RESEARCH NOTE

Genomewide analysis of the chitinase gene family in *Populus trichocarpa*

CONG JIANG, RUI FANG HUANG, JIA LIANG SONG, MIN REN HUANG and LI AN XU*

Jiangsu Key Laboratory for Poplar Germplasm Enhancement and Variety Improvement, Nanjing Forestry University, Nanjing 210037, People's Republic of China

[Jiang C., Huang R. F., Song J. L., Huang M. R. and Xu L. A. 2013 Genomewide analysis of the chitinase gene family in *Populus trichocarpa*. *J. Genet*. **92**, 121–125]

Introduction

The chitinase gene family is a large gene family in plants. As the first report of chitinase genes research on genome level in woody plants, here we report the identification of 37 *Populus trichocarpa* chitinase genes through a genomewide search. We divided these genes into five classes, based on sequence similarity and typical domain conservation. A strong correlation existed between gene structure and phylogeny, providing additional criteria to support classification. The distribution of chitinase genes on chromosomes was related to segmental duplications and tandem duplications revealing the expansion patterns of *P. trichocarpa* chitinase genes.

As ubiquitous enzymes, chitinases are widely present in bacteria, fungi, animals and plants. They catalyse the hydrolytic cleavage of the β-1,4-glycosidic bonds between biopolymers *N*-acetyl-glucosamine residues from the chitin molecule which is a major structural component of fungal cell wall, exoskeleton of insects and crustacean shells (Cohen-Kupiec and Chet 1998). Plant chitinases act as defense proteins responding to pathogenic attack (Santos *et al.* 2008). They can play two different roles in defense mechanism against fungal pathogens. One is to degrade the fungal cell wall and inhibit fungal growth, the other is to let the fungal pathogens release small fragments that act as elicitors (Roldan Serrano *et al.* 2007). Besides, a lot of events in the growth and development processes such as pollination, seed germination and somatic embryogenesis cannot be completed without the participation of chitinases (Passarinho and de Vries 2002). So far, chitinase genes have been cloned and characterized from more than 20 plant species. It has been demonstrated that plant chitinase genes can be induced by various sources of stress such as elicitors, wounding, salicylic acid, plant hormones and fungal pathogens (Graham and Sticklen 1994).

P. trichocarpa was selected as model forest species for genome sequencing in 2006 (Tuskan *et al.* 2006). The genome data provided a lot of information to carry on researches on gene families, and MADS-box, ARF and AP2/ERF gene families were studied in this species (Leseberg *et al.* 2006; Kalluri *et al.* 2007; Zhuang *et al.* 2008). In the present paper, chitinase genes in *P. trichocarpa* were identified. Phylogenetic relationships, conserved domains, gene structure, chromosomal locations and gene duplications were also surveyed. These results will provide a basis for further gene cloning and functional studies.

Materials and methods

A reiterated search in the *P. trichocarpa* genome (http://www.phytozome.net/poplar) using the keyword 'chitinase' was performed to identify *P. trichocarpa* chitinase genes. *Arabidopsis thaliana* chitinase genes were parsed from the *A. thaliana* database (http://www.arabidopsis.org/) using gene names reported earlier (Passarinho and de Vries 2002). Amino acid sequences were aligned using ClustalW (URL ftp://ftp-igbmc.u-strasbg.fr/pub/ClustalW) with default parameters. A phylogenetic tree was constructed using MEGA v. 5.0 (Kumar *et al.* 2008) using neighbour-joining method with 1000 bootstrap replicates. Protein domain analysis was carried out based on Pfam predictions (http://pfam.sanger.ac.uk/search). Gene structures were determined by comparing the transcript sequences with the corresponding genomic sequences. Information on the chromosomal location was gathered from *P. trichocarpa* database v. 1.1 (http://genome.jgi.doe.gov/Poptr1_1/Poptr1_1.info.html).

Results and discussion

In plants, chitinase gene family has been studied in several species, e.g. *A. thaliana* and *Oryza sativa* ssp. *Japonica* (Passarinho and de Vries 2002; Xu *et al.* 2007). Until now,

*For correspondence. E-mail: laxu@njfu.edu.cn.
Cong Jiang and Rui Fang Huang contributed equally to this work.

