

Contents lists available at [ScienceDirect](http://ScienceDirect.com)

Electronic Journal of Biotechnology



Research article

Over-expression of CYP78A98, a cytochrome P450 gene from *Jatropha curcas* L., increases seed size of transgenic tobacco



Yinshuai Tian, Min Zhang, Xiaole Hu, Linghui Wang, Jiao Dai, Ying Xu *, Fang Chen *

National and Local Joint Engineering Laboratory for Energy Plant Bio-oil Production and Application, Key Lab of Bio-Resources and Eco-Environment, Ministry of Education, College of Life Science, Sichuan University, Chengdu 610064, China

ARTICLE INFO

Article history:

Received 30 August 2015

Accepted 9 November 2015

Available online 26 November 2015

Keywords:

Biodiesel

Breeding

CYP78A5

Jatropha

ABSTRACT

Background: *Jatropha curcas* L. (further referred to as *Jatropha*), as a rapidly emerging biofuel crop, has attracted worldwide interest. However, *Jatropha* is still an undomesticated plant, the true potential of this shrub has not yet been fully realized. To explore the potential of *Jatropha*, breeding and domestication are needed. Seed size is one of the most important traits of seed yield and has been selected since the beginning of agriculture. Increasing the seed size is a main goal of *Jatropha* domestication for increasing the seed yield, but the genetic regulation of seed size in *Jatropha* has not been fully understood.

Results: We cloned CYP78A98 gene from *Jatropha*, a homologue of CYP78A5 in *Arabidopsis*. We found that CYP78A98 was highly expressed in male flower, female flower, stem apex, leaf and developing seed. However, its transcripts were hardly detected in root and stem. CYP78A98 protein localized in endoplasmic reticulum (ER) and the hydrophobic domain at the N-terminus was essential for the correct protein localization. Furthermore, INNER NO OUTER promoter (*pINO*) drove specific overexpression of CYP78A98 in transgenic tobacco seeds resulted in increased seed size and weight, as well as improved seed protein and fatty acid content.

Conclusions: The results indicated that CYP78A98 played a role in *Jatropha* seed size control. This may help us to better understand the genetic regulation of *Jatropha* seed development, and accelerate the breeding progress of *Jatropha*.

© 2015 Pontificia Universidad Católica de Valparaíso. Production and hosting by Elsevier B.V. All rights reserved.

1. Introduction

Seed size is one of the most important agronomic traits in crop domestication [1]. Plant seeds are formed by the coordinated growth of maternal integument and zygotic tissues developed from double fertilization and seed size is influenced by a number of factors. Parent-of-origin effectors were reported to influence seed size by the ratio of paternal genome [2]. *SHORT HYPOCOTYL UNDER BLUE1* (*SHB1*), *HAIKU2* (*IKU2*) and *MINISEED3* (*MINI3*) have been identified as zygotically acting factors [3, 4]. In addition, WRKY transcription factor *TRANSPARENT TESTA GLABRA2* (*TTG2*), *APETALA2* (*AP2*), *AUXIN RESPONSE FACTOR2* (*ARF2*) and the ubiquitin interaction motif containing DA1 protein have been demonstrated to act maternally to regulate seed size [5,6,7,8,9].

Recently, CYP78A5, a cytochrome P450 gene belongs to CYP78A subfamily from *Arabidopsis*, was reported to have a function in seed and fruit size control [10, 11]. CYP78A5 was first isolated from a screening for genes differentially expressed in early flower development between *Arabidopsis* and cauliflower [12]. CYP78A5 was

found to express in the inner integument of developing ovules and act maternally to stimulate cell proliferation in *Arabidopsis*, thus determining the growth potential of the seed coat and seed [11]. INNER NO OUTER promoter (*pINO*) drove specific overexpression of CYP78A5 in outer integuments could increase seed size in *Arabidopsis* [11]. In addition to the enlarged seed size, over-expression of CYP78A5 led to higher relative oil content [11]. More recently, it was reported that *SOD7/NGAL2* and *DPA4/NGAL3* act redundantly to regulate seed size by directly repressing CYP78A5 expression in *Arabidopsis*, indicating the important role of CYP78A5 in seed size regulation [13]. In addition, CYP78A5 could also regulate vegetative organ size and development [14]. In situ hybridization showed that CYP78A5 was strongly expressed in the peripheral region of the vegetative shoot apical meristems, defining a boundary between the central meristematic zone and the developing organ primordia [12]. The knock-out line of CYP78A5 showed a higher rate of leaf initiation but smaller size of leaves, sepals, and petals [15]. In contrast, over-expression of CYP78A5 could increase organ size [16]. The size changes in seeds and vegetative organs were due to the number of cells but not the cell size. CYP78A5 was proposed to regulate organ size through controlling the arrest of growth by proliferation [11,16]. Expression of CYP78A5 is outside of the organ proliferation region, which suggests that CYP78A5 acts in a non-cell-autonomous way

* Corresponding authors.

E-mail addresses: xuying@scu.edu.cn (Y. Xu), chenfang@scu.edu.cn (F. Chen).
Peer review under responsibility of Pontificia Universidad Católica de Valparaíso.

[12,16]. CYP78A5 protein localized to the endoplasmic reticulum (ER) and did not appear to move to the regions of cell proliferation, therefore it was proposed that CYP78A5 protein could generate a mobile growth-promoting signal to regulate cell proliferation [16,17]. However, the downstream mobile signal has not yet been identified.

Jatropha is a multi-purpose small tree or large shrub belongs to the Euphorbiaceae family and is found throughout the tropical and subtropical regions. *Jatropha* produces seeds with high oil content (40–60%) [18] and the seed oil can be easily converted into biodiesel, meeting US and European standards [19]. As a fast developing alternative eco-friendly fuel, biodiesel has gained more importance in many developed and developing countries of the world [20]. Besides its high seed oil content [21], *Jatropha* was thought as one of the most promising biodiesel plant species also due to its drought tolerance [22], rapid growth, easy propagation, and wide range of environmental adaptation.

Jatropha holds the potential to be a biodiesel producing plant, however, it should still be considered as a wild, undomesticated plant [23]. To explore the full potential of *Jatropha* and to reduce the risk of future unsustainable practices, breeding and domestication are primordial [23]. Seed size is one of the most important agronomic traits, but researches focusing on *Jatropha* seed development and seed size controlling are still quite rare. In this study, we reported the cloning and function characterization of CYP78A98 gene from *Jatropha*, a homologue of CYP78A5 in *Arabidopsis*. We found that the expression level of CYP78A98 showed significant difference in different tissues and development stages. Furthermore, *pINO* drove specific overexpression of CYP78A98 in transgenic tobacco resulted in increased seed size and weight. The transgenic seeds accumulated more protein and fatty acid.

2. Materials and methods

2.1. Plant materials

Jatropha roots, stems, leaves, flower buds, and developing fruits were collected from Xichang, Sichuan Province, China. All samples used for quantitative reverse transcriptase-polymerase chain reaction (qRT-PCR) were immediately frozen in liquid nitrogen (N₂) and stored at -80°C until needed.

Nicotiana benthamiana (*N. benthamiana*) seeds were surface-sterilized with 25% bleach and planted on Murashige-Skoog (MS) medium for germination, then transplanted into soil and maintained at 25°C with a 16 h light/8 h dark cycle.

2.2. RNA isolation and cDNA synthesis

Total RNA was isolated using RNAPrep pure Plant Kit (Tiangen, China). The quality and quantity of RNA were checked by 1.0% agarose gel electrophoresis and spectrophotometry (NanoVue ND-1000, China) respectively. First strand cDNA was synthesized using the Primescript™ RT reagent Kit (Takara, Dalian, China), according to the manufacturer's instructions.

