Characterization of the CHTBD1 gene family in Spinacia oleracea reveals candidate genes for downy mildew resistance

xiaoyu Mo
Yilin Zong
Huan Liu
Yuan Zhao
Yahan Zhu
Xiaoli Wang
Chenxi Xu
Xiaofeng Cai
xfcai@shnu.edu.cn

https://orcid.org/0000-0001-8058-0399

Research Article

Keywords: Spinacia oleracea, CHTBD1 gene family, downy mildew resistance, gene expression profiling

Posted Date: November 10th, 2025

DOI: https://doi.org/10.21203/rs.3.rs-7907263/v1

Version of Record: A version of this preprint was published at Journal of Plant Diseases and Protection on March 11th, 2026. See the published version at https://doi.org/10.1007/s41348-026-01251-0.

Characterization of the CHTBD1 gene family in *Spinacia oleracea* reveals candidate genes for downy mildew resistance

Mo Xiaoyu, Zong Yilin, Liu Huan, Zhao Yuan, Zhu Yahan, Wang Xiaoli, Xu Chenxi and Cai Xiaofeng*

Shanghai Collaborative Innovation Center of Plant Germplasm Resources, College of Life Sciences, Shanghai Normal University, Shanghai 200234, China;

* Correspondence: xfcai@shnu.edu.cn

Abstract

Spinach (*Spinacia oleracea*) is a globally cultivated leafy vegetable, and its yield is severely threatened by downy mildew (*Peronospora effusa*), yet the molecular basis of resistance remains poorly understood. Chitin-binding type 1 (CHTBD1) proteins play critical roles in pathogen recognition and immune activation across plant species. However, their genomic organization and potential involvement in disease resistance have not been explored in spinach. In this study, we identified 27 CHTBD1 homologs (SoCHIBI) in the spinach genome and performed a comprehensive analysis of their physicochemical properties, predicted subcellular localization, chromosomal distribution, gene structure, conserved motifs, phylogenetic relationships, and macrosynteny with four representative angiosperms. Promoter analysis revealed an enrichment of stress-responsive and pathogen-associated cis-acting elements. To evaluate transcriptional responses, we profiled the expression of 14 *SoCHIBI* genes using quantitative real-time PCR across six tissues in three spinach cultivars under downy mildew infection. Eight genes, including *SOV2g016080.1*, *SOV2g016120.1*, and *SOV3g012940.1*, were significantly upregulated following infection, with stronger induction observed in the resistant cultivar U9 compared to the susceptible S2. This study presents the first systematic characterization of the CHTBD1 gene family in spinach, identifies defense-associated candidate genes, and provides valuable genetic resources for future functional validation and resistance breeding.

Keywords:*Spinacia oleracea*; CHTBD1 gene family; downy mildew resistance; gene expression profiling

1 Introduction

Spinach (*Spinacia oleracea* L.), a globally cultivated leafy vegetable from the Amaranthaceae family, is highly valued for its rapid growth, high yield potential, and adaptability to diverse environmental conditions. As a major green leafy vegetable, spinach plays an essential role in both fresh consumption and food processing industries. Its large-scale production is particularly critical in temperate regions, where it supports key segments of vegetable supply chains (Gao et al., 2022).

However, spinach cultivation faces significant challenges due to downy mildew, a devastating foliar disease caused by the obligate oomycete *Peronospora effusa*. This pathogen induces leaf chlorosis, sporulation, and necrosis, resulting in substantial yield losses and reduced marketability. Of greater concern is the rapid evolutionary rate of *P. effusa*, which continuously generates novel physiological races capable of overcoming resistance genes deployed in commercial cultivars (Bhattarai et al., 2021). Numerous resistant spinach varieties have become susceptible within a few years of their commercial release, demonstrating the inherent limitations of conventional breeding approaches. The global expansion and escalating frequency of downy mildew outbreaks have emphasized the critical necessity for developing durable resistance mechanisms (Cornell Vegetable Program, 2023).

To elucidate the molecular mechanisms underlying spinach–pathogen interactions, recent transcriptomic investigations have characterized the dynamics of gene expression during pathogenic infection. Clark et al. (2024) conducted comprehensive dual transcriptome profiling of spinach and P. effusa, identifying differentially expressed receptor-like protein kinases, leucine-rich repeat-containing proteins, and ATP-binding cassette transporters implicated in host defense responses during resistant interactions. Subsequently, Leveraging six telomere-to-telomere genome assemblies, Skiadas et al. (2024) generated a chromosome-level pangenome graph of P. effusa and identified pronounced copy-number variation within effector-rich clusters, supporting a genomic basis for rapid adaptation and resistance overcoming in spinach downy mildew.

Chitin-binding type 1 (CHTBD1) proteins belong to the CBM18 family and bind chitin and its oligomers; they are conserved across plant and fungal lineages and function in plant defense by promoting immune responses such as ROS production, hypersensitive cell death, and PR-gene induction (Ali et al., 2020).Within plant innate immunity, chitin oligosaccharides from fungal cell walls are perceived by LysM-containing PRR complexes (e.g., the CEBiP–CERK1 module), which initiate

immune signalling cascades leading to ROS production and transcriptional activation of defense genes (Cai et al., 2024). To circumvent chitin perception, pathogenic fungi secrete chitin-binding effectors—most notably LysM/CBM50 and CBM18 proteins—that sequester chitin oligomers and/or shield cell-wall chitin, thereby blocking receptor engagement and suppressing chitin-triggered immunity (Volk et al., 2019).

Genome-wide analyses in pepper have identified chitin-binding proteins harboring ChtBD1 domains, classified into four classes with distinct domain organizations and tissue-specific expression, and many members are induced by P. capsici, abiotic stresses, and hormones (Ali et al., 2018). Correspondingly, genome-wide analyses in tomato identified multiple chitinase genes with a chitin-binding domain (ChtBD1) that are transcriptionally upregulated upon infection by *Sclerotinia sclerotiorum* (Cao et al., 2019). In soybean, overexpression of a fungal chitinase gene (*Trichoderma asperellum Tachi*) enhances resistance to *Sclerotinia sclerotiorum*, with transgenic plants showing increased ROS accumulation and elevated peroxidase and superoxide dismutase activities after infection (Zhang et al., 2016).

Despite accumulating evidence supporting the critical role of CHTBD1 proteins in plant immunity, the genomic organization and functional characteristics of the CHTBD1 gene family in *Spinacia oleracea* remain entirely unexplored. To address this significant knowledge deficit, we conducted a comprehensive genome-wide identification and systematic characterization of *CHTBD1* genes within the spinach genome. Our multifaceted investigation encompasses detailed structural analyses, phylogenetic reconstructions, promoter element characterization, and comparative cross-species collinearity assessments. Additionally, we performed quantitative expression profiling of selected candidate genes across multiple tissue types and cultivars under controlled *P. effusa* infection conditions using quantitative real-time PCR methodology.

This work provides a comprehensive view of the CHTBD1 gene family in spinach, offering a basis for future functional analyses and the development of resistant cultivars through molecular breeding strategies.

2 Materials and Methods

2.1 Genome-wide identification and phylogenetic analysis of the SoCHTBD1 gene family

Protein sequences of the CHTBD1 family from *Arabidopsis thaliana* were retrieved from the TAIR database(https://www.arabidopsis.org) and used as queries in BLASTP searches against the spinach reference protein database (SpinachBase, https://www.spinachbase.org). Candidate sequences were filtered based on the presence of the CHTBD1-associated domain PF01607, with domain annotation performed using the PFAM-A database (v35.0) under the WSL system. Conserved motif elements were identified using the MEME Suite v5.4.1 platform (https://meme-suite.org), and the resulting prediction files were processed by custom Python scripts to extract the physical positions of motifs and domains within each protein sequence.

Physicochemical properties, including amino acid length, molecular weight, and theoretical isoelectric point, were analyzed using the ExPASy ProtParam tool (https://web.expasy.org/protparam). Multiple sequence alignment was conducted using MUSCLE v3.8.31, followed by maximum likelihood phylogenetic reconstruction with IQ-TREE v2.2.0. Visualization of phylogenetic relationships, conserved motifs, domain composition, and gene structure was performed in R using the ggtree (Yu et al. 2017), treeio (Wang et al. 2020), gggenes (Wilke 2022), and tidyverse (Wickham et al. 2019) packages.

