Plant Omics Journal POJ

POJ 12(02):109-119 (2019) ISSN: 1836-3644
doi: 10.21475/POJ.12.02.19.p1825

Genome-wide characterization and expression patterns of chitinase genes in the pigeonpea (*Cajanus cajan* (L.) Millsp.) genome

Ajay Kumar Mahato[1,2], Ajay Kumar Sharm[2], Nagendra Kumar Singh[1]

[1]ICAR-National Research Centre on Plant Biotechnology (NRCPB), IARI, Pusa Campus, New Delhi 110012, India
[2]Meerut Institute of Engineering and Technology, A.P.J. Abdul Kalam Technical University (APJAKTU), Meerut, Uttar Pradesh 226031, India

***Corresponding author: nksingh4@gmail.com**

Abstract

Plant chitinases are involved in defense as well as a wide range of physiological functions in plants, including germination, embryogenesis, flowering, and senescence. This study was conducted to identify and annotate the chitinase-related genes from the pigeonpea genome version 2.0, their chromosomal localization and phylogenetic relationship with chitinase genes from 13 different plant species. Here, we report the identification of 34 putative chitinase genes in the pigeonpea genome. These 34 genes encode proteins belonging to two functional domain families, and are subdivided into four classes matching four of the five chitinase classes in Arabidopsis. These chitinase genes are present in clusters on the chromosome. We investigated the expression patterns of these chitinases in 29 different tissues at five developmental stages. There was clear clustering of the chitinase genes into three groups based on their expression patterns in tissues. We identified two chitinase genes C_caj-24 and C_caj-25 that were highly expressed in all tissues as well as other chitinase genes with tissue-specific expression, which suggests that they play important roles in plant defense at specific developmental stages. This information on pigeonpea chitinases could be useful for the development of pigeonpea varieties that are resistant to insect pests and fungal diseases.

Keywords: Chitinase; *Cajanus cajan*; expression pattern; phylogeny.
Abbreviations: GlcNAc_β-(1,4)-linked N-acetylglucosamine, DR_Defense response, PR_Pathogenesis-related.

Introduction

Pigeonpea (*Cajanus cajan* L. Millspaugh), which belongs to the family Fabaceae, has a diploid (2n = 22) genome size of 858 Mbp, it is a perennial legume that is widely grown in tropical and subtropical regions as a food crop and is commonly consumed in Asia, Africa, and Latin America (Greilhuber et al., 1998). Asia contributes most of the world's pigeonpea production (77.8%), with India contributing (63.4%), followed by Myanmar, Malawi, and the United Republic of Tanzania (FAOSTAT, 2016). Pigeonpea plants are susceptible to a large number of diseases, including leaf blight, seedling rot, Fusarium wilt, and leaf spot. Some diseases, such as Alternaria blight (caused by *Alternaria tenuissima* and *A. alternata*), Phyllosticta leaf spot (*Phyllosticta cajani*), and Fusarium leaf blight (*Fusarium semitectum*) can lead to significant yield losses (Reddy et al., 1993). Chitinases form the first line of a plant's defense against fungal pathogens. However, no in-depth study of the chitinase genes in the pigeonpea has been performed.

Defense response genes, including pathogenesis-related (PR) proteins, play an important role in plant disease resistance and immunity against a range of biotic and abiotic stresses. Chitinases (EC 3.2.2.14), which are also defined as "glycosyl hydrolases" are an important category of PR proteins that catalyze the breakdown of chitin, a polymer of β-(1,4)-linked N-acetylglucosamine (GlcNAc) (Alvarez and Konopka, 2007; Naseem et al., 2011) that is a major component of arthropod exoskeletons and fungal cell walls, and thus aid in a plant's defense against insect pests and fungal diseases (Liu et al., 2005; Maximov et al., 2006; Xiao et al., 2007). Based on the sequences of their catalytic domains, chitinases have been classified into two families, glycoside hydrolase-18 (GH-18) and glycoside hydrolase-19 (GH-19) (Henrissat, 1991). GH-18 chitinases are distributed in bacteria, yeast, fungi, plants, and animals, whereas GH-19 chitinases are almost exclusive to plants (Passarinho and de Vries, 2002). GH-18 and GH-19 not only differ in their protein domains and 3D structures but also in their biochemical properties, including their product form (b-anomeric and s-anomeric) (Brameld et al., 1998a). While GH-18 chitinases function by using a substrate-assisted catalysis model, GH-19 chitinases use a general acid-base mechanism (Hart et al., 1995; Brameld et al., 1998b; Garcia-Casado et al., 1998).

The expression of plant chitinase genes can be induced in various ways, by elicitors, wounding, salicylic acid, plant hormones, fungal pathogens (Graham and Sticklen, 1994), and abiotic stresses such as osmotic shock, salt, cold, and

heavy metals (Grover, 2012). Plant chitinases can be subdivided into five classes (I–V) based on their sequence and structure in Arabidopsis (Neuhaus et al., 1996). The chitinase genes of three classes (I, II, and IV) belong to the GH-19 family, while genes from classes III and V belong to the GH-18 family (Henrissat and Bairoch, 1993; Neuhaus et al., 1996; Passarinho and de Vries, 2002). Plant chitinases degrade insect and fungal pathogen chitinand release chito-oligosaccharides that act as elicitors to activate plant immunity, which is a very effective immune strategy used by plants against pathogens and herbivores (Shibuya and Minami, 2001; Passarinho and de Vries, 2002; Wan et al., 2008; Stacey and Shibuya, 1997; Felton and Korth, 2000). Chitinase expression is low in several plant organs during specific developmental stages, indicating that some chitinases are involved in plant growth and developmental process (Collinge et al., 1993; Patil et al., 2000). Chitinases have also been used as important targets for crop improvement through genetic engineering (Legrand et al., 1987; Graham and Sticklen, 1994; Van Loon and Van Strien, 1999).

The availability of the pigeonpea draft genome version 2.0 and RNA-seq data in the public domain provide an opportunity for genome-wide identification, classification, and expression analysis of the chitinase genes in pigeonpea (Singh et al., 2012; Varshney et al., 2012; Pazhamala et al., 2017; Mahato et al., 2018). Hence, the present study was aimed at the identification, categorization, and expression pattern analysis of the chitinase genes in the pigeonpea genome.

