PCP
PLANT & CELL PHYSIOLOGY

Simultaneous Targeting of Multiple Gene Homeologs to Alter Seed Oil Production in *Camelina sativa*

J. A. Aznar-Moreno and T. P. Durrett*

Department of Biochemistry and Molecular Biophysics, Kansas State University, Manhattan, KS 66506, USA
*Corresponding author: E-mail, tdurrett@ksu.edu; Fax: +1-785-532-7278.
(Received September 21, 2016; Accepted April 14, 2017)

Regular Paper

The ability to transform *Camelina sativa* easily with biosynthetic enzymes derived from other plants has made this oil seed crop an ideal platform for the production of unusual lipids valuable for different applications. However, in addition to expressing transgenic enzymes, the suppression of endogenous enzyme activity to reduce competition for common substrates or cofactors is also required to enhance the production of target compounds. As camelina possesses a relatively undifferentiated hexaploid genome, up to three gene homeologs can code for any particular enzymatic activity, complicating efforts to alter endogenous biosynthetic pathways. New genome editing technologies, such as that offered by the CRISPR/Cas (clustered regularly interspaced short palindromic repeats/CRISPR-associated protein) system, offer the capability to introduce mutations into specifically targeted genomic sites. Here, by using a carefully designed guide RNA identical to all three homeologs, we demonstrate the ability of the CRISPR/Cas genome editing system to introduce mutations in all three *CsDGAT1* or *CsPDAT1* homeologous genes important for triacylglycerol (TAG) synthesis in developing seeds. Sequence analysis from transgenic T_1 plants revealed that each *CsDGAT1* or each *CsPDAT1* homeolog was altered by multiple mutations, resulting in a genetic mosaic in the plants. Interestingly, seed harvested from both *CsDGAT1*- and *CsPDAT1*-targeted lines was often shrunken and wrinkled. Further, lipid analysis revealed that many lines produced seed with reduced oil content and altered fatty acid composition, consistent with the role of the targeted genes in seed oil biosynthesis. The CRISPR/Cas system therefore represents a useful method to alter endogenous biosynthetic pathways efficiently in polyploid species such as camelina.

Keywords: Camelina • CRISPR/Cas • Diacylglycerol acyltransferase (DGAT1) • Oil seeds • Phospholipid:diacylglycerol acyltransferase (PDAT1) • Triacylglycerol (TAG).

Abbreviations: Cas, CRISPR-associated protein; CRISPR, clustered regularly interspaced short palindromic repeats; DGAT, diacylglycerol acyltransferase; FAD, fatty acid desaturase; FAE, fatty acid elongase; FAME, fatty acid methyl ester; PAM, protospacer adjacent motif; PDAT, phospholipid:diacylglycerol acyltransferase; RNAi, RNA interference; sgRNA, single guide RNA; SNP, single nucleotide polymorphism; TAG, triacylglycerol; TALEN, transcription activator-like effector nuclease; ZFN, zinc finger nuclease.

Introduction

Seed oils represent a promising source of renewable biofuel and bioproducts (Durrett et al. 2008, Dyer et al. 2008). However, most commercially produced oil seeds synthesize a relatively small range of fatty acids offering limited functionality. In contrast, the seeds of many other plant species produce triacylglycerols (TAGs) containing unusual fatty acids that are valuable as industrial feedstocks (Badami and Patil 1980). For example, *Euonymus alatus* (Burning Bush) seeds synthesize 3-acetyl-1,2-diacylglycerols (acetyl-TAGs), unusual TAG molecules that possess an *sn*-3 acetate group instead of a long chain fatty acid (Kleiman et al. 1967, Durrett et al. 2010). The presence of the acetate group means that acetyl-TAGs possess different chemical and physical properties compared with typical TAGs. For example, acetyl-TAGs have a lower kinematic viscosity and crystallize at lower temperatures, making them useful for different applications (Durrett et al. 2010, Liu et al. 2015a, Liu et al. 2015b). However, like many of the other plants that produce unusual lipids, *E. alatus* possesses undesirable agronomic features that make it unsuitable for commercial production. Instead, the identification of enzymes responsible for the synthesis of unusual fatty acids has enabled the metabolic engineering of transgenic oil seed crops to produce these useful compounds.

In this regard, the oil seed crop *Camelina sativa* has emerged as an ideal platform for the metabolic engineering and field production of useful lipids (Bansal and Durrett 2016). Importantly, camelina is easily transformed using an *Agrobacterium* floral dip (Lu and Kang 2008), allowing the expression of novel enzymes necessary for the synthesis of unusual lipids. Further, the availability of important genomic resources including a fully sequenced camelina genome (Kagale et al. 2014), a seed transcriptome (Nguyen et al. 2013) and an expression atlas (Kagale et al. 2016) allow the easy translation of abundant knowledge from the closely related model plant *Arabidopsis thaliana*. Thus, camelina lines capable of accumulating acetyl-TAGs have been developed by expressing the *Ea*DAcT acetyltransferase isolated from *E. alatus* (Liu et al. 2015a, Liu et al. 2015b). Likewise, through

Plant Cell Physiol. 58(7): 1260–1267 (2017) doi:10.1093/pcp/pcx058, Advance Access publication on 21 April 2017, available online at www.pcp.oxfordjournals.org

