Genome-wide identification and analysis of monocot-specific chimeric jacalins (MCJ) genes in Maize (*Zea mays* L.)

Hailong Jiang[1†], Jiajian Peng[1†], Qian Li[1†], Siqian Geng[1], Hualei Zhang[1], Yuting Shu[1], Rui Wang[1], Bin Zhang[1], Changsheng Li[1] and Xiaoli Xiang[1*]

Abstract

Background The monocot chimeric jacalins (MCJ) proteins, which contain a jacalin-related lectin (JRL) domain and a dirigent domain (DIR), are specific to Poaceae. MCJ gene family is reported to play an important role in growth, development and stress response. However, their roles in maize have not been thoroughly investigated.

Results In this study, eight *MCJ* genes in the maize genome (designated as *ZmMCJs*) were identified, which displayed unequal distribution across four chromosomes. Phylogenetic relationships between the ZmMCJs were evident through the identification of highly conserved motifs and gene structures. Analysis of transcriptome data revealed distinct expression patterns among the *ZmMCJ* genes, leading to their classification into four different modules, which were subsequently validated using RT-qPCR. Protein structures of the same module are found to be relatively similar. Subcellular localization experiments indicated that the ZmMCJs are mainly located on the cell membrane. Additionally, hemagglutination and inhibition experiments show that only part of the ZmMCJs protein has lectin activity, which is mediated by the JRL structure, and belongs to the mannose-binding type. The *cis*-acting elements in the promoter region of *ZmMCJ* genes predicted their involvement response to phytohormones, such as abscisic acid and jasmonic acid. This suggests that *ZmMCJ* genes may play a significant role in both biotic and abiotic stress responses.

Conclusions Overall, this study adds new insights into our understanding of the gene-protein architecture, evolutionary characteristics, expression profiles, and potential functions of *MCJ* genes in maize.

Keywords Maize, Jacalin and dirigent proteins, Phylogenetic analysis, Subcellular localization, Agglutination activity

Background

Lectins are a group of glycoproteins that are not capable of causing an immune response [1]. They harbor at least one reversible non-catalytic site that binds to sugars, and

play important roles in various cellular processes such as, cell division, endogenous regulation [2], and responses to stress conditions [3].

Jacalin-related lectins (JRLs) are carbohydrate-binding proteins, which are lectins isolated from jackfruit [4]. In recent times, JRLs have emerged as a new subcategory of non-traditional lectins, and that are involved in response to both biotic and abiotic stressors [5]. Based on their ability to bind to sugars, the JRL family can be divided into two main types: galactose-binding lectins (storage lectins found in vesicles) [6], and mannose-binding

†Hailong Jiang, Jiajian Peng and Qian Li contributed equally to this work.

*Correspondence:
Xiaoli Xiang
xxiang@ahau.edu.cn
[1] The National Engineering Laboratory of Crop Stress Resistance Breeding, Anhui Agricultural University, Hefei, China

Jiang *et al. BMC Plant Biology*   (2024) 24:636

Page 2 of 12

lectins (inducible lectins found in the cell nucleus and cytoplasm) [7]. The mannose-binding lectins may play a widespread role in intracellular signal transduction [8, 9]. From a structural perspective, these lectins can be classified into three categories: 1) lectins with only jacalin domain, 2) hybrid lectins (contains jacalin domain and another type of domain viz. Dirigent (DIR and DIR-like), NB-ARC (nucleotide-binding domain shared with APAF-1, various R-proteins and CED-4), F-Box (a diverse class of adaptor proteins of the ubiquitin–proteasome system), GNA (Galanthus nivalis agglutinin), or PAG (polynucleotide adenosine glycosidase) [10], and 3) total lectins with multiple concatenated jacalin domains [11]. Hybrid lectins have become the focus of research due to their unique combinations of structural domains, which may give them functionalities that are not seen in their parent genes. This can often be a driving force behind adaptive evolution [12].

The unique lectin subfamily found in monocotyledonous plants, known as the *MCJs* (Dirigent and Jacalin fusion lectin) gene family, is formed through the fusion of Jacalin and Dirigent proteins [13]. Previous reports in rice and wheat suggests that the primary function of the *MCJs* family is defense against both biotic and abiotic stressors [14]. The Dirigent domain belongs to the DIR gene family, primarily involved in the biosynthesis of lignin and lignans [15]. It also regulates cell wall metabolism and is responds to various stressors, including infestations by pests [16, 17]. Presently, a total of 11 *MCJs* have been identified, including Beta-Glucosidase Aggregation Factor 1 (*BGAF1*) in maize and sorghum [18, 19], *HvJPR1* (*32* kDa *protein JRG1.2 gene, jasmonate-regulated*) in barley [20], *OsJAC1* (*MANNOSE-SPEDCIFIC JACALIN-RELATED LECTIN*) in rice [21], *Crs-1* (*creeping specific-1*) in bent grass [22], as well as *TaJA1* (*JASMONATE-INDUCED PROTEIN1*), *TaVer2* (*VERNALIZATION2*), *TaWCI-1* (*WHEAT CHEMICALLY INDUCED 1*), *TaHfr* (*HESSIAN FLY-RESPONSIVE DISEASE RESISTANCE*), *TaMCJ1* (*Monocot chimeric jacalin 1*) and *TaMCJ3* (*Monocot chimeric jacalin 31*) in wheat [14, 23, 24]. These MCJs share remarkable structural similarities and play vital roles in plant development and response to stressors [25].

In rice, *OsJAC1* overexpression enhances resistance against non-host pathogenic microorganisms, wherein OsJAC1 accumulates at the fungal infection sites [26], While the JRL domain of OsJAC1 is involved in the recognition of the mannose complexes derived from fungal hyphae, the DIR domain recruits other defense-related proteins to form disease resistance complexes, which inhibits the pathogens [27]. Similarly, barley *TaJA1* overexpression in tobacco enhances resistance against tobacco powdery mildew [28] and fungal pathogens in

barley [29]. Wheat *TaMCJ1* and *TaMCJ3* overexpression results in enhanced resistance to wildfire disease, and dehydration tolerance, respectively [14]. The maize *ZmBGAF1* (homodimer, aggregates beta-glucosidases) forms chain-like complexes, associated with pest resistance [30, 31]. The primary function of beta-glucosidases is to release toxic glycoside elements (e.g., DIMBOA,) that contributes to resistance against pests including aphid and corn borer [32].

Maize (*Zea mays*, family: Poaceae) is an important cereal worldwide [33]. This study presents the identification of MCJ gene family members in maize genome, followed by analysis of their gene structure, sequence characteristics, evolutionary relationships, and subcellular localization. Additionally, predictive protein–protein interaction networks of *ZmMCJ* gene members have been examined. The expression profiles *ZmMCJ* members have been examined across various tissues, using publicly available transcriptome sequencing data as well as by Reverse Transcription-quantitative Polymerase Chain Reaction (RT-qPCR) analysis. Blood agglutination and inhibition experiments were also carried out to investigate the lectin functions of ZmMCJ members and their potential substrates. This research sets the groundwork for further exploration of the role of the *ZmMCJ* gene family in growth and development.

Results

### Identification of *ZmMCJ* gene family members in the maize genome

To characterize the number of MCJ family members, we searched the maize MCJ proteins in the Ensembl and Phytozome databases, which were both based on the B73 genome version 4. This search resulted in the identification of eight *ZmMCJ* members with minor variations in predicted characteristics (Supplementary Table S1). The *ZmMCJ* genes are unevenly distributed on four chromosomes, with chromosome no. 6 bearing five members (Fig. 1).

