

# A Differentially Expressed Gene from a High Oil Producer Cultivar of Castor Bean (*Ricinus communis*) Is Involved in the Biosynthesis of Ricinoleic Acid

Héctor A. Rodríguez-Cabal, Claudia Y. Jaramillo-Mazo, Nicolás D. Franco-Sierra, Diego F. Villanueva-Mejía, Javier C. Alvarez\*

Department of Biological Sciences and CIBIOP Research Group, School of Sciences, Universidad EAFIT, Medellín, Colombia  
Email: \*jcorre38@eafit.edu.co

**How to cite this paper:** Rodríguez-Cabal, H.A., Jaramillo-Mazo, C.Y., Franco-Sierra, N.D., Villanueva-Mejía, D.F. and Alvarez, J.C. (2020) A Differentially Expressed Gene from a High Oil Producer Cultivar of Castor Bean (*Ricinus communis*) Is Involved in the Biosynthesis of Ricinoleic Acid. *American Journal of Plant Sciences*, 11, 393-412.

<https://doi.org/10.4236/ajps.2020.113029>

**Received:** February 6, 2020

**Accepted:** March 22, 2020

**Published:** March 25, 2020

Copyright © 2020 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## Abstract

*Ricinus communis* or castor bean is a non-edible oilseed plant widely cultivated worldwide for the high content of castor oil in its seeds and the different uses the oil has in the industry. An increase in its oil content and production efficiency is difficult, making understanding the molecular mechanisms underlying the synthesis of oils in the seed necessary. Here, a combined analysis of protein-protein interaction networks was performed using public data on differential gene expression in castor bean seeds at different stages of development. From this analysis, four key enzymes were selected and analyzed in the polyunsaturated fatty acids pathways, whose gene expression was subsequently quantified during the development of the seeds in a Colombian cultivar that produces high amounts of oils and contrasted with a lower producing cultivar. The gene coding for FAH12 was differentially expressed in the early stages of seed development in the high oil-producing cultivar and has differences in amino acids A242V and Q319H. The analysis presents this gene as one of those responsible for early ricinoleic acid synthesis, making it a candidate for use in crop genetic improvement programs to increase the oil content in castor bean.

## Keywords

Tropical Cultivars, Castor Oil, FAD2, FAH12, Oilseed Crop

## 1. Introduction

Castor bean (*Ricinus communis*) is a plant that belongs to the family Euphor-

biaceae cultivated in tropical and subtropical regions of the world to obtain fatty acids [1] [2]. It is one of the world's most important non-edible oilseed crops owing to the ability of its seeds to store lipids, proteins, and carbohydrates [3] [4]. Above 85% of Castor oil is obtained from its seeds and it has been widely used in cosmetics, plastics, and health (drugs) industries [5] and, more recently, in the industrial sector for biodiesel production [6]. However, characteristics such as growth habitat, seed size, oil content, and foliage and stem color, vary significantly among castor bean trees [7], presenting a great genotypic and phenotypic diversity that, similar to the case of the genes involved in the production of fatty acids, requires further study [8].

Germplasm collections are one of the most important sources of genetic resources in the world, using which it is possible to study genetic profiles that explain the phenotypic plasticity of plants [9]. In recent years, the Colombian Agricultural Research Corporation (in Spanish, Corporación Colombiana de Investigación Agropecuaria or AGROSAVIA) has advanced in the evaluation of wild-type castor bean materials with high oil content and adaptation to tropical climates, creating a germplasm bank suitable for studying genes related to the biosynthesis of fatty acids in tropical conditions [10].

The biosynthesis of ricinoleic acid in castor bean starts with the biosynthetic pathway of triacylglycerol (TAG) in the plastids and endoplasmic reticulum (ER), a common pathway in most flowering plants that have been studied [11]. In a posterior pathway, the enzyme oleate D12-hydroxylase (FAH12, *EC* 1.14.13.26), located in the membrane of ER [12], catalyzes the direct hydroxylation of oleic acid (C18:1) to ricinoleic acid (C18:1-OH) [13]. The gene coding this castor bean FAH12 enzyme was already isolated and characterized [14], and *Arabidopsis* plants were subsequently genetically transformed, producing low levels of ricinoleic acid [15]. Thus, several studies conducted on model plants with genes related to fatty acid production pathways in seeds, have been obtained new fatty acids and unusual fatty acids, information considered the base to new studies in non-model plants [16]. This suggests in addition to key genes, other genes are also needed for the large-scale production of ricinoleic acid in seeds, such as transcription factors or accessory enzymes [17]. Evidence of the above was found when researchers studied the biosynthesis of diacylglycerol acyltransferase (DGAT, *EC* 2.3.1.20) isoenzymes in castor bean *in vitro* plants and they identified that diricinolein and C18:1-OH-CoA are needed as substrates [18]. When these enzymes were transformed into *Arabidopsis* expressing the oleate D12-hydroxylase, the content of C18:1-OH in seeds increased by approximately 30% [19].

Therefore, it is necessary to increase knowledge about gene expression associated with the biosynthetic routes of fatty acids and their regulation. It could provide a better understanding of the mechanisms that control the oil synthesis and then move toward increasing their production efficiency [20]. Nowadays, high-throughput sequencing technologies make it possible to study the gene expression of the complete transcriptomes in plants [21]. This strategy was used in

castor bean plants to elucidate its genome [22] and organellar genome [23] and transcriptome in tissues during seed development [24] [25]. All this genetic information helps build new knowledge-based in silico reconstruction of the interaction networks between genes, proteins, and metabolites using organism models as reference [26].

In this research, a protein-protein interaction (PPI) network was reconstructed for castor bean plant based on the orthologous proteins of Arabidopsis. Gene expression values in protein-coding genes during different seed stages were uploaded to the network and biosynthetic key genes involved in the production of fatty acids were identified and verified by RT-qPCR.

## 2. Materials and Methods

### 2.1. Data Source

RNA-Seq raw data were retrieved from NCBI (National Center for Biotechnology Information) SRA database, code ERA047687 [27]. We used the transcriptome data of developing male flowers (accession: ERX021379), leaf (accession: ERX021378), germinating seed (accession: ERX021377), developing endosperm V/VI (accession: ERX021376), and developing endosperm II/III (accession: ERX021375).

For reads mapping to the reference genome (*R. communis*, <http://castorbean.jcvi.org/index.php>) was used HISAT [28]. Annotation of the coding genes was made by comparison into the TIGR database [29]. After mapping of reads in the gene models of *R. communis* genome, were obtained seven files with the read counts per gene using Cufflinks version 0.9.3 [30]. Finally, differential expression ratio for each gene was calculated contrasting different stages (II to V) of endosperm development using Cuffdiff in [30]. Read counts from germinating seed, leaf and developing male flowers were pooled and taken as control.

### 2.2. In silico Construction and Analysis of PPI Network

We conducted a combined analysis of PPI networks (interactome) for *R. communis*, which was inferred by peer-to-peer orthology analysis with the proteins reported from version TAIR10. The reconstruction of the interaction network was completed by mapping orthologous proteins against the network reported for *A. thaliana* (<ftp://ftp.arabidopsis.org/home/tair/Proteins/>) using the software OrtoMCL [31]. The interaction network was enriched with the differential expression ratios obtained in Section 2.1, having fold changes in different developmental stages of the endosperm. Finally, nodes analysis corresponding to the metabolic pathways involved in fatty acid synthesis and accumulation was performed, displaying nodes (enzymes) that exhibited overexpression in the endosperm compared to the control.




