Isolation and characterization of a new mannose-binding lectin gene from *Taxus media*

GUOYIN KAI[a], LINGXIA ZHAO[a], JINGUI ZHENG[b], LEI ZHANG[c,d], ZHIQI MIAO[a], XIAOFEN SUN[d] and KEXUAN TANG[a,d]*

[a]*Plant Biotechnology Research Center, School of Agriculture and Biology, Fudan-SJTU-Nottingham Plant Biotechnology R and D Center, School of Life Sciences and Technology, Shanghai Jiao Tong University, Shanghai 200030, People's Republic of China*

[b]*Biotechnology Center, Fujian Agricultural and Forest University, Fuzhou 350002, People's Republic of China*

[c]*School of Pharmacy, Second Military Medical University, Shanghai 200433, People's Republic of China*

[d]*State Key Laboratory of Genetic Engineering, Fudan-SJTU-Nottingham Plant Biotechnology R and D Center, School of Life Sciences, Morgan-Tan International Center for Life Sciences, Fudan University, Shanghai 200433, People's Republic of China*

**Corresponding author (Fax, 86-21-62824073; Email, kxtangl@yahoo.com)*

In this paper, we report the cloning and characterization of the first mannose-binding lectin gene from a gymnosperm plant species, *Taxus media*. The full-length cDNA of *T. media* agglutinin (TMA) consisted of 676 bp and contained a 432 bp open reading frame (ORF) encoding a 144 amino acid protein. Comparative analysis showed that TMA had high homology with many previously reported plant mannose-binding lectins and that *tma* encoded a precursor lectin with a 26-aa signal peptide. Molecular modelling revealed that TMA was a new mannose-binding lectin with three typical mannose-binding boxes like lectins from species of angiosperms. Tissue expression pattern analyses revealed that *tma* is expressed in a tissue-specific manner in leaves and stems, but not in fruits and roots. Phylogenetic tree analyses showed that TMA belonged to the structurally and evolutionarily closely related monocot mannose-binding lectin superfamily. This study provides useful information to understand the molecular evolution of plant lectins.

[Kai G, Zhao L, Zheng J, Zhang L, Miao Z, Sun X and Tang K 2004 Isolation and characterization of a new mannose-binding lectin gene from *Taxus media*; *J. Biosci.* **29** 399–407]

1. Introduction

Plant lectins or agglutinins, widespread throughout the plant kingdom, are considered as an extended group of proteins which, according to a recently updated definition, comprise all plant proteins possessing at least one non-catalytic domain that binds reversibly to specific mono- or oligosaccharides (Van Damme *et al* 1998). On the basis of recent advances in biochemistry and molecular cloning and structural analysis, virtually all currently known plant

Keywords. Evolution analysis; mannose-binding lectin; molecular modelling; *Taxus media*; TMA

Abbreviations used: AAA, *Allium ascalonicum* agglutinin; ACA, *Allium cepa* agglutinin; ACPA, *Ananas comosus* (pineapple) agglutinin; AKA, *Amorphophallus konjac* agglutinin; APA, *Allium porrum* agglutinin; ASA, *Allium sativum* agglutinin; AUA, *Allium ursinum* agglutinin; AVA, *Amaryllis vittata* agglutinin; CAA, *Crinum asiaticum* agglutinin; CTAB, cetyltrimethylammonium bromide; EHA, *Epipactis helleborine* agglutinin; GEAFP, *Gastrodia elata* antifungal protein; GNA, *Galanthus nivalis* agglutinin; LOA, *Listera ovata* agglutinin; ORF, open reading frame; PCR, polymerase chain reaction; pI, isoelectric point; RACE, rapid amplification of cDNA ends; RT-PCR, reverse transcriptase-polymerase chain reaction; TDA, *Typhonium divaricatum* agglutinin; TMA, *Taxus media* agglutinin; UTR, untrslated region; ZCA, *Zephyrathes candida* agglutinin; ZGA, *Zephyrathes grandiflora* agglutinin.

lectins can be classified into seven families of structurally and evolutionarily-related proteins (Van Damme *et al* 1998), and that most of these belong to four groups of evolutionarily related proteins: legume lectins, chitin-binding lectins, monocot mannose-binding lectins, and type 2 ribosome-inactivating proteins, among which monocot mannose-binding lectins have attracted greater attention because of their great application values in biological and biomedical research; such as in the isolation of mannose-containing glyconjugates, and in their potent inhibitory effect on human and animal retroviruses, including HIV (Barre *et al* 1996). In addition, most monocot mannose-binding lectins play an important role in the plant's defense against different kinds of plant-eating organisms; which is due to their recognition of high-mannose type glycans of plant predators (Powell *et al* 1993; Peumans and Van Damme 1995; Tang *et al* 2001; Sun *et al* 2002; Lin *et al* 2003; Yao *et al* 2003a). Up to now, monocot mannose-binding lectins have been cloned from seven families of angiosperms including Amaryllidaceae, Araceae, Alliaceae, Orchidaceae, Liliaceae, Iridaceae and Bromeliaceae (Van Damme *et al* 1991, 1994, 1996, 2000; Smeets *et al* 1997; Neuteboom *et al* 2002; Chai *et al* 2003; Fei *et al* 2003; Kai *et al* 2003; Yao *et al* 2003b; Zhao *et al* 2003a,b), among which lectins from Amaryllidaceae species have been extensively studied and well documented (Van Damme *et al* 1992, 1998; Kai *et al* 2003; Pang *et al* 2003; Zhao *et al* 2003a).