Results and Discussion

Genome-wide identification and classification of pigeonpea chitinase genes

In this study, we identified a total of 34 genes predicted to encode chitinases in the improved draft genome version 2.0 of the pigeonpea variety "Asha" (Mahato et al., 2018) based on sequence similarity and the presence of conserved domains (**Table 1**). The number of predicted chitinase genes in the pigeonpea genome is comparable to the numbers predicted in *Brassica rapa*, *Populus trichocarpa*, *Oryza sativa*, and *Hevea brasiliensis* (33–39 genes) (Jingjing et al., 2018; Jiang et al., 2013; Xu et al., 2007; Misra, 2015); higher than the numbers in *Arabidopsis thaliana* and *Musa acuminata* (24–26 genes) (Passarinho and de Vries, 2002; Backiyarani et al., 2015), but lower than the number in *Eucalyptus grandis* and *Gossypium* species (47–116 genes) (Tobias et al., 2017; Xu et al., 2016). The size of the putative pigeonpea chitinase genes ranged from 414–1059 bp, with 1–4 exons (**Table 1**) and the relative lengths of the introns and exons in the pigeonpea chitinase genes are illustrated in **Fig 1**. Thirteen pigeonpea chitinase genes belonging to the GH-18 family have no introns, another 13 genes have one intron, and the remaining eight genes have two or more introns, supporting the concept that genes related to biotic and abiotic stresses generally have fewer introns (Jeffares et al., 2008). All the identified pigeonpea chitinase genes are supported by their high sequence similarity with the pigeonpea EST/TSA sequences.

Based on the presence of functional domains, the 34 chitinase genes were divided into two families; 23 genes belong to the GH-18 family, and 11 genes belong to the GH-19 family (**Table 1**). Interestingly, only 2 of the 34 predicted chitinase genes (C_Caj_26 and C_Caj-31) contain a chitin-binding domain (CBD), and these belong to the GH-19 family (**Fig 1**). Surprisingly, four chitinase genes, all from the GH-18 family (CC_chi-9, CC_chi-10, CC_chi-20, and CC_chi-21) do not have a signal peptide (**Fig 1**). In silico analysis of the subcellular localization showed that all 34 chitinases are of the secretory type and are located in the extracellular space.

Chromosomal location and conserved motifs in the pigeonpea chitinases

Of the 34 chitinase genes identified in the improved pigeonpea draft genome (Mahato et al., 2018), only 23 (67.4%) could be mapped to chromosome pseudomolecules due to the limited genome coverage and anchoring of the pigeonpea genome scaffolds (Varshney et al., 2012). The 23 genes mapped to seven chromosomes of the pigeonpea genome, while the remaining four chromosomes contained no chitinase genes. The mapping results also showed that 17 of the 23 mapped genes were present in just six clusters, suggesting their origin by tandem duplication (**Fig 2**). The largest cluster, containing six chitinase genes, was located on chromosome 6, another cluster of three genes was located on chromosome 1, four clusters of two genes each were located on chromosomes 2, 6, and 11, and six loci, each with a single gene, were present on chromosomes 1, 3, 6, 7, and 10 (**Fig 2).** Similar clusters of chitinase genes have also been reported in cotton, poplar, eucalyptus, and rice (Xu et al., 2017; Jiang et al., 2013; Tobias et al., 2017; Xu et al., 2007). Such clustering of gene families in a genome is thought to arise thorough tandem gene duplication (Schauser et al., 2005). Analysis of the structural motifs in the pigeonpea chitinase proteins showed 10 conserved motifs (**Fig 3**). Seven of these motifs were specific to the GH-18 chitinase family, and their peptide lengths were 15 (motifs 06 and 07), 21 (motif 04), 29 (motifs 05 and 05), 42 (motif 02), and 44 (motif 01) amino acids. The three remaining motifs were specific to the GH-19 chitinase family, all with a peptide length of 41 amino acids (**Table 3**). Further analysis of the distribution of the conserved motifs in the pigeonpea chitinases revealed that in the GH-18 family, eight genes possessed all seven conserved motifs, one of which was common to the GH-18 family, while in the GH-19 family, six chitinase genes contained four motifs, three of which were specific to the GH-19 family. Two chitinase genes had only three GH-19 family-specific motifs (**Fig 3)**. Multiple sequence alignment and clustering of the 34 pigeonpea chitinase genes showed a high degree of sequence conservation among the genes, which were grouped into four clades, labeled I–IV in **Fig 3 and Fig. S1**. Clades I, II, and III belong to the GH-18 chitinase family, while clad IV includes 11 genes from the GH-19 family and 4 genes from the GH-18 family **(Fig 3)**.

Phylogenetic relationships among the pigeonpea chitinase genes

A phylogenetic analysis was carried out to investigate the evolutionary relationships among the pigeonpea chitinases and chitinases from 13 other plant species, including seven legumes: C*icer arientum*, *Phaseolus vulgaris*, *Glycine max*,

Table 1. Characteristic features of 34 EST/TSA supported chitinase genes in the pigeonpea genome, all coding for secretary type proteins.

Gene Id.	Chromosome	Chromosomal position		Protein Length (aa)	Gene Length (bp)	No. of Exons	Signal Peptide	HMMER protein domain
		Start	End					
C_Caj_1	6	7113833	7114429	298	897	2	Yes	GH-18
C_Caj_2	6	7113918	7114421	212	639	1	Yes	GH-18
C_Caj_3	6	7113918	7114421	212	639	1	Yes	GH-18
C_Caj_4	6	7113924	7114428	292	879	1	Yes	GH-18
C_Caj_5	6	7113845	7114428	295	888	1	Yes	GH-18
C_Caj_6	6	7113845	7114428	295	888	1	Yes	GH-18
C_Caj_7	2	29809512	29810411	299	900	1	Yes	GH-18
C_Caj_8	2	29809512	29810411	299	900	1	Yes	GH-18
C_Caj_9	-	-	-	177	534	1	No	GH-18
C_Caj_10	-	-	-	254	765	2	No	GH-18
C_Caj_11	-	-	-	296	891	2	Yes	GH-18
C_Caj_12	-	-	-	296	891	2	Yes	GH-18
C_Caj_13	3	27899305	27900123	272	819	1	No	GH-18
C_Caj_14	1	17482436	17482969	294	885	2	Yes	GH-18
C_Caj_15	1	17486532	17487058	290	873	3	Yes	GH-18
C_Caj_16	1	17486532	17487058	301	906	3	Yes	GH-18
C_Caj_17	1	17486532	17487058	301	906	3	Yes	GH-18
C_Caj_18	-	-	-	330	993	2	Yes	GH-18
C_Caj_19	-	-	-	330	993	2	Yes	GH-18
C_Caj_20	2	29886636	29887250	204	615	1	No	GH-18
C_Caj_21	2	29886636	29887250	204	615	1	No	GH-18
C_Caj_22	-	-	-	339	1020	1	Yes	GH-18
C_Caj_23	-	-	-	344	1035	1	Yes	GH-18
C_Caj_24	7	17048905	17049320	321	966	3	Yes	GH-19
C_Caj_25	-	-	-	318	957	3	Yes	GH-19
C_Caj_26	10	872003	872430	279	840	2	Yes	GH-19
C_Caj_27	6	9443095	9443501	227	684	2	Yes	GH-19
C_Caj_38	3	26676503	26676961	272	819	2	Yes	GH-19
C_Caj_29	11	41904049	41904513	276	831	2	Yes	GH-19
C_Caj_30	11	41904049	41904513	276	831	2	Yes	GH-19
C_Caj_31	-	-	-	352	1059	4	Yes	GH-19
C_Caj_32	-	-	-	137	414	2	Yes	GH-19
C_Caj_33	6	15379753	15380125	269	810	3	Yes	GH-19
C_Caj_34	6	15379753	15380125	269	810	3	Yes	GH-19