Phylogenetic relationships and gene structure analysis of ZmMCJ family members

Phylogenetic analysis of 36 MCJ members, including members from Wheat (16 members), Rice (4 members), Sorghum (8 members), and Maize (8 members) was carried out (Fig. 2). The MCJ proteins were classified into five Groups (I to V), wherein the eight ZmMCJ members were placed in Group I and Group III. Notably, the maize MCJ proteins exhibited a prominent presence in Group I, implying a high level of similarity for these *MCJ* genes. Furthermore, two members (ZmMCJ1, ZmMCJ7) were clustered with the MCJ proteins from wheat, rice, and sorghum, in Group III (Fig. 2). The eight maize
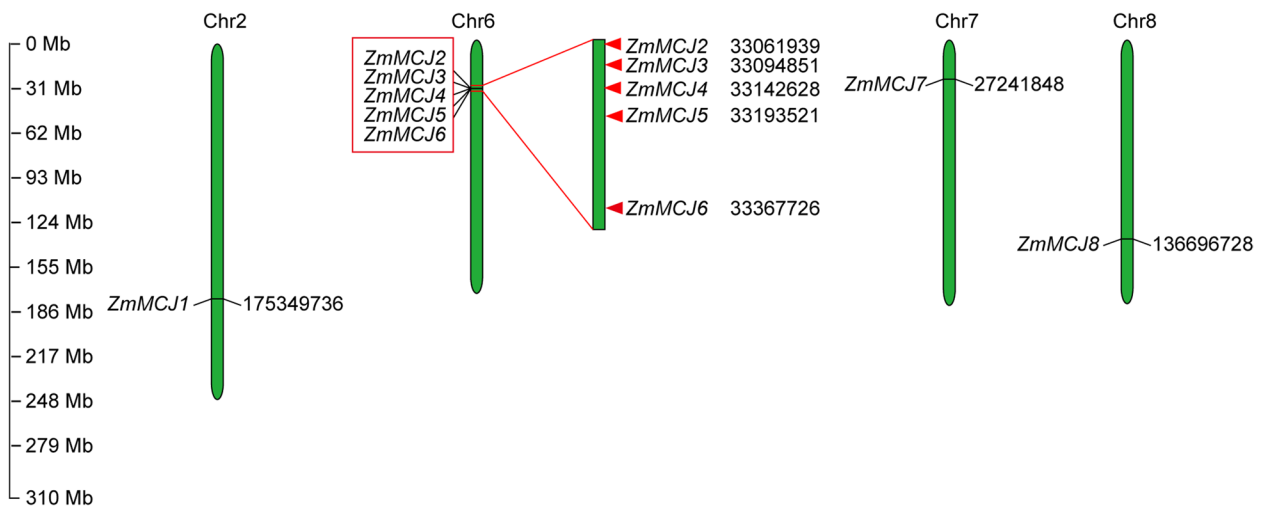
Fig. 1 Chromosomal locations of eight *ZmMCJs* in maize. The read box on chr.6 indicated the tandem duplication event
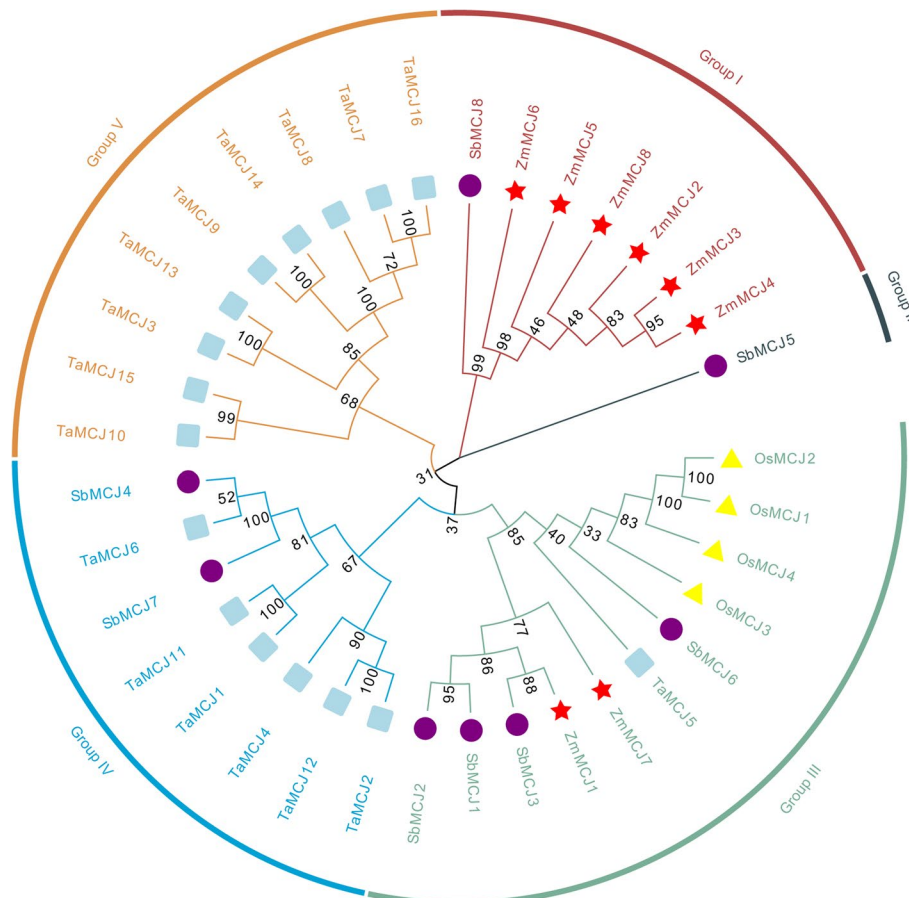
Fig. 2 Phylogenetic relationship of *ZmMCJs*, *OsMCJs*, *TaMCJs* and *SbMCJs*. An un-rooted phylogenic tree was constructed in MEGA11 based on multiple alignment of full sequences from four monocot species (*Z. mays, O. sativa, T. aestivum* and *S. bicolor*)

Jiang *et al. BMC Plant Biology* (2024) 24:636

Page 4 of 12

ZmMCJs could be further sub-categorized into Groups I to III (Fig. 3A). Through protein motif analysis of the ZmMCJ members, we identified the presence of ten conserved motifs (Fig. 3B and Fig. S1). All ZmMCJ proteins contained motif1 and motif4, with most of the members also featuring motif2, motif3, motif5, motif6, motif7, and motif9. Except for ZmMCJ4 in subgroup I, all proteins included motif3. Interestingly, only ZmMCJ1 and ZmMCJ7 of subgroup III possessed motif10. In terms of the gene structures, the ZmMCJs in subgroup I typically comprised of 4 exons, except for ZmMCJ4 which consisted of 3 exons. On the other hand, ZmMCJ6 of subgroup II contained 3 exons, while ZmMCJ5 featured 4 exons. In subgroup III, ZmMCJ7 contained 3 exons, whereas ZmMCJ1 consisted of 4 exons (Fig. 3C).

Synteny analysis of the *ZmMCJ* gene family

In order to investigate the evolution of the *ZmMCJ* gene family in maize, distribution of the eight identified ZmMCJ genes on the 10 maize chromosomes and segmental and tandem duplication events were analyzed. Homologs located on different chromosomes were considered as segmental duplications. Among the eight *ZmMCJ* genes, one set of tandem duplication event, namely *ZmMCJ2*, *ZmMCJ3 ZmMCJ4*, *ZmMCJ5* and *ZmMCJ6*, was detected (Fig. 1). One set of segmental duplicates, namely *ZmMCJ2* and *ZmMCJ8*, were detected (Fig. 4A). The Ka/Ks ratio range from 0.08 to 0.60 suggested the presence of purifying selective pressure during *ZmMCJs* gene family evolution, indicating a potential shared conserved function among these genes. But the *ZmMCJ* genes vary in gene-protein organization, suggesting some differences also in the roles. Furthermore, a comparative syntenic map for insights into relationship between the maize genome and the genomes of Wheat, Rice, and Sorghum (Fig. 4B), displayed collinearity of two maize *ZmMCJ* genes (*ZmMCJ2* and *ZmMCJ8*) with one Sorghum *SbMCJ* gene (*SbMCJ5*), suggesting potential significance of these orthologous pairs in plant evolution. These genes are all in Group I (Fig. 2), suggests they may have similar functions.

### Identification of stress-related *cis*-acting elements in the *ZmMCJ* gene promoters

The outcome of analysis of *cis*-acting elements among eight *ZmMCJ* genes exhibits presence elements associated with response to abscisic acid (ABRE, ABRE3a, ABRE4) and MeJA (CGTCA-motif) (Fig. S3). The promoter regions of eight *ZmMCJ* genes contain numerous anaerobic response elements (ARE), and most of them possess the drought-inducibility element MBS, suggesting MYB transcription factor-mediated regulation. Multiple gene promoters also harbored low temperature response element (LTR) and, four *ZmMCJ* gene promoters encompass O2 binding sites, indicating their potential significance in maize endosperm development.