However, little is known about the possible origin and molecular evolution of the carbohydrate-binding domains of plant lectins from modern flowering plants with little gene sequence information about plant lectins outside higher flowering plants (Peumans *et al* 2002). Until now, there have been no reports on molecular cloning of lectin genes from gymnosperms, including family Taxaceae, therefore the molecular evolution relationship of the plant lectins between flowering plants and gymnosperms is unclear. To investigate whether similar mannose-binding lectins also exist in gymnosperms such as Taxaceae, we employed a relatively simple cloning strategy (rapid amplification of cDNA ends, RACE) to successfully isolate a novel lectin gene from *Taxus media*, which did not belong to any previously reported families containing mannose-binding lectins in the classical taxonomical range above. *T. media* contains high content of anti-cancer drug Taxol and is a kind of valuable medicinal plant of family Taxaceae in gymnosperms. Up to now, there is no report on the molecular cloning of lectin gene from *T. media*.

In this paper, we report for the first time the molecular cloning and characterization of a novel mannose-binding lectin gene from *T. media*. The cloned gene is also the first mannose-binding lectin gene isolated from gymnosperm species. The evolutionary relationship among lectins from gymnosperm and angiosperm species was also studied.

2. Materials and methods

2.1 *Plant material*

T. media plants, provided by Professor Feng Tan at Southwest Normal University in China, were grown in pots in the greenhouse under standard conditions. All tissue materials (1 g) including leaves, fruits, stems and roots from *T. media* plants were frozen separately in liquid nitrogen for RNA extraction.

2.2 *Total RNA isolation*

Because of the high concentration of secondary products such as phenolic components and polysaccharides in *T. media* tissues, the cetyltrimethylammonium bromide (CTAB) based RNA isolation method (Jaakola *et al* 2001) was adopted in the total RNA extraction. The quality and concentration of the extracted RNA were checked by agarose gel electrophoresis and spectrophotometer (DU-640, Beckman, USA) analysis, respectively. The RNA samples were stored at – 80°C prior to RACE and reverse transcriptase-polymerase chain reaction (RT-PCR) analysis.

2.3 *Molecular cloning of* tma *full-length cDNA*

Total RNA (5 μg) was used to synthesize the first strand cDNA (3′-ready cDNA) using SMART™ RACE cDNA Amplification Kit (Clontech Laboratories Inc., USA) with the 3′ RACE CDS Primer A (5′-AAGCAGTGGTATCAACGCAGAGTAC $(T)_{30}N_{-1}N$-3′) provided by the kit. The 3′ RACE was performed with the Universal Primer A Mix (UPM, long: 5′-CTAATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGT; short: 5′-CTAATACGACTCACTATAGGGC-3′) as the reverse primer and degenerate primer (F2) (5′-ATGCAG(C/G/T)A(G/C)GAC(T/G)GCAACCTG-3′) as the forward primer that was designed and synthesized corresponding to the conserved regions [MQ(G/E/Y/D)D(C/G)NL] of previously reported plant lectins (Kai *et al* 2003; Pang *et al* 2003; Zhao *et al* 2003a). PCR was conducted in a total volume of 50 μl (containing 2 μl cDNA 20 μmol of F2 20 μmol of UPM, 10 μmol dNTPs, 1 × Ex PCR buffer and 5U Ex *Taq* Polymerase) following the protocol: cDNA was denatured at 94°C for 3 min followed by 35 cycles of amplification (94°C for 30 s, 58°C for 30 s, 72°C for 90 s) and by 10 min at 72°C. The amplified product was purified and cloned into pGEM-T Easy cloning vector (Promega, WI, USA), and transformed into *Escherichria coli* DH5*a*. Based on the colour reaction using Xgal-IPTG System and PCR identification, the positive clones were picked out and sequenced by ABI 377 Sequencer (Perkin-Elmer, USA).

The first strand cDNA (5′-ready cDNA) synthesis in 5′ RACE was performed according to the manual of the SMART™ RACE cDNA Amplification Kit (Clontech Laboratories Inc., USA) using the 5′-RACE CDS Primer [5′-$(T)_{25}N_{-1}N$-3′] provided by the kit. Based on the sequence of the 3′ RACE product, the complementary reverse gene specific primer R2 (5′-GCCGTCATTCTGCATACTCAAC-3′) was designed and synthesized. The 5′ RACE-PCR was carried out using primers R2 and UPM under the following condition: the template (the 5′-ready cDNA) was denatured at 94°C for 2 min followed by 35 cycles of amplification (94°C for 50 s, 56°C for 50 s and 72°C for 90 s) and by 7 min at 72°C. The PCR product was purified and cloned into the pGEM-T Easy cloning vector (Promega, WI, USA) followed by sequencing.

By assembling the sequences of the 3′ RACE and 5′ RACE products, the full-length cDNA sequence of *tma* was deduced, and was subsequently amplified via PCR using a pair of primers F1 (5′-GGCAATTCTCAATCG-CCATTATA-3′) and UPM, which was repeated for three times. The amplified full-length cDNA of *tma* was used for molecular characterization such as sequence homology, signal peptide prediction, presentation of conserved motifs, molecular modelling and phylogenetic tree reconstruction.

2.4 *Computer analyses*

Bioinformatic-associated analyses of TMA were performed using programs of PSI-Blast, SOPMA, CLUSTAL W1·82 and Swiss-Model. TMA and other plant lectins retrieved from GenBank were aligned with CLUSTAL W (1·82) using default parameters and the phylogenetic tree was constructed using neighbour-joining method (Thompson *et al* 1994).