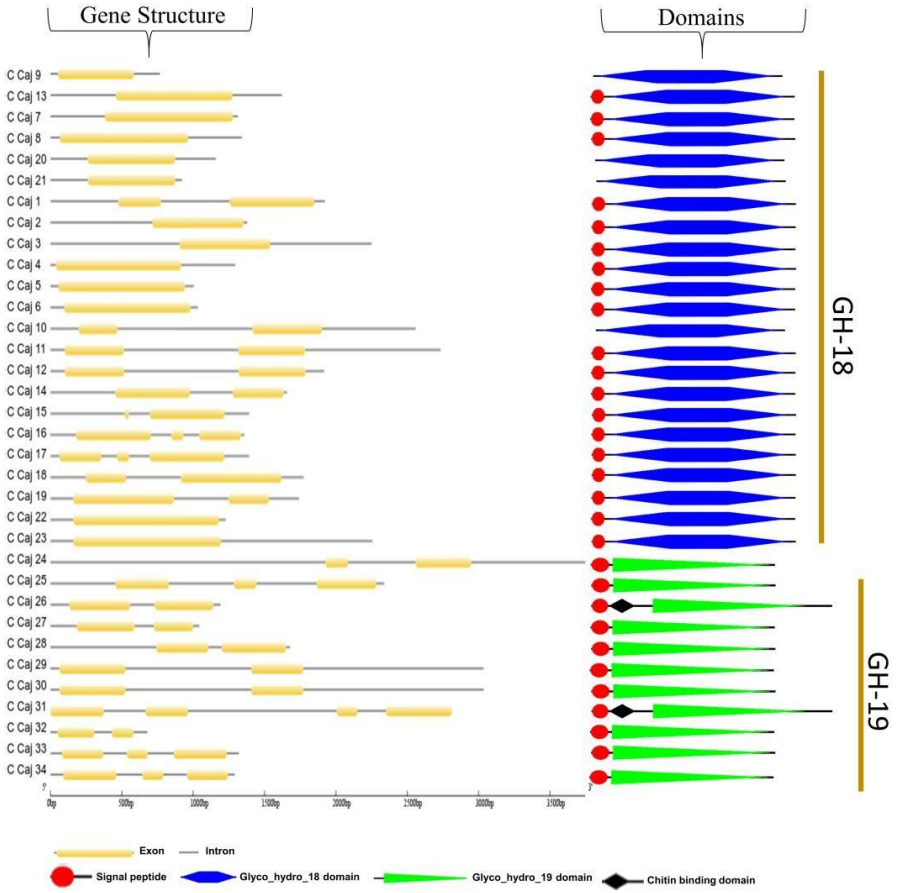

Fig 1. Gene structure and conserved domain architecture of 34 chitinase genes identified in the pigeonpea genome version 2.0. The exons are represented by boxes shaded light yellow color and introns are by black line (left panel). The protein domain architecture in respective genes are shown with different shapes, gene clusters with the glycosyl hydrolases domains 18 and 19 are also marked as GH-18 and GH-19, respectively (right panel).

Table 2. MEME identified motifs in pigeonpea chitinase gene.

Chitin class	Name	Motif	E-value	Sites	Width
	Motif 01	NGNEGTLAEACATGNYAIVIIAFLSTFGNGQTPQJNLAGHCDPS	1.1e-503	21	44
	Motif 02	DARQVASYLWNNFLGGQSSSRPLGDAVLDGIDFDIEGGSTQH	1.6e-461	18	42
	Motif 03	APQCPFPDAWLGSAJETGLFDYVWVQFYN	2.1e-409	23	29
GH-18	Motif 04	LPAIKGSSKYGGVMLWSRYYD	5.70E-263	21	21
	Motif 05	NGCTKLSSEIKSCQAKGIKVLLSJGGGAG	2.70E-270	20	29
	Motif 06	AGSGYIPPDVLTSQV	1.10E-96	20	15
	Motif 07	WDELARALKGYSKQK	8.90E-99	21	15
	Motif 08	TRKREIAAFLAQTSHETTGGWATAPDGPYAWGLCFVEEVSP	8.20E-173	9	41
GH-19	Motif 09	YPCYPGKTYYGRGPIQLSWNYNYGPAGKALGFDLLNNPELV	2.80E-187	10	41
	Motif 10	PSCHDVIVGRWKPTKADTAANRVPGYGVVTNIINGGLECGI	2.40E-125	8	41

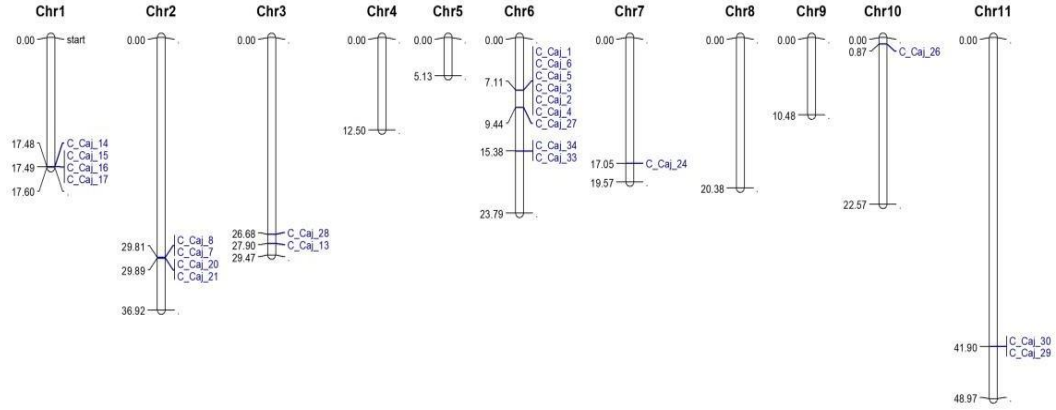

Fig 2. Distribution of chitinase genes on 11 pigeonpea chromosomes. The 23 chitinase genes of total 34 identified gene set were mapped to the 7 out of 11 chromosomes. The genes in clusters are marked with black straight line, and their mapping position on respective chromosomes is shown in Mbp.