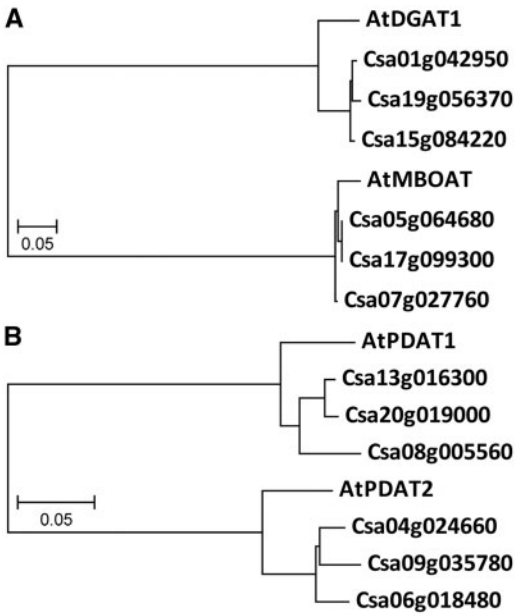

Fig. 1 Identification of DGAT1 and PDAT1 homeologous sequences in *Camelina sativa*. Phylogenetic relationships among *At*DGAT1 (A) and *At*PDAT1 (B) and closely related sequences in camelina and Arabidopsis were determined using ClustalX. *At*MBOAT (At1g57600) and *At*PDAT2 (At3g44830) are the Arabidopsis paralogs most similar to *At*DGAT1 and *At*PDAT1, respectively, and were included for comparison. Horizontal lines represent 0.05 nucleotide substitutions per site.

the transformation of biosynthetic genes from other species, transgenic camelina lines able to produce oils containing wax esters (Iven et al. 2016, Zhu et al. 2016), ricinoleic acid (Lu and Kang 2008, Snapp et al. 2014), omega-7 fatty acids (Nguyen et al. 2015) and medium chain fatty acids (Kim et al. 2015a, Kim et al. 2015b) have been created.

Enhancing the production of unusual lipids in transgenic plants often requires the elimination or reduction of endogenous enzymes that compete with the introduced biosynthetic activity (van Erp et al. 2015). For example, while expression of *Ea*DAcT in wild-type Arabidopsis plants resulted in levels of acetyl-TAGs up to 45 mol% (Durrett et al. 2010), these levels were increased up to 65 mol% by expressing *Ea*DAcT in the Arabidopsis *dgat1* mutant (Liu et al. 2015a). However, well characterized mutant collections such as those developed for Arabidopsis are not available for most species, including camelina. Further, as many important crop plants are polyploid species, genetic redundancy requires mutating multiple genes. In this regard, as camelina possesses a relatively undifferentiated hexaploid genome (Kagale et al. 2014), most enzyme activities are encoded by three similar homeologous genes. Thus camelina possesses three genes (Csa19g056370, Csa01g042590 and Csa15g084220) all highly similar to the Arabidopsis *DGAT1* gene (**Fig. 1A**) which encodes the diacylglycerol acyltransferase enzyme that catalyzes the last step in TAG biosynthesis (Routaboul et al. 1999, Zou et al. 1999, Zhang et al. 2009). Likewise, the camelina genome contains three homeologs (Csa13g016300, Csa20g019000 and Csa08g005560) for phospholipid:diacylglycerol acyltransferase (PDAT1; **Fig. 1B**) shown to also play an important part in seed oil accumulation in different species (Zhang et al. 2009, Pan et al. 2013). Similar observations in camelina have been made for other lipid biosynthesis-related genes such as *FATTY ACID DESATURASE 2* (*FAD2*) and *FATTY ACID ELONGASE 1* (*FAE1*) where all three homeologs of both genes are expressed in developing seeds (Hutcheon et al. 2010). Further, in vitro assays and the partial phenotype of mutants in individual *CsFAD2* homeologs suggest that all three homeologs encode functional enzymes that contribute additively to overall desaturase activity in developing seeds (Kang et al. 2011).

This redundant hexaploid nature of camelina has been overcome using RNA interference (RNAi) to suppress all three homeologous genes. For example, the co-expression of *Ea*DAcT combined with the RNAi-mediated suppression of *Cs*DGAT1, increased the accumulation of acetyl-TAGs in transgenic camelina seed from 64 to 85 mol% (Liu et al. 2015a). However, RNAi-mediated suppression always runs the risk of unintended off-target effects (Xu et al. 2006). Further, the transgenic nature of an RNAi approach means that it can be genetically unstable, as well as less appealing to farmers and markets more receptive to non-transgenic crops.

Over the last decade, different genome editing approaches have been developed that enable precise manipulation of specific genomic sequences, allowing the generation of targeted mutants as an alternative to RNAi-mediated gene suppression. Zinc finger nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs) have been successfully used in plants (Podevin et al. 2013, Voytas and Gao 2014). However, these methods have not been widely adopted by the plant research community because both systems require the design and construction of large modular proteins. More recently, the clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated (Cas) endonuclease system has arisen as a powerful genome editing tool (Doudna and Charpentier 2014). Briefly, this approach relies on an engineered single guide RNA (sgRNA) complementary to a specific DNA target located next to a protospacer adjacent motif (PAM). The Cas nuclease recognizes the sgRNA structure and introduces double strand breaks into the target DNA sequence. Non-homologous end-joining or homologous recombination repairs the double-stranded break, often leading to deletions or insertions that disrupt the target DNA. Recently, CRISPR/Cas has been shown to mediate targeted mutagenesis in different species, including various crops (Belhaj et al. 2013, Schaeffer and Nakata 2015). However, the ability to target multiple homeologs using a sgRNA sequence has not been tested.

