BMC Plant Biology

RESEARCH

Open Access

Identification and expression analysis of papain-like cysteine proteases gene family and response to *B. cinerea* stress in *F. vesca*

Yibo Bai[1*†], Haibin Wang[3†], Min Wu[2] and Zong-Ming Cheng[2*]

Abstract

Background Papain-like cysteine protease (PLCP), a vital subgroup of peptidases, play crucial roles in various biological processes including plant growth, seed germination, anther development, and stress responses.

Results In this study, 44 *FvPLCP* genes were identified through phylogenetic tree analysis and divided into 9 groups. Chromosomal localization revealed that the *FvPLCP* genes are unevenly distributed on 7 chromosomes. The Ka/Ks results demonstrated that *FvPLCP* genes have predominantly undergone purifying selection during evolution. The phylogenetic tree and motif analysis results indicated that *FvPLCP* genes has a conserved domain and different subfamilies have special motifs, suggesting that different subfamily members have different functions. The *cis*-element analysis indicated that *FvPLCPs* contains a significant number of *cis*-elements under biotic and abiotic stress. The single-cell transcriptome and dual transcriptome of woodland strawberry in response to *B. cinerea* indicated that the subfamily members of *FvPLCP* genes exhibit specific expression in different cell types and infection stages. RT-qPCR and strawberry fruit transient expression experiments showed that the *FvPLCP* genes may be involved in the strawberry defense process against *B. cinerea* by participating in the hormone pathway.

Conclusions These findings provided a theoretical basis for further elucidating the role of *FvPLCP* genes in response to biotic stress.

Keywords Woodland strawberry, *Botrytis cinerea*, *FvPLCP* genes, Biotic stress

†Yibo Bai and Haibin Wang contributed equally to this work.

*Correspondence:
Yibo Bai
2019204017@njau.edu.cn
Zong-Ming Cheng
zmc@njau.edu.cn
[1]National Key Laboratory for Tropical Crop Breeding, Key Laboratory of Crop Gene Resources and Germplasm Enhancement in Southern China, Ministry of Agriculture and Rual Affairs, Key Laboratory of Tropical Crops Germplasm Resources Genetic Improvement and Innovation of Hainan Province, Tropical Crops Genetic Resources Institute, Chinese Academy of Tropical Agricultural Sciences, Haikou 571101, China
[2]College of Horticulture, Nanjing Agricultural University, Nanjing 210095, China
[3]National Key Laboratory for Tropical Crop Breeding, School of Breeding and Multiplication Sanya Institute of Breeding and Multiplication, Hainan University, Sanya 572025, China

Background

The process of plant growth and development, such as seed germination, growth, flowering, and senescence, were vulnerable to various adverse effects. The health of plants will be damaged by stress factors, particularly biotic stress factors such as fungi, bacteria, and viruses, which will lead to plant death or disease. To respond to a variety of biotic stresses, such as the majority of microbial pathogens and other invaders, plants have evolved sophisticated immune defense systems.

The core of the plant immune system, papain-like cysteine proteases (PLCPs), is regulated by a series of variables in the plant defense responses [1]. These enzymes

function through multiple mechanisms: PLCP protein precursors induce the accumulation of salicylic acid (SA) by releasing the immune signal peptide Zip1 [2], thereby activating the SA signaling pathway. This SA accumulation subsequently enhances PLCP protein activity, creating a positive feedback loop that promotes the expression of SA-related immune genes [3]. However, pathogens like Phytophthora infestans can subvert this defense through effectors such as EP11, which inhibits PLCP activation in the proteolytic cascade [4]. Conversely, certain plant metabolites like maize-derived 10-OPEA (10-oxo-11-phytoenoic acid) can enhance PLCP activity, triggering programmed cell death (PCD) as a defense mechanism [5]. The activated PLCP OsXCP2a in rice demonstrates the broad regulatory capacity of these enzymes, upregulating genes associated with SA, jasmonic acid (JA), and ethylene (ET) hormone pathways, while also enriching disease resistance genes in the phenylalanine metabolic pathway [6], These findings underscore the crucial role of PLCP-dependent SA signaling in plant immunity. Additional evidence suggests nuclear involvement, with studies showing that Ralstonia solanacearum's PopP2 effector mediates nuclear translocation of RD19, implying potential nuclear regulatory functions for PLCPs in immune gene expression [7].

Structurally, PLCPs belong to the C1A family of cysteine proteases (CA clan) and function through cysteine-dependent proteolytic activity [8]. The protein structure of PLCPs has been reported to be mostly composed of an α-helix domain and a ß-sheet domain, which together constitute the papain-like folding of the protein [9]. The cysteine catalytic function of the PLCPs protein is located at the N-terminal of the α-helix domain. The substrate binding site of the PLCPs protein contains the catalytic triplet Cys-His-Asn at the deep cleavage site at the junction of the two domains, which proved that PLCPs function as cysteine proteases [10]. As stable proteolytic enzymes, PLCPs perform essential physiological functions in various organelles including lysosomes, vacuoles, and exosomes, participating in critical processes ranging from developmental transitions (senescence, vegetative and reproductive growth) to defense responses [11].

While PLCP functions have been characterized in model systems including *Arabidopsis* [12], Oryza sativa [13], Vitis vinifera [14], Ficus carica [15] and Gossypium hirsutum [16], their roles remain largely unexplored in Rosaceae species. The woodland strawberry (*Fragaria vesca*) offers an ideal model for Rosaceae studies due to its compact genome (240 Mb), diploid nature (2n = 14), rapid life cycle, and genetic tractability [17]. This study comprehensively analyzes the PLCP gene family in *F. vesca*, examining gene structures, conserved domains, protein characteristics, and evolutionary relationships. By integrating single-cell transcriptome and dual-RNAseq, we elucidate the expression dynamics of *FvPLCP* genes across cell types and during early infection stages, providing a foundation for advancing disease resistance research and breeding programs in Rosaceae crops.

Materials and Methods

Identified and classification of FvPLCP genes in woodland strawberry

All PLCP protein sequences were identified in the woodland strawberry genome (https://www.rosaceae.org/species/fragaria_vesca/genome_v4.0.a1/) by the HMM search (PF00112.26) (http://www.hmmer.org/) with default parameters [18]. To ensure the reliability of the identified PLCPs, all candidate sequences were further verified using SMART (http://smart.embl-heidelberg.de/) and the Pfam database (http://pfam.xfam.org/) [19].

Sequence alignment and phylogenetic analysis

The protein sequences of AtPLCPs were retrieved from the *Arabidopsis* information resource (TAIR) database (https://www.arabidopsis.org/) [12]. The protein sequences of OsPLCPs were retrieved from phytozome (https://phytozome-next.jgi.doe.gov). The phylogenetic tree was constructed by IQ-tree using the maximum likelihood method with 1000 bootstrap replicates [20]. The phylogenetic tree was visualized and refined by Evolview for enhanced clarity and presentation [21].

The analysis of FvPLCP genes structure and conserved motifs

The conserved motifs of the FvPLCP proteins were predicted by the MEME site (https://meme-suite.org/meme/) [22], with the maximum number of motifs set to 15. The conservative motifs were predicted by the MEME parameters. The structure analysis of *FvPLCP* genes was performed using the program GSDS2.0 website (http://gsds.cbi. pku.edu.cn/) [23].

Chromosomal locations of FvPLCP genes

To identify the chromosomal location of *FvPLCPs*, we retrieved and determined the location and length of the *FvPLCPs* from the annotated file of the *Fragaria vesca* downloaded from the GDR website. The software of MapChart was used for visualization [24]. For duplication events, *Ka* (non-synonymous substitution rate), *Ks* (synonymous substitution rate), and Ka/Ks ratios of *FvPLCP* gene duplication clusters were calculated using Perl script.

Protein properties and subcellular localization prediction

Physical properties of FvPLCP proteins, including molecular weight (MW), isoelectric points (pls), and grand average of hydropathicity (GRAVY), were analyzed by the

ExPASy website (https://web.expasy.org/ProtParam) [25]. The subcellular localization of FvPLCP proteins was predicted by the WoLF PSORT website (https://www.genscript.com/wolf-psort.html) [26].

Coexpression network of FvPLCP families

Pearson correlation coefficients (PCCs) of different *FvPLCP* genes were calculated using Cytoscape_v3.7.2 software (https://cytoscape.org) [27], and the co-expression network ($|PCCs| \geq 0.75$, $P < 0.05$) was constructed to visualize the *FvPLCP* genes.

Transcriptome analysis of FvPLCP genes during B. cinerea infection

To investigate the expression pattern of the *FvPLCP* genes in early response to *B. cinerea* stress in woodland strawberry, we selected the single-cell transcriptome and dual transcriptome data of *B. cinerea* infected woodland strawberry at 0–12 hpi (hour post infection) for analysis. For single-cell RNA-seq data, the fastQC was used to evaluate the quality of raw sequencing data, and the Cell Ranger was used to align single-cell data to the reference genome of 'Hawaii4' to form a gene-cell expression matrix (https://www.rosaceae.org). Filter out low-quality cells by removing cells with a high proportion of mitochondrial genes (> 20%) and low UMI (nFeature_RNA < 200). Predict potential double cells using DoubletFinder and filter them. The SCTransform was used to standardize library size to eliminate technical bias. Correction of batch effects between three single-cell samples using CCA + method. The dimensionality reduction analysis was performed by PCA. For bulk RNA-seq data, The FastQC was to check the quality of the raw sequencing data and filter the data to obtain clean data. Align clean reads to the reference genome using HISAT2. Generate the original gene count matrix through featureCounts. DESeq2 was used for differential expression gene analysis to identified DEGs ($|\log2FC| > 1$, $P_{adj} < 0.05$). The dataset of CRA004848 was downloaded from BIG website (https://ngdc.cncb.ac.cn/gsa) [28]. The dual transcriptom e data (PRJNA907512) was downloaded from the NCBI website (https://www.ncbi.nlm.nih.gov) [29].

