

RESEARCH ARTICLE

Development of an efficient chromatin immunoprecipitation method to investigate protein-DNA interaction in oleaginous castor bean seeds

Mohammad Enamul Haque^{1,2}, Bing Han^{1,2}, Bin Wang^{1,3}, Yue Wang¹, Aizhong Liu^{1,3*}

1 Key Laboratory of Economic Plants and Biotechnology, Yunnan Key Laboratory for Wild Plant Resources, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, China, **2** University of Chinese Academy of Sciences, Beijing, China, **3** Key Laboratory for Forest Resources Conservation and Utilization in the Southwest Mountains of China, Ministry of Education, Southwest Forestry University, Kunming, China

* liuaizhong@mail.kib.ac.cn



OPEN ACCESS

Citation: Haque ME, Han B, Wang B, Wang Y, Liu A (2018) Development of an efficient chromatin immunoprecipitation method to investigate protein-DNA interaction in oleaginous castor bean seeds. PLoS ONE 13(5): e0197126. <https://doi.org/10.1371/journal.pone.0197126>

Editor: Sue Cotterill, Saint George's University, UNITED KINGDOM

Received: January 14, 2018

Accepted: April 26, 2018

Published: May 8, 2018

Copyright: © 2018 Haque et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work was financially supported by Chinese National Key Technology R & D Program (2015BAD15B02), National Natural Science Foundation of China (31661143002, 31701465 and 31771839) and Yunnan Applied Basic Research Projects (2016FA011).

Competing interests: The authors have declared that no competing interests exist.

Abstract

Chromatin immunoprecipitation (ChIP) is usually a reliable technique to find the binding sites of a transcription factor. In the current study, we developed a suitable ChIP method using developing castor bean seeds. A castor bean seed with large and persistent endosperm contains high amounts of storage lipids (ca. 50–60%) and is often considered as a model material to studying seed biology. In oleaginous seeds, due to the rich oils which could seriously affect immunoprecipitation and DNA isolation, it is often difficult to carry out a successful ChIP experiment. Thus, the development of an efficient ChIP method for oleaginous seeds is required. In this study, we modified different steps, including tissue preparation for cross-linking, chromatin washing, sonication and immunoprecipitation of other existing methods. As exemplified by the targeted gene identification of a master regulator WRI1, which regulates fatty acid biosynthesis, we found that the improved ChIP method worked well. We analyzed percentage input and fold changes of the ChIPed DNA. We also made successful ChIP-seq libraries using this method. This method provides a technical support not only for use on castor bean seeds; it might be used equally to analyze protein-DNA interaction *in vivo* in other oleaginous seeds.

Introduction

Identification of binding sites of a given transcription factor with its target gene is an important step for gene regulatory network analysis. Protein-DNA interaction mediates diverse biological processes such as transcription, replication, DNA housekeeping, recombination and gene silencing. Many experimental methods such as chromatin immunoprecipitation (ChIP) [1–3], Bacterial one-hybrid assay [4] and yeast one-hybrid assay [5] have been broadly used to determine the protein-DNA interactions. Particularly, ChIP combined with deep sequencing (ChIP-seq) is an efficient and reliable method to find target genes and location of the DNA

Abbreviations: ChIP, chromatin immunoprecipitation; RT, room temperature; TAGs, triacylglycerols; WRI1, Wrinkled1; SDS, sodium dodecyl sulfate; EDTA, ethylenediaminetetraacetic acid; NP-40, Nonidet P-40; PMSF, phenylmethylsulfonyl fluoride; IP, immunoprecipitation; TE, tris-ethylenediaminetetraacetic acid; EP tube, Eppendorf tube; PAGE, polyacrylamide gel electrophoresis; ChIP-qPCR, chromatin immunoprecipitation-quantitative PCR; Rc, *Ricinus communis*; BCCP2, biotin carboxyl carrier protein 2; EAR1, enoyl-acyl carrier protein reductase1; DGAT1, diglyceride acyltransferase 1.

binding site for a specific transcription factor or protein associated with chromatin, and has become a critical strategy to dissect the Protein-DNA interaction *in vivo* using native cells or tissues. ChIP-chip is another technique generally used to detect protein-DNA association *in vivo* [2, 6, 7]. Due to several limitations of ChIP-chip (e.g., lower resolution, and dependant on hybridization and high coverage affinity for non-repetitive regions), ChIP-seq is becoming a more powerful technique with advantages such as nucleotide specific resolution (which is the greatest advantage over ChIP-chip), since due to the absence of the hybridization step ChIP-seq is free from noise, showing a distinct peak which is biologically meaningful; moreover, it has no limitation on genome coverage [8]. At present, although many protocols have been developed to carry out the ChIP experiment in different plant species, most of these protocols are designed for soft tissues like leaves [9–11], stems and seedlings [9, 12, 13], or for some special tissues such as Arabidopsis mesophyll protoplasts [14], wood-forming stem-differentiating xylem (SDX) tissues [15], and unicellular green algae [16]. However, little research has been done on the ChIP experiment using seeds as starting materials [17, 18]. Seeds are highly specific and special organs or tissues. It is important to dissect the molecular mechanisms underlying regulation and storage reservoir accumulation during seed development. Usually, it is difficult to carry out a successful ChIP experiment in seeds due to the high amount of storage compounds such as polysaccharides, lipids and some secondary metabolites. These storage compounds could seriously affect different steps of the experiment, especially immunoprecipitation, resulting in low amount of DNA after isolation. For oilseed crops, rich lipids within tissues also cause a difficulty in isolating quality genomic DNA and RNA from seeds [19, 20]. In particular, as the demand for plant oil increases for both food and industrial applications, it is necessary to develop and improve oilseed crops through the enhancement of oil content and improvement of fatty acid composition for human health or for industrial applications [21]. Thus, it is an immediate need to develop an efficient and reliable ChIP method to dissect the molecular regulation process of seed development (in particular for oilseeds).