2.2 Cis-regulatory element analysis

Promoter regions (2,000 bp upstream of transcription start sites) of *SoCHTBD1* genes were extracted from SpinachBase and submitted to the PlantCARE database (https://bioinformatics.psb.ugent.be/webtools/plantcare/html/) for cis-acting regulatory element prediction. The predicted elements were classified into three functional categories following established frameworks (Mengarelli and Zanor 2021). Data processing and visualization were performed in R (v4.3.0) using dplyr and ggplot2 (Wickham et al., 2019) and pheatmap (Kolde, 2019), with heatmaps and stacked bar plots constructed to illustrate cis-element distribution patterns.

2.3 Chromosomal localization and synteny analysis

Chromosomal organization and syntenic relationships were analyzed by quantifying collinear gene pair frequencies and constructing a chromosomal synteny strength matrix visualized through transparent inter-chromosomal links. Genome-wide gene density was computed using sliding window analysis and displayed as circular heatmaps, with SoCHTBD1 family members positioned on the

outermost ring. Data processing was performed in a WSL environment, and visualization was carried out using the circlize (Gu et al. 2014) and ComplexHeatmap (Gu et al. 2016) packages in R.

2.4 Plant materials and pathogen inoculation

Three spinach cultivars provided by the Development Center of Plant Germplasm Resources, College of Life Sciences, Shanghai Normal University served as differential hosts for downy mildew pathotype identification. In spring 2025, pathogen samples were collected from naturally infected plants in Shanghai greenhouses. All experiments were conducted under standardized greenhouse conditions at the Plant Germplasm Center of Shanghai Normal University. Plants displaying white or gray mildew sporulation or pale yellow lesions were classified as susceptible, while asymptomatic plants were considered resistant. Six tissue types (root, stem, functional leaf, functional petiole, bolting leaf, and bolting petiole) were collected from diseased and healthy plants, immediately sealed in plastic bags, and stored at −80°C for subsequent analyses.

2.5 Quantitative RT-PCR analysis

To investigate *SoCHTBD1* gene expression patterns under downy mildew infection, qRT-PCR (quantitative reverse transcription PCR) was performed on the six tissue types described above. Total RNA was extracted using TRIzol reagent (Thermo Fisher Scientific), quantified using a NanoDrop 2000 spectrophotometer, and reverse-transcribed using DBI HiScript III All-in-one RT SuperMix following the manufacturer's protocol. qPCR reactions were performed using DBI Hieff® SYBR® Green Master Mix on an Applied Biosystems 7500 Real-Time PCR System. Spinach *So18S* rRNA served as the internal reference gene, and relative expression levels were calculated using the 2^−ΔΔCt method (Livak and Schmittgen 2001). Primer sequences are provided in Table 1. Expression heatmaps were generated using the ComplexHeatmap package (Gu et al. 2016) in R, with positive and negative indicators denoting expression trends.

Table 1. Sequences of the primers used in this study

Gene name	Gene ID	Forward primer (5′–3′)	Reverse primer (5′–3′)
SoCHTBD1-1	SOV1g025700.1	CCCGCAGGACTTTGTTGTAGCC	AGCTTCCACCTGCTCCTCAGC
SoCHTBD1-2	SOV1g025710.1	CCCGCAGGACTTTGTTGTAGCC	TCAAGGAGCACCAACAGCAACC
SoCHTBD1-3	SOV1g025730.1	TGGAACATCAATGGCACAACAAGGT	GCTCCTCAGCACCACCACAATATG
SoCHTBD1-4	SOV1g025790.1	AAGCCCTTGCCCGTGACATTG	AGGTTTAACAGCAGCAGCGGAAG
SoCHTBD1-5	SOV2g000500.1	AACGACGCTGTTGTTGCCTTCA	TCCTCCACCACATTCCATTCCATTG
SoCHTBD1-6	SOV2g000550.1	GCGGCACCACATCTGACTACTG	CCATCACTCACTACATCGGGAACC
SoCHTBD1-7	SOV2g000570.1	TGGTAGTGGCGGTGGTGGTT	AACTAGAGCTTGCTTGGTCGATGAT
SoCHTBD1-8	SOV2g007510.1	CCACCACATTCCATTCCATTGATAGC	TTGCCAATGATGTTGTTGTTGCTTTC
SoCHTBD1-9	SOV2g016080.1	GTGGCAGTGGACCGGCATATTG	GGAGCACCAGCAGCAACCTTG
SoCHTBD1-10	SOV2g016120.1	GTGGTGGTGTTGAAGAGCAGGTC	CAAGGAGCACCAGCAGCAACC
SoCHTBD1-11	SOV2g038970.1	CCTAGTAGCTCCGGCGATGTGG	AGAACGGGCAGCAGCAATGAAAG
SoCHTBD1-12	SOV3g012940.1	ATCCCATCAAAGAATGCCTGAGTCAC	GCGGTCCTGGTAACTGTCAATCTG
SoCHTBD1-13	SOV3g037500.1	GGACCGGCATATTGTGGCGATG	GGCACCAGCTCCAGCAACAG
SoCHTBD1-14	SOV6g037200.1	CCACAGAACCCAAACCTGCT	TCATGACGACCTTGTTGGTTGA
So18s	SOV2g014030.1	CCATAAACGATGCCGACCAG	CCATAAACGATGCCGACCAG

3 Result

3.1 Genome-wide identification and sequence features of *SoCHTBD1* genes

A comprehensive genome-wide search identified 27 SoCHTBD1 family members in the spinach genome, which were systematically designated as *SoCHIBI* genes (Table 2 and supplementary material 1). These genes exhibited an uneven chromosomal distribution across the six spinach chromosomes, with chromosome 1 harboring the highest gene density (10 genes, 37.0%), followed by chromosome 2 (5 genes, 18.5%), chromosome 3 (4 genes, 14.8%), chromosomes 4 and 5 (3 genes each, 11.1% each), and chromosome 6 containing only a single gene (3.7%). This asymmetric distribution pattern suggests potential evolutionary expansion events, particularly on chromosome 1.

The SoCHTBD1 proteins displayed remarkable structural diversity in their physicochemical properties. Protein lengths ranged dramatically from 78 amino acids (SOV1g025680.1, SOV1g025700.1) to 876 amino acids (SOV2g000500.1), representing an 11-fold variation that reflects substantial functional diversification within the gene family. Correspondingly, molecular weights varied from 8.0 kDa to 93.3 kDa, with the majority of proteins (20 out of 27, 74.1%) falling within the smaller size range of 8.0 - 15.0 kDa, while only two proteins exceeded 80 kDa.

Theoretical isoelectric point (pI) analysis revealed a bimodal distribution pattern, with values ranging from 4.61 (SOV3g012940.1) to 9.12 (SOV1g025770.1, SOV1g025800.1). Notably, 14

proteins (51.9%) exhibited basic pI values (>7.0), while 13 proteins (48.1%) showed acidic characteristics (pI <7.0), indicating potential functional specialization under different pH environments. The grand average of hydropathicity (GRAVY) analysis demonstrated that most SoCHTBD1 proteins (24 out of 27, 88.9%) possessed negative GRAVY values, confirming their predominantly hydrophilic nature and supporting their predicted roles in aqueous cellular environments.

Subcellular localization predictions provided crucial insights into the functional deployment of SoCHTBD1 proteins. The majority of family members (22 out of 27, 81.5%) were predicted to localize extracellularly, consistent with their anticipated roles in pathogen recognition and cell wall reinforcement during defense responses. The remaining proteins showed diverse subcellular destinations: two proteins targeted to chloroplasts/plastids (SOV2g007510.1, SOV2g000500.1), two to vacuoles (SOV2g000550.1), and one with ambiguous prediction, suggesting functional diversification beyond traditional extracellular defense mechanisms.

Transmembrane helix analysis further supported the predominantly secretory nature of SoCHTBD1 proteins, with 15 proteins (55.6%) containing no transmembrane helices and 12 proteins (44.4%) possessing a single transmembrane helix, likely representing signal peptides for protein secretion. No proteins contained multiple transmembrane domains, confirming that SoCHTBD1 family members function primarily as soluble rather than integral membrane proteins.