RNA isolation and RT-qPCR analyses

The plant materials of woodland strawberry 'Hawaii4' used in this experiment were all preserved in the laboratory of fruit biosystems at Nanjing Agricultural University (Nanjing, Jiangsu). For the sample of scRNA-seq analysis, we divided the leaf tissue into four cell types according to the previous study [28], and extracted the RNA in different cell types using the Arcturus PicoPure RNA Isolation kit (Applied Biosystems, USA). Reverse transcribed RNA into cDNA using the Prime Script RT reagent kit with gDNA Eraser (Takara, Dalian, China), and stored it in the −80 ℃ for standby. RNA extraction from the leaves of woodland strawberry treated with *B. cinerea* at 0 hpi, 0.5 hpi, 1 hpi, 2 hpi, 3 hpi, 4 hpi, 5 hpi, 6 hpi, 7 hpi, 8 hpi, 9 hpi, 10 hpi, 11 hpi and 12 hpi, and reverse transcription into cDNA. Select different subfamily genes within the *FvPLCP* genes for RT-qPCR analysis, and the primer sequences were shown in Table S4. *Fvactin11* (*FvH4_7g22410*) was used as an internal reference gene to normalize the expression of genes by $2^{-\triangle\triangle Ct}$ method, each set of RT-qPCR data was calculated with three replicates [30].

Expression by the transient transformation in strawberry fruit

The coding sequence (CDS) of *FvSAG12* was cloned into PJ35s vector and pFGC5941 vector, respectively. Empty vectors (PJ35s and pFGC5941) were maintained as negative controls. Selected strawberry fruits at 30 days postanthesis (DPA) for inoculation. Select fruits about 30 days after flowering for injection. Prepared *B. cinerea* conidial suspension at 5×10^6 spores/mL in sterile water. Inoculate 10 strawberry fruits in each treatment and observe the changes in disease spots in different treatments.

Results

Genome-wide identification of FvPLCP genes in woodland strawberry

A total of 44 *FvPLCP* genes were identified by the Peptdase_C1 domain (PF00112.26) using the hidden Markov model (HMM) profile (Table S1). We performed the prediction analysis of the FvPLCP proteins. The protein isoelectric point ranged from 4.47 to 9.33, while its molecular weight (MW) ranged from 17.68 KDa to 57.99 KDa. The GRAVY of the 40 FvPLCP proteins was less than 0, indicating that *FvPLCP* genes were hydrophilic proteins. The results of subcellular localization prediction showed that 14 FvPLCP proteins were located in the extracellular mesenchyma, 11 FvPLCP proteins were located in the chloroplasts, 7 FvPLCP proteins were located in the nucleus, 4 FvPLCP proteins were located in the vacuoles, 3 FvPLCP proteins were located in the cytoplasm, 3 FvPLCP proteins were located in the endoplasmic reticulum.

After analyzing full-length protein sequence alignment with the IQ-tree program, the phylogenetic trees for 44 FvPLCPs, 31 AtPLCPs, and 45 OsPLCPs were constructed. Consistent with previous studies [14], the FvPLCPs were classified into nine subfamilies, which were RD21 (responsive to desiccation 21), CEP (cysteine endopeptidase), XCP (xylem cysteine peptidase), XBCP (xylem bark cysteine peptidase), THI (THI like-subfamily), SAG12 (senescence-associated gene 12), RD19 (responsive to desiccation 19), ALP (aleutian-like

protease), and CTB (cathepsin B-like) and presumably performed different biological functions (Fig. 1).

Gene structure, and conserved motif analysis of the FvPLCP genes

We used the protein sequences of FvPLCPs to investigate the gene structure and conserved motifs of *FvPLCPs* in nine different subfamilies. In our study, 15 conservative motifs were identified in *FvPLCPs*. Motifs (motif 1, 2, 3, 4, 5) which contained the Peptidase_ C1 conservative structural domain appeared in 44 *FvPLCPs*. Additionally, some motifs were rare and belong to one subfamily. For instance, 11 *FvPLCPs* include motif 12 (Granulin), which was exclusively present in the SAG12 subfamilies. Only the RD19 subfamily contains motif 14 (Papain-like-cys-pep-sf), which was dispersed throughout four members. Among the 15 motifs, motifs 1, 2, 3, 4, 5 were characterized as Peptidase_C1 domain, motifs 6 and 7 as Inhibitor_I29 domain, and motifs 15 as GRAN_2. Most *FvPLCPs* contain the Peptidase_C1 domain and the Inhibitor_I29 domain. 43 *FvPLCPs*, except *FvH4_3g14350*, include the Peptidase_C1 domain, and 43 *FvPLCPs* with motif 6 and motif 7 contain the Inhibitor_I29 domain. Only 4 *FvPLCPs* with motif 15, namely *FvH4_2g05500*,

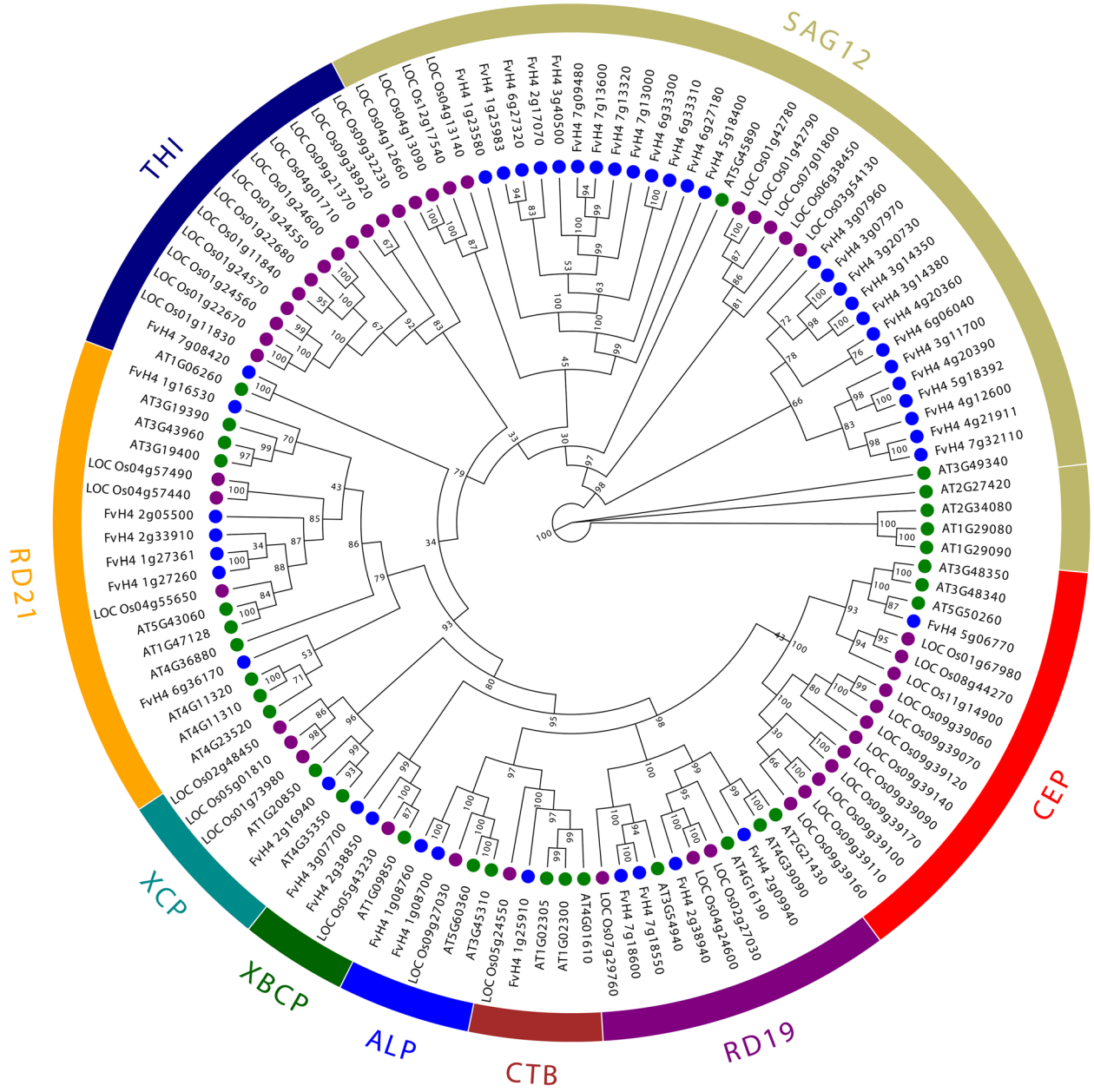

Fig. 1 Phylogenetic analysis and identification of PLCP gene families in rice, Arabidopsis, and woodland strawberry. All genes were divided into 9 subfamilies, with subfamilies distinguished by different colors and shapes. Three species were represented by dots of different colors

FvH4_2g33910, *FvH4_2g38850*, and *FvH4_3g07700*, included granulin-like domains (Fig. 2).

