Preprint: Please note that this article has not completed peer review.

Genome-wide analysis of physic nut MADS-box gene family and functional characterization of the JcMADS05 gene in transgenic rice

CURRENT STATUS: UNDER REVIEW

Yuehui Tang  
Zhoukou Normal University  
Email: yhtang2005@163.com
Corresponding Author  
ORCiD: https://orcid.org/0000-0001-5554-8688

Jian Wang  
Zhoukou normal University

Xinxin Bao  
Zhoukou Normal University

Qian Wu  
Zhoukou Normal University

Tongwen Yang  
Zhoukou Normal University

Han Li  
Zhoukou Normal University

Wenxia Wang  
Zhoukou Normal University

Yizhen Zhang  
Zhoukou Normal University

Nannan Bai  
Zhoukou Normal University

Yaxin Guan  
Zhoukou Normal University

Jiaxi Dai  
Zhoukou Normal University
Yanjie Xie  
Zhoukou Normal University  

Shen Li  
Zhoukou Normal University  

Rui Huo  
Zhoukou Normal University  

Wei Cheng  
Zhoukou Normal University  

DOI:  
10.21203/rs.2.23080/v1  

SUBJECT AREAS  
Epigenetics & Genomics  

KEYWORDS  
Physic nut, MADS-box proteins, JcMADS05, abiotic stress, seed size, transgenic plants
Abstract

**Background:** Physic nut (*Jatropha curcas*), a non-edible oilseed plant, is among the most promising alternative energy sources because of its high seed oil content, rapid growth and extensive adaptability. Proteins encoded by MADS-box genes are important transcription factors involved in the regulation of plant growth, seed development and responses to abiotic stress. However, there has been no in-depth research on the MADS-box genes and their roles in physic nut.

**Results:** In the present study, 63 *MADS-box* genes (*JcMADSs*) were identified in the physic nut genome, and classed into five groups (MIK, Mα, Mβ, Mγ, Mδ) based on phylogenetic comparison with *Arabidopsis* homologs. Expression profile analysis based on RNA-seq suggested that many *JcMADS* genes were expressed most strongly in seeds, and seven of them responded in leaves to at least one abiotic stressor (drought and/or salinity) at one or more time points. Transient expression analysis and a transactivation assay indicated that *JcMADS05* is a nucleus-localized transcriptional activator. Plants overexpressing *JcMADS05* did not show altered plant growth, but the overexpressing plants did exhibit reductions in grain size, grain length, grain width, 1000-seed weight and yield per plant. Further data on the reduced grain size in *JcMADS05*-overexpressing plants supported the putative role of *JcMADS* genes in seed development.

**Conclusions:** This study will be useful in understanding the involvement of *MADS-box* genes in growth and development in addition to their functions in abiotic stress resistance, and will ultimately form the basis for functional characterization and the exploitation of candidate genes for the genetic engineering of physic nut.

Background

The regulation of plant growth, development and stress responses is complex and is coordinated by many mechanisms. These mechanisms are under the control of a series of related genes acting through complex regulatory networks. In these processes, transcription factor (such as members of the MYB, HD-Zip, ARF, NAC, MADS-box, and ERF gene families) specifically recognize cis-regulatory elements present in the promoter regions of these genes, and regulate their expression so as to modulate a wide range of biochemical, physiological and developmental processes [1–6] (Ferrandiz
The MADS-box proteins constitute one of the largest families of transcription factors. Structurally, almost all MADS-box proteins contain a highly conserved DNA binding domain of 58 amino acids residues which binds to DNA at consensus recognition sequences known as CArG boxes [CC(A/T)6GG], regulating the transcription of downstream genes [7]. Other domains are also present in some MADS-box proteins; they include the keratin (K) domain, which is responsible for dimerization, the I (intervening) domain, which is capable of binding to form dimers, and the C-terminal (C) region, which is the most variable region [8]. Based on amino acid sequence similarities and structural features of the conserved domains, MADS-box proteins in Arabidopsis can be classified into five groups (designated MIKC, Mα, Mβ, Mγ, Mδ) [9].

In recent years, many MADS-box proteins have been identified and characterized, by means of genome-wide and expression profiles analysis studies, from a number of plant species including rice [10], wheat [11], potato [12], moso bamboo [13], Arabidopsis [9] and sheepgrass [14]. Studies on loss-of-function and gain-of-function mutants in dicots and monocots have suggested that members of MADS-box family play crucial roles in controlling plant growth and development [1]. For example, the abs stk double mutant has a reduced number of fertilized ovules and undergoes seed abortion resulting in very few seeds [15]. In rice, OsMADS87 RNAi lines produce smaller seeds and overexpression (OE) lines have larger seeds compared with wild-type seeds [16]. The Arabidopsis agl62 mutant shows accelerated endosperm cellularization, showing that AGL62 is required for suppression of cellularization [17]. Thesvp and flm mutants exhibit temperature-insensitive flowering across wide temperature ranges [18], while FUL has a role in leaf and fruit development and in flowering in Arabidopsis [19]. SEP1, SEP2, and SEP3 are critical for development of petals, stamens and carpels [20]. In addition to participating in regulating plant growth and development, evidence is accumulating to suggest that MADS-box genes are involved in plant responses to salt and drought stresses. For example, CaMADS-overexpressing Arabidopsis plants show increased resistance of cold and salt stresses [21]. Loss-of-function mutations in SVP result in sensitivity to drought stress, while the SVP-overexpressing plants are more tolerant [22]. In tomato, SIMBP11-RNAi lines are less
tolerance to salinity stress, but overexpressing this MADS-box gene confers salt stress tolerance [23].

