Metabolic engineering of hydroxy fatty acid production in plants: RcDGAT2 drives dramatic increases in ricinoleate levels in seed oil

Julie Burgal1,†, Jay Shockey1,‡, Chaofu Lu1,§, John Dyer2, Tony Larson3, Ian Graham3 and John Browse1,*

1Institute of Biological Chemistry, Clark Hall, Washington State University, Pullman, WA 99164-6340, USA
2USDA-ARS, US Arid-Land Agricultural Research Center, 21881 North Cardon Lane, Maricopa, AZ 85238, USA
3Centre for Novel Agricultural Products, Department of Biology, University of York, York YO10 5YW, UK

Summary
A central goal of green chemistry is to produce industrially useful fatty acids in oilseed crops. Although genes encoding suitable fatty acid-modifying enzymes are available from many wild species, progress has been limited because the expression of these genes in transgenic plants produces low yields of the desired products. For example, Ricinus communis fatty acid hydroxylase 12 (FAH12) produces a maximum of only 17% hydroxy fatty acids (HFAs) when expressed in Arabidopsis. cDNA clones encoding R. communis enzymes for additional steps in the seed oil biosynthetic pathway were identified. Expression of these cDNAs in FAH12 transgenic plants revealed that the R. communis type-2 acyl-coenzyme A:diacylglycerol acyltransferase (RcDGAT2) could increase HFAs from 17% to nearly 30%. Detailed comparisons of seed neutral lipids from the single- and double-transgenic lines indicated that RcDGAT2 substantially modified the triacylglycerol (TAG) pool, with significant increases in most of the major TAG species observed in native castor bean oil. These data suggest that RcDGAT2 prefers acyl-coenzyme A and diacylglycerol substrates containing HFAs, and biochemical analyses of RcDGAT2 expressed in yeast cells confirmed a strong preference for HFA-containing diacylglycerol substrates. Our results demonstrate that pathway engineering approaches can be used successfully to increase the yields of industrial feedstocks in plants, and that members of the DGAT2 gene family probably play a key role in this process.

Introduction
Ricinoleic acid (12-o-hydroxy-octadeca-cis-9-enoic acid: 18:1-OH) is an important natural product with great value as a petrochemical replacement in a variety of industrial processes. Its derivatives are found in products such as lubricants, nylon, dyes, soaps, inks, adhesives and bio-diesel (Caupin, 1997). Currently, the only commercial source of ricinoleic acid is the castor bean (Ricinus communis) seed, where it makes up approximately 90% of the total fatty acids in the oil. Industrial processes consume approximately 300 000 tonnes of castor oil each year, in more than 100 different chemical processes, and demand is growing rapidly. However, the castor bean plant is not suited to large-scale agricultural production and seeds must be harvested by hand. Currently, castor oil is produced commercially from plants grown in countries such as India, China and Brazil, where economic and political instabilities often contribute to fluctuations in the supply, price and quality of the oil (Atsmon, 1989). Furthermore, castor bean seeds contain ricin, a highly poisonous protein, and allergens that cause health problems for workers involved in the cultivation and harvesting of the plants. Because of these barriers to increased production, the use of ricinoleic acid is supply limited and the price of castor oil on world markets is twice as high as the price for commodity oils.

These considerations have encouraged chemical and biotechnology companies to investigate the engineering of
ricinoleic acid production in crop plants that are amenable to large-scale agriculture. Ricinoleic acid is formed by a hydroxylase enzyme that adds a hydroxyl group to the 12th carbon atom of oleic acid moieties (Galliard and Stumpf, 1966) esterified to the sn-2 position of phosphatidylcholine (PC) in the endoplasmic reticulum (ER) membrane (Bafor et al., 1991). On the basis of the biochemical characteristics shared between the castor bean hydroxylase and the broader family of fatty acyl desaturases (Galliard and Stumpf, 1966; Moreau and Stumpf, 1981; Bafor et al., 1991; Smith et al., 1992), van de Loo et al. (1995) screened a castor bean developing endosperm cDNA library for sequences homologous to the desaturases. They found a cDNA clone (named fatty acid hydroxylase 12, FAH12), whose predicted protein shared approximately 67% amino acid identity with the Arabidopsis oleate desaturase (FAD2), the enzyme that catalyses the desaturation of oleate (18:1) to linoleate (18:2). Expression of FAH12 in transgenic tobacco caused the accumulation of ricinoleic acid, but only to very low levels (van de Loo et al., 1995). Since that time, several laboratories have made persistent but unsuccessful attempts to identify and overcome the limitations to the high-level production and accumulation of ricinoleic acid in plants. Seed-specific over-expression of FAH12 in Arabidopsis led to higher ricinoleate levels than seen in tobacco, but the largest amount of hydroxy fatty acid (HFA) accumulated in these lines represented about 17% of total seed lipid (Broun and Somerville, 1997; Smith et al., 2003; Kumar et al., 2006; Lu et al., 2006), far less than would be necessary for practical use as a castor oil replacement.

The limited success reported for ricinoleic acid accumulation in these studies has been paralleled in projects to produce vernolic, crepenynic (Lee et al., 1998), α-eleostearic, calendic (Cahoon et al., 1999) and other industrially useful fatty acids in transgenic plants (Jaworski and Cahoon, 2003). These suboptimal results led us to consider the possibility that the accumulation of ricinoleic acid in castor bean oil required not only the evolution of a hydroxylase enzyme from an ancestral FAD2 (Broun et al., 1998), but also the co-evolution of one or more additional enzymes of the lipid biosynthetic pathway to efficiently incorporate ricinoleic acid into triacylglycerol (TAG), the principal component of seed oil. The use of a short-chain acyl-acyl carrier protein thioesterase to produce up to 60% lauric acid in canola (Voelker et al., 1996) provides one example of successful oilseed engineering. Subsequent addition of a second transgene, encoding a short-chain lysophosphatidic acid acyltransferase (LPAT), resulted in a further increase of 5% in the laurate content of the oil (Knutzon et al., 1999).