Analysis of FvPLCP gene structures revealed substantial variation in intron numbers across subfamilies, ranging from the single-intron architecture of THI members to the ten-intron organization characteristic of CTB genes. The ALP, RD21, CEP, XCP, and XBCP subfamilies have more than or equal to 3 introns, belonging to subfamilies with a large number of introns. With 24 FvPLCPs having two or fewer introns, the majority of FvPLCPs in SAG12 subfamily have comparatively few introns.

Chromosomal location and expansion of FvPLCP genes

The *FvPLCPs* were mapped on 7 chromosomes of the woodland strawberry to explore the location of *FvPLCPs.* 44 *FvPLCP* genes were unevenly distributed on 7 chromosomes. Fvb1, Fvb3, and Fvb7 have the highest proportion, with 8 *FvPLCP* genes on each chromosome. There were only three *FvPLCP* genes on Fvb5, which contains the least number of *FvPLCP* genes compared to other chromosomes. These results suggested that the woodland strawberry may contain genetic diversity in the *FvPLCP* genes.

We found five duplication events in FvPLCP gene family, including tandem duplication, transposed duplication, proximal duplication, dispersed duplication, WGD duplication, and collinearity duplication events. Dispersed duplication events accounted for the largest proportion of 78.33%. Tandem duplication and collinearity duplication accounted for the smallest proportion, which was 3.33%. It would seem that dispersed duplication events had a significant role in the expansion of *FvPLCPs*. (Figs. 3).

Evolutionary selection pressure analysis of FvPLCP genes was performed by calculating Ka/Ks ratios, where values $< 1, > 1$, and $= 1$ indicate purifying selection, positive selection, and neutral selection, respectively. The observed Ka/Ks ratios (0–0.55) for tandem duplication events strongly suggest that FvPLCP genes in woodland strawberry have predominantly undergone purifying selection during evolution. (Fig. 4).

Cis-element analysis of promoter regions in FvPLCP genes

To understand the function of *FvPLCP* genes, we selected the 2000 bp promoter of the *FvPLCP* genes for analysis. A total of 496 *cis*-elements were predicted in the *FvPLCP* promoter regions in woodland strawberry. The *FvPLCP* genes mainly contain the elements of hormone responsiveness, anaerobic induction, defense and stress responsiveness, low-temperature responsiveness, drought-inducibility, and light responsiveness, which indicated that the *FvPLCP* genes in strawberry have the function of participating in biotic and abiotic stress in plants (Fig. 5).