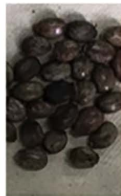
### 2.3. Plant Material

For this research, we used cultivars of castor bean (*R. communis*) classified as

experimental cultivar *Ricinus Corpoica* 03 (VERC03, Colombian collection) and experimental cultivar *Ricinus Corpoica* 12 (VERC12, a result of genetic improvement from Brazil). Considered as promising cultivars after having been studied for more than a decade and that are reported as suitable to be grown in dry agoclimatic areas [10]. These cultivars were obtained from the germplasm collection at the La Selva AGROSAVIA Research Center (Rionegro, Colombia) through a Material Transfer Agreement 20162103285. The two cultivars were cultivated in an experimental field plot in the Bajo Cauca Antioqueño, Colombia (Latitude: 7°56'29" N, Longitude: 75°8'43" W), located at 150 meters above sea level, with 85% of relative humidity and 28° Celsius of average temperature during the day, the experimental area was uniformly fertilized according to [10] and irrigation at 10-day intervals. Plant material was collected for four different stages of seed development (Figure 1 and Figure S1) under the collection framework permit granted by the National Environmental Licensing Authority (in Spanish, Autoridad Nacional de Licencias Ambientales or ANLA) under Resolution 1516 of 2014 (modified by Resolution 1312 of 2015). Fruits with different development stage, coloration and size were collected from VERC03 and VERC12 plants located in the furrows or edges of the growing fields (Figure 1). All samples collected were stored at -80°C in liquid nitrogen.

#### 2.4. Fatty Acid Composition and Oil Content

Total saturated, monounsaturated, and polyunsaturated fats in castor bean seeds during different development stages were determined using gas chromatography (GC).

Development stage	S1	S2	S3	S4
Seed photo				
Days	10	20	40	60
Length (cm)	1.38	1.20	1.43	1.35

**Figure 1.** Development stages of castor bean seeds selected from the VERC03 cultivar. S1 to S4 represent seed stages based on development time of seeds and their average size: image shows that S1 was defined as seeds with 10 days having 1.39 cm in average size; in S2 the seed having a minor size but with changes of testa color; in S3 stage the volume of seed increases dramatically until 1.43 cm and teste is mature with mosaic color brown and silver; S4 have capsules senesce and desiccate, and less seed weight and capsule size.

**Extraction Method:** In total,  $2 \pm 0.0020$  g of each sample was weighed in a glass capsule with an addition of 100 mg of butylated hydroxytoluene (BHT) and 2.0 mL of internal standard (5.00 mg/mL of triundecanoate triglyceride). This was subjected to acid hydrolysis for 1 h with HCl 4 N in SOXCAP 2047. Then, the samples were washed with distilled water to remove acid residue and then dried in a microwave oven for 45 min. The glass capsules with the dry samples were exposed to the SOXTEC 2050 equipment to extract the fat with 50 mL of hexane. The fat extract was incubated at room temperature in a desiccator and then methylated with boron trifluoride in methanol at 14%.

**Analytical Equipment:** The fatty acids were identified using an Agilent 7890B gas chromatograph system with a flame ionization detector (GC-FID), equipped with 7963A autosampler, 100:1 split injection port, and 100% biscianopropil polysiloxane TR-CN100 (60.0 m  $\times$  0.25 mm i.d.  $\times$  0.20  $\mu$ m) capillary column (Teknokroma, Spain). Temperature program conditions were: the initial temperature 100°C, ramp to 14,235°C at 83°C/min, then by a rate of 2°C/min to 220°C, the injector temperature was 250°C. Helium was used as a drag gas at a flow rate of 1.0 mL/min.

**Method of identification and quantification of compounds:** Signals were recorded and manipulated using the software OpenLab CDS ChemStation v A.01.05. Fatty acid identification was performed by comparison with the retention times of the Restek brand Food Industry FAME Mix standard (37 components). The quantification of fatty acids was performed against an internal standard, triundecanoic triglyceride (C11:0), before the calculation of total fat. These were converted to the respective triglycerides and fatty acids; subsequently, the total fats were calculated as the sum of the individual fatty acids expressed as triglyceride equivalents and saturated, monounsaturated, and polyunsaturated fats.

## 2.5. Extraction of RNA from Castor Bean Seeds and cDNA Synthesis

According to Addendum #1 (June 16, 2017) to the Framework Contract for Access to Genetic Resources and Their Derivative Products (in Spanish, *Contrato Marco de Acceso a Recursos Genéticos y Sus Productos Derivados*) #127 of May 13, 2016, between the Ministry of the Environment and Sustainable Development and EAFIT University, the total RNA was extracted from different tissues of castor bean cultivars. Equal amounts of embryo and endosperm tissue were removed from seeds collected from fruits. RNA extraction was performed using the reagent RiboZol<sup>TM</sup> (Amresco, USA), then purified with the commercial kit GeneJET Plant RNA Purification (Thermo Fisher Scientific, USA) according to the manufacturer's instructions and quantified by spectrophotometry (260/280 nm) in Nanodrop ND-2000 UV-V is Spectrophotometer (Nanodrop Technologies, USA). RNA Integrity was analyzed by electrophoresis in a 1% agarose gel. cDNA synthesis was performed using the Maxima First Strand cDNA

Synthesis Kit for RT-qPCR with dsDNAsa (Thermo Fisher Scientific) and cDNA products obtained were stored at  $-80^{\circ}\text{C}$  until the evaluation process.

## 2.6. Quantification by RT-qPCR

Primer design was made using the Primer 3 program following the parameters established by [32] seeking to amplify fragments of approximately 150 bp (Table 1) [33]. The specificity of the primers was verified using the Primer-BLAST program (NCBI).

cDNA synthesized was used to perform the PCR technique in real-time using the Maxima SYBR Green qPCR Master Mix (2 $\times$ ) kit (Thermo Scientific, USA), following the instructions of the manufacturer, in a CFX96 Touch Real-Time PCR (Bio-Rad, Foster City, CA). PCR conditions were denaturation at  $95^{\circ}\text{C}$  every 20 s, hybridization at  $60^{\circ}\text{C}$  every 30 s, and elongation at  $72^{\circ}\text{C}$  every 30 s for 39 cycles. Melting curve was also determined for each PCR product. Difference significance in the gene expression was estimated using the relative expression for each gene in the three developing stages for the two cultivars, considering the differences ( $\Delta\Delta$ ) in the crossing threshold (Ct) [34] and, assuming an efficiency of 2. All data were analyzed using the REST program, version 2009 (Relative Expression Software Tool, Gene Quantification, Munich, Germany) [35], at 5% ( $p < 0.05$ ) level of significance. As housekeeping gene, the ubiquitin was used as a reference, as previously described for castor bean [27].

## 2.7. FAD2 and FAH12 Cloning and Sequencing

Based on the analysis of PPI networks and validation by qPCR, four genes

**Table 1.** Primer sequences for RT-qPCR in castor bean.

Gene	ID_Primer	type	Sequence
Oleoyl-12-Hydroxylase ( <i>FAH12</i> )	FAH_fw	Target	5'-CATTCTCCTATGTTGCCTATGATG-3'
	FAH_rv	Target	5'-AGACCAGTGAGAATGCAGCCT-3'
Oleoyl-12-desaturase ( <i>FAD2</i> )	FAD_fw	Target	5'-GGAACGGGATGAAGTGTTTG-3'
	FAD_rv	Target	5'-TTGAATGCTAGGTACAGAGG-3'
Diacylglycerol acyltransferase ( <i>DGAT2</i> )	DGAT_fw	Target	5'-GCGAGAAGAGGGTTTGTCCGA-3'
	DGAT_rv	Target	5'-CGCCATCAGGTTTCCACCAC-3'
Ubiquitin ( <i>UBI</i> )	UBI_fw	Housekeeping	5'-CGCAAATACAACCAAGACAAGA-3'
	UBI_rv	Housekeeping	5'-CCAGCAAGCACTCTCCATCA-3'

coding enzymes involved in the biosynthesis and accumulation of fatty acids in castor bean were selected. Primer designing for the amplification of these genes was made with Primer 3 software [34] using the default parameters (Table 2).