Here we used CRISPR/Cas specifically to induce mutations in two sets of homeologous genes in camelina, *CsDGAT1* and *CsPDAT1*. In Arabidopsis, the closest homeologs of these genes are important for the last step in the synthesis of TAG in seeds (Zhang et al. 2009). Using a sgRNA sequence specific for identical regions of all three *CsDGAT1* or *CsPDAT* homeologs, we were able to introduce mutations in all three gene copies in the same plant. Sequence analysis indicated that multiple types of mutations were introduced into each homeolog, resulting in a genetic mosaic in T_1 plants. Importantly, these mutations

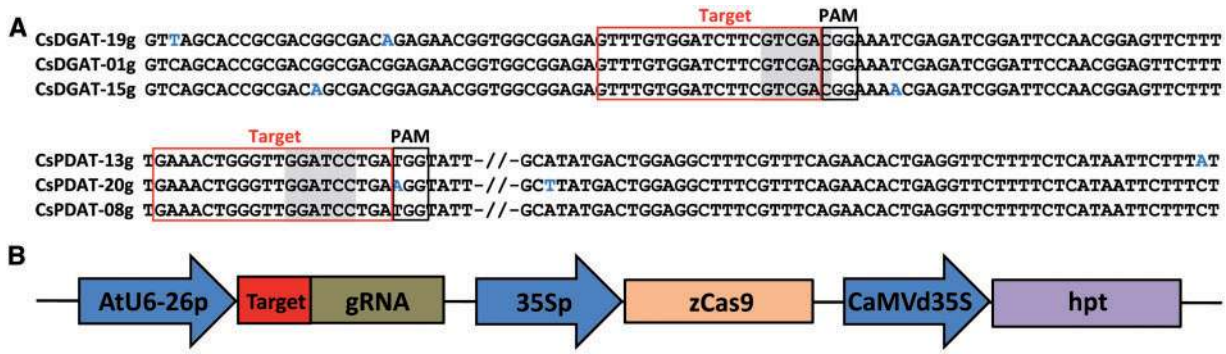

Fig. 2 Vector design for CRISPR/Cas genome editing of *DGAT1* and *PDAT1* genes in *Camelina sativa*. (A) Alignment of camelina *DGAT1* or *PDAT* homeolog sequences. The 20 nucleotide target sequences (red box) contain a restriction site (gray highlight) and are adjacent to the PAM sequence (black box). Single nucleotide polymorphisms (SNPs) that distinguish the homeologs are colored blue. Breaks in the *PDAT* sequence denote the removal of 123 nucleotides identical in all three homeologs so that more distal SNPs can be included in the figure. (B) Each target sequence and the subsequent guide RNA (gRNA) are under control of the Arabidopsis U6-26 promoter. Expression of the maize codon-optimized Cas9 (zCas9) is driven by the 35S promoter. The selectable marker hygromycin phosphotransferase (hpt) conferring hygromycin resistance is under the control of CaMVd35S promoter.

altered the seed morphology and reduced the oil content in the T_2 seed from these transgenic lines, consistent with the role of the targeted genes in seed oil biosynthesis.

Results

Design of camelina *DGAT1* and *PDAT1* guide RNA sequences

We designed the sgRNA sequences to recognize regions identical among all three *CsDGAT1* or *CsPDAT1* homeologs (**Fig. 2A**), allowing the simultaneous targeting of all three gene copies. In addition, each target sequence contained a restriction enzyme recognition site to facilitate identification of altered gene sequences. Target regions were also selected to be close to the 5′ end of each gene, as well as near single nucleotide polymorphisms (SNPs) that would allow each individual homeolog to be identified based on sequence. We also searched the camelina genome to avoid the use of sgRNA sequences that would potentially recognize off-target sites. The final sgRNA sequences were inserted into the binary transformation vector pHSE401 (Xing et al. 2014) under the control of the Arabidopsis RNA polymerase III U6-26 promoter (**Fig. 2B**). This vector also contains a maize codon optimized Cas9 gene driven by a 35S promoter, avoiding the need for transformation of multiple plasmids. After *Agrobacterium*-mediated transformation via vacuum infiltration of flowers (Lu and Kang 2008), approximately 500 T_1 seeds for each construct were screened on hygromycin plates. Six hygromycin-resistant plants from each construct were transferred to soil, after which they were genotyped to confirm the presence of Cas9 (Supplementary Fig. S1).

T_1 plants contain restriction enzyme-resistant DNA target sequences

To determine whether the introduced CRISPR/Cas systems were functioning, genomic DNA from the leaf tissue of the T_1 plants was extracted and digested with the appropriate restriction enzyme present within each target sequence (*Sal*I for *CsDGAT1*-targeted plants; *Bam*HI for *CsPDAT1*-targeted plants). The digested genomic DNA was used for PCR amplification with primers specific for each homeolog of *CsDGAT1* or *CsPDAT1* (Supplementary Table S1). These primers were designed specifically to amplify the region of the gene that contained the target sequence, as well as unique SNPs that would confirm the identity of each homeolog. Because the restriction site was included in the target sequence, only DNA from genes where the site had been edited could be amplified. Whereas no amplification occurred in digested DNA from wild-type plants, all three *CsDGAT1* or *CsPDAT1* homeologs were successfully amplified in the transgenic lines (**Fig. 3**). These results suggest that in the *CsDGAT1*- or *CsPDAT1*-targeted plants, the restriction enzyme recognition sequence has been altered, rendering the DNA resistant to cutting and allowing PCR amplification. In many of the transgenic lines, all three *CsDGAT1* or *CsPDAT1* homeologs contain restriction enzyme-resistant sequences, demonstrating that it is possible to target all three gene copies with an sgDNA sequence. However, when genomic DNA without prior restriction enzyme digestion was used as template for homeolog-specific PCR amplification, more PCR product was formed than when the gDNA was digested (**Fig. 3**). This result suggests that despite changes in some of the target genes, more wild-type copies of the *CsDGAT1* or *CsPDAT1* homeolog are present relative to the altered alleles.

We also analyzed the effect of the CRISPR/Cas system in the subsequent T_2 progeny. In the plants expressing Cas9, we detected restriction enzyme-resistant sequences (Supplementary Fig. S2). However, none of the nine Cas9-negative T_2 plants (four from *CsDGAT1*-targeted plants; five from *CsPDAT1*-targeted plants) appeared to contain mutations in their appropriate target genes (Supplementary Fig. S2). Interestingly, some of the *CsPDAT1*-targeted plants possessed aberrant phenotypes (Supplementary Fig. S3), consistent with studies in Arabidopsis indicating a role for PDAT1 in vegetative tissue (Fan et al. 2013).