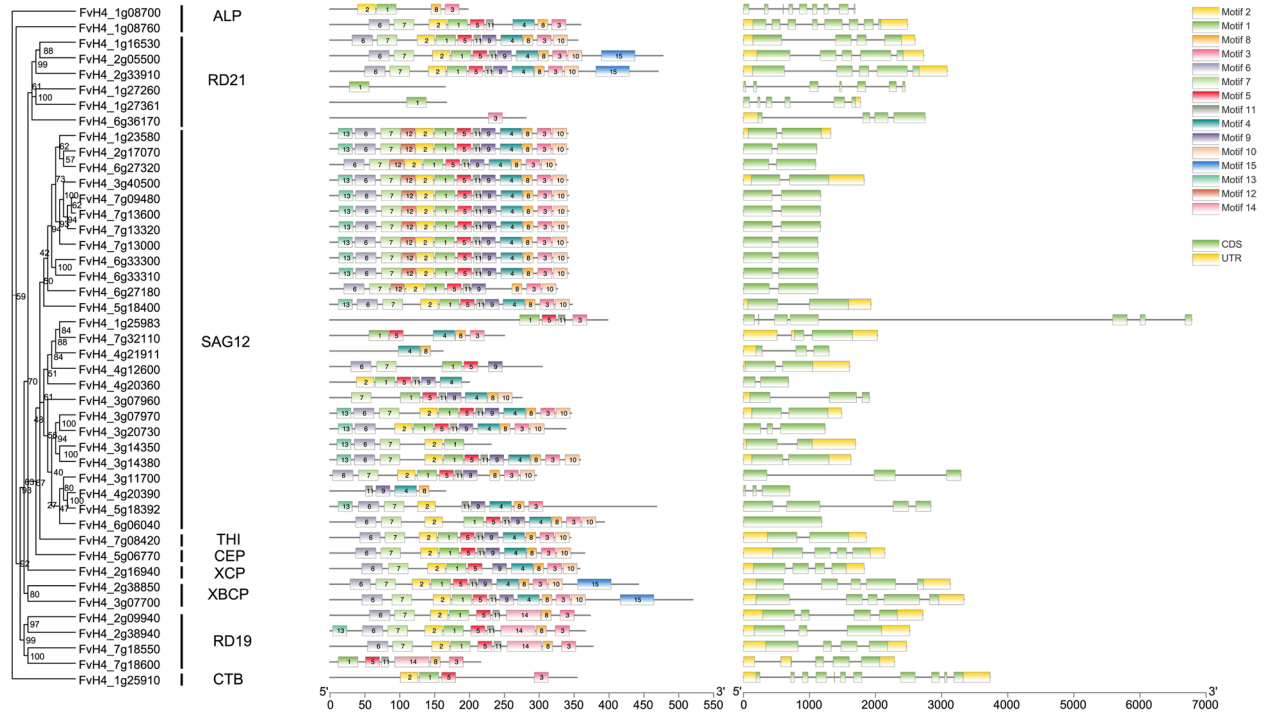

Fig. 2 Analysis of FvPCLP genes structure and conservative motifs

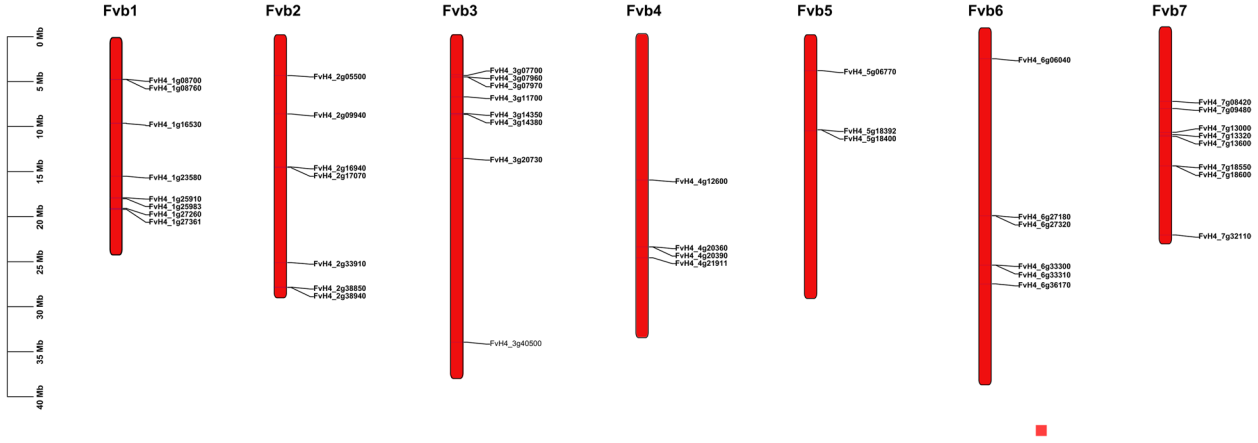

Fig. 3 Chromosomal localization analysis of FvPCLP genes

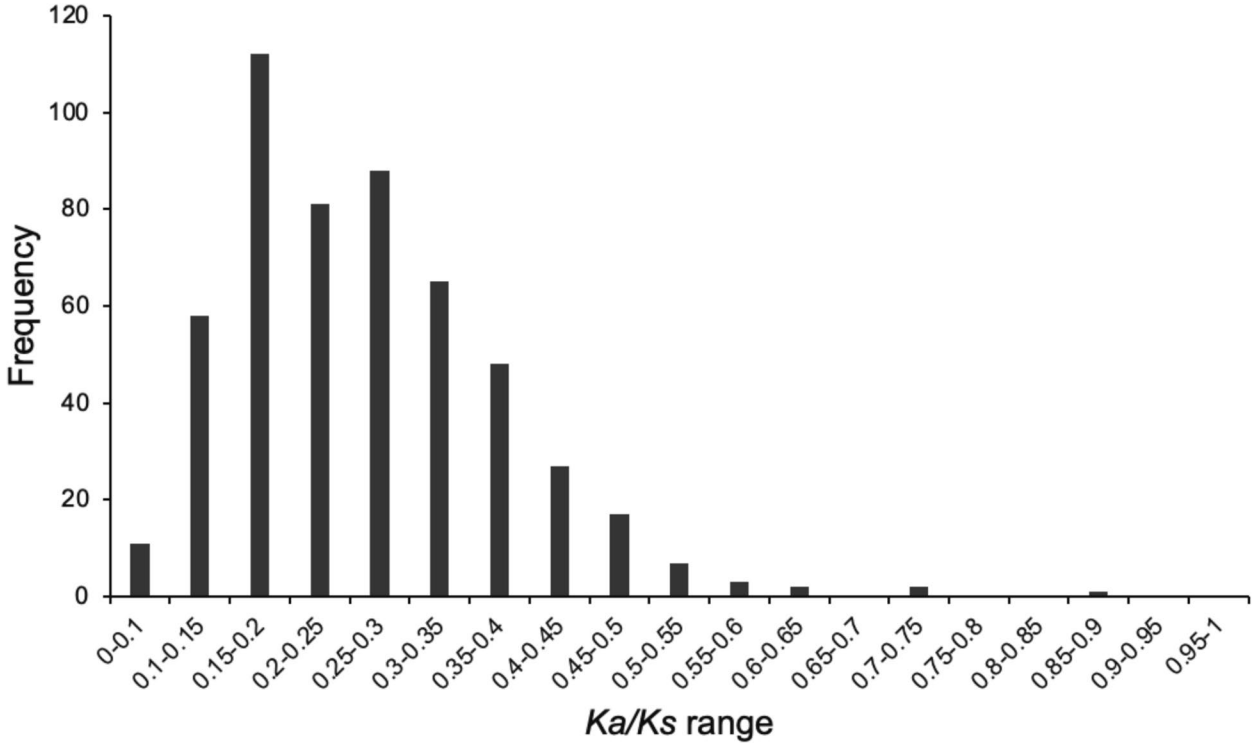

Fig. 4 The Ka/Ks range of FvPCLP gene pairs in woodland strawberry

Co-expression network of FvPLCP families with hormone-related genes

To elucidate potential regulatory relationships, we constructed a co-expression network between *FvPLCP* genes and hormone-related genes in woodland strawberry (Figs. 6 and 7, Table S3). The network comprised 31 hormone-related genes and 17 *FvPLCPs* connected by 69 co-expression edges ($|PCCs| \geq 0.85$), revealing both strong positive (31 edges) and negative (38 edges) correlations. Notably, seven *FvPLCPs* exhibited particularly high connectivity (>6 co-expression events each), with *FvH4_2g38940*, *FvH4_2g38850* and *FvH4_2g33910* showing the most extensive interactions (10 edges each), suggesting their potential as central hubs in hormone-mediated regulatory networks.

Tissue-specific expression analysis of the FvPLCP genes

To investigate the role of *FvPLCP* genes in woodland strawberry growth and development, we showed the expression profile of *FvPLCP* genes using transcriptome data from 42 different tissues and developmental stages. *FvPLCP* genes were mainly highly expressed in organs and tissues such as anther, leaf, style, ghost, embryo, pollen, carpel, seeding, and root. The expression levels of

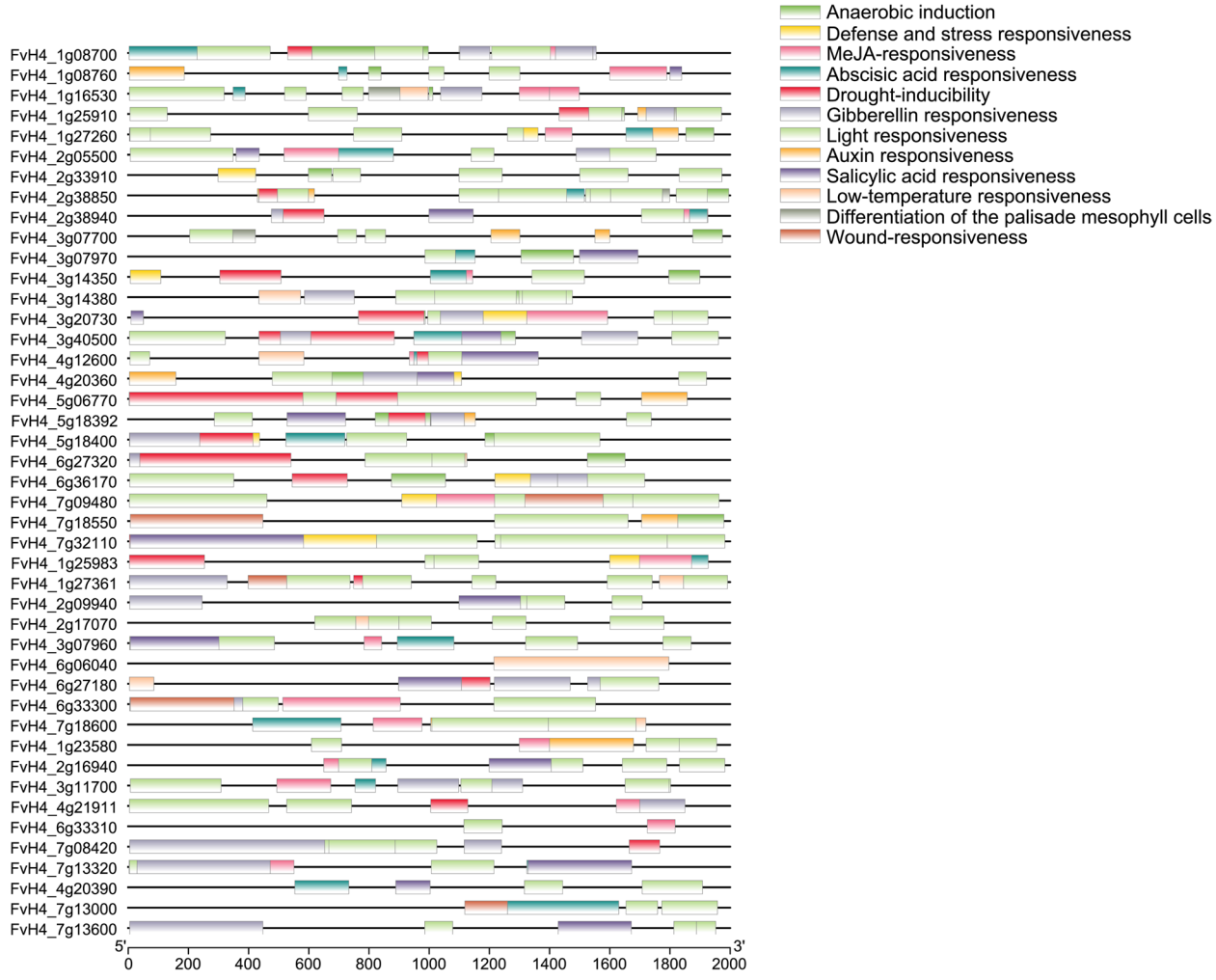

Fig. 5 Cis-element analysis of FvPCLP gene promoter regions

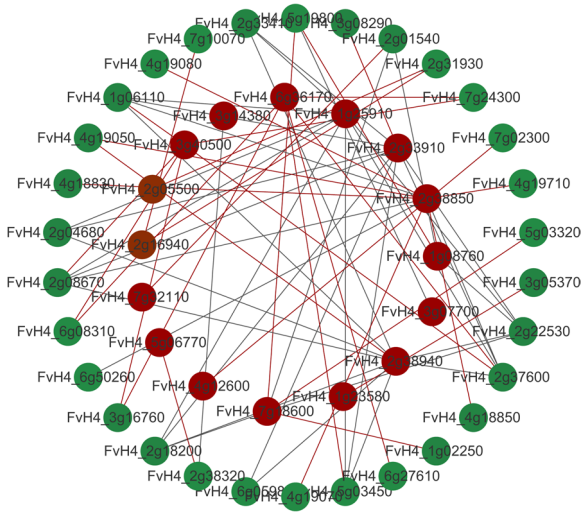

Fig. 6 Co-expression network analysis of FvPLCP genes with hormone related genes. The gray lines represent a negative correlation between gene pairs, and the red lines represent a positive correlation between gene pairs

FvPLCP genes in different tissues were different. Among them, the expression patterns were similar of *FvPLCP* genes in embryo, pollen, and seeding. These results indicated that the *FvPLCP* genes play an important role in the growth and development of strawberry.

Transcriptome and RT-qPCR analysis of FvPLCP genes and hormone related genes during B. cinerea infection

We analyzed the spatial and temporal expression patterns of *FvPLCP* genes in the early stage of *B. cinerea* infection of woodland strawberry by combining single-cell transcriptome and dual RNA-seq data. By integrating single-cell data with *Arabidopsis* and comparing homologous genes, the scRNA-seq data of strawberry was divided into 12 cell type [28]. The results of dual RNA-seq showed that the *FvPLCP* genes in RD21 (*FvH4_6g36170*), THI (*FvH4_7g08420*), SAG12 (*FvH4_6g33310* and *FvH4_3g40500*), ALP (*FvH4_1g08700*) subfamilies were mainly highly expressed within 0.5–3 hpi. The genes of RD21 (*FvH4_2g33910*), CTB (*FvH4_1g25910*),

