Gene 546 (2014) 408–416

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Tracing the origin and evolution of plant TIR-encoding genes

Xiaoqin Sun [a], Hui Pang [a], Mimi Li [a], Jianqun Chen [b,*], Yueyu Hang [a,**]

[a] Institute of Botany, Jiangsu Province and Chinese Academy of Sciences, Jiangsu Province Key Laboratory for Plant Ex Situ Conservation, Nanjing 210014, China
[b] State Key Laboratory of Pharmaceutical Biotechnology, School of Life Sciences, Nanjing University, Nanjing 210093, China

ARTICLE INFO

Article history:
Received 10 January 2014
Received in revised form 17 April 2014
Accepted 24 April 2014
Available online 28 April 2014

Keywords:
TIR-encoding genes
T genes
Disease resistance genes
Gene evolution
Monocots

ABSTRACT

Toll-interleukin-1 receptor (TIR)-encoding proteins represent one of the most important families of disease resistance genes in plants. Studies that have explored the functional details of these genes tended to focus on only a few limited groups; the origin and evolutionary history of these genes were therefore unclear. In this study, focusing on the four principal groups of TIR-encoding genes, we conducted an extensive genome-wide survey of 32 fully sequenced plant genomes and Expressed Sequence Tags (ESTs) from the gymnosperm *Pinus taeda* and explored the origins and evolution of these genes. Through the identification of the TIR-encoding genes, the analysis of chromosome positions, the identification and analysis of conserved motifs, and sequence alignment and phylogenetic reconstruction, our results showed that the genes of the TIR-X family (TXs) had an earlier origin and a wider distribution than the genes from the other three groups. TIR-encoding genes experienced large-scale gene duplications during evolution. A skeleton motif pattern of the TIR domain was present in all spermatophytes, and the genes with this skeleton pattern exhibited a conserved and independent evolutionary history in all spermatophytes, including monocots, that followed their gymnosperm origin. This study used comparative genomics to explore the origin and evolutionary history of the four main groups of TIR-encoding genes. Additionally, we unraveled the mechanism behind the uneven distribution of TIR-encoding genes in dicots and monocots.

1. Introduction

Disease resistance genes (*R* genes) in plants are important functional genes that prevent and respond to the invasion of pathogens. Most *R* genes encode members of an extremely polymorphic superfamily of nucleotide-binding leucine-rich repeat (NLR) receptors (Maekawa et al., 2011). Specific NLR proteins can be activated by specific pathogen effectors via direct interaction as receptor and ligand respectively (Dodds et al., 2006). Alternatively, a specific NLR can be activated by sensing effector-mediated alteration in a host virulence target or a decoy protein of that target (Dangl and Jones, 2001; van der Hoorn and Kamoun, 2008). In further support of the latter, experimentally established interaction networks of the *Arabidopsis thaliana* proteins and the bacterial and oomycete effectors have revealed that independently evolved effectors converge onto the hubs of the immune system network (Mukhtar et al., 2011). NLR activation coordinates effector-triggered immunity that limits pathogen proliferation.

The *R* genes encode a number of conserved domains, including the nucleotide binding site (NBS), leucine-rich repeat (LRR), protein kinase (Pkinase), toll/interleukin-1 receptor (TIR) and coiled-coil (CC) domains (Jacob et al., 2013). TIR-encoding genes represent one of the most important families of disease resistance genes in plants and act as signal mediators that respond to pathogens (Katagiri and Tsuda, 2010; Lewis et al., 2010; Römer et al., 2009; Zhang et al., 2012). These genes can be further categorized into several groups according to the variable C-terminal domains: TIR-NBS-LRR (TNL), TIR-NBS (TN), TIR-LRR (TL) and TIR-X (TX, where X is any domain other than NBS or LRR) (Nandety et al., 2013). The function of TIR-encoding genes is believed to be involved in immunity, yet some studies on Arabidopsis TIR-encoding genes suggest their possible roles beyond. Arabidopsis CHS1 which encodes a TN protein confers cold resistance by limiting chloroplast damage and cell death at low temperature (Zbierzak et al., 2013). Overexpression analysis and phylogenetic highly conservation suggest that the TNX proteins may have other nondefense roles in monocots and other plants (Nandety et al., 2013).

