Identification and Characterization of the OsCR4 Extracellular Domain-Interacting Proteins OsCIP1 and OsCIP2 in Rice

Lin-Lin Yan
Hebei Normal University
Cui-Xia Pu
Hebei Normal University
Ying Sun (yingsun@hebtu.edu.cn)
Hebei Normal University https://orcid.org/0000-0002-1444-5057

Research Article

Keywords: OsCR4, extracellular domain, yeast two-hybrid, OsCIP1, OsCIP2, rice

Posted Date: March 16th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-284840/v1

Abstract

The receptor-like kinase OsCR4 plays an important role in vegetative and reproductive growth in rice; it controls embryo morphogenesis, leaf development, and interlocking of the palea and lemma. To identify proteins capable of interacting with the OsCR4 extracellular domain (OsCR4E), we performed a yeast two-hybrid assay and obtained two candidate proteins, OsCIP1 and OsCIP2. Both proteins are cysteine-rich and harbor an N-terminal signal peptide. Localization studies showed OsCIP1-GFP accumulation at the cell surface and OsCIP2-GFP accumulation in cytoplasmic vesicles. Immunoblotting revealed the presence of full-length and truncated OsCIP1-GFP fusion proteins in tobacco leaves and rice roots, and Q62 was identified as the key site for protein cleavage. *OsCIP1* was mainly expressed in vascular bundles and the interlocking tissues of the palea and lemma, while *OsCIP2* was mainly expressed in mature seeds. Compared to wild type, *oscip1* mutant plants exhibited a short seminal root. A phylogenetic tree analysis showed that the homologs of OsCIP1 we identified all belong to the family Gramineae. Our results suggest that OsCIP1 interacts with the extracellular domain of OsCR4.

Key Message

Two cysteine-containing proteins, OsCIP1 and OsCIP2, were identified through a screen for proteins capable of interacting with the extracellular domain of OsCR4.

Introduction

Cell–cell communication is crucial for maintaining plant growth and development. Like many other multicellular organisms, flowering plants arise from a zygote and gradually develop into a highly organized individual with diverse cell types and a complex tissue structure. This is achieved through cell proliferation, cell differentiation, and apoptosis.

Plant cells communicate with neighboring cells either via the symplast pathway (delivering signals through plasmodesmata) or the apoplast pathway. Like animal cells, plant cells can produce and secrete polypeptide hormones, which travel short distances through the apoplast (extracellular space) in a non-cell autonomous way (Matsubayashi, 2014). These polypeptides are classified into two groups according to their size and features. The first includes small peptides containing 5–20 amino acids that are usually post-translationally modified (e.g., glycosylation, proline hydroxylation, and tyrosine sulfation). The second includes cysteine-rich peptides (CRPs) of about 50 amino acids that contain an even number of cysteine residues at the C-terminus for disulfide bond formation (Matsubayashi, 2014; Grienenberger and Fletcher, 2015). Mature polypeptide hormones are produced through proteolytic processing from large precursor proteins. These preproproteins possess a signal peptide at the N-terminus that targets them to the endoplasmic reticulum for secretion.

Over the last two decades, several polypeptide hormones that are vital for plant growth and development have been reported together with their receptors. CLV3, a well-known polypeptide hormone in *Arabidopsis*,

maintains stem cell homeostasis in the shoot apical meristem (SAM) by interacting with the receptor-like kinase (RLK) CLV1 (Mayer et al., 1998; Brand et al., 2000; Schoof et al., 2000). CLV3 is a 13-amino acid glycopeptide produced from a 96-amino acid preprotein that is O-arabinosylated at hydroxyproline. The bioactive polypeptide RGF1 is sulfated by a tyrosylprotein sulfotransferase then secreted into the apoplast where it regulates the activity of the root meristem via interactions with the RLK RGI1 and recruits a coreceptor PLETHORA to relay signals (Ou et al., 2016; Song et al., 2016). In rice, the polypeptide hormones FLORAL ORGAN NUMBER (FON)2/FON4 and FON1 (homologs of CLV3 and CLV1, respectively) are essential to maintain homeostasis in inflorescence and floral meristems at the reproductive stage (Nagasawa et al., 1996; Suzaki et al., 2006), while the polypeptides FON2-LIKE CLE PROTEIN (FCP)1 and FCP2 redundantly maintain homeostasis in the SAM and root apical meristem through interactions with unknown receptors during vegetative growth (Kinoshita et al., 2007; Ohmori et al., 2013, 2014). STOMAGEN and rapid alkalinization factor are CRPs that function as positive regulators of stomatal development and in root elongation, respectively (Pearce et al., 2001; Kondo et al., 2010).

Bioinformatic analysis has revealed more than 30,000 previously unannotated putative peptide-encoding sequences (Lease and Walker, 2006), compared to 600 RLK-encoding genes, in the *Arabidopsis thaliana* genome (Shiu and Bleecker, 2001), and it has been proposed that different ligands bind to the same receptor in different developmental contexts to initiate distinct downstream signaling pathways, reflecting the diversity and complexity of polypeptide signaling molecules. To date, most of these have not been functionally characterized, probably due to functional redundancy, and their receptors are largely unknown.

Crinkly4 (CR4) belongs to the Tumor Necrosis Factor (TNF) Receptor (TNFR) subfamily of RLKs. This small subfamily is characterized by a cysteine-rich TNFR domain in the extracellular region of its members. *ArabidopsisACR4*, rice *OsCR4*, and maize *CR4* are orthologs that function in epidermal cell differentiation in many organs, including leaves, roots, flowers, and seeds (Becraft et al., 1996; Jin et al., 2000; Tanaka et al., 2002; Watanabe et al., 2004; Wang et al., 2020). The polypeptide CLE40 is considered a putative ligand of ACR4 that functions in root meristem cell niche maintenance (Stahl and Simon, 2009); however, whether this ligand–receptor pair functions in epidermal cell specification is unknown. Furthermore, OsCR4 is crucial for the interlocking of the palea and lemma (Pu et al., 2012), which are both absent from *Arabidopsis*. Therefore, undiscovered ligands of OsCR4 must exist.

To search for ligands of OsCR4 during reproductive development, we screened for OsCR4 extracellular domain (OsCR4E)-interacting proteins using a cDNA library produced from a mixture of reproductive tissues using the yeast two-hybrid method. We identified two small proteins, OsCIP1 and OsCIP2, and preliminarily characterized their subcellular localization, tissue expression patterns, protein cleavage site, and biological functions. Our results provide valuable clues to understanding the regulatory mechanisms of OsCR4.

Materials And Methods

Plant growth conditions

Wild-type rice {*Oryza sativa Japonica* variety *Nipponbare* or *Hwayoung* (HW)} and transgenic rice plants were grown in a greenhouse at 28℃ under a 16-h light/8-h dark cycle or in a paddy field under natural conditions from May to October every year. *Nicotiana benthamiana* plants were grown in a growth room at 22℃ under a 16-h light (80 μmol m^{-2} s^{-1} white light)/8-h dark cycle.

Yeast two-hybrid screening

According to the user manual for "BD Matchmaker™ Library Construction & Screening Kits" (Clontech), young panicles less than 1 cm in length, anthers, unpollinated pistils, and developing seeds at 1 week after pollination were collected. The mRNA extracted from these tissues was equally and evenly blended before reverse transcription into cDNA with the specific primers in the kit. The freshly made cDNA was cloned into linear pGADT7-Rec by recombination to produce the activation domain (AD)-cDNA library, which was directly transformed into yeast strain *AH109* containing *OsCR4E-BD*, and the transformants were screened on SD/-His-Ade-Leu-Trp dropout medium. Positive transformants were screened three times and amplified by colony PCR with the primer pair *5'AD LD-insert Screening Amplimer/3'AD LD-insert Screening Amplimer*. The PCR products were separated in an agarose gel and purified before sequencing with *T7* primer. Finally, genes fused in-frame with the GAL4 AD were identified through a BLAST search against the rice nucleotide sequence database KOME (*http://cdna01.dna.affrc.go.jp/cDNA/*) and NCBI database (*https://www.ncbi.nlm.nih.gov/*).

Yeast two-hybrid assay

To construct the binding domain (*BD*) vectors, fragments of the OsCR4E lacking the signal peptide, TNFR domain, or seven repetitive b-sheets (7-Repeats) domain were, respectively, amplified using the primer pairs *OsCR4E-BD-F/OsCR4E-BD-R*, *OsCR4E-BD-F/OsCR4E-7-Repeats-R*, and *OsCR4E-TNFR-F/OsCR4E-BD-R*, and then inserted into *pGBKT7* using *Eco*RI and *Sma*I. To construct the *AD* vectors, fragments of *OsCIP1* and *OsCIP2* lacking the signal peptide were, respectively, amplified using the primer pairs *OsCIP1-AD-F/OsCIP1-AD-R* and *OsCIP2-AD-F/OsCIP2-AD-R*, and then cloned into *pGADT7* using *Eco*RI/*Bam*HI and *Eco*RI/*Sma*I. For the yeast two-hybrid assay, one *BD* vector and one *AD* vector (indicated in Fig. 1a and b) were co-transformed into yeast strain *AH109* and positive transformants were screened on SD/-Leu-Trp dropout medium. Three well-grown clones were collected and mixed evenly before gradient dilution, and then simultaneously dotted on both SD/-His-Ade-Leu-Trp and SD/-Leu-Trp dropout media. The yeast dot grown on SD/-Leu-Trp medium harbored the *AD* and *BD* vectors; growth on SD/-His-Ade-Leu-Trp medium indicated an interaction between the AD and BD proteins.