The pathways of TAG synthesis outlined in Figure 1 indicate some of the enzymes that may be targeted to test this hypothesis. On formation on PC [reaction (1) in Figure 1] (Galliard and Stumpf, 1966; Bafor et al., 1991), ricinoleate can be incorporated into TAG in several ways. Ricinoleic acid may be cleaved from the sn-2 position of PC by phospholipase A2 [reaction (2)], and then activated to a coenzyme A (CoA) thioester by a long-chain acyl-CoA synthetase (LACS) [reaction (3)] (Galliard and Stumpf, 1966), for esterification to the glycerol backbone of TAG by any of the three acyltransferases of the Kennedy pathway (Kennedy, 1961): glycerol-3-phosphate acyltransferase (GPAT) [reaction (5)] (Zheng et al., 2003), LPAT [reaction (6)] (Brown et al., 2002; Kim et al., 2005) and diacylglycerol acyltransferase (DGAT) [reaction (8)] (Hobbs et al., 1999; Zou et al., 1999). Notably, membrane-bound DGAT enzyme activity is encoded by two unrelated gene families, referred to as DGAT1 (Hobbs et al., 1999; Zou et al., 1999) and DGAT2 (Lardizabal et al., 2001; Shockey et al., 2006). Ricinoleoyl-CoA may also be produced by the reverse reaction of acyl-CoA:lysophosphatidylcholine acyltransferase (LPCAT) [reaction (4)] (Tumaney and Rajasekharan, 1999). In addition, ricinoleic acid may be transferred directly from the sn-2 position of PC to the sn-3 position of diacylglycerol (DG) via phospholipid:diacylglycerol acyltransferase (PDAT) [reaction (9)] (Dahlqvist et al., 2000), or may enter the DAG pool directly via phospholipase C [reaction (10)] or the reverse reaction of choline phosphotransferase (CPT) [reaction (11)] (Dewey et al., 1994).

Figure 1 Schematic diagram of ricinoleic acid and triacylglycerol biosynthesis in developing Ricinus communis seeds. Ricinoleic acid is synthesized from oleoyl moieties attached to endoplasmic reticulum (ER) phospholipids, primarily phosphatidylcholine (PC). Ricinoleate is removed from the phospholipid pool and channelled into triacylglycerols through any of several coenzyme A (CoA)-dependent and/or CoA-independent acyltransferase reactions (see text for details).
Full-length cDNA clones encoding several of these enzymes were identified and isolated from a castor bean seed cDNA library, and their ability to increase HFA content when co-expressed in FAH12 transgenic Arabidopsis lines was tested. In this article, it is shown that the expression of an R. communis type-2 DGAT, RcDGAT2, selectively increases the HFA content of the seed oil from 17% to nearly 30% of total fatty acids. Analysis of individual TAG species demonstrates that RcDGAT2 increases the seed oil content of individual TAG species that are also major components of native castor oil. Our results demonstrate that the previous limitations to ricinoleate accumulation in developing oilseeds were, indeed, caused partly by the lack of compatible TAG biosynthetic enzymatic machinery, and that RcDGAT2 overcomes a major bottleneck in the process of HFA accumulation in plants. The results also suggest that additional increases in HFA content will require the identification of enzyme activities that can incorporate HFA into metabolites that are present upstream of DGAT2 enzyme activity (such as diricinolein).

Results

RcDGAT2 enhances HFA accumulation in FAH12 plants

A cDNA library prepared using mRNA from developing castor bean seeds was screened for clones encoding enzymes of lipid metabolism. Degenerate primers were designed on the basis of regions of amino acid conservation amongst known and predicted DGAT, PDAT, LACS and LPAT proteins from various plant species, and were used to amplify partial cDNA sequences employing low-stringency polymerase chain reaction (PCR) protocols. Full-length cDNA sequences were obtained by 5′ and 3′ rapid amplification of cDNA ends (RACE), and four unique cDNAs were identified. These two represented classes of DGAT genes, namely RcDGAT1 (~74% amino acid identity to AtDGAT1) (He et al., 2004) and RcDGAT2 (~71% amino acid identity to AtDGAT2 and VfDGAT2) (Kroon et al., 2006; Shockey et al., 2006), one putative LPAT (RcLPAT, ~92% identity to putative Arabidopsis LPAT At5g60620) and one LACS (RcLACS4, ~85% identity to AtLACS4) (Shockey et al., 2002).

To measure the potential of these enzymes to improve HFA accumulation in Arabidopsis plants expressing FAH12, they were transformed into previously characterized FAH12 lines: CL7 and CL37 (Lu et al., 2006). The parent for these two lines was the Arabidopsis fae1 mutant, which is defective in 18-carbon fatty acid elongation (Kunst et al., 1992). The decrease in long-chain fatty acids, including 20-carbon HFAs, results in a simplified seed fatty acid profile. The CL7 and CL37 lines were selected from over 50 lines expressing the FAH12 transgene because they both consistently produced 17% ± 1% HFAs in the seeds. This level of HFA accumulation is close to the upper limit achieved by us and other research groups (Broun and Somerville, 1997; Smith et al., 2003; Kumar et al., 2006; Lu et al., 2006).

The various castor bean cDNAs were each expressed in CL7 under the control of the seed-specific phaseolin promoter (Slightom et al., 1983). Primary transgenic plants were identified by herbicide selection, and the chromosomal integration of the respective transgenes was confirmed by gene-specific PCR using small-scale genomic DNA preparations (Lukowitz et al., 2000) (see ‘Experimental procedures’). T1 plants were grown to maturity, together with CL7 controls, and samples of T1 seed were analysed for fatty acid composition by gas chromatography (GC). Four lines carrying the RcDGAT2 transgene showed substantial increases in HFA content compared with the CL7 parent (Table 1), whereas plants expressing RcDGAT1, RcLPAT1 or RcLACS4 cDNAs all had HFA levels much closer to that of parental CL7. Thus, only the RcDGAT2 clone resulted in a substantial increase in HFAs, and this cDNA was studied in detail.