Table 2. Physicochemical properties and subcellular localization predictions of SoCHTBD1 proteins in spinach

Gene ID	Number of Amino Acid	Molecular Weight	Isoelectric point (pI)	Instability Index	Aliphatic Index	Grand Average of Hydropathicity	Predictive subcellular localization	Gene localization	CDS(bp)	Number of predicted TMHs
SOV1g025680.1	78	8115.50	7.60	30.36	77.44	0.327	extr	chr1	237	1
SOV1g025700.1	78	8115.50	7.60	30.73	77.44	0.327	extr	chr1	237	1
SOV1g025710.1	78	8105.48	7.60	31.45	70.00	0.263	extr	chr1	237	1
SOV1g025720.1	78	8105.48	7.60	31.45	70.00	0.263	extr	chr1	237	1
SOV1g025730.1	78	8077.42	7.60	32.54	67.56	0.232	extr	chr1	237	1
SOV1g025740.1	78	8115.50	7.60	30.36	77.44	0.327	extr	chr1	237	1
SOV1g025750.1	78	8077.42	7.60	32.54	67.56	0.232	extr	chr1	237	1
SOV1g025760.1	78	8132.59	8.64	27.27	72.44	0.288	extr	chr1	237	1
SOV1g025770.1	88	9283.99	9.12	46.53	66.82	0.14.	extr	chr1	267	1
SOV1g025790.1	95	9584.30	8.64	45.3	66.11	0.315	extr	chr1	288	1
SOV1g025800.1	88	9268.98	9.12	47.86	65.68	0.132	extr	chr1	267	1
SOV2g000500.1	876	93394.86	8.67	36.48	60.16	-0.142	plas	chr2	2631	0
SOV2g000550.1	773	85034.98	8.88	42.08	72.60	-0.411	vacu	chr2	2322	0
SOV2g000570.1	275	28951.00	4.61	44.48	55.75	-0.117	extr	chr2	828	0
SOV2g007510.1	137	15704.85	6.54	50.11	71.82	-0.323	chlo	chr2	414	0
SOV2g016080.1	78	8008.36	8.28	25.13	67.56	0.279	extr	chr2	237	1
SOV2g016090.1	78	8008.36	8.28	25.13	67.56	0.279	extr	chr2	237	1
SOV2g016110.1	78	8008.36	8.28	25.13	67.56	0.279	extr	chr2	237	1
SOV2g016120.1	78	8107.49	6.68	18.66	73.72	0.310	extr	chr2	237	1
SOV2g016320.1	85	8923.50	6.69	44.79	67.76	0.129	extr	chr2	258	1
SOV2g038970.1	322	34652.86	6.56	42.99	58.57	-0.295	extr	chr2	969	0
SOV3g012940.1	286	30493.73	4.42	28.88	52.31	-0.258	extr	chr3	861	0
SOV3g015100.1	339	37613.34	5.23	43.42	58.97	-0.455	extr	chr3	1020	0
SOV3g037500.1	83	8625.05	5.43	37.78	69.4	0.283	extr	chr3	252	0
SOV3g037540.1	83	8724.32	8.26	35.95	63.49	0.325	extr	chr3	252	1
SOV3g037560.1	83	8724.32	8.26	35.95	63.49	0.325	extr	chr3	252	1
SOV6g037200.1	84	9177.78	8.26	41.21	75.48	0.018	extr	chr6	255	0

Note: Nucl: Nuclear; Cyto: Cytoplasm; Chlo: Chloroplast; Vacu: Vacuole; E.R.: Endoplasmic Reticulum; Mito: Mitochondria; Pero: Peroxisome; Cysk: Cytoskeleton; Extr: Extracellular; Golg: Golgi apparatus

3.2 Phylogenetic classification and evolutionary grouping

Maximum likelihood phylogenetic reconstruction of 22 representative CHTBD1 proteins from spinach, *Arabidopsis thaliana*, and *Oryza sativa* revealed distinct evolutionary lineages organized into four well-supported clades (G1–G4) (Fig. 1). This multi-species analysis provided crucial insights into the evolutionary diversification and species-specific expansion patterns within the CHTBD1 gene family.

Clade G1 represented a spinach-specific monophyletic group comprising six genes (SOV1g025800.1, SOV1g025770.1, SOV1g025790.1, SOV2g016320.1, SOV3g037560.1, and SOV6g037200.1), suggesting recent species-specific duplication events following spinach divergence from its closest relatives. Within this clade, SOV6g037200.1 and SOV3g037560.1 formed a closely related sister pair (designated subfamily I), indicating more recent duplication that may have facilitated functional specialization. The predominance of chromosome 1 genes in this clade supports the chromosomal clustering patterns observed in genome organization analysis.

Clade G2 displayed mixed phylogenetic composition, containing both spinach genes (SOV2g000570.1, SOV2g007510.1, SOV3g012940.1) and an Arabidopsis ortholog (AT3G54420.1). The close phylogenetic relationship between SOV2g000570.1 and AT3G54420.1 (designated subfamily II) indicates conservation of ancestral gene functions predating the Amaranthaceae-Brassicaceae divergence. This evolutionary pattern suggests that certain CHTBD1 functions are fundamental to plant defense mechanisms and have been preserved across taxonomically distant families.

Clade G3 comprised exclusively rice-specific genes (Os04g0493600.1, Os04g0494100.1, Os02g0605900.1), with Os04g0493600.1 and Os04g0494100.1 forming a tightly linked sister pair (subfamily III). This monocot-specific expansion indicates independent evolutionary trajectories following the monocot-dicot split, reflecting lineage-specific adaptation strategies in chitin-mediated defense responses.

Clade G4 contained solely Arabidopsis genes (AT2G43610.1, AT2G43620.1, AT2G43580.1, AT2G43590.1), with AT2G43610.1 and AT2G43620.1 representing closely related paralogs (subfamily IV). The species-specific clustering in both rice and Arabidopsis clades demonstrates convergent expansion patterns, where different plant lineages have independently amplified their CHTBD1 repertoires to meet lineage-specific pathogen pressures.

The phylogenetic topology revealed that spinach *CHTBD1* genes are distributed across both ancient (G2) and recently evolved (G1) lineages, indicating a complex evolutionary history involving both conservation of ancestral functions and acquisition of novel capabilities. The absence of spinach representatives in clades G3 and G4 suggests that certain CHTBD1 subfunctions present in monocots and Brassicaceae have been either lost in the spinach lineage or replaced by functionally equivalent genes not captured in this analysis.