Overlay assay

OsCR4E (without the signal peptide) was amplified using the primer pair *GST-CR4E-F*/ *GST-CR4E-R* and introduced into *pGEX-4T-1* using *Eco*RI and *Sma*I. *OsCIP1* and *OsCIP2* were cut from *OsCIP1-AD* and *OsCIP2-AD*, and then cloned into *pET-32a* using the same cloning site. After sequencing, the vectors were

introduced into Rosetta (*DE3*), a *BL21* derivative for the induction and expression of fusion proteins and tagged proteins. Next, 1 μg of purified GST-OsCR4E or GST was separated by SDS-PAGE and then transferred to a nitrocellulose membrane. After blocking with 5% nonfat milk, the membranes were incubated with 100 μmol/L of TRX-6His-OsCIP1, 110 μmol/L of TRX-6His-OsCIP2, and 112 μmol/L of TRX-6His at 4℃ overnight, and then detected with horseradish peroxidase (HRP)-conjugated anti-His antibody (CoWin Biosciences) after washing. The signal was developed using HRP substrate.

Subcellular localization of OsCIP1, OsCIP2, and OsCIP1 variants

To construct *OsCIP1-GFP* and *OsCIP2-GFP*, the full-length coding sequences of *OsCIP1* and *OsCIP2*, amplified from leaf cDNA with the primer pairs *OsCIP1FL-F/OsCIP1FL-R* and *OsCIP2FL-F/OsCIP2FL-R*, were introduced into the binary vector *pMDC83* using *Spe*I and *Kpn*I under the control of the *2X35S* promoter. To construct *OsCIP1Variants-GFP* (with amino acid deletions or a mutation in the middle), variants were amplified via the bridging method. Taking Os*CIP1*$^{\triangle 56\text{-}65}$ as an example, two primers, *CIP1△(56-65)-a-F* and *CIP1△(56-65)-b-R*, were designed to span the deletion at residues 56-65 with a 20-bp overlap. The N- and C-halves of *OsCIP1* were separately amplified using the primer pairs *OsCIP1FL-F/CIP1△(56-65)-b-R* and *CIP1△(56-65)-a-F/OsCIP1FL-R*, and the resulting fragments were overlapped in the middle and used as a template to amplify Os*CIP1*$^{\triangle 56\text{-}65}$ with primer pair *OsCIP1FL-F/OsCIP1FL-R*. Similarly, *CIP1*$^{\triangle 56\text{-}60}$, *CIP1*$^{\triangle 59\text{-}63}$, *CIP1*$^{\triangle 61\text{-}65}$, *CIP1*^{W61LQ62A}, and *CIP1*^{Q62A} were produced, respectively, using primers *CIP1△(56-60)-a-F* and *CIP1△(56-60)-b-R*, *CIP1△(59-63)-a-F* and *CIP1△(59-63)-b-R*, *CIP1△(61-65)-a-F* and *CIP1△(61-65)-b-R*, *CIP16162-a-F* and *CIP16162-b-R*, and *CIP162-a-F* and *CIP162-b-R*, together with primer pair *OsCIP1FL-F/OsCIP1FL-R*. Next, the above *OsCIP1* variants were introduced into the binary vector *pMDC83* in the same manner as *OsCIP1*. For the primer sequences, see Supplemental Table 1. All vectors were sequenced before *AgrobacteriumEHA105*-mediated transformation by injecting tobacco leaves or incubating rice calli to produce transgenic rice plants (Yang et al., 2004). The subcellular localization of OsCIP1-GFP, OsCIP2-GFP, and OsCIP1 variants-GFP in tobacco leaf cells and rice root cells was observed using a laser scanning confocal microscope (LSCM510; Zeiss) with excitation at 488 nm and 500-540 nm emission.

Production of *oscip1-cas* mutants

To construct the *OsCIP1-Cas9* vector, a gene-specific spacer sequence terminated with NGG was synthesized and cloned into the entry vector *pOs-SgRNA* using *Bsa*I, and then introduced into the binary vector *pH-Ubi-Cas9-7* via LR recombination (Miao et al., 2013). The resulting *OsCIP1-Cas9* vector was sequenced before transformation into rice calli (Yang et al., 2004). Gene-edited *oscip1-cas* mutants were identified by the sequencing of PCR products spanning the spacer site, and then backcrossed to HW twice to eliminate the Cas9 editor.

Phylogenetic tree construction

A BlastP search of the NCBI database, using the full-length sequence of OsCIP1 as the entry, was used to identify homologs of OsCIP1 from different species. The first 50 homologs with relatively strong

similarity to OsCIP1 were chosen to build a phylogenetic tree by the maximum likelihood method based on the JTT matrix-based mode in Mega X software. The bootstrap consensus tree inferred from 1,000 replicates was taken to represent the evolutionary history of the taxa analyzed.

Results

Yeast two-hybrid screening for OsCR4E-interacting proteins

Using the OsCR4E (amino acids 23-418, without the signal peptide) as bait, we screened a cDNA library constructed from the mRNA of rice floral organs and young seeds for interacting proteins and found 19 genes with the correct fusion in-frame. Of these, eight genes encoded proteins that harbored a signal peptide at the N-terminus, and two of them, *LOC_Os03g25350.1* (AK111061) and *LOC_Os11g33000.1* (AK242313), encoded small proteins with around 100 amino acids (Table 1). These two small proteins were named OsCR4 Interacting Protein 1 (OsCIP1) and OsCIP2, respectively. OsCIP1, which contains 114 amino acid residues with a 22-amino acid signal peptide, belongs to the protease inhibitor/seed storage/lipid transfer protein (LTP) family (putative). OsCIP2, which consists of 135 amino acids with 24-amino acid signal peptide, is predicted to be a member of the albumin seed storage protein family (Table 1). The biological functions of these two proteins are unknown.

OsCIP1 and OsCIP2 interact specifically with the OsCR4E

OsCR4 contains two subdomains: 7-Repeats and TNFR. To verify the interaction of OsCIP1 or OsCIP2 with the OsCR4E, we first performed a yeast two-hybrid assay using the OsCR4E, 7-Repeats domain, or TNFR domain as bait and OsCIP1 or OsCIP2 (without the signal peptide) as prey. Our results show that OsCIP1 interacted with the OsCR4E and TNFR domain, but not with the 7-Repeats domain (Fig. 1a). However, OsCIP2 interacted with the OsCR4E, 7-Repeats domain, and TNFR domain (Fig. 1b). There was no interaction with the empty AD or BD (negative control) (Fig. 1a and b), suggesting that the interaction of OsCIP1 or OsCIP2 with the OsCR4E in yeast was specific.

Next, we tested the interaction of OsCIP1 or OsCIP2 with the OsCR4E in an overlay assay. OsCIP1 interacted with the OsCR4E, consistent with our yeast two-hybrid results; however, in contrast to our results, OsCIP2 did not interact with the OsCR4E (Fig. 1c). This discrepancy may be due to a difference in protein conformation between bacteria-expressed OsCIP2 and yeast-expressed OsCIP2.

Subcellular localization and tissue expression patterns of OsCIP1 and OsCIP2

When tobacco epidermal cells transiently expressing and rice root cells constitutively expressing *35S:OsCIP1-GFP* and *35S:OsCIP2-GFP* were visualized by confocal microscopy, OsCIP1-GFP was mainly located at the cell surface (Fig. 2a and c), while OsCIP2-GFP was located at the cell surface and in cytoplasmic vesicles (Fig. 2b and d). Furthermore, in tobacco epidermal cells the OsCIP1-GFP signal merged completely with the propidium iodide (PI) signal (Fig. 2e), which remained outside of the plasma membrane of living cells, indicating that OsCIP1 is a secreted protein. Interestingly, three bands were

detected in either tobacco leaves or rice roots expressing OsCIP1-GFP, but not OsCIP2-GFP, by immunoblotting (Fig. 2f). According to their apparent molecular weights, we inferred that the top band corresponded to the full-length fusion protein, the bottom band corresponded to the GFP tag, and the band below the top one probably corresponded to cleaved OsCIP1 with a C-terminal GFP fusion. These results suggest that two forms of OsCIP1 were located extracellularly: the full-length protein and the cleaved protein. It is unclear, however, which form is biologically active.