To simplify subsequent genetic and biochemical analysis, homozygous lines carrying the RcDGAT2 transgene at a single insertion site were identified. Analyses of single T2 seeds indicated that lines 5-2 and 5-4 segregated approximately 3 : 1 for seeds with > 20% HFA to seeds with < 18% HFA, consistent with the segregation of a single RcDGAT2 transgene responsible for increased HFA. To confirm this, 10 T2 plants from each line were grown, and T3 seeds were analysed to identify homozygous lines 521, 522, 544 and 545. In addition, a segregant from line 5-4, which lacked the RcDGAT2 transgene, was selected and designated line 540. Six T2 plants from each of these five lines were grown together with CL7 plants, and T3 seed samples from each plant were analysed by GC. The HFA content of samples from CL7 and 540 (regenerants lacking RcDGAT2) ranged from 15.1% to 18.7%. By contrast, the 24 plants homozygous for the RcDGAT2 transgene showed significantly higher HFA levels, with a maximum of 30.8% for one of the plants in subline 544 (Figure 2). The complete fatty acid analyses for these samples are included in Table S1 (see ‘Supplementary material’).

To verify the results obtained from the transformation of RcDGAT2 into the CL7 line, the same phaseolin promoter was used to drive the expression of an RcDGAT2 cDNA in FAH12 plants of the CL37 line. In this experiment, primary transformants provided bulk T1 seed samples containing 21%–23% HFAs. The identification of homozygous T2 plants...
provided sublines containing up to 29% HFAs (data not shown). Thus, the increases in HFAs achieved in this second FAH12 transgenic line are comparable with those found in the CL7 line, and verify the link between RcDGAT2 co-expression and increased seed HFA content.

Arabidopsis also contains a gene encoding DGAT2 (locus At3g51520) (Lardizabal et al., 2001). Because RcDGAT2 was over-expressed behind the strong, seed-specific phaseolin promoter, it is possible that increased AtDGAT2 enzyme activity might also drive increased HFA content. A cDNA encoding AtDGAT2 was cloned into the phaseolin vector and used to transform CL37 plants. Analysis of the HFA content of T2 seed samples from six primary transgenics ranged from 15.3% to 18.7%, compared with an average of 17.6% for CL37 control plants (Table 2). Thus, the over-expression of AtDGAT2 does not cause a significant increase in HFA accumulation, suggesting that the increase in HFA content observed in the FAH12/RcDGAT2 double-transgenic lines is the result of the selective incorporation of HFA into TAG by the activity of the RcDGAT2 isozyme.

Increased HFA content does not compromise seed physiology

The transformation of Arabidopsis with FAH12 transgenes has occasionally provided lines with 20%–25% HFA (C. Lu and J. Browse, unpubl. data; L. Kunst, University of British Columbia, Vancouver, pers. commun.). However, these lines were invariably

### Table 1 Fatty acid composition of Arabidopsis line CL7 expressing the castor bean fatty acid hydroxylase 12 (FAH12) together with various castor bean triacylglycerol (TAG) metabolic enzymes. Seed samples from individual T1 plants were analysed by gas chromatography as described in ‘Experimental procedures’

<table>
<thead>
<tr>
<th>Transgenic line</th>
<th>16:0</th>
<th>18.0 + 18:1</th>
<th>18:2</th>
<th>18:3</th>
<th>20:1</th>
<th>18:1-OH</th>
<th>18:2-OH</th>
<th>Sum of HFAs</th>
</tr>
</thead>
<tbody>
<tr>
<td>RcDGAT2 #5-1</td>
<td>17.7</td>
<td>26.6</td>
<td>25.7</td>
<td>7.4</td>
<td>2.0</td>
<td>19.1</td>
<td>1.5</td>
<td>20.6</td>
</tr>
<tr>
<td>RcDGAT2 #5-2</td>
<td>15.5</td>
<td>29.7</td>
<td>23.0</td>
<td>6.7</td>
<td>1.6</td>
<td>21.0</td>
<td>2.5</td>
<td>23.5</td>
</tr>
<tr>
<td>RcDGAT2 #5-3</td>
<td>14.1</td>
<td>34.1</td>
<td>20.2</td>
<td>5.8</td>
<td>1.6</td>
<td>21.1</td>
<td>3.1</td>
<td>24.2</td>
</tr>
<tr>
<td>RcDGAT2 #5-4</td>
<td>14.2</td>
<td>34.8</td>
<td>20.1</td>
<td>5.8</td>
<td>1.7</td>
<td>20.6</td>
<td>2.8</td>
<td>23.4</td>
</tr>
<tr>
<td>RcDGAT1 #1</td>
<td>16.0</td>
<td>32.9</td>
<td>23.1</td>
<td>6.8</td>
<td>1.5</td>
<td>17.5</td>
<td>2.1</td>
<td>19.6</td>
</tr>
<tr>
<td>RcDGAT1 #2</td>
<td>14.3</td>
<td>38.2</td>
<td>21.9</td>
<td>5.7</td>
<td>1.4</td>
<td>15.9</td>
<td>2.6</td>
<td>18.5</td>
</tr>
<tr>
<td>RcDGAT1 #3</td>
<td>16.8</td>
<td>34.6</td>
<td>23.3</td>
<td>6.4</td>
<td>1.5</td>
<td>15.4</td>
<td>2.0</td>
<td>17.4</td>
</tr>
<tr>
<td>RcDGAT1 #4</td>
<td>16.1</td>
<td>35.9</td>
<td>23.3</td>
<td>5.9</td>
<td>1.4</td>
<td>15.3</td>
<td>2.0</td>
<td>17.3</td>
</tr>
<tr>
<td>RcDGAT1 #5</td>
<td>14.6</td>
<td>37.0</td>
<td>21.0</td>
<td>6.4</td>
<td>1.5</td>
<td>16.9</td>
<td>2.5</td>
<td>19.5</td>
</tr>
<tr>
<td>RcDGAT1 #6</td>
<td>15.6</td>
<td>37.2</td>
<td>20.6</td>
<td>6.3</td>
<td>1.8</td>
<td>16.4</td>
<td>2.1</td>
<td>18.5</td>
</tr>
<tr>
<td>RcLPAT #1</td>
<td>14.3</td>
<td>40.7</td>
<td>17.8</td>
<td>6.4</td>
<td>1.5</td>
<td>16.6</td>
<td>2.6</td>
<td>18.4</td>
</tr>
<tr>
<td>RcLPAT #3</td>
<td>13.5</td>
<td>40.3</td>
<td>18.3</td>
<td>6.9</td>
<td>1.5</td>
<td>16.4</td>
<td>2.9</td>
<td>18.1</td>
</tr>
<tr>
<td>RcLACS4 #1</td>
<td>17.8</td>
<td>36.5</td>
<td>23.9</td>
<td>6.1</td>
<td>0.5</td>
<td>15.0</td>
<td>0.2</td>
<td>15.2</td>
</tr>
<tr>
<td>CL7</td>
<td>14.3</td>
<td>40.7</td>
<td>21.2</td>
<td>5.9</td>
<td>0.9</td>
<td>19.0</td>
<td>3.5</td>
<td>17.4</td>
</tr>
</tbody>
</table>

DGAT, diacylglycerol acyltransferase; HFA, hydroxy fatty acid; LACS, long-chain acyl-CoA synthetase; LPAT, lysophosphatidic acid acyltransferase; Rc, Ricinus communis.