Table 3. Number of chitinase genes in 14 plant species based on HMMER protein domain search used for comparative and phylogenetic analysis.

S. No.	Species	GH-18 type	GH-19 type	Others	Total
1	Pigeonpea (*Cajanus cajan)*	23	11	0	34
2	Chickpea (*Cicer arientum*)	20	7	5	32
3	Common bean (*Phaseolus vulgaris*)	17	12	5	34
4	Soybean (*Glycine max*)	26	9	0	35
5	Barrel clover (*Medicago truncatula*)	16	11	10	37
6	Lotus (*Lotus japonicus*)	21	8	6	35
7	Peanut (*Arachis durensis*)	15	14	0	29
8	Mung bean (*Vignga radiata*)	14	12	0	26
9	Rice (*Oryza sativa*)	0	16	1	17
10	Sorghum (*Sorghum bicolor*)	9	14	1	24
11	Barley (*Hordeum vulgare*)	15	15	11	41
12	Maize (*Zea mays*)	21	20	6	47
13	Arabidopsis (*Arabidopsis thaliana*)	8	14	2	24
14	Grapes (*Vitis vinifera)*	20	13	6	39
Total		225	176	53	454

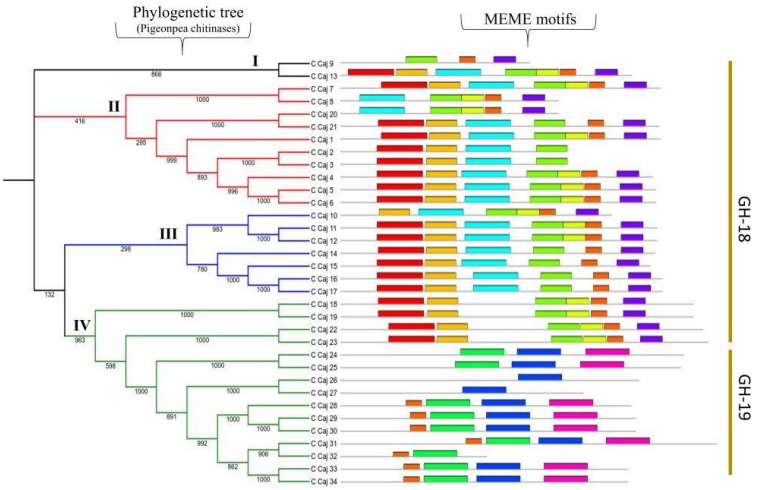

Fig 3. Phylogenetic tree of 34 chitinase genes of pigeonpea showing 4 groups which are color-coded and represented by roman number (Black=I, Red=II, Blue=II, Green= IV). The Phylogenetic tree was build using neighbor-joining method. The MEME analysis identified motifs are shown with cluster specific domain architecture and Glycoside hydrolase 18 and Glycoside hydrolase 19 family of pigeonpea chitinase genes are also marked as GH-18 and GH-19 respectively.

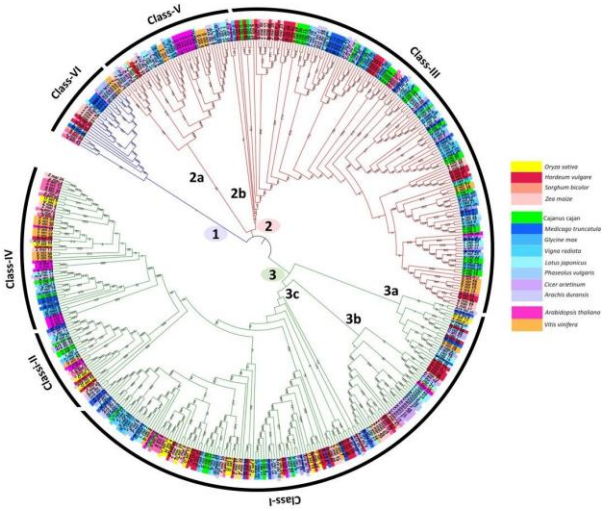

Fig 4. Phylogenetic tree of 34 chitinase proteins from pigeonpea with 13 plant species chitinases includes 7 legumes; chickpea (*Cicer arientum*), common bean *(Phaseolus vulgaris*), soybean (*Glycine max*), barrel clover (*Medicago truncatula*), lotus (*Lotus japonicus*), peanut (*Arachis durensis*), Mung bean (*Vignga radiata*), 4 cereals; rice (*Oryza sativa*), sorghum (*Sorghum bicolor*), barley (*Hordeum vulgare*), maize (*Zea mays*) and two dicot species; *Arabidopsis thaliana* and *Vitis vinifera.* The un-rooted phylogenetic tree was build using the neighbor-joining (NJ) method via ClustalW. The three major clades shaded by 1) Blue: PR-4 Like (Barwin); Red: Glycoside hydrolase 18 (GH-18); and 3) Green: Glycoside hydrolase 19 (GH-19).

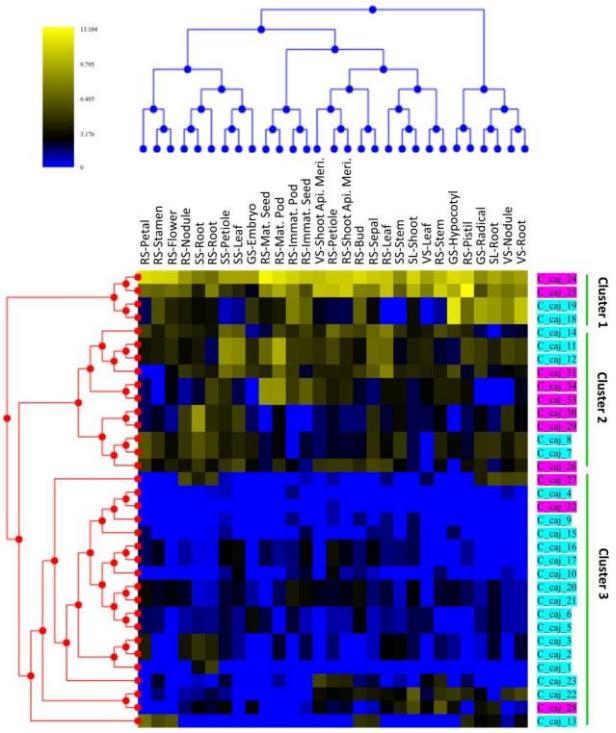

Fig 5. Expression patterns of identified pigeonpea chitinase genes in different tissues at five developmental stages (RS= Reproductive Stage, VS= Vegetative Stage, SS= Senensis Stage, SL= Seedling Stage). The pigeonpea chitinase gene of family GH-18 and GH-19 are highlighted by sky blue and purple color, respectively. The three major clusters (Clusters 1-3) based on expression levels are marked with green line.

