were also conserved in all the analyzed sequences. Mutations in Ser282 have been associated with the low erucic trait (LEA) in *Brassica napus* (Han et al., 2001; Katavic et al., 2002). However, as shown in Figure 5, Ser282 was conserved in Pennycress and Brassica plants (high erucic content, HEA) but also in Arabidopsis (low erucic producer). This observation is in line with reports that suggest that Ser282 itself is not an absolute requirement for a functional FAE1 condensing enzyme (Katavic et al., 2004). To further characterize the *TaFAE1* gene, a phylogenetic tree was constructed with 19 FAE1 deduced complete protein sequences that included 8 high erucic (>25%) accumulators, including Pennycress, as well as 11 species that did not accumulate erucic acid in seeds in significant amounts. The phylogenetic tree is shown in Figure 6. *TaFAE1* was grouped in the same clade in which *FAE1* genes from other *Brassicaceae* that accumulated high amounts of erucic acid in their seed were present. This result further strengthen the hypothesis that, although the FAE1 enzymes are highly conserved in the plant kingdom and independently of the acyl-CoA pools, certain amino acid residues or protein domains, not directly related with the active site, could be behind the high erucic acid trait phenotype of certain plants, particularly those from the *Brassicaceae* family. Recent studies, suggested a correlation between several sequence motifs and the content of erucic acid in seeds (Sun et al. 2013). A search of the seven motifs identified by Sun et al. (2013) as associated with high erucic acid accumulation in seeds showed that all of them were present in the *Ta*FAE1 sequences from both Pennycress lines while none of them were detected in the *AtFAE1* sequence (Fig. 5). Complementation analysis of *fae1* mutants from Arabidopsis using the FAE1 enzyme from Pennycress is in progress to determine the influence of specific residues on the high erucic acid trait. This information would be extremely useful to manipulate the erucic acid content in any plant

seed oils.

3.4. Expression analysis of genes involved in erucic acid accumulation in Pennycress seeds during maturation.

As a first step to understand the molecular mechanism controlling erucic acid accumulation in Pennycress seeds, we decided to monitor the expression of several genes encoding key enzymes that participate in the biosynthesis of 22:1. Taking into account the fatty acid composition of total lipids and TAG during Pennycress seed development, several genes focused our attention. On one hand, the *TaFAE1* gene that encodes the FA elongase responsible for the elongation of 18:1 CoA to 20:1 CoA and then to $22:1^{\Delta 13}$ was the first target choice. On the other hand, the microsomal Δ12 desaturase *FAD2*, that uses 18:1 esterified at sn-2 position of PC as a substrate for 18:2 biosynthesis, was a second choice. The compared analysis of the expression of both *TaFAE1* and *TaFAD2* genes might help us to analyze, at the molecular level, the changes in acyl fluxes between elongation and desaturation that seemed to occur during seed maturation as reflected by the fatty acid composition of total lipids and the TAG fraction from both Pennycress lines. The *DGAT1* gene, encoding the diacylglycerol acyl-transferase 1, that encodes the final step of the Kennedy pathway, and catalyzes the acylation of sn1,2 diacylglycerol (DAG) at the sn-3 position to produce TAG was also included in the expression analysis.

RT-PCR of the different genes analyzed from both NASC and French lines are shown in Figure 7. *TaFAE1* mRNA levels were high in the initial stages of seed development in both Pennycress lines (G and GY) but decreased in the latter stages of maturation (YG and D), (Fig. 7). In fact, *TaFAE1* mRNA was undetectable in seeds from dry pods in both French and NASC plant lines. Note that no *TaFAE1* mRNA was detected in leaves (Fig. 7), which is consistent with the specific role of FAE1 in seeds. Although the expression profile was

similar in both Pennycress lines, French plants showed a higher accumulation pattern of *TaFAE1* transcripts in seeds from green pods while the NASC plants showed higher *TaFAE1* transcript accumulation in seeds from green-yellow pods with less *TaFAE1* mRNA at the G stage (Fig. 7). Nevertheless, the expression profile of the *TaFAE1* gene in both Pennycress lines was consistent with the lipid analysis shown in Figures 2 and 3, indicating that elongase activity acts at the early stages of seed development on 18:1-CoA pools to rapidly provide 22:1-CoA for its incorporation to TAG. This result was also consistent with the presence of *TaDAGT1* mRNA already at the G or GY stages, indicating that TAG was rapidly synthesized upon seed development. It should be noted that no changes in *TaDGAT1* mRNA levels were detected during Pennycress seed maturation in both lines suggesting that its expression was controlled mainly at the post-transcriptional level. This expression pattern is consistent with that reported in Arabidopsis (Lu et al., 2003; Li et al., 2010), Tung tree (Shockey et al., 2006) or *Ricinus communis* (Kroon et al., 2006). Our expression analysis suggested that the FAE1 activity during seed maturation was controlled at least at the transcriptional level in Pennycress seeds. Overexpression of FAE1 resulted in higher levels of VLCFAs in Arabidopsis (Millar and Kunst, 1997). Similarly, seed-specific expression of *N. officinale* or *T. nudicaulis* FAE1 in Arabidopsis also resulted in higher $22{:}1^{\Delta 13}$ and $20{:}1^{\Delta 11}$accumulation in seed oil, respectively (Mietkiewska et al., 2004, 2007). All these results indicate that, independently of the availability of acyl-CoA pools for the elongase complex, any increase in the amount of the FAE1 enzyme results in an increase of erucic acid in seeds.