2.3. Cloning and sequence analysis of full-length coding sequence

In order to identify the homologue gene of CYP78A5 in *Jatropha*, we used the CYP78A5 gene from *Arabidopsis* as a query sequence to BLAST against the *Jatropha* nucleotide database (<http://www.ncbi.nlm.nih.gov/>). According to the sequence we got, gene-specific primers Jc-cyp-F and Jc-cyp-R (Table 1) were designed and used to amplify the full-length coding sequence. PCR was performed according to the following procedures: 94°C for 5 min, followed by 35 cycles of 98°C for 10 s, 55°C for 10 s, 72°C for 20 s, and a final extension at 72°C for 5 min. The products were sub-cloned into pMD19-T vectors (Takara, Dalian, China) and sequenced.

Table 1
Primers for PCR used in this paper.

Primers	Sequences (5'–3')
Jc-cyp-F	CGGGATCCATGTCCCAAGATTTCTGTTT ^a
Jc-cyp-R	CGAGCTTTC AAGCAGAAGCCCTAGAAAG ^a
Jc-qcyp-F	GCTGTCTTATGGGAAATGATCT
Jc-qcyp-R	CTGAATCGGAGACTAATCCGT
Jc-Actin-F	AGCAACTGGGATGACATGGAG
Jc-Actin-R	CTACCTTGATCTTCATGCTG
proINO-F	CCC AAGCTT TGGGTCGGTAAATATGTGATCT ^a
proINO-R	GCTCTAGAAGAGAGTGTGTGTGACAGTGAAT ^a
T-cyp-F	GTTC AATCCAAAACCGCC
T-cyp-R	ATCCTATGCTCTCTACTACC

^a Emphatic sequence of AAGCTT, TCTAGA, GGATCC and GAGCTC represented the restriction enzyme sites of Hind III, Xba I, BamH I and Sac I, respectively.

The properties of deduced amino acid sequence were estimated using ProtParam (<http://www.expasy.ch/tools/protparam.html>) [24]. Subcellular localization was predicted by Cell-PLoc2.0 (<http://www.csbio.sjtu.edu.cn/bioinf/Cell-PLoc-2/>). The deduced protein and other related proteins from different plants were aligned using the DNAMAN Software package (Version 5.2.9, Canada). Phylogenetic tree was constructed using the neighbor-joining method with MEGA 5.0 program [25]. Putative cis-acting elements in the 5'-flanking region of initiation code ATG were analyzed using PlantCARE (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html>) and PLACE (<http://www.dna.affrc.go.jp/PLACE/signalscan.html>) databases.

2.4. Subcellular localization of CYP78A98 protein

The complete ORF or 5' mutant sequence of CYP78A98 without the stop codon was amplified by L-cyp-F and L-cyp-R or L-mcyp-F and L-cyp-R, respectively (Table 1). The amplified product was double digested with BamH I and Sal I enzymes and ligated into PBI221-GFP vector digested with the same two restriction enzymes. The reconstructed vectors were used for subcellular localization. The isolation and transfection of *Arabidopsis* leaf mesophyll protoplasts were conducted as described by Yoo et al. [26]. The GFP fluorescence was examined and photographed using a Leica SP5 confocal microscope (Leica, Germany).

2.5. qPCR

RNA extraction and first-strand cDNA synthesis were performed using the methods described above. qRT-PCR was carried out using primers Jc-qcyp-F and Jc-qcyp-R (Table 1), and *Jatropha* ACTIN gene or tobacco ACTIN gene was used as internal control. The primers were tested for linearity by constructing standard curves on five serial 10-fold dilutions in order to ensure maximum specificity and efficiency during qRT-PCR. qRT-PCR was performed using a Bio-Rad iCycler MyiQ Real-Time PCR System in a 10 µL reaction volume containing 1 µL appropriate diluted cDNA from each sample, 5 µL iTaq™ Universal SYBR™ Green Supermix (Bio-Rad, USA), 0.5 µL forward primer, 0.5 µL reverse primer and 3 µL ddH₂O. PCR conditions were: 95°C for 30 s, 40 cycles of 95°C for 5 s, 55°C for 10 s and 72°C for 20 s, followed by a final melt curve profile (65–95°C). qRT-PCR analysis was performed in triplicates and the mean value was used for qRT-PCR analysis. The relative expression of CYP78A98 was calculated according to the method of 2^{-ΔΔCt}. The Ct values for both the target and the integral control gene were the means of triplicate independent PCRs.

2.6. Construction of plant expression vectors and development of transgenic tobacco lines

To generate a seed-specific over-expression construct, the full-length coding sequence of CYP78A98 and *Arabidopsis pINO* was cloned into

pMD19-T vectors (Takara, Dalian, China) using primers Jc-cyp-F, Jc-cyp-R and proINO-F, proINO-R (Table 1), respectively. The cloned CYP78A98 and *pINO* was then successively sub-cloned into the pBI121 vector between the BamH I and Sac I restriction sites and between the Hind III and Xba I restriction sites, respectively. The reconstructed vectors were named as pBI121-INO:CYP and transferred to *Agrobacterium tumefaciens* (GV3101) competent cells using the freeze-thaw method. Finally, the *A. tumefaciens* strain harboring reconstructed vector was used to generate transgenic tobacco via the leaf-disc method as described by Horsch et al. [27]. The transformants were selected using 50 mg L⁻¹ kanamycin. Leaves were collected for RNA extraction and RT-PCR to confirm the transformation when the kanamycin-resistant transgenic plantlets growth to a length of about 5 cm. T-cypA-F and T-cypA-R (Table 1) primers were used for RT-PCR. The PCR-positive plantlets were transplanted into soil for growing in the greenhouse. Seeds were harvested from inbred lines, and transgenic T2 lines were used in subsequent experiments.

2.7. Protein and fatty acid analysis

Total protein was extracted as described previously with little modification [28]. Briefly, 20 seeds from wild-type or transgenic lines were homogenized with a pellet pestle motor in 100 µL extraction buffer. The extraction buffer contains 63 mM Tris-buffer (PH 7.8), 0.5 M NaCl and 0.07% mercaptoethanol (v/v). Homogenates were centrifuged for 2 min at 15,600 × g. The supernatant was used for protein content detection. Protein content was determined using the Pierce™ BCA Protein Assay Kit (Thermo scientific, America) according to the manufacturer instructions.

Total fatty acids were extracted and analyzed as described previously [29]. Total fatty acids of 20 seeds from wild-type or transgenic lines were extracted using 1 mL of 5% (v/v) concentrated sulfuric acid in MeOH, 25 µL of BHT solution (0.2% butylated hydroxyl toluene in MeOH), and 300 µL of toluene. 10–100 µg of triheptadecanoic acid was added as a triacylglycerol internal standard. The mixture was vortexed for 30 s then heated at 90–95°C for 1.5 h. After cooling to room temperature, 1.5 mL of 0.9% NaCl was added and FAMES were extracted with 2 mL hexane. The FAME extracts were analyzed by GC-MS 2010 with a flame ionization detector (FID) on a HP-88 column. The GC conditions were: injector and flame ionization detector temperature, 260°C; oven temperature program 120°C for 2 min, then increasing at 8°C per min to 180°C and holding this temperature for 2 min, finally increasing at 12°C per min to 235°C and holding this temperature for 8 min. Total fatty acid content was estimated by comparing the total fatty acid methyl ester peak area to that of the C17:0 internal standard.