Tissue and developmental stage specific expression pattern of *ZmMCJs*

By analyzing the transcript data obtained from public RNA-seq data [21], the tissue and developmental stage specific expression patterns of the *ZmMCJ* genes was analyzed (Fig. 5A). The *ZmMCJs* were clustered into four main modules (a to d) based on expression abundance. Module a genes (*ZmMCJ5*, *ZmMCJ6*) showed high expression levels in ovule, tassel1, tassel2, tassel3, cob1, cob2, ear, and early-stage seeds (S0, S4, S6), while module b genes (*ZmMCJ1*, *ZmMCJ7*) exhibited high levels in leaf 6 and leaf 7, with *ZmMCJ1* also being highly expressed in leaf 1. Module c genes (*ZmMCJ3*, *ZmMCJ4*) displayed embryo-specific patterns with higher expression level, while module d genes (*ZmMCJ2*, *ZmMCJ8*) were expressed solely in the endosperm. *ZmMCJ2* starts expressing at 10 DAPS (days after pollination), which is the initiation stage of endosperm filling, while *ZmMCJ8* starts expressing during the middle stage of endosperm filling (14 DAP). Both *ZmMCJ2* and *ZmMCJ8* are highly enriched throughout endosperm development.
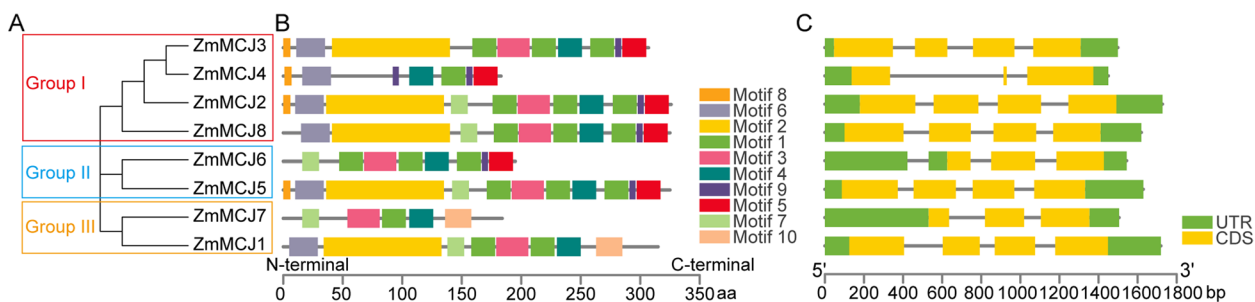
Fig. 3 Motif structure and gene structure, analysis of *ZmMCJs*. **A** Phylogenetic tree of ZmMCJ proteins. **B** 10 conserved motifs were identified in protein sequences of ZmMCJs. **C** Gene structures. Exons and introns were indicated by boxes and lines respectively
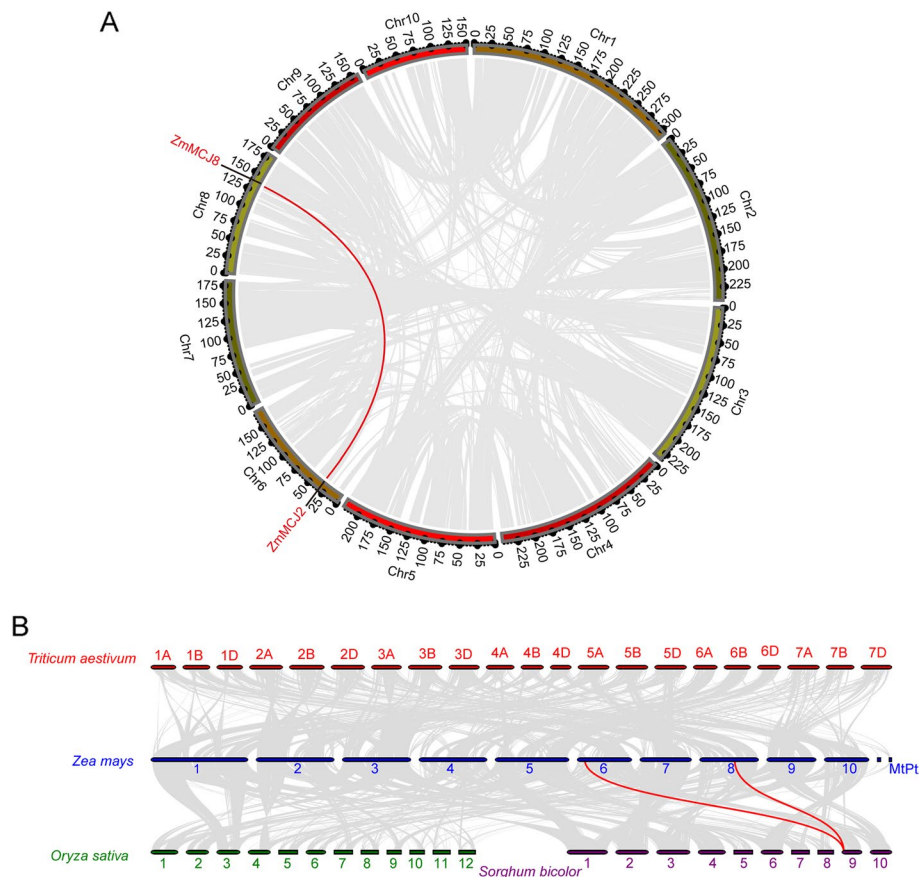
Fig. 4 Schematic diagram of the syntenic relationships of ZmMCJs. **A** Schematic diagram of the syntenic relationships of ZmMCJs in maize. The gray ribbons represent syntenic blocks in the Maize genome, and the segmental duplication events are marked in red. **B** Synteny analysis of MCJ genes among Wheat, Maize, Rice, and Sorghum. Gray lines in the background indicate collinear blocks within Maize and the indicated plant species, whereas the red lines highlight syntenic MCJ gene pairs

Validation of expression profiles of eight *ZmMCJs* (from module a to c) by RT-qPCR was done in multiple plant organs (including leaves, roots, stems, husks, tassels, sheaths, 3 DAY seeds, silks, embryos, endosperms, aleurones, seed coats, and pollen) showed consistency in some genes with the RNA-seq data (Fig. 5B). The expression levels of two genes (module a: *ZmMCJ5* and *ZmMCJ6*) from was prominent in stems, husks, and silk. Two module b genes (*ZmMCJ1* and *ZmMCJ7*) exhibit high expression levels in stems. Two module c (*ZmMCJ3* and *ZmMCJ4*) showed high expression in pollen. Two others module d (*ZmMCJ2* and *ZmMCJ8*) showed high expression in endosperm.

Subcellular localization of ZmMCJs
To determine the subcellular localization of eight ZmMCJ members, each selected MCJ protein was fused to green fluorescent protein (GFP). The free GFP was used as the control. The constitutive 35S promoter drove all gene cassettes. We transiently expressed the ZmMCJ-GFP fusion proteins in both maize yellow seedling protoplast (Fig. 6A) and tobacco leaves (Fig. 6B). All signals of the fused proteins except *35S::ZmMCJ2:GFP, 35S::ZmMCJ 6:GFP, 35S::ZmMCJ8:GFP* were localized in cell membrane, whereas the control *35S:GFP* was detected both in nuclei and the cytoplasm (Fig. 6A, B). *35S::ZmMCJ2:GF P, 35S::ZmMCJ6:GFP, 35S::ZmMCJ8:GFP* were localized as dot signal in cytoplasmic (Fig. 6A, B). They formed numerous small droplets of various sizes, which is a phenomenon characteristic of liquid–liquid phase separation [34]. These signals also similar to two Metacaspases (ZmMC1 and ZmMC2), which are mainly localized in punctate dots aggregates and are partially co-localized with autophagosomes [35]. Whether ZmMCJ2, ZmMCJ6 and ZmMCJ8 are liquid–liquid phase separation or autophagosome related, further study needs to be done.