Medicago truncatula, *Lotus japonicus*, *Arachis duranensis*, and *Vigna radiata*; four cereals: *Oryza sativa*, *Sorghum bicolor*, *Hordeum vulgare*, and *Zea mays*; and two other dicot plant species: *Arabidopsis thaliana* and *Vitis vinifera*. The tree, which was generated by the neighbor joining method, has three major clades (**Fig 4**). Clade 1 showed no further divisions, but clades 2 and 3 were further divided in to five subclades (2a, 2b, 3a, 3b, and 3c). These five subclades correspond to five of the different classes of chitinases described in Arabidopsis (Passarinho and de Vries, 2002**)**. The clade 2 genes contain GH-18 domains and its two subclades, 2a and 2b, correspond to classes III and V of the Arabidopsis chitinases, respectively. Clade 3 genes contain the GH-19 domain, and its three subclades, 3a, 3b, and 3c, correspond to classes I, II, and IV of the Arabidopsis chitinases, respectively (**Fig 4**). These groupings of chitinases into classes I–V is consistent with those reported in Arabidopsis, rice, mulberry, and poplar (Passarinho and de Vries, 2002; Xu et al., 2007; Wang et al., 2015; Jiang et al., 2013). Clades 2 and 3 include genes from all 14 plant species, including pigeonpea, but clade 1 comprises only eight plant species, excluding pigeonpea and Arabidopsis. Hence, we designated this class VI, to add to the five existing recognized classes of chitinases in Arabidopsis. Further analysis of chitinase genes in clade 1 (Class VI) showed that the genes in this clade are significantly shorter in length and are predicted to encode proteins of less than 150 amino acids. The gene annotation indicated that they belong to the pathogenesis-related (PR-4) proteins and contain a Barwin/Hevein domain, which plays a crucial role in defense against fungal pathogens and has strong antifungal activities (Ludvigsen and Poulsen, 1992; Wang et al., 2011). To assess the sequence conservation in each subclade, we selected species-specific genes from each subclade and aligned them, and the results showed that the protein sequences of GH-18, GH-19, and PR-4 in each clade and subclade across the 14 plant species are highly conserved (**Fig S2**), indicating that they are derived from different ancestral genes (Tyler et al., 2010).

Expression patterns of pigeonpea chitinase genes in different tissues

We analyzed the tissue-specific expression patterns of each of the 34 pigeonpea chitinase genes by using publicly available RNA-seq data from pigeonpea variety 'Asha' (Pazhamala et al., 2017). Two-dimensional hierarchical clustering of the RNA-seq reads from different plant tissues across developmental stages mapped onto the 34 chitinase genes showed distinct expression patterns. There were three clear clusters of genes based on chitinase gene expression in different tissues at different developmental stages (**Fig 5**). Cluster 1 consisted of the four highest expressing genes with a further subgrouping of two genes each. The first subgroup consisted of the highest expressing chitinase genes, C_caj_24 and C_caj_25, which were expressed in almost all tested tissues at all five developmental stages, whereas the second subgroup, C_caj_18 and C_caj_19, was highly expressed in the root tissues, hypocotyl, and pistil. The genes in cluster 1, which contains pigeonpea chitinase gene C_caj_18, C_caj_19, C_caj_24, and C_caj_25, showed high gene expression, and C_caj_24 showed the highest expression across different

tissues at all developmental stages and mapped to chromosome 7 of the pigeonpea genome. Chitinase genes C_caj_18 and C_caj_19 from cluster 1 showed equal expression, and their expression was highest at germination stage in the hypocotyl tissue. Cluster 2, with 11
chitinase genes, showed medium to high expression levels in several tissues, whereas cluster 3, with 19 genes, was the lowest expressing genes. In cluster 2, C_caj_29 and C_caj_30, were most highly expressed in the root tissue at the seedling stage, followed by root tissue at the reproductive stage and reproductive stage, while C_caj_1 appeared to be a root-specific chitinase and was only expressed in these tissues. Interestingly, in cluster 3, only one gene, C_caj_13, showed significant expression only in tissues at the reproductive stage (bud, flowed, petal, and stamen), clearly demonstrating that these are flower-specific chitinases. The majority of the genes in clusters 1 and 2 contain GH-19 domains, while all the genes in cluster 3 except one, C_caj_13, have a GH-18 domain (**Fig 5**). The pigeonpea chitinases that were expressed in different organs throughout development may have specific hydrolytic activities that induce signal molecules or morphogenic factors (e.g., nod factors that lead to nodule formation without evoking plant defense reaction) (Grover, 2012; Haeze and Holsters, 2002).
There is a block of four genes in cluster 1, C_caj_24, C_caj_25, C_caj_18, and C_caj_19, that are expressed at higher levels than the other genes. Interestingly, their expression was limited to 4 developmental stages (germination stage, hypocotyl; reproductive stage, pistil; seedling stage, roots; and vegetative stage, roots and root nodules), indicating their importance in these stages of plant development. In the hypocotyl, the expression of two genes, C_caj_18 and C_caj_19, was much higher than any other genes, suggesting that these genes play a crucial role during germination and are involved in a post-germination defense strategy that protects germinating seeds once its physical protective barriers are removed (Fincher, 1989; Flach et al., 1992). At senescence, the chitinase genes C_caj_29 and C_caj_30 have the highest expression in root tissue, indicating their stage-specific involvement in ethylene regulation as was reported in soybean, Brassica, and Arabidopsis (Xie et al., 1996; Hanfrey et al., 1996; Chen and Bleecker, 1995). Earlier studies showed that ethylene regulation by chitinase not only plays a crucial role in defense and senescence in leaves and flowers but also in seedling growth (Chen and Bleecker, 1995; Larsen and Chang, 2001). The expression of the pigeonpea chitinase gene C_caj_25 was higher in the pistil than all other chitinase genes in all other tissues. Similarly, C_caj-24 showed the highest expression in reproductive stage tissues (petal, stamen, and flower), while its expression is low in vegetative tissues and other developmental stages. Similar types of flower-specific chitinases have been reported in potato, tomato, rice, and Arabidopsis (Wemmer et al., 1994; Harikrishna et al., 1996; Takakura et al., 2000; Passarinho and de Vries, 2002). These constitutively expressed tissue-specific chitinases may be involved in a range of morphological and physiological processes, including seed germination, embryogenesis, flowering, and senescence, and could be used as a tool for crop improvement through marker-assisted breeding or genetic engineering (Van Loon et al., 2006; Cletus et al., 2013; Kasprezewska, 2003; Bekesiova et al., 2008).