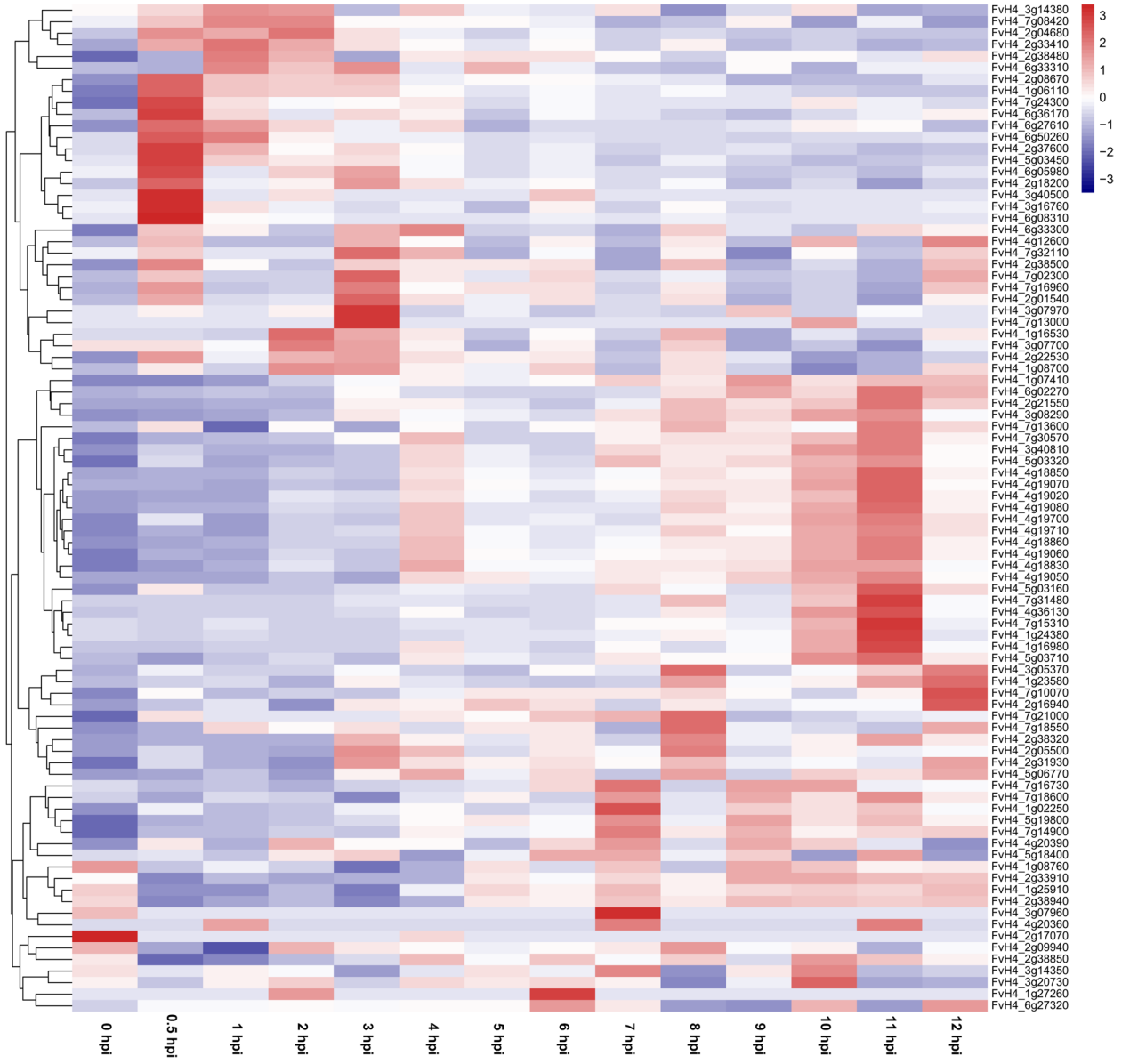

Fig. 7 Dual transcriptome analysis of FvPLCP genes and hormone related genes expression in the early stage of woodland strawberry infection with B. cinerea

CEP (*FvH4_5g06770*), XBCP (*FvH4_2g38850*), ALP (*FvH4_1g08760*), RD19 (*FvH4_2g38940*, *FvH4_7g18550* and *FvH4_7g18600*) subfamilies were highly expressed at 7–12 hpi (Fig. 8). We selected the representative genes from each subfamily for RT-qPCR analysis (Fig. 9). The results showed that the expression levels of ALP, RD21, XBCP, RD19, and CTB subfamily genes gradually increased at 7–12 hpi. The SAG12 subfamily genes were highly expressed at time points of 0.5 hpi and 3 hpi, while the THI subfamily genes exhibit periodic high expression from 0 to 5 hpi (Fig. 9). In addition, we selected hormone-related genes highly correlated with the *FvPCLP* genes for RT-qPCR analysis (Fig. S1). Among them, the expression levels of most hormone-related genes increase with the prolongation of treatment time. However, the gibberellin regulated gene (*FvH4_2g33410*), which is negatively correlated with the CTB subfamily gene (*FvH4_1g25910*), gradually increases in expression levels between 0–2 hpi and then shows a decreasing trend. Combined with the high expression of *FvH4_1g25910* at 7–12 hpi, it further confirms the possible negative correlation regulation between the two. In addition, the expression trend of the SAG12 subfamily gene (*FvH4_4g12600*) is similar to that of the DIR6 gene (*FvH4_7g02300*), which is positively correlated. This result suggests that the expression of SAG12 subfamily genes may be regulated by the salicylic acid pathway (Fig. 10).

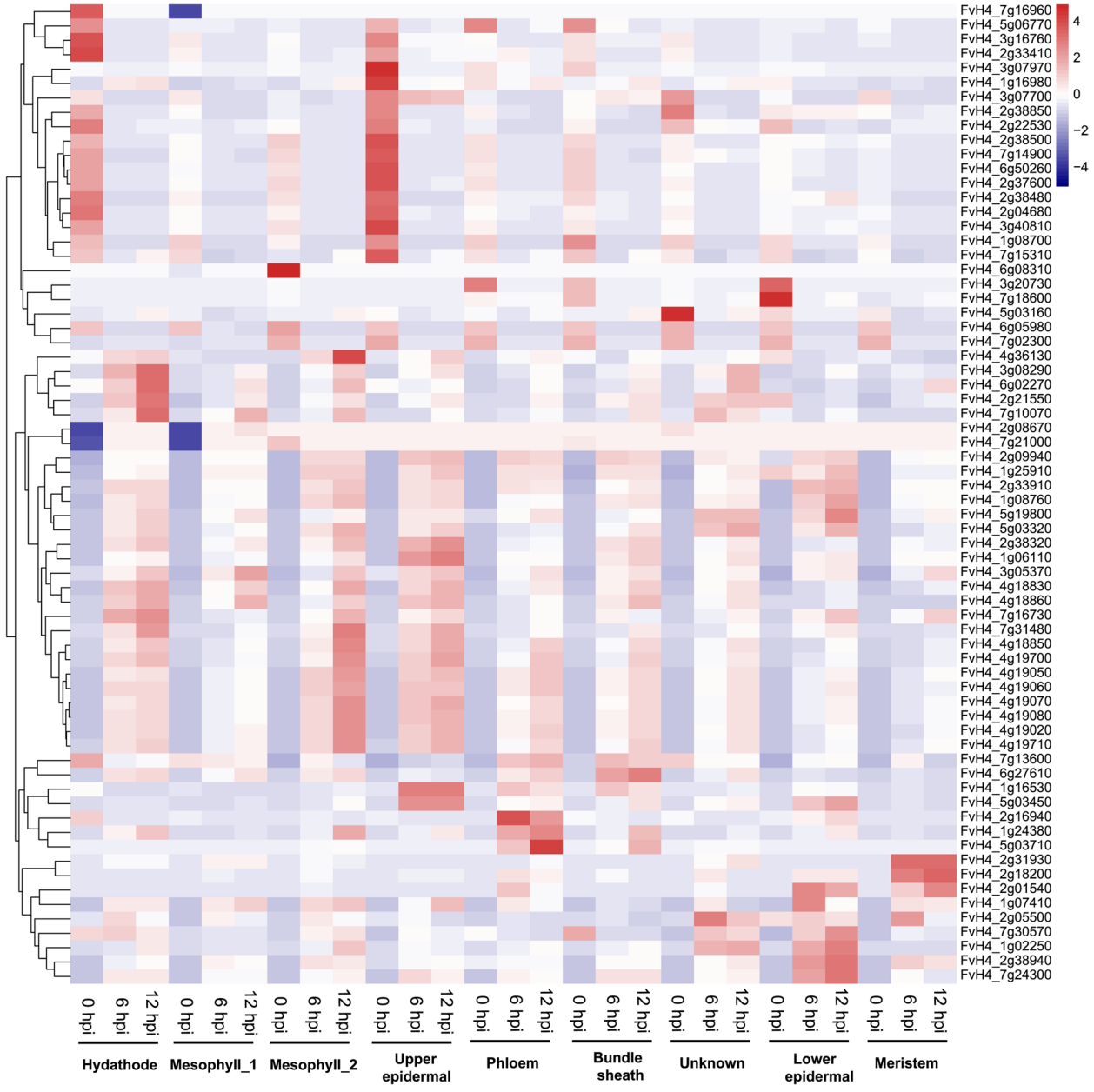

Fig. 8 Single-cell RNAseq analysis of FvPLCP genes and hormone related genes expression in the early stage of woodland strawberry infection with B. cinerea

To reveal the spatiotemporal expression of the *FvPLCP* genes in different cell types, we conducted single-cell transcriptome analysis at three timepoints (0, 6, and 12 hpi). The heatmap indicated that most *FvPLCP* genes were highly expressed between 0–12 hpi in hydathode and mesophyll cells. In addition, RD21 (*FvH4_1g16530*) and XBCP (*FvH4_3g07700*) were mainly highly expressed in upper epidermal cells, XCP (*FvH4_2g16940*) were mainly highly expressed in phloem cells, SAG12 (*FvH4_7g32110* and *FvH4_4g12600*) were mainly expressed in hydathode cells (Fig. 8). We use laser capture microdissection technology to divide strawberry leaves into four types: upper epidermal, lower epidermal, mesophyll, and vascular tissue. The RT-qPCR results indicated that most *FvPLCP* genes subfamily were highly expressed in the epidermal and mesophyll, and the expression levels gradually increased with the prolongation of treatment time. These results indicated that *FvPLCP* genes can participate in the process of strawberry response to *B. cinerea*. Notably, we found that hormone-related genes highly positively correlated with the *FvPLCP* genes were highly expressed in various tissues, and the expression increased with treatment time (Fig. S2). Similarly, Gibberellin regulated gene (*FvH4_2g33410*) negatively correlated with

