Research article

Genome-wide characterization and comparative analysis of the *MLO* gene family in cotton

Xiaoyan Wang [a,1], Qifeng Ma [b,1], Lingling Dou [b], Zhen Liu [a], Renhai Peng [a,*], Shuxun Yu [b,**]

[a] *Anyang Institute of Technology, College of Biology and Food Engineering, Anyang, Henan, 455000, PR China*
[b] *State Key Laboratory of Cotton Biology, Institute of Cotton Research of CAAS, Anyang, Henan, 455000, PR China*

ARTICLE INFO

ABSTRACT

In plants, *MLO* (*Mildew Locus O*) gene encodes a plant-specific seven transmembrane (TM) domain protein involved in several cellular processes, including susceptibility to powdery mildew (PM). In this study, a genome-wide characterization of the *MLO* gene family in G. raimondii L., G. arboreum L. and G. hirsutum L. was performed. In total, 22, 17 and 38 homologous sequences were identified for each species, respectively. Gene organization, including chromosomal location, gene clustering and gene duplication, was investigated. Homologues related to PM susceptibility in upland cotton were inferred by phylogenetic relationships with functionally characterized MLO proteins. To conduct a comparative analysis between *MLO* candidate genes from G. raimondii L., G. arboreum L. and G. hirsutum L., orthologous relationships and conserved synteny blocks were constructed. The transcriptional variation of 38 *GhMLO* genes in response to exogenous application of salt, mannitol (Man), abscisic acid (ABA), ethylene (ETH), jasmonic acid (JA) and salicylic acid (SA) was monitored. Further studies should be conducted to elucidate the functions of *MLO* genes in PM susceptibility and phytohormone signalling pathways.

1. Introduction

Powdery mildew (PM) is an obligate fungal pathogen that causes PM disease in a broad range of plants, including important crops such as pepper, tomato, apple, strawberry, and cotton (Glawe, 2008). It is difficult to diagnose at the early stages of the disease, and it can easily spread unnoticed. According to previous observations, PM disease primarily affects the leaves of sea-island cotton and upland cotton. In general, PM presents similar symptoms in cotton: white or brown spots on leaf tissues, particularly at the bottom of the plant, whereas upper leaves exert some resistance. Afterwards, tissue death of diseased spots causes infected leaves to crinkle, curl, and prematurely drop. Although blossoms and fruits are not the initial PM fungal targets, they can also become infected.

Mildew locus O (MLO) proteins belong to a plant-specific protein family containing seven transmembrane (TM) domains (Buschges et al., 1997; Devoto et al., 1999). In addition, a C-terminal calmodulin-binding domain (CaMBD) and an extracellular N-terminus (Devoto et al., 1999; Kim et al., 2002a,b) have been identified in this family. PM resistance was first characterized in barley plants in 1942 and the immunity was acquired because of the absence of a susceptibility gene which was named *Mildew Locus O* (*MLO*). Recessive *MLO* gene mutations confer durable broad-spectrum resistance to all discovered isolates of barley powdery mildew fungus *Blumeria graminis* f. sp *hordei* (Bgh) (Buschges et al., 1997; Devoto et al., 1999). Then, the discovery and identification of PM disease resistance in other plant species, such as *Arabidopsis* (Consonni et al., 2006), pea (Pavan et al., 2011) and tomato (Bai et al., 2008), has confirmed that PM resistance deriving from loss-of-function mutations in *MLO* functional orthologue is a common phenomenon. Therefore, broad-spectrum PM resistance in plants could be introduced by silencing of *MLO* gene (Pavan et al., 2010).

Calmodulin-binding of MLO proteins promotes PM susceptibility in barley (Kim et al., 2002a,b). Moreover, pharmacological studies have suggested that the influx of Ca^{2+} ions is important for

* Corresponding author.
** Corresponding author.
E-mail addresses: wxybeok@163.com (X. Wang), 13837240176@163.com (Q. Ma), xjyldll@126.com (L. Dou), liuzhen378@163.com (Z. Liu), aydxprh@163.com (R. Peng), ysx195311@163.com (S. Yu).
[1] These authors contributed equally to this work.

MLO protein function (Kim et al., 2002a,b). Therefore, Ca^{2+} may be a candidate signal because plant cells generate a transient Ca^{2+} signal in response to pathogen attack (Xu and Heath, 1998). In addition, a complex mechanism may exist during the interaction between MLO genes and PM. Nevertheless, there is limited information on the precise mechanism of MLO proteins. It has been revealed that PM fungi target MLO proteins as an access to trigger pathogenesis because vesicle-associated and actin-dependent defence pathways are negatively regulated by functional MLO proteins in the circumstance of attempted PM penetration. Studies in tomato (Bai et al., 2008), barley (Piffanelli et al., 2002), pepper (Zheng et al., 2013a,b), and grape (Feechan et al., 2009) confirmed that early stages of PM infection are associated with up-regulated expression of MLO susceptibility-related gene, with a peak at 6 h after inoculation.