Analysis of protein tertiary structures
The ZmMCJ protein consists of the JRL domain and the DIR domain. Both domains are highly conserved,

Jiang *et al. BMC Plant Biology*   (2024) 24:636

Page 6 of 12
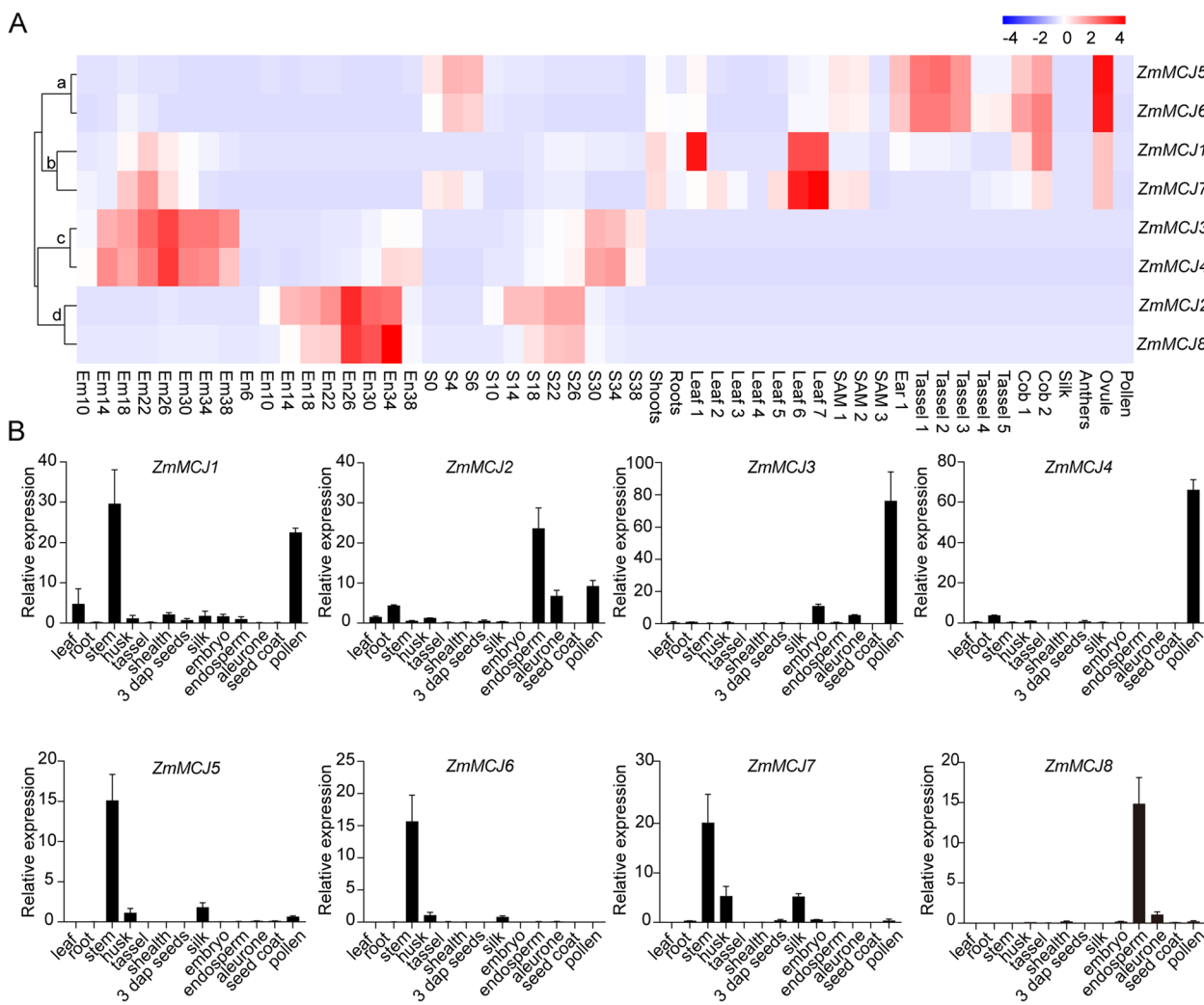
Fig. 5 The expression profiles of *ZmMCJs* genes in maize. A Heat map of *ZmMCJs* genes expression in different tissues and developmental stages. The genes are located on the right, and the tissues are indicated at the bottom of each column. The color bar represents the expression values. S0-S38: developing seed from 0 to 38 DAP (day after pollination); Em10-Em38: developing embryo from 10 to 38 DAP; En10-En38: developing endosperm from 10 to 38 DAP. S0-S38: developing seeds from 0 to 38 DAP. **B** Results of quantitative RT-PCR of eight *ZmMCJ* gens in different tissues. *Zmactin* was used as the internal reference gene. Data were shown as mean ± s.d., *n* = 3 biological replicates; each biological replicate represents an individual plant. Statistically significant differences are indicated (*p* < 0.05, Student's *t*-test)

predominantly comprising β-sheets, β-turns, and minor irregular coils. The 3D structures of the ZmMCJ1/7, ZmMCJ2/8, ZmMCJ3/4 and ZmMCJ5/6 proteins were similar to each other respectively, interestingly, this fits with their organizational positioning (Fig. 7).
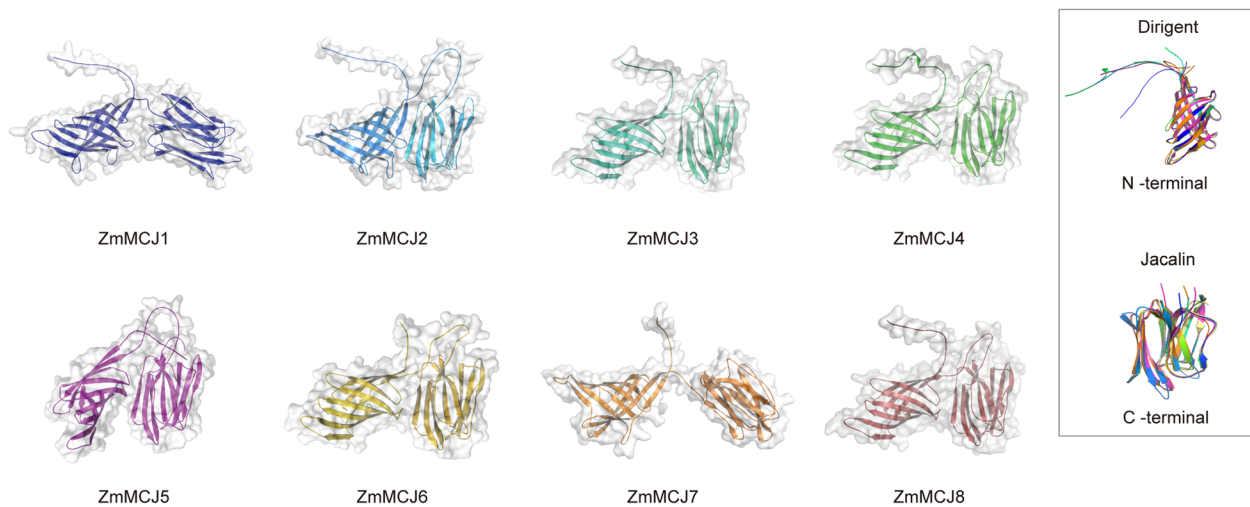
Agglutination activity and carbohydrate-binding properties

In order to explore the agglutinin function of ZmMCJ proteins, the purified proteins were used to conduct cell agglutination tests based on rabbit red blood cells. The concentration of ConA, TF, and TF-ZmMCJs was 0.5 mg/mL. ConA had the ability to clump together rabbit red

blood cells, whereas TF did not. This outcome confirms the success of the tests. Approximately 0.0625 mg/ml of TF-ZmMCJ3 was able to cause agglutination of rabbit erythrocytes, while approximately 0.0020 mg/ml of either TF-ZmMCJ5 or TF-ZmMCJ6 also resulted in agglutination of rabbit erythrocytes. Additionally, approximately 0.98 mg/ml of ConA induced the same effect (Fig. 8A). On the other hand, TF-ZmMCJ1, TF-ZmMCJ4, and TF-ZmMCJ7 did not lead to clumping of rabbit red blood cells (Fig. 8A). Analysis of carbohydrate-binding specificities of TF-ZmMCJs, inhibition assays were conducted using rabbit red blood cells. The highest concentration of DTT allowed inhibition of TF-ZmMCJ3, TF-ZmMCJ5,