The pDrive vector (QIAGEN, USA) was used to clone the four genes of interest, following the parameters set by the manufacturers. Subsequently, as a verification mechanism, a plasmid DNA extraction was made with the QIAprep Spin Miniprep kit (QIAGEN Sample and Assay Technologies) and a polymerase chain reaction for amplification of cloned genes was performed. PCR products were sequenced in triplicate by Sanger method the sequences were manually edited using Geneious 5.3.6 created by Biomatters [36] and aligned with the algorithm Clustal W [37] in the same software.

## 2.8. Phylogenetic Analysis

Phylogenetic reconstruction of the desaturated/monooxygenase protein family was performed using the cloned sequence of the *FAH12* gene from castor bean VERC03 to corroborate its identity. *FAH12* and *FAD2* gene sequences of *R. communis*, *A. thaliana*, *Glycine max*, and *Jatropha curcas* among other phylogenetically close plant species were retrieved from GenBank database (see the access numbers of the sequences in Figure 5).

Firstly, the *FAH12* sequence of castor bean VERC03 was translated with Virtual Ribosome 1.1 [38] using the standard genetic code. It was added to the genes dataset in amino acids and then aligned using MAFFT v7.271 [39]. Secondly, nucleotide sequences were aligned with RevTrans2 [40]. Thirdly, these alignments were repartitioned using GBLOCKS 1 [34], and the suggested partitions were compared to the Pfam [41] annotations for the location of the active protein site. Fourthly, the best partitioning scheme was selected according to the evolution models using Partition Finder Protein 1.0.1 [42]. Finally, three topologies were reconstructed: 1) maximum likelihood with Garli (10 independent replicates supported with 1000 bootstrap pseudoreplicates), 2) maximum likelihood with RA × ML (10 independent replicates supported with 1000 bootstrap pseudoreplicates), and 3) Bayesian inference with MrBayes (2 MCMC analyses

**Table 2.** Sequences of the primers used for the amplification of the genes to be cloned.

Gene	Enzyme code	ID_Primer	Sequence
Oleoyl-12-Hydroxylase ( <i>FAH12</i> )	1.14.18.4	FAH_fw	5-ATGGGAGGTGGTGGTCGC-3'
		FAH_rv	5'-TTAATACTTGTTCGGTACCAGA-3'
Oleoyl-12-desaturase ( <i>FAD2</i> )	1.14.19.6	FAD_fw	5'-ATGGGTGCTGGTGGCAGAAT-3'
		FAD_rv	5'-TCAAAATTTGTTGTTATACCAGA-3'
Diacylglycerol acyltransferase ( <i>DGAT2</i> )	2.3.1.20	DGAT_fw	5'-ATGGGGGAAGAAGCGAATCA-3'
		DGAT_rv	5'-TCAAAGAATTTCAAGTGAAGGT-3'
Phospholipid:diacylglycerol acyl transferase ( <i>PDAT1-2</i> )	2.3.1.158	PDAT_fw	5'-ATGTCGATTTTGAGACGGAGA-3'
		PDAT_rv	5'-CTATAGCGGCAAGTTGATCTTCT-3'

with 10,000,000 generations with 4 chains each). All topologies were contrasted using nucleotides and aminoacid seeking to determine inconsistencies.

### 3. Results

#### 3.1. *In silico* PPI Network Construction and Analysis

A combined analysis of *R. communis* PPI networks (interactome) was conducted using an *A. thaliana* network, as a reference, composed by 11,706 proteins of (TAIR), with 24,417 interactions between these proteins. Overall, it was identified 11,192 proteins orthologous to *R. communis*, with significant homology to *A. thaliana* proteins with 23,777 interactions between them. From the protein network obtained for the castor bean plant, FAH12 protein (oleate 12-hydroxylase, genome ID 28035.m000362, E.C. 1.14.18.4) was overexpressed 10.77 times in advanced seed development stages compared with the initial development stages and other tissues (**Figure 2**). Furthermore, it was evident that *FAH12* also interacts with other proteins related to the production of fatty acids, establishing a possible role of this gene in the production of fatty acids in castor bean (**Figure 2(b)**). In addition to this gene, *FAD2*, *DGAT2*, and *PDAT1-2*, genes were identified, such as possible genes involved in two physiological processes: biosynthesis and fatty acid accumulation.

#### 3.2. Fatty Acid Composition and Oil Content - VERC03 and VERC12

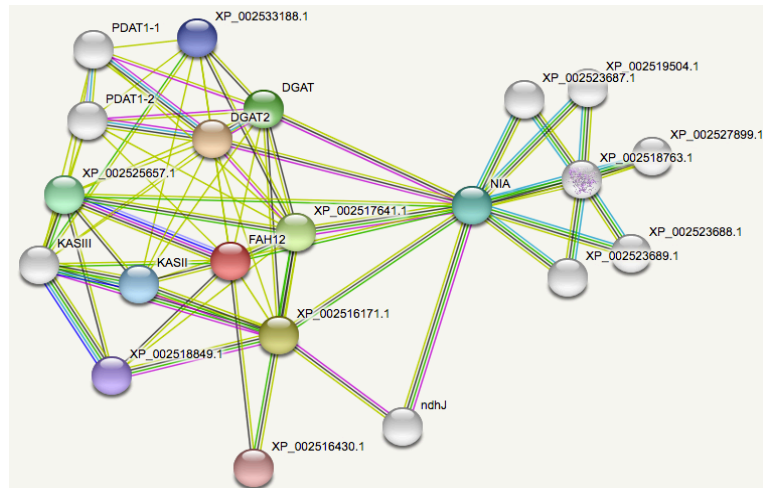
The analytes detected in the chromatogram of the samples were mainly classified into 3 groups of lipids: total saturated (SL), monounsaturated (ML), and polyunsaturated (PL). The quantity of lipids detected for each of the groups was dependent on the stage of development of the seeds, showing a better resolution in the initial stages of development, whereas, in the more advanced stages of development, peaks was hindered by the high content of carbohydrates, leading to a loss in the stability of the baseline at the end of the run (**Figure S2**).

The highest amounts of SL, ML, and PL (2.1226 g/100 g, 2.1342 g/100 g, and 2.6517 g/100 g, respectively) were detected and quantified in the fatty acids of the fruits at developmental stage II (S2) of the tropical VERC03 cultivar, revealing that the production and accumulation of triacylglycerols starts at an early stage during the formation of the fruit (**Figure 3(a)**). Likewise, a direct proportional relationship was determined between the quantities of fatty acids and the size of the fruit until developmental stage III (S3). Analyses of acid types 18:1-OH and 18:2-OH in the tropical VERC03 and VERC12 cultivars showed an increase in oleic and linoleic acids (**Figure 3(b)**).