β-Glucuronidase (GUS) staining showed that *OsCIP1* was expressed in the hook of the palea and lemma starting from the seventh stage of anther development (An7). During spikelet development, *OsCIP1* expression increased gradually and became more intense at the upper end of the hook of the palea and lemma (Fig. 3a1–5). In addition, *OsCIP1* was expressed in the dorsal large vascular bundles of mature seeds (Fig. 3b1–2), the central vascular bundles of roots, the coleoptile of seedlings at 2 days after germination (DAG) (Fig. 3c), and in the seminal root elongation zone (Fig. 3d) and mature zone (Fig. 3e) in seedlings at 3 DAG. In other words, *OsCIP1* was mainly expressed in the vascular tissues of various organs during vegetative growth and in the hook of the spikelet during reproductive growth. Whereas *OsCIP2* was not expressed in the palea and lemma of spikelets, it was expressed in mature seeds (Fig. 3f), likely reflecting its function as a seed storage protein.

Q62 is the key site in OsCIP1 cleavage

The above results indicated post-translational cleavage of OsCIP1, consistent with the processing property of small peptides. According to the molecular weight of the cleaved protein band, we inferred that cleavage occurred among the 56^{th} to 65^{th} amino acid residues (Fig. 4a, underlined). To find the cleavage site, an OsCIP1 variant expression vector, *35S:OsCIP1*$^{\Delta 56\text{-}65}$*-GFP* (10 amino acid deletion), was constructed (Fig. 4b) and the subcellular localization of OsCIP1$^{\Delta 56\text{-}65}$-GFP was observed in tobacco epidermal cells. Fluorescence was mainly detected at the cell surface and in dotted structures near the plasma membrane (Fig. 4c1), and immunoblotting revealed one band corresponding to OsCIP1$^{\Delta 56\text{-}65}$-GFP (Fig. 4d), indicating no cleavage. Next, two OsCIP1 variant expression vectors, *35S:OsCIP1*$^{\Delta 56\text{-}60}$*-GFP* and *35S:OsCIP1*$^{\Delta 61\text{-}65}$*-GFP*, were constructed (5 amino acid deletion) (Fig. 4b). Fluorescent signals from the fusion proteins were found not only on the cell surface, but also in a large number of cytoplasmic vesicles in tobacco leaf cells (Fig. 4c2–3), and neither of the variants was cleaved based on immunoblotting (Fig. 4d). Finally, the OsCIP1 variant expression vector *35S:OsCIP1*$^{\Delta 59\text{-}63}$*-GFP* lacking amino acids 59–63 was constructed (Fig. 4b). The subcellular localization of OsCIP1$^{\Delta 59\text{-}63}$-GFP in tobacco leaf cells was similar to that of OsCIP1$^{\Delta 56\text{-}60}$-GFP, and many more signals were detected in small vesicles (Fig. 4c4). On immunoblotting, a higher molecular weight band was observed for OsCIP1$^{\Delta 59\text{-}63}$-GFP (Fig. 4d). Thus, we inferred that the cleavage site was within these five amino acids. We subsequently constructed the expression vectors *35S:OsCIP1*^{W61LQ62A}*-GFP* and *35S:OsCIP1*^{Q62A}*-GFP* (Fig. 4b) and found that fluorescent signals corresponding to these mutants appeared only on the cell surface, similar to those of OsCIP1-GFP (Fig. 4c5-7). OsCIP1^{W61LQ62A}-GFP and OsCIP1^{Q62A}-GFP were not cleaved according to immunoblotting analyses (Fig. 4d), suggesting that Q62 was the cleavage site (either before

or after this amino acid; Fig. 4a and d). These results also suggest that the deletion of amino acids 56-65 or other residues in this region would affect the secretion of OsCIP1 to the cell surface due to incorrect protein folding.

Seminal root growth in *oscip1* was much slower than that in wild type

To explore the biological function of OsCIP1, we produced two mutant alleles, *oscip1-cas1* and *oscip1-cas2*, using the CRISPR/Cas9 gene editing system. One T50 nucleotide deletion in *oscip1-cas1* and the deletion of G48G49 in *oscip1-cas2* (Fig. 5a) resulted in a reading frame shift, leading to early termination of translation. In theory, *oscip1-cas1* could be translated to produce a small protein with 34 amino acids, including 18 mutated amino acids at the C-terminal end, while *oscip1-cas2* could be translated to produce a protein with 23 amino acids, including only 15 correct amino acids at the N-terminus. Therefore, both *oscip1-cas1* and *oscip1-cas2* should be loss-of-function mutants. PCR identification showed that the *Cas9* gene editor was present in *oscip1-cas1* but not in *oscip1-cas2* (Fig. 5b). The seminal root of *oscip1* was much shorter than that of wild type at 3 DAG (Fig. 5c), indicating that OsCIP1 plays a role in seminal root elongation.

Bioinformatic analysis of OsCIP1 and OsCIP2

To further understand the protein structure and function of OsCIP1, we performed a bioinformatic analysis of the OsCIP1 amino acid sequence using the SMART database (http://smart.embl-heidelberg.de/). We found that in addition to the 22- and 24-amino acid signal peptides at the N-termini of OsCIP1 and OsCIP2, respectively, each sequence contained a 75- and 74-amino acid trypsin-alpha amylase inhibitor (AAI) domain (Fig. 6a and b). An amino acid sequence alignment of mature (without the signal peptide) OsCIP1 and OsCIP2 with three known LTPs (OsLTP1, OsLTP2, and OsC6) revealed that OsCIP1 shared 17% identity with these three proteins. It also contained a conserved eight-cysteine skeleton (C1-Xn-C2-Xn-C3C4-Xn-C5XC6-Xn-C7-Xn-C8) that probably forms four disulfide bonds, like in LTPs (Fig. 6c). Further, several amino acids with similar properties to those in LTPs were found. In contrast, OsCIP2 shared 12% identity with the three LTPs and had only five conserved cysteines (Fig. 6c and d).

OsCIP1 interacts with the TNFR domain in the OsCR4E, which is believed to bind ligands in a manner similar to that of TNFR. We thus compared the mature protein sequence of OsCIP1 with the mature protein sequence of mammalian TNF (Pennica et al., 1984). We found 9.55% sequence identity and 15 conserved amino acids, including conserved cysteines at positions 5 and 7 (Fig. 6e).

In a BlastP search of the NCBI protein database using the full-length sequence of OsCIP1 as the entry, we selected the first 50 protein homologs for phylogenetic tree analysis, and we found that almost all of them belonged to the monocotyledonous family Gramineae (Fig. 6f). The most closely related protein to OsCIP1 was a homolog of unknown function from *Oryza brachyantha*. These data suggest that OsCIP1 is involved in specific biological processes of Gramineae species, such as spikelet development.

Discussion

Usually, the maturation of polypeptide hormones involves the following processing steps. First, the signal peptide is removed from the prepropeptide by signal peptidase in the endoplasmic reticulum. Second, the resulting propeptide is modified and cut into active mature peptide hormones by specific endopeptidases. Third, the hormones are folded correctly and packaged into secretory vesicles, which are secreted from the cell so that they can activate signaling pathways after binding to surface receptors on their target cells (Matsubayashi, 2014). In this study, OsCIP1 was able to bind specifically to the extracellular domain of OsCR4 in yeast two-hybrid and overlay assays, suggesting that OsCIP1 acts as an extracellular signaling molecule for OsCR4. OsCIP1 contains a conserved eight-cysteine skeleton (C1-Xn-C2-Xn-C3C4-Xn-C5XC6-Xn-C7-Xn-C8; Fig. 6c) that allows for the formation of four disulfide bonds, which stabilize the non-specific LTP (nsLTP)-like tertiary structure (Boutrot et al., 2008). However, OsCIP1 does not have lipid transfer activity *in vitro* (Fig. S1), indicating that it cannot function as an LTP. On the other hand, OsCIP1 interacted only with the TNFR domain of OsCR4, and its sequence was similar to that of mammalian TNF (Fig. 6e). Further, the only two cysteines in TNF, which should form a disulfide bond, aligned to the fifth and seventh cysteines in OsCIP1. A disulfide bond cannot be formed between these two cysteines based on the cysteine skeleton in known LTPs. Nevertheless, OsCIP1 might be cleaved between the W and Q residues, which would destroy the original disulfide bond and allow the formation of a new disulfide bond between C5 and C7 in the resulting C-terminal half (and its tertiary structure would be altered accordingly). Given that full-length or the C-terminal portion of OsCIP1 was located outside the plasma membrane (Fig. 2a), we presume that OsCIP1 is a small extracellular signaling molecule; however, further study is needed to assess whether the full-length protein or the cleaved protein is the functional form.