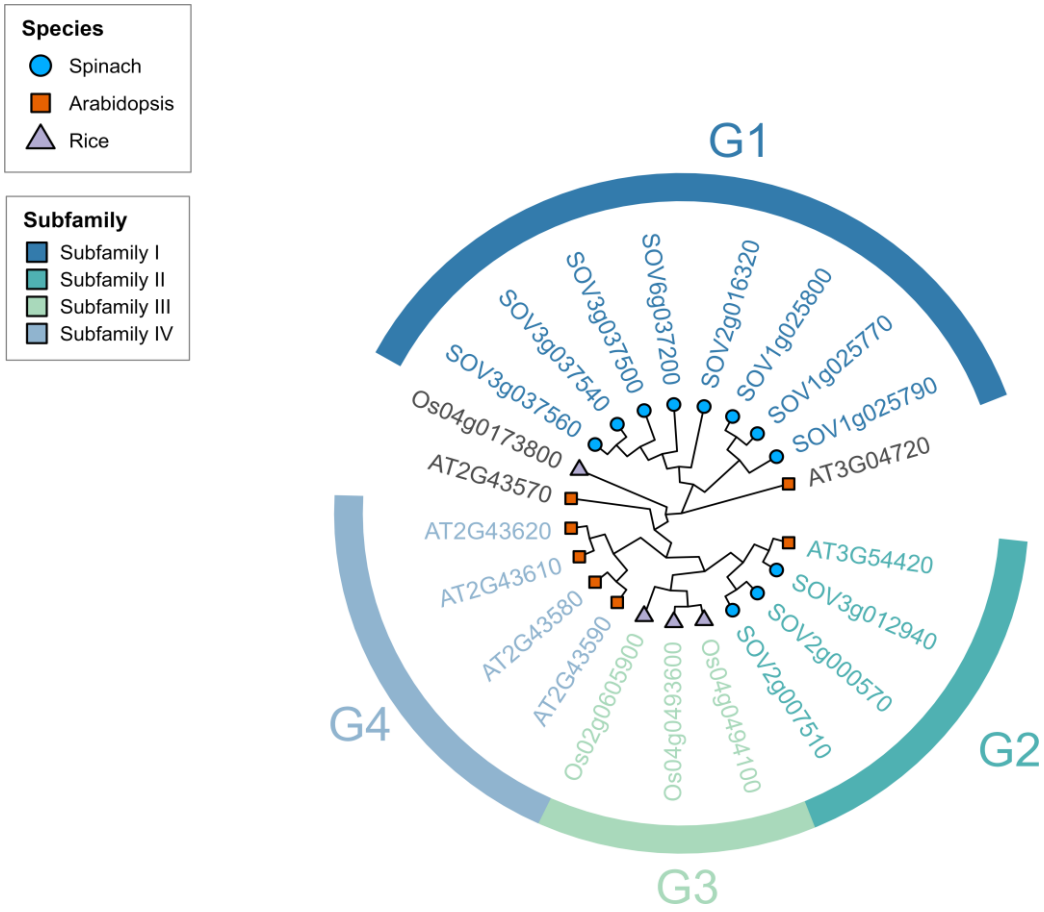

Fig 1. Phylogenetic tree of CHTBD1 proteins from three plant species

Radial maximum-likelihood tree of CHTBD1 proteins from spinach (●), Arabidopsis (■), and rice (▲). Tip labels are protein IDs. Colored outer arcs (G1–G4) delineate four major clades corresponding to Subfamily I–IV in the legend.

3.3 Conserved domain organization and motif architecture

Domain architecture analysis revealed that the SoCHTBD1 family centers on the conserved ChtBD1 domain, with most proteins (22/27, 81.5%) exhibiting single-domain architectures positioned at the N-terminus (Fig. 2). Five proteins displayed multi-domain complexity: SOV2g000550.1 and

SOV2g000500.1 contained additional chitinase class I domains, while SOV2g000500.1 uniquely harbored a tri-domain combination including a GIY-YIG endonuclease domain. SOV2g007510.1 contained only the chitinase class I domain, indicating functional specialization toward direct chitinolytic activity.

Domain patterns were correlated with phylogenetic relationships, with the spinach-specific clade G1 dominated by single-domain proteins (83.3%), whereas the mixed clade G2 comprised exclusively multi-domain variants, consistent with ancient diversification followed by additional domain acquisition.

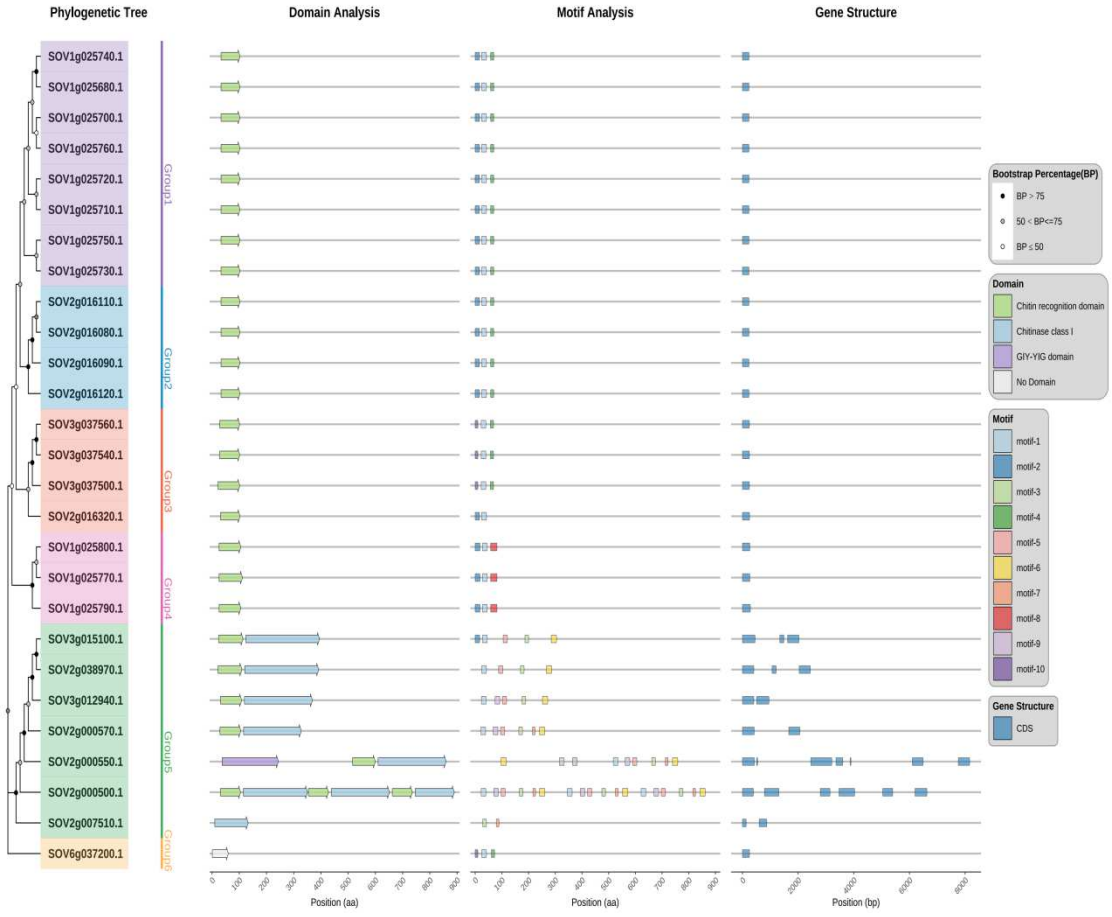

Fig 2. Domain architecture and motif organization of SoCHTBD1 proteins

(A) Maximum-likelihood tree inferred with IQ-TREE v2. Branch support from ultrafast bootstrap (UFBoot; 1,000 replicates) is binned as >75, 50–75, and ≤50. Monophyletic clades are labeled G1–G6. Clade boundaries were delimited at nodes with UFBoot >75; nodes with 50–75 were retained when at least two independent criteria agreed: concordant domain composition, similar MEME motif profiles, or similar exon–intron organization. (B) Protein domain map, range 0–900 aa. (C) MEME motifs 1–10

on protein coordinates. (D) Gene structure with CDS arrows, units in bp. Gene order is consistent across panels.

Motif analysis identified ten conserved motifs across the family. Single-domain proteins consistently exhibited a highly conserved N-terminal core comprising motifs 1, 2, and 3 in fixed sequential order, essential for chitin-binding function. Multi-domain proteins showed expanded motif repertoires, with SOV2g000500.1 containing all ten motifs distributed across the entire protein sequence, reflecting modular acquisition of supplementary functional domains during evolutionary diversification.