2.8. Statistical analysis

Experiments were arranged using a completely randomized design. Data were analyzed by analysis of variance (ANOVA) using SPSS version 17.0.

3. Results

3.1. Gene isolation and sequence analysis

Using the CYP78A5 gene from *Arabidopsis* as a query sequence, we got a homology sequence from *Jatropha* nucleotide database (<http://www.ncbi.nlm.nih.gov/>) through BLAST program. The mRNA sequence (XM_012225018) belongs to the CYP78A subfamily and contains a complete coding sequence (CDS) of 1596 bp. The CDS encodes a putative polypeptide of 531 amino acids, which has the pI and mass values of 6.69 and 60.2 kDa, respectively. This putative CYP78A gene was designated as CYP78A98 [30], according to the cytochrome nomenclature [31].

Alignment analysis of CYP78A98 protein sequence with related sequences revealed that CYP78A98 and related proteins all had a hydrophobic domain near the N-terminus, which is believed to be necessary for membrane association (Fig. 1), and this feature conformed to the predicted location of CYP78A98 protein. The heme binding site, common to all members of the cytochrome P450 superfamily, was identified at the C-terminus and this domain was largely conserved within the other members of the CYP78A subfamily. Furthermore, an oxygen binding domain was predicted (Fig. 1).

A phylogenetic tree was also constructed to determine the evolutionary relationship between CYP78A98 and the other related homology proteins (Fig. 2). Proteins from monocots and dicots were clustered into different groups and CYP78A98 protein was most closely related to protein from *Ricinus communis*. This result was consistent with the evolution process of these species.

3.2. CYP78A98 protein localizes to the ER

CYP78A98 was a membrane protein located on the ER, according to the prediction of Cell-PLoc2.0 (<http://www.csbio.sjtu.edu.cn/bioinf/Cell-PLoc-2/>). To confirm the CYP78A98 protein location, we expressed a CYP78A98-GFP fusion protein under the control of the 35S promoter in *Arabidopsis* protoplast. Intracellularly, GFP fluorescence was detected in a mesh-like pattern in the cytoplasm (Fig. 3c), which is typical for ER-associated proteins [32]. The hydrophobic domain near the N-terminus was conserved across different plants and was believed responsible for the correct protein location. In order to confirm the function of the hydrophobic domain, we fused the hydrophobic domain mutant protein (73 amino acids at N-terminus were deleted) with GFP. The mutant protein adopted localization and formed a small spherical body (Fig. 3m), which indicated that the hydrophobic domain near the N-terminus was essential for the correct protein location.

3.3. Expression patterns of CYP78A98

qRT-PCR was performed to analyze the expression patterns of CYP78A98. *Jatropha* is a monoecious plant and CYP78A98 was most highly expressed in male flowers. Relatively higher levels of transcripts were detected in leaves, stem apexes, female flowers and developing seeds, however, CYP78A98 transcripts were hardly detected in stem and root (Fig. 4a).

Given the high expression level of CYP78A98 in male and female flowers, CYP78A98 must have an essential role in flower development. To get a further insight into the function CYP78A98 may have during the flower development, the development process of male flowers before pollen dispersal and that of female flowers before pollination were both divided into 3 stages. We examined the expression level of CYP78A98 gene at different female and male flower developmental stages (Fig. 4b). Compared with female flower, male flower had a relatively higher level of expression at all stages examined. The amounts of CYP78A98 transcripts kept increasing during the male flower development and remained at a high level before pollen dispersal. However, expression of CYP78A98 in female flower had an opposite tendency (Fig. 4b).

Considering the function of CYP78A subfamily in seed size control, we further analyzed the transcription pattern of CYP78A98 during *Jatropha* seed development. The process of *Jatropha* seed development was divided into 6 stages according to the days of after pollination (DAP). Seed size increased rapidly during the first 3 stages, and marginal increase in seed size was observed during 30 DAP to seed maturation. CYP78A98 was highly expressed when seed size was rapidly increasing at the early development stages. The expression level decreased sharply at 30 DAP, at which stage seed had reached its final size (Fig. 4c).

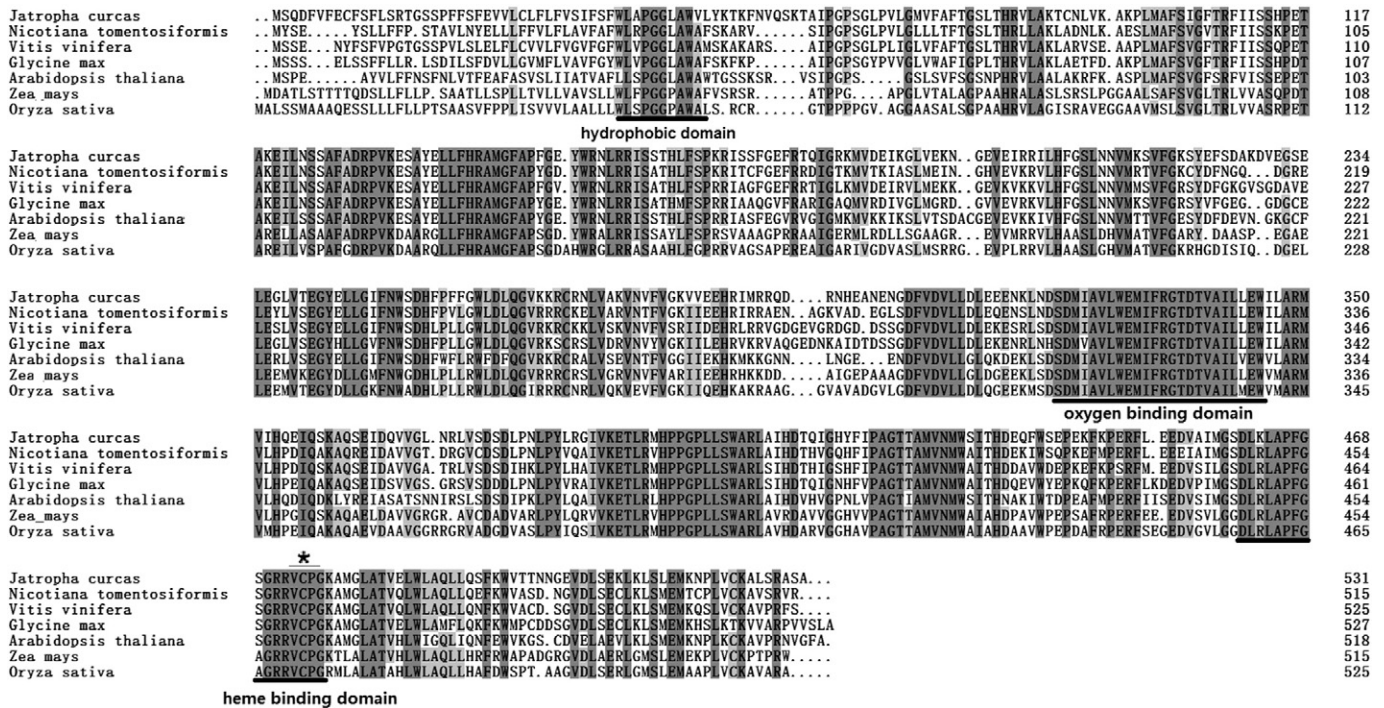


Fig. 1. Alignment of the deduced CYP78A98 protein with closely related protein from other plants. Identical and similar amino acids were shaded in dark gray and gray, respectively. The function domains are underlined. The conserved cysteine residues in heme binding domain are indicated by asterisk (*). *Nicotiana tomentosiformis* (XP_009617550.1), *Vitis vinifera* (XP_002265310.1), *Glycine max* (XP_003550706.1), *Arabidopsis thaliana* (ABJ17153.1), *Zea mays* (XP_008656985.1), *Oryza sativa* (NP_001060215.1), *Jatropha curcas* (XP_012080408.1).