2.5 *Tissue expression analyses*

In order to investigate the expression pattern of *tma* in different tissues including leaves, fruits, stems and roots of *T. media*, semi-quantitative One-step RT-PCR was carried out according to the manufacturer' instruction (Takara, Japan). Aliquots of total RNA (0·5 μg) extracted from leaves, fruits, stems and roots of *T. media*, individually were used as templates in one-step RT-PCR reaction with the forward primer KF1 (5′-ATGGGAGAATCATCTG-TAAC-3′) and reverse primer KR1 (5′-TCAGAGAAG-TTTTTTCTTGTT-3′) specific to coding sequence of *TMA* using one-step RNA PCR kit. Meanwhile, the RT-PCR reaction for the housekeeping gene (actin gene) using specific primers actF (5′-GTGACAATGGAACTGGAA-TGG-3′) and actR (5′-AGACGGAGGATAGCGTGAGG-3′) designed according to the conserved regions of plant actin genes was performed as an internal control to estimate whether equal amounts of RNA among samples were used in semi-quantitative RT-PCR (Kai *et al* 2004). Amplifications were performed under the following condition: 50°C for 30 min and 94°C for 2 min followed by 25 cycles of amplification (94°C for 30 s, 60°C for 30 s and 72°C for 2 min). The amplified products were separated on a 1% agarose gel and analysed with Gene analysis software package (Gene company, USA).

3. Results

3.1 *Cloning of the full-length cDNA of* tma

Based on sequences of the conserved regions of mannose-binding lectins from other plant species, a degenerate primer (F2) was designed and used for the amplification of 3′-end of *T. media* agglutinin cDNA (*tma*). A single fragment of about 480 bp was amplified using primers F2 and UPM, and a 3′ untranslated region (UTR) of 109 bp was found downstream from the stop codon in the amplified sequence. Complementary reverse-specific primer R2 was designed and used for the cloning of the 5′-end sequence of *tma* by 5′ RACE based on the sequence of 3′ RACE fragment. A 400 bp fragment was amplified in which a 5′ UTR of 62 bp was found upstream of the first ATG codon. Based on the sequences of the 3′ and 5′ RACE products, the full-length cDNA fragment was deduced and amplified by PCR using a pair of primers F1 and UPM. The cloned full-length cDNA of *tma* was 676 bp and contained a 432 bp ORF encoding a protein of 144 amino acids, with an isoelectric point of 9·48 and calculated molecular weight of 15,166 Da (figure 1). The richest amino acid in the deduced TMA was Ala (13% by frequency), followed by Leu (11%), Ser (10%), Val (8%), Gly (8%), and Asn (7%). Acidic and basic amino acids constituted 4% and 9% of the polypeptide, respectively. Nineteen per cent of the total amino acids were charged and the per centages of polar and hydrophobic amino acids were 34% and 40%, respectively.

3.2 *Sequence analyses of TMA*

Sequence comparison by performing Blast Search in GenBank database (http://www.ncbi.nih.gov) revealed that TMA had high homology with many other mannose-binding lectins such as *Zephyrathes grandiflora* agglutinin (ZGA), *Zephyrathes candida* agglutinin (ZCA), *Epipactis helleborine* agglutinin (EHA), *Listera ovata* agglutinin (LOA), *Amaryllis vittata* agglutinin (AVA), *Galanthus nivalis* agglutinin (GNA) and *Crinum asiaticum* agglutinin (CAA), suggesting that TMA belonged to the monocot mannose-binding lectin superfamily. Furthermore, TMA also showed high homology to *Gastrodia elata* antifungal protein (GEAFP). On the amino acid level, TMA was 50%,

51%, 50%, 50%, 49%, 46%, 47% and 49% identical to ZGA, ZCA, EHA, LOA, AVA, GNA, GEAFP and CAA, respectively, and was 61%, 60%, 63%, 66%, 61%, 61%, 59% and 60% similar to ZGA, ZCA, EHA, LOA, AVA, GNA, GEAFP and CAA respectively (figure 2). Signal peptide of TMA was predicted by the signal peptide prediction software (http://www.cbs.dtu.dk/services/signalP)

```
  1                              ggcaattctcaatcgcc
 18 attatatcagatcagcagcttagcctcgctaacagatcaaacata
 63 atgggagaatcatctgtaactacagccgccattgtagcgctggta
    M  G  E  S  S  V  T  T  A  A  I  V  A  L  V
108 atattacttacgtttgcaaatccttgctcatcaaaatctgttctc
    I  L  L  T  F  A  N  P  C  S  S  K  S  V  L
153 aaaagcggggaatcactggcggctggagagtctctgcaatacgct
    K  S  G  E  S  L  A  A  G  E  S  L  Q  Y  A
198 caatatatattggtaatgcaaggcgactgcaacctggttttgtat
    Q  Y  I  L  V  M  Q  G  D  C  N  L  V  L  Y
243 gcaaacaaggtgaaagtgttgtgggcgtcaaggacaaatggaaaa
    A  N  K  V  K  V  L  W  A  S  R  T  N  G  K
288 ggcggcgccgcctcctgcaagttgagtatgcagaatgacggcaat
    G  G  A  A  S  C  K  L  S  M  Q  N  D  G  N
333 ttggtgatatatgcggcgaccacacctgtatgggctagtagaacc
    L  V  I  Y  A  A  T  T  P  V  W  A  S  R  T
378 tccagagcctttgcttcctacaagctcaatcttcagggcgatggt
    S  R  A  F  A  S  Y  K  L  N  L  Q  G  D  G
423 aatgttgttatttatgggccgtctggagctatttgggctactaat
    N  V  V  I  Y  G  P  S  G  A  I  W  A  T  N
468 acagctcagaacaagaaaaaacttctctgaactttctggtggggg
    T  A  Q  N  K  K  K  L  L  *
513 aacttgtttttagatactacaggagttaacgatccagtaacttg
558 gcacacttcttgacgtgttaataataggaagatacaatcttcttc
603 cgttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa
648 aaaaaaaaaaaaaaaaaaaaaaaaaaaaa
```