The fatty acid composition from total lipids and TAG, as well as the expression profile of the *TaFAE1* gene suggested that FAE1 elongase activity was very high at the early stages of seed development. Only the DAG fraction showed an increase in 18:2 fatty acids. Since an increase in elongase activity might result in a change of acyl fluxes with less 18:1 CoA

pools available for its incorporation into PC and further desaturation by the micsosomal Δ12 desaturase FAD2, we decided to monitor the expression of the *TaFAD2* gene during Pennycress seed development. As shown in Figure 7, *TaFAD2* mRNA levels remained unaffected during all the steps of seed development. This observation suggests that the coordination of Δ12 desaturase activity during seed maturation was controlled at the post-transcriptional level. This is not striking since post-transcriptional regulation has been shown to control microsomal desaturase activity in many plant species (O´Quin et al., 2010). Nevertheless, the expression analysis of *TaFAE1* and *TaFAD2* suggests the existence of a complex mechanism, acting on different genes at different levels during seed development to maintain appropriate levels of erucic acid in Pennycress seeds. The identification of control elements acting on this regulation would be extremely useful to manipulate the erucic acid content not only in the Pennycress oil but in any plant seed oils. To shed light into this question we decided to include the *WRI1* gene in our analysis. *WRI1* encodes a transcription factor from the APETALA2/ethylene-responsive element binding family that governs the flux of carbon through glycolisis and fatty acid biosynthesis in developing seeds by regulating the expression of a group of target genes that encode key enzymes of these pathways (Baud et al., 2007; To et al., 2012). These include the *FAE1* and *DGAT1*, which are critical for determining the composition and/or quantity of seed storage oil in Arabidopsis (To et al., 2012; Li-Beisson et al., 2013). When the *TaWRI1* gene expression was analyzed, a similar pattern of expression to that of the *TaFAE1* gene (Fig. 7) was observed, with higher accumulation of the *TaWRI1* RNA at the earlier stages of seed development. This result suggests that, as occurred in Arabidopsis, the *Ta*WRI1 is a major transcriptional factor controlling oil biosynthesis in Pennycress seeds. Overall, the expression analysis identified several genes like *TaFAE1*, *TaDGAT1* or *TaWRI1* as potential candidates for genetic manipulation in order to increase oil-seed content in this plant

species.

4. Conclusion

Pennycress has become an attractive oilseed feedstock for biofuel production because of its excellent low temperature and energetic properties. Understanding the biochemical and molecular pathways involved in oil biosynthesis in Pennycress seeds together with the agronomic analysis of new varieties would be necessary to guide future crop breeding efforts directed towards an increase in oil yield and quality as well as other important agronomic traits. In this work we have characterized oil and erucic acid biosynthesis both at the biochemical and molecular levels in two Pennycress lines of European origin. Our data showed differences between both lines in some agronomic interesting traits like vernalization (winter cycle in the case of French and spring cycle in the case of NASC), plant size (taller in the NASC plants), flowering time (delayed in the French plants) or seed production (higher in the French line). Both lines showed a high amount (35-40%) of erucic acid in their seed oil. Our lipid analysis showed that TAG was the major reservoir of $22{:}1^{\Delta 13}$ in the seed and that it was incorporated to TAG very rapidly during seed development. This analysis also suggested that a change in acyl fluxes towards elongation occurred during Pennycress seed maturation. This pattern of erucic acid accumulation was consistent with the transcriptional control of the *TaFAE1* gene that encodes the main component of the elongase complex responsible of $22{:}1^{\Delta 13}$ biosynthesis. *TaWRI1*, the master regulator of seed-oil biosynthesis, might participate in this regulation. Overall, these data might guide the biotechnological manipulation of Pennycress seeds to maximize the production of fatty acids for their use in biofuel production.

Acknowledgements

Authors wish to thank M. Josefa Salvador, Marina de la Vega and Ricardo Gracia for skilful technical assistance in growing Pennycress in growth chambers and field plots. This work was supported by the Spanish Ministry of Economy and Competitiveness (MINECO AGL2014-55300-R) and Aragón Government (Grant E33).

References

Baud, S., Mendoza, M.S., To, A., Harscoët, E., Lepiniec, L., Dubreucq, B., 2007. WRINKLED1 specifies the regulatory action of LEAFY COTYLEDON2 towards fatty acid metabolism during seed maturation in Arabidopsis. Plant J. 50, 825-838.

Bates, P.D., Stymne, S., Ohlrroge, J., 2013. Biochemical pathways in seed oil synthesis. Curr. Op. Plant Biol.16, 358-364.

Bligh, E.G., Dyer, W.S., 1959. A rapid method of total lipid extraction and purification. Can. J. Biochem. Physiol. 37, 911-917.

Bouaid, A., Martínez, M., Aracil, J., 2009. Production of biodiesel from bioethanol and Brassica carinata oil: oxidation stability study. Biores. Tech. 100, 2234-2239.

Cahoon, E.B., Shockey, J.M., Dietrich, C.R., Gidda, S.K., Mullen, R.T., Dyer, J.M., 2007. Engineering oilseeds for sustainable production of industrial and nutricional feedstocks: solving bottlenecks in fatty acid flux, Curr. Op. Plant Biol. 10, 236-244.

Dorn, K.M., Fankhauser, J.D., Wyse, D.L., Marks, M.D., 2013. De novo assembly of the pennycress (*Thlaspi arvense*) transcriptome provides tools for the development of a winter cover crop and biodiesel feedstock. Plant J. 75, 1028-1038.

Evangelista, R.L., Isbell, T.A., Cermak, S.C., 2012. Extraction of pennycress (*Thlaspi arvense* L.) seed oil by full pressing. Ind. Crop Prod. 37, 76-81.

Fan, J., Shonnard, D.R., Kalnes, T.N., Johnsen, P.B., Rao, S., 2013. A life cycle assessment of pennycress (*Thlaspi arvense* L.) –derived jet fuel and diesel. Biomass and Bioenergy 55, 87-100.

Frohlich, A., Rice, B., 2005. Evaluation of Camelina sativa oil as a feedstock for biodiesel production. Ind, Crops and Prod. 21, 25-31.

Furmanek, T., Demski, K., Banas, W., Haslam, R., Napier, J., Stymne, S., Banas, A., 2014. The utilization of the Acyl-CoA and the involvement PDAT and DGAT in the biosynthesis of erucic acid-rich tryacylglycerols in Crambe seed oil. Lipids 49, 327-333.

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

Gasic, K., Hernández, A., Korban, S., 2004. RNA extraction from different apple tissues rich in polyphenols and polysaccharides for cDNA library construction. Plant Mol, Biol. Rep 22, 437-438.

Ghanevati, M., Jaworski, J.G., 2001. Active-site residues of a plant membrane-bound fatty acid elongase beta-ketoacyl-CoA synthase FAE1 KCS. Biochim. Biophys. Acta 1530, 77-85.

Ghanevati, M., Jaworski, J.G., 2002. Engineering and mechanistic studies of the Arabidopsis FAE1 beta-ketoacyl-CoA synthase, FAE1 KCS. Eur. J. Biochem. 269, 3531-3539.