Figure 2 Hydroxy fatty acid content of seeds from Arabidopsis lines co-expressing castor bean fatty acid hydroxylase 12 (FAH12) and type-2 diacylglycerol acyltransferase (DGAT2). Each data point represents the hydroxy fatty acid content in a seed sample from an individual homozygous T1 plant. Horizontal bars indicate the mean for each dataset.

© 2008 The Authors
Journal compilation © 2008 Blackwell Publishing Ltd, Plant Biotechnology Journal, 6, 819–831
unstable, either producing seed that failed to germinate or producing progeny that contained < 20% HFA in the seed oil. These observations suggest the possibility that high levels of HFAs damage the metabolism and physiology of seeds. The high levels of HFA (25%–30%) in our FAH12/RcDGAT2 transgenics are clearly heritable. Furthermore, seedlings of FAH12/RcDGAT2 double-transgenic lines germinated and grew as well as parental FAH12 controls (lines CL7 or CL37), and subsequent plant growth, development and seed production were normal. The seed weight and total seed lipid content of transgenic and non-transgenic segregants from CL37 plants transformed with the phaseolin-RcDGAT2 were also compared. The 100-seed weight averaged across six double-transgenic lines was 2.19 ± 0.27 mg (average ± standard error), compared with 2.03 ± 0.01 mg for plants of the CL37 parental line and 2.00 ± 0.06 mg for segregants lacking the RcDGAT2 gene. The total fatty acid content of three double-transgenic lines was 6.70 ± 0.72 μg per seed, compared with 6.30 ± 0.30 and 5.69 ± 0.32 μg for CL37 and three non-RcDGAT2 segregant lines, respectively. These results indicate that the increased HFA accumulation provided by RcDGAT2 has no detrimental effect on the size or total TAG accumulation in comparison with parental controls.

RcDGAT2 expression in seeds enhances the synthesis of castor oil-like TAG species

To further understand the changes in TAG fatty acid composition that contribute to the increased HFA content of the seed oil, quantitative analysis was carried out of the TAG and DAG molecular species in oil from the parental CL7 line and the 544, which expressed the RcDGAT2 transgene. For this purpose, high-pressure liquid chromatography (HPLC) coupled with tandem mass spectroscopy (LC/MS/MS; see ‘Experimental procedures’) was used. This protocol identifies all the major molecular species of TAG, with a detection limit for individual species of approximately 0.2% of the total TAGs. Forty-seven individual TAG species were resolved, identified and quantified by this technique; an annotated example of the analysis protocol is included in Figure S1 (see ‘Supplementary material’).

A comparison of the TAG molecular species in lines CL7 and 544 is shown in Figure 3a. The data are plotted with triricinolein (RRR) on the left, followed by TAG species that contain two HFAs, TAG species containing one HFA and, finally, TAGs containing no HFA. Calculating the fold change between line 544 and the CL7 parent for each TAG species (Figure 3b) makes the differences clear. TAG species with bars above the broken line are more abundant in line 544 than CL7, and those below the broken line are less abundant in line 544. As a point of reference, Figure 3c shows the TAG species profile (mol.%) of castor oil. The most obvious relative change between the single- and double-transgenic lines is a nearly 17-fold increase in triricinolein (RRR) content from < 0.1% in CL7 to 1.5% in line 544. Other ricinoleate-enriched TAG species, such as RRLn, RRL, RRP and RRS (see Figure 3 for abbreviations), are also highly enriched in line 544 compared with line CL7. Indeed, nearly every major TAG component in castor oil is enriched in line 544 relative to CL7, with the only exception being ROO, which is present in line 544 at about 80% of CL7 levels. Taken together, these data strongly suggest that RcDGAT2 acts as a major determinant of TAG fatty acid composition in castor oil, and that the encoded enzyme retains the ability to selectively enrich HFAs in seed oil when transgenically expressed in Arabidopsis.

Table 2 Fatty acid composition of Arabidopsis thaliana lines co-expressing castor bean hydroxylase and Arabidopsis type-2 diacylglycerol acyltransferase (DGAT2)

<table>
<thead>
<tr>
<th>Transgenic line</th>
<th>16:0</th>
<th>18:0</th>
<th>18:1</th>
<th>18:2</th>
<th>18:3</th>
<th>20:1</th>
<th>18:1-OH</th>
<th>18:2-OH</th>
<th>Sum of HFAs</th>
</tr>
</thead>
<tbody>
<tr>
<td>AtDGAT2 #1</td>
<td>13.9</td>
<td>4.3</td>
<td>37.5</td>
<td>19.2</td>
<td>6.1</td>
<td>0.4</td>
<td>16.0</td>
<td>2.7</td>
<td>18.6</td>
</tr>
<tr>
<td>AtDGAT2 #2</td>
<td>13.5</td>
<td>4.7</td>
<td>34.6</td>
<td>22.4</td>
<td>7.4</td>
<td>0.3</td>
<td>13.3</td>
<td>3.8</td>
<td>17.1</td>
</tr>
<tr>
<td>AtDGAT2 #3</td>
<td>12.3</td>
<td>4.8</td>
<td>39.7</td>
<td>18.0</td>
<td>6.4</td>
<td>0.3</td>
<td>14.0</td>
<td>4.4</td>
<td>18.5</td>
</tr>
<tr>
<td>AtDGAT2 #4</td>
<td>12.7</td>
<td>4.7</td>
<td>35.1</td>
<td>21.6</td>
<td>7.2</td>
<td>0.3</td>
<td>13.1</td>
<td>5.3</td>
<td>18.4</td>
</tr>
<tr>
<td>AtDGAT2 #5</td>
<td>12.9</td>
<td>4.6</td>
<td>34.8</td>
<td>21.5</td>
<td>7.1</td>
<td>0.3</td>
<td>14.5</td>
<td>4.3</td>
<td>18.7</td>
</tr>
<tr>
<td>AtDGAT2 #6</td>
<td>13.7</td>
<td>4.0</td>
<td>36.0</td>
<td>21.3</td>
<td>9.4</td>
<td>0.4</td>
<td>11.9</td>
<td>3.4</td>
<td>15.3</td>
</tr>
<tr>
<td>CL37</td>
<td>12.0</td>
<td>4.7</td>
<td>38.7</td>
<td>20.2</td>
<td>6.5</td>
<td>0.4</td>
<td>12.1</td>
<td>5.4</td>
<td>17.6</td>
</tr>
</tbody>
</table>