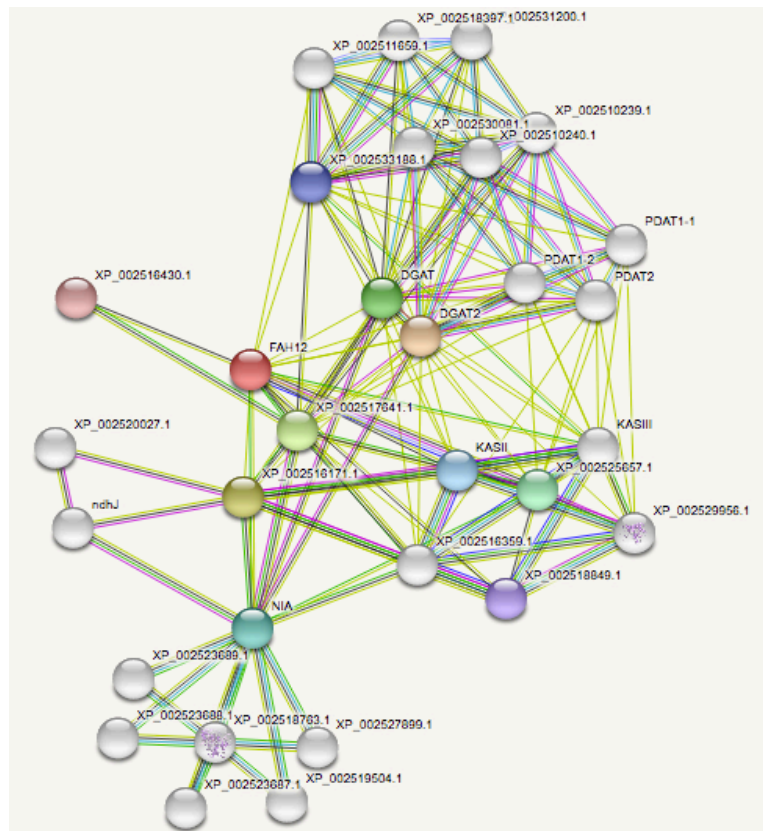
#### 3.3. Quantification by RT-qPCR

The results of gene expression obtained with the RT-qPCR technique for the key biosynthetic *FAH12*, *FAD2*, and *DGAT* genes, in three different fruit developmental stages of the VERC03 and VERC12 cultivars, showed to *FAH12* and *FAD2* transcripts being up-regulated more than 3-folds and 1-fold respectively,





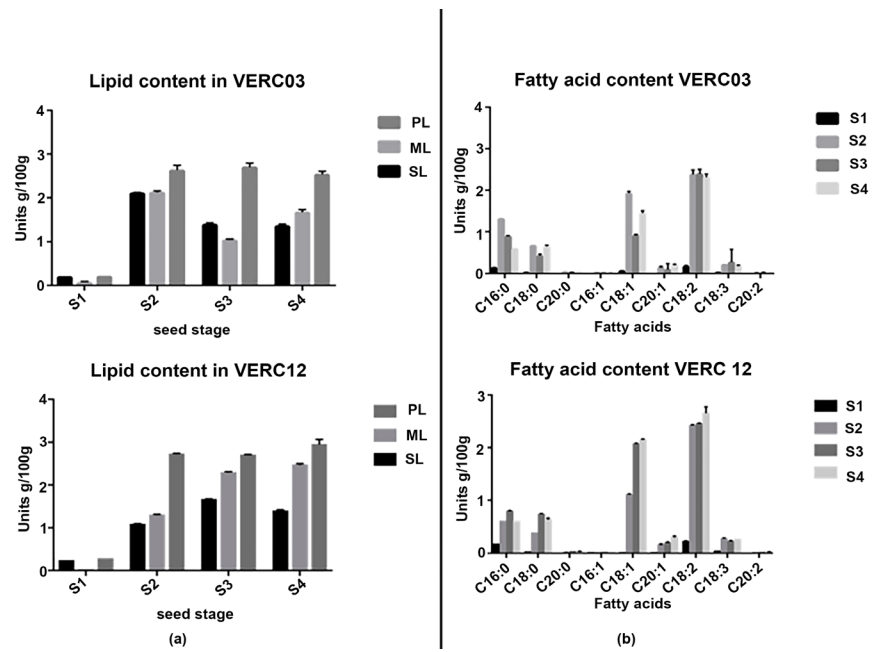
(a)



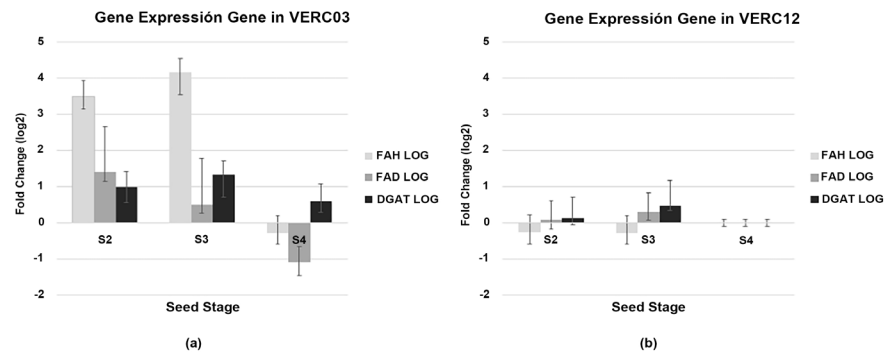
(b)

**Figure 2.** Castor bean interaction protein-protein networks. (a) Some genes involved in the fatty acid pathway of the castor bean enriched with expression data, a green node corresponds to FAH12up-regulated (10.77 times). (b) Gene interaction network related to the FAH12gene (red sphere) using the string interaction platform.

in the stages S2 and S3 for only VERC03 cultivar (**Figure 4**). Expression of these genes in cultivar VERC12 (improved commercial cultivar from Brazil-Embrapa) showed to be statistically invariable during the three developmental stages with



**Figure 3.** Lipid content determined by GC. (a) Total lipid content in g/100 g of sample for four developmental stages from castor bean seed in VERC03 and VERC12 cultivars. (b) Fatty acid content in g/100 g of sample in VERC03 and VERC12 cultivars.



**Figure 4.** Differential expression of FAH12, FAD2, and DGAT2 genes in three stages of the development of the castor bean seed. S2 = Development stage 2; S3 = Development stage 3; S4 = Development stage 4. Relative expression level was calculated comparing to housekeeping ubiquitin gene expression using the  $2^{-\Delta\Delta Ct}$  method.

respect to the UBI housekeeping gene and S1 stage as a reference. DGAT transcript was invariable in all stages and cultivars (Figure 4).