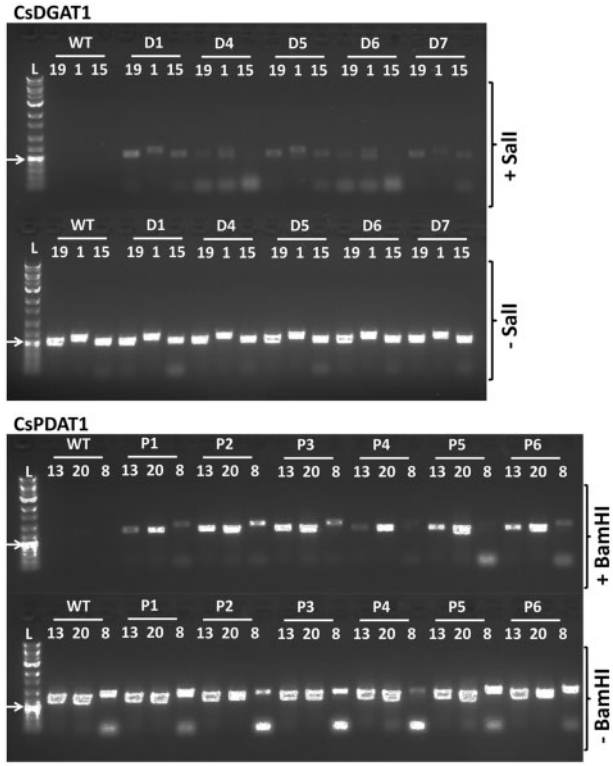

Fig. 3 Transgenic T_1 plants possess modified *CsDGAT1* or *CsPDAT1* genes. Genomic DNA extracted from leaf tissue from T_1 transgenic lines with (upper panel) and without (lower panel) digestion with *Sal*I (*CsDGAT1*) or *Bam*HI (*CsPDAT1*) was amplified with homeolog-specific primers. Individual homeologs are indicated by the number of the chromosome on which they are located. Wild-type (WT) camelina genomic DNA was used as a negative control. Arrows adjacent to the DNA size ladders (L) indicate 500 bp.

Multiple *DGAT1* or *PDAT1* mutations are present in Cas9 transgenic lines

To confirm that the restriction enzyme-resistant amplicons (**Fig. 3**) were the result of Cas9-mediated DNA editing, the PCR products were directly sequenced. Interestingly, the sequence quality of the reads was dramatically reduced for all the PCR products in the region targeted by the sgRNA (Supplementary Fig. S4). We hypothesized that this poor sequence quality was due to the presence of multiple DNA sequence variants caused by different mutations present in individual gene sequences. To confirm this hypothesis, homeolog-specific PCR products were cloned and subsequently sequenced. An alignment of the sequences confirmed that multiple insertions/deletions (indels), probably introduced by error-prone DNA repair, were present in each of the *DGAT1* or *PDAT1* homeologs (**Fig. 4**). The sequencing data also provide additional evidence that all three *DGAT1* or *PDAT1* homeologs in a particular transgenic line were simultaneously targeted for mutation. For the most part, these indels were relatively small, though a large 16 bp deletion and a 31 bp insertion were particularly noticeable in two of the *PDAT1* mutant alleles. The mutations all occurred at the 3′ end of the target region, consistent with Cas9 inducing a double-stranded break upstream of the PAM (Jinek et al. 2012).

T_2 seeds from transgenic lines possess aberrant phenotypes

Many of the T_1 transgenic lines targeting the *CsDGAT1* or *CsPDAT1* homeologs produced T_2 seed that was wrinkled and darker in appearance than wild-type seeds. (Supplementary Fig. S5). Further, at least half of the *CsDGAT1*- and *CsPDAT1*-targeted lines possessed smaller T_2 seed compared with seed from wild-type plants (**Table 1**). These results for the *DGAT1* mutated lines are consistent with observations that Arabidopsis *dgat1* mutants possess wrinkled seeds (Routaboul et al. 1999). However, while at least two of the *CsPDAT1*-targeted lines produced aberrant and smaller T_2 seed (**Table 1**; Supplementary Fig. S5), no phenotype was found in Arabidopsis *pdat1* T-DNA insertion lines (Mhaske et al. 2005). Although all lines displayed restriction enzyme-resistant amplicons (**Fig. 3**), *CsDGAT1*-targeted line D4 and *CsPDAT1*-targeted line P1 demonstrated the strongest seed phenotype with the most shrunken seeds (Supplementary Fig. S5).

T_2 seeds from *CsDGAT1*- or *CsPDAT1*-targeted lines possess altered oil content

Because DGAT1 and PDAT1 enzymes are important for the synthesis of seed oil in other species (Katavic et al. 1995, Routaboul et al. 1999, Zou et al. 1999, Zhang et al. 2009, Pan et al. 2013), we quantified the fatty acid content and composition in seed from the mutated lines. Similar to observations in Arabidopsis, seed from *CsDGAT1*-targeted transgenic lines tended to possess a lower oil content, with line D showing a 31% reduction relative to the wild type (**Fig. 5A**). The oil content of some *CsPDAT1*-targeted lines (lines P1 and P6) was also reduced relative to the wild-type control. Only seed from *CsDGAT1*-targeted plants displayed different fatty acid profiles, with some lines possessing higher levels of linoleic acid (18:2) instead of linolenic acid (18:3). In contrast, seed from lines with mutations in *CsPDAT1* homeologs possessed a fatty acid profile similar to that of wild-type seed (**Fig. 5B**; Supplementary Table S2).