3.4. Analysis of putative cis-acting elements in the CYP78A98 5'-flanking region

To elucidate the mechanism underlying the expression patterns of CYP78A98, a 1500 bp sequence of 5'-flanking region of CYP78A98 was obtained from *Jatropha* genome database (<http://www.kazusa.or.jp/jatropha/>). The Plant CARE and PLACE databases were used to analyze the putative cis-acting elements of this region. Identified cis-elements are shown in (Table 2). Several putative cis-acting elements may be involved in plant hormone mediated development regulation,

including TATC-box and CARE-motif involved in the gibberellins responsiveness, ARFAT-element involved in auxin responsiveness, and CPBSPOR-element is a cytokinin-dependent protein binding site. GT1-motif, ACE-element and AT1-motif were identified as light responsive element. In addition, leaf development related elements and pollen-specific expression elements were also identified, such as the POLLEN1LELAT52 element which involved in pollen specific activation, HD-zip 2 mediated leaf morphology development and so on. These data suggested that CYP78A98 may play an important role in plant hormone mediated leaf and pollen development control.

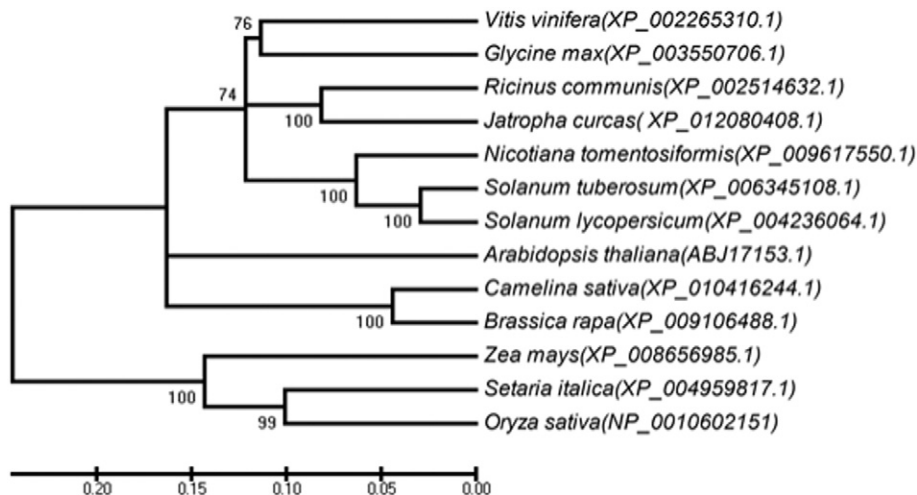


Fig. 2. Phylogenetic analysis of CYP78A98. The phylogenetic tree was constructed using the software MEGA 5.1. Neighbor joining method was used and bootstrap analysis was performed with 1000 replications. Bootstrap values above 70 were indicated at branch points.

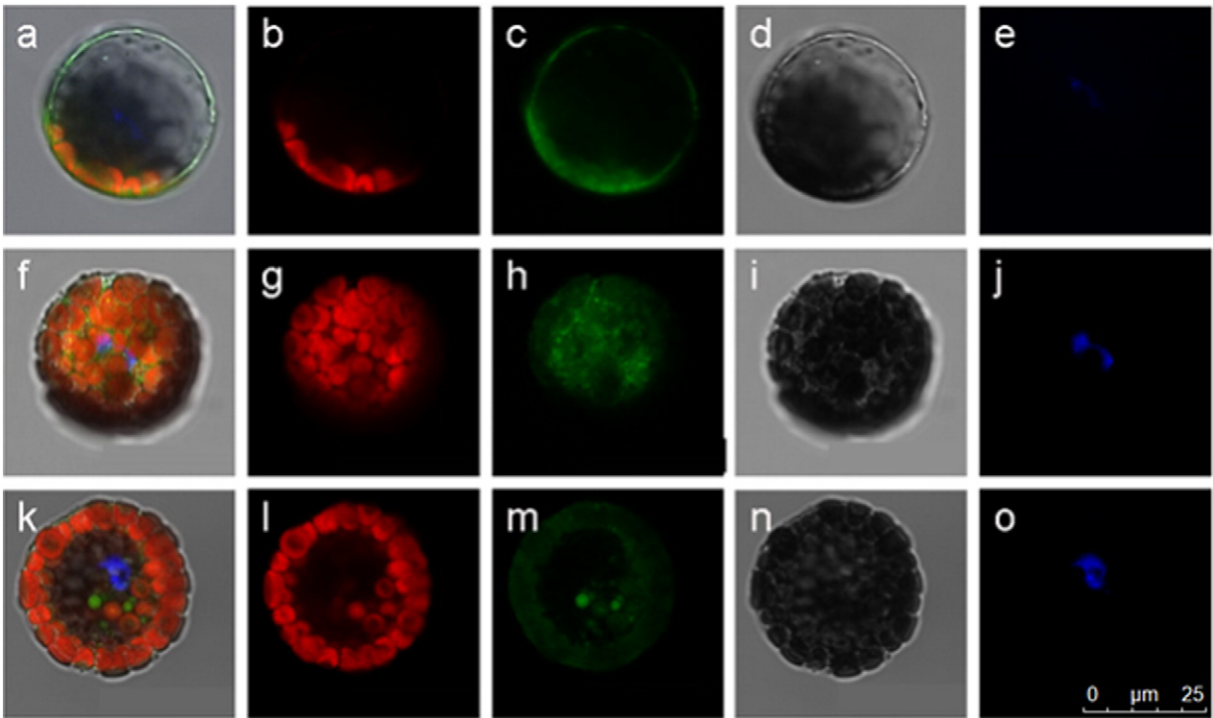


Fig. 3. Subcellular localization of CYP78A98 protein. (a–e) CYP78A98-GFP fusion construct; (f–j) pBI221-GFP empty vector; (k–o) mutated CYP78A98-GFP fusion construct. Merged images (a, f and k); auto fluorescence of chloroplasts (b, j and l); GFP fluorescence (c, h and m); bright field (d, i and n); DAPI staining (e, j and o).

3.5. Identification of transgenic plants

To further illuminate the function of CYP78A98 in seed size control, we constructed the pBI121-INO:CYP vector (Fig. 5a) and transgenic tobacco plants over-expressing CYP78A98 under the control of pINO were obtained by kanamycin selection and confirmed by RT-PCR. Seeds were harvested from inbred lines. Transcript abundance of transgenic T2 lines were analyzed by qRT-PCR and three of the transgenic lines (OE-2, OE-5 and OE-6) with high CYP78A98 expression level were selected and used for subsequent experiments (Fig. 5b). The transgenic lines had no observed phenotype.

3.6. Over-expression of CYP78A98 increases seed size and weight

To further investigate the function of CYP78A98 in seed size control, we examined the seed size of the transgenic lines and wild type plants. Compared with wild type plants, seeds from all transgenic lines show larger size. We further determined the average seed mass of

transgenic plants. Transgenic tobacco had larger seeds (Fig. 6a) and transgenic seeds were 1.4- to 1.6-fold heavier than the wild type (Fig. 6b). By contrast, the number of seeds per fruit was not changed for wild type plants and transgenic lines (Fig. 6c).