### 3.4. Gene Sequencing

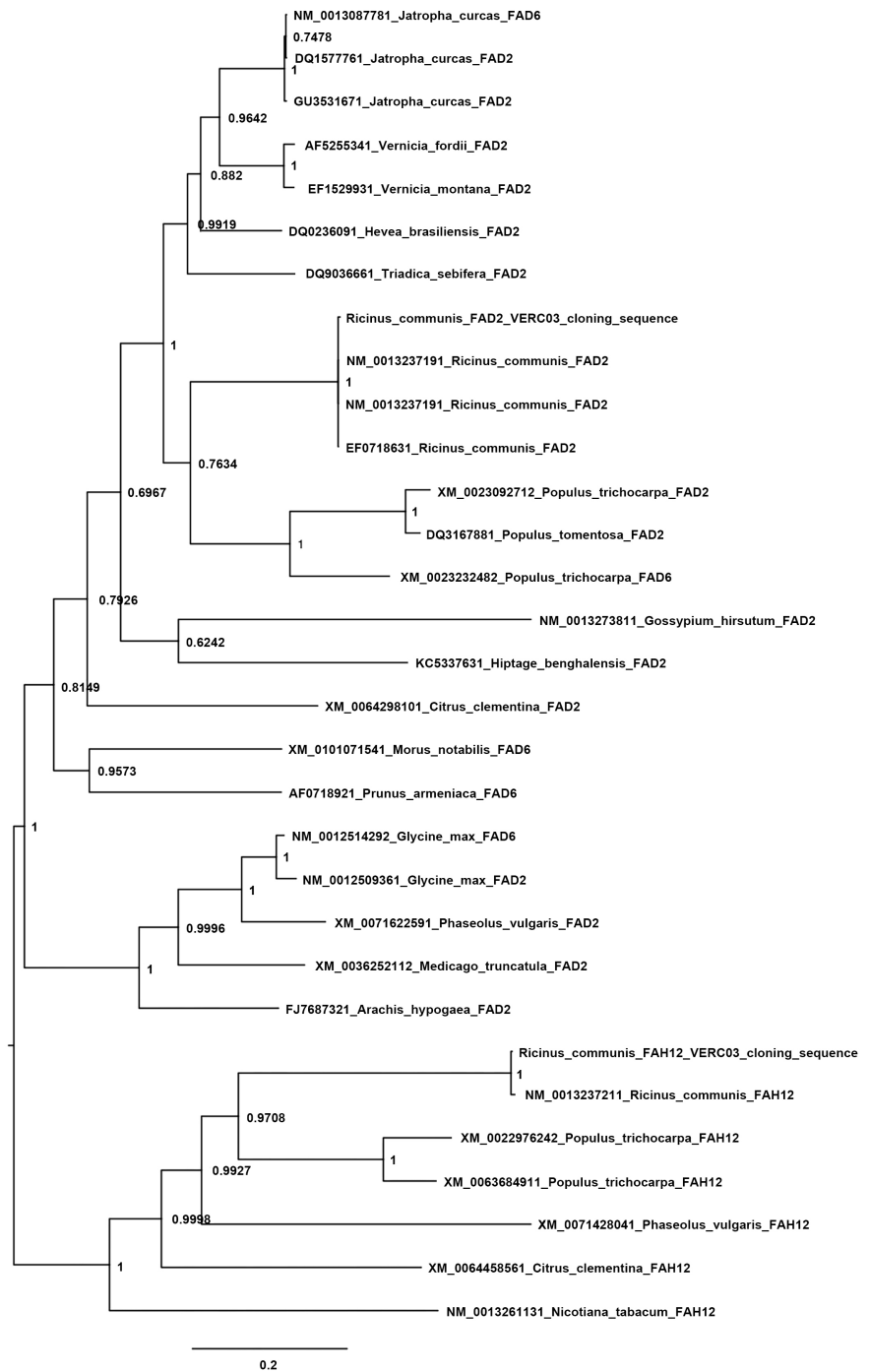
To validate the identity of FAH12 and FAD2 genes, two fragments of 1166 bp and 1151 bp respectively, were obtained from amplification using cDNA from VERC03 cultivar, and then sequenced. The sequences obtained were aligned and comparative with sequences reported in the GenBank. Differences were found in the amino acids A242V and Q319H in FAH12 and F88V in FAD2 (Figure S3). The sequences were reported under the GenBank code MH990329 (FAH12) and MK033959 (FAD2).

### 3.5. Phylogenetic Analysis

A phylogenetic reconstruction was made for the desaturase/monooxygenase protein family using sequences cloned from *FAH12* and *FAD2* of the VERC03 castor bean cultivar and accessions retrieved from GenBank. Two groups were formed with a significant value of maximum likelihood, one with sequences of *FAH12* genes and another derived from this, with *FAD2* group in the top. We corroborated the identity of the clones sequenced and they were located in the same clade with each reference sequence for *R. communis* (Figure 5).

## 4. Discussion

In this research, we studied the gene expression related with the biosynthesis pathway of triacylglycerols in fruits of two castor bean tropical cultivars (VERC-03, originally from Colombia, and VERC-12, originally from Brazil), which were chosen as promising materials adapted to warm agro-ecological zones according to previous agronomic trials developed by the Colombian Agricultural and Livestock Research Corporation (AGROSAVIA, before CORPOICA) (CITA 10), which showed that VERC03 cultivar has slightly higher yields per hectare compared with VERC12 cultivar. In studies conducted by other authors [27], made evidence the need to evaluate in Castor bean plants, the gene expression in different fruit development stages to understand deeply the production of fatty acids, looking forward to determine exceptional conditions in the production of seeds with high oil concentrations for its commercialization. Likewise, previous studies led to the need for deepening the understanding of genetic processes related to the biosynthesis of fatty acids in castor bean cultivars, looking forward to finding out key genes involved in seed oil production, accumulation of long-chain acids in different seeds stages and thereby contribute to crop genetic improvement with high precision and less uncertainty. From the protein interaction network obtained for castor bean, it was determined the proteins related to the biosynthesis pathways of modification and accumulation of fatty acids, which made it possible to determine that the protein *FAH12* (oleate 12-hydroxylase, genome ID 28035.m000362, E.C. 1.14.18.4), a precursor of the pathway responsible for the production of ricinoleic acid, it is key in the production of ricinoleic acid in castor bean plants. It metabolizes the transition of 2-oleoyl phosphatidylcholine to ricinoleoyl phosphatidylcholine [43]. Additionally, in our research was found *FAH12* gene differential expressed in the developmental stage III of the seed gene in contrast to the values obtained in developmental stage IV in a cultivar VERC03. The characterization of *FAH12* in castor bean began to take relevance from 1995 where [14] determined an increase in the differential expression of this gene in the seeds with respect to the leaves. In 2007, [20] studied the expression profiles of genes involved in fatty acid and triacylglycerol synthesis in *R. communis*, wherein they determined an increased expression of *FAH* during seed development; therefore, a direct relationship between the increase in the expression of this gene and the production



**Figure 5.** Phylogenetic tree of FAH12 and FAD2 sequences obtained from tropical castor bean seed VERC03 against other phylogenetically close plant species. Alignment was performed with MAFFT v7.271 and the reconstruction of the tree was performed with maximum likelihood using RAXML software (10 independent replicates supported with 1000 bootstrap pseudoreplicates).

of ricinoleic acid was further determined [20]. Later, were evaluated fatty acid profiles in *Arabidopsis* plants with the FAH12 transgene to increase ricinoleate levels [15] [18] [44].

In our case, a greater amount of oleic and linoleic acid was evidenced in an intermediate-advanced phase of seed development, a condition equivalent to what was observed in the results of the study of gene expression in different phases of development of castor bean fruits of this same cultivar [19]. With regard to VERC12 cultivar, it was detected a progressive accumulation of polyunsaturated fatty acids during fruit maturation, which translates into a greater amount of monounsaturated and polyunsaturated fats in the last stage of development evaluated. On the other hand, seed lipids were characterized on natural castor bean mutant deficient in ricinoleic acid synthesis (OLE-1) and it was identified high levels of palmitic and linoleic acid in the initial development seed stages, but during seed maturation, the ricinoleic acid content came up to 80% of the total fatty acids [44].