The castor bean (*Ricinus communis* L.) is one of the most important non-edible oilseed crops and its seed oil is widely used in industry, such as for the production of lubricating oil and biodiesel [22, 23]. Its seed contains more than 60% storage lipids (in the form of triacylglycerols) and is rich in ricinoleic acids. Unlike Arabidopsis and other dicotyledonous oilseed crops such as rapeseed, sunflower, groundnut and soybean (their oils are accumulated in cotyledons and their endosperms are ephemeral in the early stage of seed development), castor bean seed oils and other storage reservoirs (such as storage proteins and polysaccharides) are accumulated in the endosperm and its endosperm is persistent through seed development. Thus, the castor bean is often considered as a model dicotyledonous plant for studying the molecular regulation of seed development and storage material biosynthesis in the field of seed biology [24, 25]. Due to the lack of an efficient and suitable ChIP method, studies on dissecting the molecular mechanism of seed development and storage material biosynthesis have been largely restricted for the castor bean.

Previously, we tried to use different ChIP methods [2, 10, 12] in the experiment of castor bean seed ChIP-seq—commonly used on leaves, shoots and other soft tissues and could not obtain enough targeted ChIP-ed DNA from developing seeds for further analysis due to the rich presence of lipids and polysaccharides. Here, we modified the common ChIP method by optimizing several experimental steps, and have developed a suitable and efficient method to dissect the molecular mechanism of seed development and oil biosynthesis in castor bean seeds. We have resolved some solution formulations and some difficulties such as sonication, washing of the nuclear pellet, immunoprecipitation, washing of beads-immune complexes, etc. As the master regulator WRI1 of oil biosynthesis [26] exemplified, we have demonstrated a successful and detailed protocol to conduct the ChIP experiment using developing castor

bean seeds. We trust this method to not only provide an optimized experimental procedure for conducting ChIP in castor bean seeds, but also be suitable for use in investigating the protein-DNA interaction in other oleaginous seed crops.

Materials and methods

Plant materials

In this study, we used castor bean variety ZB107 as plant materials (kindly provided by the Zibo Academy of Agriculture Sciences, Shandong, China). Seeds were cultivated in the experimental field of the Kunming Institute of Botany, Chinese Academy of Sciences, during the months of April to October, 2016, when we maintained the plant's growth under natural conditions. During flowering, we performed hand pollination to prevent cross pollination and to maintain proper records from the day after pollination (DAP). Endosperms were isolated from the seed samples collected from 7, 15, 25, 35 and 45 DAP respectively, and snap-frozen immediately to extract RNA.

Total RNA purification and complementary DNA synthesis

Total RNA was purified from the collected endosperm samples using TRIzol (Invitrogen, Carlsbad, CA), followed by RNeasy Mini Kit (Qiagen, Valencia, CA), according to the manufacturer's instructions. Then we checked the concentration and quality of total RNA by NANODROP-2000 spectrophotometer (Thermo Scientific) and by 1.5% agarose gel as well. Complementary DNA for different stages of seed samples were synthesized from 1 μ g of total RNA by transcript one-step gDNA removal and cDNA synthesis supermix (TIANGEN, Beijing, China), according to the manufacturer's instructions, and then the cDNA samples were stored at -20°C until use.

Lipid analysis

Total lipid was extracted according to Xu et al. and Zhang et al. [27, 28]. At first, the seed samples were collected from five developmental stages and dried in a vacuum incubator. The dried samples for three replications were measured accurately. The seed tissues were ground with a mortar and pestle to fine powder using liquid nitrogen, transferred into 10 ml glass tubes, and then added 2 ml of hexane:isopropanol (3:2, v/v) and mixed well. To collect the supernatant, the mixtures were centrifuged at 5000 g for 5 min. After that, total lipids were extracted by drying the supernatants in a vacuum incubator at 50°C to remove the dissolved hexane phase, and determining gravimetrically.

Primer design

For primer design, we obtained the CDS sequences (for RcWRI1 and RcEF1 α) and promoter sequences (for RcBCCP2, RcEAR1 and RcDGAT1) from Phytozome v12.1 through query of the specific gene's protein sequence from Arabidopsis. Prime Primer 5.0 (Premier, Palo Alto, CA, USA) was used to design primers with amplicon sizes between 100–180 bp. For ChIP qPCR, we designed primers from the promoter sequences of RcBCCP2 and RcEAR1 containing AW box. All the primers are listed in S1 Table.

Quantitative real-time PCR analysis

Quantitative real-time PCR was performed by the CFX96 system (Bio-Rad, Hercules, CA, USA) to analyze tissue specific expression of the WRI1 transcription factor and ChIP-ed DNA enrichment. For ChIP qPCR, the ChIP-ed DNA, 10% input DNA and negative control DNA

(no Ab sample) were diluted three times with deionized water. PCR amplifications were carried out in 20 μ L reaction mixture containing the following: 0.5 μ L of DNA template, 8.7 μ L of deionized water, 10 μ L of TransStart Tip Green qPCR SuperMix (TransGen Biotech, Beijing, China), and 0.8 μ L of forward and reverse primers. The two-step PCR program consisted of initiation at 95°C for 30 s, then denaturation at 95°C for 15 s, annealing at 60°C for 30 s, and extension at 95°C for 10 s with 40 cycles. The melting curve was obtained from 65°C to 95°C, at an increment of 0.5°C every 5 s. The *RcEF1 α* was used as an internal control for gene expression normalization of *RcWRI1* and the reaction was performed three times. In the case of ChIP qPCR, the enrichments were analyzed in two biological and three technical replications.