However, although various members of MADS-box family have been cloned and functionally studied, little is known about the members of this family and their roles in many taxa, including the Euphorbiaceae.

Physic nut, a small woody member of the Euphorbiaceae, is a non-food oilseed crop grown mainly in tropical and subtropical regions and its seed oil is used widely in industry [24]. Importantly, it has drawn much attention because it is suitable for biodiesel production owing to its rapid growth, ease of propagation, high oil content and extensive adaptability [24]. Further research is thus necessary in order to elucidate the molecular mechanisms that underlie the regulation of growth and development of physic nut. The recent completion of the physic nut genome sequence allows us to explore all the JcMADS genes at the genome level. However, there has as yet been no systematic study on the identification, classification, expression profiles or functions of MADS-box genes in this species. To address this deficiency, we firstly searched for and identified 63 MADS-box genes in the physic nut genome (hereafter referred to as JcMADS genes). Secondly, we investigated their phylogenetic relationships, conserved motifs, chromosomal distribution, expression profiles and potential roles in physic nut development. Finally, JcMADS05 was chosen for further functional analysis, and we tested its effects in transgenic crop. This study focuses on the functional roles of JcMADS genes in the development of physic nut.

Results
Identification of MADS-box proteins in physic nut
All Arabidopsis MADS-box protein sequences were used as queries in a BlastP search to identify physic nut proteins. A Hidden Markov Model (HMM) search was also performed against the physic nut protein database using the MADS-domain PF03106. In total, 63 putative MADS proteins were identified in physic nut, with the presence of the MADS-box domain in each of them being confirmed by a SMART website search. These genes were named sequentially from JcMADS01 to JcMADS63 according to their chromosomal locations (Additional File 1). The 63 JcMADS genes ranged in length from 195 (JcMADS35) to 1164 (JcMADS34), thus potentially the proteins encoded would be from 64 to
387 amino acids; their GenBank accession numbers are given in Additional File 1.

**Phylogenetic Relationships Of The Jcmads Proteins**

To clarify the phylogenetic relationships of the physic nut MADS family proteins with the previously reported members of the family in Arabidopsis, an unrooted tree was constructed by MEGA6 using the neighbor-joining method (Fig. 1). On the basis of full length amino acid sequence conserved domain and similarity, we subdivided the 171 typical members of the MADS gene family into 5 groups (designated MIKC, Mα, Mβ, Mγ, Mδ), according to the previous classification of MADS proteins from Arabidopsis [9]. Of the 63 inferred physic nut JcMADS proteins, thirty-two were assigned to group MIKC (JcMADS03, 04, 05, 07, 12, 16, 17, 19, 20, 22, 24, 25, 26, 27, 28, 29, 31, 32, 37, 38, 42, 44, 46, 47, 49, 50, 52, 53, 54, 55, 56, 63), thirteen to group Mα (JcMADS01, 06, 08, 11, 13, 14, 33, 35, 39, 43, 57, 58, 59), four to group Mβ (JcMADS09, 48, 51, 60), six to group Mγ (JcMADS10, 15, 34, 36, 41, 62) and eight to group Mδ (JcMADS02, 18, 21, 23, 30, 40, 45, 61). In the phylogenetic tree, some members of the JcMADS gene family formed related sister pairs (Fig. 1): JcMADS11 and 35, 13 and 14, 02 and 45, 10 and 36, 07 and 55, 32 and 54. There were also triplets (JcMADS06, 57 and 58; 34, 41 and 62; 09, 51 and 60; 03, 04 and 05), and a set of quadruplets in the case of JcMADS01, 08, 33 and 39. The tree indicated that proteins in group MIKC were the most numerous; it contained 39 AtMADS and 32 JcMADS proteins

**Conserved Motifs In Jcmads Proteins**

The structures of proteins encoded by JcMADS genes were analyzed using the MEME online software tool. Twenty conserved motifs, which we named motifs 1–20, were identified in the 63 JcMADS proteins (Fig. 2 and Additional File 2). As expected, motif 1 and motif 4 corresponded to the typical MADS-box domain, and motif 1 was found in all the JcMADS proteins. Motif 9, specifying the K domain, was found in most MIKC type proteins; the exceptions were JcMADS17, 31, 37, 52, 56 and 63, which were relatively short amino acid sequences. In addition to these known functional motifs, some of unknown function were also detected. Examples included motifs 5, 10 and 20 (detected only in group Mγ), and motif 13 (found only in groups MIKC and Mβ). Motif 12 was found only in group MIKC, while motifs 14 and 19 was detected only in group Mδ. Additionally, most conserved motifs detected in
JcMADS proteins were clade-specifically assigned in different groups, implying similarity of function within a given group.

Chromosomal localization of JcMADS genes

We mapped 62 of the 63 JcMADS genes (all except JcMADS63) to LGs based on a previously published report on the physic nut genome [25]. As shown in Fig. 3, we found that LGs 4 and 7 had more members of the JcMADS gene family than other LGs, with eleven and nine JcMADS genes respectively. They were followed by LGs 2, 3, 5 and 10, each of which had six JcMADS genes. In addition, there was five JcMADS genes on each of LGs 6 and 9, three on LG8, three on LG11 and two on LG1. The results also indicated that most JcMADS genes were on the lower and middle regions of the LGs. Tandem duplications, defined as tandem repeats which are separated by < 4 non-homologous spacers or are genes located within 50 kb of each other [26], were found among these members of the JcMADS gene family. The gene pairs present as tandem repeats (T) were T1 (JcMADS03 and 04), T2 (JcMADS06 and 07), T3 (JcMADS12, 13 and 14), T4 (JcMADS19 and 20), T5 (JcMADS26 and 27), T6 (JcMADS30 and 31), T7 (JcMADS33 and 34), T8 (JcMADS37 and 38), T9 (JcMADS43 and 44), T10 (JcMADS46 and 47), T11 (JcMADS50, 51 and 52) and T12 (JcMADS55, 57 and 58), on LG2, LG2, LG3, LG4, LG5, LG5, LG6, LG7, LG7, LG8, LG9 and LG10 respectively.