Figure 1. The full-length cDNA sequence and deduced amino acid sequence of *T. media* agglutinin (TMA). The start codon (atg) is in italics and the stop codon (tga) is in italics and underlined. The conserved mannose-binding sites (QDNY) are underlined. Asterisk mark (*) represents that stop codon can not be translated into any amino acid.

and the result showed that *tma* encoded a precursor lectin with a signal peptide with the cleavage site being between Ser^{26} (S) and Lys^{27} (K) (before KSVL), which was very similar to some mannose-binding lectins from angiosperm species such as *G. nivalis*, and *Z. grandiflora* (Van Damme *et al* 1992; Kai *et al* 2003).

NCBI conserved domain search results suggested that TMA belonged to mannose-binding lectin superfamily and contained consensus sequence motif (QDNY) which was involved in alpha-D-mannose recognition (figure 2). The conserved motifs were often conservative during evolution, while some variations on un-conserved domain can form the molecular foundation for the diversity of the structures and functions.

The secondary structure of TMA was analysed by SOPMA (Geourjon and Deléage 1995) and the result showed that the putative TMA peptide contained 21% of alpha helices, 32% of extended strands, 9% of beta turns and 38% of random coils (figure 3A). The random coil and extended strands constituted interlaced domain of the main part of the secondary structure.

3.3 *Molecular modelling of TMA*

Swiss-Model structure prediction for the three-dimensional structure (Guex and Peitisch 1997) revealed that TMA had similar structure to those of many other mannose-binding lectins and also had three specific conservative mannose-binding sites with the same amino acid sequences (QDNY) symmetrically arranged on TMA molecular (figure 3B), in consistent with many other mannose-binding lectins, such as ZGA, ZCA and GNA (Van Damme *et al* 1991; Kai *et al* 2003; Pang *et al* 2003; Zhao *et al* 2003a,b), indicating TMA is a mannose-specific lectin. The three conserved mannose-binding sites may play an important role in sugar-binding activity and thus are preserved in evolution. As shown in figure 3B, it was very noteworthy that ***b***-sheets occurred predominantly and the overall folding was typically built from ***b***-sheets connected by turns and loops, which created very tight structural scaffolds. Molecular modelling provides new evidence for an evolutionary relationship among the plant lectins. Purification of TMA protein from *T. media* and X-ray crystal structure analysis will further elucidate the detailed structure of TMA in the future.

3.4 *Molecular evolution analyses*

To investigate the evolutionary relationship among different plant lectins, a phylogenetic tree was constructed based on the deduced amino acid sequences of TMA and other plant lectins. As shown in dendrogram (figure 4), all the monocot mannose-binding lectins were grouped into four