Groeneveld, J.H., Klein, A.M., 2014. Pollination of two oil-producing plant species:

camelina (*Camelina sativa* L Krantz) and pennycress (*Thlaspi arvense* L.) double cropping in Germany. GCB Bioenergy 6, 242-251.

Groeneveld, J.H., Klein, A.M., 2015. Pennycress-corn double cropping increases ground beetle diversity. Biomass and Bioenergy 77, 16-25.

Guan, R., Lager. I., Li, X., Stymne, S., Zhu, L.-H., 2014. Bottlenecks in erucic acid accumulation in genetically engineered ultrahigh erucic acid *Crambe abyssinica*. Plant Biotechnol. J. 12, 193-203.

Guindon, S., Gascuel O. 2003. A simple, fast and accurate algorithm to estimate large phylogenies by maximum likehood. Systematic Biol. 52, 696-704.

Han, J., Lühs, W., Sonntag, K., Zähringer, U., Borchardt, D.S., Wolter, F.P., Heinz, E., Frentzen M., 2001. Functional characterization of beta-ketoacyl-CoA synthase genes from *Brassica napus* L. Plant Mol. Biol. 46, 229-239.

James, D.W., Lim, E., Keller, J., Plooy, I., Dooner, H.K., 1995. Directed tagging of the Arabidopsis fatt acid elongation 1 (*FAE1*) gene with the maize transposon activator. Plant Cell 7, 309-319.

Katavic, V., Friesen, W., Barton, D.L., Gossen, K.K., Giblin, E.M., Luciw, T., An, J., Zou, J.-T., Mackenzie, S.L., Keller, W.A., Males, D., Taylor, D.C., 2001. Improving erucic acid content in rapeseed through biotechnology: What can the *Arabidopsis FAE1* and the yeast *SLC1-1* genes contribute. Crop Sci. 41, 739-747.

Katavic, V., Mietkiewska, E., Barton, D.L., Giblin, E.M., Reed, D.W., Taylor, D.C., 2002. Restoring enzyme activity in non-functional low erucic acid *Brassica napus* fatty acid elongase 1 by a single amino acid substitution. Eur. J. Biochem. 269, 5625-5631.

Katavic, V., Barton, D.L., Giblin, M., Reed, D.W., Kumar, A., Taylor, D.C., 2004. Gaining insight into the role of serine 282 in *B. napus* FAE1 condensing enzyme. FEBS Lett. 562, 118-124.

Kroon, J.T.M., Wei, W., Simon, W.J., Slabas, A.R. (2006). Identification and functional expression of a type 2 acyl-CoA:diacylglycerol acyltransferase (DGAT2) in developing castor bean seeds which has high homology to the major triglyceride biosynthetic enzyme of fungi and animal. Phytochemistry 67, 2541-2549.

Li, R., Yu, K., Hildebrand, D.F. 2010. DGAT1, DGAT2 and PDAT expression in seeds and other tissues of epoxy and hidroxy fatty acid accumulating plants. Lipids 45, 145-157.

Li, Y.H., Beisson, F., Pollard, M., Ohlrogge, J., 2006. Oil content of Arabidopsis seeds: The influence of seed anatomy, light and plant-to plant variation. Phytochemistry 67, 904–915.

Li-Beisson, Y., Shorrosh, B., Beisson, F., Andersson, M.X., Arondel, V., Bates, P.D., Baud, S., Bird, D., Debono, A., Durret, T.P., Franke, R.B., Graham, I.A., Katayama, K., Kelly, A.A., Larson, T., Markham, J.E., Miquel, M., Molina, I., Nishida, I., Rowland, O., Samuels, L., Schmid, K.M., Wada, H., Welti, R., Xu, C., Zallot, R., Ohlrogge, J., 2013. Acyl-lipid metabolism. Arabidopsis Book 11, e0161.

Lu, C.L., Bayon de Noyer, S., Hobbs, D.H., Kang, J., Wen, Y., Krachtus, D., Hills, M. 2003. Expression pattern of diacylglycerol acyltransferase-1, an enzyme involved in

triacylglycerol biosynthesis in Arabidopsis thaliana. Plant Mol Biol. 52, 31-41.

Lu, C., Xin, Z., Ren, Z., Miquel, M., 2009. An enzyme regulating triacylglycerol composition is encoded by the ROD1 gene of Arabidopsis. Proc. Natl, Acad. Sci. USA 106, 18837-18842.

Mietkiewska, E., Giblin, E.M., Wang, S., Barton, D.L., Dirpaul, J., Brost, J.M., Katavic, V., Taylor, D.C., 2004. Seed-specific heterologous expression of a nasturtium FAE gene in Arabidopsis results in a dramatic increase in the proportion of erucic acid. Plant Physiol. 136, 2665-2675.

Mietkiewska, E., Brost, J.M., Giblin, E.M., Barton, D.L., Taylor, D.C., 2007. *Teesdelia nudicalis FAE1* complements the *fae1* mutation in transgenic Arabidopsis thaliana plants and shows a preference for elongating oleic acid to eicosenoic acid. Plant Sci. 173, 198-205.

Millar, A.A., Kunst, L., 1997. Very-long-chain fatty acid biosynthesis is controlled through the expression and specificity of the condensing enzyme. Plant J. 12, 101-111.

Moser, B.R., Knothe, G., Vaughn, S.F., Isbell, T.A., 2009a. Production and evaluation of biodiesel from field Pennycress (*Thlaspi arvense* L.) oil. Energy Fuels 23, 4149-4155.

Moser, B.R., Shah, S.N., Winkler-Moser, J.K., 2009b. Composition and physical properties of cress (*Lepidium sativum* L.) and field pennycress (*Thlaspi arvense* L.) oils. Ind. Crops Prod. 30, 199-205.

Moser, B.R., 2012. Biodiesel from alternative feedstocks: camelina and field pennycress. Biofuels 3, 193-209.

O´Quin, J.B., Bourassa, L., Zhang, D.Y., Shockey, J.M., Gidda, S.K., Fosnot, S., Chapman, K.D., Mullen, R.T., Dyer, J.M. 2010. Temperature-sensitive post-translational regulation of plant omega-3 fatty acid desaturases is mediated by the endoplasmic reticulum-associated degradation pathway. J. Biol. Chem. 285, 21781-21796.