Since the HvMLO gene was first identified in barley (Buschges et al., 1997), MLO genes have been discovered in Arabidopsis thaliana (Devoto et al., 2003), Oryza sativa (Liu and Zhu, 2008), Vitis vinifera (Angela Feechan et al., 2008), Triticum aestivum (Konishi et al., 2010), Glycine max (Deshmukh et al., 2014), Cucumis sativus (Zhou et al., 2013), Malus domestica (Pessina et al., 2014) and Solanum lycopersicum (Chen and Zhang, 2014). More detailed studies have uncovered that medium-sized gene family of MLO is plant-specific and MLO-based PM resistance is not confined to monocotyledones, but is also discovered in distantly related dicotyledones (Acevedo-Garcia et al., 2014). For example, mutant alleles of AtMLO2, one of 15 MLO genes present in A. thaliana, caused partial resistance to the adapted strains such as Golovinomyces orontii and G. cichoracearum. Complete PM resistance was produced when two other homologous genes AtMLO6 and AtMLO12 were also mutated (Consonni et al., 2006). Subsequently, two studies showed that loss-of-function of the SlMLO1 gene was the cause of resistance to PM disease in tomato (Bai et al., 2008; Zheng et al., 2013a,b). It was demonstrated that pea PM resistance was associated with loss-of-function mutations in an MLO-homologous locus (Pavan et al., 2011). The results of virus-induced gene silencing suggested that both CaMLO1 and CaMLO2 are involved in the susceptibility of pepper to the PM fungus Leveillula taurica (Zheng et al., 2013a,b). Recently, NtMLO1, which is predicted to be an orthologue of tomato SlMLO1 and pepper CaMLO2, was shown to be involved in PM susceptibility (Appiano et al., 2015).

Apart from susceptibility/resistance to PM disease in both monocotyledonous and dicotyledonous plants, increasing reports have suggested that MLO may be involved in a variety of developmental processes. Leaf mesophyll cells in MLO barley mutants have been shown to undergo spontaneous cell death, which is an indication of accelerated leaf senescence (Buschges et al., 1997; Piffanelli et al., 2002). MLO family members in Arabidopsis presented tissue-specific expression patterns and silencing of AtMLO7 involved in pollen tube reception by the embryo sac led to decreased fertility (Zhongying Chen et al., 2006). Two additional Arabidopsis genes, AtMLO4 and AtMLO11, control root architecture, as null mutants generate asymmetrical root growth and exaggerated curvature (Chen et al., 2009). More results have revealed that MLO family members are involved in diverse abiotic stresses for Capsicum annuum CaMLO2 intensely induced upon exogenous treatment of pepper leaves with the phytohormone abscisic acid (ABA) and drought stress, is shown to act as a suppressor of ABA signalling to prevent water loss from leaves under drought conditions (Lim and Lee, 2014).

MLO genes have been intensively studied in many monocots and dicots, but very little research has focused on cotton. Published genomic data on G. raimondii (DD; 2n = 26) (Paterson et al., 2012), and G. arboretum (AA; 2n = 26) (Li et al., 2014) as well as Gossypium hirsutum (AADD; 2n = 52) (Li et al., 2015) provide an opportunity to conduct a comprehensive overview of the MLO gene family in diploid and tetraploid cotton species. In this study, we characterized the MLO gene family in these three species with respect to their structural, genomic and gene-expression features. Moreover, we assessed the orthologous relationships between the G. raimondii L., G. arboreum L. and G. hirsutum L. genomes.

2. Materials and methods

2.1. In silico identification and annotation

Genomic databases of G. raimondii L. (D5, JGI__v2.1), G. arboreum L. (A2, BGI _v1.0) and G. hirsutum L. (AD1, BGI _v1.0), available at the CottonGen website (https://www.cottongen.org/) (Yu et al., 2014), were downloaded for the identification of.

MLO homologue nucleotide and protein sequences. Then, several local BLAST searches using the Arabidopsis AtMLO1 amino acid sequence as a query were performed. Candidates with an E-value less than $1.0e^{-20}$ were estimated to be MLO homologues, and their gene coding regions, genomic DNA and deduced amino acid sequences were acquired. Conserved MLO domains within the acquired MLO sequences were confirmed by searching NCBI's conserved domain database (http://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi) and Pfam's protein domain families (http://pfam.xfam.org/). The presence and number of TM helices in the proteins of interest were predicted using the online software TMHMM (http://www.cbs.dtu.dk/services/TMHMM-2.0/).

2.2. Gene organization

The chromosomal localization of each MLO gene in G. raimondii L., G. arboreum L. and G. hirsutum L. was deduced based on the available genomic information at the CottonGen database. Mapchart 2.2 software was used to visualize the distribution of MLO genes on the chromosomes, with the exception of a small portion of genes that have not been localized to a chromosome (Voorrips, 2002). Introns and exons of each MLO gene were determined by comparing the cDNAs with their corresponding genomic DNA sequences. Intron/exon composition and position were analysed by the Gene Structure Display Server (GSDS) tool (http://gsds.cbi.pku.edu.cn/).

The MLOs of G. hirsutum were first aligned by Clustal W2 at EMBL-EBI (http://www.ebi.ac.uk/Tools/msa/clustalw2/). Gene duplication events were identified when the following conditions were fulfilled: (1) the alignment covered more than 80% of the longer gene, (2) the identity of the aligned regions was greater than 80% of the alienable region, and (3) only one duplication event was taken into account for tightly linked genes (Gu et al., 2002).

2.3. Phylogenetic analysis

A total of 38 GhMLO amino acid sequences together with 36 other MLO homologues from 9 dicot and monocot species were used to construct phylogenetic trees. Amino acid sequences of 36 known MLOs from A. thaliana (Devoto et al., 2003), Hordeum vulgare (Buschges et al., 1997), O. sativa (Liu and Zhu, 2008), Zea mays (Devoto et al., 2003), S. lycopersicum (Chen and Zhang, 2014), Pisum sativum (Pavan et al., 2011), V. vinifera (Angela Feechan et al., 2008), Malus domestica (Pessina et al., 2014) and C. annuum (Zheng et al., 2013a,b) were obtained based on the published information. A total of 74 MLO protein sequences were included to perform multiple alignments using ClustalW (Thompson et al., 1994) with the default parameters. A neighbour-joining phylogenetic tree was constructed by MEGA 6.0 software (Tamura et al., 2013) with the pairwise deletion option and Poisson correction model. Bootstrapping (1000