Western blotting

For western blotting, proteins were purified from 0.5 g of young leaves and 1 g of seeds of 35 DAP castor bean according to Fido et al. [29] and quantified by the Bradford method [30]. Proteins were also purified from decrosslinked samples of IP and mock control. In this case, a 20 μ L solution from each of the samples was taken and precipitated by 3 volumes of acetone at -20°C for 30 min. The tubes were centrifuged at 12000 rpm for 5 min to pellet the proteins, and then the pellets were washed with 70% ethanol, dried briefly and dissolved with 20 μ L of 2x Laemmli Buffer. Proteins were also purified from another set of IP and mock control samples. Before being run on SDS-PAGE, 2x Laemmli Buffer was also added to 40 μ g of crude protein extracts from the leaf and seed samples, and the protein samples were denatured at 95°C for 10 min and run on SDS-PAGE. After that the proteins were transferred on PVDF membrane using Trans-Blot[®] SD Semi-Dry Transfer Cell (BIO-RAD). The membrane was blocked in 5% BSA in TBST at room temperature for 1 hour, and then added primary antibody and incubated at room temperature for 3 hours with gentle shaking. The membrane was rinsed 3 times for ten minutes with TBST. Then it was incubated with the secondary antibody for 2 hours at room temperature and washed with TBST for three times in the same way as previous washing steps. Finally the image was developed by ECL Prime Western Blotting Detection Reagent (GE Healthcare Life Sciences) using MicroChemi 4.2 (DNR Bio-imaging Systems, ISRAEL) according to the manufacturer's instructions. In this study, we used rabbit anti-WRI1 antibody (Willget Biotech Co., China) as the primary antibody and goat anti-rabbit IgG/HRP (Biosharp, China) or Anti-Rabbit IgG HRP (Rabbit TrueBlot) (Rockland Immunochemicals Inc.) as the secondary antibody.

ChIP experiment

The overall procedure of our ChIP experiment is diagrammatically shown in Fig 1. Based on the existing ChIP experiment methods [2, 10, 12] which are usually used in soft tissues such as leaves and seedlings, we modified the solution formulations and some critical steps. First, crosslinking is a critical step that largely affects the final acquisition of targeted DNA fragments [3, 10]. The seed tissues were cut very thin to increase the penetration of crosslinking agents to the tissues. Before crosslinking, we washed the dissection tissues quickly with fixation buffer without formaldehyde to remove most lipids and other debris. Second, to further clean the lipids and other debris the nuclear pellets were washed several times using extraction buffer 2 and 3 during chromatin extraction. Third, after sonication, we added centrifugation of the chromatin solution several times to further remove debris. Fourth, during the washing of beads we extended the time and number of washes of the beads with washing buffers. For preparing fixation buffers we applied 1x PBS (pH 7.4) with 0.1 M sucrose and 1x Protease inhibitor, as PBS and sucrose induce crosslinking in the hard lipid tissues. For extraction buffer 1, 2 and 3, compared with the previous buffer systems described by Kaufmann et al. [2] we added 1

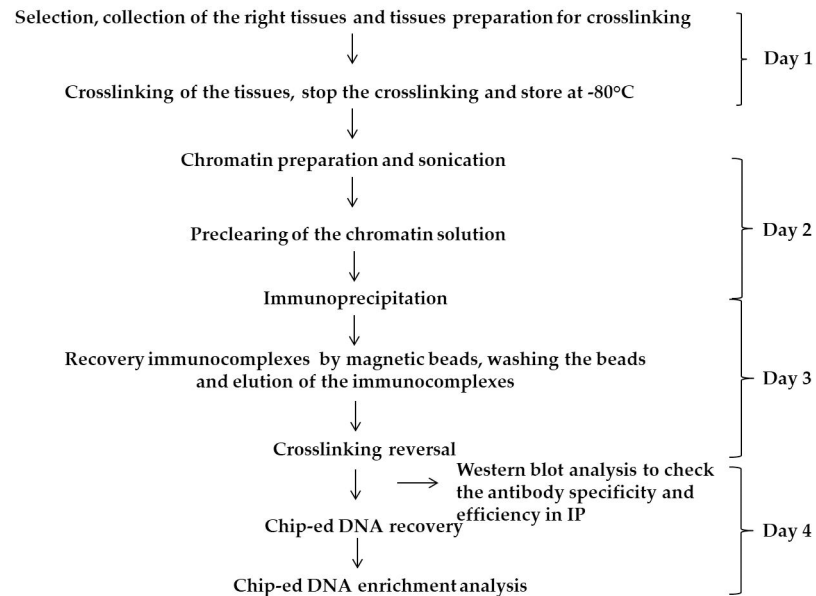


Fig 1. Diagrammatic representation of the modified ChIP experiment that was used in the castor bean oilseed.

<https://doi.org/10.1371/journal.pone.0197126.g001>

mM PMSF and 1 mM EDTA, and increased the concentration of 2-methyl 2, 4-pentanediol up to 1.5 M to easily remove lipids from the nuclear extract and precipitate proteins completely during several centrifugations. In addition, we added 0.1% Triton X-100 in buffer 3 to help extensively wash lipids off the nuclear pellet. For the magnetic beads washing buffers and elution buffers we adopted the buffer combination described by Ricardi et al. [10] and Bowler et al. [12]. The detailed procedure of the ChIP experiment is given in S1 File.

ChIP qPCR analysis and performed ChIP-seq

ChIP qPCR data were analyzed by percentage input [31] and fold enrichment using the $2^{(-\Delta\Delta Ct)}$ method [32]. Input samples were used as positive controls and no antibody-treated samples were used as negative controls. *RcDGAT1* was used as a negative control locus. After checking the enrichment, the DNA samples were sequenced by Illumina HiSeq 2000 on the BGI platform (BGI, Shenzhen, China).

