

The *Endo-β-Mannanase* gene families in *Arabidopsis*, rice, and poplar

Joshua S. Yuan · Xiaohan Yang · Jingru Lai ·
Hong Lin · Zong-Ming Cheng · Hiroyuki Nonogaki ·
Feng Chen

Received: 22 March 2006 / Revised: 18 June 2006 / Accepted: 19 June 2006 / Published online: 8 August 2006
© Springer-Verlag 2006

Abstract Mannans are widespread hemicellulosic polysaccharides in plant cell walls. Hydrolysis of the internal β -1,4-D-mannopyranosyl linkage in the backbone of mannans is catalyzed by *endo*- β -mannanase. Plant *endo*- β -mannanase has been well studied for its function in seed germination. Its involvement in other plant biological processes, however, remains poorly characterized or elusive. The completed genome sequences of *Arabidopsis* (*Arabidopsis thaliana*), rice (*Oryza sativa*), and poplar (*Populus trichocarpa*) provide an opportunity to conduct comparative genomic analysis of *endo*- β -mannanase genes in these three species. *In silico* sequence analysis led to the identification of eight, nine and 11 *endo*- β -mannanase genes in the genomes of *Arabidopsis*, rice, and poplar, respectively. Sequence comparisons revealed the conserved amino acids and motifs that are critical for the active site of *endo*- β -mannanases. Intron/exon structure analysis in conjunction with phylogenetic analysis implied that both intron gain and intron loss

has played roles in the evolution of *endo*- β -mannanase genes. The phylogenetic analysis that included the *endo*- β -mannanases from plants and other organisms implied that plant *endo*- β -mannanases have an ancient evolutionary origin. Comprehensive expression analysis of all *Arabidopsis* and rice *endo*- β -mannanase genes showed divergent expression patterns of individual genes, suggesting that the enzymes encoded by these genes, while carrying out the same biochemical reaction, are involved in diverse biological processes.

Keywords Endo- β -mannanase · Gene family · Comparative genome analysis · Evolution

Introduction

The plant cell wall is a composite of interwoven polymers that includes cellulose, hemicellulose, and pectin (Carpita and Gibeaut 1993). Selective modification of cell wall architecture is important for many aspects of plant growth, development, and responses to environmental factors (Carpita and McCann 2000). Many different types of cell wall glycosyl hydrolases are involved in altering the mechanical properties of cell walls (Cosgrove 1999). Mannans are widespread hemicellulosic polysaccharides in plant cell walls (Handford et al. 2003), and may have an important structural role (Whitney et al. 1998). The key enzyme for modifying plant mannans is *endo*- β -mannanase (EC 3.2.1.78), which randomly hydrolyzes the internal β -1,4-D-mannopyranosyl linkage in the mannan backbone.

The majority of the plant *endo*- β -mannanases that have been characterized have a role in seed biology. In the seeds

J. S. Yuan · X. Yang · J. Lai · Z.-M. Cheng · F. Chen (✉)
Department of Plant Sciences, University of Tennessee,
2431 Joe Johnson Drive,
Knoxville, TN 37996, USA
e-mail: fengc@utk.edu

H. Lin
Crop Diseases, Pests and Genetics, USDA-ARS,
9611 S. Riverbend Avenue,
Parlier, CA 93648, USA

H. Nonogaki
Department of Horticulture, Oregon State University,
4017 Agricultural and Life Science Building,
Corvallis, OR 97331, USA

of tomato and other Solanaceae species, the mannan-rich endosperm serves as a physical barrier to radicle protrusion. The weakening of the micropylar endosperm that covers the radicle is required for successful germination of such seeds (Groot and Karssen 1987; Chen and Bradford 2000). A number of cell wall modifying enzymes appear to have a mechanistic role in endosperm weakening, and *endo*- β -mannanase is one of them (Bradford et al. 2000). Some other *endo*- β -mannanase isoforms are involved in endosperm degradation after the completion of seed germination, which provides energy for seedling growth (Nonogaki and Morohashi 1996). In addition to its role in seed germination, *endo*- β -mannanase has also been associated with tissue softening in fruiting ripening (Bewley et al. 2000; Brummell et al. 2004). Recently, *endo*- β -mannanase was shown to be involved in anther and pollen development, suggesting that this enzyme plays a role in biological events other than seed germination and fruit development (Filichkin et al. 2004).

The tomato seed *LeMAN1* is the first *endo*- β -mannanase gene isolated from plants, and the protein it encodes appears to have a role in post-germination reserve mobilization (Bewley et al. 1997). Based on the sequence information of *LeMAN1*, a number of *endo*- β -mannanase genes have subsequently been isolated from several plant species. The plant *endo*- β -mannanase genes that have been characterized include *LeMAN2* (Nonogaki et al. 2000), *LeMAN4a* (Bourgault and Bewley 2002), and *LeMAN5* (Filichkin et al. 2004) from tomato, *manA* and *manB* from coffee (*Coffea arabica*) (Marraccini et al. 2001), *LsMAN1* from lettuce (*Lactuca sativa*) (Wang et al. 2004), and *DfMAN* from *Datura ferox* (Arana et al. 2005). Most of the products from these genes have roles in seed germination. It should be noted that all these *endo*- β -mannanase genes are from dicot species.

The recent completion of the genome sequence for *Arabidopsis* (*Arabidopsis thaliana*) (Arabidopsis Genome Initiative 2000), rice (*Oryza sativa*) (International Rice Genome Sequencing Project 2005), and poplar (*Populus trichocarpa*) provides a good opportunity for performing comparative genomic studies of *endo*- β -mannanase genes in these three species. Using an *in silico* approach, 8, 9, and 11 *endo*- β -mannanase genes were identified in the genomes of *Arabidopsis*, rice, and poplar, respectively. Sequence comparisons showed that the proteins encoded by these genes are highly conserved. Genomic organization analysis and phylogenetic analysis revealed the potential mechanisms responsible for the evolution of the gene family within and among species. Divergent expression patterns of individual *Arabidopsis* and rice *endo*- β -mannanase genes, determined by RT-PCR analysis, suggested that the enzymes encoded by these genes have roles in diverse biological processes.

Bioinformatics analyses

Sequence retrieval and analysis

The protein sequence of tomato *endo*- β -mannanase *LeMAN1* (accession: AAB87859) was used initially as a query sequence to search against the *Arabidopsis* (<http://www.Arabidopsis.org>), the rice (<http://www.tigr.org/tdb/e2k1/osa1>), and the poplar genome database (<http://genome.jgi-psf.org/Poptr1/Poptr1.home.html>), respectively, using the BLASTP algorithm (Altschul et al. 1990). The newly identified *endo*- β -mannanase-like sequences were used reiteratively to search the same sequence database. The cutoff *e* value was set to be e^{-6} . Gene structures were determined by refined analyses of the locations of introns and exons, and by comparing gene models with the cDNA and/or EST sequences when the information was available. In addition, TBLASTN searches were performed to identify the *endo*- β -mannanase sequences in the three genomes that may not have been annotated.

To identify *endo*- β -mannanase-like sequences from a gymnosperm species, the sequence of *LeMAN1* was used as a query to search against the loblolly pine (*Pinus taeda* L.) expression database (http://www.tigr.org/tigrscripts/tgi/T_index.cgi?species=pinus) using TBALSTN algorithm. To identify the fungi and microbial *endo*- β -mannanases that are related to plant *endo*- β -mannanases, the *LeMAN1* sequence was used as a query to search the fungi and microbial genomes at NCBI (<http://www.ncbi.nlm.nih.gov/BLAST/>).

Subcellular localization and signal peptide cleavage sites of the predicted *endo*- β -mannanase proteins were predicted by using SignalP 3.0 server available at the website <http://www.cbs.dtu.dk/services/SignalP> (Bendtsen et al. 2004).

Analysis of chromosomal location and intron/exon structure

The chromosome locations of the *Arabidopsis*, rice, and poplar *endo*- β -mannanase genes were generated by Map Viewer (<http://www.ncbi.nlm.nih.gov/mapview/static/MVstart.html>). Intron/exon structures were drawn using the Gene Structure Draw software (<http://warta.bio.psu.edu/cgi-bin/Tools/StrDraw.pl>).

Multiple sequence alignment and phylogenetic analysis

Multiple protein sequence alignments were constructed using ClustalW software (Thompson et al. 1994), and displayed using GeneDoc (<http://www.psc.edu/biomed/genedoc/>). To make phylogenetic trees, multiple sequence alignments performed using ClustalW were saved as NEXUS files and executed by PAUP (version 3.0) to generate UPGMA trees. The phylogenetic trees were viewed using the Tree-

View software (<http://www.taxonomy.zoology.gla.ac.uk/rod/treeview.html>).

Motif search

To discover conserved motifs in *endo*- β -mannanases, the sequences of all predicted *endo*- β -mannanases in *Arabidopsis*, rice, and poplar, LeMAN1, and the related bacterial, fungal, and blue mussel *endo*- β -mannanases were used as input for the MEME program (<http://www.meme.nbcrc.net/meme/meme-intro.html>).

The putative *endo*- β -mannanases in *Arabidopsis*, rice, and poplar

To identify the complete *endo*- β -mannanase gene families in the three species, the sequence of LeMAN1, the prototype plant *endo*- β -mannanase (Bewley et al. 1997), was initially used to search the genome sequence databases of the three species using the BLASTP algorithm (Altschul et al. 1990). The new *endo*- β -mannanase-like sequences detected in each species were in turn used reiteratively to search the respective sequence database. Because of the potential errors associated with automatic annotation (Rouze et al. 1999), refined sequence analysis of all putative *endo*- β -mannanase genes was performed. Different gene models assigned by different gene identification programs were compared with known plant *endo*- β -mannanases to determine the locations of introns and exons in the genomic sequences of individual *endo*- β -mannanase genes and to deduce the structures of the corresponding *endo*- β -mannanases. In the end, the cDNA and/or EST information available for *endo*- β -mannanase genes was used to verify the predictions of gene structures. Through the exhaustive sequence search, eight (*AtMAN1* to *AtMAN7* and *AtMANP*), nine (*OsMAN1* to *OsMAN8* and *OsMANP*), and 11 (*PtMAN1* to *PtMAN8*, and *PtMANP1* to *PtMANP3*) *endo*- β -mannanase genes were identified in the genome of *Arabidopsis*, rice, and poplar, respectively (Table 1). TBLASTN searches did not identify additional genes.

The length of *endo*- β -mannanases ranges from 408aa to 448aa in *Arabidopsis*, from 380aa to 492aa in rice, and from 416aa to 519aa in poplar (Table 1). All *endo*- β -mannanases, except *OsMAN7* and *OsMAN8*, are predicted to contain a signal peptide, and are, therefore, secreted proteins. *OsMAN7* and *OsMAN8* are shorter at the N terminus compared to other predicted proteins. All *Arabidopsis* *endo*- β -mannanases except *AtMANP* have corresponding cDNAs or ESTs. For rice *endo*- β -mannanase genes, all except *OsMAN5* and *OsMANP*, have corresponding ESTs or full-length cDNAs available. For poplar *endo*- β -mannanase genes, only three of them,

PtMANP1, *PtMANP2*, and *PtMANP3*, do not have ESTs or cDNAs.