Studies of the TIR-encoding genes (mostly involving TNLs) have primarily focused on gene cloning, the functional verification of the genes and the molecular mechanisms of disease resistance (Anderson et al., 1985; Franchel et al., 2012; Gassmann et al., 1999; Lawrence et al., 1995; Meyers et al., 1999; Parker et al., 1997; Whitham et al., 1994). In contrast, the analysis of genetic polymorphisms and gene evolution has been limited to only a few genes or individual plant species (reviewed in Jacob et al., 2013). The functions of TXs and TNs (the

Abbreviations: CC, coiled-coil; HNL, α/β-hydrolase-NBS-LRR; LRR, leucine-rich repeat; NBS, nucleotide binding site; NLR, nucleotide-binding leucine-rich repeat; Pkinase, protein kinase; PNL, protein-kinase-NBS-LRR; TIR, Toll-interleukin-1 receptor.

* Correspondence to: J. Chen, State Key Laboratory of Pharmaceutical Biotechnology, School of Life Sciences, Nanjing University, 22 Hankou Road, Nanjing 210093, China.
** Correspondence to: Y. Hang, Jiangsu Province Key Laboratory for Plant Ex Situ Conservation, Institute of Botany, Jiangsu Province and Chinese Academy of Sciences, 1 Qianhuhoucun Road, Nanjing 210014, China.
E-mail addresses: chenjq@nju.edu.cn (J. Chen), hangyueyu@gmail.com (Y. Hang).

http://dx.doi.org/10.1016/j.gene.2014.04.060
0378-1119/

other two groups of TIR-encoding genes) remain unclear; however, the diversification and conservation of these genes suggest that some of these proteins have important functions (Nandety et al., 2013). For example, *Arabidopsis* CHS1, which encodes a TN protein, confers cold resistance (Zbierzak et al., 2013). Overexpression analysis of *Arabidopsis* TXs and TNs in tobacco or *Arabidopsis* suggested that TXs and TNs might play roles in disease resistance (Nandety et al., 2013).

Previous studies have shown that the TLs and TXs originated from eubacteria (Yue et al., 2012) and that the TNs and TNLs were first detected in the moss *Physcomitrella patens* (Akita and Valkonen, 2002; Xue et al., 2012; Yue et al., 2012). The frequency of the appearance of conserved motifs in the NBS domain of TNLs has changed during the evolutionary progression from mosses to tracheophytes and from gymnosperms to spermatophytes (Yue et al., 2012). Some studies have also proposed that TNLs evolved differently during the early differentiation of dicots and monocots. Indeed, TNLs are completely absent in monocots, while these genes differentiated and expanded to form an important family of resistance genes in dicots (Bai et al., 2002; Cannon et al., 2002; Pan et al., 2000). In support of this hypothesis, TNLs could not be identified using either degenerate PCR or database searching in nine cereal genera (Cannon et al., 2002; Meyers et al., 1999); furthermore, TNLs could not be detected in monocot species other than cereals, including *Dracaena marginata* Lam., *Sansevieria trifasciata* Hort ex D. Prain, *Spathiphyllum* sp., *Carex blanda* Dewey, *Musa acuminata* Colla and *Elaeis guineensis* Jacg (Tarr and Alexander, 2009). In contrast, dozens to hundreds of TNLs have been discovered in dicots (reviewed by Jacob et al., 2013), indicating that the TNLs have undergone massive expansion in these plants. Unfortunately, few studies have addressed the evolution of TNs, TLs or TXs. The few reports describing the phylogenetic relationships and chromosomal positions of TNs, TXs and TNLs in *A. thaliana* have suggested that TXs were derived from TNs or TNLs and co-evolved as functional units (Meyers et al., 2002). Data from poplar and grapevine also support this finding (Yang et al., 2008).