3.7. Over-expression of CYP78A98 produces seeds with more protein and fatty acid

Protein and seed oil are main seed reserves, seed mass increase in transgenic tobacco may be due in part to an increase in seed reserves. We extracted seed protein from 20 seeds each of wild type plants or transgenic plants and examined the protein content. Seeds from transgenic plants accumulated more protein than those from wild type, when expressed as μg/seed (Fig. 6d). Nonetheless, no difference in protein content was observed when expressed on the basis of seed weight, and the increase in protein content is obviously due to the increase of seed mass.

We next measured the content and composition of total fatty acid methyl-ester of Ws wild type and transgenic tobacco seeds. Tobacco

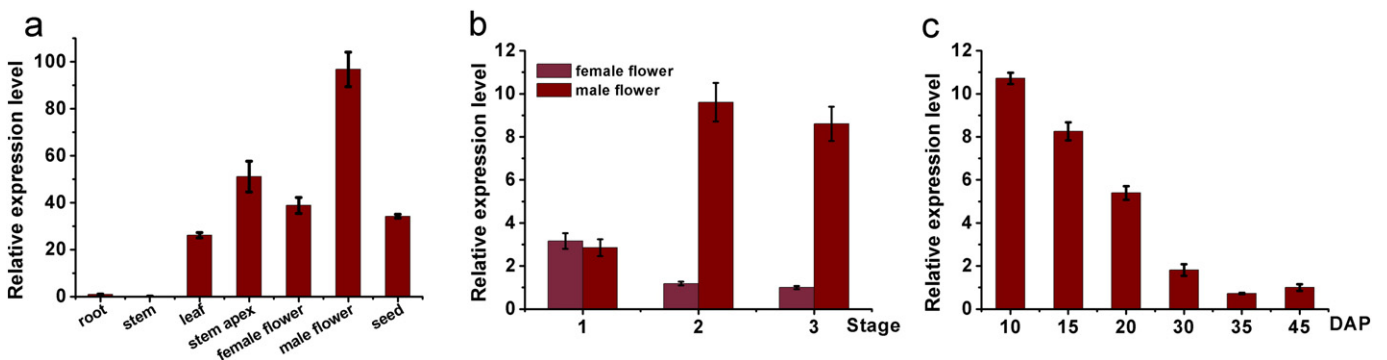


Fig. 4. Expression patterns of CYP78A98. a: Relative expression level of CYP78A98 in different tissue; b: expression profile of CYP78A98 in different development stages of male and female flower; c: relative expression level of CYP78A98 during Jatropha seed development. The error bars represent the standard deviation from three independent experiments.

Table 2
Details of CYP78A98 5'-flanking cis-elements region involved in development regulation.

Name of cis-element	Sequence	Function
TATC-box	TATCCCA	Gibberellin responsiveness
CAREs	CAACTC	gibberellin-upregulated expression
ARFAT	TGCTCT	auxin response factor binding site
CPBCSPOR	TATTAG	Cytokinin-dependent protein binding
GT1-motif	GGTTAA	Light responsiveness
ACE	AAAACGTTTA	Light responsiveness
AT1-motif	AATTATTTTTATT	Light responsiveness
POLLEN1LELAT52	AGAAA	Pollen specific activation
GTGANTG10	GTGA	Late pollen gene expression
HD-zip 2	CAAT(G/C)ATTG	Leaf morphology development
HD-zip 1	CAAT(A/T)ATTG	Differentiation of the palisade mesophyll cells

seed fatty acids are mainly composed of palmitic (C16:0), stearic (C18:0), oleic (C18:1), linoleic (C18:2), and linolenic (C18:3) [33]. As shown in (Fig. 5e), transgenic seeds accumulated more total fatty acids per seed than wild type plant seeds. By contrast, the fatty acid composition, including palmitic (C16:0), stearic (C18:0), oleic (C18:1), linoleic (C18:2), and linolenic (C18:3), was not significantly altered between wild type and transgenic plants (Table 3).

4. Discussion

Plant seed is the main food source for mankind and livestock, as well as a renewable source of energy [1]. Crop plants have undergone a long selection for larger seed size, since the beginning of agriculture [34]. Understanding how seed size is regulated at the genetic level is a key goal for plant scientists, as seed size is an important component of overall seed yield. In this study, CYP78A98, a CYP78A subfamily member from *Jatropha* was identified and characterized. This could help us to get an insight into the genetic regulation of seed size

in *Jatropha*, which may accelerate the breeding progress of this promising biodiesel plant.

Cytochrome P450s are hemethiolate enzymes involved in numerous biosynthetic and xenobiotic pathways found in all organisms from bacteria to humans [30]. All plant P450s described so far are bound to membrane, usually anchored on the cytoplasmic surface of the ER through a short hydrophobic domain of their N-terminus [35]. We identified a transmembrane domain at the N-terminus of CYP78A98 protein and CYP78A98 protein was predicted to localize on the ER. Furthermore, transient expression of CYP78A98-GFP fused protein in *Arabidopsis* protoplast confirmed CYP78A98 protein location and the function of hydrophobic domain at the N-terminus (Fig. 3). These results suggested that CYP78A98 may function in the same way as other P450 proteins. Threonine containing oxygen binding pocket (A/G)Gx(D/E)T(T/S) is usually, but not always, conserved in eukaryotic cytochrome P450s [36]. The first and last residues of this domain are not conserved in the CYP78A98 protein and other related proteins, indicating that these enzymes may not use molecular oxygen in catalysis.

The expression pattern of CYP78A98 in vegetative organ was similar to that of CYP78A5 in *Arabidopsis* [12]. CYP78A98 was highly expressed in stem apex and leaf, and this pattern was consistent with the function of CYP78A5 in leaf size and initiation control [16,37].

Expression of CYP78A5 has been detected in flower organs, including growing petal, gynoecium and the base of the flower [16]. The activity of CYP78A5 was demonstrated to coordinate flower growth within individual inflorescences [38]. We detected high expression level of CYP78A98 in both male and female flowers during their early development stages (Fig. 4b), suggesting that CYP78A98 may have a similar role as CYP78A5 in *Jatropha* flower development. Furthermore, we found two cis-elements involved in pollen specific expression in the 5'-flanking region of CYP78A98 (Table 2), and the exact expression pattern of CYP78A98 in different flower organs needs to be further examined.

In angiosperms, seed development begins with double fertilization, during which one of the two sperm cells fuses with the egg cell to form the diploid embryo and the other one fertilizes the diploid central cell to generate the triploid endosperm [39]. The maternal integuments surrounding the developing embryo and endosperm form the seed coat [40]. Seed development is influenced by the coordinated growth of the embryo, endosperm, and the integuments [1,41]. Plant final seed size is mainly attained during the rapid proliferation of the endosperm and the seed coat cells [42]. Zygotic and maternal factors affecting seed size function during this process. *Jatropha* seed usually has a development duration of 60–65 d, but seed size increases immediately after pollination and reaches its final size at about 25–30 DAP. Expression pattern of CYP78A98 was in correspondence with the rapid increase of seed size (Fig. 4c), indicating that CYP78A98 functions during the rapid increase of seed size.

The organ size regulation mechanism of CYP78A5 has been illuminated extensively; however, the genetic network of CYP78A5 is still elusive. In order to understand the genetic network of CYP78A98, we analyzed its 5'-flanking region (Table 2). Several cis-acting elements predicted to be involved in plant hormone mediated development regulation were identified, which suggested that CYP78A98 may act downstream of plant hormone to control organ size. In addition, light responsive elements were also identified. The functions of these elements need to be studied in the future.