big clusters. The first cluster comprised TMA, ZGA, CAA, LOA, EHA, *Typhonium divaricatum* agglutinin (TDA), GEAFP, *Amorphophallus konjac* agglutinin (AKA) and *Ananas comosus* (pineapple) agglutinin (ACPA), which seemed to be made up of some taxonomically unrelated families such as Taxaceae, Araceae, Amaryllidaceae, Orchidaceae and Bromeliaceae. All the lectins from family Alliaceae including *Allium cepa* agglutinin (ACA), *Allium ascalonicum* agglutinin (AAA), *Allium sativum* agglutinin (ASA), *Allium porrum* agglutinin (APA) and *Allium*

```
ZGA        ---MTTVSINVLLLATTLGLLMTTGARA--ANVLQSGQQLGTGGSLDYAAYNFRMQADCN 55
CAA        -----MARNYILLPAAIFGLLLTS-CMA--ANVLYSGQALSTGASLSFAAYQFVMQADCN 52
LOA        -VSFLSMSTILLVSTMALMSLLATPVSA--SDRLNAGQSLGTGGSLALGGYIFIIQGDCN 57
EHA        MASFLSLRTILLLSTAALMSLLATPVSA--ADRLNAGQSLGTGGSLALGGYIFKIQADCN 58
GEAFP      MAASASTAVILFFAVTTVMSLSAIPAFA--SDRLNSGHQLDTGGSLAQGGYLFIIQNDCN 58
TMA        --MGESSVTTAAIVALVILLTFANPCSS--KSVLKSGESLAAGESLQYAQYILVMQGDCN 56
ZCA        ----MTRITFL-VVAAIFLGVITPSCLSVCDSILYSGETLSAGQSLNNGPYTFIMQEDCN 55
AVA        ----MAKTSFLIVVAAIVLGVITPSCLS--DNILYSGETLSAGESLNYGPYTFIMQQDCN 54
GNA        --------TSLLILATIFLGVITPSCLS--DNIMYSGETLSTGEFLNYGSYVFIMQEDCN 50
Consensus                                      G  L  G  L    Y    Q DCN

ZGA        LVLYDNG-NAVWASGTNGRG--SGCTCRMQYDGNLVVYTNA-NRAVWASNTNRGTG-SYQ 110
CAA        LVLYDNG-NPVWASGTNGRG--SGCQCRMQYDGNLVVYTSG-NRAVWASNTNRGTG-RYQ 107
LOA        LVLYDNN-RAVWASGTNGRG--SNCILSMQRDGNLVIYS-S-GRAIWASNTNRQNG-NYY 111
EHA        LVLYDNN-RAVWASGTNGRA--TDCILSMQRDGNLVIYS-G-SRVIWASNTNRQSG-NYY 112
GEAFP      LVLYDNN-RAVWASGTNGKA--SGCMLKMQNDGNLVIYS-G-SRAIWASNTNRQNG-NYY 112
TMA        LVLYANKVKVLWASRTNGKGGAASCKLSMQNDGNLVIYA-A-TTPVWASRTSRAFA-SYK 113
ZCA        LVLYNVD-KPIWASNTGGLA--RGCHLSMQRDGNLVVYSQR-NRPIWASNTGGHNAANYV 111
AVA        LVLYDVD-KPIWASNTGGLA--RGCHLSMQSDGNLVVYSPSGNRPIWASNTQGENG-NYV 110
GNA        LVLYDVD-KPIWATNTGGLS--RSCYLNMQTDGNLVVYNPS-NKPIWASNTGGQNG-NYV 105
Consensus  LVLY       WA  T G      C   MQ DGNLV Y        WAS T

ZGA        LILQRDRNVVIYDPSGRAIWATGTNVAAAN-VTLSAARPTN--LTITAAKLTNNNNVPAA 167
CAA        CILQRDRNVVVYDPQGRPIWATGTNVAAASNVTLAAAAAAHGPTNVTVAALN-----PAA 162
LOA        LILQRDRNVVIYDNSNNAIWATGTNVGNAAIAVIPHN-------NGTAAASG-------A 157
EHA        LILQRDRNVVIYDNSNNAIWATGTNVGNAAIAVIPHN-------NGTAAASG-------A 158
GEAFP      LILQRDRNVVIYDNSNNAIWATHTNVGNAEITVIPHS-------NGTAAASG-------A 158
TMA        LNLQGDGNVVIYGPSG-AIWATNT------------------------------------ 136
ZCA        LILQKDRNVVIYGPAK---WATGT---YTGAVGISGFETG------TAGIIK-------P 152
AVA        CIVQKDRNVVIYGTAR---WATGT---YTGTVGISGFATG------MAGIIK-------S 151
GNA        CILQKDRNVVIYGPAR---WATGTNIHGAGIVGVPGSAPQN----STAEMIK-------L 151
Consensus     Q D NVV Y       WAT T

ZGA        ATEKASTSNGSNPTQVPVTTAGKN 191
CAA        AKTTAKT-----------TAAGGN 175
LOA        KQNKVRELNP-------------- 167
EHA        QQNKVRELYP-------------- 168
GEAFP      AQNKVNELYISMY----------- 171
TMA        AQNKKKLL---------------- 144
ZCA        VTVKSAK----------------- 159
AVA        VTDKSAK----------------- 158
GNA        VRKYLITK---------------- 159
Consensus
```

Figure 2. Multiple alignment of TMA with other mannose-binding lectins. The completely identical amino acids were indicated with single capital letter. ZGA (AAQ18904); ZCA (AAM94381); EHA (AAC48927); LOA (AAC37423); AVA (AAP57409); GNA (AAA33349); CAA (AAO59506); AVA (AAP57409) and GEAFP (CAB94239). Mannose-binding sites (QDNY) are underlined.

ursinum agglutinin (AUA) were grouped into the second cluster as expected. The third cluster grouped most lectins of Amaryllidaceae family such as *Zephyranthes candida*, *Amaryllis vittata*, *Hippeastrum rutilum*, *Narcissus hybrid*, *Galanthus nivalis*, *Clivia miniata* and *Lycoris radiata*. As expected, the fourth cluster contained most lectins from Araceae species including *Polygonatum cyrtonema*, *Polygonatum multiflorum*, *Arisaema heterophyllum*, *Pinellia ternata* and *Crocus sativus*. All the above four groups of lectins were derived from a common ancestor in evolution, suggesting that these lectins shared a common evolutionary ancestor in plants. It is very surprising and interesting that some non-Orchidaceae lectins such as ZGA and CAA from family Amaryllidaceae are very homologous to the typical Orchidaceae lectins such as LOA and EHA, while TMA from gymnosperm species also has much higher similarity to LOA and EHA than those in other families. The finding that mannose-binding lectin exists in gymnosperms suggests that the mannose-binding lectins are much more widely distributed than previously believed.

3.5 *Tissue expression pattern analyses*

To investigate *tma* expression pattern in various tissues of *T. media*, total RNA was isolated from different tissues including leaves, fruits, stems and roots, and subjected to one-step RT-PCR analysis using the primers KF1 and KR1. The result showed that *tma* is expressed in a tissue-specific manner. *Tma* expression could be detected in leaves and stems, but not in fruits and roots (figure 5). Therefore, the *tma* was considered to be a non-constitutively expressing gene.