Jiang *et al. BMC Plant Biology*   (2024) 24:636

Page 7 of 12
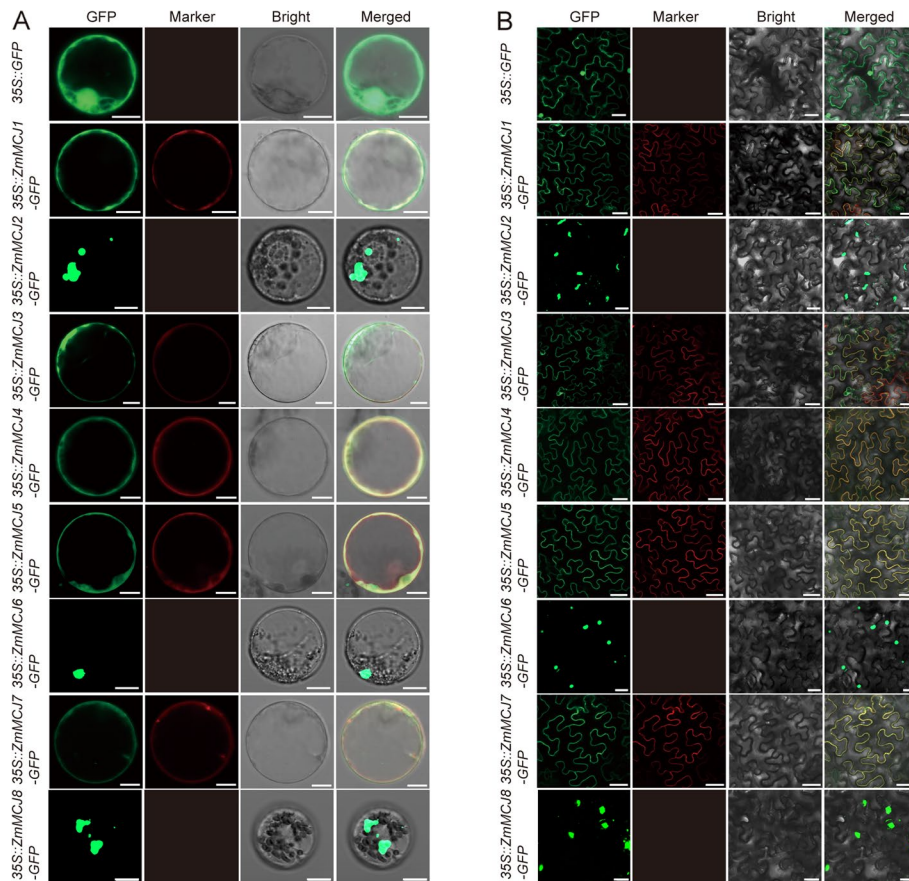
Fig. 6 Subcellular localization of eight ZmMCJs. **A** The constructs were transiently expressed in maize leaves protoplast. **B** The constructs were transiently expressed in *N. benthamiana* leaves via *Agrobacteria* infiltration. The *ZmMCJ* genes were cloned from maize inbred line B73 and used to construct CaMV35S::*ZmMCJs*-GFP vectors in which GFP was fused at the C-terminus. The mCherry-*OsRAC1* was used as cell membrane marker. Pictures were taken using a confocal fluorescence microscope. Scale bars = 50 μm

Fig. 7 Predicated structures of eight ZmMCJ proteins. Dirigent domain is on the N-terminal of MCJ proteins. Jacalin domain is on the C-terminal of MCJ proteins
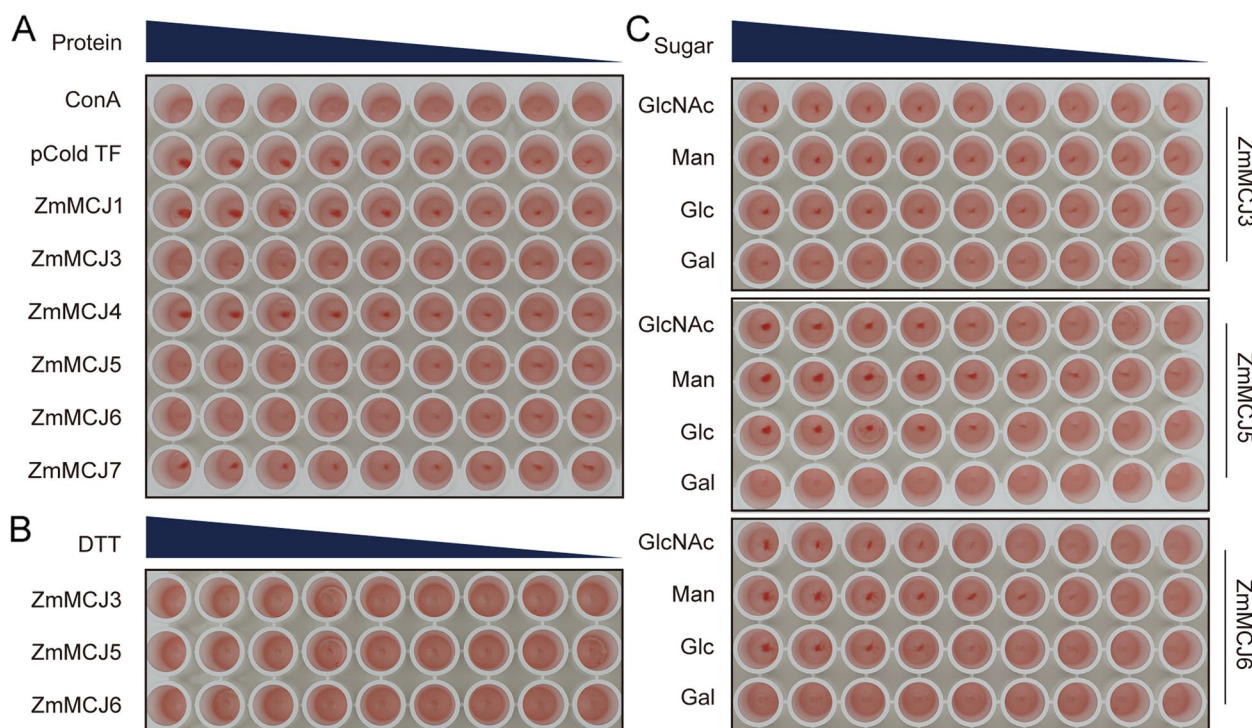
Fig. 8 Agglutination activity and carbohydrate-binding properties of TF-ZmMCJs. **A** Agglutination activity test. The initial concentration of ConA, TF and TF-ZmMCJs is 0.5 mg/mL, and it is diluted 2 times each time. **B** The initial concentration of DTT is 2 mM, and it is diluted 2 times each time, The concentration of TF-ZmMCJs is 0.1 mg/mL. **C** Carbohydrate-binding property test: GlcNAc, Gal, Glc and Man as inhibitors (final concentration approximately 66.7 mM); the carbohydrate-binding specificities of TF-ZmMCJ3 was inhibited by Man when diluted in 192-fold increments, by Glc in 192-fold increments, by Gal or GlcNAc in 768-fold increments

and TF-ZmMCJ6, as shown in Fig. 8B. Furthermore, the minimum inhibitory concentrations of GlcNAc, Man, and Gac for TF-ZmMCJ3 were 0.0010, 0.0005, and 0.0010 mg/ml, respectively (Fig. 8C). Therefore, Man was the most effective inhibitor for TF-ZmMCJ3. For TF-ZmMCJ5, the minimum inhibitory concentrations of GlcNAc, Man, and Glc were 8, 4, and 8 times higher, respectively. Among them, Man had the lowest concentration. Overall, TF-ZmMCJ3, TF-ZmMCJ5, and TF-ZmMCJ6 demonstrated the highest affinity for Man and were unaffected by Gal. These lectins (ZmMCJ3, ZmMCJ5, and ZmMCJ6) belong to a group known as mannose-binding jacalin-related lectins [36].

Discussion

In recent years, numerous studies have focused on the analysis of MCJ gene family in plant species. These genes play various roles in plant growth, development, and response to stress [25, 37]. In this study, eight MCJ family members were identified in maize, which is two times of rice (four) and Moso bamboo (*Phyllostachys edulis*) [5], half of wheat (16), and the same as sorghum (eight) [29]. This suggests that certain MCJ family members have been relatively conserved throughout evolution. Most of

ZmMCJ genes were located as a gene cluster on chromosome 6. This suggests that duplication events may have occurred more frequently on chromosome 6 during the evolution of gene families, possibly influencing their functions [38, 39].