Endo- β -mannanase pseudo genes

Among the *endo*- β -mannanase genes in the three species, one gene from *Arabidopsis* (*AtMANP*), one from rice (*OsMANP*), and three from poplar (*PtMANP1*, *PtMANP2*, and *PtMANP3*) showed major changes compared to conserved structures of other plant *endo*- β -mannanases, and are therefore predicted to be putative pseudo genes. *AtMANP* has a premature stop codon resulting from a point mutation (G255 to A) in the genomic sequence (Fig. 1a). For *OsMANP*, a point mutation (T640 to C) in its genomic sequence eliminates the conserved intron II 5'-splice site, which leads to a transcript encoding a truncated protein (Fig. 1b). Three poplar *endo*- β -mannanase genes appear to be pseudo genes. *PtMANP1* and *PtMANP3* are most similar to each other among the poplar *endo*- β -mannanases. Alignment of the genomic sequences of *PtMANP1* and *PtMANP3* revealed an insertion of four bases in *PtMANP1* in a highly conserved coding region. This insertion leads to a frame shift and results in premature termination of translation of *PtMANP1* (Fig. 1c). Both *PtMANP2* and *PtMANP3* contain multiple stop codons in the predicted coding regions.

Chromosome localization and genomic environment of *endo*- β -mannanase genes

Mapping of the eight *Arabidopsis* *endo*- β -mannanase genes revealed that these genes are scattered on all five chromosomes (Fig. 2). A cluster of two *endo*- β -mannanase genes (*AtMAN3* and *AtMAN4*) were found on the distal region of the long arm of chromosome 3. The nine *endo*- β -mannanase genes in rice are distributed on seven out of a total of 12 chromosomes (Fig. 2). A cluster of two *endo*- β -mannanase genes (*OsMAN3* and *OsMAN4*) are located on the short arm of chromosome 3. There are 19 chromosomes in the poplar genome. Eight poplar *endo*- β -mannanase genes are distributed on six chromosomes. The remaining three genes are assigned to three scaffolds that have not yet been assigned to any chromosome.

The size of *endo*- β -mannanase protein family in *Arabidopsis* is smaller than that of some other types of cell wall hydrolases, such as xyloglucan endotransglycosylases/hydrolase (XTH), which has 33 members in the *Arabidopsis* genome (Yokoyama and Nishitani 2004). XTHs in *Arabidopsis* and rice are featured by the presence of multiple tandem duplications (Yokoyama et al. 2004), suggesting that the modest size of the *endo*- β -

Table 1 Identification of *endo*- β -mannanase genes in *Arabidopsis*, rice and poplar

Gene	ID	Chr ^a	Strand ^b	Location	Determination of gene structure	Exons ^c	Protein size
<i>AtMAN1</i>	At1 g02310	1	–	458133–460696	Prediction and EST	5	411
<i>AtMAN2</i>	At2 g20680	2	+	8927548–8930254	Prediction and EST	5	433
<i>AtMAN3</i>	At3 g10890	3	–	3407460–3409005	Prediction and EST	4	414
<i>AtMAN4</i>	At3 g10900	3	–	3410257–3412075	Prediction and EST	4	408
<i>AtMAN5</i>	At4 g28320	4	–	14018189–14020223	Prediction and EST	5	431
<i>AtMAN6</i>	At5 g01930	5	–	361085–362895	Prediction and EST	5	448
<i>AtMAN7</i>	At5 g66460	5	–	26555621–26558262	Prediction and EST	5	431
<i>AtMANP</i>	At3 g30540	3	–	12147271–12148954	Prediction	4	
<i>OsMAN1</i>	LOC_Os01 g47400	1	+	27061482–27065327	Prediction and EST	4	432
<i>OsMAN2</i>	LOC_Os01 g54300	1	–	31234538–31232538	Prediction and EST	5	446
<i>OsMAN3</i>	LOC_Os03 g61270	3	+	34733350–34734934	Prediction and EST	3	469
<i>OsMAN4</i>	LOC_Os03 g61280	3	+	34738807–34742060	Prediction and EST	3	462
<i>OsMAN5</i>	LOC_Os05 g25480	5	+	14639095–14641413	Prediction	3	492
<i>OsMAN6</i>	LOC_Os06 g20620	6	+	11876530–11881234	Prediction and EST	5	441
<i>OsMAN7</i>	LOC_Os11 g02600	11	–	821829–819185	Prediction and EST	5	380
<i>OsMAN8</i>	LOC_Os12 g02520	12	–	859531–857268	Prediction and EST	5	383
<i>OsMANP</i>	LOC_Os02 g52800	2	–	32230727–32229097	Prediction	5	
<i>PtMAN1</i>	eugene3.00021716	2	+	14463129–14465446	Prediction and EST	5	416
<i>PtMAN2</i>	grail3.0002058401	5	+	15933413–15935576	Prediction and EST	5	421
<i>PtMAN3</i>	fgenes1_pm.C_LG_VI000038	6	+	616330–618903	Prediction and EST	5	441
<i>PtMAN4</i>	fgenes1_pg.C_LG_VI000925	6	+	7764093–7766592	Prediction and EST	6	519
<i>PtMAN5</i>	grail3.0019022001	7	+	11001899–11004288	Prediction and EST	5	436
<i>PtMAN6</i>	eugene3.00161335	16	+	13250582–13253003	Prediction and EST	6	472
<i>PtMAN7</i>	eugene3.00570156	S ^d _57	+	1336744–1339388	Prediction and EST	5	436
<i>PtMAN8</i>	grail3.0142001501	S_142	+	189350–193508	Prediction and EST	5	436
<i>PtMANP1</i>	eugene3.00160095	16	+	720779–723344	Prediction	5	
<i>PtMANP2</i>	fgenes1_pg.C_LG_XIX000624	19	+	8198172–8199788	Prediction	5	
<i>PtMANP3</i>	eugene3.01250004	S-125	+	45486–47108	Prediction	5	

^aChromosome number^bOrientation of open read frame^cNumber of exons^dScaffold

mannanase protein families in *Arabidopsis* and rice is partly due to the low occurrence of local tandem duplication.

The plant genome contains many different types of cell wall hydrolases (Henrissat et al. 2001), and it has been proposed that concerted action of multiple cell wall glycosyl hydrolases is critical for cell wall disassembly (Rose and Bennett 1999). In microbes, such as bacteria, cell wall hydrolases sometimes occur in clusters (Tamaru and Doi 2000), which provides a mechanism for concerted regulation of gene expression. To understand whether this phenomenon also occurs in plants, the genes that are adjacent to every *Arabidopsis endo*- β -mannanase gene were identified and the presumed biochemical functions of the respective products from these genes was analyzed. No cell wall hydrolase genes were found to be closely linked to *endo*- β -mannanase genes, suggesting that plants employ a different mechanism from the one used by microbes in regulating the coordinated expression of multiple cell wall hydrolases.

Sequence alignment and conserved structures of plant *endo*- β -mannanases

Plant *endo*- β -mannanases belong to the glycosyl hydrolases superfamily, which is the most widespread group of enzymes and have been divided into more than 70 families based on protein sequence similarities (Henrissat 1991). The plant *endo*- β -mannanases that have been characterized belong to the glycosyl hydrolase family 5 (GH5). This is in contrast to bacterial *endo*- β -mannanases, which belong to both GH5 and GH26, depending on the protein sequence of the particular enzyme (Hogg et al. 2003). Pairwise comparisons showed that the overall protein sequence similarities of all *endo*- β -mannanases in *Arabidopsis*, rice, and poplar range from 40 to 94%. The high-level sequence similarities suggest that all plant *endo*- β -mannanases belong to GH5.

LeMAN4a, a tomato *endo*- β -mannanase, is the first plant *endo*- β -mannanase whose three-dimensional (3-D) crystal structure has been solved (Bourgault et al. 2005). To

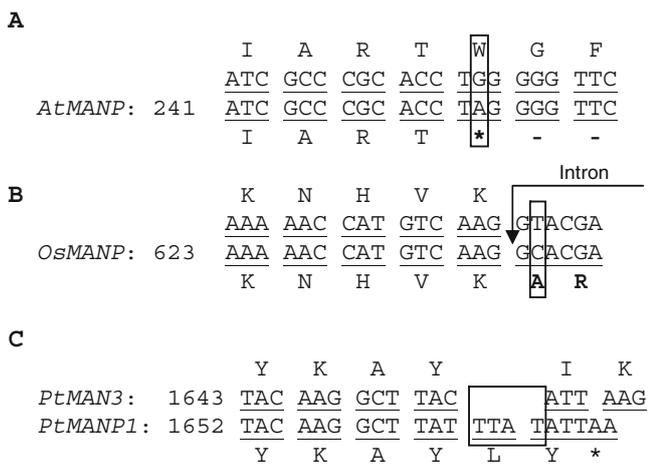


Fig. 1 Putative pseudo *endo*- β -mannanase genes from *Arabidopsis* (*AtMANP*), rice (*OsMANP*), and poplar (*PtMANP1*). **a** A G225 to A mutation occurs in *AtMANP* genomic DNA. While “TGG” encodes a W in other intact *endo*- β -mannanase genes, the mutation generates a stop codon “TAG”. **b** A T640 to C mutation in *OsMANP* genomic DNA leads to the elimination of a conserved intron cleavage site. **c** Insertion of “TTAT” in *PtMANP1* results in a frame shift and premature termination of protein translation

gain structural information on the predicted *endo*- β -mannanases, sequence alignment was made with LeMAN4a and all *endo*- β -mannanases from *Arabidopsis*, rice, and poplar. The alignment made with LeMAN4a and two representative *endo*- β -mannanases from *Arabidopsis*, rice, and poplar, respectively, is presented to illustrate the general observations (Fig. 3). LeMAN4a was shown to adopt the $(\beta/\alpha)_8$ fold that is common to the members of GH5 (Bourgault et al. 2005). Among the α -helices of LeMAN4a, α A appears to be unique to LeMAN4a as it is not present in the 3-D structures of other *endo*- β -mannanases. In LeMAN4a, three noncanonical helices, including α A, α B, and α C, surround the active site (Bourgault et al. 2005). The amino acids comprising α A, α B, and α C in LeMAN4a are conserved in plant *endo*- β -mannanases (Fig. 4).