Results and discussion

Selection of tissues and preparation for crosslink

As most of the transcription factors are usually tissue specific in their expression, it is inevitable that a right tissue is selected where the desired transcription factor is abundant for its regulation. It is also a key step for the success of the ChIP experiment, or the recovery quality and quantity of the targeted DNA over the control. Here, to find out the right seed developmental stage for master transcription factor WRI1 [26, 33], which regulates many fatty acid biosynthetic genes for lipid accumulation during seed development, we applied a real-time qPCR technique to investigate the expression of WRI1 at different stages of seed development, including 7, 15, 25, 35 and 45 DAPs. As shown in Fig 2A, WRI1 is expressed highly at 35 DAP, implying that WRI1 is rapidly expressed between 25 to 35 DAP. Correspondingly, the oil content is rapidly accumulated between 25 to 35 DAP (Fig 2B). Meanwhile, we can observe visible endosperm development at 35 DAP (Fig 2C). Thus, we collected the seed samples from 35

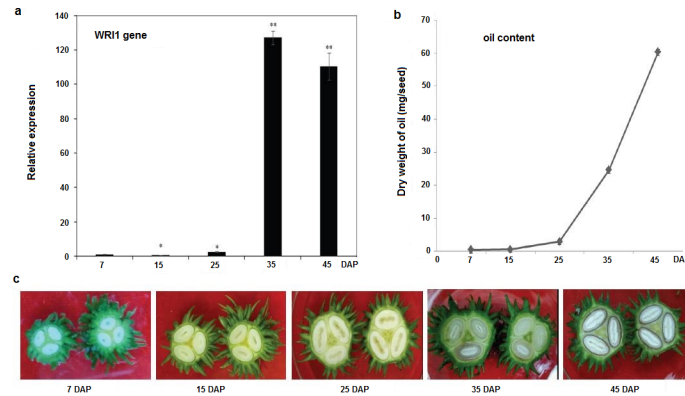


Fig 2. Relative expression of the WR1 transcription factor, oil contents and different developmental stages of the castor bean seed. **a** qPCR was performed to analyze the relative expression of the WR1 transcription factor in different seed stages. **b** and **c** Oil contents and morphology of the different developmental stages of the seeds. Tissues from different developmental stages (7, 15, 25, 35, 45 DAP) of the seeds were used to check the expression levels, analyze oil content and study seed morphology. DAP indicates day after pollination, error bars indicate the SEM. Statistical significance was carried out by pair-sample t-test. Different asterisks indicate different significant levels: one asterisk (*) for $P < 0.05$ and two asterisks (**) for $P < 0.01$. qPCR data were normalized to *RcEF1 α* , a housekeeping gene.

<https://doi.org/10.1371/journal.pone.0197126.g002>

DAP. Similarly, storage lipids accumulation is concentrated from the mid to late phase of seed development in most oilseeds [27, 34–36], thus sampling the developing seed tissues at the mid-phase of seed development might be critical for most oilseed crops, to dissect the target genes of a given transcription factor functionally involved in regulating lipids (or other storage materials) accumulation during seed development.

The crosslinking reaction is a critical step in making a successful ChIP experiment. According to our experience, for the seed tissues, it is necessary to cut the seeds into very thin pieces prior to crosslinking. In this experiment, we made homogenous tissues for generating uniform and complete crosslinking reactions, because the dissection tissues could be easily infiltrated. Although there have been several reports on the optimization of crosslinking using different concentrations of formaldehyde in different types of tissues [3, 11, 37], we here optimized the sonication conditions using a standard concentration (1%) of formaldehyde [2, 11] and obtained relatively ideal results.

Chromatin processing and immunoprecipitation

Chromatin processing, including chromatin washing and fragmentation, is always considered to be a key step in obtaining a good result for all ChIP experiments. In this method, we optimized several critical steps (of washing) to clear the supernatant. These improved steps enhance the binding efficiency of the antibody to the beads and reduce more background during IP. However, this procedure has a drawback as it causes a bit of loss in some chromatin. For a study of protein binding sites genome-wide, it is important to shear the chromatin efficiently in desired lengths. Different ChIP methods have mentioned their sonication conditions differently; in most of the cases the fragment size ranges from 200–2000 bp [15, 16, 31]. However, the sonication conditions depend on the tissue, chromatin sample volume and sample concentration. In our experiment, we tried to optimize the shearing conditions using Bioruptor® Pico (DIAGENODE, BELGIUM) with a combination of 20 sec ON/30 sec OFF for several times, and obtained maximum fragments in the range of 100–500 bp in 20 times on agarose gel (Fig 3A). We have also checked the DNA size of the IP sample on the Agilent 2100 Bioanalyzer, and found the pick range to be between 100–500 bp (see S1 Fig). Thus, we

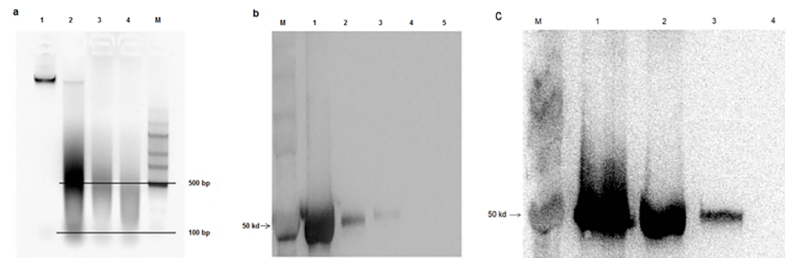


Fig 3. Sonication efficiency and antibody specificity. **a** Chromatin samples were fragmented at 20 sec ON/30 sec OFF using Bioruptor® Pico. Samples were decrosslinked, DNA was purified by phenol:chloroform:isoamyl alcohol extraction and loaded on 1.5% agarose gel. Lane 1: non-sonicated; lane 2: sonicated 7 times; lane 3: sonicated 15 times; lane 4: sonicated 20 times; lane M: 5000 bp DNA marker. Scales indicate the DNA fragments to be between 100–500 bp. **b** Western blot figure showing lane M: protein molecular weight marker; lane 1: crude protein extract (40 µg); lane 2: IP in sample 1; lane 3: IP in sample 2; lane 4: mock control in sample 1; lane 5: mock control in sample 2. Arrow indicates the protein marker band 50 KD. **c** Western blot showing the lane M protein weight marker; lane 1: crude protein extract (40 µg) from endosperms of 35 DAP castor bean seeds; lane 2: IP sample; lane 3: IP sample (20 µl from the IP sample used for lane 2); lane 4: mock control sample. Arrow indicates the protein marker band 50 KD.