Materials and Methods

Genome-wide identification of chitinase-encoding genes

The improved draft genome version 2.0 of pigeonpea (GenBank accession number AFSP02000000) (Mahato et al., 2018) was downloaded from NCBI. Gene prediction was performed using repeat masked genome via FGENESH from the Molquest package, version 4.5. (http://www.softberry.com). Full length genes were extracted, and their protein sequences were subjected to analysis by HMMER (Eddy, 1991) searching for the HMMs Glyco_hydro_18 (PF00704) and Glyco_hydro_19 (PF00182) to identify the chitinase genes in the pigeonpea genome.
The HMMER-searched chitinase protein sequences were re-annotated by a BLASTp (Altschul, 1990) search against the NCBI-nr database, and domain reconfirmation was carried out using SMART (Letunic et al., 2014), InterProScan (Quevillon et al., 2005), and NCBI-CDD search (Marchler-Bauer et al., 2004). Putative chitinase protein sequences were manually curated based on the annotations generated from the sequence and protein domain searches. In these filtered pigeonpea chitinase proteins, the signal peptide was predicted using SignalP (Petersen et al., 2011), and their subcellular localization was predicted using ProtComp (http://linux1.softberry.com/berry.phtml).

Chromosomal mapping and analysis of conserved motifs

The candidate pigeonpea chitinase genes were mapped to the pigeonpea genome pseudomolecules (Varshney et al., 2012) submitted to GenBank (accession number AGCT00000000.1) using a BLASTn search. Using an in-house PERL script, the mapping result was tabulated and formatted, which was used to generate a physical map of the chitinase genes on the pigeonpea genome with Map-chart (Voorrips, 2002). The exon/intron structures were analyzed by aligning the genomic DNA sequences with their corresponding coding sequences using the Gene Structure Display Server (GSDS) program (http://gsds1.cbi.pku.edu.cn/). Conserved motifs in the candidate protein sequence were identified by using locally configured Multiple Expectation Maximization for Motif Elicitation (MEME) V.4.12.0 (Bailey et al., 2009) with the following options: 1) mode = anr (any number of repetitions), 2) number of motifs = 10, 3) minimum motif width = 6, and 4) maximum motif width = 50.

Sequence alignment and phylogenetic tree construction

To generate a phylogenetic tree of the chitinase genes in the pigeonpea genome, the candidate chitinase protein sequences were aligned using MUSCLE (Edger, 2004), and a phylogenetic tree was generated by the neighbor-joining (NJ) method using ClustalX (Higgins and Sharp 1998) with 1000 bootstrap replicates. To confer the phylogenetic relationships among the pigeonpea chitinase proteins and other plant species, we downloaded the GFF3 file and predicted chitinase sequences from 13 plant species, including seven legumes: chickpea (C*icer arientum*), common bean (*Phaseolus vulgaris*), soybean (*Glycine max*), barrel clover (*Medicago truncatula*), lotus (*Lotus japonicus*),

peanut (*Arachis duranensis*), and Mung bean (*Vigna radiata*) downloaded from the Legume information system (https://legumeinfo.org/); four cereals: rice (*Oryza sativa*), sorghum (*Sorghum bicolor*), barley (*Hordeum vulgare*), and maize (*Zea mays*) downloaded from phytozome (https://phytozome.jgi.doe.gov/); and two other dicot species: *Arabidopsis thaliana* from tair (https://www.arabidopsis.org/) and *Vitis vinifera* from phytozome (https://phytozome.jgi.doe.gov/). The species-specific chitinase genes were extracted from the GFF3 file of the 13 plant species using the keyword "chitinase," and their corresponding protein sequences were extracted from the downloaded protein sequence file using an in-house SHELL script. These extracted chitinase proteins were re-validated using BLASTp and InterProScan searches, and species-specific chitinase genes were finally filtered based on the annotation. The filtered chitinase proteins, along with the pigeonpea chitinase proteins, were aligned with MUSCLE using default parameters, and a phylogenetic tree was constructed using the NJ method with 1000 bootstrap replications using ClustalX, the final tree was viewed and edited in itol (https://itol.embl.de/).

Expression analysis of pigeonpea chitinase genes

To estimate the tissue specific expression of the chitinases, the reads per kilobase of transcript model pre-million mapped reads (RPKM) method was used. We downloaded 331 GB of pigeonpea RNA-seq data (Pazhamala et al., 2017) from the NCBI-SRA database submitted under BioProject accession number PRJNA354681. This RNA-seq data is from different tissues (embryo, radical, hypocotyl, root, root nodule, stem, leaf, shoot apical meristem, bud, flower, sepal, petal, petiole, stamen, pistil, immature seed, and mature seed) collected at five developmental stages (germination, seedling, vegetative, reproductive, and senescence). The downloaded RNA-seq data were filtered using Trimmomatic (Bolger et al., 2014). High quality RNA sequencing reads were mapped, and a read count table was created using Bowtie 2 (Langmed and Salzberg, 2012) and RSEM (Le and Dewey, 2011). The read count data were normalized, and the hierarchical clustering method was used to generate a heat map, illustrating the gene expression profiles of the 34 pigeonpea chitinase genes in various tissues during the plant life cycle.

Conclusions

In the present study, the identification, classification, and expression analysis of pigeonpea chitinase genes at the whole genome level were conducted using an in-silico approach. In total of 34 chitinase genes were identified, and their gene structure, functional domain identification, chromosomal mapping, and phylogenetic relationship with chitinase genes in other plant species were investigated. These results provide the first genome-wide analysis of pigeonpea chitinase genes and shed light on their expression in diverse tissues at different developmental stages. This information will be very useful for advancing our knowledge and utilization of pigeonpea chitinase genes for variety improvement.