When a glycosyl hydrolase binds its substrate, it binds multiple sugar groups. The subsites in the enzyme that bind to different sugar groups are numbered -4, -3, -2, -1, 1 and 2 from the nonreducing end to the reducing end of polysaccharides (Davies et al. 1997). *Endo*- β -mannanase cleaves the mannosyl linkage between -1 and 1 subsites through a so-called “retaining mechanism”, for which two catalytic residues, one acid/base and one nucleophile, are critical (Bourgault et al. 2005). In LeMAN4a, these two residues were determined to be E204 and E318 (Bourgault et al. 2005), which are conserved in plant *endo*- β -mannanases (Fig. 3). E204 is also located in the signature sequence common to all GH5 enzymes ([LIV]-[LIVMFYWGA](2)-[DNEQG]-[LIVMGST]-{SENR}-N-E-[PV]-[RHDNSTLIVFY]). In

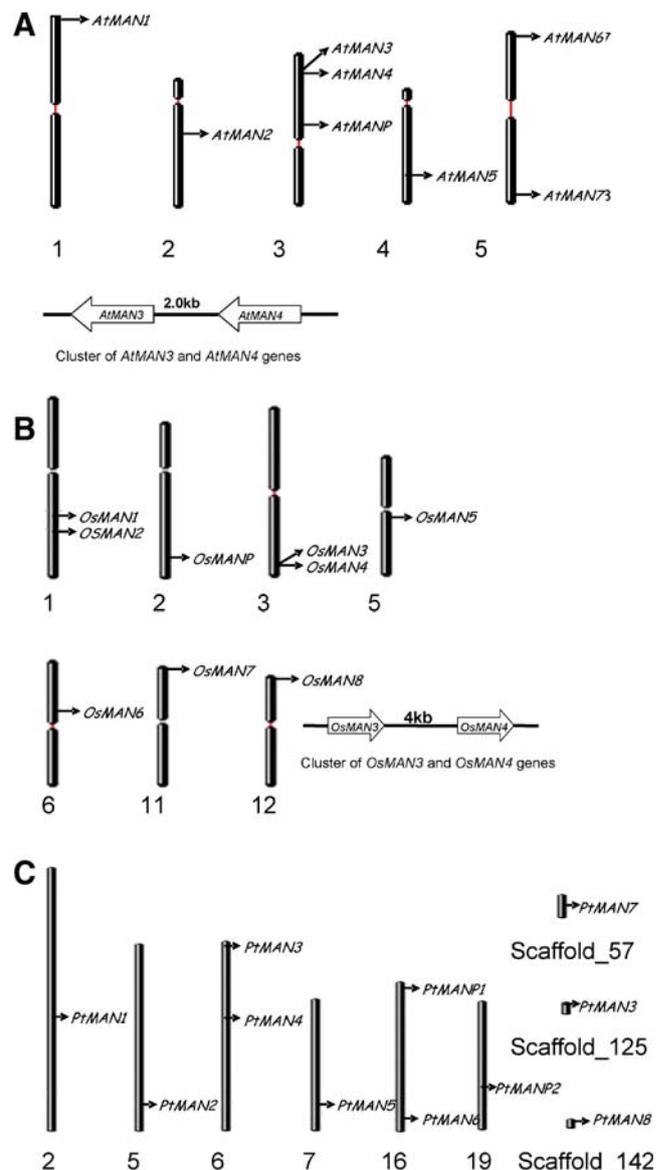


Fig. 2 Chromosomal distribution of *endo*- β -mannanase genes. **a** The chromosomal positions of *endo*- β -mannanase genes in *Arabidopsis*. *AtMAN3* and *AtMAN4* are clustered. **b** The chromosomal positions of *endo*- β -mannanase genes in rice. *OsMAN3* and *OsMAN4* are clustered. **c** The chromosomal positions of *endo*- β -mannanase genes in poplar. Three poplar *endo*- β -mannanase genes were assigned to scaffolds

the *endo*- β -mannanases from *Arabidopsis*, rice, and poplar, the signature sequence is [AGS]-W-[EQ]-L-[MI]-N-E-P-[RHQ]-[CS].

In addition to *endo*- β -mannanases, *exo*-1,3-glucanases, *endo*-1,6-glucanases, xylanases, and endoglycoceramidases also belong to GH5 (Henrissat 1991; Henrissat and Bairoch 1996; Davies et al. 1997; Henrissat et al. 2001). The presence of the signature sequence in all these GH5 hydrolases suggests that amino acids comprising the active site as well as the peripheral amino acids outside the active site all play roles in determining enzymatic specificity.

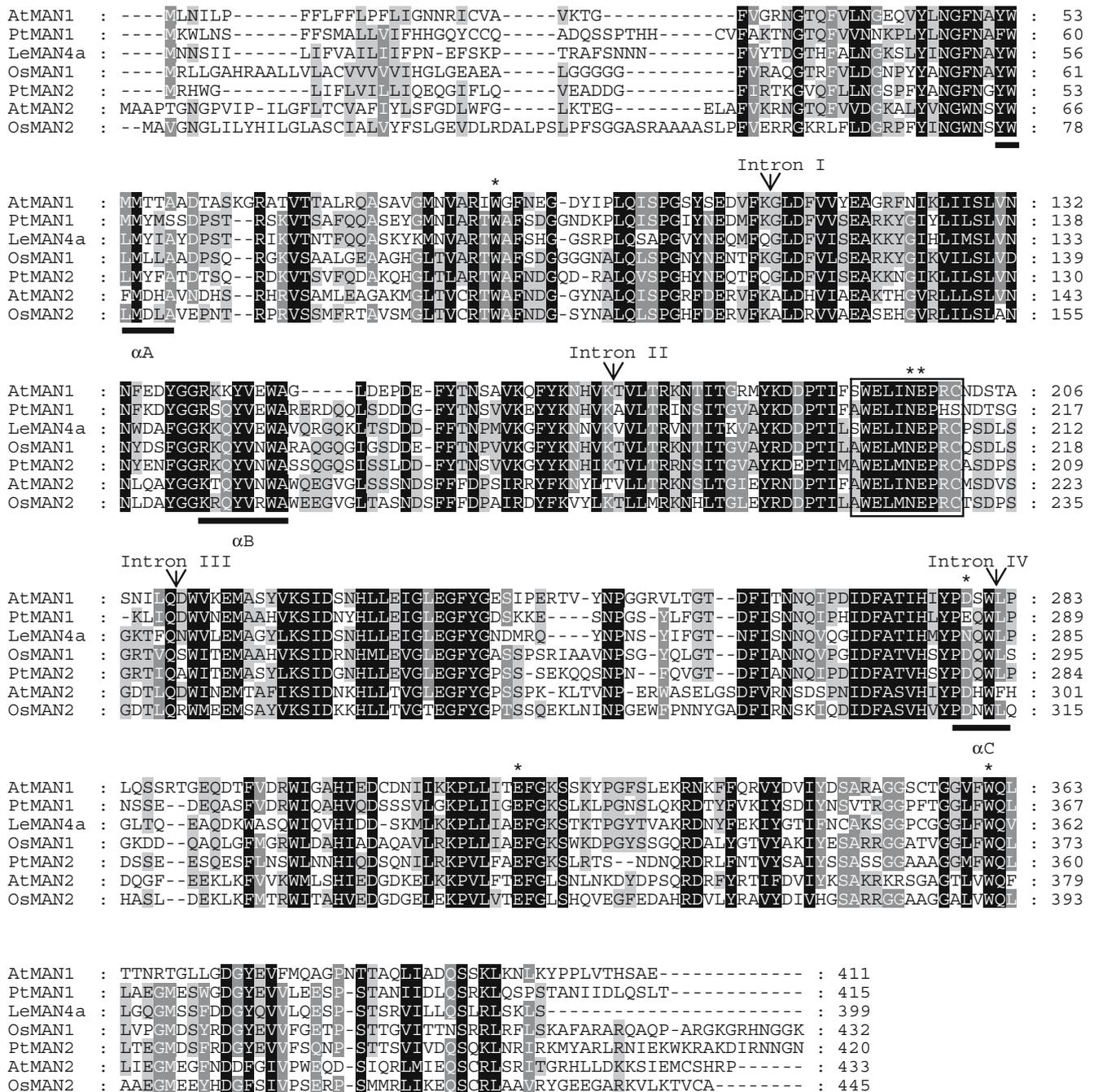


Fig. 3 The sequence alignment of representative *endo*- β -mannanases from *Arabidopsis*, rice and poplar with the tomato LeMAN4a. The proteins are less conserved at the N terminus. The GH5 signature sequences are boxed. The conserved amino acids critical for enzyme

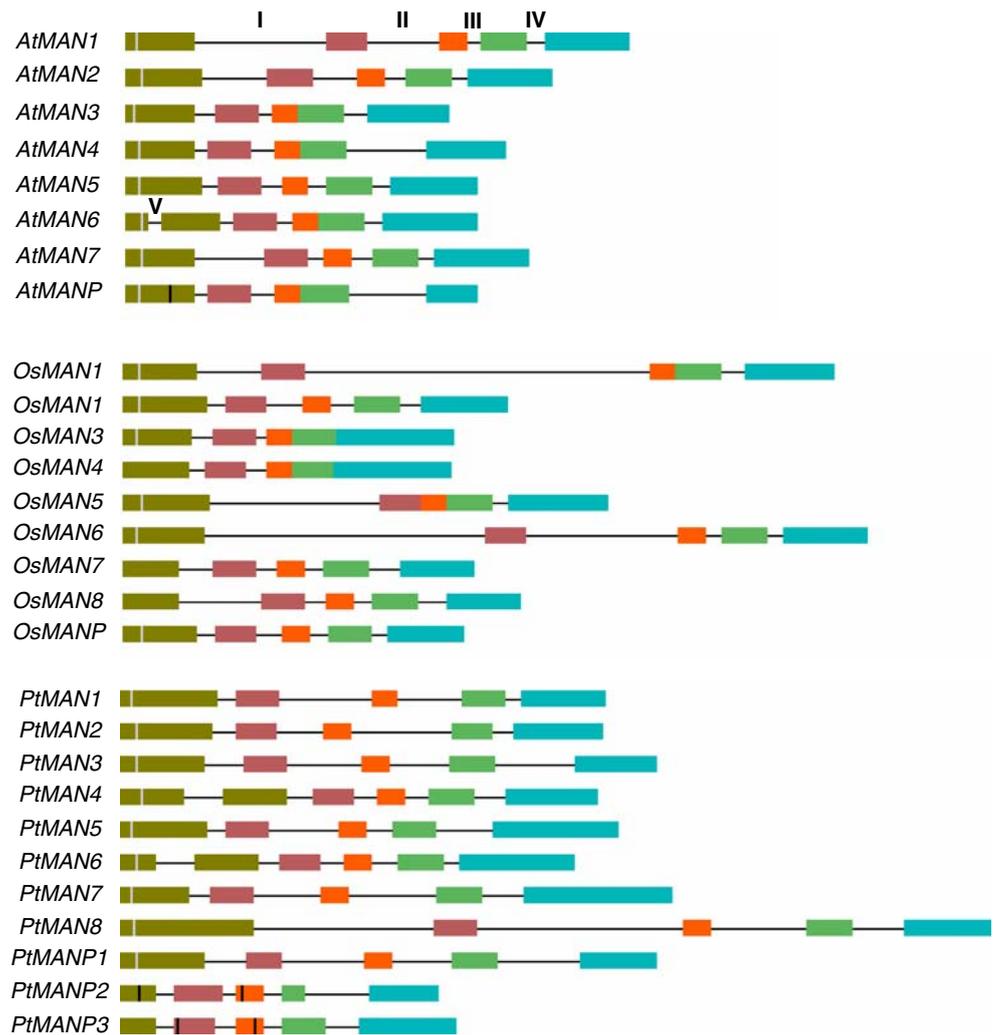
activity are indicated with asterisks. Intron positions are indicated with arrows. The regions that correspond to α A, α B, and α C, the three featured α -helices in the 3-D structure of LeMAN4a, are underlined

Among the other amino acids involved in the enzyme active site of LeMAN4a, W88, N203, W283, and W360 are also strictly conserved among the *endo*- β -mannanases in *Arabidopsis*, rice, and poplar (Fig. 3).