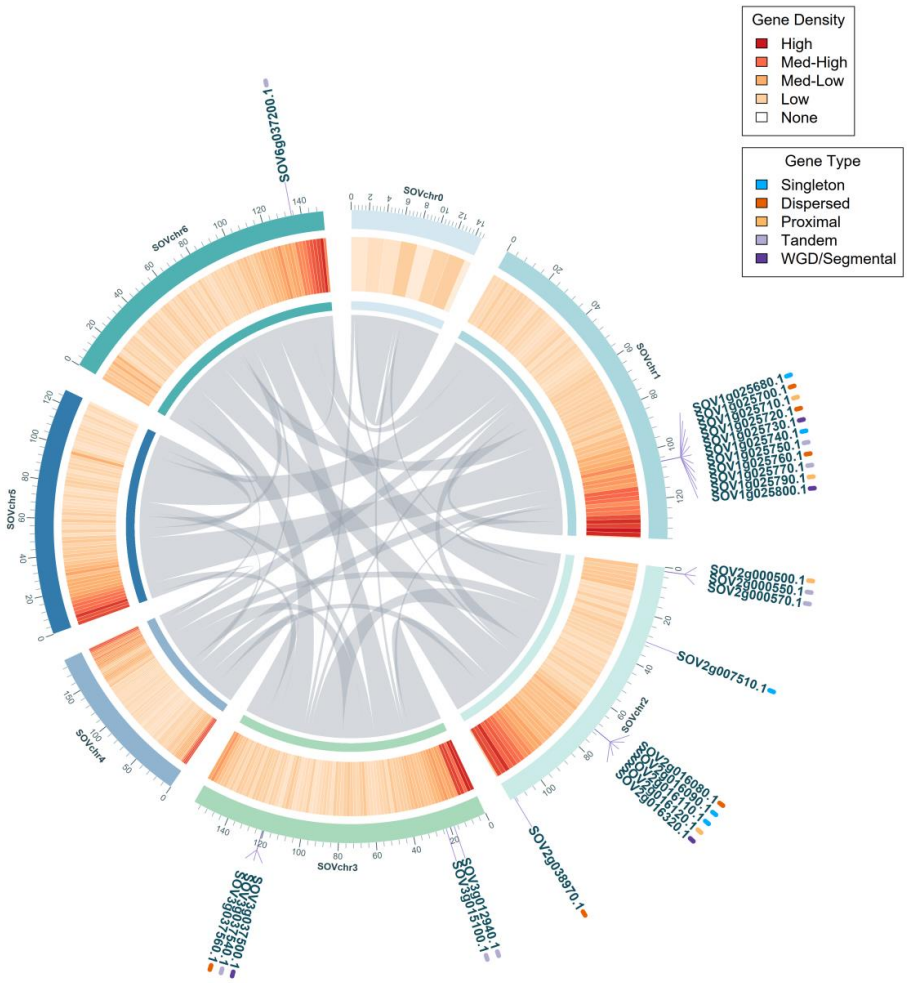

Fig. 3a Chromosomal distribution and duplication patterns of *SoCHTBD1* genes

From the outermost ring inward, the tracks are: chromosome scale with labels SOVchr1–SOVchr6 and tick marks in megabases; a gene-density heat track with a low-to-high gradient; and gray synteny

ribbons linking collinear blocks that contain *SoCHTBD1* genes. Colored squares next to the outer gene labels indicate duplication class—singleton, dispersed, proximal, tandem, and whole-genome or segmental duplication.

3.4 Chromosomal Mapping and Synteny Analysis

Chromosomal mapping revealed uneven distribution of SoCHTBD1 loci across six spinach chromosomes, with notable enrichment on chromosomes 1 and 2 (Fig. 3a). Genome-wide density analysis showed *SoCHTBD1* genes distributed across varying density regions without clear positional preferences, suggesting neutral chromosomal integration patterns.

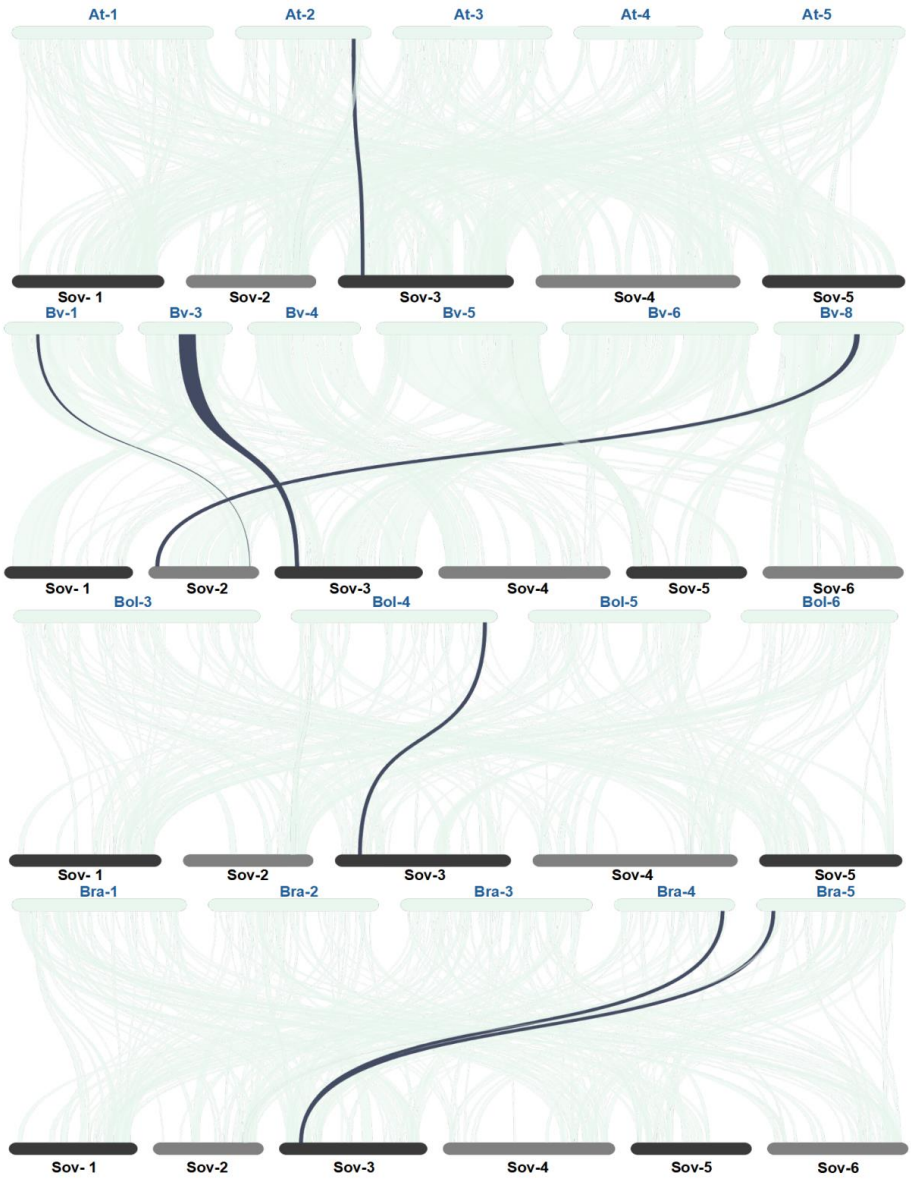

Fig. 3b Cross-species syntenic relationships of spinach SoCHTBD1 loci
Comparisons are shown as spinach–Arabidopsis, spinach–*Beta vulgaris* (sugar beet), spinach–*Brassica*

oleracea, and spinach–*Brassica rapa* from top to bottom. Curved ribbons represent orthologous syntenic blocks; dark ribbons mark blocks that harbor CHTBD1 orthologs, while pale ribbons indicate the genome-wide syntenic background.

Extensive inter-chromosomal syntenic links demonstrated significant segmental collinearity, supporting the inferred duplication mechanisms (Fig. 3a). Cross-species synteny analysis revealed evolutionary conservation patterns correlating with phylogenetic distances (Fig. 3b). Syntenic relationships were strongest between spinach and sugar beet (Amaranthaceae family), with three confirmed orthologous pairs, moderate with Brassica species (two pairs each for B. rapa and B. oleracea), and weakest with Arabidopsis (single orthologous pair). Key syntenic anchors included SOV3g012940.1, SOV2g038970.1, and SOV2g000550.1, suggesting these represent evolutionarily important CHTBD1 functions preserved across taxonomic boundaries.

3.5 Promoter Architecture and Regulatory Elements

PlantCARE analysis of 2-kb upstream promoter regions identified diverse *cis*-acting regulatory elements across *SoCHTBD1* genes, revealing complex transcriptional control networks (Fig. 4). Light-responsive elements were most abundant, followed by phytohormone-responsive and abiotic stress-related motifs, indicating multifaceted environmental regulation of *CHTBD1* gene expression.

Stress-responsive elements showed distinct distribution patterns: ARE (hypoxia response), MBS (drought response), and LTR (low-temperature response) occurred frequently, with *SOV2g000570.1* displaying the strongest enrichment, suggesting heightened abiotic stress sensitivity. Hormone-responsive motifs revealed sophisticated regulatory integration: ABRE (ABA-responsive) and jasmonate-associated elements (CGTCA/TGACG) were widely distributed, with *SOV3g037540.1* and *SOV3g037560.1* showing the highest enrichment levels. Many genes contained ABRE-CGTCA/TGACG combinations, indicating coordinated ABA-jasmonate signaling pathways.