Discussion

Camelina is an allohexaploid species with three highly related and undifferentiated subgenomes (Kagale et al. 2014). Therefore, three homeologous copies exist for many genes. In many cases, it appears that all three homeologs are expressed and contribute to enzyme activity. For example, all three *CsFAE1* homeologs encode functional enzymes, and mutations in individual genes only result in a partial phenotype (Kang et al. 2011). This genetic redundancy means that suppressing or eliminating endogenous enzyme activity requires targeting all three encoding genes. Here we demonstrate that all three of the camelina *DGAT1* homeologs (Csa19g056370, Csa01g042590; Csa15g084220; **Fig. 1A**) or *PDAT1* homeologs (Csa13g016300, Csa20g019000, Csa08g005560; **Fig. 1B**) could be mutated using

Fig. 4 Individual homeologs in CRISPR/Cas lines contain multiple types of mutations. Sequences of CRISPR/Cas-induced mutations in T_1 transgenic lines targeting *CsDGAT1* or *CsPDAT1*. Target regions are underlined and the PAM is highlighted in gray. Insertions and deletions are indicated by red text and dashes, respectively. Single nucleotide polymorphisms (SNPs) that distinguish the homeologs are represented by blue letters. Nucleotide changes (+, insertion; −, deletion) are indicated to the right of each mutated sequence.

Table 1 Seed size analysis of the different CsDGAT1 and CsPDAT1 CRISPR/Cas9 transgenic lines

Line	Average (mm)	SD
Wild type	1.77	0.20
DGAT1-D1*	1.56	0.14
DGAT1-D4*	1.50	0.23
DGAT1-D5*	1.59	0.17
DGAT1-D6	1.75	0.17
DGAT1-D7	1.72	0.21
PDAT1-P1*	1.68	0.30
PDAT1-P2	1.73	0.21
PDAT1-P3*	1.71	0.19
PDAT1-P4	1.78	0.19
PDAT1-P5	1.75	0.18
PDAT1-P6*	1.66	0.21

*Statistically significantly difference from wild-type seed (Student's *t*-test. $P < 0.05$).

Cas9 and a sgRNA (**Figs. 3**, **4**). This ability to target all three homeologs simultaneously is perhaps not surprising given that previous studies have shown that similar gene paralogs in rice could be edited by targeting regions of identity (Endo et al. 2015). Our result provides evidence that targeting multiple functional homeologs in polyploid species is feasible and overcomes one of the main challenges with altering enzyme activity in camelina.

Poor sequence quality in the target site (Supplementary Fig. S4) and individual gene sequences (**Fig. 4**) suggests that a variety of gene editing events are occurring in the different cells from which the genomic DNA was extracted for analysis. The considerably larger amount of total amplifiable DNA than the restriction enzyme-resistant amplifiable DNA (**Fig. 3**) suggests that editing events are only occurring in a minority of the targeted genes in all cells and might reflect the larger genome size of camelina. The leaf tissue from which the DNA was extracted therefore exists as a genetic mosaic, with different cells

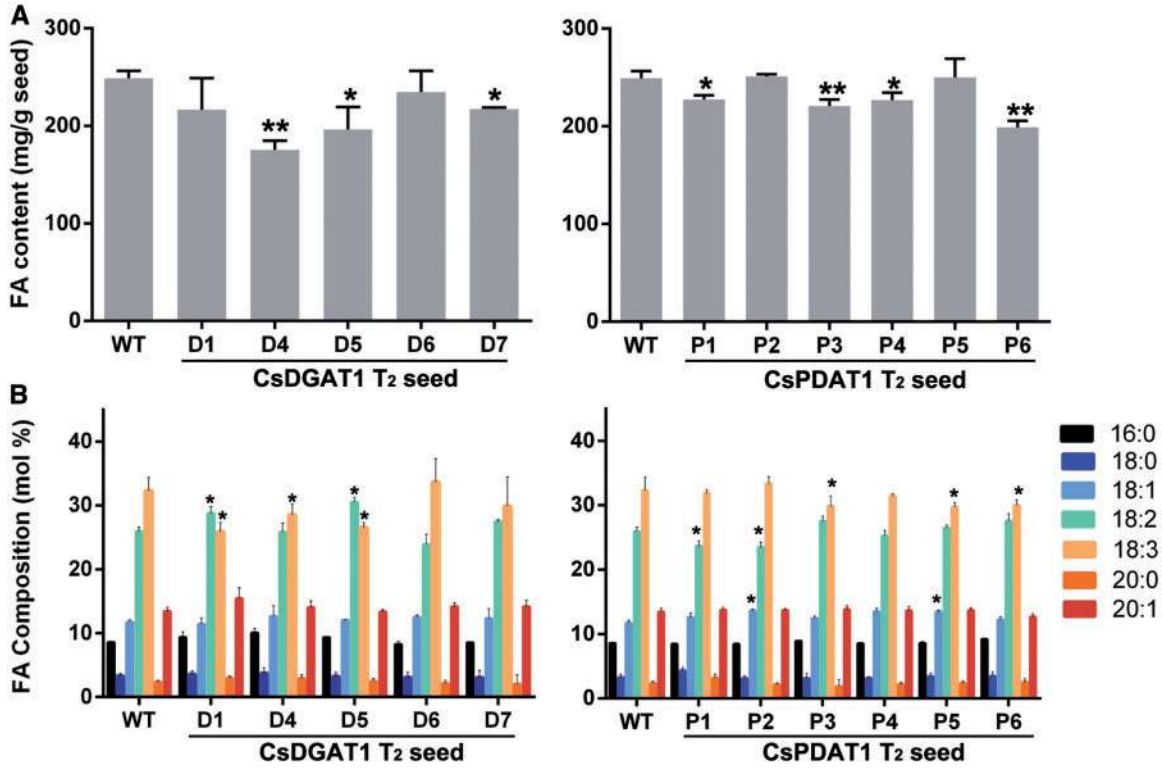

Fig. 5 T_2 seed from *CsPDAT1* and *CsDGAT1* CRISPR/Cas lines possess reduced oil content. Mean fatty acid content (A) and composition (B) of T_2 seed harvested from transgenic lines and wild-type (WT) control plants grown together. Error bars represent the SEM for three independent analyses. Asterisks indicate statistically significant differences (Student's *t*-test with Holm–Sidak correction for multiple comparisons; $^*P < 0.05$; $^{**}P < 0.01$).