For LeMAN4a, L398 was shown to be important for catalytic activity, as deletion of L398 severely reduced enzyme activity (Bourgault and Bewley 2002). From the crystal structure, L398 was proposed to be critical for

maintaining the integrity of the enzyme (Bourgault et al. 2005), as protein missing L398 and S399 would be unstable. Interestingly, all predicted *endo*- β -mannanases are longer at the C terminal than LeMAN4a (Fig. 3). The amino acids in the predicted *endo*- β -mannanases corresponding to L398 and S399 in LeMAN4a are conserved, supporting that this region is important for *endo*- β -mannanase activity.

Fig. 4 The intron/exon structures of *endo-β-mannanase* genes in *Arabidopsis*, rice, and poplar. The colored boxes and lines represent exons and introns, respectively. Corresponding exons in different genes are highlighted in the same color. The signal peptide cleavage site in the first exon is indicated by a white line. Mutation in putative pseudo genes are illustrated by black lines



Intron/exon structure of *endo-β-mannanase* genes

In general, there are a total of five positions in which introns may occur within the *endo-β-mannanase* genes from *Arabidopsis*, rice, and poplar. Based on the location where an intron may occur, the introns are named Intron I to Intron V. Four of the five introns (Intron I to Intron IV) are highly conserved, as they have the same placement on the amino acid sequences (Fig. 3). Intron V is less conserved.

The majority of the 28 *endo-β-mannanase* genes from *Arabidopsis*, rice, and poplar have four introns (19/28) (Fig. 4). The only reported plant *endo-β-mannanase* gene with analyzed intron/exon organization is the tomato *LeMAN5*, which also contains four introns (Filichkin et al. 2004). An intron-less gene was presumed to encode the tomato *LeMAN2* (the term *LeMAN2* was given to the cDNA), but this protein is most likely the product of *LeMAN5* (the term given to the genomic DNA) (H.

Nonogaki, unpublished result). Taken together, these data suggest that four introns (I, II, III, and IV) is a predominant form for stereotypical plant *endo-β-mannanase* genes. In addition, there are three *endo-β-mannanase* genes containing five introns, four containing three, and three containing two.

Examining the frequency of occurrence of the five introns, Intron I appears to be the most conserved, because it is in all genes. Intron II, which is missing only in *OsMAN5*, is also highly conserved. In contrast, Intron V is the least conserved intron. It occurs only in three genes: *AtMAN6*, *PtMAN4*, and *PtMAN6*. Intron/exon organization patterns of *endo-β-mannanase* genes differ in the three species. All poplar genes contain Introns I, II, III, and IV. For *Arabidopsis* genes, variation occurs with Intron III, which exists in four (*AtMAN3*, *AtMAN4*, *AtMAN6*, and *AtMANP*) out of the eight genes. For the rice genes, variation occurs with both Intron III and Intron IV. Four genes (*OsMAN1*, *OsMAN3*, *OsMAN4*, and *OsMAN5*) do not contain Intron

III and two genes (*OsMAN3* and *OsMAN4*) do not contain Intron IV. The two genes that do not contain Intron IV also lack Intron III.

In comparison with the conserved nature of intron positions, the sizes of individual introns in different genes vary. Intron II, for example, has a size of 1,662 bases in *OsMAN1* but only 132 bases in *OsMAN3*.

Relatedness of *endo*- β -mannanases in *Arabidopsis*, rice, and poplar

Phylogenetic analyses were performed to infer the evolutionary relationships among the individual *endo*- β -mannanases in *Arabidopsis*, rice, and poplar. Variations in intron/exon structures of *endo*- β -mannanases were also analyzed in association with their phylogeny. In this analysis, the common ancestor gene of all *endo*- β -mannanases in *Arabidopsis*, rice, and poplar is presumed to contain four introns (Introns I, II, III, and IV). In *Arabidopsis*, the eight *endo*- β -mannanases are grouped into two clades (Fig. 5a), with one clade containing *AtMAN2* and *AtMANP* and the other clade containing six genes. In the clade of larger size, there are two gene pairs: *AtMAN3* and *AtMAN4*, and *AtMAN6* and *AtMAN7*, which likely resulted from recent duplication events. As a tandem repeat, *AtMAN3* and *AtMAN4* are likely the result of a recent local duplication, a mechanism that has played a significant role in gene family evolution (Cannon et al. 2004). Both *AtMAN3* and *AtMAN4* contain three introns, suggesting that the ancestral gene of *AtMAN3* and *AtMAN4* contained three introns. In contrast, the other two gene pairs in *Arabidopsis*, *AtMAN6* and *AtMAN7*, and *AtMAN2* and *AtMANP*, do not have the same intron–exon pattern, suggesting that intron gain and/or intron loss occurred after the duplication of the genes.

The rice genes are also grouped into two clades. One clade contains *OsMAN2* and *OsMAN6*. The large clade contains seven genes, which can be further divided into two subclades. One subclade contains *OsMAN5*, *OsMAN7*, and *OsMAN8*. The other subclade contains *OsMAN1*, *OsMANP*, *OsMAN3*, and *OsMAN4* (Fig. 5b). Unlike the *Arabidopsis* genes, the majority of which appear to have evolved by sequential duplication, the majority of rice genes come from two lineages. The ancestral genes of these two lineages may have diverged early in the evolutionary history. *OsMAN3* and *OsMAN4*, localized in a cluster on chromosome 3, appear to be most closely related, suggesting they have resulted from a recent tandem duplication event. *OsMAN3* and *OsMAN4* both have two introns, suggesting that their ancestral gene had only two introns.

The phylogeny of the poplar *endo*- β -mannanases is similar to that of the *Arabidopsis* genes. Eleven genes are grouped into two clades (Fig. 5c). Three genes, *PtMAN3*,

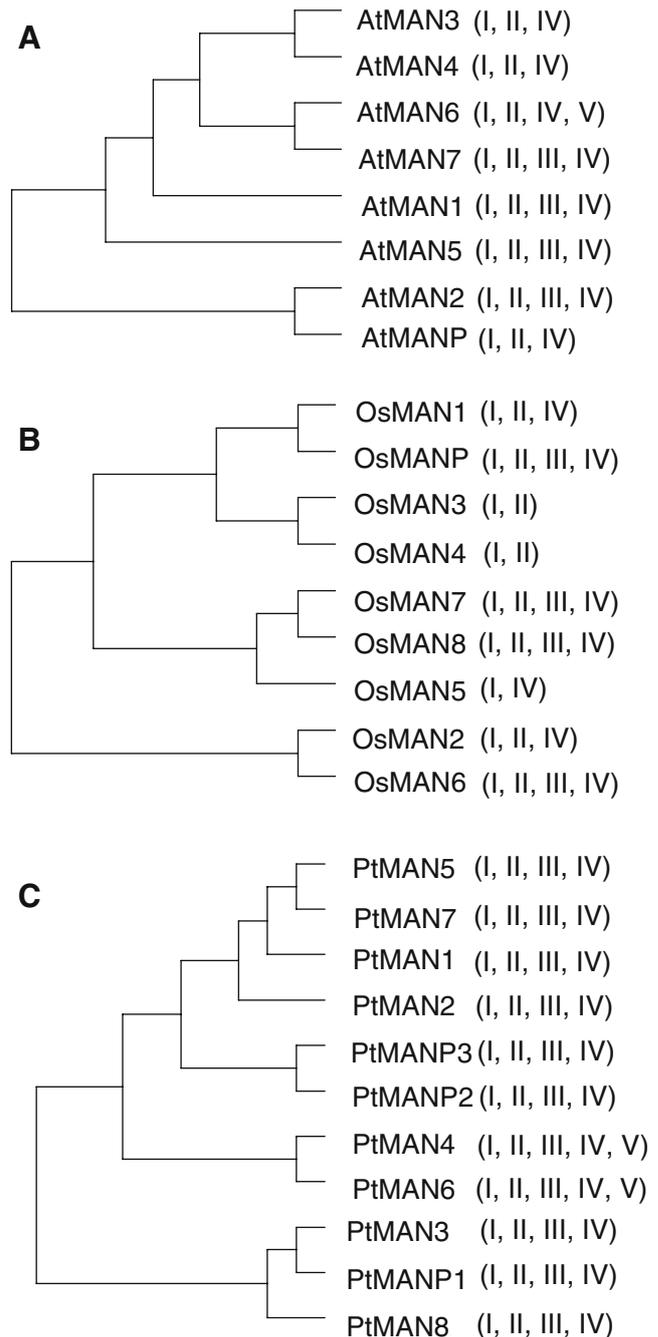


Fig. 5 Phylogenetic analysis of *endo*- β -mannanases. The introns contained in each gene are also shown. **a** Phylogenetic tree constructed from *endo*- β -mannanases in *Arabidopsis*; **b** phylogenetic tree constructed from *endo*- β -mannanases in rice; **c** phylogenetic tree constructed from *endo*- β -mannanases in poplar

PtMANP1, and *PtMAN8*, are grouped into one clade. The eight genes in the other clade are likely resulted from sequential duplication. *PtMAN4* and *PtMAN6* are most similar to each other. Both genes contain five introns, suggesting that intron V has existed before the divergence of the two genes. Different from the *Arabidopsis* and rice *endo*- β -mannanase genes, for which intron loss has played

an important role in the evolution of the gene families, poplar genes do not seem to have undergone intron loss.

As described earlier, *Arabidopsis*, rice, and poplar all contain putative pseudo *endo-β-mannanase* genes (Fig. 1). Interestingly, each of these pseudo genes, except *PtMANP2* and *PtMANP3*, is highly related at the sequence level with a functional *endo-β-mannanase* gene. Because these pseudo genes and their closely related intact genes are likely the result of gene duplication, the existence of multiple pseudo-intact gene pairs suggest that the duplicated genes have been negatively selected during evolution.