HFA, hydroxy fatty acid.

© 2008 The Authors
Journal compilation © 2008 Blackwell Publishing Ltd, Plant Biotechnology Journal, 6, 819–831
A summary of the quantitative analysis of TAGs is shown in Figure 4. Here, the HFAs (18:1-OH and 18:2-OH) are represented by ‘H’ and all other (non-hydroxy) fatty acids are represented by ‘x’. As noted above, tri-HFA molecular species are only a small percentage of the total. Most of the increase in HFAs is accounted for by a nearly three-fold increase in di-HFA molecular species, and there is also an increase in mono-HFA species. By contrast, molecular species containing no HFA are reduced from 48% of the total in CL7 TAGs to only 23% in line S44.

Additional insight into the substrate selectivity of RcDGAT2 and the positional distribution of HFA in the TAG pool was gained through LC/MS/MS analysis of the DAG pools present in the different transgenic seeds. These analyses identified 12 molecular species of DAG that were detectable in seeds of CL7 and line S44 (Figure 5). Because of the small size of the DAG pool (relative to TAGs), the detection limit for individual DAG species was approximately 1% of total DAGs. The DAG species included di-racinooleoyl DAG and five that contained one HFA, but other molecular species that were predicted...
precursors of the TAGs shown in Figure 3 could not be detected. These may be present at levels below the detection limit.

RcDGAT2 has high activity on diricinolein

The ability of the castor bean DGAT2 enzyme to use diricinolein as a DAG substrate was tested directly by expression of the cDNA in Saccharomyces cerevisiae strain H1228, a knockout strain lacking the sole yeast DGAT gene, DGAT1, which is responsible for approximately 70% of total TAG synthesis in stationary phase yeast cells (Oelkers et al., 2002; Sandager et al., 2002). Microsomes from yeast strains expressing either RcDGAT2 or AtDGAT2 were isolated and used in assays containing radioactive oleoyl-CoA as the acyl donor and different molecular species of DAG as acyl acceptors. Relative incorporation of radioactivity into different lipid products was analysed by thin layer chromatography (TLC) separation and subsequent exposure to autoradiography film. For quantification, the appropriate TAG species (dihydroxy-TAG (retention factor (Rf) = 0.48) for assays containing diricinolein and total neutral TAG (Rf = 0.7) for assays containing either diolein or dilinolein) were scraped from the TLC plate and analysed by liquid scintillation counting. AtDGAT2 was essentially inactive in these preparations with these particular substrates. By contrast, RcDGAT2 was active towards all three DAG substrates tested, but produced nearly 10-fold more TAG when diricinolein was used as the acyl acceptor relative to diolein or dilinolein (Figure 6). Notably, RcDGAT2 has also been shown to be active with ricinoleoyl-CoA as an acyl donor (Kroon et al., 2006).
Discussion

Engineering oilseed crops to produce large amounts of industrially useful fatty acids has been proven to be very difficult. Biotechnology researchers have identified the enzymes responsible for the synthesis of many derivatized fatty acids and cloned the corresponding genes. However, the expression of these genes in crop plants, or in Arabidopsis as a model system, has provided lines with only low yields of the target fatty acids (van de Loo et al., 1995; Broun and Somerville, 1997; Lee et al., 1998; Cahoon et al., 1999; Jaworski and Cahoon, 2003; Smith et al., 2003). When the oleate hydroxylation from castor bean, FAH12, was expressed in Arabidopsis, the stable transgenic lines contained no more than 17%–18% HFAs. In this study, however, the castor bean DGAT2 has been identified as a gene that is able to increase the HFA content from 17% to nearly 30% in transgenic Arabidopsis.

The biochemistry of TAG synthesis in seeds is complex (Figure 1) and the incorporation of unusual fatty acids, such as ricinoleate, may be limited by the activities or substrate specificities of one or more of the many enzymes involved. Previous studies have suggested that DGAT is rate limiting for TAG synthesis in oilseeds (Ichihara et al., 1998), and the DGAT activity in microsomal preparations from developing castor bean seeds has been shown to favour 18:1-OH-containing substrates (Wiberg et al., 1994; Vogel and Browse, 1996). However, the interpretation of these data is complicated by the fact that plants contain at least two types of membrane-bound enzyme that catalyse DGAT activity, and the biochemical studies cited above were carried out prior to the cloning of either DGAT1 (He et al., 2004) or DGAT2 (Kroon et al., 2006), so that the relative contribution of each type of enzyme was not determined in the early studies.

The completion of the genome sequences for Arabidopsis, rice, yeast (S. cerevisiae) and many other eukaryotic organisms has allowed for the tentative identification of candidate sets of orthologous genes for many of the enzymatic steps involved in TAG biosynthesis. This expanding database of genomic information was used in this study to identify cDNAs corresponding to four different classes of lipid biosynthetic enzyme from castor bean, including RcDGAT1, RcDGAT2, RclPAT1 and RclLACS4. Since the time of initiation of our studies, projects aimed at the sequencing of the castor bean genome have been completed, and a 4× coverage of the genome is currently available (http://castorbean.tigr.org/); complete sequencing of the castor bean genome will greatly accelerate the determination of the size of relevant gene families and the analysis of the biochemical roles of each of the candidate enzymes listed in Figure 1.