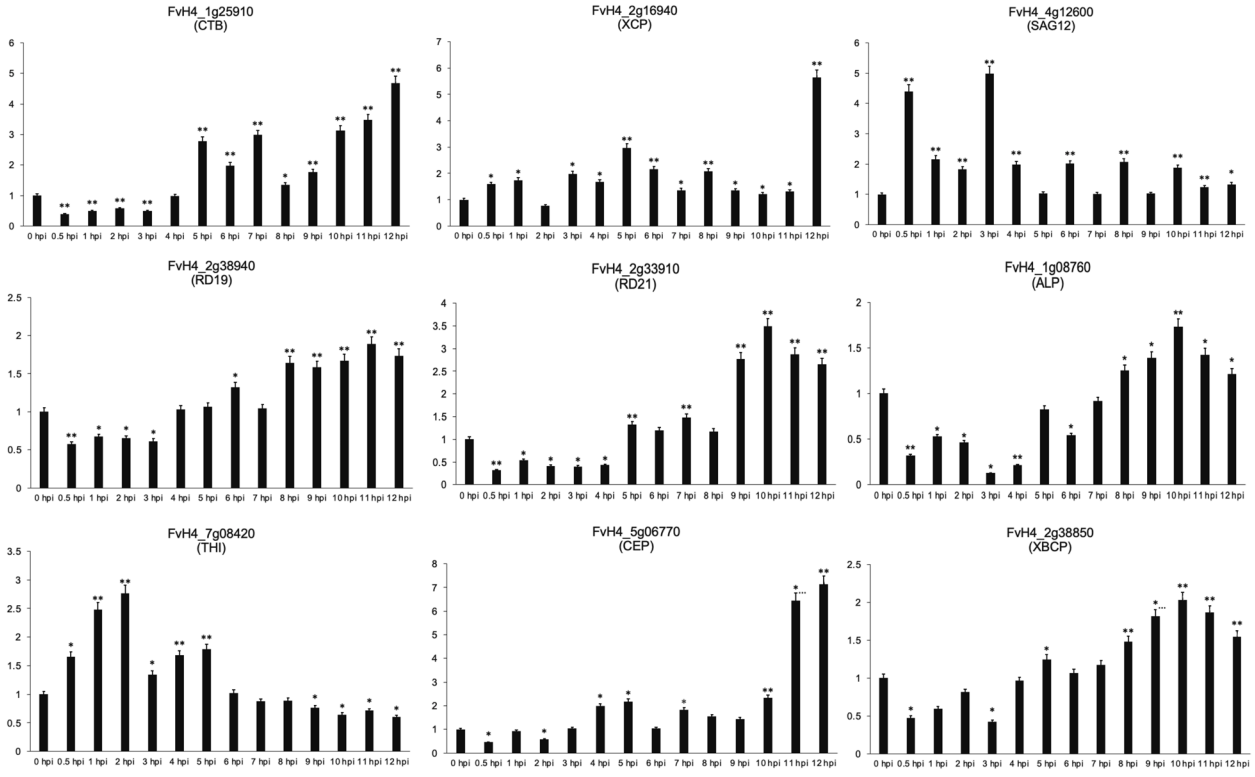

Fig. 9 Relative expression of the FvPLCP genes at different infection stages under the invasion of B. cinerea into woodland strawberry

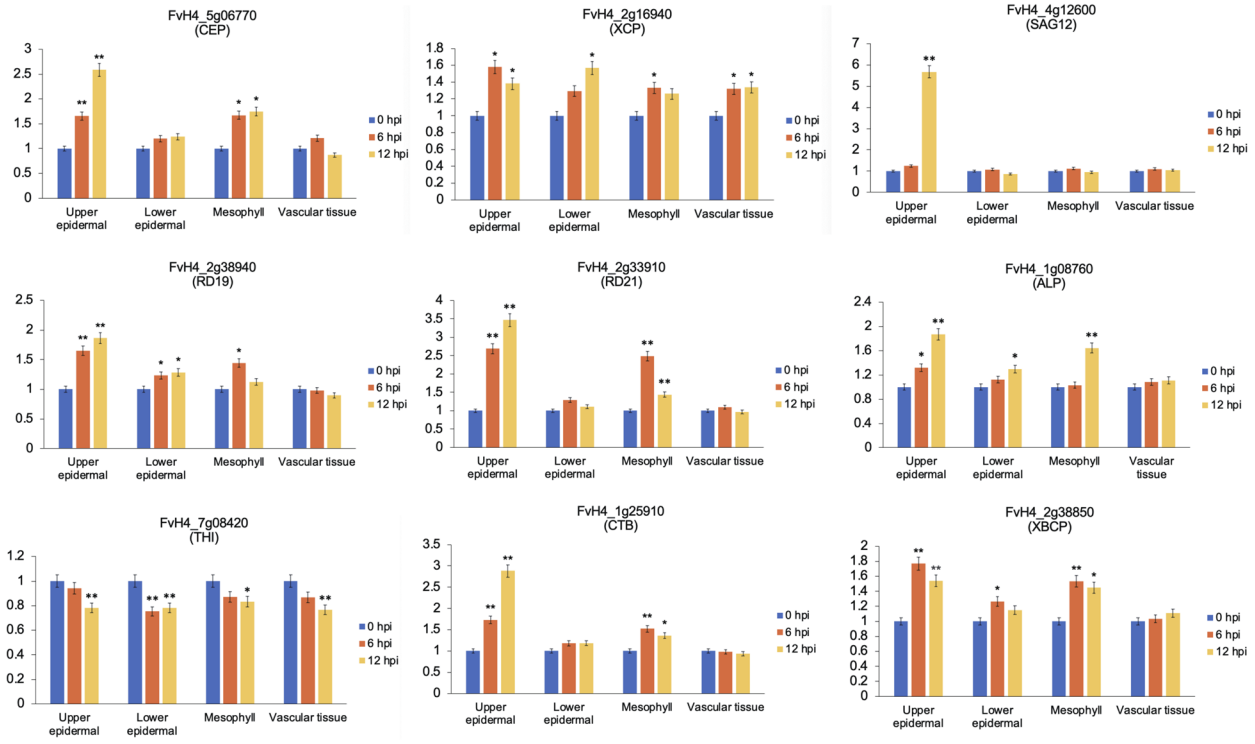

Fig. 10 Relative expression of FvPLCP genes at different cell types under the invasion of B. cinerea into woodland strawberry