CR4 family members regulate epidermal cell differentiation and promote cuticle formation (Czyzewicz et al., 2016). ALE1, a secretory serine protease, is also necessary for epidermal cell differentiation and cuticle formation (Tanaka et al., 2001), similar to ACR4 in *Arabidopsis* (Watanabe et al., 2004). Thus, ALE1 might produce active signal molecules by cleaving extracellular propeptides, and then cooperate with the RLK ACR4 to regulate leaf epidermal cell specification. Although the homolog of ALE1 has not been identified in rice, and despite the fact that its substrates are unknown, the characterization of OsCIP1 as a cleaved protein and interacting partner of the OsCR4E provides tantalizing clues regarding the significance of these molecular interactions.

The biological function of OsCIP1 might be related to spikelet development. *OsCIP1* was expressed in the hook of the palea and lemma from an early stage of spikelet development, and the expression level increased gradually with further spikelet development (Fig. 3a1–5). This pattern overlaps with the tissue expression pattern of *OsCR4* (Pu et al., 2012), suggesting that the interaction between OsCIP1 and OsCR4 is involved in the development of the palea and lemma, and that it might function in the interlocking of the palea and lemma by promoting epidermal cell specification. However, the *oscip1* mutant did not show the open-hull phenotype displayed by *OsCR4* RNAi plants due to an abnormal epidermis cell morphology in the palea and lemma (Pu et al., 2012). This may be due to the presence of OsCIP1 homologs in rice. Indeed, a homolog of OsCIP1 exists in rice according to our phylogenetic tree analysis, and almost all of

the homologs belong to the family Gramineae (Fig. 6f), suggesting that OsCIP1 is involved in Gramineae-specific biological processes. In addition, *OsCIP1* was strongly expressed in large vascular bundles on the dorsal side of seeds and in vascular bundles of the coleoptile and seminal root (Fig. 3b-e). The dorsal vascular bundle is responsible for nutrient transport to seeds, implying that OsCIP1's function is related to vascular tissue development or nutrient transport. The shorter seminal root in the *oscip1* mutant (Fig. 5c and d) is consistent with the phenotype of the *oscr4* mutant (Wang et al., 2020) and similar to the short main root observed in the *acr4* mutant (De Smet et al., 2008). ACR4 regulates *Arabidopsis* root meristem cell activity by receiving the peptide ligand CLE40 (Stahl and Simon, 2009). Whether OsCIP1 regulates cell division in rice roots (like CLE40) or whether it regulates cell elongation should be clarified.

OsCIP2 does not belong to the nsLTP family because there is no conserved eight-cysteine skeleton in its protein sequence (Fig. 6c). OsCIP2 interacted with the full-length extracellular domain and 7-Repeats domain of OsCR4 in yeast cells. The 7-Repeats domain is involved in the internalization of ACR4 (Gifford et al., 2005). The location of most of the OsCIP2 in cytoplasmic vesicles was similar to that of ACR4 (Gifford et al., 2005) and OsCR4 (Fig. S2). The expression of *OsCIP2* in mature seeds (Fig. 3g) is consistent with the *OsCR4* expression pattern (Pu et al., 2012). Therefore, we speculate that OsCIP2 is involved in the differentiation of aleurone layer cells by regulating OsCR4 internalization and vesicle traffic.

In maize, CR4 cooperates with the large transmembrane protein DEK1 and a vacuole sorting protein, SAL1, to regulate aleurone cell fate determination (Tian et al., 2007). In this process, two signaling molecules, DEK1 and CR4, are maintained at a certain concentration on the plasma membrane through SAL1-mediated internalization and are consequently recycled or degraded (Tian et al., 2007). It would be interesting to investigate whether OsCIP2 is involved in this process.

In *Arabidopsis*, ALE2, another RLK with a cysteine-containing sequence in its extracellular domain, functions in epidermal cell specification (Tanaka et al., 2007), similar to ACR4. The presence of multiple membrane signaling molecules, whether it be in the monocot maize or in the dicot *Arabidopsis*, means that diverse intercellular communications are required for epidermal cell fate determination. OsCIP1 and OsCIP2, as cysteine-rich proteins that interact with the extracellular domain of OsCR4, are worth studying to determine their roles in cell–cell communication.

Abbreviations

RLK	Receptor-like kinase	GUS	β-Glucuronidase
7-Repeats	Seven repetitive β-sheets domain	LTP	Lipid transfer protein
aa	Amino acid	OsCIP1	OsCR4 interacting protein 1
AAI	Trypsin-alpha amylase inhibitor	OsCIP2	OsCR4 interacting protein 2
AD	Activation domain	OsCR4E	OsCR4 extracellular domain
BD	Binding domain	TNF	Tumor necrosis factor
CRP	Cysteine-rich peptides	TNFR	Tumor necrosis factor receptor
DAG	Days after germination		

Declarations

Funding This study was supported by the National Natural Science Foundation of China (30570153), Hebei Provincial Education Department Graduate Innovation Funding Project (CXZZBS2018098), and China Postdoctoral Science Foundation (2012M520594).

Conflicts of interest/Competing interests The authors have no conflicts of interest to declare that are relevant to the content of this article.

Availability of data and materials All data generated or analyzed during this study are included in this published article and its supplementary information files.

Code availability Not applicable for this section

Author contribution statement

All authors contributed to the study conception and design. Material preparation, data collection, and analysis were performed by LLY and CXP. The first draft of the manuscript was written by LLY and CXP, and revised by YS. All authors commented on the previous versions of the manuscript. All authors have read and approved the final version of the manuscript.

Ethics approval Not applicable for this section

Consent for publication Not applicable for this section

Acknowledgements

We thank Dr. Li-Jia Qu (Peking University) for kindly providing the plasmids *pOs-SgRNA* and *pH-Ubi-Cas9-7*, Dr. Cui-Feng Li (Nankai University) for providing the lipid transfer protein CaMBP-10 and assisting with

the lipid-binding assay, and Liang Lu (Hebei Normal University) for providing advice on the phylogenetic tree analysis.

References

Becraft PW, Stinard PS, **and** McCarty DR (1996) CRINKLY4: A TNFR-like receptor kinase involved in maize epidermal differentiation. Science 273:1406–1409. https://doi.org/10.1126/science.273.5280.1406

Boutrot F, Chantret N, **and** Gautier MF (2008) Genome-wide analysis of the rice and *Arabidopsis* non-specific lipid transfer protein (nsLtp) gene families and identification of wheat nsLtp genes by EST data mining. BMC Genom 9:86. https://doi.org/10.1186/1471-2164-9-86

Brand U, Fletcher JC, Hobe M, Meyerowitz EM, **and** Simon R (2000) Dependence of stem cell fate in *Arabidopsis* on a feedback loop regulated by CLV3 activity. Science 289:617–619. https://doi.org/10.1126/science.289.5479.617

Czyzewicz N, Nikonorova N, Meyer MR, Sandal P, Shah S, Vu LD, Gevaert K, Rao AG, **and De** Smet, **I**. (2016). The growing story of (ARABIDOPSIS) CRINKLY 4. Journal of Experimental Botany **67**, 4835–4847. https://doi.org/10.1093/jxb/erw192

De Smet I, Vassileva V, **De** Rybel B, Levesque MP, Grunewald W, **Van** Damme D, **Van** Noorden G, Naudts M, **Van** Isterdael G, **De** Clercq R, Wang JY, Meuli N, Vanneste S, Friml J, Hilson P, Jurgens G, Ingram GC, Inze D, Benfey PN, **and** Beeckman T (2008). Receptor-like kinase ACR4 restricts formative cell divisions in the *Arabidopsis* root. Science 322, 594–597. https://doi.org/10.1126/science.1160158

Gifford ML, Robertson FC, Soares DC, **and** Ingram GC (2005) ARABIDOPSIS CRINKLY4 function, internalization, and turnover are dependent on the extracellular crinkly repeat domain. Plant Cell 17:1154–1166. https://doi.org/10.1105/tpc.104.029975

Grienenberger E, **and** Fletcher JC (2015) Polypeptide signaling molecules in plant development. Curr Opin Plant Biol 23:8–14. https://doi.org/10.1016/j.pbi.2014.09.013

Jin, P., Guo, T., **and** Becraft, P.W. (2000). The maize CR4 receptor-like kinase mediates a growth factor-like differentiation response. Genesis **27**, 104–116. https://doi.org/10.1002/1526-968x(200007)27:3<104::aid-gene30>3.0.co;2-i

Kinoshita A, Nakamura Y, Sasaki E, Kyozuka J, Fukuda H, **and** Sawa S (2007) Gain-of-function phenotypes of chemically synthetic CLAVATA3/ESR-related (CLE) peptides in *Arabidopsis thaliana* and *Oryza sativa*. Plant & Cell Physiology **48**, 1821–1825. https://doi.org/10.1093/pcp/pcm154

Kondo T, Kajita R, Miyazaki A, Hokoyama M, Nakamura-Miura T, Mizuno S, Masuda Y, Irie K, Tanaka Y, Takada S, Kakimoto T, **and** Sakagami Y (2010) Stomatal density is controlled by a mesophyll-derived signaling molecule. Plant Cell Physiol 51:1–8. https://doi.org/10.1093/pcp/pcp180