containing either wild-type versions of the three *CsDGAT1* or *CsPDAT1* homeologs or different mutations in at least one of the targets. Currently we cannot distinguish between different possibilities where cells might only contain a mutation in one of the homeologs, leaving two unaltered (which would be consistent with the large proportion of non-edited DNA) or whether a minority of cells contain mutations in more than one homeolog, with the majority of cells being unaltered by Cas9. Here, the use of egg cell-specific promoters to drive the expression of Cas9 and sgRNA in germline cells offers a promising approach to minimize the generation of plants with a genetic mosaic in the target gene, as well as enable the efficient generation of mutant lines where Cas9 has been segregated away (Wang et al. 2015, Mao et al. 2016).

Despite the presumed mosaic nature of the gene editing events, seed from plants where either *CsDGAT1* or *CsPDAT1* was targeted possessed altered physiology (Supplementary Fig. S5) and reduced oil content (**Fig. 5**). As all the observed mutations result in coding frameshifts close to the 5′ end of the gene (**Fig. 4**), the encoded enzymes are probably non-functional. The wrinkled appearance, reduced oil content and altered fatty acid composition of some *CsDGAT1*-targeted lines therefore are consistent with similar observations of Arabidopsis *dgat1* mutants (Katavic et al. 1995, Routaboul et al. 1999). Interestingly, the seed from some *CsPDAT1*-targeted lines also possessed altered physiology and reduced oil content. These phenotypes contrast with the lack of an observable seed phenotype in Arabidopsis *pdat1* T-DNA insertion mutants (Mhaske et al. 2005). One explanation could simply be that PDAT1 activity plays a more important role in TAG synthesis in other species such as camelina, than it does in Arabidopsis. In flax (*Linum usitatissimum*), for example, seed oil is mainly synthesized by a pair of PDATs which are predominantly expressed during seed development and display high selectivity for substrates containing 18:3 (Pan et al. 2013) Because camelina oil also contains high level of the polyunsaturated fatty acid 18:3, our results here might suggest a similar important role for CsPDAT1 homeologs in seed oil biosynthesis.

In summary, we have demonstrated the feasibility of using the CRISPR/Cas genome editing system as an effective tool to target endogenous biosynthetic pathways in an oilseed crop, thus providing a more suitable background for the production of unusual lipids. Importantly, as camelina is a hexaploid species, we show that the CRISPR/Cas system is able to edit three homeologous genes with a single target sequence. Phenotypes of reduced oil content and wrinkled seeds in T_2 seeds are consistent with the role of DGAT1 and PDAT1 in synthesizing TAG and demonstrate that this genome editing system can be used to alter camelina biosynthetic pathways for the production of unusual lipids.

Materials and Methods

Phylogenetic analysis

Sequences for camelina *DGAT1* and *PDAT1* homeologs were identified using TBLASTN to search the currently available camelina genome (http://www.camelinadb.ca/) with the Arabidopsis

DGAT1 (At2g19450) and PDAT1 (At5g13640) protein sequences. Translated amino acid sequences were aligned using ClustalX v.2.0.10 under the default settings (Larkin et al. 2007), and the alignments used to generate a phylogenetic tree with a Neighbor–Joining algorithm (Saitou and Nei 1987); the final phenogram was created with MEGA 5.0 (Tamura et al. 2011). Sequences slightly more distantly related to *At*DGAT1 (*At*MBOAT, At1g57600; camelina homeologs Csa07g027760, Csa17g099300 and Csa05g064680) and *At*PDAT1 (*At*PDAT2, At3g44830; camelina homeologs Csa04g024660, Csa09g035780 and Csa06g018480) were included to compare the homology analysis.

Plasmid construction and generation of transgenic plants

Guide RNA sequences were amplified with specific oligonucleotides (Supplementary Table S3) from genomic DNA using Phusion DNA polymerase (ThermoScientific) and cloned into the vector pHSE401 (Xing et al. 2014) using a PCR splicing approach (Li et al. 2013) as outlined in Supplementary Fig. S6. The fidelity of the introduced sequence was confirmed by sequencing before electroporation of the plasmids into *Agrobacterium tumefaciens* strain GV3101. Wild-type *Camelina sativa* 'Suneson' plants were transformed using an *Agrobacterium*-mediated floral vacuum infiltration method (Lu and Kang 2008). T_1 seeds were selected on Murashige and Skoog (MS) plates containing 25 mg l^{-1} hygromycin, and resistant seedlings were transferred to soil.

Genotyping of transgenic camelina plants

Genomic DNA from camelina leaves was extracted according to Edwards et al. (1991). Digestion of 200 ng of genomic DNA with 20 U of *Sal*I or 20 U of *Bam*HI (New England Biolabs) was carried out overnight at 37°C. Target sites were amplified by PCR using primers specific for each *CsDGAT1* or *CsPDAT1* homeolog (Supplementary Table S1). Restriction enzyme-resistant PCR products were excised from the gel, purified from the agarose and ligated into pGEM-T easy (Promega). Plasmid DNA was extracted from individual transformants and sequenced using a T7 universal primer (Genewiz).