<https://doi.org/10.1371/journal.pone.0197126.g003>

recommend dissolving the chromatin pellet with lysis buffer according to the primary tissue taken for the experiment and the efficiency of sonication. Lipids make the chromatin solution more viscous, which generates heat in the solution, reduces fragmentation efficiency and denatures protein during sonication. Generally smaller fragments are expected in a ChIP experiment; however, DNA fragments within 100–500 bp are able to meet the purposes of the ChIP experiment.

We have mentioned another key step of the ChIP experiment in lipid seeds, which is washing the beads following IP. We washed the beads twice with each of the washing buffers carefully (see [methods](#)), to thoroughly remove lipids and other seed-derivatives which were attached to the beads during IP. However, excessive washing could cause the loss of DNA and this has generally happened if the tissues are not crosslinked properly. After the final wash of the beads, we performed decrosslinking over-night at 65°C [37] to recover the maximum amount of DNA. We purified protein from 20 µl of decrosslinked solution of IP and mock control samples, that was run on the SDS-PAGE along with 40 µg of crude protein extract (from 35 DAP seed stage), following western blot analysis to check the efficiency of IP and antibody specificity. We got the protein bands at the position of 50 KD on the membrane ([Fig 3B](#)), lane 1 for crude protein sample, lanes 2 and 3 for IP of sample 1 and sample 2, and lanes 4 and 5 for the mock controls of sample 1 and 2, respectively. However we confused with the protein size at the position of 50 KD as the Ig heavy chain has also the similar molecular weight. We used goat anti-rabbit IgG/HRP secondary antibody which could bind to Ig heavy chain also, even the membrane was washed extensively before and after antibodies treatment which may reduce the presence of nonspecific binding. To confirm specific protein for WR11, we further checked the IP following western blot of another set of IP and mock control samples along with a crude protein extract using anti-Rabbit IgG HRP (Rabbit TrueBlot) secondary antibody (which is unable to recognize denatured immunoglobulins) and we found the protein band at the position of 50 KD for WR11 ([Fig 3C](#)), suggesting that our ChIP method work efficiently. For lane 3 in [Fig 3C](#), proteins of 20 µl IP's sample was loaded which was taken from the IP sample loaded on lane 2, as we have mentioned to perform western blot using small amount of IP sample after decrosslink. Before the ChIP experiment we also checked the antibody specificity by western blot using the crude protein extract from the endosperms of 35 DAP seeds and young leaves (see [S2 Fig](#)), and we found the antibody to be specific.

DNA enrichment analysis in ChIP samples by qPCR

Several techniques have been used to validate the ChIP-ed DNA [11]; however, we performed qPCR followed by analysis of input percentage and fold changes using ChIP-ed DNA, input DNA and mock control DNA. We designed the primers for qPCR to contain AW box, the binding site of the WRI1 transcription factor [38]. Then we analyzed the enrichment of the binding site in the promoter of *RcBCCP2* and *RcEAR1*, the two key enzymes in fatty acid biosynthesis [39–41]. We also analyzed the enrichment of binding loci in the promoter of *RcDGAT1* as a negative control, the non-target gene of the WRI1 transcription factor [38]; however, it is a key-enzyme gene responsible for TAG assembly in the Kennedy pathway [42]. We performed qPCR of two samples using three technical replications. Percentage input and fold changes were normalized to the input DNA (10%) for each of the promoter regions. In the case of percentage input, the *RcBCCP2* region showed input percentage values of 0.7 and 0.32 in sample 1 and sample 2 respectively, whereas the *RcEAR1* region showed values of 0.99 and 0.33 for sample 1 and sample 2 respectively. In the case of the *RcDGAT1* region, % input values in ChIP-ed DNA were not significant and more or less similar to the no antibody mock control for both of the samples (Fig 4A). We also analyzed the fold changes in each of the