In addition to the important role of CYP78A5 in seed size control, other genes from CYP78A subfamily have also been reported to be involved in seed size and development control [43,44,45,46,47,48,49]. To investigate the effect of CYP78A98 in seed size control, *pINO* droved CYP78A98 overexpression tobaccos were generated. The transgenic plants showed larger seeds than that of wild type plants (Fig. 6a,b), and the increase in seed size is due in part to the increased accumulation of protein and fatty acid (Fig. 6d,e). Similar results were

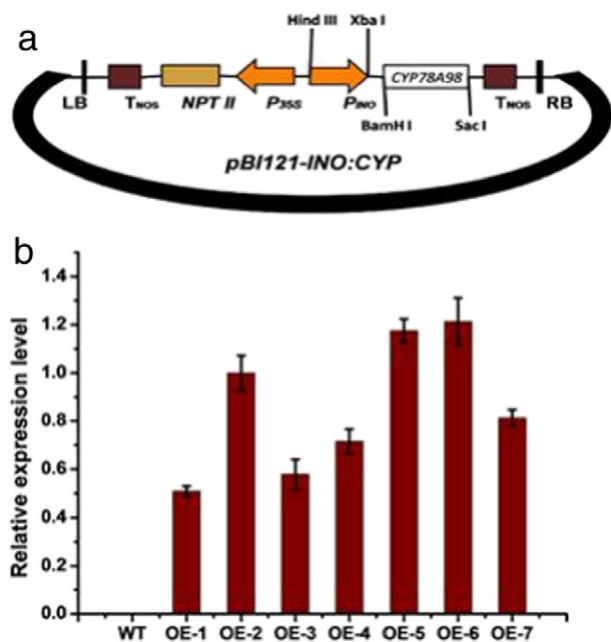


Fig. 5. Generation and transcripts detection of CYP78A98 overexpression tobacco lines. a: Map of the binary vector used to overexpress CYP78A98. P35S: 35S promoter, NPT II: neomycin phosphotransferase gene, pINO: INNER NO OUTER promoter, TNOS: NOS terminator. b: Transcription levels of CYP78A98 in wild-type (WT) tobacco plants and seven homozygous overexpression (OE) lines (named OE-1 to OE-7). The error bars represent the standard deviation from three independent experiments. Asterisks indicate statistically significant differences from the wild-type at *P < 0.05 and **P < 0.01 (Student's t test).

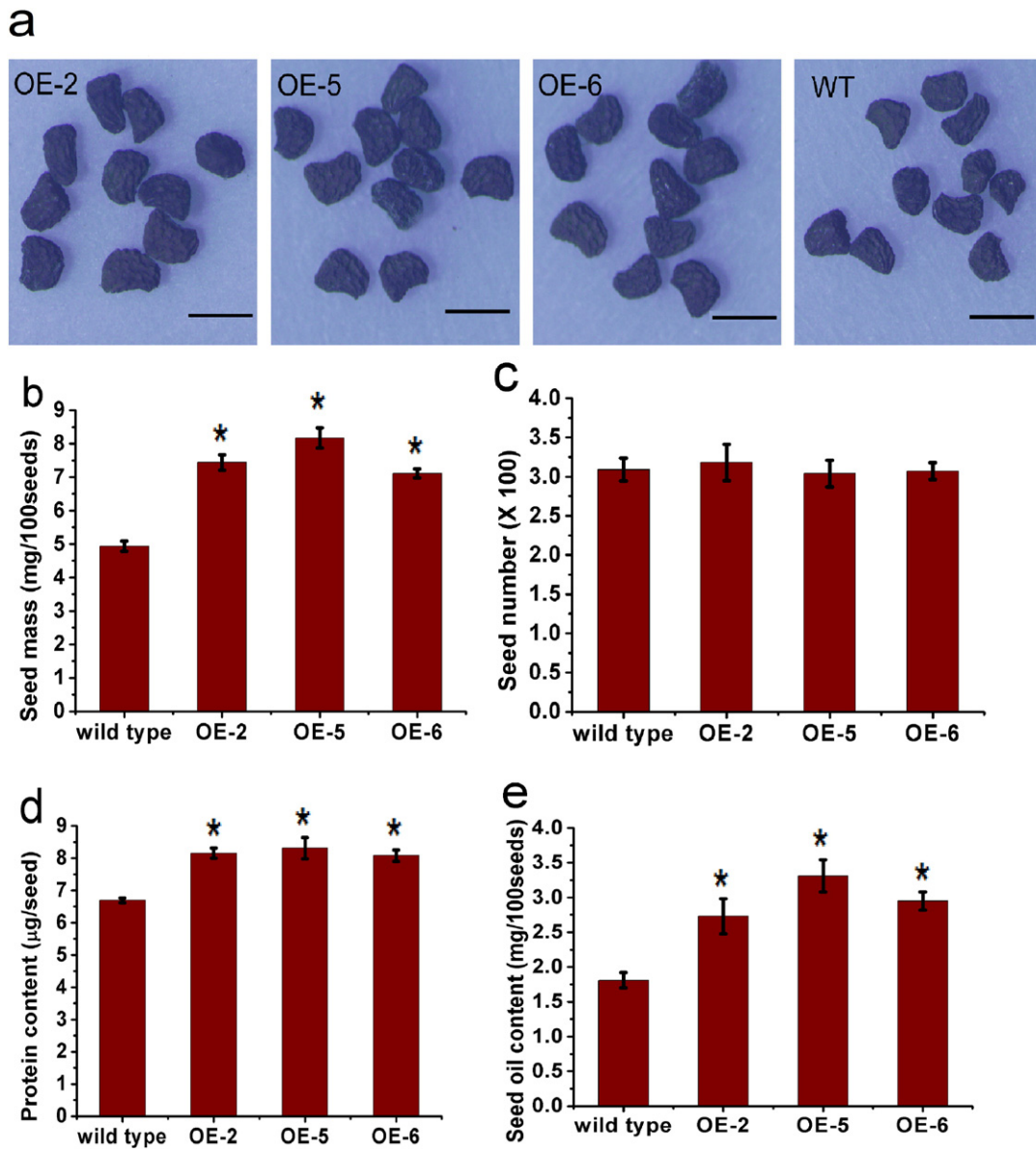


Fig. 6. Seed phenotype of wild type plants and *CYP78A98* overexpression lines. a: mature tobacco seeds; b: average seed mass per 100 seed; c: seed number per fruit; d: total protein content per seed; e: total fatty acid methyl-ester content per seed. The error bars represent the standard deviation from three independent experiments. Asterisks indicate statistically significant differences from the wild-type at * $P < 0.05$ (Student's *t* test).

observed in *Arabidopsis* when *CYP78A5* was over-expressed [11], indicating that *CYP78A* subfamily genes have a conserved function in different plants during the process of evolution.

In conclusion, *CYP78A98*, a member of *CYP78A* subfamily gene from *Jatropha*, was isolated and characterized. *CYP78A98* was found to be highly expressed in developing male and female flowers, implying that the *CYP78A98* may function in the development of *Jatropha*

flower. Furthermore, elements involved in plant hormone regulation were identified in the 5'-flanking region of *CYP78A98*, suggesting that *CYP78A98* may act downstream of plant hormone pathway to regulate plant development. Over-expression of *CYP78A98* in tobacco resulted larger seeds, indicating that *CYP78A98* had a role in seed size control. In addition, the transgenic lines accumulated more protein and fatty acid. In the future, more efforts may be required to elucidate the mechanism of *CYP78A98* in development control.

Table 3

Fatty acid profiles of tobacco seeds from overexpression lines and wild type seeds.