Lease KA, **and** Walker JC (2006) The *Arabidopsis* unannotated secreted peptide database, a resource for plant peptidomics. Plant Physiol 142:831–838. https://doi.org/10.1104/pp.106.086041

Matsubayashi Y (2014) Posttranslationally modified small-peptide signals in plants. Annu Rev Plant Biol 65:385–413. https://doi.org/10.1146/annurev-arplant-050312-120122

Mayer KF, Schoof H, Haecker A, Lenhard M, Jurgens G, **and** Laux T (1998) Role of WUSCHEL in regulating stem cell fate in the *Arabidopsis* shoot meristem. Cell 95:805–815. https://doi.org/10.1146/annurev-arplant-050312-120122

Miao J, Guo D, Zhang J, Huang Q, Qin G, Zhang X, Wan J, Gu H, **and** Qu LJ (2013) Targeted mutagenesis in rice using CRISPR-Cas system. Cell Res 23:1233–1236. https://doi.org/10.1038/cr.2013.123

Nagasawa N, Miyoshi M, Kitano H, Satoh H, **and** Nagato Y (1996) Mutations associated with floral organ number in rice. Planta 198:627–633. https://doi.org/10.1007/BF00262651

Ohmori Y, Yasui Y, **and** Hirano HY (2014) Overexpression analysis suggests that FON2-LIKE CLE PROTEIN1 is involved in rice leaf development. Genes Genet Syst 89:87–91. https://doi.org/10.1266/ggs.89.87

Ohmori Y, Tanaka W, Kojima M, Sakakibara H, **and** Hirano HY (2013) WUSCHEL-RELATED HOMEOBOX4 is involved in meristem maintenance and is negatively regulated by the CLE gene FCP1 in rice. Plant Cell 25:229–241. https://doi.org/10.1105/tpc.112.103432

Ou Y, Lu X, Zi Q, Xun Q, Zhang J, Wu Y, Shi H, Wei Z, Zhao B, Zhang X, He K, Gou X, Li C, **and** Li J (2016) RGF1 INSENSITIVE 1 to 5, a group of LRR receptor-like kinases, are essential for the perception of root meristem growth factor 1 in *Arabidopsis thaliana*. Cell Res 26:686–698. https://doi.org/10.1038/cr.2016.63

Pearce G, Moura DS, Stratmann J (2001) RALF, a 5-kDa ubiquitous polypeptide in plants, arrests root growth and development. Proc Natl Acad Sci USA 98:12843–12847. https://doi.org/10.1073/pnas.201416998, **and** Ryan, C.A., **Jr**

Pennica D, Nedwin GE, Hayflick JS, Seeburg PH, Derynck R, Palladino MA, Kohr WJ, Aggarwal BB, **and** Goeddel DV (1984) Human tumour necrosis factor: precursor structure, expression and homology to lymphotoxin. Nature 312:724–729. https://doi.org/10.1038/312724a0

Pu CX, Ma Y, Wang J, Zhang YC, Jiao XW, Hu YH, Wang LL, Zhu ZG, Sun D, **and** Sun Y (2012) Crinkly4 receptor-like kinase is required to maintain the interlocking of the palea and lemma, and fertility in rice, by promoting epidermal cell differentiation. The Plant Journal: for Cell Molecular Biology 70:940–953. https://doi.org/10.1111/j.1365-313X.2012.04925.x

Schoof H, Lenhard M, Haecker A, Mayer KF, Jurgens G, **and** Laux T (2000) The stem cell population of *Arabidopsis* shoot meristems in maintained by a regulatory loop between the CLAVATA and WUSCHEL

genes. Cell 100:635–644. https://doi.org/10.1016/s0092-8674(00)80700-x

Shiu SH, **and** Bleecker AB (2001) Plant receptor-like kinase gene family: diversity, function, and signaling. Sci STKE **2001**, re22. https://doi.org/10.1126/stke.2001.113.re22

Song W, Liu L, Wang J, Wu Z, Zhang H, Tang J, Lin G, Wang Y, Wen X, Li W, Han Z, Guo H, **and** Chai J (2016) Signature motif-guided identification of receptors for peptide hormones essential for root meristem growth. Cell Res 26:674–685. https://doi.org/10.1038/cr.2016.62

Stahl Y, **and** Simon R (2009) Is the *Arabidopsis* root niche protected by sequestration of the CLE40 signal by its putative receptor ACR4? Plant Signaling Behavior 4:634–635. https://doi.org/10.1016/j.cub.2009.03.060

Suzaki T, Toriba T, Fujimoto M, Tsutsumi N, Kitano H, **and** Hirano HY (2006) Conservation and diversification of meristem maintenance mechanism in *Oryza sativa*: Function of the FLORAL ORGAN NUMBER2 gene. Plant Cell Physiology 47:1591–1602. https://doi.org/10.1093/pcp/pcl025

Tanaka H, Watanabe M, Watanabe D, Tanaka T, Machida C, **and** Machida Y (2002) ACR4, a putative receptor kinase gene of *Arabidopsis thaliana*, that is expressed in the outer cell layers of embryos and plants, is involved in proper embryogenesis. Plant Cell Physiology 43:419–428. https://doi.org/10.1093/pcp/pcf052

Tanaka H, Onouchi H, Kondo M, Hara-Nishimura I, Nishimura M, Machida C, **and** Machida Y (2001) A subtilisin-like serine protease is required for epidermal surface formation in *Arabidopsis* embryos and juvenile plants. Development 128:4681–4689. https://dev.biologists.org/content/128/23/4681.long

Tanaka H, Watanabe M, Sasabe M, Hiroe T, Tanaka T, Tsukaya H, Ikezaki M, Machida C, **and** Machida Y (2007) Novel receptor-like kinase ALE2 controls shoot development by specifying epidermis in *Arabidopsis*. Development 134:1643–1652. https://doi.org/10.1242/dev.003533

Tian Q, Olsen L, Sun B, Lid SE, Brown RC, Lemmon BE, Fosnes K, Gruis DF, Opsahl-Sorteberg HG, Otegui MS, **and** Olsen OA (2007) Subcellular localization and functional domain studies of DEFECTIVE KERNEL1 in maize and *Arabidopsis* suggest a model for aleurone cell fate specification involving CRINKLY4 and SUPERNUMERARY ALEURONE LAYER1. Plant Cell 19:3127–3145. https://doi.org/10.1105/tpc.106.048868

Wang J, Yan LL, Yue ZL, Li HY, Ji XJ, Pu CX, **and** Sun Y (2020) Receptor-like kinase OsCR4 controls leaf morphogenesis and embryogenesis by fixing the distribution of auxin in rice. Journal of Genetics Genomics = Yi chuan xue bao 47:577–589. https://doi.org/10.1016/j.jgg.2020.08.002

Watanabe M, Tanaka H, Watanabe D, Machida C, **and** Machida Y (2004) The ACR4 receptor-like kinase is required for surface formation of epidermis-related tissues in *Arabidopsis thaliana*. The Plant Journal: for Cell Molecular Biology 39:298–308. https://doi.org/10.1111/j.1365-313X.2004.02132.x

Yang YZ, Peng H, Huang HM, Wu JX, Ha SR, Huang DF, **and** Lu TG (2004) Large-scale production of enhancer trapping lines for rice functional genomics. Plant Sci 167:281–288. https://doi.org/10.1016/j.plantsci.2004.03.026

Tables

Due to technical limitations, table 1 is only available as a download in the Supplemental Files section.