A previous study also constructed a phylogenetic tree based on JRL proteins using the maximum likelihood method, which classified them into eight groups [37]. It implied that JRLs in maize were more closely related to sorghum, which was consistent with our study. The expansion of gene families and genetic evolution in plant genomes are influenced by fragment replication and tandem replication [40, 41]. Gene expression is regulated by *cis*-acting elements. Previous research has analyzed the *cis*-acting elements of *PeD-J* genes in bamboo [5]. Furthermore, it has been demonstrated that MCJs play a role in pest resistance [31], indicating their potential involvement in maize defense against pests and diseases [18]. Here, ZmMCJ2, ZmMCJ6 and ZmMCJ8 proteins were localized as dot signal in cytoplasmic. These might be liquid–liquid phase separation phenomenon, which might be involved in selective autophagy of invading pathogens [34, 42]. These signals are also like endosomes localized, which might have a novel role in defense response [35,

Jiang *et al. BMC Plant Biology*   (2024) 24:636

Page 9 of 12

43]. The expression of *ZmMCJ* genes is tissue-specific in plant. Notably, *ZmMCJ3* and *ZmMCJ4* were high expressed in embryo, suggesting their importance in embryo growth and development. On the other hand, the expression of *ZmMCJ2* and *ZmMCJ8* gradually increased during the grain filling stage, which may be related to endosperm development. Lastly, the *ZmMCJ1* and *ZmMCJ7* genes exhibited high expression levels in leaves.

The JRL domain exhibits the common β-prism fold found in the JRL family, with the active site typically located within the concave cavity at the top of the prism [44]. It consists of three four-stranded β-sheets, forming three Greek key motifs [45]. The DIR domain adopts a β-barrel three-dimensional structure composed of eight antiparallel β-strands [46]. The 3D structure of ZmMCJs show that ZmMCJs might exist as heterodimer or homodimer in vivo [31].

Hemagglutination testing has been widely used for testing the lectin activity [21, 47]. Our data show that ZmMCJ3, ZmMCJ5 and ZmMCJ6 indeed exhibits the agglutination activity, whereas ZmMCJ1, ZmMCJ4, ZmMCJ7 do not have this function. The data suggest that ZmMCJ3, ZmMCJ5 and ZmMCJ6 are functional mannose-specific protein. They may play important role on autoimmune disease and against pathogens [36].

Methods

Identification of *ZmMCJ* genes in maize
The presence of MCJ genes in maize was investigated by HMMER3 [48] based on the search for dirigent domain (PF03018) and the jacalin domain (PF01419) in the maize genome database (https://www.maizegdb.org/). Hits with *E*-values less than or equal to 10^{-10} were considered as potential candidate for further analysis. Sequences containing both dirigent and jacalin domains (MCJ genes), were identified using Venn online software (http://bioinformatics.psb.ugent.be/webtools/Venn/) [49]. The MCJ genes were also searched in the genome databases of rice, wheat, and sorghum online at Ensembl (http://www.ensembl.org/index.html) and Phytozome (https://phytozome.jgi.doe.gov/pz/portal.html) databases. The basic physicochemical properties and signal peptide information of each MCJ family member was predicted using ExPASy ProtParam tool (https://web.expasy.org/protparam/). The MCJ genes identified were designated from *ZmMCJ1* to *ZmMCJ8*, based on their respective genomic locations within the maize chromosomes.

Sequence alignment, phylogenetic relationships analysis
The ZmMCJ full-length protein sequences were aligned using the DNAMAN software version 10 (Lynnon Biosoft, San Ramon, CA, USA). To visually represent the conservation pattern of these sequences, sequence logos were created using the WebLogo tool [50]. For intraspecific phylogenetic trees (for relationships and similarity among different ZmMCJ full-length proteins) and interspecific phylogenetic trees (for relationships among MCJ protein domain sequences from different plant species), multiple sequence alignments were performed using Clustal W program using the default parameters [51] (gap opening penalty 10, gap extension penalty 0.2, gap distance 4, no end gaps and protein weight matrix using Gonnet). Further, the neighbor-joining (NJ) method in MEGA11 was employed to construct these phylogenetic trees [52]. The resulting trees were assessed for their confidence coefficients using 1,000 bootstraps with the pairwise deletion option.

Chromosomal location and collinearity analyses
The details of physical position of the *ZmMCJ* genes were obtained from the maize genome database, while the block duplication events were visualized by TBtools [53]. For the construction of the syntenic map, *MCScanX* was employed [54], and the resulting map was visualized using TBtools. In *MCScanX*, tandem genes are assessed by gene rank according to chromosomal positions [54]. Using TBtools, the non-synonymous (Ka) and synonymous (Ks) substitution rates of the collinear gene pairs were estimated, for to predict the divergence time (t) and evolutionary rate (Ka/Ks ratio) [53].

Motifs and gene structures of *ZmMCJ* genes
The MEME suite (http://meme-suite.org/) was utilized to detect and visualize the conserved patterns found in *ZmMCJ* genes, whereas TBtools software was used to analyze the gene structure, including introns and exons.

Promoter motif analysis
To identify *cis*-acting elements in the promoter region, the 1,400 bp upstream sequence (preceding the ATG start site) of *ZmMCJ* genes was retrieved and analyzed at the PlantCare database (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/), and the results were visualized using TBtools.

Heat map analysis of RNA-Seq gene expression data
The maize reference genome B73 was utilized for analysis of the expression pattern of eight *ZmMCJ* genes [21]. R package 'heatmap', was used for generating a heat map and hierarchical clustering of *ZmMCJ* genes across 52 different samples (tissues and developmental stages). The gene expression values, normalized using $\log_2$ (RPKM+1), were employed in this analysis. The sample IDs were assigned following the previously described procedure [21].

Jiang *et al. BMC Plant Biology*   (2024) 24:636

Page 10 of 12

Total RNA extraction and RT-qPCR analysis

Trizol reagent (Invitrogen Inc., Waltham, MA) was used to extract total RNA from maize tissues, and after *DNase I* digestion (Qiagen Inc., Germantown, MD, USA), the RNA was purified using the RNeasy Mini Kit. For endosperm RNA extraction, the same procedure as above, with the exception of using the RNA extract buffer prior to Trizol reagent [55]. Reverse transcription was conducted with ImProm-II™ Reverse Transcriptase (Promega, A3800). RT-qPCR analysis was performed with SYBR Green I using BIO-RAD CFX96 system. The PCR program included an initial step at 95 °C for 120 s, followed by 40 cycles of 95 °C for 5 s, 60 °C for 30 s, and 95 °C for 5 s. A melt curve analysis was performed from 65 °C to 95 °C, with an increment of 0.5 °C for 5 s. The $2^{-\Delta\Delta Ct}$ analysis method [56] was employed to analyze the expression levels of the target genes relative to the maize Actin gene (*ZmActin*, Zm00001d012277). The PCR primers used are listed in Table S3. Primers were designed by Primer Premier 5.0 to ensure primer specificity [57, 58].

Subcellular localization analysis of ZmMCJ proteins

The ZmMCJ protein sequences were analyze by online tools like wolf-psort (https://psort.hgc.jp/) or Predict-NLS (https://rostlab.org/owiki/index.php/ PredictNLS) to anticipate the nuclear localization signal (NLS). The C-terminal of each *ZmMCJ* was fused to an enhanced GFP (*eGFP*) in pCAMBIA2300 plasmid. Transformation of maize protoplasts involved polyethylene glycol (PEG)-mediated techniques, while Agrobacterium-mediated approaches were employed for tobacco (Nicotiana benthamiana) leaf epidermal cells. As a reference, a pCAMBIA2300-*GFP* vector encompassing a 35 S::*GFP* construct was used. For reliable results, a minimum of three replicates were used. A confocal microscope (Germany, Zeiss, LSM880) was used imaging of the eGFP signal.