Keywords. chromosomal location; duplication; exon–intron structure; phylogeny.

there is no report about the whole chitinase gene family research in woody plants, although this kind of gene in *Populus* was studied several years ago (Davis *et al.* 1991). As a model woody plant, the complete genome of *P. trichocarpa* is available, providing a good opportunity for gene family studies. Our work was just carried out based on the information from the genome database.

P. trichocarpa chitinases were identified by genome search using 'chitinase' as a keyword. Their sequences were compared with sequence in GenBank. Most of the closest sequences in other species were chitinase or chitinase-like genes which indicated the high reliability of the genes identified. After removing the incomplete genes, we finally obtained 37 chitinases with complete ORF. This number is larger than that for *A. thaliana* (24 genes) and similar to that for rice (37 genes) (Xu *et al.* 2007). Their sequences were downloaded in *P. trichocarpa* genome database to investigate the phylogenetic relationships of *P. trichocarpa* chitinases. The *P. trichocarpa* chitinase genes could be divided into five classes: classes I, II, III, IV and V, as the phylogenic tree of the *P. trichocarpa* chitinases revealed (figure 1a). In detail, 11 and three chitinases were grouped into classes I and II, respectively. Thirteen chitinases belonged to class III. Classes IV and V both contained five members. From the phylogenetic tree, we found that classes I and III chitinase genes in *Populus* were much more than those in *A. thaliana*. The gene numbers in the other three classes of *Populus* were roughly the same as in *A. thaliana.* From the phylogenetic tree of the *P. trichocarpa* chitinases, all the five classes could be segregated into two categories which have relatively independent evolutionary history. They could be easily distinguished by their domains. One category, characterized by glyco_hydro_19 domain, comprised classes I, II and IV. Another one included classes III and V, which possessed glyco_hydro_18 domain. Glycoside hydrolase family 19 genes are almost exclusively present in plants, whereas glycoside hydrolase family 18 genes are distributed in various organisms, including animals, plants, fungi and bacteria.

The transcript sequences were compared with the sequence of the corresponding genomic clone to determine the structure of *P. trichocarpa* chitinase genes. High variation was observed in both gene size and numbers of exons and introns. Moreover, a strong correlation existed between exon–intron structure and phylogeny, providing additional criteria to support the classification (figure 1b). The genomic length of the chitinase genes ranged from 720 to 1978 bp and the protein coding region varies from 516 to 1513 bp. Based on different exon–intron structures, *P. trichocarpa* chitinase genes could be divided into three groups: genes with no intron, with only one intron, and with two and more introns. Chitinase genes in *O. sativa* were also divided into