Expression profile of JcMADS genes under non-stressed growth condition

To clarify the roles of the JcMADS in regulating physic nut development, we examined the expression profiles of JcMADS genes in roots, stem cortex, leaves, and seeds (S1 and S2) under non-stressed growth conditions based on data from RNA sequencing (RNA-seq) (Additional File 3 and Fig. 4). The result suggested that fifty of the predicted JcMADS genes were expressed in at least one of the organs examined, while thirteen (JcMADS06, 13, 17, 35, 39, 40, 49, 52, 57, 58, 60, 61 and 63) were not expressed in any of these tissues. Of the 50 JcMADS genes for which expression was detected, two (JcMADS09 and 47) were highly expressed across all the tissues sampled, ten (JcMADS03, 05, 11, 12, 14, 23, 34, 37, 46 and 62) were expressed only in seeds, thirteen (JcMADS04, 15, 21, 24, 25, 29, 38, 41, 43, 48, 50, 51 and 54) exhibited highest expression in seeds, four (JcMADS08, 16, 42 and 44) preferred to be expressed in roots, and one (JcMADS28) was most strongly expressed in the stem
As shown in Fig. 4, most of the JcMADS genes were expressed more highly in seeds at the S1 stage than at the S2 stage. It was noteworthy that nine genes (JcMADS03, 05, 12, 14, 25, 34, 46, 48 and 62) was detected expression only in seeds at S1 stage. Based on the results of expression pattern analysis, the JcMADS05 gene was chosen for functional research.

Expression profile of JcMADS under abiotic stress conditions
Many studies have suggested that some MADS-box genes encode proteins involved in the regulation of abiotic stresss [21, 27-28]. We therefore further investigated the patterns of expression of JcMADS genes in leaves after 2 d, 4 d and 7 d of drought stress and after 2 h, 2 d and 4 d of salinity stress according to data from RNA-Seq. As shown in Fig. 5, the transcript abundances of seven JcMADS genes indicated at least a twofold enhancement or reduction compared with the control in response to at least one stress at one or more time points. Of these seven genes detected as having differential expression, three (JcMADS42, 43 and 47) exhibited significantly induced or inhibited expression in response to drought and salinity stresses, three (JcMADS22, 30 and 53) showed differential expression only in response to drought stress, and JcMADS15 responded solely to salt stress.

JcMADS05 is a nucleus-localized transcriptional activator
To confirm the subcellular localization of the protein encoded by JcMADS05 gene, the 35S:JcMADS05-YFP fusion construct and the 35S:YFP empty vector were introduced into Arabidopsis protoplasts. The fluorescence signals from the protoplasts were then observed immediately by confocal laser-scanning microscopy. As shown in Fig. 6, we observed that the yellow fluorescent signal was distributed throughout the whole of the cell when the 35S:YFP vector was used, whereas in protoplasts harboring the construct 35S:JcMADS05-YFP a strong yellow fluorescent signal was detected in the nuclei. These findings indicate that JcMADS05 gene is located in the nucleus.

A dual-luciferase assay was used to examine the transcription activation activity of JcMADS05 protein. The full-length CDS of JcMADS05 was attached to the vector pBD, then the pBD-JcMADS05 fusion effector vector and the p5 × GAL-Reporter vector were transformed into Arabidopsis protoplasts. The results indicated that the LUC/REN ratio was significantly lower in the control protoplasts (pBD) than
in the pBD-JcMADS05 group. Our data suggest that the full-length JcMADS05 has transactivation activity (Fig. 7). Based on the above results, we drew the conclusion that JcMADS05 functions as a transcription activator.

Phenotypic analysis of transgenic rice plants expressing JcMADS05

To investigate the role of JcMADS05 gene in regulating plant development, and to assess the feasibility of using JcMADS genes to control seed size in an important crop plant, we overexpressed this gene in rice. Three independent transgenic lines (OE1, OE2 and OE3) were confirmed as expressing JcMADS05 expression using semi-quantitative RT-PCR, and selected for further study. Expression of JcMADS05 were detected in transgenic lines, whereas no expression was found in WT (wild type) plants (Fig. 8C). Phenotypic analysis showed that the growth and flower structure of transgenic plants overexpressing JcMADS05 were similar to those of WT plants (Fig. 8A and B).

Statistical analysis indicated that there was no obvious difference in root and shoot lengths in the transgenic plants compared to the WT plants (Fig. 8D and E). Taken together, these results led to the conclusion that JcMADS05 did not have any major effect on the growth of the transgenic plants.

Overexpression of JcMADS05 reduces the grain size in transgenic rice

As described above, JcMADS05 expression was most strongly detected in seed, suggesting that JcMADS05 might have an important role in seed growth and development. To test this, we examined the effects of JcMADS05 overexpression on rice grain size. We found that JcMADS05 transgenic plants produced dramatically smaller seeds than the WT lines (Fig. 9A). The results also showed that JcMADS05 transgenic plants had a significant reduction in both grain length and width compared to the WT plants (Fig. 9B and C). We also detected a significant reduction in 1000-seed weight and yield per plant in JcMADS05 transgenic lines (Fig. 9D and E). Our data suggested that overexpressing JcMADS05 significantly altered seed size in transgenic plants.