Robson, P., Jensen, E., Hawkins, S., White, S.R., Kenobi, K., Clifton-Brown, J., Donnison, I., Farrar, K., 2013. Accelerating the domestication of a bioenergy crop: identifying and modelling morphological targets for sustainable yield increase in *Miscanthus*. J. Exp. Bot. 64, 4143-4155.

Sedbrook, J.C., Phippen, W.B., Marks, M.D., 2014. New approaches to facilitate rapid domestication of a wild plant to an oilseed crop: example pennycress (*Thlaspi arvense* L.). Plant Sci. 227, 122-132.

Sharma, N., Cram, D., Huebert, T., Zhou, N., Parkin, I.A.P., 2007. Exploiting the wild crucifer *Thlaspi arvense* to identify conserved and novel genes expressed during a plant's response to cold stress. Plant Mol. Biol. 63, 171-184.

Shockey, J.M., Gidda, S.K., Chapital, D.C., Kuan, J.-C., Dhanoa, P.K., Bland, J.M., Rothstein, S.J., Mullen, R.T., Dyer, J.M. (2006). Tung tree DGAT1 and DGAT2 havew nonredundant function in triaacylglycerol biosynthesis and are localized to different subdomains of the endoplasmic reticulum. The Plant Cell 18, 2294-2313.

Sun, X., Pang, H., Li, M., Guo, H., Yan, Q., Hang, Y., 2013. Evolutionary pattern of the FAE1 gene in Brassicaceae and its correlation with the erucic acid trait. PLOS One 8, 1-

12.

To, A., Joubès, J., Barthole, G., L'Ecureuil, A., Scagnelli, A., Jasisnski, S., Lepiniec, L., Baud, S., 2012. WRINKLED transcription factors orchestrate tissue-specific regulation of fatty acid biosynthesis in Arabidopsis. Plant Cell 24, 5007-5023.

Tsogtbaatar, E., Cocuraon, J.-C., Corchado Sonera, M., Alonso, A.P., 2015. Metabolite fingerprinting of pennycress (*Thlaspi arvense* L.) embryos to assess active pathways during oil biosynthesis. J, Exp. Bot. 66, 4267-4277.

Vaughn, S.F., Isbell, T.A., Weisleder, D., Berhow, M.A., 2005. Biofumigant compounds released by field pennycress (*Thlaspi arvense* L.) seedmeal. J. Chem. Ecol. 31, 167-177.

Von Wettstein-Knowles, P.M., 1982. Elongase and epicuticular wax biosynthesis. Physiol. Veg. 20, 797-809.

Warwick, S.I., Francis, A., Susko, D.J., 2002. The biology of Canadian weeds. 9. *Thlaspi arvense* L. (updated). Can. J. Plant. Sci. 8, 803-823.

FIGURE LEGENDS

Figure 1. Plant growth comparison of the different Pennycress accessions used in this work. (A) Pennycress plants grown in the culture chamber. Left, French line plants that did not develop flowers showing that vernalization was required; middle, French line plants showing normal growth but delayed flowering with respect to NASC plants; right, NASC plants. All Pennycress lines were sown at the same date. (B,C) Photographs showing the differences in flowering between both plant lines: (B) French plants and (C) NASC plants. Plants from both lines were sown at the same time and grown in a culture chamber, under controlled environment. (D) Germination assays of Pennycress seeds from both NASC and French plants. Seeds were germinated onto a wet Whatman paper in the absence of any chemical. Left plate, NASC seeds; Right plate, French seeds. (E) Seeds obtained from both Pennycress lines; left plate, French seeds; right plate, NASC seeds.

Figure 2. Fatty acid composition analysis of Pennycress (*Thlaspi arvense* L.) seeds during maturation. (A) Total lipid on a fresh weight basis, (B) Erucic acid, 22:1, (C) linoleic acid, 18:2; and (D) linolenic acid, 18:3, content. Data are expressed as molar percentages from the quantitative analysis of peak areas obtained from the GC analysis. Data were obtained from three independent experiments. SD±3. Open circles, NASC line; closed circles, French line. The photograph shows the maturation stages of Pennycress pods from which the seeds were analyzed: Green (G), Green-Yellow (GY), Yellow-Green (YG) and Dry (D).

Figure 3. Fatty acid composition of the TAG fraction of the Pennycress seed oil from French (A) and NASC (B) plants. G, Green seed, GY, Green-yellow seed, YG, yellow-green

seed and D, Dry seed developmental stages. Data are means of two independent biological experiments.

Figure 4. Fatty acid composition analysis of DAG fraction from Pennycress oil from French (A) and NASC (B) plants. Closed circles, erucic acid, 22:1; open circles, linolenic acid, 18:3; open triangles, linoleic acid, 18:2. Data are expressed as molar percentages from the quantitative analysis of peak areas obtained from the GC analysis. Data were obtained from at least three independent experiments. SD±3.

Figure 5. Comparison of the deduced protein sequence of the *Thlaspi arvense* FAE1 protein from both NASC and French lines with FAE1 proteins from *Arabidopsis thaliana* (low erucic acid producer) and *Brassica napus* (High erucic acid producer). Sequences were aligned using CLUSTALW2. GenBank accession numbers are: *Thlaspi arvense* NASC line, (KT223024); *Thlaspi arvense* French line, (KT223025); *Arabidopsis thaliana* (U29142.1) and *Brassica napus* (U50771.1). Identical residues are indicated by an asterisk below the sequence. Conservative substitutions were indicated as a double point below the sequence. Semi-conservative substitutions were indicated as a single point below the sequence. Residues involved in the active site of the enzyme are indicated with a cross above the sequence. Sequence motifs related with the high erucic acid trait are indicated by a line over the sequence.

Figure 6. Phylogenetic tree of the FAE1 protein from different plant species. The names and Genbank accession numbers of the sequences used for the analysis are: *Arabidopsis thaliana* (U29142.1); *Brassica napus* (U50771.1); *Brassica oleracea* (GU325726.1); *Brassica rapa* (GU325723.1); *Crambe hispanica* (KF664157); *Crambe abyssinica* (KC565738); *Isatis*

tinctoria (AY888038.1); *Lepidium campestre* (FJ907545.1); *Limnanthes douglasii* (AF247134); *Marchantia polymorpha* (AY308831); *Nastortium officinale* (AY0826190); *Orychophragmus violaceus* (AY888042.1)*; Simondsia chinensis* (U37088); *Sinapsis arvense* (AY888041.1); *Sinapsis alba* (AY888040.1); *Teesdalia nudicaulis* (*EF186003.1); Thlaspi arvense* (KT223024); *Tropaelum majus* (AY082610); *Zea mays* (AJ29770).