	Wild type	OE-2	OE-5	OE-6
16:0	16.65 ± 0.52	17.86 ± 0.60	16.61 ± 1.04	17.75 ± 0.59
18:0	2.66 ± 0.14	2.57 ± 0.08	2.57 ± 0.18	2.84 ± 0.15
18:1	14.24 ± 0.64	13.83 ± 0.11	14.66 ± 0.53	13.85 ± 0.48
18:2	65.35 ± 0.30	64.25 ± 0.79	65.00 ± 0.70	64.43 ± 0.68
18:3	1.55 ± 0.05	1.54 ± 0.05	1.61 ± 0.05	1.68 ± 0.06

Numbers in each column refer to the relative molar ratios of the different fatty acids with the total being 100%. Means and standard deviation of three independent samples are presented.

Conflict of interest

The authors declare no conflict of interest.

Financial support

This research is supported by National Key Technology R&D Program of 12th Five-Year Plan of China (No.2011BAD22B08).

References

- [1] Li N, Li Y. Maternal control of seed size in plants. *J Exp Bot* 2015;66:1087–97. <http://dx.doi.org/10.1093/jxb/eru549>.
- [2] Xu W, Dai M, Li F, Liu A. Genomic imprinting, methylation and parent-of-origin effects in reciprocal hybrid endosperm of castor bean. *Nucleic Acids Res* 2014;42:6987–98. <http://dx.doi.org/10.1093/nar/gku375>.
- [3] Luo M, Dennis ES, Berger F, Peacock WJ, Chaudhury A. *MINISEED3* (*MINI3*), a *WRKY* family gene, and *HAIKU2* (*IKU2*), a leucine-rich repeat (*LRR*) *KINASE* gene, are regulators of seed size in *Arabidopsis*. *Proc Natl Acad Sci U S A* 2005;102:17531–6. <http://dx.doi.org/10.1073/pnas.0508418102>.
- [4] Zhou Y, Zhang X, Kang X, Zhao X, Zhang X, Ni M. Short hypocotyl under blue1 associates with *Miniseed3* and *Haiku2* promoters *in vivo* to regulate *Arabidopsis* seed development. *Plant Cell* 2009;21:106–17. <http://dx.doi.org/10.1105/tpc.108.064972>.
- [5] Johnson CS, Kolevski B, Smyth DR. Transparent testa glabra 2, a trichome and seed coat development gene of *Arabidopsis*, encodes a *WRKY* transcription factor. *Plant Cell* 2002;14:1359–75. <http://dx.doi.org/10.1105/tpc.001404>.
- [6] Jofuku KD, Omidyar PK, Gee Z, Okamoto JK. Control of seed mass and seed yield by the floral homeotic gene *APETALA2*. *Proc Natl Acad Sci U S A* 2005;102:3117–22. <http://dx.doi.org/10.1073/pnas.0409893102>.
- [7] Ohto MA, Fischer RL, Goldberg RB, Nakamura K, Harada JJ. Control of seed mass by *APETALA2*. *Proc Natl Acad Sci U S A* 2005;102:3123–8. <http://dx.doi.org/10.1073/pnas.0409858102>.
- [8] Schruff MC, Spielman M, Tiwari S, Adams S, Fenby N, Scott RJ. The auxin response factor 2 gene of *Arabidopsis* links auxin signalling, cell division, and the size of seeds and other organs. *Development* 2006;133:251–61. <http://dx.doi.org/10.1242/dev.02194>.
- [9] Li Y, Zheng L, Corke F, Smith C, Bevan MW. Control of final seed and organ size by the *DA1* gene family in *Arabidopsis thaliana*. *Genes Dev* 2008;22:1331–6. <http://dx.doi.org/10.1101/gad.463608>.
- [10] Chakrabarti M, Zhang N, Sauvage C, Munos S, Blanca J, Canizares J, et al. A cytochrome P450 regulates a domestication trait in cultivated tomato. *Proc Natl Acad Sci U S A* 2013;110:17125–30. <http://dx.doi.org/10.1073/pnas.1307313110>.
- [11] Adamski NM, Anastasiou E, Eriksson S, O'Neill CM, Lenhard M. Local maternal control of seed size by *KLUH/CYP78A5* dependent growth signaling. *Proc Natl Acad Sci U S A* 2009;106:20115–20. <http://dx.doi.org/10.1073/pnas.0907024106>.
- [12] Zondlo SC, Irish VF. *CYP78A5* encodes a cytochrome P450 that marks the shoot apical meristem boundary in *Arabidopsis*. *Plant J* 1999;19:259–68. <http://dx.doi.org/10.1046/j.1365-3113.1999.00523.x>.
- [13] Zhang Y, Du L, Xu R, Cui R, Hao J, Sun C, et al. Transcription factors *SOD7/NGAL2* and *DPA4/NGAL3* act redundantly to regulate seed size by directly repressing *KLU* expression in *Arabidopsis thaliana*. *Plant Cell* 2015;27:620–32. <http://dx.doi.org/10.1105/tpc.114.135368>.
- [14] Stransfeld L, Eriksson S, Adamski NM, Breuninger H, Lenhard M. *KLUH/CYP78A5* promotes organ growth without affecting the size of the early primordium. *Plant Signal Behav* 2010;5:982–4. <http://dx.doi.org/10.4161/psb.5.8.12221>.
- [15] Wang JW, Schwab R, Czech B, Mica E, Weigel D. Dual effects of miR156-targeted *SPL* genes and *CYP78A5/KLUH* on plastochron length and organ size in *Arabidopsis thaliana*. *Plant Cell* 2008;20:1231–43. <http://dx.doi.org/10.1105/tpc.108.058180>.
- [16] Anastasiou E, Kenz S, Gerstung M, MacLean D, Timmer J, Fleck C, et al. Control of plant organ size by *KLUH/CYP78A5*-dependent intercellular signaling. *Dev Cell* 2007;13:843–56. <http://dx.doi.org/10.1016/j.devcel.2007.10.001>.
- [17] Rolland-Lagan AG. Organ size: The role of mobile signals. *Curr Biol* 2010;20:268–9. <http://dx.doi.org/10.1016/j.cub.2010.02.017>.
- [18] Koh MY, Mohd. Ghazi TI. A review of biodiesel production from *Jatropha curcas* L. oil. *Renew Sustain Energy Rev* 2011;15:2240–51. <http://dx.doi.org/10.1016/j.rser.2011.02.013>.
- [19] Berchmans HJ, Hirata S. Biodiesel production from crude *Jatropha curcas* L. seed oil with a high content of free fatty acids. *Bioresour Technol* 2008;99:1716–21. <http://dx.doi.org/10.1016/j.biortech.2007.03.051>.
- [20] Kumar N, Reddy MP, Sujatha M. Genetic transformation of *Jatropha curcas*: Current status and future prospects. *Jatropha*, Challenges for a New Energy Crop; 2013 535–46. http://dx.doi.org/10.1007/978-1-4614-4915-7_28.
- [21] Achten WMJ, Verchot L, Franken YJ, Mathijs E, Singh VP, Aerts R, et al. *Jatropha* bio-diesel production and use. *Biomass Bioenergy* 2008;32:1063–84. <http://dx.doi.org/10.1016/j.biombioe.2008.03.003>.
- [22] Openshaw K. A review of *Jatropha curcas*: An oil plant of unfulfilled promise. *Biomass Bioenergy* 2000;19:1–15. [http://dx.doi.org/10.1016/S0961-9534\(00\)00019-2](http://dx.doi.org/10.1016/S0961-9534(00)00019-2).
- [23] Achten WM, Nielsen LR, Aerts R, Lengkeek AG, Kjaer ED, Trabucco A, et al. Towards domestication of *Jatropha curcas*. *Biofuels* 2010;1:91–107.