Contributions

Mahato A. K. and Singh N. K. conceived and designed research. Mahato A. K. carried out all bioinformatics analysis and wrote the manuscript. Singh N. K. and Sharma A. K. revised and finalized the manuscript. All authors read and approved the manuscript.

Acknowledgments

This work was financially supported by Indian Council of Agricultural Research- Network Project on Transgenics in Crops (ICAR-NPTC) (Grant No. 3052), India.

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol. 215:403-410.

Alvarez FJ, Konopka JB (2007) Identification of an N-acetylglucosamine transporter that mediates hyphal induction in Candida albicans. Mol Biol Cell. 18(3):965–975.

Backiyarani S, Uma S, Nithya S, Chandrasekar A, Saraswathi MS, Thangavelu R, et al. (2015) Genome-wide analysis and differential expression of chitinases in banana against root lesion nematode (Pratylenchus coffeae) and Eumusa leaf spot (Mycosphaerella eumusae) pathogens. Appl Biochem Biotechnol. 175:3585–3598.

Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, et al. (2009). MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res. 37: 202– 208.

Bekesiova B, Hraska S, Libantova J, Moravcikova J, Matusikova I (2008) Heavy-metal stress induced accumulation of chitinase isoforms in plants. Mol Biol Rep. 35:579–588.

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 30:2114–2120.

Brameld KA, Goddard III WA (1998) The role of enzyme distortion in the single displacement mechanism of family 19 chitinases. Proc Natl Acad Sci USA. 95(8):4276–4281.

Brameld KA, Goddard III WA (1998). Substrate distortion to a boat conformation at subsite ÿ1 is critical in the mechanism of family 18 chitinases. J Am Chem Soc. 120:3571-3580.

Chen J, Piao Y, Liu Y, Li X, Piao Z (2018) Genome-wide identification and expression analysis of chitinase gene family in Brassica rapa reveals its role in clubroot resistance. Plant Sci. 270:257-267.

Chen QG, Bleecker AB (1995) Analysis of ethylene signal-transduction kinetics associated with seedling-growth response and chitinase induction in wild-type and mutant arabidopsis. Plant Physiol. 108(2):597-607.

Cletus J, Balasubramanian V, Vashist D, Sakthivel N (2013) Transgenic expression of plant chitinases to enhance disease resistance. Biotechnol lett. 35:1719–32.

Collinge DB, Kragh KM, Mikkelsen JD, Nielsen KK, Rasmussen U, Vad K (1993) Plant chitinases. Plant J. 3:31-40.

Eddy SR (2011) Accelerated profile HMM searches. PLoS Comput. Biol. 7(10):e1002195.

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput, Nucleic Acids Res. 32: 1792-1797.

Felton GW, Korth KL (2000) Trade-offs between pathogen and herbivore resistance. Curr Opin Plant Biol. 3(4):309-314.

Fincher GB (1989) Molecular and cellular biology associated with endosperm mobilization in germinating cereal grains. Annu Rev Plant Physiol Plant Mol Biol. 40:305–346.

Flach J, Pilet PE, Jolles P (1992) What,s new in chitinase research?. Experientia. 48:701–716

Garcia-Casado G, Collada C, Allona I, Casado R, Pacios LF, Aragoncillo C, et al. (1998) Site-directed mutagenesis of active site residues in a class I endochitinase from chestnut seeds. Glycobiology. 81(1):1021–1028.

Graham LS, Sticklen MB (1994) Plant chitinases. Canadian J Bot. 1994;721(1):1057–1083.

Greilhuber J, Obermayer R (1998) Genome size variation in Cajanus cajan (Fabaceae): a reconsideration. Plant Syst Evol. 212:135–141.

Grover A (2012) Plant Chitinases: Genetic Diversity and Physiological Roles. Critical Reviews in Plant Sciences. 31(1):57-73.

Haeze WD, Holsters M (2002) Nod factor structures, responses, and perception during initiation of nodule development. Glycobiology. 12: 79–105.

Hanfrey C, Fife M, Buchanan-Wollaston V (1996) Leaf senescence in Brassica napus: expression of genes encoding pathogenesis-related proteins. Plant Mol Biol. 30:597 609.

Harikrishna K, Jampates-Beale R, Milligan SB, Gasser CS (1996) An endochitinase gene expressed at high levels in the stylar transmitting tissue of tomatoes. Plant Mol Biol. 30(5):899-911.

Hart PJ, Pfluger HD, Monzingo AF, Hollis T, Robertus JD (1995) The refined crystal structure of an endochitinase from Hordeum vulgare L. seeds at 1.8 A resolution. J Mol Biol. 2481(1):402–413.

Henrissat B (1991) A classification of glycosyl hydrolases based on amino acid sequence similarities. Biochem J. 280:309-316.

Henrissat B, Bairoch A (1993) New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. Biochem J. 293 (3):781- 788.

Higgins DG, Sharp PM (1988). CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. Gene. 73, 237-244.

Jeffares DC, Penkett CJ, Bähler, J (2008) Rapidly regulated genes are intron poor. Trends Genet. 24(8):375–378.

Jiang C, Huang RF, Song JL, Huang MR, Xu LA (2013) Genomewide analysis of the chitinase gene family in Populus trichocarpa. Journal of Genetics. 92: 121–125.

Kasprezewska A (2003) Plant Chitinases-Regulation and Function. Cell Mol Biol. 8:809–824.

Langmead B, Salzberg S (2012) Fast gapped-read alignment with Bowtie 2. Nature Methods. 9:357-359.

Larsen PB, Chang C (2001) The Arabidopsis eer1 mutant has enhanced ethylene responses in the hypocotyl and stem. Plant Physiol. 1251(1):1061–1073.

Legrand M, Kaufmann S, Geoffroy P, Fritig B (1987) Biological functions of pathogenesis-related proteins: four tobacco pathogenesis-related proteins are chitinases. Proc Natl Acad Sci. 84(19): 6750–6754.

Letunic I, Bork P (2017) 20 years of the SMART protein domain annotation resource; Nucleic Acids Res. 46(D1):493–D496.

Li B, Dewey CN (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics. 12:323.

Liu JJ, Ekramoddoullah AK, Zamani A (2005) A class IV chitinase is up- regulated by fungal infection and abiotic stresses and associated with slow- canker-growth resistance to Cronartium ribicola in western white pine (Pinus monticola). Phytopatholog. 95(3): 284-291.

Ludvigsen S, Poulsen FM (1992) Secondary structure in solution of barwin from barley seed using 1H nuclear magnetic resonance spectroscopy. Biochemistry. 31:8771–8782.