4. Discussion

The full-length cDNA of a lectin was successfully isolated from *T. media* and Blast search result showed that *tma* had very low identity to all the other known lectin genes in GenBank at nucleotide level (data not shown). However amino acid sequence comparison revealed that TMA shared comparatively higher similarities to other lectins (figure 2). This result suggested that lectin genes

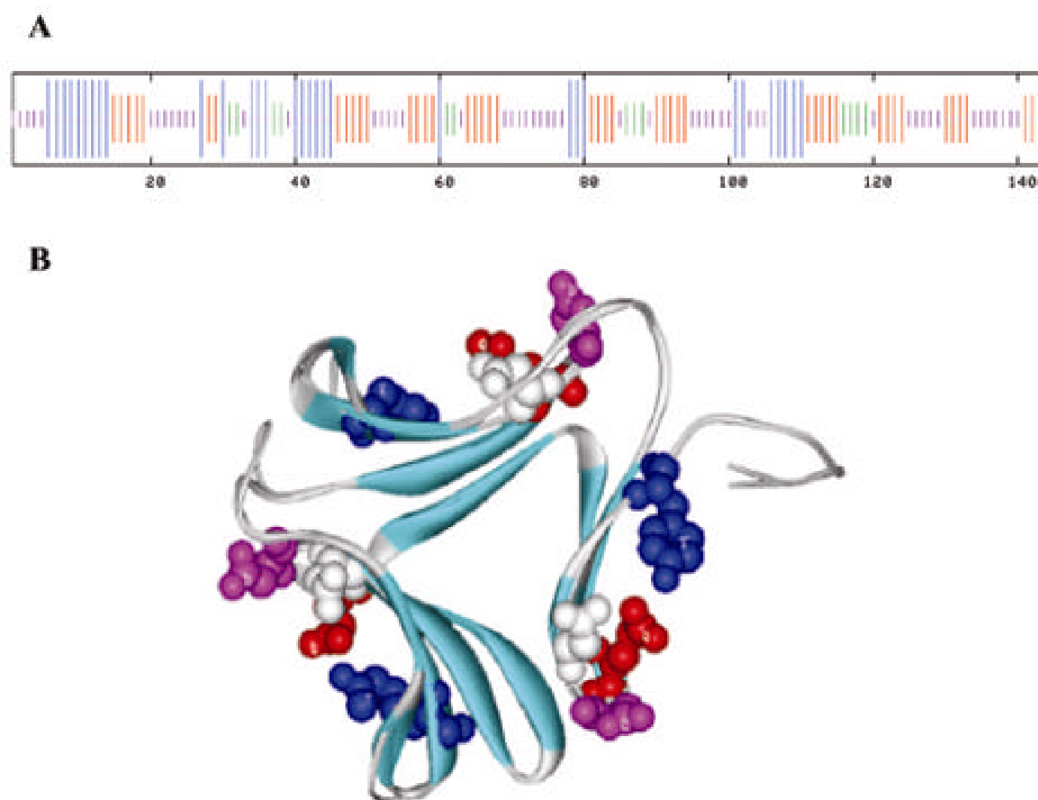

Figure 3. Secondary and three-dimensional structures of TMA. (**A**) The secondary structure of TMA. The helices, sheets, turns and coils are indicated respectively as blue, red, green and purple vertical lines. (**B**) The three-dimensional structure of TMA. ***b***-sheets are indicated as patches. Turns and loops are indicated as lines, and the amino acids constituting mannose-binding sites are indicated as balls.

