Phytocytokine genes newly discovered in Malus domestica and their regulation in response to Erwinia amylovora and acibenzolar-S-methyl

Marie-charlotte Guillou, Matthieu Gaucher, Emilie Vergne, Jean-Pierre Renou, Marie-Noëlle Brisset, Sébastien Aubourg

Received: 8 July 2024 | Accepted: 6 November 2024

DOI: 10.1002/tpg2.20540

The Plant Genome

ORIGINAL ARTICLE

Phytocytokine genes newly discovered in *Malus domestica* and their regulation in response to *Erwinia amylovora* and acibenzolar-S-methyl

Marie-Charlotte Guillou | **Matthieu Gaucher** | **Emilie Vergne** | **Jean-Pierre Renou** | **Marie-Noëlle Brisset** | **Sébastien Aubourg**

Institut Agro, Université Angers, INRAE, IRHS, SFR QUASAV, Angers, France

Correspondence
Sébastien Aubourg, Institut Agro, Université Angers, INRAE, IRHS, SFR QUASAV, F-49000, Angers, France.
Email: sebastien.aubourg@inrae.fr

Present address
Marie-Charlotte Guillou, Department of Plant Physiology and Biochemistry, University of Hohenheim, 70593 Stuttgart, Germany.

Assigned to Associate Editor Awais Khan.

Funding information
Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement; Université Angers

Abstract

Phytocytokines belong to a category of small secreted peptides with signaling functions that play pivotal roles in diverse plant physiological processes. However, due to low levels of sequence conservation across plant species and poorly understood biological functions, the accurate detection and annotation of corresponding genes is challenging. The availability of a high-quality apple (*Malus domestica*) genome has enabled the exploration of five phytocytokine gene families, selected on the basis of their altered expression profiles in response to biotic stresses. These include phytosulfokine, inflorescence deficient in abscission/-like, pathogen-associated molecular pattern induced secreted peptide, plant peptide containing sulfated tyrosine, and C-terminally encoded peptide. The genes encoding the precursors of these five families of signaling peptides were identified using a customized bioinformatics protocol combining genome mining, homology searches, and peptide motif detection. Transcriptomic analyses showed that these peptides were deregulated in response to *Erwinia amylovora*, the causal agent of fire blight in pome fruit trees, and in response to a chemical elicitor (acibenzolar-S-methyl). Finally, gene family evolution and the orthology relationships with *Arabidopsis thaliana* homologs were investigated.

Plain Language Summary

Orchards are suffering damages from the pathogenic bacterium *Erwinia amylovora* responsible for fire blight. Phytocytokines, small peptides secreted by plant cells, play

Abbreviations: ASM, acibenzolar-S-methyl; CDS, coding sequence; CEP, C-terminally encoded peptide; CRP, cysteine-rich peptide; *Ea*, *Erwinia amylovora*; IDA/IDL, inflorescence deficient in abscission/-like; PAMP, pathogen-associated molecular pattern; PIP, PAMP-induced secreted peptide; PSK, phytosulfokine; PSY, plant peptide containing sulfated tyrosine; PTI, PAMP-triggered immunity; PTMP, post-translationally modified peptide; TAG, tandemly arrayed gene cluster.

Marie-Charlotte Guillou and Matthieu Gaucher have contributed equally to this work.

a role in defense responses. Characterizing them is crucial for understanding how plants respond to pathogen attacks. Our aim is to identify genes encoding phytocytokines in *Malus domestica* potentially involved in defense mechanisms in response to Erwinia infection. We measured gene activity and used bioinformatics methods to detect apple genes similar to phytocytokines from other plants. We identified and exhaustively annotated five phytocytokine families comprising 75 genes, with 30 showing deregulated activity by Erwinia. This work overcomes the limitations of automatic annotation software, revealing the diversity, genomic organization, and evolution of these gene families and suggests candidates for the selection of apple trees more tolerant to pathogens and for the development of new biocontrol strategies.

1 | INTRODUCTION

In response to recurrent pathogen challenges, plants have developed intricate systems characterized by sophisticated perception and defense mechanisms. These plant defense responses are driven by complex regulatory networks involving numerous proteins and hormones, leading to massive changes in gene expression (Buscaill & Rivas, 2014). Small peptides secreted by plant cells play a crucial role in plant growth and development and are pivotal actors in the regulation of stress tolerance (Gust et al., 2017; Marmiroli & Maestri, 2014; Murphy et al., 2012). Functional peptides are generated from prepropeptides through several maturation steps and can be classified into two major groups based on structural features (Matsubayashi, 2011; Murphy et al., 2012): The first group comprises the small post-translationally modified peptides (PTMPs) that usually contain up to 20 amino acids (aa) and are characterized by enzymatic post-translational modifications. The second group includes the cysteine-rich peptides (CRPs), larger peptides (ranging from 40 to 160 aa) characterized by the presence of cysteine residues involved in disulfide bond formation. Plasma membrane-localized receptors, such as receptor-like kinases (RLKs), that contain a variable number of extracellular leucine-rich repeats, can perceive PTMPs and, more rarely, CRPs (Tavormina et al., 2015). The triggered underlying signaling pathways include, among others, PAMPs (pathogen-associated molecular patterns)-triggered immunity (PTI), the first line of inducible defense in an infection context (Couto & Zipfel, 2016; De Smet et al., 2009). These peptides with signaling functions are also named peptidic hormones or phytocytokines by analogy with animal cytokines (Gust et al., 2017).