Although there is some evidence that castor bean DGAT1 and LPAT may also be important for the accumulation of unusual fatty acids in plants (Lassner et al., 1995; Knutzon et al., 1999; He et al., 2004; Milcamps et al., 2005), our transformation of RcDGAT1 and RclPAT clones into CL37 plants did not lead to any substantial increase in HFAs above the parental level of 17%. Because RcDGAT2 clearly increased HFA levels, our efforts were focused on this gene, and RcDGAT1, RclPAT1 and RclLACS were not investigated any further. In enzyme assays, recombinant RcDGAT2 exhibited an approximately 10-fold higher preference for 18:1-OH-DAG over 18:1- or 18:2-DAG, suggesting that this protein may indeed have co-evolved in castor bean for the efficient transfer of HFA into oil. Other recent analyses also suggest that RcDGAT2 is important for the accumulation of unusual fatty acids. For example, Kroon et al. (2006) recently characterized RcDGAT2 gene expression in castor bean by real-time reverse transcriptase-polymerase chain reaction (RT-PCR), which demonstrated that RcDGAT2 transcripts were enriched 18-fold in castor bean seeds relative to leaves, whereas RcDGAT1 transcripts showed no such up-regulation. Furthermore, RcDGAT2 could utilize ricinoleate in acyl donor and acyl acceptor substrates (Kroon et al., 2006), a requirement for the synthesis of triricinolein, the major TAG species in castor oil (Figure 3). In this study, it has been shown that RcDGAT2 facilitates the synthesis of products containing at least two HFAs (Figures 3 and 4). In addition to studies with castor bean DGAT2, the DGAT2 enzyme from the tung tree (Vernicia fordii), whose seed oil contains approximately 80% of an unusual fatty acid called α-eleostearic acid, incorporates eleostearic acid into TAG more efficiently than does tung tree DGAT1 (Shockey et al., 2006).

By contrast, our data have shown that Arabidopsis DGAT2 does not utilize 18:1-OH-DAG effectively, in either purified biochemical assays or in planta, in comparison with lines that express RcDGAT2. The low overall activity of AtDGAT2 in the DGAT enzyme assays described here may be a result of the suboptimal experimental conditions, enzyme instability or the enzyme may be specialized for a subset of the fatty acids that were not utilized in the assay. For example, wild-type Arabidopsis seeds contain approximately 25% of 20- and 22-carbon fatty acids, but AtDGAT2 has not yet been tested with substrates containing these very-long-chain fatty acids. Whether DGAT2 enzymes in plants are preferentially utilized for the incorporation of unusual fatty acids into storage oil, or whether they may also be involved in the processing of ‘normal’ fatty acids, in conjunction with DGAT1 enzyme activity, remains to be determined.
The amino acid sequence of RcDGAT2 is approximately 70% identical to that of other plant DGAT2 enzymes (Figure 7). Like many other DGAT2 isoforms, RcDGAT2 contains two closely spaced membrane-spanning domains (Shockey et al., 2006) and a C-terminal aromatic ER retrieval motif (McCartney et al., 2004). The only obvious unique feature of RcDGAT2 is a stretch of 12 residues in the N-terminal region that contains a total of nine asparagine residues, including six in a row. The functional significance of this region, as well as the identification of the determinants that impart high affinity for HFA substrates, awaits future experimentation.

Our results indicate that HFA accumulation in the seed oil of FAH12 transgenic plants is not limited by the activity of the hydroxylase, but rather by the inability of TAG synthesis enzymes, in this case the endogenous DGAT activity, to efficiently incorporate HFAs into TAG. In the absence of RcDGAT2, HFAs may feedback inhibit the hydroxylase, or be subject to β-oxidation to set up a futile cycle of HFA synthesis and degradation (Eccleston and Ohlrogge, 1998). Irrespective of the mechanism that operates to limit HFA accumulation in FAH12 plants, the expression of RcDGAT2 partially overcomes this limitation and increases HFAs from 17% to nearly 30% of the fatty acid content of the seed oil. This is significant because oils containing 25%–40% HFA can be employed for several industrial applications, including use as a lubricant. More importantly, the results presented here support the proposal that additional enzymes have co-evolved with the castor bean hydroxylase for the selective incorporation of HFAs into storage oil. That engineering and expression of more transgenes from castor bean might produce additional improvements in Arabidopsis HFA content is supported by the work of Abbadi et al. (2004) and Qi et al. (2004), who engineered oilseeds to produce significant quantities of very-long-chain polyunsaturated fatty acids through the use of multiple elongases and desaturases.

The removal of unusual fatty acids from their site of synthesis in the phospholipid pool of the ER membrane also limits the accumulation of novel fatty acids in transgenically produced oils (Cahoon et al., 2006). Therefore, enzymes that promote the removal of ricinoleic acid from phospholipids, whilst favouring HFA accumulation in the acyl-CoA and DAG pools (Tumaney and Rajasekharan, 1999; Dahlqvist et al., 2000; Shockey et al., 2002), will also be attractive targets for future metabolic engineering programmes. Other important genes of interest are isoforms of Kennedy pathway acyltransferases, such as GPATs and LPATs (Knutzon et al., 1999), which have been used successfully to produce > 60% lauric acid when expressed with a short-chain acyl-acyl carrier protein thioesterase in transgenic canola. The prospects for future gene-stacking experiments are certainly bolstered by our data, which demonstrate clearly that the expression of just one additional enzyme from castor bean, together with FAH12, can successfully boost the accumulation of HFAs in transgenic plants.