Pathogen-responsive TCA elements (salicylic acid-responsive) were prominently present in defense-related candidates including *SOV2g016320.1* and *SOV2g016120.1*, supporting their potential roles in biotic stress responses. The integration of multiple hormone pathways (ABA, JA, SA) with abiotic stress signals suggests that *SoCHTBD1* genes function within complex regulatory networks that coordinate plant responses to diverse environmental challenges.

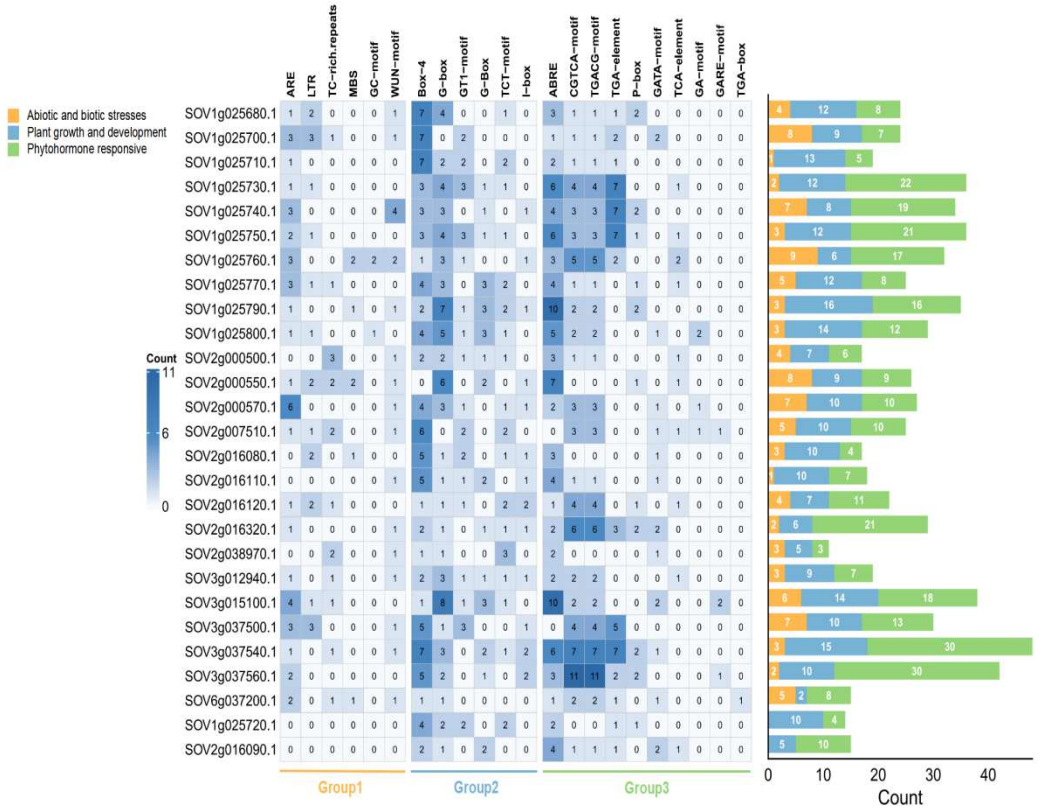

Fig. 4 Cis-regulatory element distribution patterns in *SoCHTBD1* gene promoters

Heatmap shows counts of cis-elements within the 2 kb upstream region for each gene. Columns are element families related to stress, development, and phytohormone signaling; color indicates count per family. The stacked bar at the right summarizes totals per gene, partitioned into three categories: abiotic and biotic stresses, plant growth and development, and phytohormone responsive. Gene clade assignments (Group 1, Group 2, Group 3) are indicated along the bottom.

3.6 Tissue-Specific Expression and Pathogen Response

Quantitative RT-PCR analysis across six tissue types revealed distinct tissue-specific and pathogen-responsive expression patterns for *SoCHTBD1* genes (Fig. 5). In susceptible cultivar S2, pathogen infection predominantly induced gene expression in above-ground tissues, with strongest responses in functional leaves and petioles, moderate induction in bolting organs, and minimal changes in roots and stems. This tissue-specific pattern aligns with the foliar nature of downy mildew infection and supports the predicted extracellular localization of most SoCHTBD1 proteins.

Key responsive genes showed differential induction patterns: *SOV3g037560.1*, *SOV3g037540.1*, and *SOV3g015100.1* exhibited consistently strong upregulation in functional leaves and petioles with

sustained moderate responses in bolting tissues. *SOV2g000570.1* displayed pronounced petiole-specific induction, while *SOV2g016320.1* and *SOV2g016120.1* showed moderate but consistent leaf responses. These expression profiles correlate with the presence of pathogen-responsive TCA elements in their promoters.

Comparative analysis between resistant (U9) and susceptible (S2) cultivars revealed genotype-dependent expression differences (Fig. 5b). The resistant cultivar maintained higher basal expression levels of defense-associated genes, particularly in leaf tissues. Strongly infection-responsive genes (*SOV3g037560.1*, *SOV3g037540.1*, *SOV3g015100.1*) showed enhanced expression in the resistant genotype, suggesting constitutive priming mechanisms. This differential expression pattern indicates that elevated *SoCHTBD1* gene activity contributes to enhanced pathogen resistance through improved chitin recognition and defense activation.

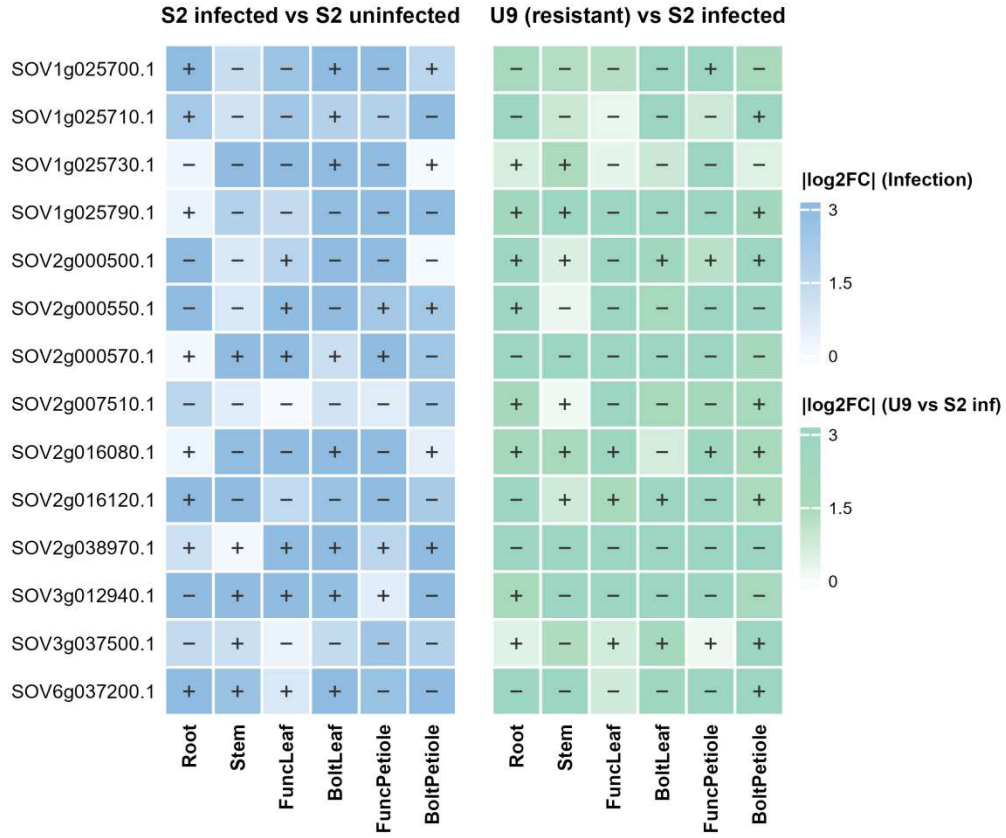

Fig. 5 Expression profiles of *SoCHTBD1* genes in response to downy mildew infection

Heatmaps summarize log2 fold change derived from ΔΔCt for selected genes. Left panel: S2 infected versus S2 uninfected. Right panel: U9 resistant versus S2 infected. Columns denote tissues (Root, Stem, Functional leaf, Bolting leaf, Functional petiole, Bolting petiole). Color intensity encodes the

magnitude of log2 fold change. Symbols mark direction and significance: plus indicates significant up-regulation, minus indicates significant down-regulation, dash indicates not significant.