Figures

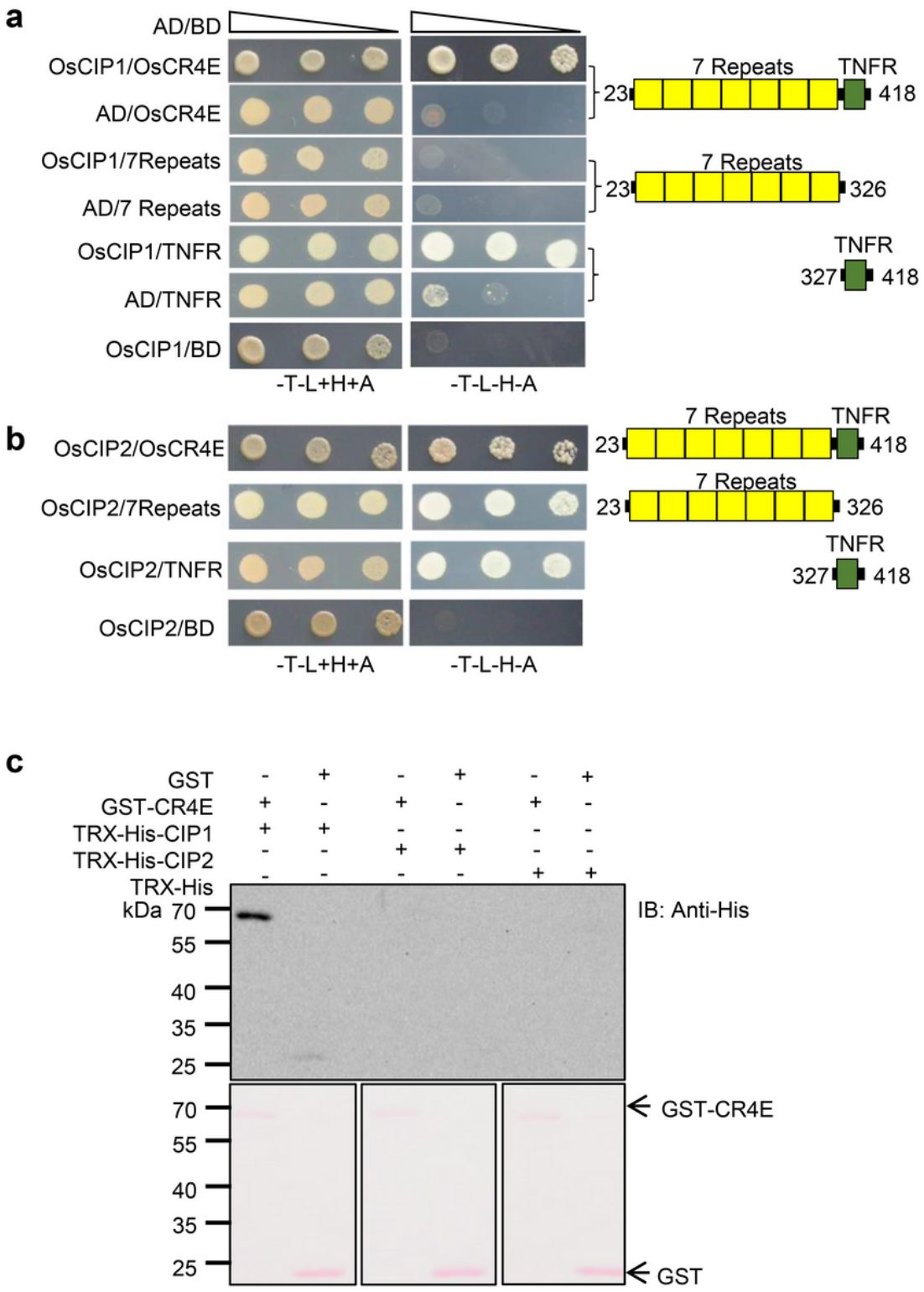

Figure 1

OsCIP1 and OsCIP2 interact with the OsCR4E Verification of the interaction between the OsCR4E and OsCIP1 (a) or OsCIP2 (b) by yeast-two hybrid analysis. The complete OsCR4E or one of two domains (7-Repeats or TNFR) was included in the BD vector as bait, and OsCIP1 or OsCIP2 was included in the AD vector as prey. Three gradiently diluted clones grown on medium without Trp, Leu, His, and Ade (-T-L-H-A) for 4 days after inoculation were considered as proof of an interaction, while clones showing growth on

medium without Trp and Leu (-T-L+H+A) were used as the plasmid co-transformation control. Co-transformation with the empty vector pGBKT7 (BD) or pGADT7 (AD) was used as a negative control for the bait-prey interaction. A schematic diagram of the OsCR4E and of the 7-Repeats and TNFR domains are provided on the right. The numbers indicate amino acid positions. (c) OsCIP1, but not OsCIP2, interacted with the OsCR4E in an overlay assay. Three parallel nitrocellulose membranes loaded with GST-CR4E and GST were, respectively, incubated with TRX-6His-OsCIP1, TRX-6His-OsCIP2, and TRX-6His, and the His-fused proteins were detected with HRP-labeled anti-His antibody (upper panel). Ponceau S (P.S) staining indicates equal loading (bottom panel).

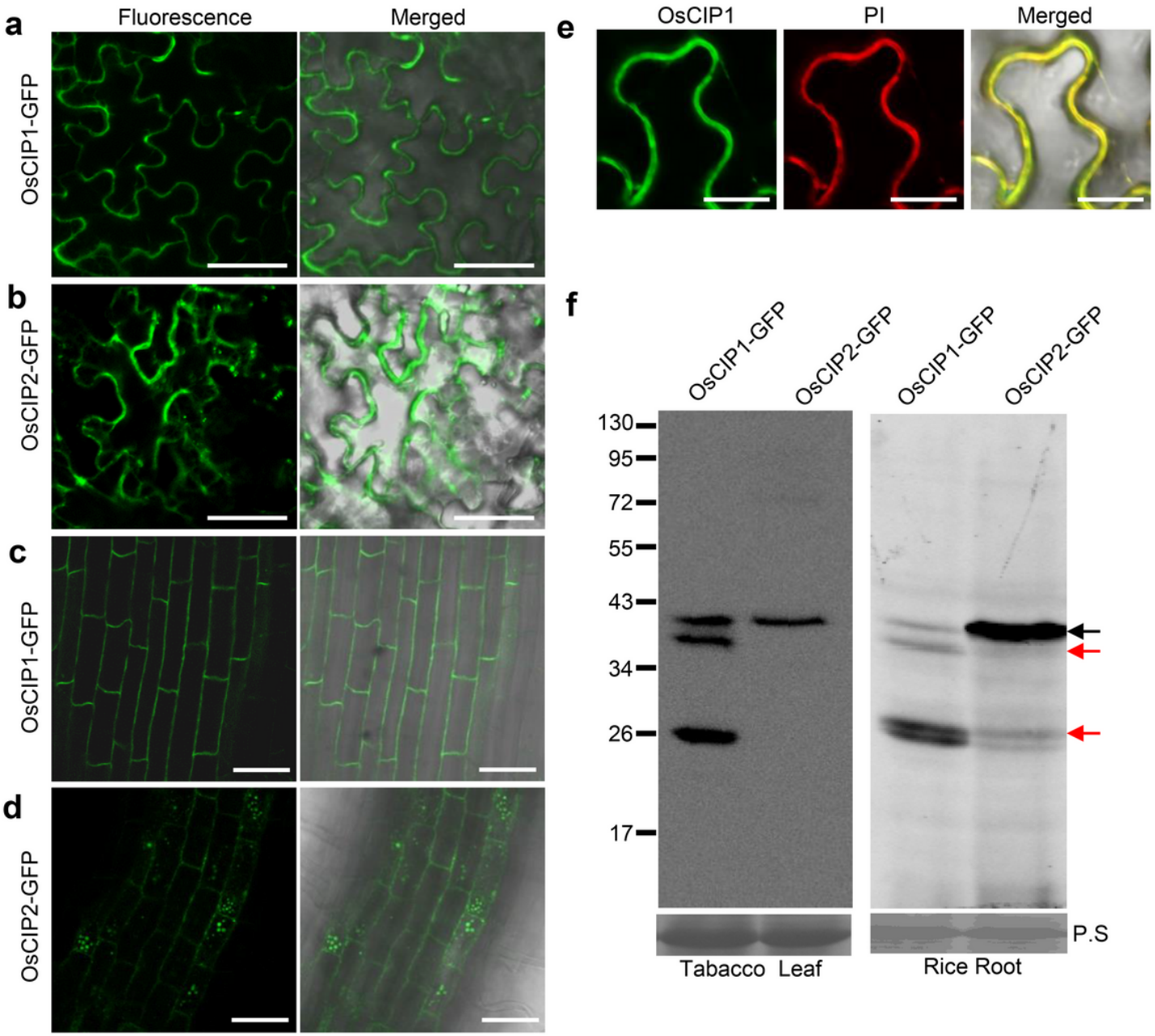

Figure 2

Subcellular localization and immunoblot assays for OsCIP1 and OsCIP2 The subcellular localization patterns of OsCIP1-GFP and OsCIP2-GFP in epidermal cells of tobacco leaves (a and b) and transgenic

rice root cells (c and d) are shown. (e) OsCIP1-GFP co-localized with PI staining. (f) Immunoblot analyses of the expression of OsCIP1-GFP and OsCIP2-GFP fusion proteins in tobacco leaf and transgenic rice roots using anti-GFP polyclonal antibody are shown. The black arrow marks the full-length fusion protein, and the red arrows mark the truncated GFP fusion protein or GFP tag. Bars = 50 μm in (a-d), 20 μm in (e).