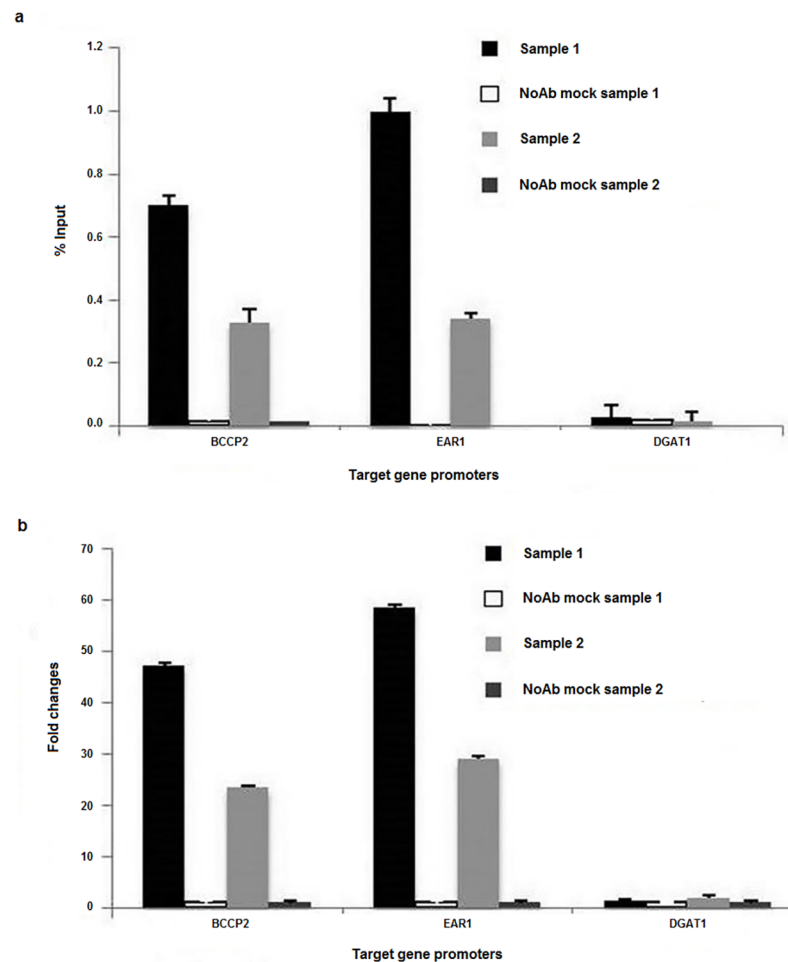


Fig 4. Percentage input and fold changes analysis. **a** percentage input and **b** fold changes of the target regions in the promoters of *RcBCCP2*, *RcEAR1* and *RcDGAT1*. *RcDGAT1* was used as a negative control locus. The ChIP results were obtained from two biological and three technical replications. Error bars indicate the SEM values.

<https://doi.org/10.1371/journal.pone.0197126.g004>

Table 1. Clean reads, mapped reads and uniquely mapped reads of sequencing data.

Sample name	Clean Reads Number	Clean Rate (%)	Mapped Reads	Mapped Rate (%)	Uniquely Mapped Reads	Uniquely Mapped Rate (%)
Sample 1	21,912,297	99.87	20,286,833	92.58	12,663,834	57.79
Sample 2	21,926,008	99.93	19,699,186	89.84	13,015,131	59.36
Input sample 1	21,774,009	99.24	20,438,626	93.87	14,211,823	65.27

<https://doi.org/10.1371/journal.pone.0197126.t001>

samples: more than 20-fold for both of the target regions. On the other hand, the non-target region (*RcDGAT1*) showed no fold changes over the control in both of the samples (Fig 4B), suggesting that our developed ChIP method was efficient and suitable for conducting protein-DNA interaction in the developing castor bean seeds.

Quality analysis of ChIP-sequencing libraries

To test whether the quality of the obtained ChIP libraries is sufficient for ChIP-seq analysis high throughput DNA sequencing was applied. After sequencing, the raw reads were mapped to the reference genome of the castor bean using SOAP2 (version 2.21t) [43], to analyze uniquely mapped reads and quality of the reads. From the analyzed data, we found the mapped rates for sample 1, sample 2 and input-sample 1 to be 92.58%, 89.84% and 93.87% respectively. Meanwhile, the uniquely mapped rates for sample 1, sample 2 and input-sample 1 were 57.79%, 59.36% and 65.27% respectively (Table 1), suggesting that the quality of the libraries was good for further ChIP-seq analysis.

Conclusions

Due to the presence of lipids and other metabolic compounds, it is difficult to perform ChIP on seed tissues using other methods mentioned for soft tissues like leaves, stems, seedlings, etc. In this study, we provide a simple and optimized chromatin immunoprecipitation method for lipid seeds to analyze protein-DNA interaction *in vivo*. We optimized some important steps such as tissue preparation for crosslinking, sonication and immunoprecipitation to minimize the problems generated by lipids and other metabolic compounds and we also optimized the solution formulations. We have shown sound results, as exemplified by the investigation of the targeted gene's enrichment of the WR11 transcription factor in lipid castor bean seeds. This method provides a technical support not only for use in castor bean seeds, but also for equal use in the global genetic analysis of other oleaginous seeds.

Supporting information

S1 File. The detailed procedure of the ChIP experiment.
(PDF)

S1 Fig. DNA fragments checked by Agilent 2100 Bioanalyzer. Figure showing maximum fragments ranges between 100–500 bp. Peaks at 35 and 10380 showing lower and upper markers.
(TIF)

S2 Fig. Western blot analysis to check anti-WR11 antibody specificity. Figure showing the lane M protein weight marker; lane 1 and 2: crude protein extract from young leaves; lane 3 and 4: crude protein extract from endosperms of 35 DAP castor bean seeds. In all cases, 40 µg of crude protein extracts were run on gel. Arrow indicates the protein marker band 50 KD.
(TIF)

S1 Table. Primer sequences used for qPCR analysis.
(PDF)

Acknowledgments

We thank Mr. Chao Sun for cultivating the castor bean and maintaining the experimental materials. Also, we extend our thanks to Dr. Wei Xu for his constructive comments.

Author Contributions

Conceptualization: Aizhong Liu.

Data curation: Bin Wang.

Funding acquisition: Aizhong Liu.

Investigation: Mohammad Enamul Haque, Bing Han.

Methodology: Yue Wang.

Supervision: Aizhong Liu.

Writing – original draft: Mohammad Enamul Haque.