**Experimental procedures**

**Identification of castor bean genes and vector construction**

Degenerate primers were designed according to regions of amino acid conservation between related enzymes from various plant species, based on either published sequences (Brown et al., 1994, 2002; Hobbs et al., 1999; Lardizabal et al., 2001; Shockey et al., 2002) or predicted proteins found by BLAST analysis of the Plant Gene Indices at the Institute for Genomic Research (TIGR) (http://www.tigr.org/tdb/tgi/plant.shtml). Sequences were aligned using the CLUSTALX algorithm (Thompson et al., 1997). Initial castor bean amplicons were isolated from an R. communis seed cDNA library. Gene-specific primers were used to perform 5’ and 3’ RACE reactions. Sequence information from the RACE products was used to design full-length cDNA primers. After obtaining the full-length cDNA sequences, the open reading frames (ORFs) were amplified from the library with...
PfuUltra polymerase (Stratagene, Cedar Creek, TX, USA) using end-specific primers containing restriction sites appropriate for cloning into pBlUESCRIPT SK’ (Stratagene) or the seed-specific plant expression vector, pOEA2, which contains a multiple cloning site immediately 3’ to the seed-specific promoter phaseolin (Slightom et al., 1983). Complete details concerning the oligonucleotide primers used in gene cloning and plasmid constructions are available on request. The sequences for \textit{RcDGAT1}, \textit{RcDGAT2}, \textit{RcLACS4} and \textit{RcLPAT} have been deposited in the GenBank database under accession numbers EU391591, EU391592, EU391593 and EU391594, respectively.

**Plant material, growth conditions and transformation**

\textit{Arabidopsis thaliana} plants were grown at 22 °C under a 16-h day/8-h night photoperiod. A marker-less castor bean hydroxylase construct was developed and used to transform the \textit{Arabidopsis fae1} mutant (line AC56) (Kunst et al., 1992) employing the floral dip method of Clough and Bent (1998). The resulting pool of \textit{T1} plants was screened for the presence of HFAs by GC analysis of bulk seed fatty acids. Lines CL7 and CL37 were identified in this manner; subsequent generations of each plant line were analysed until homozygous individuals were identified (Lu et al., 2006). The lines were transformed with the castor bean constructs described above, and \textit{T1} seeds were harvested and subsequently screened on soil watered with Finale herbicide (Farnam Companies, Phoenix, AZ, USA; 430 μLα). \textit{T2} seed was screened by bulk seed fatty acid analysis and compared with untransformed CL7 and CL37 seed. The presence of the transgenes was verified by single-leaf genomic DNA extraction (Lukowitz et al., 2000) and gene-specific PCR.

**Average seed weight**

Five hundred seeds from each line were counted with the aid of the Syngene Bio Imaging System and the software packages Gene Snap and Gene Tools (Synoptics, Cambridge, Cambridgeshire, UK). Each seed sample was then weighed, and the value was divided by five to obtain the 100-seed weight.

**GC analysis of fatty acids**

The fatty acid composition of seeds was determined by GC after derivatization with 2.5% (v/v) H2SO4 in methanol for 1 h (Miquel and Browse, 1992). For bulk seed analysis, 50–100 seeds were used per sample. The total lipid content was determined by performing quantitative GC using exactly 20 seeds and 20 μg of 17:0 fatty acid as an internal standard. GC samples were analysed on a 15 m × 0.25 mm (inside diameter) AT-WAX Alltech column (Alltech, Deerfield, IL, USA), using the following programme: initial temperature, 190 °C for 2 min; ramp increase at 10 °C/min to 230 °C; final temperature hold, 4 min.

**LC/MS/MS of seed neutral lipids**

Lipids were extracted for LC/MS/MS analysis by transferring 50 \textit{Arabidopsis} seeds to a 1.5-ml microfuge tube and adding 10 μL 1,1,1-13C-triolein (5 mg/mL in chloroform) as an internal standard; 400 μL of hexane-isopropanol (3 : 2, v/v) was added to each tube and the seeds were ground using a plastic pestle. Samples were snap frozen in liquid nitrogen and then incubated at 4 °C for 60 min. Tissue debris was removed by centrifugation at 20 000 × g for 5 min in a microfuge, and the supernatant was transferred to a new 1.5-ml microfuge tube. The remaining pellet was washed three times with 100 μL of hexane-isopropanol (3 : 2) and the supernatants were combined. Samples were clarified by the addition of 350 μL of 6.7% sodium sulphate (w/w), vortexing and centrifugation for 30 s at 20 000 × g. The supernatant was transferred to a tapered HPLC vial and the solvent was removed by drying down on a speed-vac. The lipid residue was reconstituted in 100 μL of chloroform; 10 μL was injected on to an LCQ MS (Thermoquest, Hemel Hempstead, Hertfordshire, UK) equipped with a C30 HPLC column (YMC; 250 × 4.6 mm, 5 μm particle size; Shimogyo-ku, Kyoto, Japan) held at 30 °C. A tertiary separation gradient was used at 1 mL/min with the following solvents: A, 20 mM ammonium formate in 80% (v/v) methanol; B, methanol; C, tetrahydrofuran. All solvents additionally contained 0.2% (v/v) formic acid. The run gradient was as follows: 0–5 min isocratic 5% A, 95% B; 5–45 min to 5% A, 35% B, 60% C; then hold isocratic for 45–50 min. A 10-min re-equilibration time was used between injections. The column eluent was introduced without splitting into an atmospheric pressure chemical ionization (APCI) source with the following conditions: vaporizer temperature, 350 °C; sheath gas (N2) flow, 60 units; auxiliary gas flow, 60 units; source current, 5 μA; capillary voltage, 32 V; capillary temperature, 150 °C. Full-scan MS data were collected over the range 450–1500 m/z, and MS2 fragmentation data were collected in data-dependent mode at 60% normalized collision energy and an isolation width of 4 m/z. Individual TAGs were identified by reconciling DAG fragments produced as daughter-ion spectra in data-dependent MS2 mode with \textit{TAG} parent-ion spectra. Similarly, DAGs were identified by reconciling fatty acid fragments from MS2 data with DAG parent-ion spectra. TAGs and DAGs were quantified by comparison of integrated full-scan MS peak areas of the ammoniated molecular ions ([M + NH₄]+) with a known amount of 13C-triolein internal standard [see Figure S1 (‘Supplementary material’ for an example of \textit{TAG} identification and quantification].