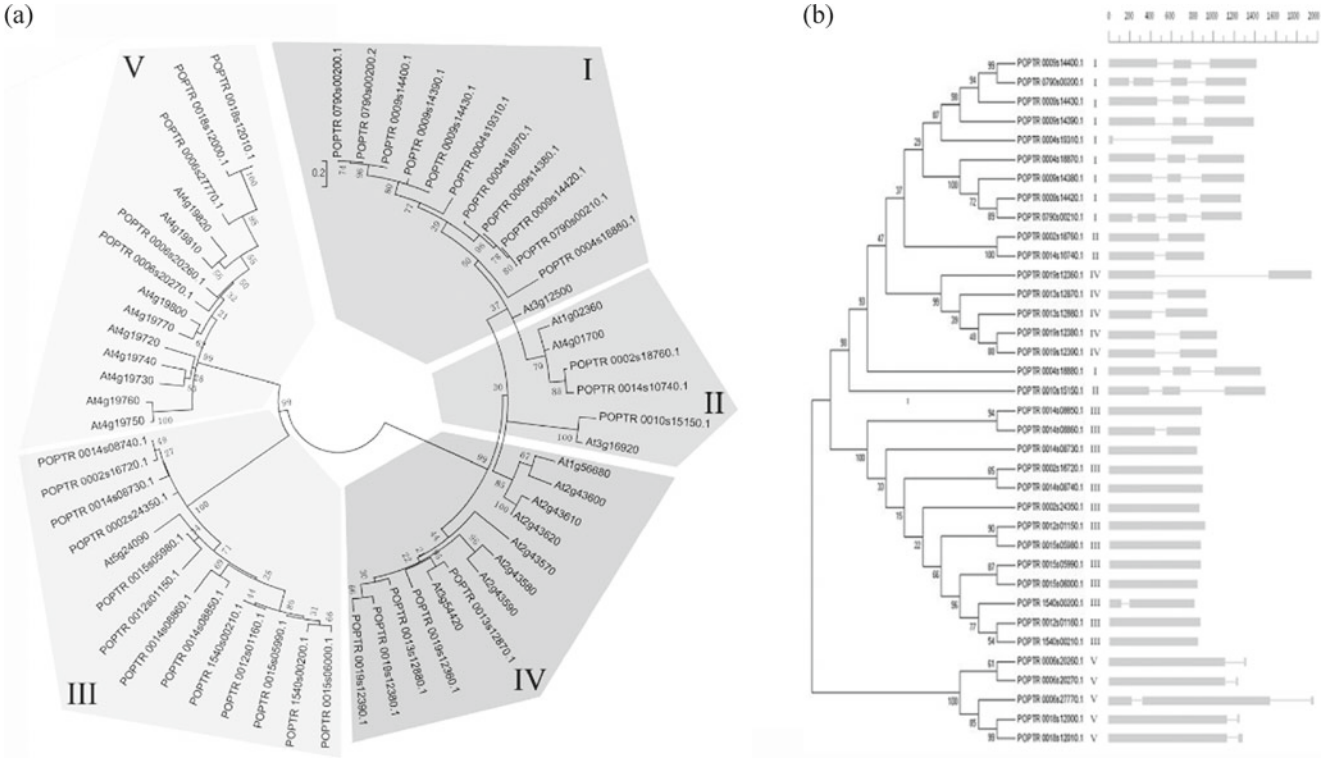

Figure 1. (a) Phylogenetic tree based on the amino acid sequence alignment of *P. trichocarpa* and *A. thaliana* chitinase genes. (b) Phylogenetic relationships among *P. trichocarpa* chitinase genes and their gene structures. Exons are illustrated by gray boxes and introns by lines. The classes to which the genes belonged are also marked.

these three groups. On the other hand, *A. thaliana* chitinase genes only had two kinds of exon–intron structures, with no genes without a single intron. For *P. trichocarpa*, most class I chitinase genes had two introns and their positions were relatively fixed. The lengths of the exons and introns were stable too. There were three members in class II. One member had two introns whereas the other two members had only one intron. All of the class IV chitinase genes had one intron, although its position was more or less unchanged in different genes, its size ranged from 119 to 1150 bp. Class V chitinase genes also had only one intron. Unlike class IV group, both the positions and sizes of the intron were relatively stable. Almost all members of class III group lacked an intron. The only exon was about 900 bp long. Genes belonging to the same class had similar exon–intron structures. However, this correlation was inconsistent in *A. thaliana* and *O. sativa*. In *A. thaliana*, one class II gene, one class III gene and two class V genes have two introns, the other genes have only one intron. In *O. sativa*, most class III genes did not contain any intron and most class V genes had only one intron, which is similar to that in *P. trichocarpa*. Moreover, the exon–intron structures of glycoside hydrolase family 19 members (classes I, II and IV) in *O. sativa* do not significantly correlate with their amino acid sequences.

The analysis on the exon–intron structures of *P. trichocarpa* chitinase genes provide a way to find out which class of chitinase genes might be of a more ancient origin. Comparison of intron positions in orthologous genes from various animals and plants showed that the common ancestor of eukaryotic lineages possessed an intron-rich genome (Rogozin *et al.* 2005). Based on this theory, one of the class II genes and all of the class I genes might have existed for the longest time and then intron loss occurred, and other members of class II and class IV genes appeared. Previous research proposed a similar conclusion (Ohme-Takagi *et al.* 1998). Class III and V chitinase genes appeared to be derived from an ancestor different from classes I, II, and IV. It was reported that the class III chitinase genes preceded the divergence between fungi and plants because plant class III chitinase has high sequence similarity with fungal chitinases (Hamel *et al.* 1997).