Figure 7. *TaFAE1, TaFAD2, TaDGAT1 and TaWRI1* gene expression during seed maturation in both Pennycress lines. Bold letters correspond to the different seed maturation stages tested: Green (G), Green-Yellow (GY), Yellow-Green (YG) and Dry (D). (L) shows the expression in leaf tissue. ACTIN was used as housekeeping gene. All PCRs were done under non-saturating conditions.

Table 1. Seed oil and size characteristics of two *Thlaspi arvense* L. lines

	% oil[1]	% 22:1[2]	Seed wt (g)[3]	Pods/plant[4]	Seeds/pod[4]	Seeds/plant[5]
NASC Line	39.08±5.08	35.12±2.08	0.123±0.003*	349±23.06*	9.32±1.88	3253
French Line	41.18±7.61	36.34±5.03	0.101±0.004*	490±27.96*	9.21±1.64	4513

[1] Calculated as total amount of oil on a seed dry weight basis (dwb).
[2] Data were obtained from three biological replicates.
[3] 100 seeds were weighted. Data were obtained from 10 independent determinations.
[4] Data were obtained from 10 different plants.
[5] Data were deducted from results of seeds/pod and pod/plant.
*. Differences were statistically significant at $p<0,05$.

Table 2. Fatty acid composition of seeds from *Thlaspi arvense* L. lines grown in a culture chamber and in field plots.

Fatty acid	NASC chamber	NASC field	French chamber	French field
16:0	4.26±0.76	3.61±0.05	4.09±0.13	3.23±0.07
16:1	0.37±0.050	0.34±0.001	0.40±0.04	0.33±0.007
18:0	0.68±0.23	0.52±0.001	0.61±0.12	0.48±0.001
18:1	3.14±1.94	11.64±0.07	13.16±3.01	11.39±0.05
18:2	20.26±0.51	19.54±0.25	20.71±1.34	19.54±0.01
18:3	13.02±1.47	11.17±0.26	13.14±0.76	11.69±0.04
20:0	0.22±0.09	0.24±0.007	0.18±0.01	0.23±0.001
20:1	10.26±1.08	11.09±0.007	8.94±0.60	9.83±0.07
22:0	0.42±0.16	0.13±0.007	0.42±0.19	0.15±0.007
22:1	35.12±2.08	37.23±0.55	36.34±5.02	39.68±0.08
24:0	0.15± 0.03	0.09±0.001	0.11±0.04	0.07±0.007
24:1	3.87± 0.69	4.18±0.07	3.36±0.03	3.17±0.03

Data (mole % from total fatty acids) are means of four independent biological determinations for chamber experiments and two independent biological determinations for field experiments.

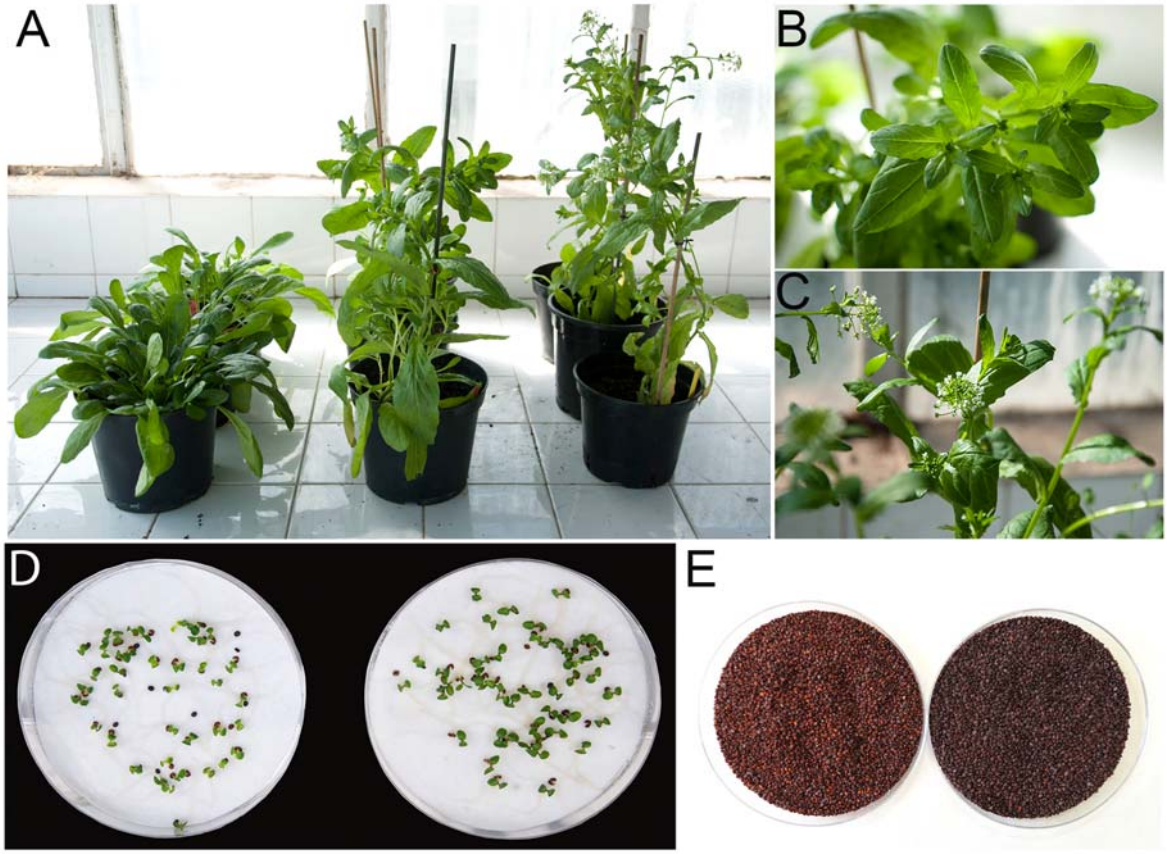

Fig. 1. Plant growth comparison of the different Pennycress accessions used in this work. (A) Pennycress plants grown in the culture chamber. Left, French line plants that did not develop flowers showing that vernalization was required; middle, French line plants showing normal growth but delayed flowering with respect to NASC plants; right, NASC plants. All Pennycress lines were sown at the same date. (B,C) Photographs showing the differences in flowering between both plant lines: (B) French plants and (C) NASC plants. Plants from both lines were sown at the same time and grown in a culture chamber, under controlled environment. (D) Germination assays of Pennycress seeds from both NASC and French plants. Seeds were germinated onto a wet Whatman paper in the absence of any chemical. Left plate, NASC seeds; Right plate, French seeds. (E) Seeds obtained from both Pennycress lines; left plate, French seeds; right plate, NASC seeds.