Seed analyses

Approximately 100 randomly selected seeds were photographed with a Leica M165FC microscope. Seed sizes were calculated using imageJ software (version 1.48v) following a previously established method (Herridge et al. 2011). Fatty acid methyl esters (FAMEs) were prepared using a well-established method (Miquel and Browse 1992) with minor modifications. Twenty randomly picked seeds per replicate were quickly homogenized with a polytron (PT2500E, Kinematica AGd) in 1.5 ml of toluene. A 100 µg aliquot of triheptadecanoin was added as an internal standard and total lipids were transmethylated for 1 h at 80°C with 3 ml of 2.5% (v/v) sulfuric acid in methanol. After cooling, 2 ml of water and 1 ml of hexane were added, and FAMEs extracted in the organic fraction were analyzed by gas chromatography with a flame ionization detector on a 30 × 0.25 mm (internal diameter) DB-23 column (Agilent Technologies), using the following program: initial temperature, 190°C for 2 min; ramp increase at 5°C min^{-1} to 240°C; final temperature hold, 4 min. FAMEs were identified by comparison with a 37 component standard mix (Supelco # 47885-U) and chromatogram peak areas were corrected for FID response as described previously (Li et al. 2006).

Supplementary data

Supplementary data are available at PCP online.

Funding

This work was supported by the U.S. Department of Agriculture [under Award Number 2015-67013-22815].

Acknowledgments

We would like to thank Garrison Olds for assistance with careful plant care and seed harvesting. Contribution number 17-323-J from the Kansas Agricultural Experiment Station.

Disclosures

The authors have no conflicts of interest to declare.

References

Badami, R.C. and Patil, K.B. (1980) Structure and occurrence of unusual fatty acids in minor seed oils. *Prog. Lipid Res.* 19: 119–153.

Bansal, S. and Durrett, T.P. (2016) *Camelina sativa*: an ideal platform for the metabolic engineering and field production of industrial lipids. *Biochimie* 120: 9–16.

Belhaj, K., Chaparro-Garcia, A., Kamoun, S. and Nekrasov, V. (2013) Plant genome editing made easy: targeted mutagenesis in model and crop plants using the CRISPR/Cas system. *Plant Methods* 9: 39.

Doudna, J.A. and Charpentier, E. (2014) The new frontier of genome engineering with CRISPR–Cas9. *Science* 346: 1258096.

Durrett, T.P., Benning, C. and Ohlrogge, J. (2008) Plant triacylglycerols as feedstocks for the production of biofuels. *Plant J.* 54: 593–607.

Durrett, T.P., McClosky, D.D., Tumaney, A.W., Elzinga, D.A., Ohlrogge, J. and Pollard, M. (2010) A distinct DGAT with *sn*-3 acetyltransferase activity that synthesizes unusual, reduced-viscosity oils in *Euonymus* and transgenic seeds. *Proc. Natl. Acad. Sci. USA* 107: 9464–9469.

Dyer, J.M., Stymne, S., Green, A.G. and Carlsson, A.S. (2008) High-value oils from plants. *Plant J.* 54: 640–655.

Edwards, K., Johnstone, C. and Thompson, C. (1991) A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Nucleic Acids Res.* 19: 1349.

Endo, M., Mikami, M. and Toki, S. (2015) Multigene knockout utilizing off-target mutations of the CRISPR/Cas9 system in rice. *Plant Cell Physiol.* 56: 41–47.

Fan, J., Yan, C. and Xu, C. (2013) Phospholipid:diacylglycerol acyltransferase-mediated triacylglycerol biosynthesis is crucial for protection against fatty acid-induced cell death in growing tissues of Arabidopsis. *Plant J.* 76: 930–942.

Herridge, R.P., Day, R.C., Baldwin, S. and Macknight, R.C. (2011) Rapid analysis of seed size in Arabidopsis for mutant and QTL discovery. *Plant Methods* 7: 1–11.

Hutcheon, C., Ditt, R., Beilstein, M., Comai, L., Schroeder, J., Goldstein, E., et al. (2010) Polyploid genome of *Camelina sativa* revealed by isolation of fatty acid synthesis genes. *BMC Plant Biol.* 10: 233.

Iven, T., Hornung, E., Heilmann, M. and Feussner, I. (2016) Synthesis of oleyl oleate wax esters in *Arabidopsis thaliana* and *Camelina sativa* seed oil. *Plant Biotechnol. J.* 14: 252–259.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A. and Charpentier, E. (2012) A Programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337: 816–821.

Kagale, S., Koh, C., Nixon, J., Bollina, V., Clarke, W.E., Tuteja, R., et al. (2014) The emerging biofuel crop *Camelina sativa* retains a highly undifferentiated hexaploid genome structure. *Nature Communications* 5: 3706.

Kagale, S., Nixon, J., Khedikar, Y., Pasha, A., Provart, N.J., Clarke, W.E., et al. (2016) The developmental transcriptome atlas of the biofuel crop *Camelina sativa*. *Plant J.* 88: 879–894.

Kang, J., Snapp, A.R. and Lu, C. (2011) Identification of three genes encoding microsomal oleate desaturases (FAD2) from the oilseed crop *Camelina sativa*. *Plant Physiol. Biochem.* 49: 223–229.

Katavic, V., Reed, D.W., Taylor, D.C., Giblin, E.M., Barton, D.L., Zou, J.T., et al. (1995) Alteration of seed fatty acid composition by an ethyl methanesulfonate-induced mutation in *Arabidopsis thaliana* affecting diacylglycerol acyltransferase activity. *Plant Physiol.* 108: 399–409.

Kim, H.J., Silva, J.E., Iskandarov, U., Andersson, M., Cahoon, R.E., Mockaitis, K., et al. (2015a) Structurally divergent lysophosphatidic acid acyltransferases with high selectivity for saturated medium chain fatty acids from Cuphea seeds. *Plant J.* 84: 1021–1033.

Kim, H.J., Silva, J.E., Vu, H.S., Mockaitis, K., Nam, J.-W. and Cahoon, E.B. (2015b) Toward production of jet fuel functionality in oilseeds: identification of FatB acyl-acyl carrier protein thioesterases and evaluation of combinatorial expression strategies in Camelina seeds. *J. Exp. Bot.* 66: 4251–4265.