Phylogenetic analysis of all *endo-β-mannanases* in *Arabidopsis*, rice, and poplar implies that the existence of *endo-β-mannanases* predated the divergence of monocots and dicots (Fig. 6). Orthological relationships are evidently identified for two groups of genes (Fig. 6, indicated as a and b). In each group, the *Arabidopsis* gene(s) is/are more related to the poplar gene(s) than to the rice gene(s), consistent with the evolutionary relationships between the three plant species.

The evolution of plant *endo-β-mannanases*

It is known that the thickened secondary cell walls of gymnosperms contain large amounts of mannans (Lundqvist et al. 2002). However, no *endo-β-mannanase* gene has been isolated from a gymnosperm. To understand whether *endo-β-mannanases* also occur in gymnosperms, the protein sequence of LeMAN1 was used as a query sequence to search against an expression database of loblolly pine (*Pinus taeda* L.). From this analysis, four *endo-β-mannanase* sequences that are significantly similar to the known plant *endo-β-mannanases* were identified.

As a group of enzymes widely distributed, *endo-β-mannanases* are present in many microbes including both fungi and bacteria, which use the enzymes to degrade the mannans in plant cell walls. Because of their commercial value, many microbial *endo-β-mannanase* genes have been cloned and characterized. Microbial *endo-β-mannanases* belong to two families: GH5 and GH26 (Hogg et al. 2003). The proteins from the two families do not share significant sequence similarities, implying that the enzymes of the two families evolved independently. Even within bacterial GH5 *endo-β-mannanases*, the protein sequences display a great variation. Sequence search against microbial databases using LeMAN1 yields some fungal and bacterial *endo-β-mannanases* that exhibit significant similarity to the plant *endo-β-mannanases*. It should be noted that although the amount of genome sequence information available for microbes is large, the number of microbial *endo-β-mannanases* related to plant *endo-β-mannanases* identified is rather small. An *endo-β-mannanase* and its coding gene

were recently isolated from blue mussel (*Mytilus edulis*) (Xu et al. 2002)—the function of the enzyme is for blue mussel to digest mannans-containing food such as sea-weed (Xu et al. 2001). This was the first time for an animal *endo-β-mannanase* to be cloned and characterized.

To understand the evolutionary relationships between plant *endo-β-mannanases* and the proteins from non-plant sources, a comprehensive phylogenetic analysis was performed (Fig. 7). The plant *endo-β-mannanases* from both angiosperms and gymnosperms are highly related. For the four *endo-β-mannanase* sequences identified in loblolly pine, three forms a subclade, while the fourth is situated in a distinct clade. This suggests that the existence of *endo-β-mannanases* predated the divergence of angiosperm and gymnosperm linkages. The related *endo-β-mannanases* from bacteria and fungi are grouped into a single clade, suggesting that the diversification of plant *endo-β-mannanases* occurred after the divergence of plants from other lineages.

Examination of the aligned protein sequences revealed a number of conserved motifs within plant *endo-β-mannanases* and *endo-β-mannanases* from fungi, bacteria and mussel (Fig. 8). All of the analyzed proteins contain an N-terminal signal peptide with no significant sequence similarity. Among all the motifs detected, motif 1 (YKDDPTIMAWELMNEPRCQSDPSGDT) is the most conserved, as it is present in all *endo-β-mannanases*. Not surprisingly, the GH5 signature sequence is located in motif 1. Certain motifs appear to be species-specific. For instance, motif 8 (RDxFFNTVYDKIYNSAKRGGAGAGGLFWQL) occurs only in plants, suggesting that this motif likely evolved after the divergence of plants and other lineages. The mussel protein has fewer conserved motifs. The presence of closely related *endo-β-mannanases* in plants, fungi, bacteria, and animals suggest that these cell-wall degrading enzymes have an ancient evolutionary origin.

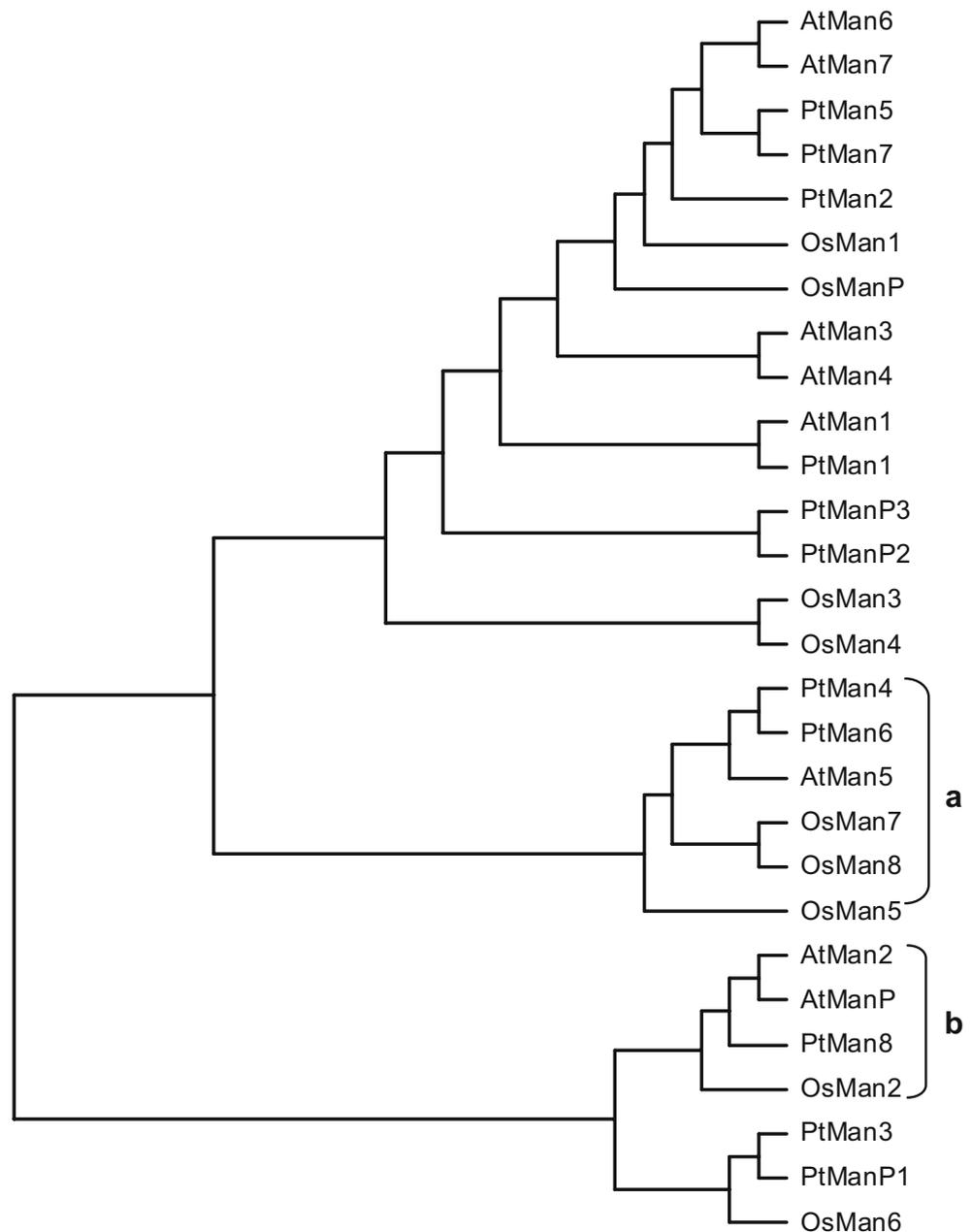
Expression patterns of individual *Arabidopsis* and rice *endo-β-mannanase* genes

The expression patterns of *endo-β-mannanase* genes has been studied in *Arabidopsis* and rice by Yuan et al. (Joshua S. Yuan, Xiaohan Yang, Jingru Lai, Hong Lin, Zong-ming Cheng, Hiroyuki Nonogaki, and Feng Chen, unpublished) and is summarized below.

Plant materials

Arabidopsis (Col 0) plants were grown on soil under natural light condition at 22°C in a greenhouse from January to March 2005 at Oregon State University. The

Fig. 6 Phylogenetic analysis of all *endo*- β -mannanases in *Arabidopsis*, rice, and poplar. The phylogenetic tree was generated from the corresponding alignment of multiple protein sequences

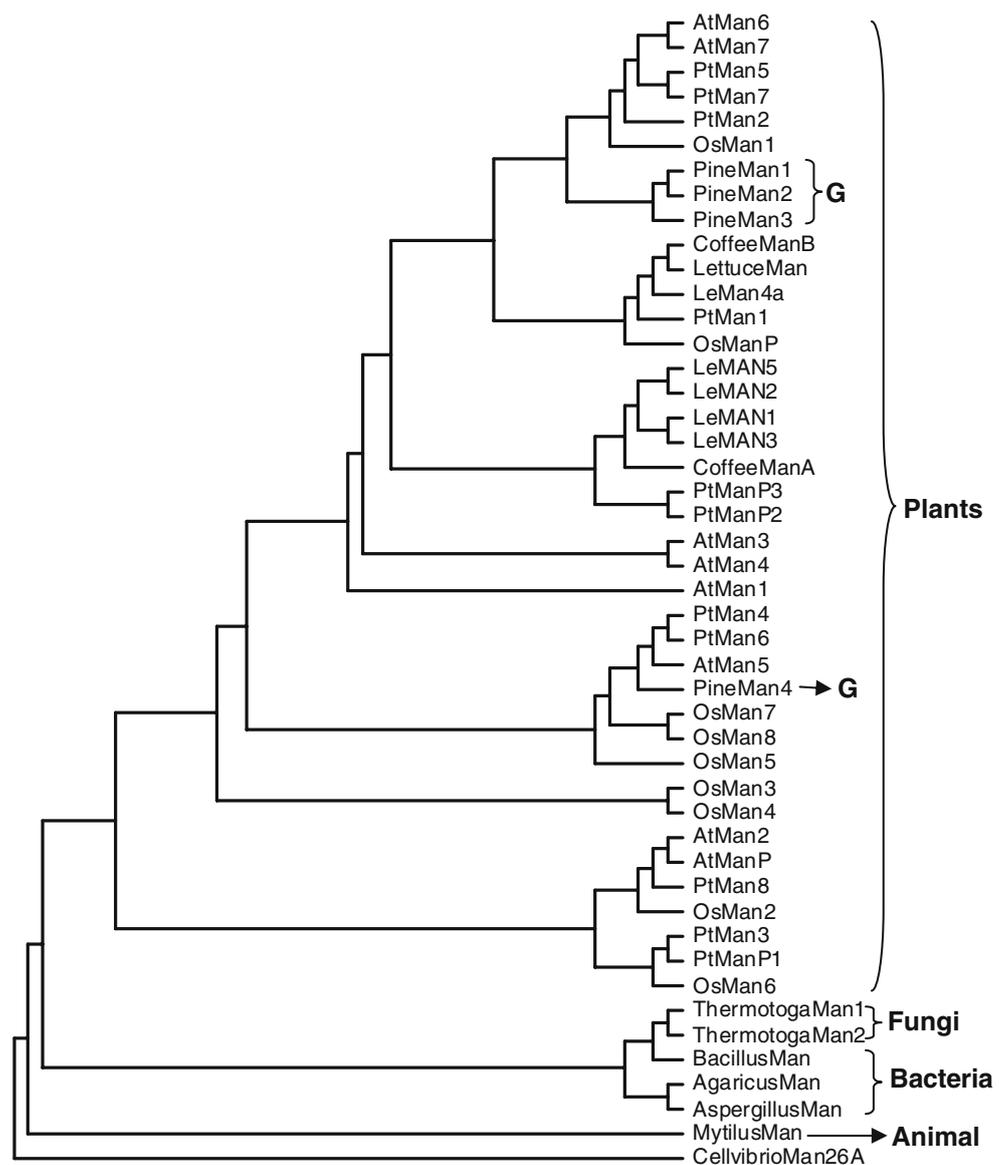


tissues of the inflorescence (mixture of flower buds with approximately 1 to 3 mm in length), silique (approximately 2 to 3 cm in length), main stem, cauline leaf, rosette leaf, and root were collected from mature flowering plants and used for RNA extraction. For the sample of germinating seeds, 20 mg of seeds were imbibed on wet filter paper at 4°C for 24 h to break dormancy. The seeds were next transferred to 22°C for 18 h and then collected for RNA extraction.