4 Discussion

4.1 Novelty and Impact

As one of the world's most economically important leafy vegetables, spinach faces persistent production losses from downy mildew caused by *Peronospora effusa* (Gao et al., 2022; Clark et al., 2024). The pathogen evolves rapidly: new physiological races repeatedly emerge and quickly overcome deployed resistance genes (Skiadas et al., 2024). Recent genome assemblies and pangenome analyses reveal clustered effector repertoires with extensive copy-number variation in P. effusa, variation that likely underpins rapid adaptation and resistance breakdown (Skiadas et al., 2024; Fletcher and Michelmore, 2023). Despite these advances, the genomic organization and functional roles of chitin-binding type 1 (CHTBD1) proteins in spinach immunity had not been systematically investigated prior to this study.

This investigation represents the first systematic genome-wide identification and analysis of the *CHTBD1* gene family in spinach, establishing a molecular foundation for understanding chitin-mediated defense mechanisms in this economically significant crop. Our identification of 27 *SoCHTBD1* genes and their detailed structural and regulatory analyses provide important genomic resources for spinach improvement initiatives. Critically, we identified eight genes exhibiting significant pathogen-responsive expression ($\log_2FC \geq 1.5$, $p < 0.01$), with enhanced induction patterns in resistant cultivars, providing initial molecular evidence for CHTBD1 involvement in spinach downy mildew resistance. This work establishes foundations for future functional investigations while providing immediate targets for marker-assisted breeding applications.

4.2 Gene Family Evolution

Our identification of 27 *SoCHTBD1* genes indicates a modestly sized family without evidence of recent large-scale expansion. Comparative context supports species-specific patterns: Arabidopsis encodes 24 chitinase genes distributed across five classes (Passarinho and de Vries, 2002)., while tomato exhibits a substantially larger family of 43 members (Cao et al. 2019). This variation likely reflects distinct pathogen pressures and environmental challenges encountered by different crop species, with spinach maintaining essential chitin-binding functions while avoiding excessive redundancy.

Phylogenetic reconstruction resolved SoCHTBD1 proteins into distinct evolutionary lineages, revealing both ancient conservation and recent species-specific expansions. The spinach-specific clade G1, containing six genes clustered on chromosome 1, suggests recent tandem duplication events facilitating rapid functional diversification. Conversely, the mixed clade G2, harboring both spinach and Arabidopsis orthologs, indicates conservation of ancestral functions predating the Amaranthaceae-Brassicaceae divergence. This phylogenetic architecture demonstrates how gene family evolution balances functional conservation with adaptive innovation (Panchy et al. 2016).

Gene duplication analysis identified diverse expansion mechanisms contributing to current family architecture. Dispersed duplication emerged as the predominant mode, contributing to widespread chromosomal distribution, while whole-genome duplication events provided evidence for ancient polyploidy contributions (Qiao et al. 2019). Extensive inter-chromosomal collinearity suggests multiple duplication mechanisms collectively shaped the SoCHTBD1 family, consistent with established plant gene family evolution patterns (Lynch and Conery, 2000). ross-species synteny analysis revealed conservation patterns that track the phylogenetic relationships. Sugar beet showed the strongest conservation with spinach, with three orthologous gene pairs identified., reflecting their close relationship within the Amaranthaceae, while progressively weaker relationships with Brassicaceae species indicate that *CHTBD1* gene family evolution has followed species divergence patterns while maintaining core functional constraints.

4.3 Functional Mechanisms

Chitin is a central MAMP in plant–fungus interactions: plants perceive chitin via LysM-type PRR complexes, whereas fungi counter by modifying cell-wall polymers and secreting chitin-binding effectors that sequester oligomers and shield hyphae, thereby suppressing chitin-triggered immunity (Sánchez-Vallet et al., 2014). However, five proteins display complex multi-domain architectures that suggest functional diversification: SOV2g000500.1 uniquely harbors a tri-domain combination including a GIY-YIG endonuclease domain, potentially conferring nucleic acid processing functions alongside chitin-mediated defense activities.

The conserved four-disulfide “hevein” core within ChtBD1 domains mediates specific binding to β(1→4)-linked N-acetylglucosamine oligomers (chitin), a polymer abundant in fungal cell walls and invertebrate exoskeletons (Asensio et al., 2000). Our motif analysis has identified ten conserved motifs across the family, with single-domain proteins consistently exhibiting highly conserved N-terminal

arrangements that are essential for chitin-binding function, while multi-domain proteins show expanded motif repertoires reflecting modular acquisition of supplementary functional domains.

Promoter analysis revealed numerous stress- and hormone-responsive cis-elements (e.g., ARE, MBS, LTR; ABRE, CGTCA/TGACG, TCA), indicating that *SoCHTBD1* genes are embedded in transcriptional networks integrating environmental and hormonal cues. This is consistent with models in which salicylic acid exerts transcriptional control over jasmonate signaling via canonical promoter motifs (e.g., TGACG/as-1, GCC-box), coordinating defense outputs alongside ABA, JA, and SA pathways (Caarls et al., 2015). The integration of ABA, jasmonate, and salicylic acid signaling suggests a coordinated hormone-mediated defense response.

Our expression profiling has revealed distinct tissue-specific and pathogen-responsive patterns that support functional predictions. The preferential expression in above-ground tissues aligns with the foliar nature of downy mildew infection, while enhanced expression in resistant cultivars suggests constitutive priming mechanisms. Key candidates including *SOV2g016080.1*, *SOV2g016120.1*, and *SOV3g012940.1* have demonstrated significant pathogen-responsive upregulation, which correlates with the presence of pathogen-responsive regulatory elements in their promoters.

4.4 Cross-Species Comparisons

Cross-species functional studies provide robust support for the conserved importance of CHTBD1 proteins in plant defense mechanisms. In pepper, *CaChiIII7*, which harbors the conserved ChtBD1 domain, was strongly induced upon *Colletotrichum acutatum* infection, with gene silencing resulting in reduced resistance and increased fungal biomass accumulation (Ali et al. 2020). Similarly, In soybean, overexpression of the fungal chitinase gene *Tachi* from *Trichoderma asperellum* confers enhanced resistance to *Sclerotinia sclerotiorum*; transgenic plants display increased ROS accumulation and elevated peroxidase and superoxide dismutase activities after infection, whereas catalase shows no significant change (Zhang et al., 2016), thereby confirming the functional relevance of ChtBD1-mediated pathogen recognition across taxonomically distant species.

The tissue-specific expression patterns observed in spinach are consistent with reports from pepper, where several ChtBD1-containing genes show preferential expression in aerial tissues and are inducible by pathogen and stress cues (Ali et al., 2018). This conservation of expression patterns across taxonomically distant species reinforces the fundamental role that ChtBD1-mediated recognition plays in apoplastic defense responses.

Structural studies have established that the conserved four-disulfide hevein core in ChtBD1 domains binds chitin oligomers with high specificity, providing a molecular rationale for chitin-associated defense (Asensio et al., 2000). Consistent with this mechanism, functional work in crops shows that enhancing the activity of ChtBD1-containing chitinases increases resistance and elevates immune readouts such as ROS bursts and PR-gene induction (Ali et al., 2020; Zhang et al., 2016). These observations indicate that ChtBD1-mediated functions are transferable across species, aligning with the regulatory circuitry inferred from our promoter analyses.

5 Conclusion

This study establishes the first comprehensive molecular framework for understanding CHTBD1-mediated defense mechanisms in spinach through genome-wide identification of 27 *SoCHTBD1* genes, combined with phylogenetic, domain, and expression analyses that reveal evolutionary patterns characterized by both ancient functional conservation and recent adaptive expansions.

Building upon this framework, we further identified eight pathogen-responsive key genes, among which *SOV2g016080.1*, *SOV2g016120.1*, and *SOV3g012940.1* exhibited the most significant disease resistance-associated characteristics. Notably, the enhanced expression patterns of these candidate genes in resistant cultivars, together with the enrichment of defense-related regulatory elements in their promoter regions, collectively provide direct molecular evidence for *CHTBD1* gene involvement in spinach downy mildew resistance.