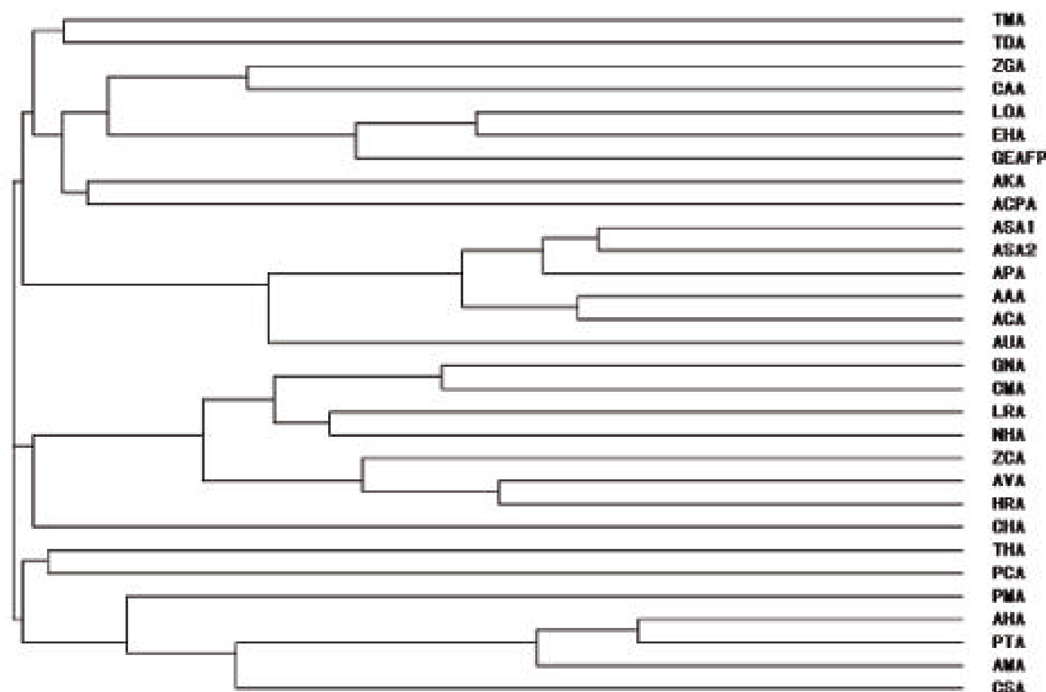

Figure 4. Phylogenetic tree analysis of TMA and other mannose-binding lectins. ZGA (AAQ18904); ZCA (AAM94381); EHA (AAC48927); LOA (AAC37423); AVA (AAP57409); GNA (AAA33349); CAA (AAO59506); GEAFP (CAB94239); *L. radiata* agglutinin (LRA, AAP20877); *Clivia miniata* agglutinin (CMA, AAA19910); *Allium sativum* (garlic) agglutinin (ASA1, AAB64238); *Allium sativum* (garlic) agglutinin (ASA2, AAB64237); AKA (AAP04617); *Tulipa hybrid* agglutinin (THA, AAC49386); APA (AAC37361); TDA (AAQ55289); AAA (S39488); ACA (S39487); *Cymbidium hybrid* agglutinin (CHA, S43463); AUA (AAA16280); *Narcissus hybrid* agglutinin (NHA, AAA33546); *Polygonatum cyrtonema* agglutinin (PCA, AAM28644); ACPA (AAM28277); *Polygonatum multiflorum* agglutinin (PMA, AAC49412); *Hippeastrum rutilum* agglutinin (HRA, AAN73326); *Arisaema heterophyllum* agglutinin (AHA, AAP50524); *Arum maculatum* (cuckoopint) agglutinin (AMA, AAC48998); *Pinellia ternata* agglutinin (PTA, AAP20876); *Crocus sativus* agglutinin (CSA, AAK29077).

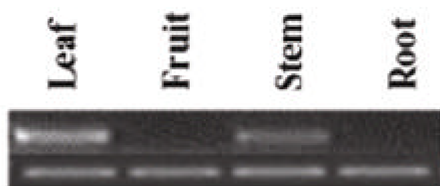

Figure 5. Expression analyses of *tma* in different *T. media* tissues. Total RNA (0·5 μg/sample) was isolated from leaf, fruit, stem and root respectively, and subjected to one-step RT-PCR amplification (upper panel). *Actin* gene was used as the control to show the normalization of the amount of templates in PCR reactions (lower panel).

in different species are very diverse at nucleotide level but their deduced proteins are fairly conservative, especially the conserved motifs, which may play an important role in biological functions and thus are preserved in evolution. Molecular characterization of TMA such as mannose-binding site analysis, signal cleavage site prediction, secondary and three-dimensional structures analysis revealed that TMA had many typical characters possessed by monocot mannose-binding lectins, implying that TMA might have similar functions with many other mannose-binding lectins such as inhibition of fungal growth or insect feeding. The cloning of *T. media* agglutinin gene (*tma*) will also enable us to test its potential functions for pest resistance by transferring the gene into tobacco in the future.

To our knowledge this is the first report of a mannose-binding lectin gene isolated from gymnosperms. Interestingly, molecular evolution analyses result reveal that some non-Orchidaceae lectins from higher plants and even TMA from gymnosperms show more closer evolutionary relationship to typical Orchidaceae lectins than those in other families; implying that Orchidaceae lectins may be much lower in higher flowering plant species which is consistent with previous study (Peuman *et al* 2002). Considering the relatively low similarity between

tma and other mannose-binding lectin genes, the divergence of their respective genes from a common ancestor is not a recent event. This unexpected finding extends the existence of the monocot mannose-binding lectin superfamily to family Taxaceae, a taxonomically unrelated family with those of angiosperms. The identification of TMA for the first time as a new member of the family of monocot mannose-binding lectins not only suggests that these lectins are more widespread among plants than previously believed but also provides new insights into further understanding of the molecular evolution of monocot mannose-binding lectins.

Acknowledgements

We are grateful for anonymous referees' helpful comments on the MS. This work was funded by China National "863" High-Tech Program (No. 2002AA212191), Opening Foundation from Key Laboratory of Medicinal Plant biotechnology of Jiangsu province in China (KJS03080), Shanghai Science and Technology Committee, and UK/CHINA Science and Technology Collaboration Fund.

References

Barre A, Van Damme E J M, Peumans W J and Rouge P 1996 Structure-function relationship of monocot mannose binding lectins; *Plant Physiol.* **112** 1531–1540

Chai Y R, Pang Y Z, Liao Z H, Zhang L, Sun X F, Lu Y Q, Wang S G and Tang K X 2003 Molecular cloning and characterization of a mannose-binding lectin gene from *Crinum asiaticum*; *J. Plant Physiol.* **160** 913–920

Fei J, Liao Z H, Chai Y R, Pang Y Z, Yao J H, Sun X F and Tang K X 2003 Molecular cloning and characterization of a novel mannose-binding lectin gene from *Amorphophallus konjac; Mol. Biol. Rep.* **30** 177–183

Geourjon C and Deléage G 1995 SOPMA: Significant improvement in protein secondary structure prediction by consensus prediction from multiple alignments; *Cabios* **11** 681–684

Guex N and Peitisch M C 1997 SWISS-MODEL and the Swiss-Pdbviewer: An environment for comparative protein modeling; *Electrophoresis* **18** 2714–2723

Jaakola L, Pirtila A M, Halonen M and Hohtola A 2001 Isolation of high quality RNA from bilberry (*Vaccinium myrtillus* L.) fruit; *Mol. Biotechnol.* **19** 210–213