The chromosomal localization of chitinase genes were analysed to investigate their genomic organization in *P. trichocarpa*. Thirty genes could be positioned on eight of 19 chromosomes and the other genes were located on short scaffolds (figure 2). It was quite different from those in *A. thaliana* and *O. sativa*. In *A. thaliana*, chitinase genes are distributed in all the five chromosomes and in *O. sativa*, only one chromosome does not contain any chitinase gene. It seemed that the distribution of chitinase genes in the other two species was more dispersed than that in *P. trichocarpa*.

In detail, relatively high density of chitinase genes was observed on chromosomes 9, 14 and 19, while chromosomes 1, 3, 5, 7, 8, 11, 12, 13, 15, 16 and 17 contained no chitinase genes. The localization of different classes of chitinase genes was also revealed. Class I genes were located on chromosomes 4 and 9. Class II genes were distributed on chromosomes 2, 10 and 14. Class III genes were positioned on chromosomes 2 and 14. Class IV genes were on chromosome 19, and class V genes on chromosomes 6 and 18. Often these genes were found in clusters. A total of seven clusters were found (figure 2). The number of genes in these clusters ranged from two to six. The cluster which contained the most members was located in a 34 kb region appeared at the top of chromosome 9 where six putative genes were organized in tandem. Five chitinase genes were organized into a cluster which was located in a 45 kb region on chromosome19. On the chromosome14, there were two clusters spanning a region of 65 kb with 10 unrelated genes in between. The other three clusters containing only two genes were located in chromosomes 4, 6 and 18 respectively. The clusters were caused by tandem duplication and were seen in all classes except class II. The largest two clusters belonged to class I and class IV, respectively. It seemed that the tandem duplication occurred in these two classes in maximum frequency. Besides, clusters also existed among chitinase genes in other plant species indicating widespread occurrence of tandem duplication contributing to chitinase gene amplification. In Salicaceae, a whole-genome duplication event is believed to have occurred 60–65 Mya (million years ago), affecting roughly 92% of the *Populus* genome. Nearly 8000 pairs of paralogous genes of similar age were identified (Tuskan *et al.* 2006; Jansson and Douglas 2007). At least 25 mapped chitinase genes were located inside the duplicate blocks. However, the number of genes in corresponding paralogous segments was not always the same. This might be caused by gene deletion and tandem duplication events. These two situations were determined by the order of tandem duplication and segmental duplication occurrences. Sequence similarity comparison would help us to find out which one occurred earlier (Lohmann *et al.* 2010). POPTR_0002s18760.1 on chromosome 2 and POPTR_0014s10740.1 on chromosome 14 were the direct results of segmental duplication. Tandem duplication did not occur. Another segmental duplication occurred between chromosomes 2 and 14 which involved three genes, POPTR_0002s16720.1, POPTR_0014s08730.1 and POPTR_0014s08740.1. As POPTR_0014s08740.1 had higher similarity to POPTR_0002s16720.1 compared with POPTR_0014s08730.1, the tandem duplication probably occurred before genome segmental duplication. The formation of the clusters on chromosomes 4 and 9 was relatively complicated. It was caused by genome segmental duplication and two tandem duplications. Comparison of the sequences similarity indicated that tandem duplication appeared to predate the genome segmental duplication whereas the other tandem duplication occurred after the genome segmental duplication. Gene deletion and the latter tandem duplication caused the number of genes on chromosome 9 to be more than that on chromosome 4. The cluster on chromosome 19 was rather special. They were located outside duplicate blocks. Therefore, they did not take part in the segmental