To study the molecular mechanism of JcMADS05 gene regulates grain size, we further tested the expression of grain-size-related genes (Fig. 9F). The results showed that expression of some positive regulatory factors, such as GS2, SMG11, was significantly lower in transgenic plants than that in wild type, while expression of some negative regulatory factors, such as OsMKP1, GW2, was obviously
higher than that in wild-type. Taken together, these data supported a putative role for JcMADS genes in seed development.

Discussion

Increasing evidence suggests that the MADS-box genes are involved in a series of plant physiological phenomena. Up to now, many studies on the functions of the MADS-box genes have been focused on the model plants rice and Arabidopsis [27]. The molecular mechanisms involved in seed development in the biofuel plant physic nut, and more specifically the identities, expression profiles and functions of its MADS-box genes remain poorly understood. We therefore characterized and examined expression profiles of MADS-box genes in this species, and chose one (designated JcMADS05) that was most strongly expressed in seed for further functional analysis by transgenically expressing it in rice.

In the present study, a total of 63 MADS-box genes were identified in the physic nut genome. Compared with Arabidopsis (genome size 125 Mb) and rice (genome size 466 Mb), the MADS-box family seems to have relatively fewer members in the physic nut genome (genome size 320 Mb) [25]. One possible explanation for the smaller number of JcMADS genes may be that MADS-box family genes in the physic nut genome did not undergo a chromosomal segment duplication event during the early evolution of the species [25], whereas such duplications made major contributions to the expansion of both rice MADSs and Arabidopsis MADSs [9-10]. Our phylogenetic tree showed that there were twenty-two MADS genes in the Mβ group in Arabidopsis, whereas there were only four JcMADS genes in group Mβ (Fig. 1). These finding suggests that the members of the group may have been either acquired in the Arabidopsis lineage or lost in the physic nut after divergence from the last common ancestor shared by Arabidopsis and physic nut. Motif analysis indicated that the distribution of protein motifs across the different groups was diverse, but members of the same group had a similar motif complement (Fig. 2), supporting their strong evolutionary conservation. Similar results have been observed in a variety of plants, including bread wheat [29], potato [12], moso bamboo [13], Arabidopsis [9] and sheepgrass [14]. Our results show that the evolution and classification of the members of MADS-box family is quite highly conserved in the physic nut, as it is in other plant
species.

Preliminary predictions about the biological functions of genes and their products can be made by analysis of gene expression profiles, we therefore detected the expression of 63 MADS-box genes sequencing-based transcriptome data. Our results show that JcMADS25 expression was highest in seeds. Its Arabidopsis homolog AGL11(STK) is essential for seed development [15], and its homolog in oil palm SHELL controls seed oil yield [30]. It can therefore be inferred from the high levels of JcMADS25 expression in physic nut seeds that it may be involved in regulation of the development of physic nut seed. JcMADS16 was most highly expressed in roots, and in Arabidopsis, its homologous AGL12 is also preferentially expressed in root tissues and is essential for root development [31–32]. The results indicated that JcMADS16 may play an important in physic nut root development. TT16, a MADS-box transcription factor, which affects seed development [33], and its homolog JcMADS05, are preferentially expressed in seeds, suggesting that JcMADS05 may have a significant regulatory role in seed development. JcMADS09 and 47 are expressed in all tissues (Fig. 4), indicating that they may participate in the overall development of the physic nut plant. It is worth noting that many JcMADS genes showed preferential expression in seeds, implying that they may all be very important for physic nut plants in seed development. Overall, we deduce that JcMADS genes may have functions in each growing stages of physic nut plants; further study is required to confirm their roles.

Research increasingly have suggested that MADS-box genes participate in responses to various abiotic stresses in many plant species [21, 27–28]. For example, CaMADS, which is strongly induced by salinity stress, and by abscisic acid, acts as a mediator that has positive feedback effects on cold and salinity stress signaling pathways in pepper [21]. OsMADS26 is a regulator of drought stress response in rice [28]. AGL22 gene plays a crucial role in connecting changes in the initiation of drought stress responses and primary metabolism [34]. Although some studies have begun to identify certain genes of the MADS-box family as important molecular components of abiotic stress responses, we have hitherto lacked integrated information about the responses of members of MADS-box family to drought and salt stresses in physic nut. In this study, RNA-based sequencing data in response to drought and salt stress, combined with qRT-PCR analysis, enabled us to identify JcMADS genes that
respond to abiotic stress. For example, expression of JcMADS42, 43 and 47 was induced or inhibited by salt and drought stresses at one or more time points, whereas JcMADS22, 30 and 53 responded only to drought stress (Fig. 5). Collectively, we preliminary judgment these JcMADS genes may play important roles in the regulation of abiotic stress responses in physic nut, and their functions merit further investigation.

Grain size, which is an important agronomic trait in many crops, is determined by grain width, grain length and grain thickness [35]. Although some genes that manipulate seed size have been identified in crop plants [36], the molecular mechanisms that regulate seed size are still poorly understood. In our work, we found that JcMADS05, a gene in group MIKC, was most highly expressed in the seeds of physic nut (Fig. 4), and in order to investigate its function we tested its effects in transgenic rice. Our results showed that the JcMADS05 transgenic plants had smaller, shorter and narrower grains and lower 1000-grain weight compared with wild-type plants (Fig. 9). Furthermore, our results also suggested that JcMADS05 overexpressing plants reduced expression of GS2 and SMG11, and increased expression of OsMKP1 and GW2 (Fig. 9F). SMG11 overexpressing plants increases grain size by altering expression of several grain-size-related genes [35]. Gain of function mutations in OsMKP1 reduces grain size, conversely, results in larger grain [37]. Loss of GW2 function increase grain size and grain weight [38]. Overexpression of GS2 increases grain size and weight [39]. Collectively, JcMADS05 overexpressing plants reduces grain size, at least in part, by influencing expression of grain-size-related genes. These findings further support a role for JcMADS05 in negative regulation of grain size. Given that reporter gene studies in Arabidopsis protoplasts suggested that JcMADS05 is likely to act as a transcriptional activator, its function as a negative regulator of grain size could be mediated via activation of other repressors. In summary, the results provide a novel resource for future studies on MADS-box genes in physic nut, especially with respect to their effects on seed size.