Seeds of *O. sativa* ssp. Japonica cv. Nipponbare were obtained from Dale Bumpers National Rice Research Center

(Stuttgart, AR, USA). Dehulled seeds were germinated on filter paper. After 4 days, the seedlings were transferred to soil and placed in a growth chamber. The plants were grown under 14 h light/10 h dark photoperiod. The temperatures were fluctuated between 26°C (day) and 22°C (night). The RH was 80% and light intensity was 400 $\mu\text{mol m}^{-2} \text{s}^{-1}$. Leaves, stems, and roots for RNA extraction were collected from 4-week-old seedlings, which were about 15 cm tall. For the seed sample, rice seeds were dehulled and then placed on wet filter paper in growth chamber (26°C). After 18 h, seeds were collected for RNA extraction.

Fig. 7 The phylogenetic analysis of *endo*- β -mannanases from plants and non-plant sources. In addition to the predicted proteins in *Arabidopsis*, rice, and poplar, other *endo*- β -mannanases used for phylogenetic analysis include the tomato LeMAN1 (AAB87859), LeMAN2 (AAG00315), LeMAN3 (AAG14352), LeMAN4a (AAK97760), LeMAN5 (AAM26920), CoffeeManA (CAC08208), CoffeeManB (CAC08442), LettuceMan (CAC51690), PineMan1 (TC57177), PineMan2 (TC57555), PineMan3 (TC78369), PineMan4 (TC67610), ThermotogaMan1 (CAB56856), ThermotogaMan2 (AAD36302), BacillusMan (AAU23418), AgaricusMan (CAB76904), AspergillusMan (AAA67426), MytilusMan (CAC81056), and CellvibrionMan26A (CAA57670). *G* represents gymnosperm



Gene expression analysis by RT-PCR

Preparation of total RNA and cDNA from *Arabidopsis* tissues were done as previously described (Liu et al. 2005). Semi-quantitative RT-PCR conditions were as follows: the initial denaturation at 94°C (4 min), seven touchdown cycles (94°C, 15 s), x (x represents an optimal temperature, see below)+7°C· x °C (15 s), and 72°C, (30 s) [one cycle for each annealing temperature from ($x+7$) to x] and 20 cycles at 94°C [15 s, x °C (15 s) and 72°C (30 s)] followed by extension at 72°C (7 min). Optimal temperature (x) for *AtMAN1*, *AtMAN2*, *AtMAN4*, and *AtMANP* was 60°C, for *AtMAN3* and *AtMAN7*, $x=62$ °C, and for *AtMAN5* and *AtMAN6*, $x=64$ °C. Primers specific for anactin gene *ACT2* (An et al. 1996) were used as a control in the semi-quantitative PCR. The annealing temperature for *ACT2* was

60°C. For amplification of each gene fragment from genomic DNA, 10 more cycles at 94°C (15 s), 54°C (15 s) and 72°C (30 s) followed by extension at 72°C (7 min) were added. All primers were tested with genomic DNA to confirm their effectiveness before RT-PCR.

The amplification for *AtMAN4/P* and *AtMAN6* resulted in fragments of two sizes. In both cases, the fragment of smaller size was expected. The larger fragment for *AtMAN4/P* was likely due to non-specific amplification. In the case of *AtMAN6*, the larger fragment was likely the product of gene At5g01920 that encodes a protein kinase. At5g01920 is situated on the same region as *AtMAN6*, but is encoded by the DNA strand opposite to the coding sequence of *AtMAN6*. The pair of primers designed for *AtMAN6* also amplified the cDNA of At5g01920. As in this region At5g01920 has no intron, but *AtMAN6* contains

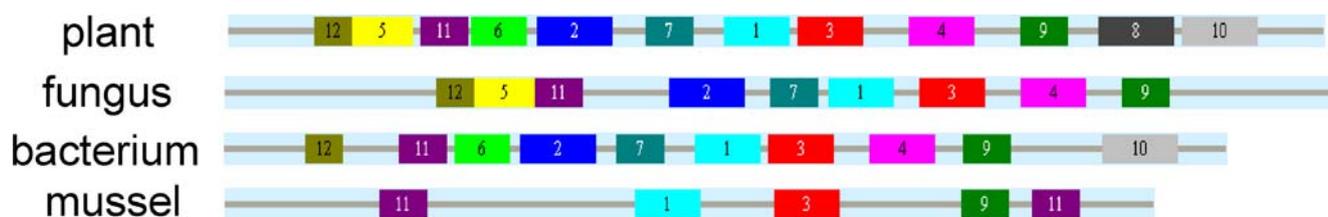


Fig. 8 A schematic illustration of the presence of conserved motifs in *endo*- β -mannanases. A total of 12 motifs were identified in all the proteins examined. While all 12 motifs are present in plant proteins, fungal proteins contain nine motifs (motifs 1, 2, 3, 4, 5, 7, 9, and 12),

bacterial proteins 10 motifs (motifs 1, 2, 3, 4, 6, 7, 9, 10, 11 and 12), and mussel protein five motifs (motifs 1, 3, 9 and two motifs 11). Positions of conserved motifs are determined based on the information in the multiple sequence alignment of *endo*- β -mannanases

introns, the amplicons for At5g01920 using the same pair of primers designed for *AtMAN6* should be larger than those for *AtMAN6*. To further verify this, we designed a second pair of primers for *AtMAN6*: forward primer 5'-AGGAC-CAACTTGGATTGAGAATTGTGCTC-3' and reverse primer 5'-GAGGGCATCAAGA GCAGTCT GATTCTTTT-3'. This pair of primers would amplify a cDNA fragment of 1,322 bp for *AtMAN6* and a larger fragment for At5g01920. PCR using the second pair of primers showed same amplification patterns as those using the first pair of primers, suggesting that the larger fragment in *AtMAN6* expression analysis was the product of At5g01910.

Semi-quantitative RT-PCR expression analysis of the rice genes was performed essentially as previously described (Chen et al. 2003). The primer sequences and sizes of the products are shown in Table 2. All primers were tested with genomic DNA to confirm their effectiveness.

Total RNA was isolated with Trizol (Invitrogen, Carlsbad, CA, USA). One microgram total RNA was synthesized into first-strand cDNA in a 20- μ l reaction using iScript cDNA synthesis kit (Bio-Rad Laboratories). One microliter of cDNA were used in each PCR reaction under the following conditions: an initial denaturing step at 95°C for 2 min followed by 30 cycles of 95°C for 45 s, 54°C for 45 s and 72°C for 60 s, and finally followed by an extension step of 72°C for 10 min.

Expression patterns of *Arabidopsis* and rice *endo*- β -mannanase genes

None of the *endo*- β -mannanase genes identified in *Arabidopsis*, rice, or poplar has been studied for their biological functions. To obtain information on the biological processes in which the products of these *endo*- β -mannanase genes are

Table 2 The primers used for RT-PCR analysis of *Arabidopsis* and rice genes

Gene name	Primers	Length (genomic DNA)	Length (cDNA)
<i>AtMAN1</i>	F: TCAGATCTCTCCTGGCTCCT R: GCCCTTGACTGTCGTAGAT	1,936	762
<i>AtMAN2</i>	F: TGGACCATGCTGTGAATGAT R: GCTTCCTCTGGCTGATTTG	1,644	903
<i>AtMAN3</i>	F: AATGGGAAACCATTTTACGC R: CAATGGCATCGAGTTGTGA	703	883
<i>AtMAN4/P</i>	F: AATGGGAAACCATTTTACGC R: CAATGGCATCGATATTGTGA	901	706
<i>AtMAN5</i>	F: CAATGACGGTGGCTACAATG R: GCCTGCACCTGACTTCTTTC	1,204	820
<i>AtMAN6</i>	F: GAGAGGCAAAGTGACGGAAG R: GTAAACCGCATAGCCGTCAT	1,207	953
<i>AtMAN7</i>	F: CGTTAGCTGAGGCAAGAAGG R: GTTCCCAACTTGTCCCTGA	1,219	912
<i>ACT2</i>	F: GCCATCCAAGCTGTTCTCTC R: GAACCACCGATCCAGACACT	706	629
<i>OsMAN1</i>	F: CGTCAAGTCCATCGACCGCAA R: AACGCCTTGCTGAGGAAGCG	662	455
<i>OsMAN2</i>	F: GATCGACAAGAAGCACCTCCTCA R: ATCATGGACGGCCTCTCGCT	642	492
<i>OsMAN3</i>	F: TTCACCAGCTCCGTCGTCAA R: GCGTCCAGTTGCGGAAGAAG	511	429
<i>OsMAN4</i>	F: CTGGACGCGCTCCACATTG R: CGCAGCTAAGCGAGTTGCCT	400	400
<i>OsMAN5</i>	F: CACAAGATCAGGCTGATTCTTCG R: ATGAACGCGTCCCTCGACGT	785	673
<i>OsMAN6</i>	F: TGAACCTGCTATATTGCTTGGGAA R: ACATGTGTGTTCTTCAACCATCAA	713	556
<i>OsMAN7</i>	F: GGATGCAACAACATATTGATGATGC R: AAACGATGTCGTACCAACCGTGA	330	330
<i>OsMAN8</i>	F: TGTGGATAAATGGATGCAAC R: CATTCCAGACACCTAAGTG	310	310
<i>OsMANP</i>	F: TACGTGAAATCCGTGGATCCG R: TAGACTTGCTGCGCTCGAGC	470	470
<i>OsActin</i>	F: GACTCTGGTGATGGTGTGACCCAC R: CTGCTGGAATGTGCTGAGAGATGC	602	602

F forward primer, R reserve primer

involved, comprehensive gene expression analyses using semi-quantitative RT-PCR were performed for all *Arabidopsis* and rice genes. In *Arabidopsis*, the analyses were performed with RNA isolated from inflorescence, siliques, stems, cauline leaves, rosette leaves, and roots from mature plants, as well as germinating seeds (Fig. 9a). In rice, gene expression analyses were performed with RNA isolated from leaves, roots, and stems from 1-month old seedlings, and germinating seeds (Fig. 9b).