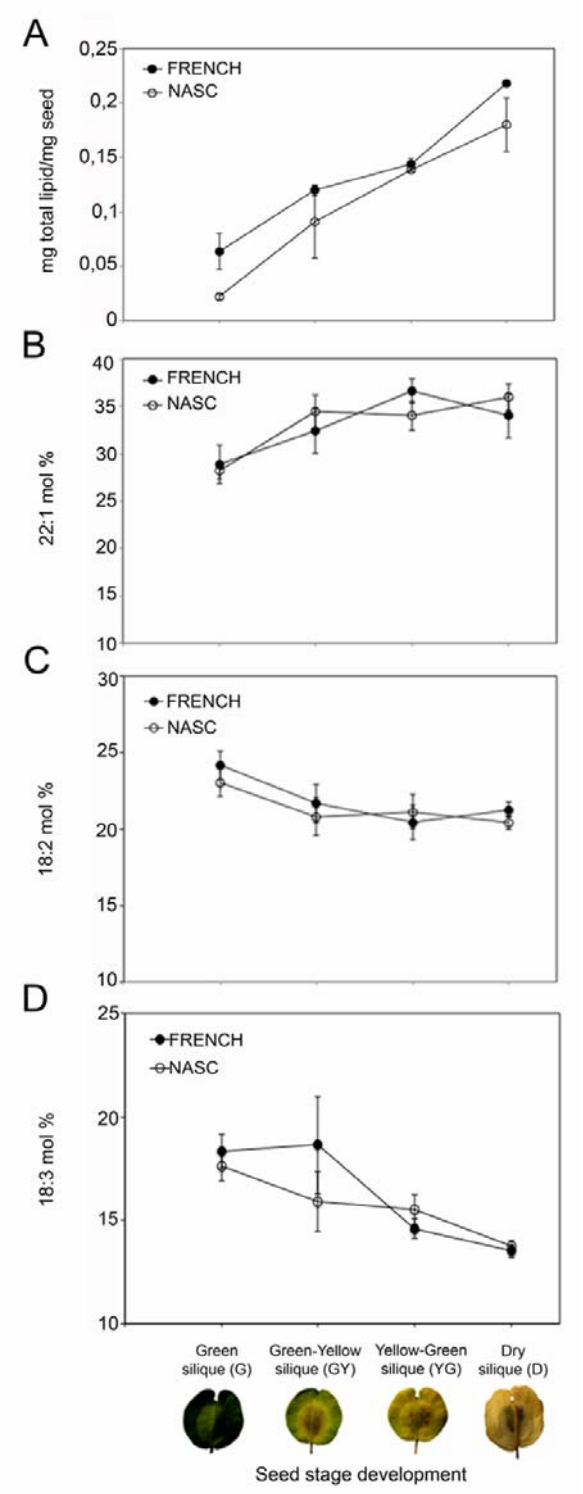

Fig. 2. Fatty acid composition analysis of Pennycress (*Thlaspi arvense* L.) seeds during maturation. (A) Total lipid on a fresh weight basis, (B) Erucic acid, 22:1, (C) linoleic acid, 18:2; and (D) linolenic acid, 18:3, content. Data are expressed as molar percentages from the quantitative analysis of peak areas obtained from the GC analysis. Data were obtained from three independent experiments. SD±3. Open circles, NASC lines; closed circles, French lines. The photograph shows the maturation stages of Pennycress pods from which the seeds were analyzed: Green (G), Green-Yellow (GY), Yellow-Green (YG) and Dry (D).