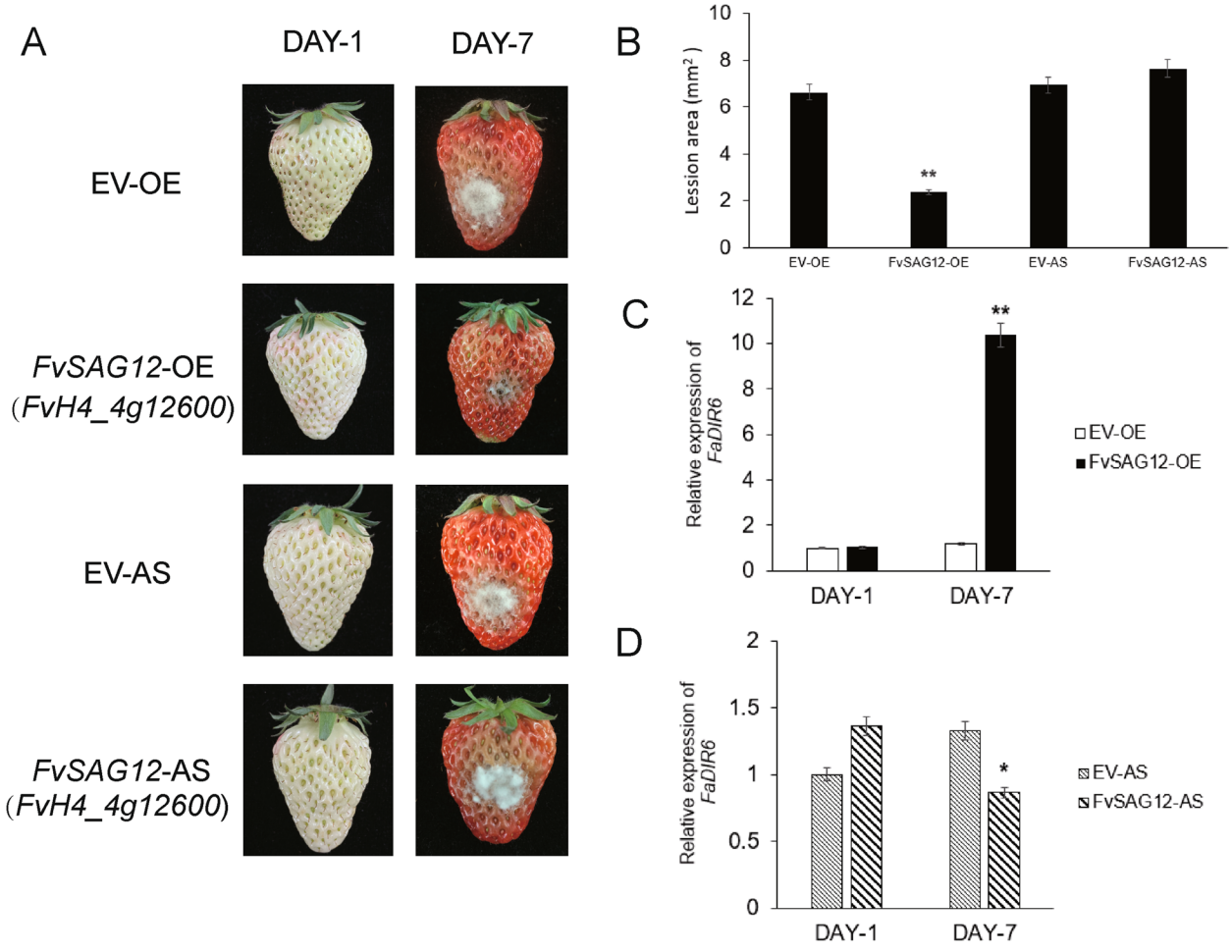

Fig. 11 Phenotype produced by transient overexpression or silencing of FvSAG12 in strawberry fruits. (A) Lesion development after inoculation of strawberry fruits with Botrytis cinerea. DAY-1 represented the first day after inoculation with B. cinerea. (B) Measurement of lesion area using ImageJ. (n=10, Student's t-test, *P<0.05 and **P<0.01). (C) Relative expression levels of overexpression of FaSAG12 in fruit from different treatments (n=10, Student's t-test, *P<0.05 and **P<0.01). (D) Relative expression levels of silencing FaSAG12 in fruit from different treatments (n=10, Student's t-test, *P<0.05 and **P<0.01)

CTB subfamily gene (*FvH4_1g25910*) showed a decreasing trend with prolonged treatment time in various tissues. This result indicated that *FvH4_2g33410* negatively regulates the response of strawberry to infection by *B. cinerea*.

FvSAG12 gene positively regulates the response process of strawberry to B. cinerea

To further investigate the role of *FvPLCP* genes in strawberry response to *B. cinerea* stress, we selected the SAG12 subfamily gene (*FvH4_4g12600*) for transient transformation of strawberry fruit. Functional analysis revealed that FvSAG12-overexpressing plants exhibited significantly enhanced resistance to B. cinerea, as evidenced by 64% smaller lesion areas compared to controls at 7 dpi ($p < 0.01$). Molecular characterization showed a concomitant 10.36-fold upregulation of FaDIR6 in transgenic plants, confirming the positive regulatory relationship between these defense-related genes (Fig. 11).

Discussion

Throughout their life cycle, plants encounter diverse biotic and abiotic stresses that activate sophisticated defense mechanisms, with proteases serving as crucial mediators of protein turnover and stress signaling [31]. As a type of cysteine hydrolase in plants, the PLCP gene family has been widely studied in multiple species [8]. In plants, the PLCP gene family has been comprehensively reported in nine species, *Arabidopsis thaliana* [32], *Oryza sativa* [13], *Glycine max* [33], *Vitis vinifera* [14], *Ficus carica* [15], *Gossypium hirsutum* [16], *Carica papaya* [34], *Hevea brasiliensis* [35], *Ricinus communis* and *Jatropha curcas* [6], with varying gene family size. 97 *GmPLCPs* were present in *Glycine max*, 45 *OsPLCPs* were present in *Oryza sativa*, and 31 *AtPLCPs* were present in *A. thaliana*. 44 *PLCP* genes were identified in strawberry, and each gene family member contained a conserved peptidase-C1 domain. Evolutionary tree analysis divided 44 *PLCP* genes into 9 subfamilies, namely

RD21, THI, SAG12, ALP, CTB, CEP, XBCP, RD19, and XCP subfamilies. Among them, each subfamily member contains different motifs. For example, motif 12 (Granulin) only exists in the SAG12 subfamily, while motif 14 (Papain-like-cys-pep-sf) only exists in the RD19 subfamily. These results suggested that the PLCP gene family in strawberry was highly conserved in angiosperms, combined with motif analysis, it indicated that the *FvPLCP* genes may have different functions. Comparative genomic analysis revealed that *FvH4_1g25983*, a SAG12 subfamily member, possesses unique structural features distinguishing it from other *FvPLCP* genes. Phylogenetic and syntenic analyses suggested that its extended CDS sequence may result from misannotation of adjacent open reading frames.

As the main way of gene family amplification, gene duplication plays an important role in plant evolution [36]. The high retention rate of ancient repetitive events and existing duplicated gene pairs contributes to the enrichment of duplicated genes in the plant genome [37]. These duplicates contribute to the evolution of new functions [38]. The study of the GmPLCP gene family revealed many tandem duplication events and two significant segment duplication events [39–41]. 37 *GmPLCPs* were engaged in the two massive segment duplication events, which were crucial in the expansion of the GmPLCP gene family. Therefore, there were significantly more *GmPLCPs* in *Glycine max* than in other species, including *AtPLCPs* in *A. thaliana* and *VvPLCPs* in *V. vinifera* [32, 41]. In our study, proximal duplication events played a dominant role in various duplication events, laying the foundation for the expansion of the *FvPLCP* gene family in strawberry. The Ka/Ks ratio was used to evaluate the selection pressure of gene family members [42]. Interestingly, the Ka/Ks result of the *FvPLCP* gene pairs is less than 1, indicating that the PLCP gene family is mainly influenced by purification selection to remove harmful mutations [43].

Members of the PLCP family have been reported to be involved in programmed cell death, immune responses, and seed germination in plants [44]. For example, citrus fruits infected with CLas exhibit an accumulation of PLCP protein, the expression of the *PLCP* gene was activated after salicylic acid treatment of citrus [45]. The *cis*-elements located on the gene promoter can regulate gene expression, thereby regulating plants in response to various external stress processes [16]. We found that the *PLCP* gene in woodland strawberry contains 12 *cis*-elements, most of which were hormone-responsive elements, and biotic and abiotic stress-responsive elements, indicating that the *FvPLCP* genes may be involved in plant stress response processes. To verify the response process of *PLCP* genes in biotic stress, we analyzed single-cell transcriptome and RNA-seq data from the early stage of *B. cinerea* infection in woodland strawberry. The results showed that subfamily members of FvPLCPs expressed specificity in different cell types, such as XCP subfamily members mainly expressed in phloem cells and RD21 subfamily members mainly expressed in upper epidermis cells. In addition, different members of the PLCP subfamily were expressed at different treatment periods. For instance, THI subfamily members were mainly expressed at 0.5–3 hpi, while XBCP subfamily members were mainly expressed at 7–12 hpi. We screened hormone-related differentially expressed genes (DEGs), mainly including SA, JA and ET-related genes, and found that genes of different subfamilies of *FvPLCP* genes were highly correlated with hormone-related DEGs. RT-qPCR results showed that hormone-related DEGs had the same expression pattern as *FvPLCP* genes. Transient expression experiments in strawberry fruit showed that *FvSAG12* and *FaDIR6* genes could respond to the process of strawberry defense against *B. cinerea*. These results indicated that the *FvPCLP* genes may be involved in the strawberry resistance to *B. cinerea* by participating in the SA, JA and ET hormone signaling pathways. These results not only laid the foundation for further research on the function of *FvPLCP* genes, but also had great significance for agricultural production and genetic improvement of strawberry stress resistance.

Fundings

This research was supported in part by the Priority Academic Program Development of Jiangsu Higher Education Institutions and the State Key Laboratory of Crop Genetics and Germplasm Enhancement (ZW201813), Hainan Provincial Natural Science Foundation Youth Fund (No.325QN455) and Central Public-interest Scientific Institution Basal Research Fund for Chinese Academy of Tropical Agricultural Sciences (No.1630032024023).

Abbreviations

PLCP	Papain-like cysteine protease
Ka	Nonsynonymous substitution
Ks	Synonymous substitution
PCC	Pearson correlation coefficient
GRAVY	Grand average of hydropathicity
Hpi	Hour post infection
DEGs	Differentially expressed genes

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12870-025-07291-2.

Supplementary Material 1

Supplementary Material 2

Supplementary Material 3

Supplementary Material 4

Supplementary Material 5

Acknowledgements
Not applicable.

Authors' contributions
YBB and ZMC conceived this research. YBB, HBW, and MW analyzed the data. YBB wrote the manuscript. ZMC modified the manuscript. The authors read and approved the final manuscript.

Data availability
The plant materials of woodland strawberry 'Hawaii 4' used in this experiment were all preserved in the laboratory of fruit biosystems at Nanjing Agricultural University (Nanjing, Jiangsu). Full-length RNA-seq reads can be retrieved from NCBI SRA database under project accession PRJNA907512 and BIG database under project accession CRA004848. The datasets supporting the conclusions of this article are included in the article and its additional files. The datasets generated and analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate
The experimental research on plants performed in this study complies with relevant institutional, national, and international guidelines.

Consent for publication
Not applicable.

Competing of interests
The authors declare no competing interests.