Expression of the eight *Arabidopsis endo-β-mannanase* genes in each biological sample were analyzed in seven PCR reactions (Fig. 9a). Due to the high sequence similarity, one pair of primers was designed to determine the expression of both *AtMAN4* and *AtMANP*. *AtMANP* is the only *Arabidopsis endo-β-mannanase* gene that does not have corresponding cDNAs or ESTs in the public database; therefore, the expression detected by the *AtMAN4/AtMANP* primers most likely represents *AtMAN4* expression. All other genes, except *AtMAN5*, which showed expression only in stems, showed expression in more than one tissue. All genes except *AtMAN6* showed expression in inflorescence. While no gene showed expression in cauline leaves, two genes, *AtMAN2* and *AtMAN3*, were expressed in rosette leaves. The differential expression of *endo-β-mannanase* genes in rosette and cauline leaves suggest that the expression of relevant genes are under developmental regulation. Three genes (*AtMAN2*, *AtMAN3*, and *AtMAN7*) were expressed in germinating seeds. Whether these genes are involved in regulating seed germination by enhancing the growth potential of the embryo and/or weakening endosperm, as that has been demonstrated in the germinating seeds of a number of plant species including tomato, coffee, and *D. ferox* (Bradford et al. 2000), remains to be investigated.

The expression of the nine rice *endo-β-mannanase* genes was also analyzed. In contrast to the *Arabidopsis* genes, none of which were expressed in all the tissues examined, five rice *endo-β-mannanase* genes showed ubiquitous expression (*OsMAN1*, *OsMAN2*, *OsMAN4*, *OsMAN6*, and *OsMAN8*). Except *OsMAN1*, the four other genes showed differences in gene expression levels among different tissues. *OsMAN5*, *OsMAN7*, and *OsMANP* did not show expression in any of the tissue examined. While most of the genes showed comparable levels of expression in different tissues, *OsMAN6* and *OsMAN8* showed lower levels of expression in roots than that in other tissues. Six genes (*OsMAN1*, *OsMAN2*, *OsMAN3*, *OsMAN4*, *OsMAN6*, and *OsMAN8*) were expressed in germinating seeds, and *OsMAN3* showed seed-specific expression (Fig. 9b). Recently, it was reported that *endo-β-mannanase* activity was detected in germinated rice seeds (Wang et al. 2005). It will be interesting to determine which rice *endo-β-mannanase* gene(s) is responsible for the activity detected.

Biological functions of plant *endo-β-mannanases*

Mannans are present in the endosperm of seeds of a large number of plant species, such as carob (*Ceratonia siliqua*), guar (*Cyamopsis tetragonolobus*), fenugreek (*Trigonella foenum-graecum*), and tomato. In these seeds, mannans serve as food reserves and/or play a role in regulating seed germination. Mannans are also abundant in the lignified secondary walls of gymnosperms (Lundqvist et al. 2002), where they may have a structural role in cross-linking cellulose fibrils. Mannans are also present in a variety of other tissues (Bacic et al. 1988). However, the biological roles of these mannans are unclear.

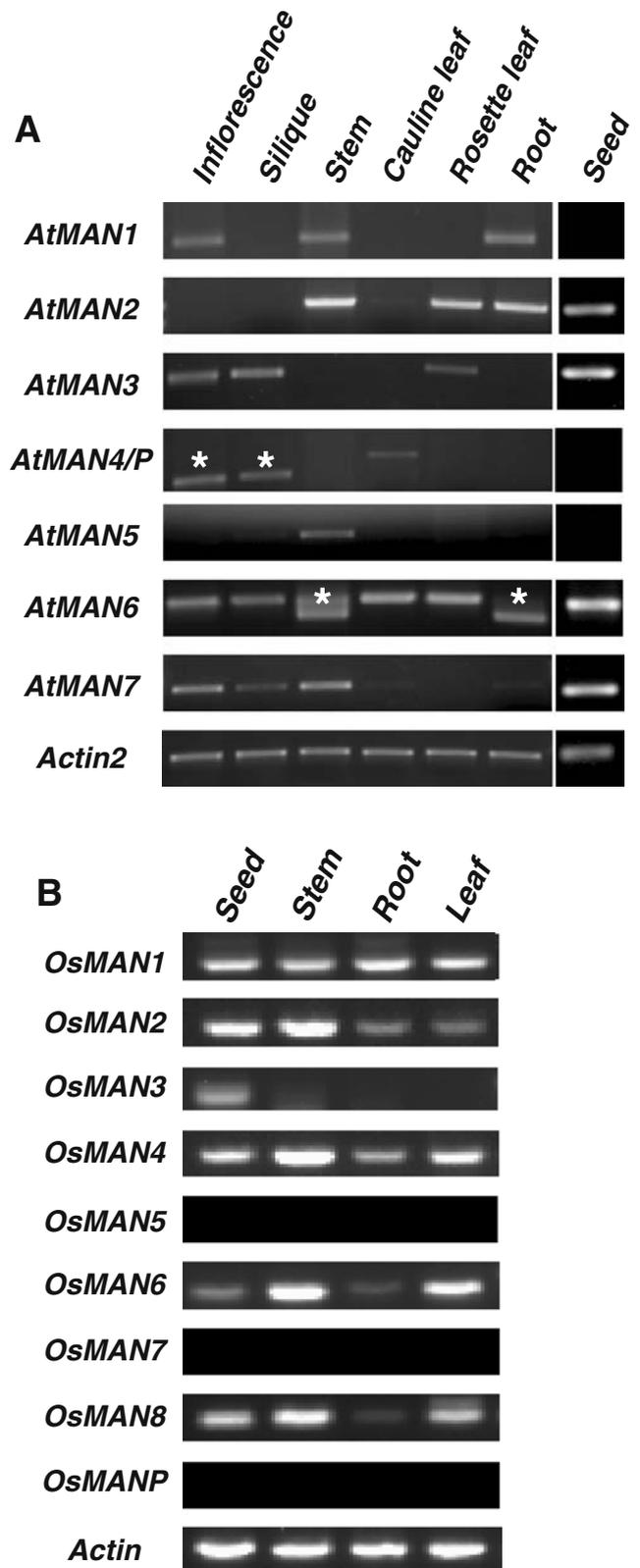
Mannans may have roles in plant development. This notion is supported by the wide distribution of mannans in the plant body. In a recent study, various tissues of *Arabidopsis* were found to contain mannans (Handford et al. 2003). Using immunofluorescence light microscopy and immunogold electron microscopy, Handford et al. (2003) showed that mannans are abundantly present in the thickened secondary cell walls of the xylem of various tissues and in the thickened epidermal cell walls of both leaves and stems. Low levels of mannans were observed in most other cell types examined. The widespread distribution of mannans in *Arabidopsis* suggests that mannans in different tissues may have specific roles.

The biological function of mannans is likely to be regulated by their biosynthesis and degradation. The mannan backbone is synthesized through the action of β -mannan synthase (ManS) (Dhugga et al. 2004). ManS is phylogenetically related to group A of the cellulose synthase-like sequences (Csl). *Arabidopsis* and rice each contains only two Csl sequences (Dhugga et al. 2004). Compared to the eight and nine *endo-β-mannanases* in *Arabidopsis* and rice, respectively, the biochemistry of mannan biosynthesis has much less diversity, suggesting that mannan degradation controlled by *endo-β-mannanases* may be more important to the biological roles of mannans.

The function of *endo-β-mannanases* in monots is intriguing. No *endo-β-mannanase* from a monocot species has been characterized in terms of gene function. The cell walls of monocots have a number of major structural and compositional differences from those of dicots and are termed type II walls (Carpita 1996). Nonetheless, small amounts of glucomannan have been found to be tightly bound to the cellulose microfibrils in the cell walls of grass species (Carpita 1996). The presence of a family of *endo-β-mannanases* in the rice genome that are closely related to their counterparts in dicots suggests that the functions of these enzymes are likely conserved in monocots and dicots.

Why do plant genomes contain multiple *endo-β-mannanases*? This very same question has been raised for

Fig. 9 Expression analysis of *Arabidopsis* and rice *endo- β -mannanase* genes. **a** RT-PCR analysis with *Arabidopsis* *endo- β -mannanase* genes. Total RNA was extracted from inflorescence, siliques, stems, cauline leaves, rosette leaves and roots from mature plants, as well as germinating seeds, and used for RT-PCR. The expression of eight *Arabidopsis* *endo- β -mannanase* genes in each sample was examined in seven PCR reactions, with *AtMAN4* and *AtMANP* examined using one pair of primers. The reactions for both *AtMAN4/P* and *AtMAN6* produced two fragments of different sizes. In both cases, the smaller size fragments are the products of the corresponding *endo- β -mannanase* genes (labeled with asterisks). The expression of an actin gene was used as an internal control. **b** RT-PCR analysis with rice *endo- β -mannanase* genes. Total RNA was extracted from leaves, stems, roots, and germinating seeds, and used for RT-PCR. The expression of nine rice *endo- β -mannanase* genes in each sample was examined in nine PCR reactions. The expression of an actin gene was used as an internal control



other cell wall hydrolases, for example, polygalacturonases (Hadfield and Bennett 1998) and XTHs (Yokoyama et al. 2004), which also exist as protein families with multiple members. Individual XTHs in rice exhibited organ- and growth stage-specific expression (Yokoyama et al. 2004), suggesting these genes have diverse roles in plant growth and development. The divergent expression patterns of the *Arabidopsis* and rice *endo- β -mannanase* genes suggests that these genes are also likely to be involved in diverse biological processes (Fig. 9). Although the expression of poplar *endo- β -mannanase* genes has not yet been analyzed, it is worth mentioning that *PtMAN6* has been suggested to be a potential target for a *MicroRNA* gene ptr-miR160 (Lu et al. 2005). MicroRNAs are 21–24 nt small RNA species that play critical roles in various biological processes by regulating gene expression (Bartel and Bartel 2005). It remains to be determined whether *PtMAN6* is the real target for ptr-miR160, and if so, what the biological significance of the regulation might be.