Conclusions
In this study, we identified 63 JcMADS genes in the physic nut genome, and characterized their expression profiles under normal growth and abiotic stress conditions. Transgenic expression in rice of one of the genes (JcMADS05) reduced grain size, 1000-grain weight and yield per plant, supporting
the hypothesis that some members of the family participate in the regulation of physic nut seed development. These findings provide insights facilitating prediction of the functions of MADS-box genes in stress tolerance and seed development, and comprehensive analysis of the gene family produced results that will be helpful in screening genes for further functional characterization and for the genetic improvement of agronomic traits in physic nut.

Methods

Plant materials

An inbred cultivar of J. curcas, GZQX0401, was used in this research, since its genome has been fully sequenced [25]. Seeds of the cultivar was obtained from South China Botanical Garden, Chinese Academy of Sciences, Guangzhou, China. The wild-type rice (Oryza sativa L.) line used was japonica cv. Zhonghua 11 (ZH11). Seeds were germinated and plants cultivated in soil in basins in a greenhouse under natural sunlight at Zhoukou Normal University, China.

Identification Of Mdas-box Gene In Physic Nut

One hundred and eight previously identified Arabidopsis MADS-box protein sequences were used as queries in a search against the physic nut genome database. In addition, the HMM profile of the conserved MADS-box domain (PF00319), obtained from the Pfam database (Pfam 32.0, http://pfam.xfam.org/), was used to carry out a BLASTP search against the physic nut genome database with the expected value (e-value) cut-off set at 0.01. All protein sequences identified as containing the MADS domain were confirmed through the SMART service (http://smart.embl-heidelberg.de/). The theoretical pI and molecular weight of all confirmed JcMADS proteins were determined using the ExPASy ProtParam tool (http://expasy.org/).

Phylogenetic Analysis

MADS protein sequences from Arabidopsis were downloaded from the TAIR database (https://www.arabidopsis.org/), and sequences for Jatropha curcas were downloaded from GenBank (http://www.ncbi.nlm.nih.gov/). The ClustalX (1.83) program was used to carry out alignment of multiple amino acid sequences. Phylogenetic trees comparing physic nut and Arabidopsis MADS proteins were constructed using the neighbor-joining method and 1000 bootstraps, and the results were displayed with the Mega Software Version 6.

Conserved Motif And Chromosomal Distribution
The conserved motifs of individual MADS-box proteins were identified using the MEME server (http://alternate.meme-suite.org/). MEME was run with the following parameters: site distribution (Zero or one occurrence per sequence), motif number (20), motif width (between 6 and 100 wide). Chromosomal locations of JcMADS genes were obtained as described by Wu [25], and linkage maps of the JcMADS genes were drawn with the MapChart software package.

Expression profile analysis of JcMADS genes
To detected the transcript abundance of JcMADS genes in physic nut, root, leaf and stem cortex tissue from two-week-old seedlings, and seeds at day 14 (early developmental stage S1) and day 35 (filling and maturation stage S2) after pollination, were collected and stored in an ultra-low temperature freezer for further digital gene expression analysis. Salinity and drought stresses were initiated at the six-leaf stage. For salinity stress, Hoagland solution containing 100 mM NaCl was used to irrigate seedlings daily. To apply drought stress, stop watering. Leaves were collected at 2 days, 4 days and 7 days after drought stress and 2 hours, 2 days and 4 days after salinity stress. Samples were immediately stored in an ultra-low temperature freezer for further digital gene expression and qRT-PCR analyses; raw sequence data were obtained based on standard protocols and submitted to the sequence read archive (SRA) at NCBI (accession nos. are PRJNA257901 (for the drought stress data) and PRJNA244896 (for the salt stress data)).

Subcellular Localization And Transcriptional Activation Analysis
The amplified coding region of JcMADS05 without the termination codon was inserted into the pSAT6-eYFP-N1 vector to generate 35S::JcMADS05-YFP. The 35S::JcMADS05-YFP fusion expression construct and the 35S::YFP empty vector were transformed into Arabidopsis protoplasts using the polyethylene glycol-mediated method. Subcellular localization of the control YFP and JcMADS05-YFP fusion proteins was observed under a confocal laser scanning microscope. Arabidopsis protoplasts were prepared following Tang [40].

For the transactivation assay of JcMADS05, the full-length JcMADS05 gene was fused to the pBD vector to generate the construct pBD-JcMADS05. This construct and the p5 × GAL-Reporter vector were introduced into Arabidopsis protoplasts. A ProteoPrep® Total Extraction Sample Kit (Sigma) was
used to extract total protein from Arabidopsis protoplasts according to the manufacturer’s instructions, then the fluorescence activity of proteins was analyzed using the enzyme-labeled instrument. The LUC/REN ratio was used to measure the transcriptional activation of JcMADS05.

Gene Cloning And Plant Transformation
The coding sequence of JcMADS05 was amplified by RT-PCR from total RNA isolated from physic nut seed with the JcMADS05-F and JcMADS05-R primers given in Table S1, and cloned into the pMD18-T vector. Successful amplification of the target gene was confirmed by DNA sequencing. And then the target sequence was excised from the pMD18-T vector after digestion with Kpn I and Xba I, then cloned into the pCAMBIA1301 vector at the Kpn I/Xba I site under the control of the CaMV 35S promoter. The resulting constructs were introduced into Agrobacterium (strain EHA105) by the freeze–thaw procedure, and strains harboring the constructs were transformed into calli of rice cv. ZH11, according to the method of Tang [40].