The results obtained here for *DGAT2* in the VERC03 cultivar are comparable with those reported in 2010 by [27], who revealed the highest expression of this gene in the fruits at intermediate stages of development. In the same way, analyses of the expression of cDNAs in *FAH12* transgenic plants revealed that the castor bean type-2 acyl-coenzyme A: diacylglycerol acyltransferase (RcDGAT2) could enhance HFAs from 17% to nearly 30%. These results indicate the probable importance of members of the DGAT2 gene family plays a key role in the process to increase the hydroxy fatty acids in castor bean [19].

Moreover, our sequence analysis of *FAH12* gene led to identify two mutations, one of them (H319Q) directly related to changes in its secondary structure and also to enhance ricinoleic acid production as well as was reported by [44], who identified H319Q residue located at *FAH12* His box III (domain IX), recognized as the enzyme catalytic site.

Finally, the phylogenetic reconstruction of *FAD2* and *FAH12* was consistent with what was reported by [14]. Our ML tree clustered both gene sequences with the same genes from *Populus trichocarpa*, which was also to be expected because of the taxonomic closeness of the two species. Their results suggest that *FAD2* and *FAH12*, despite belonging to the same protein family, have a different function based on similar reaction mechanisms, which may imply an enzymatic evolution.

## 5. Conclusion

We successfully used interactome analysis, gas chromatography, qRT-PCR and gene sequencing to develop efficient methods for predicting orthology genes, analytes determination, differential expression of genes and cloning key genes involved in initial and advanced seed development stages, regarding to clarify the biosynthesis pathways and accumulation of fatty acids for a couple of tropical cultivars of castor bean. Despite the fact that we found four important genes involved in the pathway of ricinoleic synthesis in castor bean, it was possible to determine that *FAH12* gene was expressed more than 3 folds in advanced seed development stages compared with the initial stages in the high oil-producing VERC03 cultivar. Additionally, it was found special differences in some aminoa-

cid residues for FAH12 protein, differences involved in structural changes and enhanced ricinoleic acid production. The results obtained here are of vital importance for future genetic characterization of castor bean promising materials, thinking in using those genes and its identified mutations as molecular markers in markers-assisted breeding in castor bean plants. Finally, we suggest FAH12 gene as a key gene for plant breeding programs that seek to enhance ricinoleic acid production.

### Acknowledgements

This work was supported by “Component No. 1 - Biotechnology”, project code 4600 000480 financed by Colombian general system of royalties (SGR, Spanish acronym), in partnership with Secretaría de Agricultura y Desarrollo Rural of Departamento de Antioquia, Universidad Pontificia Bolivariana, Universidad Nacional de Colombia, and Universidad EAFIT, and performed in the molecular biology, computational biology and plant biotechnology laboratories of EAFIT University. We especially thank Dr. Orville Hernandez of the Corporation for Biological Research for the support in qPCR analysis; Lawyer Lina Chamorro Gutierrez for the development of the MTA between EAFIT and AGROSAVIA (before CORPOICA), and for having managed the addendum #1 (June 16, 2017) to the Framework Contract for Access to Genetic Resources and Their Derivative Products #127 of May 13, 2016, between the Ministry of the Environment and Sustainable Development and EAFIT University; the EAFIT Biology student Luis Arteaga for support in the phylogenetic section analysis; as well as AGROSAVIA (before CORPOICA) for the transfer of plant material.

### Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

### References

- [1] Davis, C.C., Latvis, M., Nickrent, D.L., Wurdack, K.J. and Baum, D.A. (2007) Floral Gigantism in Rafflesiaceae. *Science*, **315**, 1812. <https://doi.org/10.1126/science.1135260>
- [2] Wurdack, K.J., Hoffmann, P. and Chase, M.W. (2005) Molecular Phylogenetic Analysis of Uniovulate Euphorbiaceae (Euphorbiaceae Ssensu Stricto) Using Plastid RbcL and TrnL-F DNA Sequences. *American Journal of Botany*, **92**, 1397-1420. <https://doi.org/10.3732/ajb.92.8.1397>
- [3] Roetheli, J.C., Glaser, L.K. and Brigham, R.D. (1990) Castor: Assessing the Feasibility of US Production. *Workshop Summary*, Plainview, September 1990, 18-19.
- [4] Johnson Jr., W. (2007) Final Report on the Safety Assessment of *Ricinus communis* (Castor) Seed Oil, Hydrogenated Castor Oil, Glyceryl Ricinoleate, Glyceryl Ricinoleate Se, Ricinoleic Acid, Potassium Ricinoleate, Sodium Ricinoleate, Zinc Ricinoleate, Cetyl Ricinoleate, Ethyl Ric. *International Journal of Toxicology*, **26**, 31-77. <https://doi.org/10.1080/10915810701663150>
- [5] Stephen, G. (2009) An Investigation into the Components of Triricinoleic Acid

Production in the Developing Castor Bean Endoplasmic Reticulum an Investigation into the Components of Triricinoleic Acid Production in the Developing Castor Bean Endoplasmic Retic. Doctoral E-Theses, Durham University, Durham.

- [6] Lima Da Silva, N., Wolf Maciel, M.R., Batistella, C.B. and Filho, R.M. (2006) Optimization of Biodiesel Production from Castor Oil. *Applied Biochemistry and Biotechnology*, **130**, 405-414. <https://doi.org/10.1385/ABAB:130:1:405>
- [7] Brigham, R.D. (1993) Castor: Return of an Old Crop. *New Crops*. Wiley, New York, 380-383.
- [8] Qiu, L., Yang, C., Tian, B., Yang, J.-B. and Liu, A. (2010) Exploiting EST Databases for the Development and Characterization of EST-SSR Markers in Castor Bean (*Ricinus communis* L.). *BMC Plant Biology*, **10**, Article No. 278. <https://doi.org/10.1186/1471-2229-10-278>
- [9] Allard, R.W., PDeT, A., Ashri, A. and Barton, J.H. (1991) Managing Global Genetic Resources; the US National Plant Germplasm System: Elements of the National Plant Germplasm System. The National Academies Press, Washington DC.
- [10] Navas, A. (2011) Evaluación de Cultivares Foráneos y Generación de Variedades Colombianas de Higuierilla Para La Producción de Biodiesel y Otros Usos En La Industria. CI La Selva, Rionegro.
- [11] Bafor, M., Smith, M., Jonsson, L., Stobart, K. and Stymne, S. (1991) Ricinoleic Acid Biosynthesis and Triacylglycerol Assembly in Microsomal Preparations from Developing Castor-Bean (*Ricinus communis*) Endosperm. *The Biochemical Journal*, **280**, 507-514. <https://doi.org/10.1042/bj2800507>
- [12] Galliard, T. and Stumpf, P.K. (1966) Fat Metabolism in Higher Plants XXX. Enzymatic Synthesis of Ricinoleic Acid by a Microsomal Preparation from Developing *Ricinus communis* Seeds. *Journal of Biological Chemistry*, **241**, 5806-5812.
- [13] Morris, L.J. (1967) The Mechanism of Ricinoleic Acid Biosynthesis in *Ricinus communis* Seeds. *Biochemical and Biophysical Research Communications*, **29**, 311-315. [https://doi.org/10.1016/0006-291X\(67\)90454-8](https://doi.org/10.1016/0006-291X(67)90454-8)
- [14] van de Loo, F.J., Broun, P., Turner, S. and Somerville, C. (1995) An Oleate 12-Hydroxylase from *Ricinus communis* L. Is a Fatty Acyl Desaturase Homolog. *Proceedings of the National Academy of Sciences of the United States of America*, **92**, 6743-6747. <https://doi.org/10.1073/pnas.92.15.6743>
- [15] Broun, P. and Somerville, C. (1997) Accumulation of Ricinoleic, Lesquerolic, and Densipolic Acids in Seeds of Transgenic Arabidopsis Plants That Express a Fatty Acyl Hydroxylase cDNA from Castor Bean. *Plant Physiology*, **113**, 933-942. <https://doi.org/10.1104/pp.113.3.933>
- [16] Thelen, J.J. and Ohlrogge, J.B. (2002) Metabolic Engineering of Fatty Acid Biosynthesis in Plants. *Metabolic Engineering*, **4**, 12-21. <https://doi.org/10.1006/mben.2001.0204>
- [17] Lu, C., Fulda, M., Wallis, J.G. and Browse, J. (2006) A High-Throughput Screen for Genes from Castor That Boost Hydroxy Fatty Acid Accumulation in Seed Oils of Transgenic Arabidopsis. *Plant Journal*, **45**, 847-856. <https://doi.org/10.1111/j.1365-3113X.2005.02636.x>
- [18] Kroon, J.T.M., Wei, W., Simon, W.J. and 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 Animals. *Phytochemistry*, **67**, 2541-2549. <https://doi.org/10.1016/j.phytochem.2006.09.020>
- [19] Bursal, J., Shockey, J., Lu, C., Dyer, J., Larson, T., Graham, I. and Browse, J. (2008)