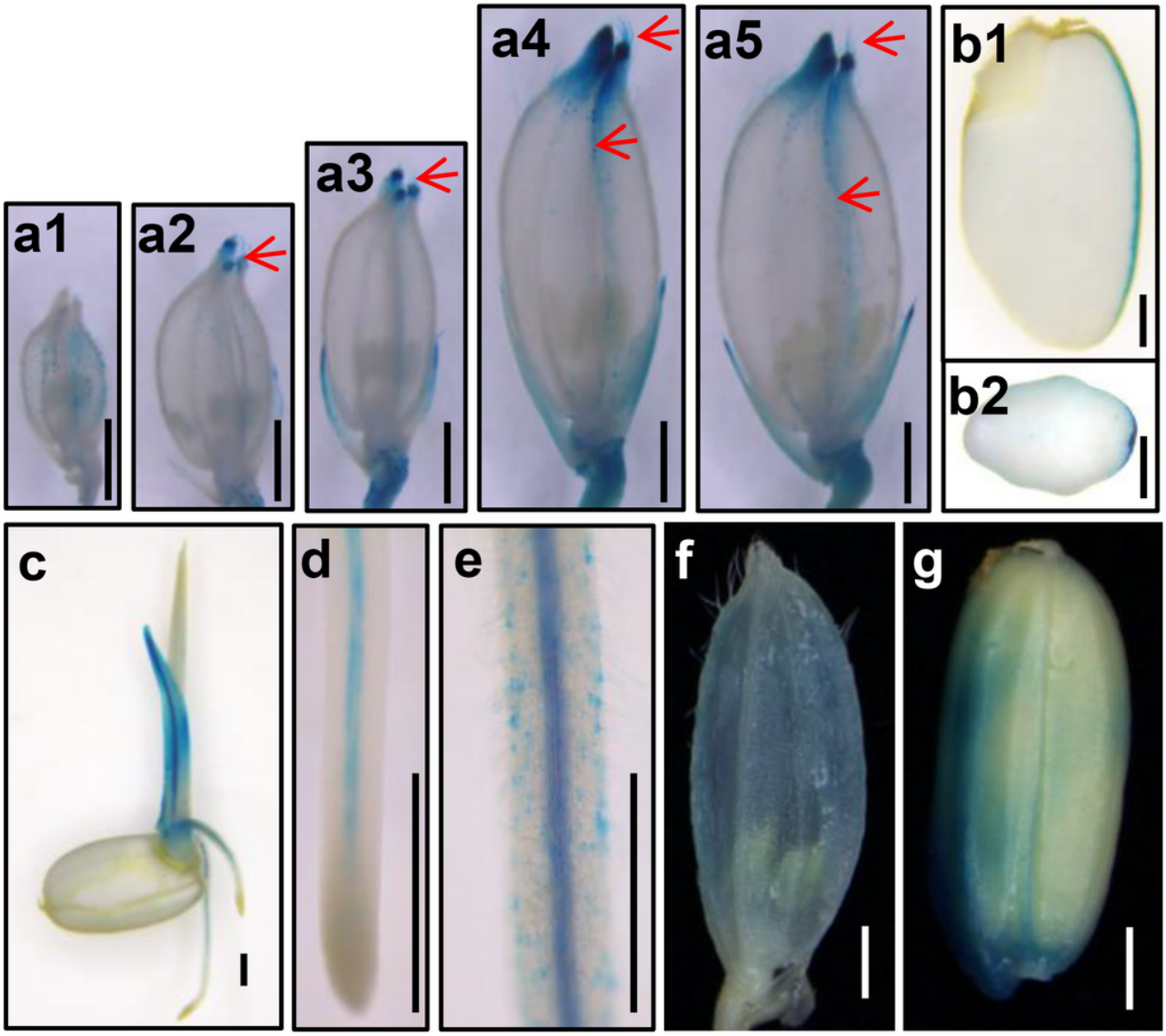

Figure 3

The tissue expression patterns of OsCIP1 and OsCIP2 (a-e) GUS staining for proOsCIP1:GUS/Nip. (a1-5) Spikelets at anther stages 7, 8, 9, 10, and 11, respectively. (b1-2) A longitudinal section (b1) and transverse section (b2) of a mature seed. (c) Seedlings at 2 DAG. (d and e) Seminal roots at 3 DAG. (f and g) GUS staining for proOsCIP2:GUS/Nip. (f) A spikelet at anther stage 9. (g) Mature seed. Bars = 1 mm.

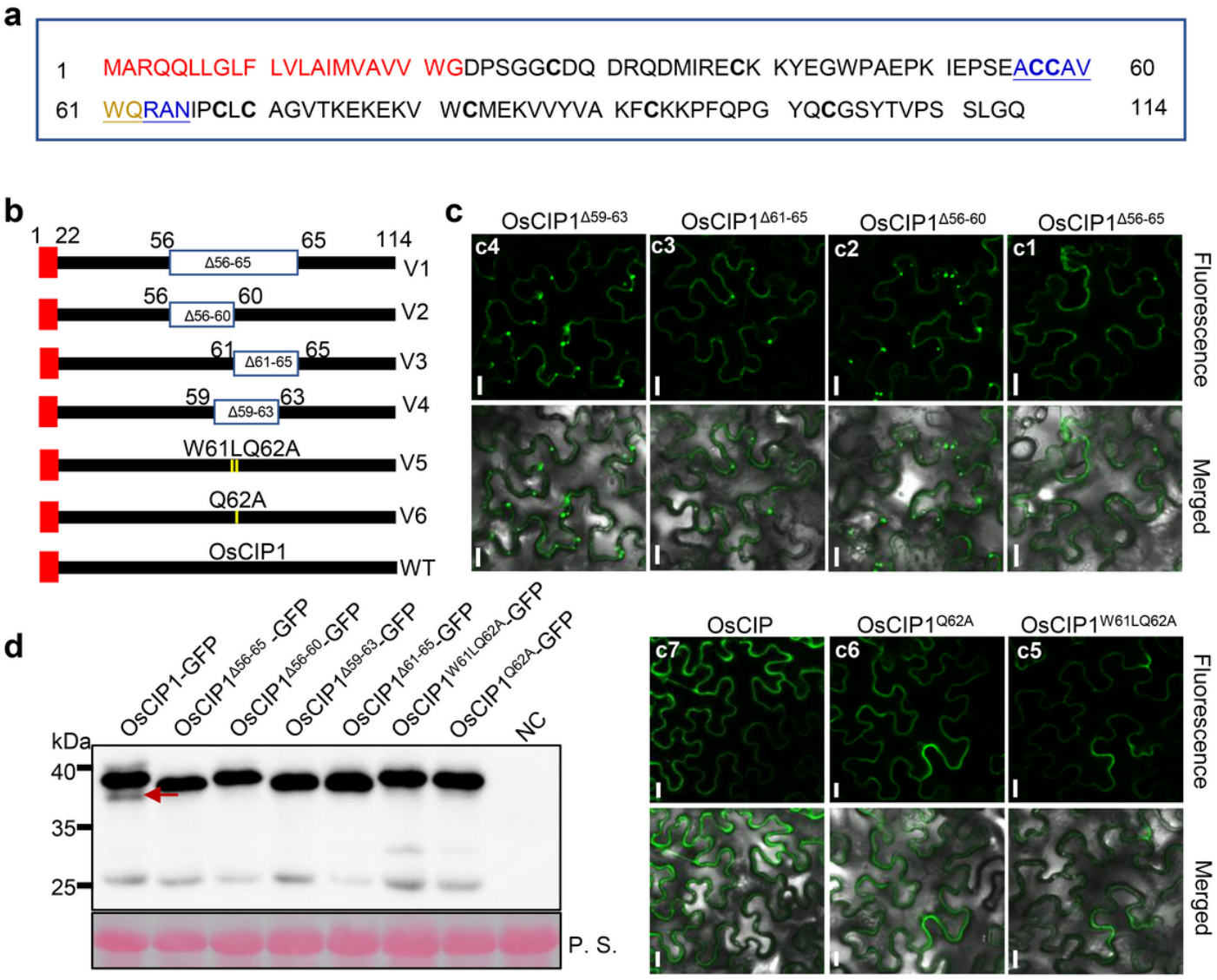

Figure 4

Subcellular localization and immunoblot assays for OsCIP1 and OsCIP1 variants (a) The amino acid sequence of OsCIP1. The signal peptide (SP) is shown in red. The variant amino acid position is underlined. (b) Schematic diagram of the OsCIP1 and OsCIP1-V variants. The red box indicates the SP, the empty box indicates deleted amino acids, the yellow bar indicates mutated amino acids, and numbers indicate the amino acid positions. (c) Subcellular localization of wild-type OsCIP1 and the indicated OsCIP1 variants with a GFP fusion at the C-terminus (driven by the 35S promoter) in tobacco leaf epidermal cells. Bars = 20 μm. (d) Detection of wild-type OsCIP1 and the OsCIP1 variants in (c) by immunoblotting using anti-GFP antibody. The truncated form of OsCIP1-GFP is marked with a red arrow. P.S staining shows equal loading.