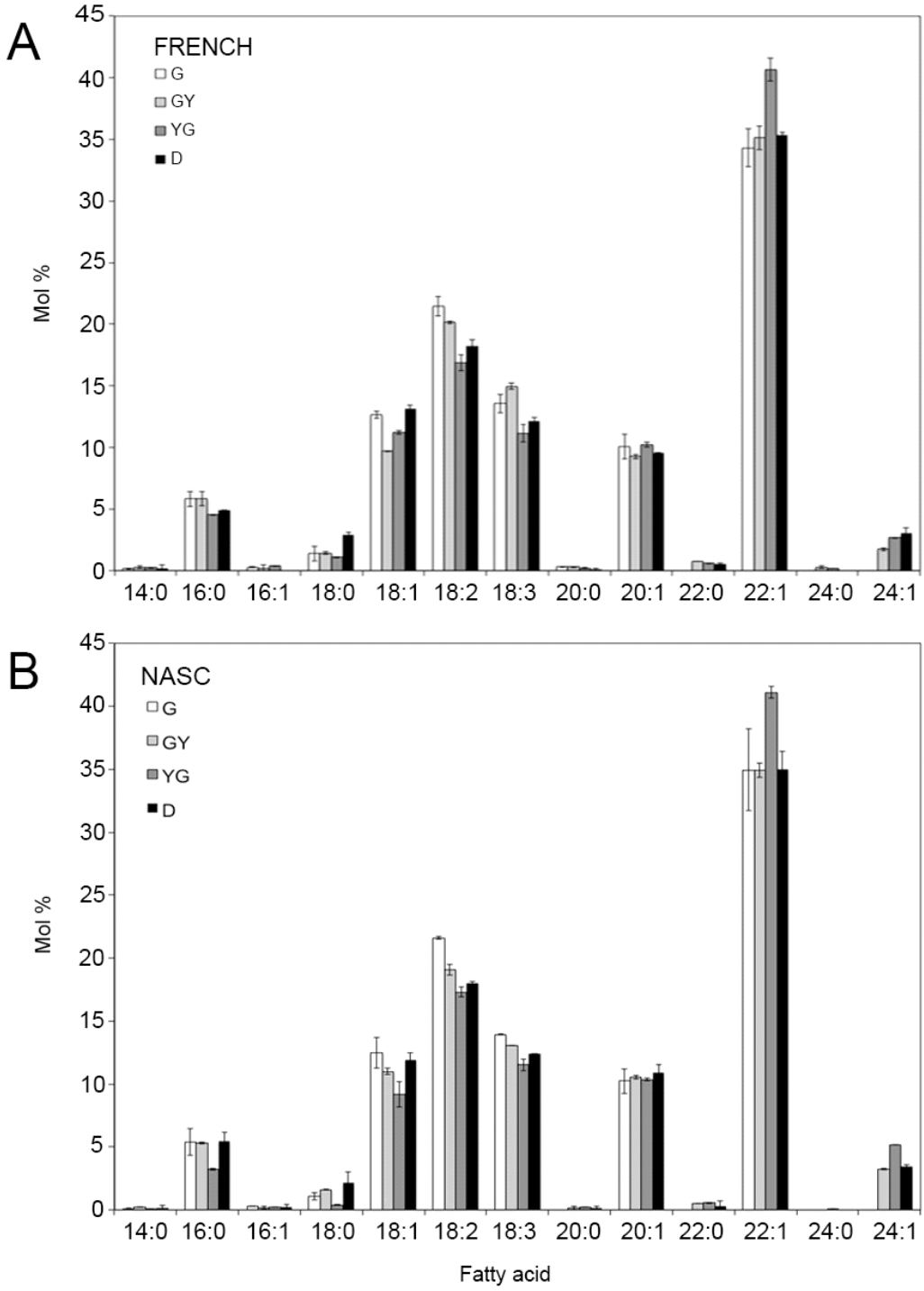

Fig. 3. Fatty acid composition of the TAG fraction of the Pennycress seed oil from French (A) and NASC (B) plants. G, Green seed, GY, Green-yellow seed, YG, yellow-green seed and D, Dry seed developmental stages. Data are means of two independent biological experiments.

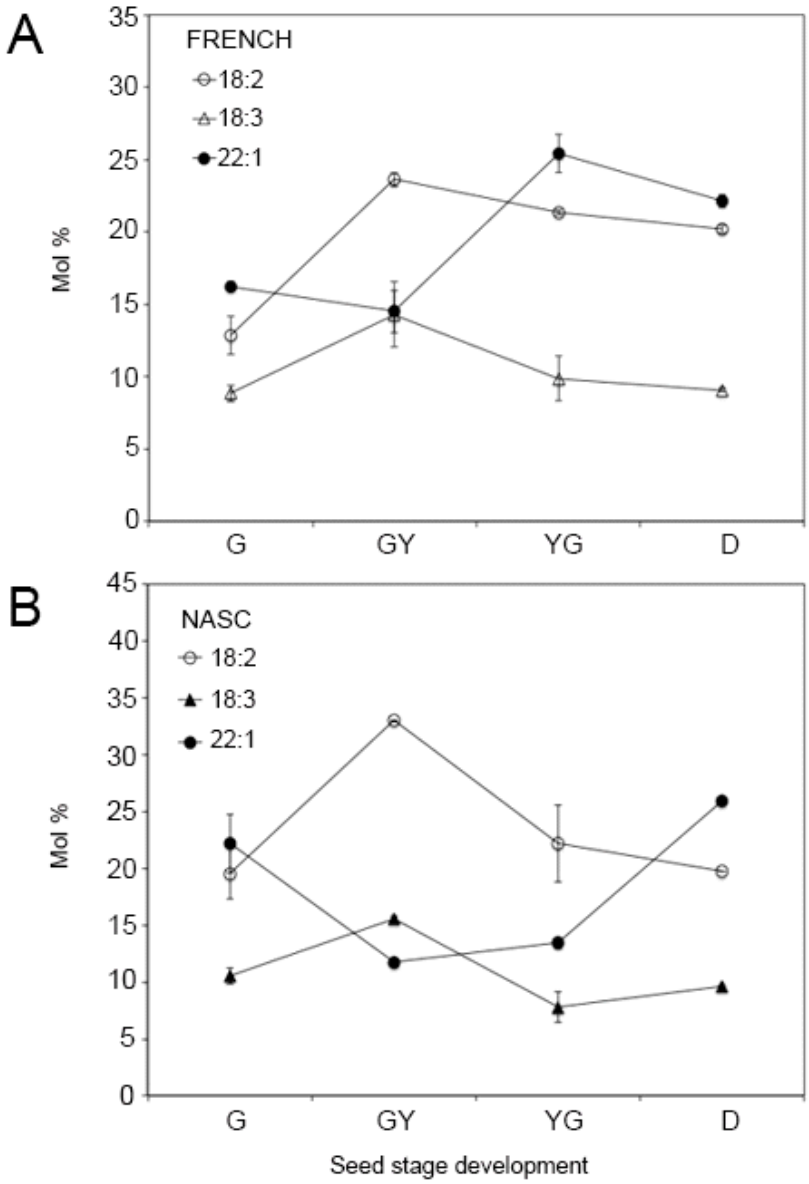

Fig. 4. Fatty acid composition analysis of DAG fraction from Pennycress oil from French (A) and NASC (B) plants. Closed circles, erucic acid, 22:1; open circles, linolenic acid, 18:3; open triangles, linoleic acid, 18:2. Data are expressed as molar percentages from the quantitative analysis of peak areas obtained from the GC analysis. Data were obtained from at least three independent experiments. SD±3.

Fig. 5. Comparison of the protein sequence of the *Thlaspi arvense* FAE1 protein from either NASC and French lines with FAE1 proteins from *Arabidopsis thaliana* (low erucic acid producer) and Brassica napus (High erucic acid producer). Sequences were aligned using CLUSTALW2. GenBank accession numbers are: *Thlaspi arvense* NASC lines, (KT223024); *Thlaspi arvense* French lines, (KT223025); *Arabidopsis thaliana* (U29142.1) and *Brassica napus* (U50771.1). Identical residues are indicated by an asterisk below the sequence. Conservative substitutions were indicated as a double point below the sequence. Residues involved in the active site of the enzyme are indicated with a cross above the sequence. Sequence motifs related with the high erucic acid trait are indicated by a line over the sequence.