- [24] Wei L, Yin L, Hu X, Xu Y, Chen F. A novel chloroplastic isopentenyl diphosphate isomerase gene from *Jatropha curcas*: Cloning, characterization and subcellular localization. *Electron J Biotechnol* 2014;17:296–303. <http://dx.doi.org/10.1016/j.ejbt.2014.09.003>.
- [25] Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 2011;28:2731–9. <http://dx.doi.org/10.1093/molbev/msr121>.
- [26] Yoo SD, Cho YH, Sheen J. *Arabidopsis* mesophyll protoplasts: A versatile cell system for gene expression analysis. *Nat Protoc* 2007;2:1565–72. <http://dx.doi.org/10.1038/nprot.2007.199>.
- [27] Horsch R, Fry J, Hoffmann N, Eichholtz D, Sa Rogers, Fraley R. A simple and general method for transferring genes into plants. *Science* 1985;227:1229–31. <http://dx.doi.org/10.1126/science.227.4691.1229>.
- [28] Heath JD, Weldon R, Monnot C, Meinke DW. Analysis of storage proteins in normal and aborted seeds from embryo-lethal mutants of *Arabidopsis thaliana*. *Planta* 1986;169:304–12. <http://dx.doi.org/10.1007/BF00392124>.
- [29] Li Y, Beisson F, Pollard M, Ohlrogge J. Oil content of *Arabidopsis* seeds: The influence of seed anatomy, light and plant-to-plant variation. *Phytochemistry* 2006;67:904–15. <http://dx.doi.org/10.1016/j.phytochem.2006.02.015>.
- [30] Nelson DR. The cytochrome p450 homepage. *Hum Genomics* 2009;4:59–65.
- [31] Nelson DR. Cytochrome P450 nomenclature. *Methods Mol Biol* 1998;107:15–24. <http://dx.doi.org/10.1385/0-89603-519-0-15>.
- [32] Haseloff J, Siemering KR, Prasher DC, Hodge S. Removal of a cryptic intron and subcellular localization of green fluorescent protein are required to mark transgenic *Arabidopsis* plants brightly. *Proc Natl Acad Sci U S A* 1997;94:2122–7. <http://dx.doi.org/10.1073/pnas.94.6.2122>.
- [33] Usta N, Aydoğar B, Çon AH, Uğuzdoğan E, Özkal SG. Properties and quality verification of biodiesel produced from tobacco seed oil. *Energy Convers Manag* 2011;52:2031–9. <http://dx.doi.org/10.1016/j.enconman.2010.12.021>.
- [34] Fan C, Yu S, Wang C, Xing Y. A causal C–A mutation in the second exon of *GS3* highly associated with rice grain length and validated as a functional marker. *Theor Appl Genet* 2009;118:465–72. <http://dx.doi.org/10.1007/s00122-008-0913-1>.
- [35] Williams PA, Cosme J, Sridhar V, Johnson EF, McRee DE. Mammalian microsomal cytochrome P450 monooxygenase: Structural adaptations for membrane binding and functional diversity. *Mol Cell* 2000;5:121–31. [http://dx.doi.org/10.1016/S1097-2765\(00\)80408-6](http://dx.doi.org/10.1016/S1097-2765(00)80408-6).
- [36] Durst F, Nelson DR. Diversity and evolution of plant P450 and P450-reductases. *Drug Metabol Drug Interact* 1995;12:189–206. <http://dx.doi.org/10.1515/DMDI.1995.12.3-4.189>.
- [37] Kazama T, Ichihashi Y, Murata S, Tsukaya H. The mechanism of cell cycle arrest front progression explained by a *KLUH/CYP78A5*-dependent mobile growth factor in developing leaves of *Arabidopsis thaliana*. *Plant Cell Physiol* 2010;51:1046–54. <http://dx.doi.org/10.1093/pcp/pcq051>.
- [38] Eriksson S, Stransfeld L, Adamski NM, Breuninger H, Lenhard M. *KLUH/CYP78A5*-dependent growth signaling coordinates floral organ growth in *Arabidopsis*. *Curr Biol* 2010;20:527–32. <http://dx.doi.org/10.1016/j.cub.2010.01.039>.
- [39] Chaudhury AM, Koltunow A, Payne T, Luo M, Tucker MR, Dennis E, et al. Control of early seed development. *Annu Rev Cell Dev Biol* 2001;17:677–99. <http://dx.doi.org/10.1146/annurev.cellbio.17.1.677>.
- [40] Sundaresan V. Control of seed size in plants. *Proc Natl Acad Sci U S A* 2005;102:17887–8. <http://dx.doi.org/10.1073/pnas.0509021102>.
- [41] Berger F, Grini PE, Schnittger A. Endosperm: An integrator of seed growth and development. *Curr Opin Plant Biol* 2006;9:664–70. <http://dx.doi.org/10.1016/j.pbi.2006.09.015>.
- [42] Boissard-Lorig C, Colon-Carmona A, Bauch M, Hodge S, Doerner P, Bancharel E, et al. Dynamic analyses of the expression of the *HISTONE: YFP* fusion protein in *Arabidopsis* show that syncytial endosperm is divided in different domains. *Plant Cell* 2001;13:495–509. <http://dx.doi.org/10.1105/tpc.13.3.495>.
- [43] Fang W, Wang Z, Cui R, Li J, Li Y. Maternal control of seed size by *EOD3/CYP78A6* in *Arabidopsis thaliana*. *Plant J* 2012;70:929–39. <http://dx.doi.org/10.1111/j.1365-3113.2012.04907.x>.
- [44] Sotelo-Silveira M, Cucinotta M, Chauvin AL, Chavez Montes RA, Colombo L, Marsch-Martinez N, et al. Cytochrome P450 *CYP78A9* is involved in *Arabidopsis* reproductive development. *Plant Physiol* 2013;162:779–99. <http://dx.doi.org/10.1104/pp.113.218214>.
- [45] Sotelo-Silveira M, Cucinotta M, Colombo L, Marsch-Martinez N, de Folter S. Toward understanding the role of *CYP78A9* during *Arabidopsis* reproduction. *Plant Signal Behav* 2013;8:e25160. <http://dx.doi.org/10.4161/psb.25160>.
- [46] Xu F, Fang J, Ou S, Gao S, Zhang F, Du L, et al. Variations in *CYP78A13* coding region influence grain size and yield in rice. *Plant Cell Environ* 2015;38:800–11. <http://dx.doi.org/10.1111/pce.1245>.
- [47] Yang W, Gao M, Yin X, Liu J, Xu Y, Zeng L, et al. Control of rice embryo development, shoot apical meristem maintenance, and grain yield by a novel cytochrome p450. *Mol Plant* 2013;6:1945–60. <http://dx.doi.org/10.1093/mp/sst107>.
- [48] Nagasawa N, Hibara K, Heppard EP, Vander Velden KA, Luck S, Beatty M, et al. *GIANT EMBRYO* encodes *CYP78A13*, required for proper size balance between embryo and endosperm in rice. *Plant J* 2013;75:592–605. <http://dx.doi.org/10.1111/tpj.12223>.
- [49] Ma M, Wang Q, Li Z, Cheng H, Li Z, Liu X, et al. Expression of *TaCYP78A3*, a gene encoding cytochrome P450 *CYP78A3* protein in wheat (*Triticum aestivum* L.), affects seed size. *Plant J* 2015;83:312–25. <http://dx.doi.org/10.1111/tpj.12896>.