Individual *endo- β -mannanases* may also display specificity towards mannans with different structures. Depending on the composition of sugars in the backbone and the branches of mannan polysaccharides, plant mannans can be divided into four subgroups: pure mannans, glucomannans, galactomannans, and galactoglucomannans. Whether the individual *endo- β -mannanases* in *Arabidopsis*, rice, and poplar have preference for using mannans of different structure as substrate remains to be determined. Interestingly, a recent study showed that *LeMAN4* from ripe tomato fruit can act as a mannan transglycosylase in the presence of mannan-derived oligosaccharides, in addition to its mannan hydrolases activity (Schröder et al. 2006). This new finding suggests that the roles of *endo- β -mannanases* in seed germination, fruit ripening, and other plant developmental processes may need to be reinterpreted, as the different activity of *endo- β -mannanases* would result in different modifications of the structure of plant cell walls. The rich genetic and genomic resources for *Arabidopsis*, rice, and poplar will help identify the biological roles of individual *endo- β -mannanases*. The comparative genomic studies presented in this paper provide a valuable roadmap into the further functional investigation of these genes.

References

- Altschul SF, Stephen F, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215:403–410
- An YQ, McDowell JM, Huang S, McKinney EC, Chambliss S, Meagher RB (1996) Strong, constitutive expression of the *Arabidopsis ACT2/ACT8* actin subclass in vegetative tissues. *Plant J* 10:107–121
- Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408:796–815
- Arana MV, de Miguel LC, Sanchez RA (2005) A phytochrome-dependent embryonic factor modulates gibberellin responses in the embryo and micropylar endosperm of *Datura ferox* seeds. *Planta* (in press). DOI 10.1007/s00425-005-0134-7
- Bacic A, Harris PJ, Stone BA (1988) Structure and function of plant cell walls. In: Preiss J (ed) *Carbohydrates. The biochemistry of plants*, vol. 14. Academic, New York, pp 297–371
- Bartel B, Bartel DP (2005) MicroRNAs: at the root of plant development? *Plant Physiol* 132:709–717
- Bendtsen JD, Nielsen H, Heijne G, Brunak S (2004) Improved prediction of signal peptides: Signal P 3.0. *J Mol Biol* 340:783–795
- Bewley JD, Burton RA, Morohashi Y, Fincher GB (1997) Molecular cloning of a cDNA encoding a (1 \rightarrow 4)-beta-mannan endohydrolase from the seeds of germinated tomato (*Lycopersicon esculentum*). *Planta* 203:454–459
- Bewley JD, Banik M, Bourgault R, Feurtado JA, Toorop P, Hilhorst HW (2000) Endo- β -mannanase activity increases in the skin and outer pericarp of tomato fruits during ripening. *J Exp Bot* 51:529–538
- Bourgault R, Bewley JD (2002) Variation in its C-terminal amino acids determines whether ENDO- β -MANNANASE is active or inactive in ripening tomato fruits of different cultivars. *Plant Physiol* 130:1254–1262
- Bourgault R, Oakley AJ, Bewley JD, Wilce MCJ (2005) Three-dimensional structure of (1,4)- β -D-mannan mannanohydrolase from tomato fruit. *Protein Sci* 14:1233–1241
- Bradford KJ, Chen F, Cooley MB, Dahal P, Downie B, Fukunaga KK, Gee OH, Gurusinge S, Mella RA, Nonogaki H, Wu CT, Yim KO (2000) Gene expression before radicle emergence in imbibed tomato seeds. In: Black M, Bradford KJ, Vazquez-Ramos J (eds) *Seed biology: advances and applications*. CABI, Wallingford, UK, pp 231–251
- Brummell DA, Cin VD, Crisosto CH, Labavitch JM (2004) Cell wall metabolism during maturation, ripening and senescence of peach fruit. *J Exp Bot* 55:2029–2039
- Cannon SB, Mitra A, Baumgarten A, Young ND, May G (2004) The roles of segmental and tandem gene duplication in the evolution of large gene families in *Arabidopsis thaliana*. *BMC Plant Biol* 4:10
- Carpita NC (1996) Structure and biogenesis of the cell walls of grasses. *Annu Rev Plant Physiol Plant Mol Biol* 47:445–476
- Carpita NC, Gibeaut DM (1993) Structural models of primary cell walls of flowering plants: consistency of molecular structure with the physical properties of the walls during growth. *Plant J* 3:1–30
- Carpita N, McCann M (2000) The cell wall. In: Buchanan BB, Gruissem W, Jones RL (eds) *Biochemistry and molecular biology of plants*. American Society of Plant Biologist, Rockville, MD, pp 52–108
- Chen F, Bradford KJ (2000) Expression of an expansin is associated with endosperm weakening during tomato seed germination. *Plant Physiol* 124:1265–1274
- Chen F, Tholl D, D'Auria JC, Farooq A, Pichersky E, Gershenzon J (2003) Biosynthesis and emission of terpenoid volatiles from *Arabidopsis* flowers. *Plant Cell* 15:481–494
- Cosgrove DJ (1999) Enzymes and other agents that enhance cell wall extensibility. *Annu Rev Plant Physiol Plant Mol Biol* 50:391–417
- Davies GJ, Wilson KS, Henrissat B (1997) Nomenclature for sugar-binding subsites in glycosyl hydrolases. *Biochem J* 321:557–559

- Dhugga KS, Barreiro R, Whitten B, Stecca K, Hazebroek J, Randhawa GS, Dolan M, Kinney AJ, Tomes D, Nichols S, Anderson P (2004) Guar seed β -mannan synthase is a member of the cellulose synthase super gene family. *Science* 303:363–366
- Filichkin SA, Leonard JM, Monteros A, Liu PP, Nonogaki H (2004) A novel *endo*- β -mannanase gene in tomato *LeMAN5* is associated with anther and pollen development. *Plant Physiol* 134:1080–1087
- Groot SPC, Karssen CM (1987) Gibberellins regulate seed germination in tomato by endosperm weakening—a study with gibberellin-deficient mutants. *Planta* 171:525–531
- Hadfield KA, Bennett AB (1998) Polygalacturonases: many genes in search of a function. *Plant Physiol* 117:337–343
- Handford MG, Baldwin TC, Gobet F, Prime TA, Miles J, Yu X, Dupree P (2003) Localisation and characterization of cell wall mannan polysaccharides in *Arabidopsis thaliana*. *Planta* 218:27–36
- Henrissat B (1991) A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem J* 280:309–316
- Henrissat B, Bairoch A (1996) Updating the sequence-based classification of glycosyl hydrolases. *Biochem J* 316:695–696
- Henrissat B, Coutinho PM, Davies GJ (2001) A census of carbohydrate-active enzymes in the genome of *Arabidopsis thaliana*. *Plant Mol Biol* 47:55–72
- Hogg D, Pell G, Dupree P, Goubet F, Martin-Orue SM, Armand S, Gilbert HJ (2003) The molecular architecture of *Cellvibrio japonicus* mannanases in glycoside hydrolase families 5 and 26 points to differences in their role in mannan degradation. *Biochem J* 371:1027–1043
- International Rice Genome Sequencing Project (2005) The map-based sequence of the rice genome. *Nature* 436:793–800
- Liu P-P, Koizuka N, Homrichhausen TM, Hewitt JR, Martin RC, Nonogaki H (2005) Large scale screening of *Arabidopsis* enhancer-trap lines for seed germination-associated genes. *Plant J* 41:936–944
- Lu SF, Sun YH, Shi R, Clark C, Li L, Chiang VL (2005) Novel and mechanical stress-responsive MicroRNAs in *Populus trichocarpa* that are absent from *Arabidopsis*. *Plant Cell* 17:2186–2203
- Lundqvist J, TeLeMAN A, Junel L, Zacchi G, Dahlman O, Tjerneld F, Stalbrand H (2002) Isolation and characterization of galactoglucomannan from spruce (*Picea abies*). *Carbohydr Polym* 48:29–39
- Marraccini P, Rogers WJ, Allard C, André ML, Caillet V, Lacoste N, Lausanne F, Michaux S (2001) Molecular and biochemical characterization of ENDO- β -MANNANASES from germinating coffee (*Coffea arabica*) grains. *Planta* 213:296–308
- Nonogaki H, Morohashi Y (1996) An ENDO- β -MANNANASE develops exclusively in the micropylar endosperm of tomato seeds prior to radicle emergence. *Plant Physiol* 110:555–559
- Nonogaki H, Gee OH, Bradford KJ (2000) A germination-specific ENDO- β -MANNANASE gene is expected in the micropylar endosperm cap of tomato seeds. *Plant Physiol* 123:1235–1245
- Rose JKC, Bennett AB (1999) Cooperative disassembly of the cellulose-xyloglucan network of plant cell walls: parallels between cell expansion and fruit ripening. *Trends Plant Sci* 4:176–183
- Rouze P, Pavy N, Rombauts S (1999) Genome annotation: which tools do we have for it? *Curr Opin Plant Biol* 2:90–95
- Schröder R, Wegrzyn TF, Sharma NN, Atkinson RG (2006) *LeMAN4* *endo*- β -mannanase from ripe tomato fruit can act as a mannan transglycosylase or hydrolase. *Planta* DOI 10.1007/s00425-006-0286-0
- Tamaru Y, Doi RH (2000) The *engL* gene cluster of *Clostridium cellulovorans* contains a gene for cellulosomal ManA. *J Bacteriol* 182:244–247
- Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22:4673–4680
- Wang AX, Li JR, Bewley JD (2004) Molecular cloning and characterization of an *endo*- β -mannanase gene expressed in the lettuce endosperm following radicle emergence. *Seed Sci Res* 14:267–276
- Wang AX, Wang XF, Ren YF, Gong XM, Bewley JD (2005) Endo- β -mannanase and β -mannosidase activities in rice grains during and following germination, and the influence of gibberellin and abscisic acid. *Seed Sci Res* 15:219–227
- Whitney SEC, Brigham JE, Darke AH, Reid JSG, Gidley MJ (1998) Structural aspects of the interaction of mannan-based polysaccharides with bacterial cellulose. *Carbohydr Res* 307:299–309
- Xu B, Häggglund P, Stålbrand H, Janson J-C (2001) Endo- β -1,4-Mannanases from blue mussel, *Mytilus edulis*: purification, characterization, and mode of action. *J Biotechnol* 92:267–277
- Xu B, Sello D, Janson J (2002) Cloning and expression in *Pichia pastoris* of a blue mussel (*Mytilus edulis*) β -mannanase gene. *Eur J Biochem* 269:1753–1760
- Yokoyama R, Nishitani K (2004) Genomic basis for cell-wall diversity in plants. A comparative approach to gene families in rice and *Arabidopsis*. *Plant Cell Physiol* 45:1111–1121
- Yokoyama R, Rose JKC, Nishitani K (2004) A surprising diversity and abundance of xyloglucan endotransglucosylase/hydrolases in rice. Classification and expression analysis. *Plant Physiol* 134:1088–1099