Kleiman, R., Miller, R.W., Earle, F.R. and Wolff, I.A. (1967) (*S*)-1,2-diacyl-3-acetins: optically active triglycerides from *Euonymus verrucosus* seed oil. *Lipids* 2: 473–478.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., et al. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* 23: 2947–2948.

Li, J.-F., Norville, J.E., Aach, J., McCormack, M., Zhang, D., Bush, J., et al. (2013) Multiplex and homologous recombination-mediated genome editing in *Arabidopsis* and *Nicotiana benthamiana* using guide RNA and Cas9. *Nat. Biotechnol.* 31: 688–691.

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

Liu, J., Rice, A., McGlew, K., Shaw, V., Park, H., Clemente, T., et al. (2015a) Metabolic engineering of oilseed crops top produce high levels of novel acetyl glyceride oils with reduced viscosity, freezing point and calorific value. *Plant Biotechnol. J.* 13: 858–865.

Liu, J., Tjellström, H., McGlew, K., Shaw, V., Rice, A., Simpson, J., et al. (2015b) Field production, purification and analysis of high-oleic acetyl-triacylglycerols from transgenic *Camelina sativa*. *Ind. Crops Prod.* 65: 259–268.

Lu, C. and Kang, J. (2008) Generation of transgenic plants of a potential oilseed crop *Camelina sativa* by *Agrobacterium*-mediated transformation. *Plant Cell Rep.* 27: 273–278.

Mao, Y., Zhang, Z., Feng, Z., Wei, P., Zhang, H., Botella, J.R., et al. (2016) Development of germ-line-specific CRISPR–Cas9 systems to improve the production of heritable gene modifications in Arabidopsis. *Plant Biotechnol. J.* 14: 519–532.

Mhaske, V., Beldjilali, K., Ohlrogge, J. and Pollard, M. (2005) Isolation and characterization of an *Arabidopsis thaliana* knockout line for phospholipid:diacylglycerol transacylase gene (At5g13640). *Plant Physiol. Biochem.* 43: 413–417.

Miquel, M. and Browse, J. (1992) Arabidopsis mutants deficient in polyunsaturated fatty acid synthesis. Biochemical and genetic characterization of a plant oleoyl-phosphatidylcholine desaturase. *J. Biol. Chem.* 267: 1502–1509.

Nguyen, H.T., Park, H., Koster, K.L., Cahoon, R.E., Nguyen, H.T.M., Shanklin, J., et al. (2015) Redirection of metabolic flux for high levels of omega-7 monounsaturated fatty acid accumulation in camelina seeds. *Plant Biotechnol. J.* 13: 38–50.

Nguyen, H.T., Silva, J.E., Podicheti, R., Macrander, J., Yang, W., Nazarenus, T.J., et al. (2013) Camelina seed transcriptome: a tool for meal and oil improvement and translational research. *Plant Biotechnol. J.* 11: 759–769.

Pan, X., Siloto, R.M.P., Wickramarathna, A.D., Mietkiewska, E. and Weselake, R.J. (2013) Identification of a pair of phospholipid:diacylglycerol acyltransferases from developing flax (*Linum usitatissimum* L.) seed catalyzing the selective production of trilinolenin. *J. Biol. Chem.* 288: 24173–24188.

Podevin, N., Davies, H.V., Hartung, F., Nogue, F. and Casacuberta, J.M. (2013) Site-directed nucleases: a paradigm shift in predictable, knowledge-based plant breeding. *Trends Biotechnol.* 31: 375–383.

Routaboul, J.M., Benning, C., Bechtold, N., Caboche, M. and Lepiniec, L. (1999) The TAG1 locus of Arabidopsis encodes for a diacylglycerol acyltransferase. *Plant Physiol. Biochem.* 37: 831–840.

Saitou, N. and Nei, M. (1987) The Neighbor–Joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4: 406–425.

Schaeffer, S.M. and Nakata, P.A. (2015) CRISPR/Cas9-mediated genome editing and gene replacement in plants: Transitioning from lab to field. *Plant Sci.* 240: 130–142.

Snapp, A.R., Kang, J., Qi, X. and Lu, C. (2014) A fatty acid condensing enzyme from *Physaria fendleri* increases hydroxy fatty acid accumulation in transgenic oilseeds of *Camelina sativa*. *Planta* 240: 599–610.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28: 2731–2739.

van Erp, H., Shockey, J., Zhang, M., Adhikari, N.D. and Browse, J. (2015) Reducing isozyme competition increases target fatty acid accumulation in seed triacylglycerols of transgenic Arabidopsis. *Plant Physiol.* 168: 36–46.

Voytas, D.F. and Gao, C.X. (2014) Precision genome engineering and agriculture: opportunities and regulatory challenges. *PLoS Biol.* 12: e1001877.

Wang, Z.-P., Xing, H.-L., Dong, L., Zhang, H.-Y., Han, C.-Y., Wang, X.-C., et al. (2015) Egg cell-specific promoter-controlled CRISPR/Cas9 efficiently generates homozygous mutants for multiple target genes in *Arabidopsis* in a single generation. *Genome Biol.* 16: 1–12.

Xing, H.-L., Dong, L., Wang, Z.-P., Zhang, H.-Y., Han, C.-Y., Liu, B., et al. (2014) A CRISPR/Cas9 toolkit for multiplex genome editing in plants. *BMC Plant Biol.* 14: 1–12.

Xu, P., Zhang, Y., Kang, L., Roossinck, M.J. and Mysore, K.S. (2006) Computational estimation and experimental verification of off-target silencing during posttranscriptional gene silencing in plants. *Plant Physiol.* 142: 429–440.

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

Zhu, L.-H., Krens, F., Smith, M.A., Li, X., Qi, W., van Loo, E.N., et al. (2016) Dedicated industrial oilseed crops as metabolic engineering platforms for sustainable industrial feedstock production. *Sci. Rep.* 6: 22181.

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