Kai G Y, Miao Z Q, Qiu C X, Zhang L, Zhao L X, Li Z G, Xu T F, Zhang L D, Gong Y F, Zhao D L, Liu D H, Sun X F and Tang K X 2004 Molecular cloning and characterization of a taxadienol acetyl transferase cDNA from *Taxus media*; *Plant Sci.* **167** 759–764

Kai G Y, Zheng J G, Zhang L, Pang Y Z, Liao Z H, Li Z G, Zhao L X, Sun X F and Tang K X 2003 Molecular cloning of a new lectin gene from *Z. grandiflora*; *DNA Sequence* **14** 335–338

Lin J, Yao J H, Zhou X W, Sun X F and Tang K X 2003 Expression and purification of a novel mannose-binding lectin from *Pinellia ternata*; *Mol. Biotechnol.* **25** 215–221

Neuteboom L W, Kunimitsu W Y and Christopher D A 2002 Characterization and tissue-regulated expression of genes involved in pineapple (*Ananas comosus* L.) root development; *Plant Sci.* **163** 1021–1035

Pang Y Z, Shen G A, Liao Z H, Yao J H, Fei J, Sun X F, Tan F and Tang K X 2003 Molecular cloning and characterization of a novel lectin gene from *Zephyranthes candida*; *DNA Sequence* **14** 163–167

Peumans W J, Barre A, Bras J, Rouge P, Proost P and Van Damme E J M 2002 The liverwort contains a lectin that is structurally and evolutionary related to the monocot mannose-binding lectins; *Plant Physiol.* **129** 1054–1065

Peumans W J and Van Damme E J M 1995 Lectins as Plant Defense Proteins; *Plant Physiol.* **109** 347–352

Powell K S, Gatehouse A M R, Hilder V A and Gatehouse J A 1993 Antimetabolic effects of plant lectins and plant and fungal enzymes on the nymphal stages of two important rice pest, *Nilaparvata lugens* and *Nephotettix cinciteps; Entomol. Exp. Applicata* **66** 119–126

Smeets K, Van Damme E J M, Van Leuven F and Peumans W J 1997 Isolation, characterization and molecular cloning of a leaf-specific lectin from ramsons (*Allium ursinum* L.); *Plant Mol. Biol.* **35** 531–535

Sun X F, Wu A Z and Tang K X 2002 Transgenic rice lines with enhanced resistance to the small brown planthopper; *Crop Prot.* **21** 511–514

Tang K X, Zhao Z, Sun X F, Wan B, Qi H X and Lu X G 2001 Production of transgenic rice homozygous lines with enhanced resistance to rice brown planthopper; *Acta Biotechnol.* **21** 117–128

Thompson J D, Higgins D G, Gibson T J and Wolfinger R D 1994 Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice; *Nucleic Acids Res.* **22** 4673–4680

Van Damme E J M, Astoul C H, Barre A, Rouge P and Peumans W J 2000 Cloning and characterization of a monocot mannose-binding lectin from *Crocus vernus* (family: Iridaceae); *Eur. J. Biochem.* **267** 5067–5077

Van Damme E J M, Brike F, Winter H C, Van Leuven F, Goldstein I J and Peumans W J 1996 Molecular cloning of two different mannose-binding lectins from tulip bulbs; *Eur. J. Biochem.* **236** 419–427

Van Damme E J M, Goldstein I J, Vercammen G and Peumans W J 1992 Lectins of members of the *Amaryllidaceae* are encoded by multigene families which show extensive homologies; *Physiol. Plant.* **86** 245–252

Van Damme E J M, Kaku H, Perini F, Goldstein I J, Peeters B, Yagi F, Decock B and Peumans W J 1991 Biosynthesis, primary structure and molecular cloning of snowdrop (*Galanthus nivalis* L.) lectin; *Eur. J. Biochem.* **202** 23–30

Van Damme E J M, Peumans W J, Barre A and Rougé P 1998 Plant Lectins: A Composite of Several Distinct Families of Structurally and Evolutionary Related Proteins with Diverse Biological Roles; *Crit. Rev. Plant Sci.* **17** 575–692

Van Damme E J M, Smeets K, Torrekens S, Van Leuven F and Peumans W J 1994 Characterization and molecular cloning of mannose-binding lectins from the *Orchidaceae* species *Listera ovata*, *Epipactis helleborine* and *Cymbidium hybrid*; *Eur. J. Biochem.* **221** 769–777

Yao J H, Pang Y Z, Qi H X, Wan B L, Zhao X Y, Kong W W, Sun X F and Tang K X 2003a Transgenic tobacco expressing *Pinellia ternata* agglutinin confers enhanced resistance to aphids; *Transgenic Res.* **12** 715–722

Yao J H, Zhao X Y, Zhang H Y, Liao Z H, Chen F, Song J, Sun X F and Tang K X 2003b Molecular cloning and characterization of a novel lectin gene from *Pinellia ternata*; *Cell Res.* **13** 301–308

Zhao X Y, Yao J H, Sun X F and Tang K X 2003a Molecular cloning and characterization of a novel lectin gene from *Lycoris radiata; DNA Sequence* **14** 223–226

Zhao X Y, Yao J H, Liao Z H, Zhang H Y, Chen F, Wang L, Lu Y Q, Sun X F, Yu S Q and Tang K X 2003b Molecular Cloning of a Novel Mannose-binding Lectin Gene from *Arisaema heterophyllum*; *Plant Sci*. **165** 55–60

MS received 22 June 2004; accepted 29 September 2004

*e*Publication: 28 October 2004

Corresponding editor: K MUNIYAPPA