Phytocytokines play diverse roles in plant immunity through several families identified across various species (Hou et al., 2021). Among PTMPs, phytosulfokines (PSKs) are short five-aa peptides present in higher plants with two sulfated tyrosine residues. They were initially identified as cell proliferation inducers (Matsubayashi & Sakagami, 1996) and are now also known to be involved in the regulation of PTI (Amano et al., 2007; Sauter, 2015). The PAMP-induced secreted peptide (PIP) family, with PIP1 and PIP2 being well studied, amplifies plant immunity through the RLK7 receptor (Hou et al., 2014). The plant peptide containing sulfated tyrosine (PSY) family has been identified in diverse plant species, including *Oryza sativa*, *Musa acuminata*, *Solanum lycopersicum*, and *Triticum aestivum* (Pruitt et al., 2017). Among them, PSY1 has been shown to suppress PTI responses, promoting *Arabidopsis thaliana* resistance to *Alternaria brassicicola* but susceptibility to *Pseudomonas syringae* pv. *tomato* and *Fusarium oxysporum* (Mosher et al., 2013; Shen & Diener, 2013). Members of C-terminally encoded peptide (CEP) and inflorescence deficient in abscission/-like (IDA/IDL) families have known roles in the regulation of plant lateral root growth and the abscission of floral organs after pollination, respectively (Butenko et al., 2003; Delay et al., 2013). However, recent studies also provide evidence that CEPs and their receptors promote immunity in an N status-dependent manner and that IDA peptides modulate plant immunity in tissues undergoing cell separation (Lalun et al., 2023; Rzemieniewski et al., 2022).

The sequence conservation of phytocytokines vary widely across plant species, and their functional aspects are often understudied. Some phytocytokines, including RALFs (rapid alkalinization factor) and plant elicitor peptide, are widely distributed among various plant species (Gust et al., 2017) but others are exclusive to specific plant families, such as systemin in *Solanaceae* (Pearce et al., 1991) or SERINE RICH ENDOGENOUS PEPTIDES (SCOOPs) in *Brassicaceae* (Gully et al., 2019), or even to one species such as Zip1 in *Zea mays* and its wild ancestor (Depotter et al., 2022). The sequence diversification of phytocytokines may result from an ongoing evolutionary competition, wherein plants must constantly adapt to changing environments by selecting new versions of these signaling molecules (Furumizu & Shinohara, 2024). Whether signaling peptides serve comparable roles in dicots and monocots remains largely unknown.

So far in *Malus domestica* (apple), only CLAVATA3 (CLV3)/embryo surrounding region (CLE) and CEP phytocytokines have been identified (Li et al., 2018; Yu et al., 2019; Zhang et al., 2022). In the present work, we conducted a genome-wide investigation to identify apple homologs of five major PTMP families known for their signaling functions in other species and especially in *Arabidopsis*: PSK, IDA/IDL, PIP, PSY, and CEP (see Tavormina et al. [2015] for review). The selection of these PTMP families was also motivated by the transcriptional induction of some of their members in response to the necrogenic bacteria *Erwinia amylovora* (*Ea*), the causal agent of fire blight (for a review on the disease, see Vanneste [2000]) and therefore their possible role in defense responses. We analyzed their chromosome location, gene structures, and conserved motifs. In addition to the study of their expression in *Ea* infected leaves, we investigated their responses to the chemical elicitor acibenzolar-S-methyl (ASM, also called BTH). This functional salicylate analog is known to significantly modify fire blight susceptibility in controlled conditions as well as in orchard, and to induce a large reprogramming of apple gene expression (Aćimović et al., 2015; Brisset et al., 2000; Dugé de Bernonville et al., 2014; Warneys et al., 2018; Yuan et al., 2023). Lastly, we explored the orthology relationship between *Arabidopsis* and *Malus* phytocytokine genes, providing a basis for future research into the functional roles of phytocytokines in apple immunity.

Core Ideas

- Five signaling peptide families in apple were identified.
- Transcriptomic revealed their deregulation in response to the causal agent of fire blight and a chemical elicitor.
- Phylogeny and transcriptomic profiling were used to explore orthology relationships between apple and *Arabidopsis*.