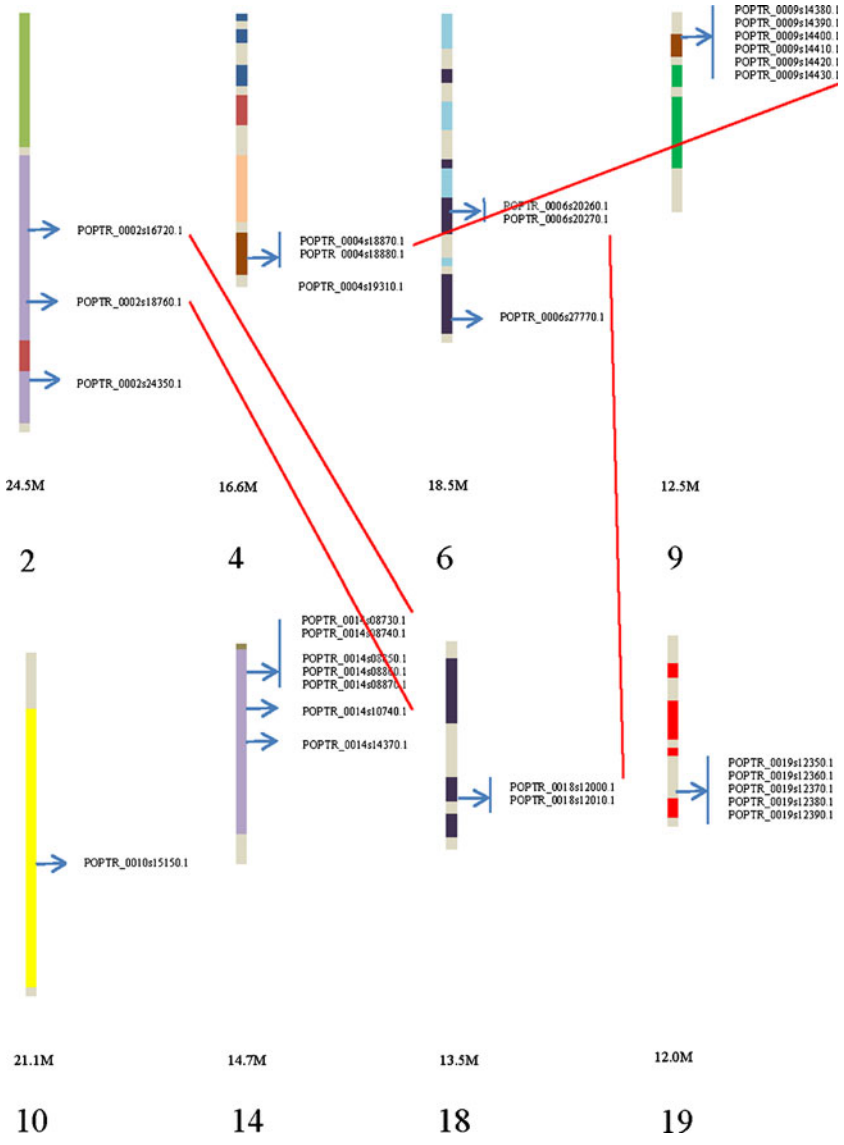

Figure 2. Genomic localization of *P. trichocarpa* chitinase genes. Homologous genome blocks shared by two chromosomes are shown in same colour. The number and length of chromosomes are indicated below each chromosomes. Related genes are connected by red lines. The chromosomes in gray were not involved in genome segmental duplication events.

duplication. The other genes did not have homologous genes in the corresponding duplicate block. The possible reason is that the genes were lost after the segmental genome duplication.

Expression analysis of chitinase genes have been performed in various species. Most genes are induced by biotic and abiotic stresses. Even though the functional differences between classes were not so clear. Here, expression analysis of five genes belong to five classes respectively was carried out to find the potential functions of them.

POPTR_0009s14390.1, a member of class I chitinase genes, was highly expressed after inoculation with a pathogenic fungus *Marssonina brunnea*. The expression level increased continuously and peaked late during the infectious process. POPTR_0010s15150.1, which represents class II chitinase genes, was induced by methyl jasmonate (MeJA) rather than other treatments. POPTR_0012s01150.1 is a class III chitinase gene. It was sensitive to elicitor and wound inducements but insensitive to salicylid acid (SA) and MeJA treatments. When *Populus* was infected by *M. brunnea*, the expression of this gene increased slightly. POPTR_0013s12870.1, a class IV chitinase gene, showed higher expression after the plant was treated with elicitor and wounding. However, this gene did not response to any other treatments. As a class V chitinase gene, POPTR_0006s20260.1 was induced by MeJA and fungal

infection. From the results, we found genes in different classes have a lot of overlapping functions, despite considerable functional divergence.