Because few studies have investigated the origin and evolutionary history of the TIR-encoding genes and have instead mainly focused on TNLs, the evolution of the four major groups of TIR-encoding genes remains unclear. In this study, the origin, evolutionary history and relationships between the four groups of TIR-encoding genes in the plant kingdom were investigated using 32 whole-genome sequenced plants, including two members of Chlorophyta.

2. Materials and methods

2.1. Data sampling

A total of 32 different whole-genome sequenced species were selected, including two species in Chlorophyta, one species in Bryophyta, one species in Pteridophyta and 28 species in Angiospermae. The TIR-encoding genes in *Pinus taeda* L. (Meyers et al., 2002) were also incorporated into the dataset to represent the gymnosperms. These species were sampled because they represent the five major taxonomic phyla (Chlorophyta, Bryophyta, Pteridophyta, Gymnospermae and Angiospermae) in the plant kingdom. The complete genome sequences and corresponding annotation information were downloaded from online databases (Table S1).

The complete set of TIR-encoding genes was identified in a reiterative manner. Three analytical steps were followed to compile the final set of sequences. First, for each species, protein entries matching the TIR domain (Pfam: PF01582) were identified as TIR-encoding genes using BLASTP with an *E*-value cutoff of 10^{-4}. The presence of the TIR

Table 1
Distribution of TIR-encoding genes in the sequenced plant genomes.

Phylum	Class	Species	All	TNL	TN	TL	TX	
							All	T[a]
Chlorophyta		*Chlamydomonas reinhardtii*	2	0	0	0	2	2
		Volvox carteri	1	0	0	0	1	1
Bryophyta		*Physcomitrella patens*	13	7	3	0	3	3
Pteridophyta		*Selaginella moellendorffii* Hieron.	1	0	1	0	0	0
Gymnospermae		*P. taeda* L.	5	2	2	0	1	1
Angiospermae	Dicotyledoneae	*Aquilegia coerulea* E. James	1	0	0	0	1	1
		Vitis vinifera L.	31	15	6	1	9	7
		Manihot esculenta L.	45	26	3	2	14	8
		Populus trichocarpa (Torr. & Gray)	177	84	28	9	56	33
		Ricinus communis L.	43	24	7	0	12	5
		Glycine max Merr.	182	53	92	0	37	36
		Medicago truncatula Gaertn.	272	26	164	0	82	82
		Lotus *corniculatus* L. var. *japonicus* Regel	157	25	47	1	84	68
		Prunus persica (L.) Batsch	169	128	15	1	25	19
		Cucumis sativis L.	25	12	8	0	5	4
		Eucalyptus grandis Hill ex Maiden	754	281	204	11	258	194
		Citrus clementina Hort. ex Tan.	112	57	29	3	23	16
		Citrus sinensis (L.) Osb.	113	33	33	4	43	27
		Carica papaya L.	11	6	1	0	4	2
		Arabidopsis thaliana L.	162	102	19	0	41	23
		Arabidopsis lyrata L.	158	90	23	0	45	30
		Thellungiella halophila (C. A. Mey.) O. E. Schulz	98	52	7	0	39	16
		Brassica rapa L.	155	91	23	0	41	22
		Mimulus guttatus DC.	2	0	0	0	2	2
		Solanum lycopersicum L.	37	18	5	2	12	6
		Solanum tuberosum L.	102	44	15	2	41	24
	Monocotyledoneae	*Musa acuminata* Colla	1	0	0	0	1	1
		Brachypodium distachyon (L.) Beauv	1	0	0	0	1	1
		Oryza sativa L.	2	0	0	0	2	2
		Sorghum bicolor (L.) Moench	2	0	0	0	2	2
		Setaria italica (L.) Beauv.	2	0	0	0	2	2
		Panicum virgatum L.	4	0	0	0	4	4
		Zea mays L.	3	0	0	0	3	3
Total		33	2843	1176	735	36	896	647

[a] The column shows the distribution of T genes, a sub-category of TXs which contain only the TIR domain(s).