Homology modeling

The 3D structures of ZmMCJ proteins were predicted using homology modeling, which is based on their similarity to proteins that already have known structures. To find the appropriate templates for homology modeling, we retrieved them from the AlphaFold database (https://alphafold.ebi.ac.uk/).The template used for ZmMCJ1-ZmMCJ18 are B4G256, C0HFQ1, C0PDX9, B4F7S2 、B4FV87, A0A1D6LHB1, B7ZZ04, C0PDX9, respectively. The EasyModeller software (http://softwaretopic.informer.com/easy-modeller-for-windows/) was used to model the main chain, side chain, and loop region of the proteins' tertiary structures, followed by refinement and optimization using the SAVES server (https://servicesn. mbi.ucla.edu/SAVES/). Visualization and manipulation of the predicted models was done using Discovery Studio, version 2016 (BIOVIA, Pairs, France).

Purification of recombinant protein

To produce the recombinant TF -tagged proteins, the full-length CDS were amplified and inserted into cloning sites *Kpn I* and *EcoR I* in the pCold-TF expression vector (TAKARA) by homologous recombination, and used for overexpression in *Escherichia coli* DE3 (BL21) cells, induced by 1.0 mM IPTG at 16 6 lsessionat100 rpm/min. The cells were collected and sonicated, and the supernatants containing the corresponding fusion proteins were purified with NI-HA beads (QIAGEN). The proteins were eluted with 0 mM, 25 mM, 50 mM, 100 mM, 250 mM, 350 mM imidazole in 30 mM Tris–HCl buffer. The eluted proteins were analyzed on a 10% SDS-PAGE gel. The ZmMCJs proteins were eluted at 250 mM imidazole. The proteins were desalting concentration using Amicon Ultra 50 ml (UFC800308, Millipore, Merck).

Agglutination activity and carbohydrate-binding property tests

Agglutination activity and carbohydrate-binding properties were tested using a two-fold serial dilution procedure [59]. ConA (*Canavalia ensiformis* lectin, which specifically binds mannose and glucose), and pCold TF empty vector (Trigger Factor (TF)-tagged protein) were used as positive and negative controls, respectively.

Statistical analysis

The significance of difference between groups was analyzed by Independent Sample T Test analysis, and results presented as means ± standard deviation (± sd). Statistically significant differences were calculated by Student's t test. Figures are made using software GraphPad Prism 9 and Adobe Illustrator 2024.

Conclusions

A total of eight *ZmMCJ* genes were identified in the maize genome and were found to be distributed unevenly across four chromosomes, with highest number on Chromosome 6. Depending on the nature of the *cis*-acting elements, regulation of *ZmMCJ* genes seems responsive to phytohormones and environmental factors. RNA-seq data and RT-qPCR analysis indicate specificity to tissues different developmental stages, implying the potential involvement of *ZmMCJs* in tissue differentiation and seed development. This investigation offers valuable insights to enhance our understanding of the evolutionary mechanisms and functional characteristics of the MCJ gene family in maize.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12870-024-05354-4.

Supplementary Material 1.

Supplementary Material 2.

Supplementary Material 3.

Supplementary Material 4.

Authors' contributions
Xiaoli Xiang and Changsheng Li performed data collection and processing, participated in study design and interpretation, performed experiments, and wrote the manuscript. Hailong Jiang, Jiajian Peng, and Qian Li participated in all the experiments. Siqian Geng, and Hualei Zhang assisted in the interpretation of the results. Yuting Shu, Rui Wang, and Bin Zhang are responsible for the completeness of the data and accuracy of the data analysis. The manuscript was reviewed and edited by all authors. The authors read and approved the final manuscript.

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

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

References

1. Souza MA, Carvalho FC, Ruas LP, Ricci-Azevedo R, Roque-Barreira MC. The immunomodulatory effect of plant lectins: a review with emphasis on ArtinM properties. Glycoconj J. 2013;30(7):641–57.
2. Xin Z, Wang A, Yang G, Gao P, Zheng ZL. The Arabidopsis A4 subfamily of lectin receptor kinases negatively regulates abscisic acid response in seed germination. Plant Physiol. 2009;149(1):434–44.
3. Ray S, Kapoor S, Tyagi AK. Analysis of transcriptional and upstream regulatory sequence activity of two environmental stress-inducible genes, NBS-Str1 and BLEC-Str8, of rice. Transgenic Res. 2012;21(2):351–66.
4. Witzel K, Matros A, Bertsch U, Aftab T, Rutten T, Ramireddy E, Melzer M, Kunze G, Mock HP. The Jacalin-Related Lectin HvHorcH Is Involved in the Physiological Response of Barley Roots to Salt Stress. Int J Mol Sci. 2021;22(19):10248–27.
5. Ma R, Huang B, Chen J, Huang Z, Yu P, Ruan S, Zhang Z. Genome-wide identification and expression analysis of dirigent-jacalin genes from plant chimeric lectins in Moso bamboo (Phyllostachys edulis). PLoS ONE. 2021;16(3):e0248318.
6. Van Damme EJ, Hause B, Hu J, Barre A, Rouge P, Proost P, Peumans WJ. Two distinct jacalin-related lectins with a different specificity and subcellular location are major vegetative storage proteins in the bark of the black mulberry tree. Plant Physiol. 2002;130(2):757–69.
7. Peumans WJ, Hause B, Van Damme EJ. The galactose-binding and mannose-binding jacalin-related lectins are located in different sub-cellular compartments. FEBS Lett. 2000;477(3):186–92.
8. Peumans JW, Hause B, Damme EJV. The galactose-binding and mannose-binding jacalin-related lectins are located in different sub-cellular compartments. FEBS Letters. 2000;4777(2000):186–92.
9. Van Holle S, Van Damme EJ. Distribution and evolution of the lectin family in soybean (Glycine max). Molecules. 2015;20(2):2868–91.
10. Vd EJM. N L, WJ P. Plant Lectins. Adv Botanical Res. 2008;48:107–209.
11. Van Holle S, Van Damme EJM. Messages from the past: new insights in plant lectin evolution. Front Plant Sci. 2019;10:36.
12. Crow KD, Wagner GP, Investigators ST-NY. Proceedings of the SMBE Tri-National Young Investigators' Workshop 2005. What is the role of genome duplication in the evolution of complexity and diversity? Mol Biol Evol. 2006;23(5):887–92.
13. Ma QH, Zhen WB, Liu YC. Jacalin domain in wheat jasmonate-regulated protein Ta-JA1 confers agglutinating activity and pathogen resistance. Biochimie. 2013;95(2):359–65.
14. Ma QH, Han JQ. Identification of monocot chimeric jacalin family reveals functional diversity in wheat. Planta. 2021;253(2):30–40.
15. Davin LB, Lewis NG. Dirigent proteins and dirigent sites explain the mystery of specificity of radical precursor coupling in lignan and lignin biosynthesis. Plant Physiol. 2000;123(2):453–62.
16. Paniagua C, Bilkova A, Jackson P, Dabravolski S, Riber W, Didi V, Houser J, Gigli-Bisceglia N, Wimmerova M, Budinska E, et al. Dirigent proteins in plants: modulating cell wall metabolism during abiotic and biotic stress exposure. J Exp Bot. 2017;68(13):3287–301.
17. Khan A, Li RJ, Sun JT, Ma F, Zhang HX, Jin JH, Ali M, Haq SU, Wang JE, Gong ZH. Genome-wide analysis of dirigent gene family in pepper (Capsicum annuum L.) and characterization of CaDIR7 in biotic and abiotic stresses. Sci Rep. 2018;8(1):5500–21.
18. Esen A, Blanchard D. A Specific b-Glucosidase-Aggregating Factor Is Responsible for the b-Glucosidase Null Phenotype in Maize. Plant Physiol. 2000;122(2):563–72.
19. Kittur FS, Yu HY, Bevan DR, Esen A. Homolog of the maize beta-glucosidase aggregating factor from sorghum is a jacalin-related GalNAc-specific lectin but lacks protein aggregating activity. Glycobiology. 2009;19(3):277–87.
20. Mueller-Uri F, Parthier B. Jasmonate-induced alteration of gene expression in barley leaf segments analyzed by in-vivo and in-vitro protein synthesis. Planta. 1998;176(2):241–7.
21. Jiang JF, Han Y, Xing LJ, Xu YY, Xu ZH, Chong K. Cloning and expression of a novel cDNA encoding a mannose-specific jacalin-related lectin from Oryza sativa. Toxicon. 2006;47(1):133–9.
22. Li HM, Rotter D, Bonos SA, Meyer WA, Belanger FC. Identification of a gene in the process of being lost from the genus Agrostis. Plant Physiol. 2005;138(4):2386–95.
23. Williams CE, Collier CC, Nemacheck JA, Liang C, Cambron SE. A lectin-like wheat gene responds systemically to attempted feeding by avirulent first-instar Hessian fly larvae. J Chem Ecol. 2002;28(7):1411–28.
24. Yong WD, Xu YY, Xu WZ, Wang X, Li N, Wu JS, Liang TB, Chong K, Xu ZH, Tan KH, et al. Vernalization-induced flowering in wheat is mediated by a lectin-like gene VER2. Planta. 2003;217(2):261–70.
25. Esch L, Schaffrath U. An update on jacalin-like lectins and their role in plant defense. Int J Mol Sci. 2017;18(7):1592–11.
26. Huwa N, Weiergraber OH, Fejzagic AV, Kirsch C, Schaffrath U, Classen T. The Crystal Structure of the Defense Conferring Rice Protein OsJAC1 Reveals a Carbohydrate Binding Site on the Dirigent-like Domain. Biomolecules. 2022;12(8):1126–44.
27. Huwa N, Weiergraber OH, Fejzagic AV, Kirsch C, Schaffrath U, Classen T. The Crystal Structure of the Defense Conferring Rice Protein OsJAC1 Reveals a Carbohydrate Binding Site on the Dirigent-like Domain. Biomolecules. 2022;12(8): 1126–18.
28. Ma QH, Tian B, Li YL. Overexpression of a wheat jasmonate-regulated lectin increases pathogen resistance. Biochimie. 2010;92(2):187–93.