References

1. Kuo MH, Allis CD. *In vivo* cross-linking and immunoprecipitation for studying dynamic protein-DNA associations in a chromatin environment. *Methods*. 1999; 19:425–433. <https://doi.org/10.1006/meth.1999.0879> PMID: 10579938
2. Kaufmann K, Muiño JM, Østerås M, Farinelli L, Krajewski P, Angenent GC. Chromatin immunoprecipitation (ChIP) of plant transcription factors followed by sequencing (ChIP-seq) or hybridization to whole genome arrays (ChIP-CHIP). *Nat. Protoc*. 2010; 5:457–472. <https://doi.org/10.1038/nprot.2009.244> PMID: 20203663
3. Das PM, Ramachandran K, vanWert J, Singal R. Chromatin immunoprecipitation assay. *Biotechniques*. 2004; 37:961–969. PMID: 15597545
4. Meng X, Brodsky MH, Wolfe SA. A bacterial one-hybrid system for determining the DNA-binding specificity of transcription factors. *Nat. Biotechnol*. 2005; 23:988–994. <https://doi.org/10.1038/nbt1120> PMID: 16041365
5. Deplancke B, Dupuy D, Vidal M, Walhout AJ. A gateway-compatible yeast one-hybrid system. *Genome Res*. 2004; 14:2093–2101. <https://doi.org/10.1101/gr.2445504> PMID: 15489331
6. Kim TH, Ren B. Genome-Wide Analysis of Protein-DNA Interactions. *Annual Rev. Genom. Hum. Genet*. 2006; 7:81–102.
7. Sandmann T, Jakobsen JS, Furlong EEM. ChIP-on-chip protocol for genome-wide analysis of transcription factor binding in *Drosophila melanogaster* embryos. *Nat. Protoc*. 2007; 1:2839–2855.
8. Park PJ. ChIP-seq: advantages and challenges of a maturing technology. *Nat. Rev. Genet*. 2009; 10:669–680. <https://doi.org/10.1038/nrg2641> PMID: 19736561
9. Saleh A, Alvarez-Venegas R, Avramova Z. An efficient chromatin immunoprecipitation (ChIP) protocol for studying histone modifications in Arabidopsis plants. *Nat. Protoc*. 2008; 3:1018–1025. <https://doi.org/10.1038/nprot.2008.66> PMID: 18536649
10. Ricardi MM, González RM, Lusem ND. Protocol: fine-tuning of a Chromatin Immunoprecipitation (ChIP) protocol in tomato. *Plant Methods*. 2010; 6:11. <https://doi.org/10.1186/1746-4811-6-11> PMID: 20380723
11. Haring M, Offermann S, Danker T, Horst I, Peterhansel C, Stam M. Chromatin immunoprecipitation: optimization, quantitative analysis and data normalization. *Plant Methods*. 2007; 3:11. <https://doi.org/10.1186/1746-4811-3-11> PMID: 17892552
12. Bowler C, Benvenuto G, Laflamme P, Molino D, Probst V, Tariq M, et al. Chromatin techniques for plant cells. *Plant J*. 2004; 39:776–789. <https://doi.org/10.1111/j.1365-313X.2004.02169.x> PMID: 15315638
13. Yamaguchi N, Winter CM, Wu M-F, Kwon CS, William DA, Wagner D. Protocols: Chromatin Immunoprecipitation from Arabidopsis Tissues. *Arabidopsis Book*. 2014, 12:e0170. <https://doi.org/10.1199/tab.0170> PMID: 24653666

14. Lee JH, Jin S, Kim SY, Kim W, Ahn JH. A fast, efficient chromatin immunoprecipitation method for studying protein-DNA binding in *Arabidopsis* mesophyll protoplasts. *Plant Methods*. 2017; 13:42. <https://doi.org/10.1186/s13007-017-0192-4> PMID: 28539971
15. Li W, Lin Y-C, Li Q, Shi R, Lin C-Y, Chen H, et al. A robust chromatin immunoprecipitation protocol for studying transcription factor–DNA interactions and histone modifications in wood-forming tissue. *Nat. Protoc*. 2014; 9:2180–2193. <https://doi.org/10.1038/nprot.2014.146> PMID: 25144269
16. Strenkert D, Schmollinger S, Schroda M. Protocol: methodology for chromatin immunoprecipitation (ChIP) in *Chlamydomonas reinhardtii*. *Plant Methods*. 2011; 7:35. <https://doi.org/10.1186/1746-4811-7-35> PMID: 22050920
17. Mönke G, Seifert M, Keilwagen J, Mohr M, Grosse I, Hähnel U, et al. Toward the identification and regulation of the *Arabidopsis thaliana* ABI3 regulon. *Nucleic Acids Res*. 2012; 40:8240–8254. <https://doi.org/10.1093/nar/gks594> PMID: 22730287
18. Yang H, Liu X, Xin M, Du J, Hu Z, Peng H, et al. Genome-Wide Mapping of Targets of Maize Histone Deacetylase HDA101 Reveals Its Function and Regulatory Mechanism during Seed Development. *Plant Cell*. 2016; 28:629–645. <https://doi.org/10.1105/tpc.15.00691> PMID: 26908760
19. Santos CD Jr, Dias ACC, Amaral IMR, Bonetti AM, Campos TA. New efficient DNA extraction method to access the microbiome of *Ricinus communis* seeds. *Genet Mol Res*. 2013; 12:3128–3135. <https://doi.org/10.4238/2013.February.28.23> PMID: 23479168
20. Lan T, Yao B, Shen Y, Wang X. Isolation of high-quality total RNA from lipid-rich seeds. *Analyt. Biochem*. 2013; 438:11–13. <https://doi.org/10.1016/j.ab.2013.03.012> PMID: 23524018
21. Schmid KM, Ohlrogge JB. Lipid metabolism in plants. In: Vance DE, Vance JE, editors. *Biochemistry of Lipids, Lipoproteins and Membranes*. Elsevier Science BV; 2008. pp. 97–130.
22. Ogunniyi DS. Castor oil: a vital industrial raw material. *Bioresour. Technol*. 2006; 97:1086–1091. <https://doi.org/10.1016/j.biortech.2005.03.028> PMID: 15919203
23. Houston NL, Hajduch M, Thelen JJ. Quantitative Proteomics of Seed Filling in Castor: Comparison with Soybean and Rapeseed Reveals Differences between Photosynthetic and Nonphotosynthetic Seed Metabolism. *Plant Physiol*. 2009; 151:857–868. <https://doi.org/10.1104/pp.109.141622> PMID: 19675154
24. Xu W, Yang T, Dong X, Li D-Z, Liu A. Genomic DNA Methylation Analyses Reveal the Distinct Profiles in Castor Bean Seeds with Persistent Endosperms. *Plant Physiol*. 2016; 171:1242–1258. <https://doi.org/10.1104/pp.16.00056> PMID: 27208275
25. Hajduch M, Matusova R, Houston NL, Thelen JJ. Comparative proteomics of seed maturation in oil-seeds reveals differences in intermediary metabolism. *Proteomics*. 2011; 11:1619–1629. <https://doi.org/10.1002/pmic.201000644> PMID: 21413150
26. Focks N, Benning C. wrinkled1: a novel, low-seed-oil mutant of *Arabidopsis* with a deficiency in the seed-specific regulation of carbohydrate metabolism. *Plant Physiol*. 1998; 118:91–101. PMID: 9733529
27. Xu Ronghua, Wang R, Liu A. Expression profiles of genes involved in fatty acid and triacylglycerol synthesis in developing seeds of *Jatropha* (*Jatropha curcas* L.). *Biomass Bioenergy*. 2011; 35:1683–1692.
28. Zhang Y, Mulpuri S, Liu A. High light exposure on seed coat increases lipid accumulation in seeds of castor bean (*Ricinus communis* L.), a nongreen oilseed crop. *Photosynth. Res*. 2016; 128:125. <https://doi.org/10.1007/s11120-015-0206-x> PMID: 26589321
29. Fido RJ, Mills ENC, Rigby NM, Shewry PR. Protein Extraction from Plant Tissues. In: Cutler P, editors. *Protein Purification Protocols*. Humana Press, *Methods Mol. Biol*. 2004; 244:21–27. PMID: 14970539
30. Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem*. 1976; 72:248–254. PMID: 942051
31. Lin X, Tirichine L, Bowler C. Protocol: Chromatin immunoprecipitation (ChIP) methodology to investigate histone modifications in two model diatom species. *Plant Methods*. 2012; 8:48. <https://doi.org/10.1186/1746-4811-8-48> PMID: 23217141
32. Mukhopadhyay A, Deplancke B, Walhout AJM, Tissenbaum HA. Chromatin immunoprecipitation (ChIP) coupled to detection by quantitative real-time PCR to study transcription factor binding to DNA in *Caenorhabditis elegans*. *Nat. Protoc*. 2008; 3:698–709. <https://doi.org/10.1038/nprot.2008.38> PMID: 18388953
33. Baud S, Wuillème S, To A, Rochat C, Lepiniec L. Role of WRINKLED1 in the Transcriptional Regulation of Glycolytic and Fatty Acid Biosynthetic Genes in *Arabidopsis*. *Plant J*. 2009; 60:933–947. <https://doi.org/10.1111/j.1365-3113X.2009.04011.x> PMID: 19719479
34. Cagliari A, Margis-Pinheiroa M, Loss G, Mastroberti AA, de JE, Mariath A, et al. Identification and expression analysis of castor bean (*Ricinus communis*) genes encoding enzymes from the