Acknowledgements

This research was supported by the Doctorate Fellowship Foundation of Nanjing Forestry University, the Graduate Research Innovation Project of Jiangsu Province, China and the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD).

References

Cohen-Kupiec R. and Chet I. 1998 The molecular biology of chitin digestion. *Curr. Opin. Biotech.* **9**, 270–277.

Davis J. M., Clarke H. R. G., Bradshaw Jr H. D. and Gordon M. P. 1991 *Populus* chitinase genes: structure, organization, translated sequences to herbaceous plant chitinases. *Plant Mol. Biol.* **17**, 631–639.

Graham L. S. and Sticklen M. B. 1994 Plant chitinases. *Canad. J. Bot.* **72**, 1057–1083.

Hamel F., Boivin R., Tremblay C. and Bellemare G. 1997 Structural and evolutionary relationships among chitinases of flowering plants. *J. Mol. Evol.* **44**, 614–624.

Jansson S. and Douglas C. J. 2007 *Populus*: a model system for plant biology. *Annu. Rev. Plant Biol.* **58**, 435–458.

Kalluri U. C., DiFazio S. P., Brunner A. M. and Tuskan G. A. 2007 Genomewide analysis of Aux/IAA and ARF gene families in *Populus trichocarpa*. *BMC Plant Biol.* **7**, 59.

Kumar S., Nei M., Dudley J. and Tamura K. 2008 Mega: a biologist-centric software for evolutionary analysis of DNA and protein sequences. *Brief Bioinform*. **9**, 299–306.

Leseberg C. H., Li A., Kang H., Duvall M. and Mao L. 2006 Genomewide analysis of the MADS-box gene family in *Populus trichocarpa*. *Gene* **378**, 84–94.

Lohmann G. V., Shimoda Y., Nielsen M. W., Jorgensen F. G., Grossmann C., Sandal N. *et al.* 2010 Evolution and regulation of the Lotus japonicus LysM receptor gene family. *Mol. Plant Microbe Interact.* **23**, 510–521.

Ohme-Takagi M., Meins Jr F. and Shinshi H. 1998 A tobacco gene encoding a novel basic class II chitinase: a putative ancestor of basic class I and acidic class II chitinase genes. *Mol. Gen. Genet.* **259**, 511–515.

Passarinho P. A. and de Vries S. C. 2002 Arabidopsis chitinases: a genomic survey. In *The Arabidopsis Book*, pp. 1–25. The American Society of Plant Bilogist, Rockville, USA.

Rogozin I. B., Sverdlov A. V., Babenko V. N. and Koonin E. V. 2005 Analysis of evolution of exon-intron structure of eukaryotic genes. *Brief Bioinform.* **6**, 118-134.

Roldan Serrano A., Luna del Castillo J., Jorrin Novo J., Fernández Ocaña A. and Gómez Rodríguez M. 2007 Chitinase and peroxidase activities in sunflower hypocotyls: effects of BTH and inoculation with Plasmopara halstedii. *Biol. Plant.* **51**, 149–152.

Santos P., Fortunato A., Ribeiro A. and Pawlowski K. 2008 Chitinases in root nodules. *Plant Biotech.* **25**, 299–307.

Tuskan G. A., Difazio S., Jansson S., Bohlmann J., Grigoriev I., Hellsten U. *et al.* 2006 The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* **313**, 1596–1604.

Xu F., Fan C. and He Y. 2007 Chitinases in *Oryza sativa* ssp. *japonica* and *Arabidopsis thaliana*. *J. Genet. Genomics* **34**, 138–150.

Zhuang J., Cai B., Peng R. H., Zhu B., Jin X. F., Xue Y. *et al.* 2008 Genomewide analysis of the AP2/ERF gene family in *Populus trichocarpa*. *Biochem. Biophy. Res. Commun.* **371**, 468–474.

Received 8 May 2012, in final revised form 12 December 2012; accepted 3 January 2013
Published on the Web: 2 April 2013