Mahato AK, Sharma AK, Singh NK (2018) An improved draft of the pigeonpea (Cajanus cajan (L.) Millsp.) genome. Data Brief. 16: 376–380.

Marchler-Bauer A, Bryant SH (2004) CD-Search: protein domain annotations on the fly. Nucleic Acids Res. 32:327–331.

Maximova SN, Marelli JP, Young A, Pishak S, Verica JA, Guiltinan MJ (2006) Over-expression of a cacao class I chitinase gene in Theobroma cacao L. enhances resistance against the pathogen, Colletotrichum gloeosporioides. Planta. 224(4):740–749.

Misra B (2015) Molecular Evolution and Functional Divergence of Chitinase Gene Family in Hevea brasiliensis. The Winnower. DOI: 10.15200/winn.144125.54243.

Naseem S, Gunasekera A, Araya E, Konopka JB (2011) N-acetylglucosamine (GlcNAc) induction of hyphal morphogenesis and transcriptional responses in Candida albicans are not dependent on its metabolism. J Biol Chem. 286(33):28671–28680.

Neuhaus, JM, Fritig, B, Linthorst, HJM, Meins, FJ (1996) A revised nomenclature for chitinase genes. Plant Mol Biol Rep. 14(2):102-104.

Passarinho PA, de Vries SC (2002) Arabidopsis chitinases: a genomic survey. The Arabidopsis book/American Society of Plant Biologists. 1:e0023.doi: 10.1199/tab.0023.

Patil RS, Ghormade V, Deshpande MV (2000) Chitinolytic enzymes: an exploration. Enzyme Microb Technol. 26:473-483.

Pazhamala LT, Purohit S, Saxena RK, Garg V, Krishnamurthy L, Verdier J, et al. (2017) Gene expression atlas of pigeonpea and its application to gain insights on genes associated with pollen fertility implicated in seed formation. J Exp Bot. 68(8):2037–2054.

Petersen N, Brunak S, Heijne GV, Nielsen H (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Meth. 8:785-786.

Quevillon E, Silventoinen V, Pillai S, Harte N, Mulder N, et al. (2005) InterProScan: protein domains identifier. Nuc Acids Res. 33:116–120.

Reddy MV, Raju TN, Nene YL, Ghanekar AM, Amin KS, Arjunan G, et al. (1993) Variability in sterility mosaic pathogen in pigeonpea in India. Indian Phytopathol. 46:206– 212.

Schauser L, Wieloch W, Stougaard J (2005) Evolution of NIN-like proteins in Arabidopsis, rice, and Lotus japonicus. J Mol Evol. 60:229–237.

Shibuya N, Minami E (2001) Oligosaccharide signalling for defence responses in plant. Physiol Mol Plant Pathol. 59: 223-233.

Singh N K, Gupta DK, Jayaswal PK, Mahato AK, Dutta S, Singh S, et al. (2012). The first draft of the pigeonpea genome sequence. J Plant Biochem Biotechnol. 21: 98–112.

Stacey G, Shibuya N (1997) Chitin recognition in rice and legumes. In: Opportunities for Biological Nitrogen Fixation in Rice and Other Non-Legumes. Springer. pp 161-169.

Takakura Y, Ito T, Saito H, Inoue, T, Komari T, Kuwata S (2000) Flower- predominant expression of a gene encoding a novel class I chitinase in rice (Oryza sativa L.). Plant Mol Biol. 42:883-897.

Tobias PA, Christie N, Naidoo S, Guest DI, Külheim C (2017) Identification of the Eucalyptus grandis chitinase gene family and expression characterization under different biotic stress challenges. Tree Physiol. 37(5):565-582.

Tyler L, Bragg JN, Wu J, Yang X, Tuskan GA, Vogel JP (2010) Annotation and comparative analysis of the glycoside hydrolase genes in Brachypodium distachyon. BMC Genomics. 11: 600.

Van Loon L, Van Strien E (1999) The families of pathogenesis-related proteins, their activities, and comparative analysis of PR-1 type proteins. Physiol Mol Plant. 55(2):85-97.

Van Loon LC, Rep M, Pieterse CM (2006) Significance of inducible defense- related proteins in infected plants. Annu Rev Phytopathol. 44:135–62.

Varshney RK, Chen W, Li Y, Bharti AK, Saxena RK, Schluete JA, et al. (2012) Draft genome sequence of pigeonpea (Cajanus cajan), an orphan legume crop of resource-poor farmers. Nat. Biotechnol. 30:83–89.

Voorrips RE (2002) MapChart: software for the graphical presentation of linkage maps and QTLs. J Hered. 93:77–78.

Wan J, Zhang X-C, Neece D, Ramonell KM, Clough S, Kim S-y, et al. (2008) A LysM receptor-like kinase plays a critical role in chitin signaling and fungal resistance in Arabidopsis. Plant Cell. 20(2):471-481.

Wang J, Bi Y, Zhang Z, Zhang H, Ge Y (2011) Reduction of latent infection and enhancement of disease resistance in muskmelon by preharvest application of harpin. J AgricFood Chem. 59:12527–12533.

Wang X, He N, Zeng Q, Xiang Z (2015) Identification and expression analyses of chitinase genes in mulberry (Morus L.) plants. Plant Omics J. 8(2):183- 190.

Wemmer T, Kaufmann H, Kirch H-H, Schneider K, Lottspeich F, Thompson RD (1994) The most abundant soluble basic protein of the stylar transmitting tract in potato (Solanum tuberosum L.) is an endochitinase. Planta. 194:264-27.

Xiao YH, Li XB, Yang XY, Luo M, Hou L, Guo SH, et al. (2007) Cloning and characterization of a balsam pear class I chitinase gene (Mcchit1) and its ectopic expression enhances fungal resistance in transgenic plants. Biosci Biotechnol Biochem. 71(5)1211–1219.

Xie ZP, Staehelin C, Wiemken A, Boller T, (1996) Ethylene responsiveness of soybean cultivars characterized by leaf senescence, chitinase induction and nodulation. J Plant Physiol. 149:690–694.

Xu F, Fan C, He Y (2007) Chitinases in Oryza sativa ssp. japonica and Arabidopsis thaliana. J Genet Genomics. 34(2):138-150.

Xu J, Xu X, Tian L, Wang G, Zhang X, Wang X, et al. (2016) Discovery and identification of candidate genes from the chitinase gene family for Verticillium dahliae resistance in cotton. Nat Sci Reports. 6:29022.