- Metabolic Engineering of Hydroxy Fatty Acid Production in Plants: RcDGAT2 Drives Dramatic Increases in Ricinoleate Levels in Seed Oil. *Plant Biotechnology Journal*, **6**, 819-831. <https://doi.org/10.1111/j.1467-7652.2008.00361.x>
- [20] Chen, G.Q., Turner, C., He, X., Nguyen, T., McKeon, T.A. and Laudencia-Chingcuanco, D. (2007) Expression Profiles of Genes Involved in Fatty Acid and Triacylglycerol Synthesis in Castor Bean (*Ricinus communis* L.). *Lipids*, **42**, 263-274. <https://doi.org/10.1007/s11745-007-3022-z>
- [21] Imadi, S.R., Kazi, A.G., Ahanger, M.A., Gucl, S. and Ahmad, P. (2015) Plant Transcriptomics and Responses to Environmental Stress: An Overview. *Journal of Genetics*, **94**, 525-537. <https://doi.org/10.1007/s12041-015-0545-6>
- [22] Chan, A.P., Crabtree, J., Zhao, Q., Lorenzi, H., Orvis, J., Puiu, D., Melake-Berhan, A., Jones, K.M., Redman, J., Chen, G., Cahoon, E.B., Gedil, M., Stanke, M., Haas, B.J., Wortman, J.R., Fraser-Liggett, C.M., Ravel, J. and Rabinowicz, P.D. (2010) Draft Genome Sequence of the Oilseed Species *Ricinus communis*. *Nature Biotechnology*, **28**, 951-956. <https://doi.org/10.1038/nbt.1674>
- [23] Rivarola, M., Foster, J.T., Chan, A.P., Williams, A.L., Rice, D.W., Liu, X., Melake-Berhan, A., Creasy, H.H., Puiu, D., Rosovitz, M.J., Khouri, H.M., Beckstrom-Sternberg, S.M., Allan, G.J., Keim, P., Ravel, J. and Rabinowicz, P.D. (2011) Castor Bean Organelle Genome Sequencing and Worldwide Genetic Diversity Analysis. *PLoS ONE*, **6**, e21743. <https://doi.org/10.1371/journal.pone.0021743>
- [24] Brown, A.P., Kroon, J.T.M., Swarbreck, D., Febrer, M., Larson, T.R., Graham, I.A., Caccamo, M. and Slabas, A.R. (2012) Tissue-Specific Whole Transcriptome Sequencing in Castor, Directed at Understanding Triacylglycerol Lipid Biosynthetic Pathways. *PLoS ONE*, **7**, e30100. <https://doi.org/10.1371/journal.pone.0030100>
- [25] Chandrasekaran, U., Xu, W. and Liu, A. (2014) Transcriptome Profiling Identifies ABA Mediated Regulatory Changes towards Storage Filling in Developing Seeds of Castor Bean (*Ricinus communis* L.). *Cell & Bioscience*, **4**, 33. <https://doi.org/10.1186/2045-3701-4-33>
- [26] Bassel, G.W., Gaudinier, A., Brady, S.M., Hennig, L., Rhee, S.Y. and De Smet, I. (2012) Systems Analysis of Plant Functional, Transcriptional, Physical Interaction, and Metabolic Networks. *The Plant Cell*, **24**, 3859-3875. <https://doi.org/10.1105/tpc.112.100776>
- [27] Cagliari, A., Margis-Pinheiro, M., Loss, G., Mastroberti, A.A., de Araujo Mariath, J.E. and Margis, R. (2010) Identification and Expression Analysis of Castor Bean (*Ricinus communis*) Genes Encoding Enzymes from the Triacylglycerol Biosynthesis Pathway. *Plant Science*, **179**, 499-509. <https://doi.org/10.1016/j.plantsci.2010.07.015>
- [28] Perteau, M., Kim, D., Perteau, G.M., Leek, J.T. and Salzberg, S.L. (2016) Transcript-Level Expression Analysis of RNA-Seq Experiments with HISAT, String Tie and Ballgown. *Nature Protocols*, **11**, 1650-1667. <https://doi.org/10.1038/nprot.2016.095>
- [29] Childs, K.L., Hamilton, J.P., Zhu, W., Ly, E., Cheung, F., Wu, H., Rabinowicz, P.D., Town, C.D., Buell, C.R. and Chan, A.P. (2007) The TIGR Plant Transcript Assemblies Database. *Nucleic Acids Research*, **35**, D846-D851. <https://doi.org/10.1093/nar/gkl785>
- [30] Trapnell, C., Hendrickson, D.G., Sauvageau, M., Goff, L., Rinn, J.L. and Pachter, L. (2013) Differential Analysis of Gene Regulation at Transcript Resolution with RNA-Seq. *Nature Biotechnology*, **31**, 46-53. <https://doi.org/10.1038/nbt.2450>
- [31] Fischer, S., Brunk, B.P., Chen, F., Gao, X., Harb, O.S., Iodice, J.B., Shanmugam, D.,