Phenotypic Analysis And Evaluation Of The Yield-related Traits
Thirty individual plants of each of the JcADS05-overexpressing (OE1, OE2, and OE3) and wild-type lines were used to measure root and shoot lengths, 1000-grain weight, grain length and width, and grain yield per plant and to examine flower structure. Each line contained three independent biological replicates

Rna Isolated And Qrt-pcr Analysis
Leaves from 2-week-old wild type and transgenic rice seedlings were sampled and stored at – 80℃ until required for use. Total RNA from different organs was extracted using a HiPure Plant RNA Mini Kit (Code No.R4151-02, Magen, http://www.magentec.com.cn/about.php). The cDNA synthesized using a PrimeScript™ IV 1st strand cDNA Synthesis Mix (TAKARA, Beijing, China). qRT-PCR was performed on a Mini Option real-time PCR system (LightCycler 480). Cycling conditions were as follows: 95 °C for 30 s, 95 °C for 5 s, 60 °C for 20 s, and 72 °C for 20 s. The reaction was carried out for 40 cycles. We used the $2^{-\Delta\Delta CT}$ method to detect relative transcript abundance, and rice OsUbiquitin gene was used for normalization. The primers used employed in Additional File 4.

Statistical analysis
In the research, each experiment contained three biological replicates. Statistical analysis was
performed using SAS software package according to the Duncan multiple range test (Duncan, 1955).

Declarations

**Abbreviations**

HD, homeodomain; ORF, open reading frame; CDS, coding sequence.

**Ethical approval and consent to participate**

Not applicable

**Consent to publish**

Not applicable

**Availability of data and materials**

Raw data supporting findings of this study have been submitted to NCBI’s sequence read archive (SRA) (accession nos. for the salinity and drought stress data: PRJNA244896 and PRJNA257901, respectively). Acquired sequences of *JcMADS* proteins are available from DDBJ/EMBL/GenBank under accession no. AFEW0000000NCBI. Other relevant data obtain during the research are included in this published article and associated supplementary information files.

**Competing interests**

The authors declare that they have no competing interests.

**Funding**

This research was supported by the Foundations of He’nan Science and Technology Committee (Grant No 182102110200), which provided funding for the research of JcMDAS protein, and publication of our paper.

**Authors’ contributions**

The research was conceived and designed by XB and YT. The experiments were performed by QW, HL, WW, YZ, NB, YG and JD, and the data were analyzed by WJ, YX, SL, RH and WC. The manuscript was written and revised by YT and TY. All the authors read and approved the final manuscript.

**Acknowledgements**

The authors give many thanks to the Key Laboratory of Plant Resources Conservation and Sustainable Utilization, South China Botanical Garden, Chinese Academy of Sciences, Guangzhou, China for
providing the physic nut seeds.

**Authors' information**

1. Key Laboratory of Plant Genetics and Molecular Breeding, Zhoukou Normal University, Henan, Zhoukou, China.
2. Henan Key Laboratory of Crop Molecular Breeding and Bioreactor, Henan, Zhoukou, China.
3. School of Journalism and Communication, Zhoukou Normal University, Henan, Zhoukou, China.

**References**

1. Ferrandiz C, Fourquin C: Role of the FUL-SHP network in the evolution of fruit morphology and function. J Exp Bot. 2014; 65(16):4505-
2. Gong X, Zhao L, Song X, Lin Z, Gu B, Yan J, Zhang S, Tao S, Huang X: Genome-wide analyses and expression patterns under abiotic stress of NAC transcription factors in white pear (*Pyrus bretschneideri*). BMC Plant Biol. 2019;19(1):161.
3. Li SB, Xie ZZ, Hu CG, Zhang JZ: A review of auxin response factors (ARFs) in plants. Front Plant Sci. 2016; 7:47.
4. Roy S: Function of MYB domain transcription factors in abiotic stress and epigenetic control of stress response in plant genome. Plant Signal Behav. 2016; 11(1):e1117723.
5. Tang Y, Wang J, Bao X, Liang M, Lou H, Zhao J, Sun M, Liang J, Jin L, Li G et al: Genome-wide identification and expression profile of HD-ZIP genes in physic nut and functional analysis of the *JcHDZ16* gene in transgenic rice. BMC Plant Biol. 2019; 19(1):298.
6. Xie Z, Nolan TM, Jiang H, Yin Y: AP2/ERF Transcription factor regulatory networks in hormone and abiotic stress responses in *Arabidopsis*. Front Plant Sci. 2019; 10:228.
7. Messenguy F, Dubois E: Role of MADS box proteins and their cofactors in combinatorial control of gene expression and cell development. Gene. 2003; 316:1-
8. Smaczniak C, Immink RG, Angenent GC, Kaufmann K: Developmental and evolutionary diversity of plant MADS-domain factors: insights from recent studies. Development. 2012; 139(17):3081–

9. Parenicova L, de Folter S, Kieffer M, Horner DS, Favalli C, Busscher J, Cook HE, Ingram RM, Kater MM, Davies B et al: Molecular and phylogenetic analyses of the complete MADS-box transcription factor family in Arabidopsis: new openings to the MADS world. Plant Cell. 2003; 15(7):1538–

10. Arora R, Agarwal P, Ray S, Singh AK, Singh VP, Tyagi AK, Kapoor S: MADS-box gene family in rice: genome-wide identification, organization and expression profiling during reproductive development and stress. BMC Genomics. 2007; 8:242.