Therefore, this research not only provides important theoretical foundations for understanding plant chitin-mediated immunity mechanisms, but also establishes valuable genetic resources for spinach disease resistance breeding. These findings will promote the development of more precise and durable downy mildew resistance improvement strategies, which holds significant implications for enhancing the sustainable development of the spinach industry.

Acknowledgments: This research was funded by Shanghai Engineering Research Center of Plant Germplasm Resources (17DZ2252700).

Conflicts of Interest: The authors declare no conflicts of interest.

Data availability

All data generated or analyzed during this study are included in this published article and its supplementary information files.

References

Ali M, Luo D-X, Khan A, Haq SU, Gai W-X, Zhang H-X, Cheng G-X, Muhammad I, Gong Z-H (2018) Classification and genome-wide analysis of chitin-binding proteins gene family in pepper (*Capsicum annuum* L.) and transcriptional regulation to *Phytophthora capsici*, abiotic stresses and hormonal applications. *Int J Mol Sci* 19(8):2216. https://doi.org/10.3390/ijms19082216

Ali M, Li QH, Zou T, Wei AM, Gombojab G, Lu G, Gong ZH (2020) Chitinase gene positively regulates hypersensitive and defense responses of pepper to *Colletotrichum acutatum* infection. *Int J Mol Sci* 21:6624. https://doi.org/10.3390/ijms21186624

Asensio JL, Cañada FJ, Siebert H-C, Laynez J, Poveda A, Nieto PM, Soedjanaamadja UM, Gabius H-J, Jiménez-Barbero J (2000) Structural basis for chitin recognition by defense proteins: GlcNAc residues are bound in a multivalent fashion by extended binding sites in hevein domains. *Chemistry & Biology* 7:529–543.

Bandi V, Gutwin C (2020) Interactive exploration of genomic conservation. in: Proceedings of the 46th Graphics Interface Conference (GI'20). Canadian Human-Computer Communications Society, Waterloo, Canada. https://synvisio.github.io/

Bhattarai G, Clark KJ, Correll JC, Koike ST, Smith S, Shi A, Mou B (2021) High resolution mapping and candidate gene identification of downy mildew race 16 resistance in spinach. *BMC Genomics* 22:478. https://doi.org/10.1186/s12864-021-07788-8

Caarls L, Pieterse CMJ, Van Wees SCM (2015) How salicylic acid takes transcriptional control over jasmonic acid signaling. *Frontiers in Plant Science* 6:170. https://doi.org/10.3389/fpls.2015.00170

Cai M, Liu Y, Gan L, Wang Y, Zhang H, Han F, Guo J, Xie J, Liu Y (2024) Receptor-like proteins: decision-makers of plant immunity. *Phytopathol Res* 6:7. https://doi.org/10.1186/s42483-024-00279-0

Cao J, Tan X (2019) Comprehensive analysis of the chitinase family genes in tomato (*Solanum lycopersicum*). *Plants* 8:52. https://doi.org/10.3390/plants8030052

Clark KJ, Feng C, Anchieta AG, Van Deynze A, Correll JC, Klosterman SJ (2024) Dual transcriptional characterization of spinach and *Peronospora effusa* during resistant and susceptible race-cultivar interactions. *BMC Genomics* 25:937. https://doi.org/10.1186/s12864-024-10809-x

Cornell Vegetable Program (2023) Spinach downy mildew – disease factsheet. Cornell University. https://www.vegetables.cornell.edu/pest-management/disease-factsheets/spinach-downy-mildew/. Accessed 6 Oct 2024

Fletcher K, Michelmore R (2023) Genome-enabled insights into downy mildew biology and evolution. *Annu Rev Phytopathol* 61:165–183. https://doi.org/10.1146/annurev-phyto-021622-103440

Gao S, Zhang H, Zhou Y, Mou B (2022) Fine mapping and identification of a candidate gene of downy mildew resistance, *RPF2*, in spinach (*Spinacia oleracea* L.). *Int J Mol Sci* 23:14872. https://doi.org/10.3390/ijms232314872

Gu Z, Eils R, Schlesner M (2016) Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* 32:2847–2849. https://doi.org/10.1093/bioinformatics/btw313

Gu Z, Gu L, Eils R, Schlesner M, Brors B (2014) Circlize implements and enhances circular visualization in R. *Bioinformatics* 30:2811–2812. https://doi.org/10.1093/bioinformatics/btu393

Kolde R (2019) pheatmap: Pretty Heatmaps. R package version 1.0.12. https://CRAN.R-project.org/package=pheatmap

Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2−ΔΔCt method. *Methods* 25:402–408. https://doi.org/10.1006/meth.2001.1262

Lynch M, Conery JS (2000) The evolutionary fate and consequences of duplicate genes. *Science* 290:1151–1155. https://doi.org/10.1126/science.290.5494.1151

Mengarelli DA, Zanor MI (2021) Genome-wide characterization and analysis of the CCT motif family genes in soybean (*Glycine max*). *Planta* 253:15. https://doi.org/10.1007/s00425-020-03537-5

Panchy N, Lehti-Shiu M, Shiu SH (2016) Evolution of gene duplication in plants. *Plant Physiol* 171:2294–2316. https://doi.org/10.1104/pp.16.00523

Passarinho PA, de Vries SC (2002) *Arabidopsis* chitinases: a genomic survey. *Arabidopsis Book* 1:e0023. https://doi.org/10.1199/tab.0023

Qiao X, Li Q, Yin H, Qi K, Li L, Wang R, Zhang S, Paterson AH (2019) Gene duplication and evolution in recurring polyploidization–diploidization cycles in plants. *Genome Biol* 20:38. https://doi.org/10.1186/s13059-019-1650-2

Sánchez-Vallet A, Mesters JR, Thomma BPHJ (2014) The battle for chitin recognition in plant–microbe interactions. *FEMS Microbiology Reviews* 39:171–183. https://doi.org/10.1093/femsre/fuu003

Skiadas P, Riera Vidal S, Dommisse J, Warris S, van den Burg HA, de Ridder D, Govers F (2024) Pangenome graph analysis reveals extensive effector copy-number variation in spinach downy mildew. PLoS Genet 20:e1011452. https://doi.org/10.1371/journal.pgen.1011452

Volk H, Marton K, Flajšman M, Radišek S, Tian H, Hein I, Podlipnik Č, Thomma BPHJ, Košmelj K, Javornik B, Berne S (2019) Chitin-binding protein of *Verticillium nonalfalfae* disguises fungus from plant chitinases and suppresses chitin-triggered host immunity. Mol Plant Microbe Interact 32:1378–1390. https://doi.org/10.1094/MPMI-03-19-0079-R

Wang LG, Lam TT, Xu S, Dai Z, Zhou L, Feng T, Guo P, Dunn CW, Jones BR, Bradley T, Zhu H, Guan Y, Jiang Y, Yu G (2020) Treeio: an R package for phylogenetic tree input and output with richly annotated and associated data. Mol Biol Evol 37:599–603. https://doi.org/10.1093/molbev/msz240

Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, Grolemund G, Hayes A, Henry L, Hester J, Kuhn M, Pedersen TL, Miller E, Bache SM, Müller K, Ooms J, Robinson D, Seidel DP, Spinu V, Takahashi K, Vaughan D, Wilke C, Woo K, Yutani H (2019) Welcome to the tidyverse. J Open Source Softw 4:1686. https://doi.org/10.21105/joss.01686

Wang Y, Tang H, DeBarry JD, Tan X, Li J, Wang X, Lee TH, Jin H, Marler B, Guo H, Kissinger JC, Paterson AH (2012) MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res 40:e49. https://doi.org/10.1093/nar/gkr1293

Wilkins D (2022) Gggenes: Draw Gene Arrow Maps in ggplot2. R package version 0.5.1. https://github.com/wilkox/gggenes

Yu G, Smith DK, Zhu H, Guan Y, Lam TT (2017) Ggtree: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. Methods Ecol Evol 8:28–36. https://doi.org/10.1111/2041-210X.12628

Zhang F, Ruan X, Wang X, Liu C, Wan C, Liu W, Jiang L (2016) Overexpression of a chitinase gene from *Trichoderma asperellum* increases disease resistance in transgenic soybean. Appl Biochem Biotechnol 180:1542–1558. https://doi.org/10.1007/s12010-016-2186-5