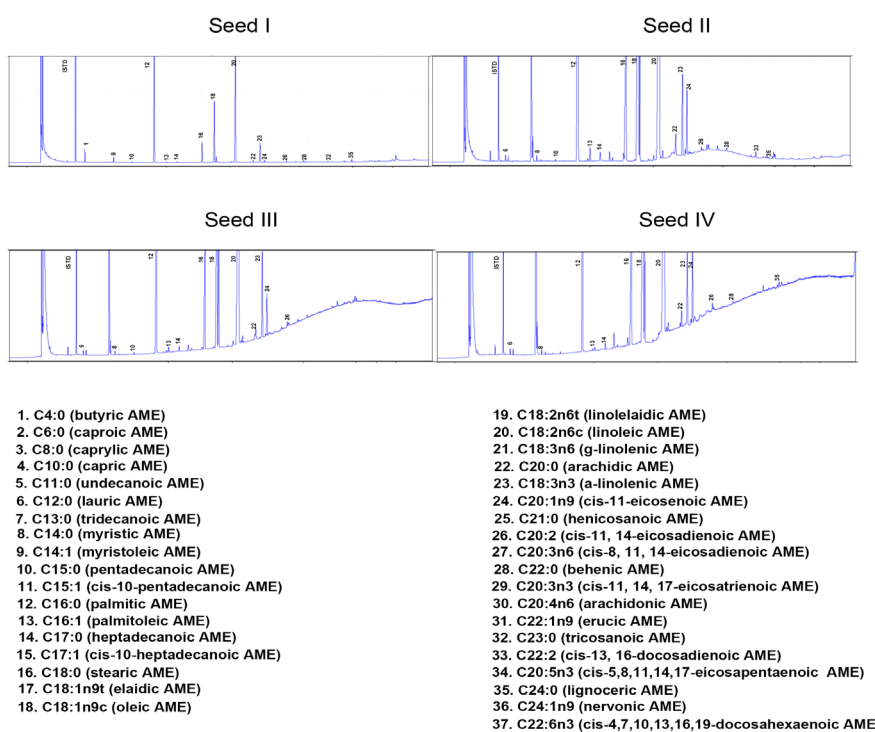
- Roos, D.S. and Stoeckert, C.J. (2011) Using OrthoMCL to Assign Proteins to OrthoMCL-DB Groups or to Cluster Proteomes into New Ortholog Groups. *Current Protocols in Bioinformatics*, 1-23. <https://doi.org/10.1002/0471250953.bi0612s35>
- [32] Bustin, S., Benes, V., Garson, J., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M.W., Shipley, G.L., Vandesompele, J. and Wittwer, C.T. (2009) The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments. *Clinical Chemistry*, **55**, 611-622. <https://doi.org/10.1373/clinchem.2008.112797>
- [33] Rozen, S. and Skaletsky, H. (2000) Primer3 on the www for General Users and for Biologist Programmers. In: Misener, S. and Krawetz, S.A., Eds., *Bioinformatics Methods and Protocols*, Humana Press, Totowa, 365-386. <https://doi.org/10.1385/1-59259-192-2:365>
- [34] Talavera, G. and Castresana, J. (2007) Improvement of Phylogenies after Removing Divergent and Ambiguously Aligned Blocks from Protein Sequence Alignments. *Systematic Biology*, **56**, 564-577. <https://doi.org/10.1080/10635150701472164>
- [35] Pfaffl, M.W., Horgan, G.W. and Dempfle, L. (2002) Relative Expression Software Tool (REST©) for Group-Wise Comparison and Statistical Analysis of Relative Expression Results in Real-Time PCR. *Nucleic Acids Research*, **30**, e36. <https://doi.org/10.1093/nar/30.9.e36>
- [36] Geneious. Bioinformatics Software for Sequence Data Analysis. <https://www.geneious.com>
- [37] Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J. and Higgins, D.G. (2007) Clustal W and Clustal X Version 2.0. *Bioinformatics*, **23**, 2947-2948. <https://doi.org/10.1093/bioinformatics/btm404>
- [38] Wernersson, R. (2006) Virtual Ribosome—A Comprehensive DNA Translation Tool with Support for Integration of Sequence Feature Annotation. *Nucleic Acids Research*, **34**, W385-W388. <https://doi.org/10.1093/nar/gkl252>
- [39] Katoh, K. and Standley, D.M. (2013) MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution*, **30**, 772-780. <https://doi.org/10.1093/molbev/mst010>
- [40] Wernersson, R. and Pedersen, A.G. (2003) RevTrans: Multiple Alignment of Coding DNA from Aligned Amino Acid Sequences. *Nucleic Acids Research*, **31**, 3537-3539. <https://doi.org/10.1093/nar/gkg609>
- [41] Finn, R.D., Coghill, P., Eberhardt, R.Y., Eddy, S.R., Mistry, J., Mitchell, A.L., Potter, S.C., Punta, M., Qureshi, M., Sangrador-Vegas, A., Salazar, G.A., Tate, J. and Bateman, A. (2016) The Pfam Protein Families Database: Towards a More Sustainable Future. *Nucleic Acids Research*, **44**, D279-D285. <https://doi.org/10.1093/nar/gkv1344>
- [42] Lanfear, R., Frandsen, P.B., Wright, A.M., Senfeld, T. and Calcott, B. (2016) PartitionFinder 2: New Methods for Selecting Partitioned Models of Evolution for Molecular and Morphological Phylogenetic Analyses. *Molecular Biology and Evolution*, **34**, 772-773. <https://doi.org/10.1093/molbev/msw260>
- [43] Lin, J.T., Woodruff, C.L., Lagouche, O.J., McKeon, T.A., Stafford, A.E., Goodrich-Tanrikulu, M., Singleton, J.A. and Haney, C.A. (1998) Biosynthesis of Triacylglycerols Containing Ricinoleate in Castor Microsomes Using 1-Acyl-2-Oleoyl-Sn-Glycero-3-Phosphocholine as the Substrate of Oleoyl-12-Hydroxylase. *Lipids*, **33**, 59-69. <https://doi.org/10.1007/s11745-998-0180-3>
- [44] Venegas-Calación, M., Sánchez, R., Salas, J.J., Garcés, R. and Martínez-Force, E.

(2016) Molecular and Biochemical Characterization of the OLE-1 High-Oleic Castor Seed (*Ricinus communis* L.) Mutant. *Planta*, **244**, 245-258.  
<https://doi.org/10.1007/s00425-016-2508-4>

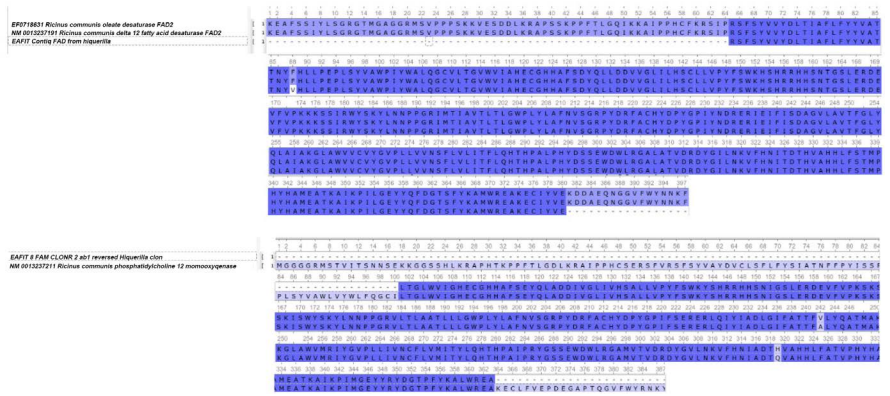
## Supplemental Material

Development stage	S1	S2	S3	S4
Seed photo				
Days	10	20	40	60

**Figure S1.** Development stages of the castor bean seeds selected from the cultivar VERC12. S1 to S4 represent seed stages based on development time of seeds and their average size: image shows that S1 was defined by 10 days of fruit development with a white testa color; in S2 the seed have a minor size but have changes of testa color; in S3 was defined by 40 days of fruit development with a mosaic brown and black testa color; and S4 was defined by 60 days of fruit development with capsules senesce and desiccate, and less seed weight and capsule size.



**Figure S2.** Chromatograms obtained from different development stages of the castor bean seeds selected from the cultivar VERC03.



**Figure S3.** Multiple Sequence Alignment with MAFFT software of FAD2 (A, mutation found in F88V) and FAH12 (B, mutations found in A242V and Q349H) sequences from *Ricinus communis* VERC03 (sequencing by EAFIT university).