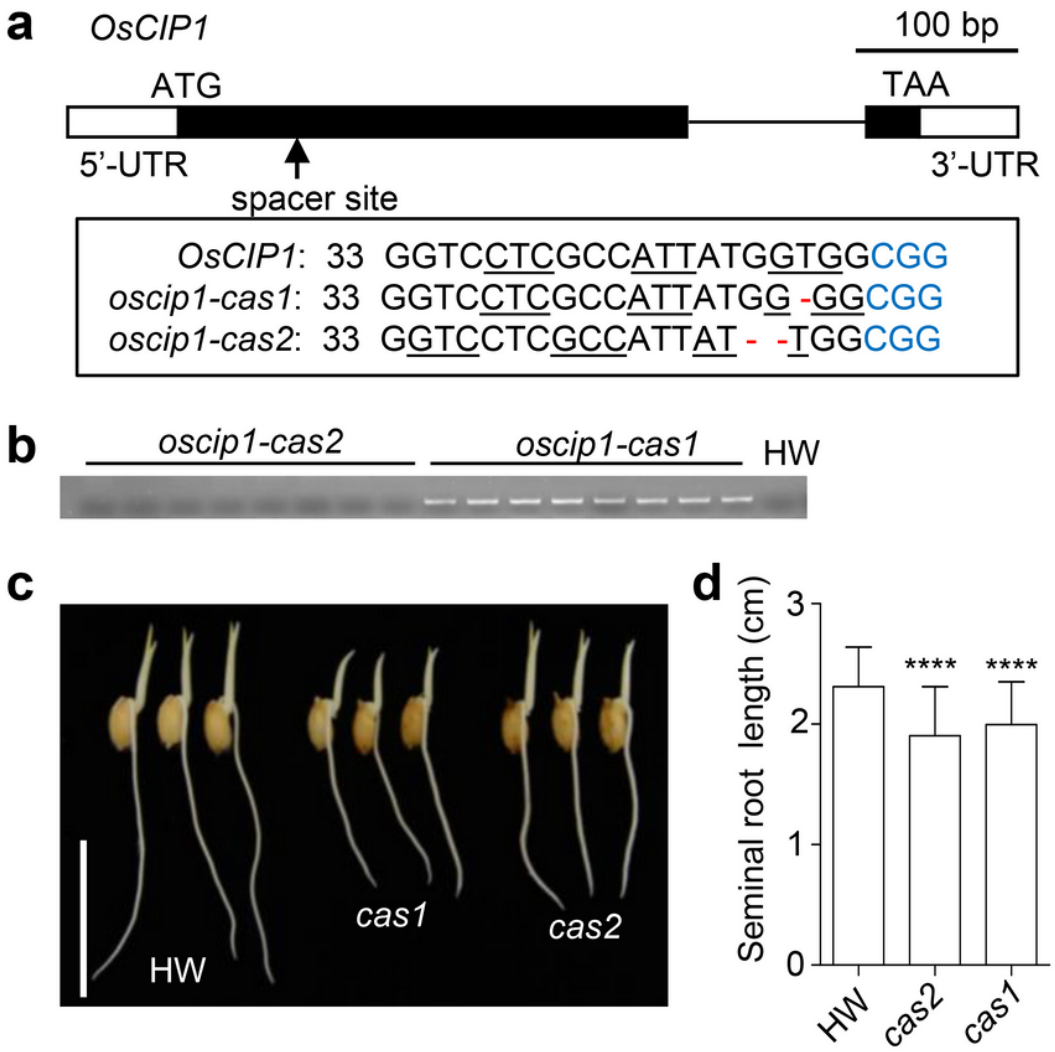

Figure 5

The seminal root of oscip1-cas is much shorter than that of HW (a) Schematic diagram of the OsCIP1 gene structure (upper). Boxes indicate exons and lines indicate introns. Empty box, 5'- or 3'-UTR. Filled box, the coding region. The arrow marks the CRISPR/Cas9 target. The nucleotide sequence at the Cas9 target site in two oscip1 mutant alleles was compared with wild type (lower). The red dashes indicate deleted amino acids, and underlining indicates the coding frame. The editing target sequence is in blue. (b) Detection of the CRISPR/Cas9 vector in oscip1-cas1 and -cas2 by PCR using the vector-specific primer U3-F and the CRISPR/Cas9 target primer 25350-1-R. (c) Dark-grown HW and oscip1 mutant seedlings at 3 DAG. Bar = 2 cm. (d) Statistical analysis of the seminal root length in (c). Error bars indicate the means ± standard deviation (n>24). *** indicates a significant difference from HW (P<0.0001).

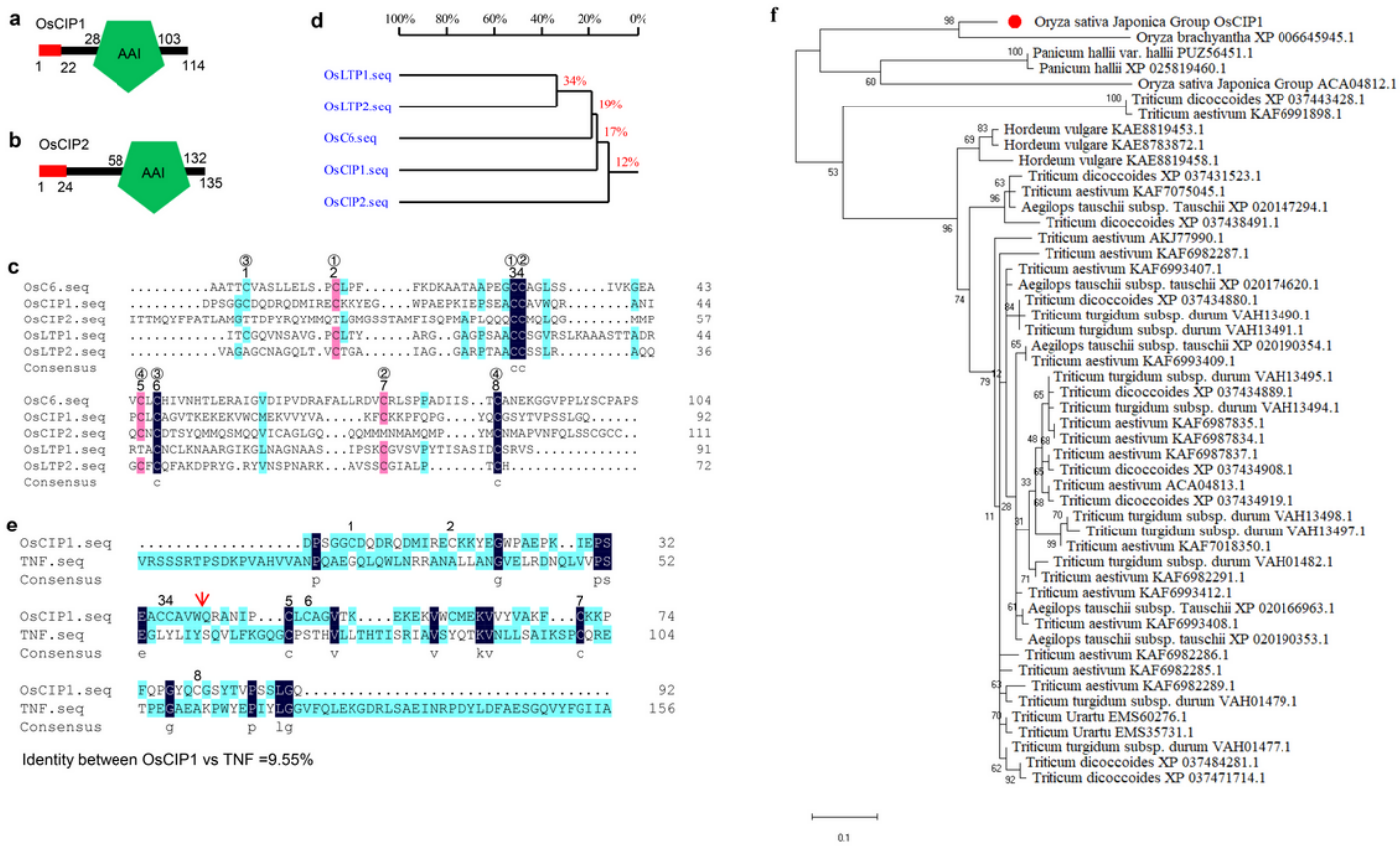

Figure 6

Amino acid sequence, protein structure, and phylogenetic tree analyses (a and b) The predicted domains in OsCIP1 and OsCIP2. Red boxes indicate signal peptides and green boxes represent the AAI domain. Numbers mark the amino acid positions. (c) Protein sequence alignment of OsCIP1, OsCIP2, and three known LTPs: OsLTP1, OsLTP2, and OsC6. Numbers mark the positions of eight conserved cysteines in the nsLTP and circled numbers mark the four pairs of cysteine disulfide bonds formed in nsLTP1. (d) Homolog assay using the five proteins in (c). (e) Protein sequence alignment of OsCIP1 and TNF. The red arrow indicates the cleavage site and numbers mark the positions of the eight conserved cysteines in OsCIP1. (f) Phylogenetic tree showing the OsCIP1 and OsCIP1 homologs.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Table1.xlsx
- SupplementalFigures.pptx
- supplementalmaterials.docx