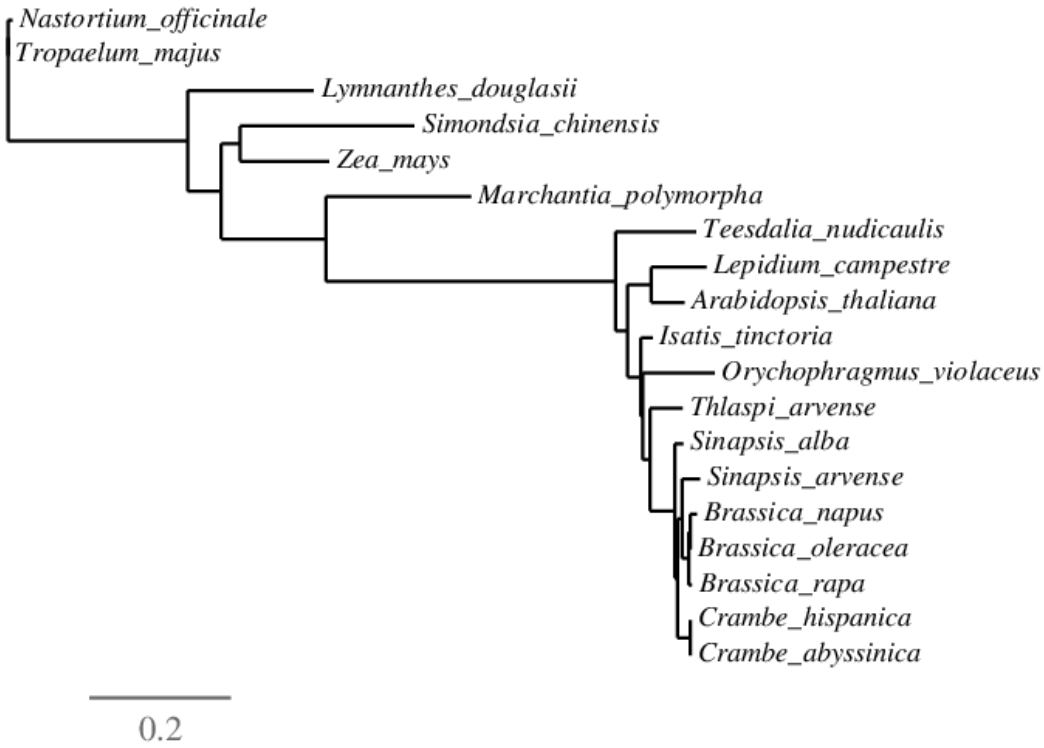

Fig. 6. Phylogenetic tree of the FAE1 protein from different plant species. The names and Genbank accession numbers of the sequences used for the analysis are: *Arabidopsis thaliana* (U29142.1); *Brassica napus* (U50771.1); *Brassica oleracea* (GU325726.1); *Brassica rapa* (GU325723.1); *Crambe hispanica* (KF664157); *Crambe abyssinica* (KC565738); *Isatis tinctoria* (AY888038.1); *Lepidium campestre* (FJ907545.1); *Limnanthes douglasii* (AF247134); *Marchantia polymorpha* (AY308831); *Nastortium officinale* (AY0826190); *Orychophragmus violaceus* (AY888042.1)*; Simondsia chinensis* (U37088); *Sinapsis arvense* (AY888041.1); *Sinapsis alba* (AY888040.1); *Teesdalia nudicaulis* (*EF186003.1); Thlaspi arvense* (KT223024); *Tropaelum majus* (AY082610); *Zea mays* (AJ29770).

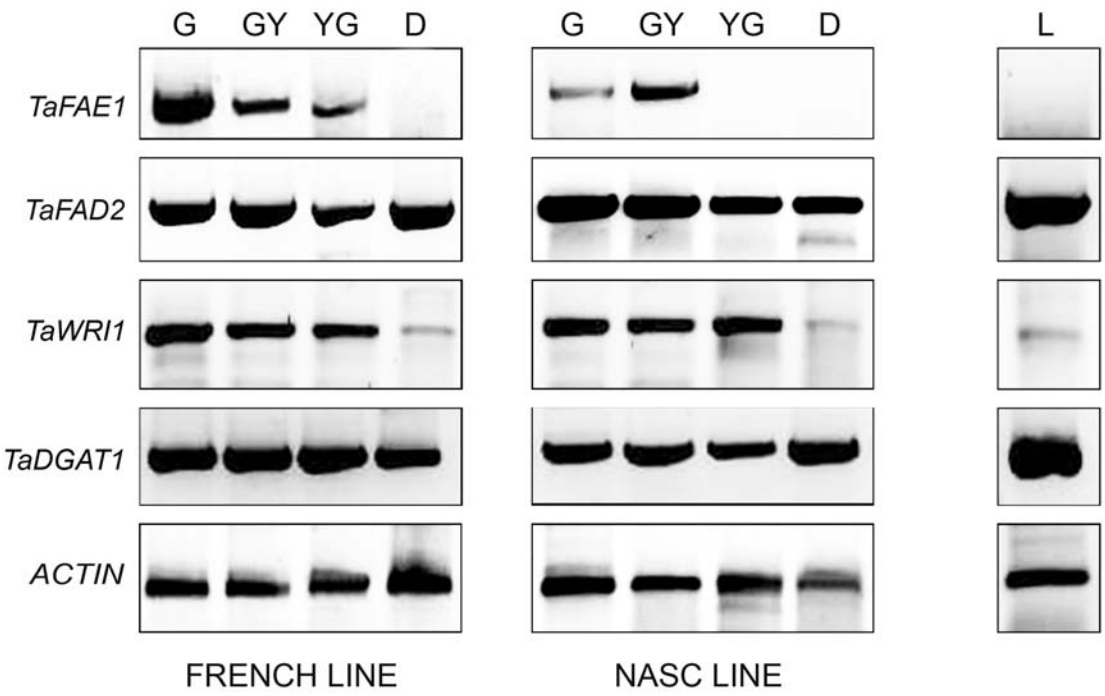

Figure 7. *TaFAE1, TaFAD2, TaDGAT1 and TaWRI1* gene expression during seed maturation in both Pennycress lines. Bold letters correspond to the different seed maturation stages tested: Green (G), Green-Yellow (GY), Yellow-Green (YG) and Dry (D). (L) shows the expression in leaf tissue. ACTIN was used as housekeeping gene. All PCRs were done under non-saturating conditions.