- triacylglycerol biosynthesis pathway. *Plant Sci.* 2010; 179:499–509. <https://doi.org/10.1016/j.plantsci.2010.07.015> PMID: 21802608
35. Kanai M, Mano S, Kondo M, Hayashi M, Nishimura M. Extension of oil biosynthesis during the mid-phase of seed development enhances oil content in *Arabidopsis* seeds. *Plant Biotechnol. J.* 2016; 14:1241–1250. <https://doi.org/10.1111/pbi.12489> PMID: 26503031
 36. Bhardwaj HL, Hamama AA. Accumulation of glucosinolate, oil, and erucic acid in developing Brassica seeds. *Ind. Crops Prod.* 2003; 17:47–51.
 37. Orlando V, Strutt H, Paro R. Analysis of chromatin structure by *in vivo* formaldehyde cross-linking. *Methods.* 1997; 11:205–214. <https://doi.org/10.1006/meth.1996.0407> PMID: 8993033
 38. Maeo K, Tokuda T, Ayame A, Mitsui N, Kawai T, Tsukagoshi H, et al. An AP2-type transcription factor, WRINKLED1, of *Arabidopsis thaliana* binds to the AW-box sequence conserved among proximal upstream regions of genes involved in fatty acid synthesis. *Plant J.* 2009; 60:476–487. <https://doi.org/10.1111/j.1365-313X.2009.03967.x> PMID: 19594710
 39. Ohlroggeav J, Browse J. Lipid Biosynthesis. *Plant Cell.* 1995; 7:957–970. <https://doi.org/10.1105/tpc.7.7.957> PMID: 7640528
 40. Fukuda N, Ikawa Y, Aoyagi T, Kozaki A. Expression of the genes coding for plastidic acetyl-CoA carboxylase subunits is regulated by a location-sensitive transcription factor binding site. *Plant Mol. Biol.* 2013; 82:473–483. <https://doi.org/10.1007/s11103-013-0075-7> PMID: 23733600
 41. Kim HU, Lee K-R, Shim D, Lee JH, Chen GQ, Hwang S. Transcriptome analysis and identification of genes associated with ω -3 fatty acid biosynthesis in *Perilla frutescens* (L.) var. *Frutescens*. *BMC Genomics.* 2016; 17:474. <https://doi.org/10.1186/s12864-016-2805-0> PMID: 27342315
 42. Routaboul JM, Benning C, Bechtold N, Caboche M, Lepineic L. The TAG1 locus of *Arabidopsis* encodes for a diacylglycerol acyltransferase. *Plant Physiol. Biochem.* 1999; 37:831–840. PMID: 10580283
 43. Li R, Yu C, Li Y, Lam TW, Yiu SM, Kristiansen K, Wang J. SOAP2: an improved ultrafast tool for short read alignment. *BIOINFORMATICS.* 2009; 25:1966–1967. <https://doi.org/10.1093/bioinformatics/btp336> PMID: 19497933