**Expression of \textit{RcDGAT2} and \textit{AtDGAT2} in yeast**

\textit{RcDGAT2} and \textit{AtDGAT2} were cloned into pYES2 (Invitrogen, Carlsbad, CA, USA) and transformed into \textit{S. cerevisiae} mutant strain H1228 (kindly provided by Sten Stynne and Ulf Stahl, Swedish University of Agricultural Sciences, Alnarp) (Sandager et al., 2002). The wild-type yeast strains G175 and H1228 bearing empty pYES2 plasmid were used as controls. Transformants were grown in liquid galactose medium lacking uracil at 30 °C, harvested by centrifugation, washed with sterile water and resuspended in lys buffer (5 mL phosphate-buffered saline solution, pH 7.2, containing complete, mini ethylenediaminetetraacetic acid (EDTA)-free protease inhibitor cocktail (Roche, Penzberg, Germany)]. Cells were lysed with 600-μm glass beads using a tabletop bead-beater (4 × 1 min). Unlysed cells and cell debris were removed by centrifugation (700 × g) at 4 °C for 15 min, followed by sedimentation of membranes at 100 000 × g at 4 °C for 30 min. The resulting pellet was resuspended in ice-cold phosphate-buffered saline solution containing 20% glycerol. Samples were stored at −80 °C before use in DGAT enzyme assays.
DGAT enzyme assay

The DGAT enzyme assay mixture consisted of yeast microsomal membranes (250 μg protein), 120 μM DAG, 18 μM [1-14C]-oleoyl-CoA (American Radiolabelled Chemicals, St. Louis, MO, USA; 110 000 dpm/μg), 2.03 × 10^5 Bq/nmol), 5 mM ATP, 5 mM reduced coenzyme A (CoASH) and 1 mM MgSO_4 in a final volume of 500 μL. sn-1,2-Diolein, sn-1,2-diricinolein (Nu-Chek, Elysian, MN, USA) and sn-1,2-dilinolein (Turner et al., 2003) (generously provided by Tom McKeon, WRRC, USDA-ARS, Albany, CA, USA) were used as DAG substrates. Assays were performed at 25 °C for 15 min with shaking at 100 r.p.m.

Lipids were extracted as described previously (Miquel and Browse, 1992), and separated on TLC plates developed in hexane–ethyl ether–acetic acid (35 : 70 : 1.5, v/v/v). The lipid spots were visualized by light staining in iodine vapour. Lipid standard migration measurements were determined by viewing TLC plates treated with a lipid charring solution (3% cupric acetate, 8% phosphoric acid), followed by baking the plates for 10 min at 180 °C. Radioimages were taken by exposure to X-Omat Blue XB-1 autoradiographic film (Kodak, Rochester, New York, USA). TAG fractions were scraped and analysed for radioactivity by liquid scintillation counting. For assays using dioloein and dilinolein as DAG substrates, the neutral TAG fraction was used for DGAT2 activity quantification. For the diricinolein assays, the 2-OH TAG fraction was quantified.

Acknowledgements

We thank Dr Sten Stymne and Dr Ulf Stahl for providing yeast strains, and Dr Tom McKeon (USDA-ARS, Albany) for the generous gift of diricinolein. We thank Dr Jim Wallis and Chris Skidmore (Washington State University, Pullman) for assistance with plant growth, lipid analysis and manuscript preparation. This work was supported by Dow Chemical Co. & Dow AgroSciences, the National Research Initiative of the USDA Cooperative State Research Education & Extension Service (Grant DBI-06-03263), the National Science Foundation (Grant DBI-0701919), the Agricultural Research Center (administered by the Biotechnology and Biological Sciences Research Council of the UK).

References


Atsmon (1989) Castor Oil: its Use in Medicine (administered by the Biotechnology and Biological Sciences Research Council of the UK).

Hydroxy fatty acid production in plants 829


© 2008 The Authors
Supplementary material

The following supplementary material is available for this article:

Figure S1 Liquid chromatography/mass spectroscopy (LC/MS) identification and quantification of triacylglycerols (TAGs).
(a) Total ion chromatogram of TAGs derived from plant line 544. Individual peaks are labelled with the retention time and mass of the corresponding ammoniated molecular ion. The internal standard, 13C-labelled triolein, has a retention time of 32.35 and an ammoniated molecular ion mass of 905.5. (b) Expanded region of the total ion chromatogram showing just 2 min of separation near the 30-min mark. Peaks are labelled as in (a). The peak at 30.20 min, with a prominent molecular ion at 918.3, is used as an example below for TAG identification. (c) Full MS scan at the retention time of 30.20 min, showing the prominent ammoniated molecular ion at 918.3 and a second ammoniated molecular ion at 870.6, representing a minor TAG species co-eluting at the same retention time position. (d) Fragmentation analysis of the molecular ion at 918.3 by collision-induced dissociation and collection of fragments by MS2. Prominent neutral loss ions are observed at 603.4 (loss of ricinoleic acid), 617.3 (loss of stearic acid) and 621.6 (loss of linoleic acid), indicating that the parental molecular ion (mass of 918.3) represents a TAG composed of R, L and S (see Figure 3 legend for abbreviations). Although the MS2 fragment of lowest abundance probably represents the loss of a fatty acid from the sn-2 position (thereby defining the fatty acid at sn-2), no discrimination of sn-2 vs. sn-1 and sn-3 was made in our analysis because of a lack of appropriate molecular standards. (e) Fragmentation analysis of the parental ion at 870.6, showing prominent neutral loss ions at 573.3 (loss of linoleic acid), 575.4 (loss of linolenic acid) and 597.4 (loss of palmitic acid), thereby identifying the TAG as LnLP. Once all molecular ions are identified, the TAGs can be quantified by integrating the respective peak areas in the total ion chromatogram (b), and comparing the peak areas to a known amount of internal standard. The ionization efficiencies of different unsaturated TAG species are essentially the same, allowing direct comparison of TAG species with the internal standard (unpublished data).

Table S1 Fatty acid compositions of seed samples from four Arabidopsis lines. Lines CL7 (parental) and 540 carry only the fatty acid hydroxylase 12 (FAH12) transgene. Lines 521, 522, 544 and 545 are FAH12/Ricinus communis type-2 acyl-coenzyme A:diacylglycerol acyltransferase (RcDGAT2) double transgenics

This material is available as part of the online article from: http://www.blackwell-synergy.com/doi/abs/10.1111/j.1467-7652.2008.00361.x
(This link will take you to the article abstract).

Please note: Blackwell Publishing is not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.