11. Schilling S, Kennedy A, Pan S, Jermiin LS, Melzer R: Genome-wide analysis of MIKC-type MADS-box genes in wheat: pervasive duplications, functional conservation and putative neofunctionalization. New Phytol. 2019; 225(1):511–

12. Gao H, Wang Z, Li S, Hou M, Zhou Y, Zhao Y, Li G, Zhao H, Ma H: Genome-wide survey of potato MADS-box genes reveals that StMADS1 and StMADS13 are putative downstream targets of tuberigen StSP6A. BMC Genomics. 2018; 19(1):726.

13. Zhang Y, Tang D, Lin X, Ding M, Tong Z: Genome-wide identification of MADS-box family genes in moso bamboo (Phyllostachys edulis) and a functional analysis of PeMADS5 in flowering. BMC Plant Biol. 2018; 18(1):176.

14. Jia J, Zhao P, Cheng L, Yuan G, Yang W, Liu S, Chen S, Qi D, Liu G, Li X: MADS-box family genes in sheepgrass and their involvement in abiotic stress responses. BMC Plant Biol. 2018; 18(1):42.

15. Chiara M, Marta Adelina M, Elisabetta C, Arp S, Kater MM, Raffaella B, Lucia C: The MADS box genes SEEDSTICK and ARABIDOPSISIB sister play a maternal role in fertilization and seed development. Plant J. 2012; 70(3):409–
16. Chen C, Begcy K, Liu K, Folsom JJ, Wang Z, Zhang C, Walia H: Heat stress yields a unique MADS box transcription factor in determining seed size and thermal sensitivity. Plant Physiol. 2016; 171(1):606-

17. Kang I-H, Steffen JG, Portereiko MF, Lloyd A, Drews GN: The AGL62 MADS domain protein regulates cellularization during endosperm development in Arabidopsis. Plant Cell. 2008; 20(3): 635-

18. Lee JH, Ryu HS, Chung KS, Pose D, Kim S, Schmid M, Ahn JH: Regulation of temperature-responsive flowering by MADS-box transcription factor repressors. Science. 2013; 342(6158):628-

19. Balanza V, Martinez-Fernandez I, Ferrandiz C: Sequential action of FRUITFULL as a modulator of the activity of the floral regulators SVP and SOC1. J Exp Bot. 2014; 65(4):1193-

20. Pelaz S, Ditta GS, Baumann E, Wisman E, Yanofsky MF: B and C floral organ identity functions require SEPALLATA MADS-box genes. Nature. 2000; 405(6783):200-

21. Chen R, Ma J, Luo D, Hou X, Ma F, Zhang Y, Meng Y, Zhang H, Guo W: CaMADS, a MADS-box transcription factor from pepper, plays an important role in the response to cold, salt, and osmotic stress. Plant Sci. 2019; 280:164-

22. Wang Z, Wang F, H Y, Y J, R Z, S H, and Z, J: The flowering repressor SVP confers drought resistance in Arabidopsis by regulating abscisic acid catabolism. Mol plant. 2018; 1184-

23. Guo X, Chen G, Cui B, Gao Q, Hu Z: Solanum lycopersicum agamous-like MADS-box protein AGL15-like gene, SIMBP11, confers salt stress tolerance. Mol Breeding. 2016; 36(9):125.

24. Openshaw K: A review of Jatropha curcas: an oil plant of unfulfilled promise. Biomass & Bioenergy, 2000; 19(1):1-
25. Wu P, Zhou C, Cheng S, Wu Z, Lu W, Han J, Chen Y, Chen Y, Ni P, Wang Y et al: Integrated genome sequence and linkage map of physic nut (*Jatropha curcas*), a biodiesel plant. Plant J. 2015; 81(5):810-821.

26. Cannon SB, Mitra A, Baumgarten A, Young ND, May G: The roles of segmental and tandem gene duplication in the evolution of large gene families in *Arabidopsis thaliana*. BMC Plant Biol. 2004; 4(1):10.

27. Castelán-Muñoz N, Herrera J, Cajero-Sánchez W, Arrizubieta M, Trejo C, García-Ponce B, Sánchez MdIP, Álvarez-Buylla ER, Garay-Arroyo A: MADS-Box genes are key components of genetic regulatory networks involved in abiotic stress and plastic developmental responses in plants. Front Plant Sci. 2019; 10:853.

28. Khong GN, Pati PK, Richaud F, Parizot B, Bidzinski P, Mai CD, Bes M, Bourrie I, Meynard D, Beeckman T et al: OsMADS26 negatively regulates resistance to pathogens and drought tolerance in rice. Plant Physiol. 2015; 169(4):2935–

29. Ma J, Yang Y, Luo W, Yang C, Ding P, Liu Y, Qiao L, Chang Z, Geng H, Wang P et al: Genome-wide identification and analysis of the MADS-box gene family in bread wheat (*Triticum aestivum*). PLoS One. 2017; 12(7):e0181443.

30. Singh R, Low ET, Ooi LC, Ong-Abdullah M, Ting NC, Nagappan J, Nookiah R, Amiruddin MD, Rosli R, Manaf MA et al: The oil palm SHELL gene controls oil yield and encodes a homologue of SEEDSTICK. Nature. 2013; 500(7462):340–

31. Burgeff C, Liljegren SJ, Tapia-Lopez R, Yanofsky MF, Alvarez-Buylla ER: MADS-box gene expression in lateral primordia, meristems and differentiated tissues of *Arabidopsis thaliana* Planta. 2002; 214(3):365–372.