Jiang *et al. BMC Plant Biology*   (2024) 24:636

Page 12 of 12

29. Weidenbach D, Esch L, Moller C, Hensel G, Kumlehn J, Hofle C, Huckel-hoven R, Schaffrath U. Polarized Defense Against Fungal Pathogens Is Mediated by the Jacalin-Related Lectin Domain of Modular Poaceae-Specific Proteins. Mol Plant. 2016;9(4):514–27.

30. Poulton JE. Cyanogenesis in plants. Plant Physiol. 1990;94(2):401–5.

31. Blanchard DJ, Cicek M, Chen J, Esen A. Identification of beta-glucosidase aggregating factor (BGAF) and mapping of BGAF binding regions on Maize beta -glucosidase. J Biol Chem. 2001;276(15):11895–901.

32. Czjzek M, Cicek M, Zamboni V, Burmeister WP, Bevan DR, Henrissat B, Esen A. Crystal structure of a monocotyledon (maize ZMGlu1) beta-glucosi-dase and a model of its complex with p-nitrophenyl beta-D-thiogluco-side. Biochem J. 2001;354(1):37–46.

33. Ranum P, Pena-Rosas JP, Garcia-Casal MN. Global maize production, utilization, and consumption. Ann N Y Acad Sci. 2014;1312:105–12.

34. Guan B, Xue HW. Arabidopsis AUTOPHAGY-RELATED3 (ATG3) facilitates the liquid-liquid phase separation of ATG8e to promote autophagy. Sci Bull (Beijing). 2022;67(4):350–4.

35. Luan QL, Zhu YX, Ma S, Sun Y, Liu XY, Liu M, Balint-Kurti PJ, Wang GF. Maize metacaspases modulate the defense response mediated by the NLR protein Rp1-D21 likely by affecting its subcellular localization. Plant J. 2021;105(1):151–66.

36. Priscila Marcelino Dos Santos Silva WFdO. Priscilla Barbosa Sales Albu-querque, Maria Tereza Dos Santos Correia, Luana Cassandra Breitenbach Barroso Coelho: Insights into anti-pathogenic activities of mannose lectins. Int J Biol Macromol. 2019;140:234–44.

37. Min Song WX, Xiang Y, Jia H, Zhang L, Zhengqiang Ma. Association of jacalin-related lectins with wheat responses to stresses revealed by transcriptional profiling. Plant Mol Biol. 2014;84:95–110.

38. Dufault-Thompson K, Hall B, Jiang X. Taxonomic distribution and evolu-tionary analysis of the equol biosynthesis gene cluster. BMC Genomics. 2022;23(1):182.

39. Wu D, Jiang B, Ye CY, Timko MP, Fan L. Horizontal transfer and evolution of the biosynthetic gene cluster for benzoxazinoids in plants. Plant Com-mun. 2022;3(3):100320.

40. Panchy N, Lehti-Shiu M, Shiu SH. Evolution of Gene Duplication in Plants. Plant Physiol. 2016;171(4):2294–316.

41. Sankoff D, Zheng C. Whole Genome Duplication in Plants: Implications for Evolutionary Analysis. Methods Mol Biol. 2018;1704:291–315.

42. Klionsky DJ, Abdalla FC, Abeliovich H, Abraham RT, Acevedo-Arozena A, Adeli K, Agholme L, Agnello M, Agostinis P, Aguirre-Ghiso JA, et al. Guide-lines for the use and interpretation of assays for monitoring autophagy. Autophagy. 2012;8(4):445–544.

43. Sun Y, Ma S, Liu X, Wang GF. The maize ZmVPS23-like protein relocates the nucleotide-binding leucine-rich repeat protein Rp1-D21 to endosomes and suppresses the defense response. Plant Cell. 2023;35(6):2369–90.

44. Raval S, Gowda SB, Singh DD, Chandra NR. A database analysis of jacalin-like lectins: sequence–structure–function relationships. Glycobiology. 2004;14(12):1247–63.

45. Barre A, Bourne Y, Van Damme EJ, Peumans WJ, Rougé P. Mannose-binding plant lectins: different structural scaffolds for a common sugar-recognition process. Biochimie. 2001;83(7):645–51.

46. Gasper R, Effenberger I, Kolesinski P, Terlecka B, Hofmann E, Schaller A. Dirigent Protein Mode of Action Revealed by the Crystal Structure of AtDIR6. Plant Physiol. 2016;172(4):2165–75.

47. Zhang W, Peumans WJ, Barre A, Astoul CH, Rovira P, Rouge P, Proost P, Truffa-Bachi P, Jalali AA, Van Damme EJ. Isolation and characterization of a jacalin-related mannose-binding lectin from salt-stressed rice (Oryza sativa) plants. Planta. 2000;210(6):970–8.

48. Mistry J, Finn RD, Eddy SR, Bateman A, Punta M. Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. Nucleic Acids Res. 2013;41(12):e121.

49. Cheng Y, Jiang S, Zhang X, He H, Liu J. Whole-Genome Re-Sequencing of Corylus heterophylla Blank-Nut Mutants Reveals Sequence Variations in Genes Associated With Embryo Abortion. Front Plant Sci. 2019;10:1465.

50. Crooks GE, Hon G, Chandonia JM, Brenner SE. WebLogo: a sequence logo generator. Genome Res. 2004;14(6):1188–90.

51. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sen-sitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 1994;22(22):4673–80.

52. Tamura K, Stecher G, Kumar S. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. Mol Biol Evol. 2021;38(7):3022–7.

53. Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, Xia R. TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. Mol Plant. 2020;13(8):1194–202.

54. Wang Y, Tang H, Debarry JD, Tan X, Li J, Wang X, Lee TH, Jin H, Marler B, Guo H, et al. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res. 2012;40(7):e49.

55. Yang T, Guo L, Ji C, Wang H, Wang J, Zheng X, Xiao Q, Wu Y. The B3 domain-containing transcription factor ZmABI19 coordinates expression of key factors required for maize seed development and grain filling. Plant Cell. 2021;33(1):104–28.

56. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods. 2001;25(4):402–8.

57. Singh VK, Mangalam AK, Dwivedi S, Naik S. Primer premier: program for design of degenerate primers from a protein sequence. Biotechniques. 1998;24(2):318–9.

58. Lalitha S. Primer Premier 5. Biotech Software & Internet Report. 2000;1(6):270–2.

59. Debray H, Decout D, Strecker G, Spik G, Montreuil J. Specificity of twelve lectins towards oligosaccharides and glycopeptides related to N-glycosylproteins. Eur J Biochem. 1981;117(1):41–55.

Publisher's Note