Received: 1 June 2025 / Accepted: 22 August 2025
Published online: 02 October 2025

References

1. Misas-Villamil JC, van der Hoorn RA, Doehlemann G. Papain-like cysteine proteases as hubs in plant immunity. New Phytol. 2016;212(4):902–7.
2. Ziemann S, van der Linde K, Lahrmann U, Acar B, Kaschani F, Colby T, et al. An apoplastic peptide activates salicylic acid signalling in maize. Nat Plants. 2018;4(3):172–80.
3. van der Linde K, Hemetsberger C, Kastner C, Kaschani F, van der Hoorn RA, Kumlehn J, et al. A maize cystatin suppresses host immunity by inhibiting apoplastic cysteine proteases. Plant Cell. 2012;24(3):1285–300.
4. Paulus JK, Kourelis J, Ramasubramanian S, Homma F, Godson A, Hörger AC, et al. Extracellular proteolytic cascade in tomato activates immune protease Rcr3. Proc Natl Acad Sci U S A. 2020;117(29):17409–17.
5. Christensen SA, Huffaker A, Kaplan F, Sims J, Ziemann S, Doehlemann G, et al. Maize death acids, 9-lipoxygenase-derived cyclopente(a)nones, display activity as cytotoxic phytoalexins and transcriptional mediators. Proc Natl Acad Sci U S A. 2015;112(36):11407–12.
6. Zou Z, Huang Q, Xie G, Yang L. Genome-wide comparative analysis of papain-like cysteine protease family genes in castor bean and physic nut. Sci Rep. 2018;8(1):331.
7. Bernoux M, Timmers T, Jauneau A, Brière C, de Wit PJ, Marco Y, et al. RD19, an Arabidopsis cysteine protease required for RRS1-R-mediated resistance, is relocalized to the nucleus by the Ralstonia solanacearum PopP2 effector. Plant Cell. 2008;20(8):2252–64.
8. Rawlings ND, Barrett AJ, Bateman A. MEROPS: the peptidase database. Nucleic Acids Res. 2010;38(Database issue):D227-233.
9. Turk V, Turk B, Turk D. Lysosomal cysteine proteases: facts and opportunities. EMBO J. 2001;20(17):4629–33.
10. Richau KH, Kaschani F, Verdoes M, Pansuriya TC, Niessen S, Stüber K, et al. Subclassification and biochemical analysis of plant papain-like cysteine proteases displays subfamily-specific characteristics. Plant Physiol. 2012;158(4):1583–99.
11. Ahmad R, Zuily-Fodil Y, Passaquet C, Bethenod O, Roche R, Repellin A. Identification and characterization of MOR-CP, a cysteine protease induced by ozone and developmental senescence in maize (*Zea mays* L.) leaves. Chemosphere. 2014;108:245–50.
12. Lamesch P, Berardini TZ, Li D, Swarbreck D, Wilks C, Sasidharan R, Muller R, Dreher K, Alexander DL, Garcia-Hernandez M et al: The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. Nucleic Acids Res 2012, 40(Database issue):D1202–1210.
13. Niño MC, Kang KK, Cho YG. Genome-wide transcriptional response of papain-like cysteine protease-mediated resistance against *Xanthomonas oryzae* pv. *oryzae* in rice. Plant Cell Rep. 2020;39(4):457–72.
14. Kang J, Gong P, Ge M, Sadeghnezhad E, Liu Z, Zhang M, et al. The PLCP gene family of grapevine (*Vitis vinifera* L.): characterization and differential expression in response to *Plasmopara viticola*. BMC Plant Biol. 2021;21(1):499.
15. Zhai Y, Cui Y, Song M, Vainstein A, Chen S, Ma H. Papain-like cysteine protease gene family in fig (*Ficus carica* L.): genome-wide analysis and expression patterns. Front Plant Sci. 2021;12:681801.
16. Zhang Xu. Z, Sun H, Sun L, Shaban M, Yang X, Zhu L: Genome-Wide Identification of Papain-Like Cysteine Proteases in Gossypium hirsutum and Functional Characterization in Response to Verticillium dahliae. Front Plant Sci. 2019;10:134.
17. Shulaev V, Sargent DJ, Crowhurst RN, Mockler TC, Folkerts O, Delcher AL, et al. The genome of woodland strawberry (*Fragaria vesca*). Nat Genet. 2011;43(2):109–16.
18. Potter SC, Luciani A, Eddy SR, Park Y, Lopez R, Finn RD. HMMER web server: 2018 update. Nucleic Acids Res. 2018;46(W1):W200-w204.
19. Letunic I, Bork P. 20 years of the SMART protein domain annotation resource. Nucleic Acids Res. 2018;46(D1):D493-d496.
20. Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. Iq-tree: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 2015;32(1):268–74.
21. Subramanian B, Gao S, Lercher MJ, Hu S, Chen WH. Evolview v3: a webserver for visualization, annotation, and management of phylogenetic trees. Nucleic Acids Res. 2019;47(W1):W270-w275.
22. Bailey TL, Johnson J, Grant CE, Noble WS. The MEME suite. Nucleic Acids Res. 2015;43(W1):W39-49.
23. Hu B, Jin J, Guo AY, Zhang H, Luo J, Gao G. GSDS 2.0: an upgraded gene feature visualization server. Bioinformatics. 2015;31(8):1296–7.
24. Voorrips RE. Mapchart: software for the graphical presentation of linkage maps and QTLs. J Hered. 2002;93(1):77–8.
25. Zhu K, Fan P, Liu H, Zhao J, Tan P, Mo Z, et al. Pecan kinome: classification and expression analysis of all protein kinases in *Carya illinoinensis*. For Res. 2021;1(1):1–12.
26. Yu CS, Cheng CW, Su WC, Chang KC, Huang SW, Hwang JK, et al. CELLO2GO: a web server for protein subCELlular LOcalization prediction with functional gene ontology annotation. PLoS ONE. 2014;9(6):e99368.
27. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003;13(11):2498–504.
28. Bai Y, Liu H, Lyu H, Su L, Xiong J, Cheng ZM. Development of a single-cell atlas for woodland strawberry (*Fragaria vesca*) leaves during early *Botrytis cinerea* infection using single cell RNA-seq. Hortic Res. 2022;9:uhab055.
29. Bai Y, Wang H, Zhu K, Cheng ZM. The dynamic arms race during the early invasion of woodland strawberry by *Botrytis cinerea* revealed by dual dense high-resolution RNA-seq analyses. Hortic Res. 2023;10(12):uhad225.
30. Livak K, Schmittgen T. Analysis of relative gene expression data using real-time quantitative PCR and the 2-(-Delta Delta C(T)) method. Methods. 2001;25(4):402–8.
31. van der Hoorn RA. Plant proteases: from phenotypes to molecular mechanisms. Annu Rev Plant Biol. 2008;59:191–223.
32. Lohman KN, Gan S, John MC, Amasino RM. Molecular analysis of natural leaf senescence in *Arabidopsis thaliana*. Physiol Plant. 1994;92(2):322–8.
33. Yuan S, Ke D, Li R, Li X, Wang L, Chen H, et al. Genome-wide survey of soybean papain-like cysteine proteases and their expression analysis in root nodule symbiosis. BMC Plant Biol. 2020;20(1):517.
34. Liu J, Sharma A, Niewiara MJ, Singh R, Ming R, Yu Q. Papain-like cysteine proteases in *Carica papaya*: lineage-specific gene duplication and expansion. BMC Genomics. 2018;19(1):26.
35. Zou Z, Xie G, Yang L. Papain-like cysteine protease encoding genes in rubber (*Hevea brasiliensis*): comparative genomics, phylogenetic, and transcriptional profiling analysis. Planta. 2017;246(5):999–1018.
36. Zhu K, Fan P, Liu H, Tan P, Ma W, Mo Z, et al. Insight into the CBL and CIPK gene families in pecan (*Carya illinoinensis*): identification, evolution and expression patterns in drought response. BMC Plant Biol. 2022;22(1):221.
37. Murat F, Van de Peer Y, Salse J. Decoding plant and animal genome plasticity from differential paleo-evolutionary patterns and processes. Genome Biol Evol. 2012;4(9):917–28.

38. Panchy N, Lehti-Shiu M, Shiu SH. Evolution of gene duplication in plants. Plant Physiol. 2016;171(4):2294–316.
39. Velasco R, Zharkikh A, Troggio M, Cartwright DA, Cestaro A, Pruss D, et al. A high quality draft consensus sequence of the genome of a heterozygous grapevine variety. PLoS ONE. 2007;2(12):e1326.
40. Schlueter JA, Dixon P, Granger C, Grant D, Clark L, Doyle JJ, et al. Mining EST databases to resolve evolutionary events in major crop species. Genome. 2004;47(5):868–76.
41. Schmutz J, Cannon SB, Schlueter J, Ma J, Mitros T, Nelson W, et al. Genome sequence of the palaeopolyploid soybean. Nature. 2010;463(7278):178–83.
42. Zhang LJ, Zhao XQ, Wang J, Wong GK, Yu J. Kaks_calculator: calculating ka and ks through model selection and model averaging. Genomics Proteomics Bioinformatics. 2006;4(4):259–63.
43. Hou J, Wei S, Pan H, Zhuge Q, Yin T. Uneven selection pressure accelerating divergence of *Populus* and *Salix*. Hortic Res. 2019;6:37.
44. Zhao P, Zhou XM, Zhang LY, Wang W, Ma LG, Yang LB, et al. A bipartite molecular module controls cell death activation in the basal cell lineage of plant embryos. PLoS Biol. 2013;11(9):e1001655.
45. Clark K, Franco JY, Schwizer S, Pang Z, Hawara E, Liebrand TWH, et al. An effector from the Huanglongbing-associated pathogen targets citrus proteases. Nat Commun. 2018;9(1):1718.

Publisher's Note