32. Pacheco-Escobedo MA, Ivanov VB, Ransom-Rodriguez I, Arriaga-Mejia G, Avila H, Baklanov IA, Pimentel A, Corkidi G, Doerner P, Dubrovsky JG et al: Longitudinal zonation pattern in *Arabidopsis* root tip defined by a multiple structural change
algorithm. Ann Bot. 2016;118(4): 763–

33. Prasad K, Ambrose BA: Shaping up the fruit: control of fruit size by an Arabidopsis B-sister MADS-box gene. Plant Signal Behav. 2010; 5(7):899–

34. Bechtold U, Penfold CA, Jenkins DJ, Legaie R, Moore JD, Lawson T, Matthews JS, Violet-Chabrand SR, Baxter L, Subramaniam S et al: Time-series transcriptomics reveals that AGAMOUS-LIKE22 affects primary metabolism and developmental processes in drought-stressed Arabidopsis. Plant Cell. 2016; 28(2):345–

35. Fang N, Xu R, Huang L, Zhang B, Duan P, Li N, Luo Y, Li Y: SMALL GRAIN 11 controls grain size, grain number and grain yield in rice. Rice (New York, NY). 2016; 9(1):64.

36. Li N, Li Y: Signaling pathways of seed size control in plants. Curr Opin Plant Biol. 2016; 33:23–

37. Xu R, Yu H, Wang J, Duan P, Zhang B, Li J, Li Y, Xu J, Lyu J, Li N et al: A mitogen-activated protein kinase phosphatase influences grain size and weight in rice. Plant J. 2018; 95(6):937–

38. Song XJ, Huang W, Shi M, Zhu MZ, Lin HX: A QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase. Nature genetics. 2007; 39(5):623–

39. Hu J, Wang Y, Fang Y, Zeng L, Xu J, Yu H, Shi Z, Pan J, Zhang D, Kang S et al: A rare allele of GS2 enhances grain size and grain yield in rice. Mol Plant. 2015; 8(10):1455–

40. Tang Y, Bao X, Zhi Y, Wu Q, Guo Y, Yin X, Zeng L, Li J, Zhang J, He W et al: Overexpression of a MYB family gene, OsMYB6, increases drought and salinity stress tolerance in transgenic rice. Front Plant Sci. 2019; 10:168.

Supplementary Files Legend

Additional File 1 Summary details of 63 JcMADS genes encoding MADS-box proteins in physic nut.

Additional File 2 The amino acid sequence of each conserved motif within each protein is shown by
a colored box.

**Additional File 3** Levels of expression of the 63 *JcMADS* genes in the organs tested (root, stem cortex, leaf, and seeds) based on RNA-seq data.

**Additional File 4** Primers used in this study.

Figures
Unrooted phylogenetic tree of the MADS-box family of proteins in physic nut and Arabidopsis. The amino acid sequences were aligned using ClustalW and the phylogenetic tree was constructed using the neighbor-joining method. Bootstrap values were calculated for 1000 replicates.
Figure 2
Conserved motifs within each MADS-box protein in physic nut. Motifs were determined using MEME suite version 5.1.0. Grey lines represent non-conserved sequence, and the relative position of each conserved motif within each protein is shown by a colored box at the bottom.

Figure 3

Chromosomal locations of physic nut JcMADS genes. In total, 62 JcMADS genes were mapped to 11 linkage groups (LGs). The chromosome number is indicated at the top of each chromosome. The scale is in centiMorgans (cM). T, tandem duplication.
Patterns of expression of each JcMADS gene in physic nut roots, stem cortex, leaves, and seeds at an early developmental stage (S1) and filling stage (S2), with a colored scale indicating expression levels shown at the bottom.
Levels of expression of the 63 JcMADS genes in physic nut leaves exposed to drought and salinity stresses: log2 ratios of signals from treated versus control leaves are presented as a heat map based on transcriptomic data, with the color scale shown at the bottom. NA: not available.
The product of the JcMADS05 gene is localized in the nucleus. Scale bar, 10 μM.

Transcriptional activity analysis of JcMADS05 in Arabidopsis protoplasts. (A) Construction of reporter and effector vectors. (B) A dual-luciferase assay suggested that JcMADS05 has transcriptional activity. Each experiment included three biological replicates, each with two technical replicates (means of n = 6±SD; asterisks above the bars indicate significant differences from controls at p < 0.01).
Characterization of JcMADS05 transgenic plants (OE1, OE2 and OE3) and their growth phenotypes. (A) Growth phenotype of two-week-old wild-type and JcMADS05 transgenic plants under normal growth conditions. (B) Flower structure in wild-type and transgenic plants. (C) Levels of JcMADS05 transcript in wild-type and transgenic lines. (D) Root length in two-week-old wild-type and transgenic plants. (E) Shoot length in two-week-old wild-type and transgenic plants. Data presented in (D) and (E) are the means of n = 30 ± SD from three independent experiments.
JcMADS05 regulates grain size. (A) Grains from wild-type and JcMADS05 transgenic lines. Scale bar, 1 cm. (B) Grain length in wild-type and JcMADS05 transgenic lines. (C) Grain width in wild-type and JcMADS05 transgenic lines. (D) 1000-grain weight. (E) Yield per plant. All phenotypic data in B-E were measured on plants grown at a spacing of 16×20 spacing in paddies under normal cultivation conditions. Values represent means of n = 30±SD (Duncan test: **, P < 0.01). (F) Expression levels of seed-size-lated genes. Each experiment included three biological replicates, each with two technical replicates (means of n = 6±SD; asterisks above the bars indicate significant differences from controls at p < 0.01).

Supplementary Files
This is a list of supplementary files associated with this preprint. Click to download.
Additional File 2.tif
Additional File 3.xlsx
Additional File 4.xlsx