2 | MATERIALS AND METHODS

2.1 | Identification of phytocytokine families in the apple genome

Starting from full-length precursor sequences of IDA, PIP, PSK, PSY, and CEP families described in *Arabidopsis* (Amano et al., 2007; Butenko et al., 2003; Hou et al., 2014; Ohyama et al., 2008; Stührwohldt et al., 2021), hidden Markov models (HMM) were defined after multiple alignment (MUSCLE v2.1, Madeira et al., 2022) with HMMbuild and applied with HMMsearch (HMMER package v3.4; Eddy, 2011) to screen the GDDH13 v1.1 proteome (Daccord et al., 2017; https://iris.angers.inra.fr/gddh13/). In parallel, *Arabidopsis* and *Malus* candidate proteins were used as queries for TBLASTN searches against GDDH13 genome in an iterative way as described in Van Canh and Aubourg (2024). All hits obtained were manually checked to determine whether they corresponded to new candidate genes (previously annotated or not). Hits with insignificant similarities were discarded. Subsequently, the candidate proteins from each family were further screened based on criteria including protein length (maximum 200), presence of an N-terminal signal peptide (SignalP v6.0; Nielsen, 2017) necessary for secretion into apoplast, and conservation of the C-terminal motif characteristic of the respective phytocytokine family. Curated results are available in a specific track of GDDH13 JBrowse (https://iris.angers.inra.fr/gddh13/). The schematic representation of *Md*PSK, *Md*IDA, *Md*PIP, *Md*PSY, and *Md*CEP chromosomal positioning was drawn using MG2C (Chao et al., 2021).

2.2 | Sequence feature analyses

The conserved protein domain of *Md*PSK, *Md*IDA, *Md*PIP, *Md*PSY, and *Md*CEP proteins were investigated by multiple alignment analysis (MUSCLE v2.1). The signal peptides were predicted using the SignalP v6.0 software and the conserved motifs were detected using HMMsearch (HMMER package v3.4; Eddy, 2011) with cut-off value of 1e-3 and manual check of results. The motifs LOGO were displayed by WebLogo (v2.8.2; Crooks et al., 2004) using amino acid Clustal color code. The physical and chemical properties of each peptide were analyzed with ProtParam tool from Expasy (v3.0; Duvaud et al., 2021).

2.3 | Phylogenetic analyses

The phylogenetic analyses were performed with the protein sequences of *Md*PSK, *Md*IDA, *Md*PIP, *Md*PSY, and *Md*CEP and their homologous sequences from *Arabidopsis*. Full-length protein sequences were aligned by MUSCLE v2.1 (Madeira et al., 2022) using standard parameters. The phylogenetic trees were generated via the neighbor-joining method with bootstrap values calculated from 1000 replicates with ClustalX (v2.0) and edited with ITol software (v6.9).

2.4 | Transcriptomic analyses

2.4.1 | Biological material

The experiments were carried out on young trees of MM106 and Evereste chosen for their difference in *Ea* resistance (Durel et al., 2009; Venisse et al., 2002), and both grafted

onto MM106 rootstock, or on apple seedlings from open-pollinated cv. Golden Delicious collected at the experimental unit of INRAE Angers (France), and chosen for their overall susceptibility to fire blight (Chavonet et al., 2022; Dugé de Bernonville et al., 2014). Plants were grown in individual pots (0.23 L for seedlings and 1 L for grafted plants, Soparco) under greenhouse conditions (natural photoperiod supplemented with artificial light if needed, 17°C night and 20–23°C day) for 3–4 weeks before use.

Grafted plants were inoculated by *Ea* or mock through vacuum-infiltration as described in Dugé de Bernonville et al. (2012). The virulent CFBP1430 strain of *Ea* was grown at 26°C for 12 h on solid King's B medium (King et al., 1954) prior to each experiment, and bacterial suspensions were prepared at 10^7 CFU/mL in sterile water containing 0.01% of Silwet to improve infiltration efficacy. Three youngest fully expanded leaves (one leaf per plant from three different plants per modality of infiltration —*Ea* or mock) were collected at 6 and 24 hpi, then pooled, and immediately frozen in liquid nitrogen and stored at −80°C until subsequent RNA extraction. Three independent experiments were performed as biological replicates.

Apple seedlings were treated with Bion 50WG (50% ASM, 0.2 mg a.i./L of water) or water, applied to runoff on the whole plants using an Aeryo-1.4 spray gun (Deltalyo). Five youngest fully expanded leaves (one leaf per plant from five different plants per modality of treatment—ASM or water) were collected at 3 days post-treatment, then pooled, and immediately frozen in liquid nitrogen and stored at −80°C until RNA extraction. Three independent experiments were performed as biological replicates.

2.4.2 | RNA extraction and reverse transcription

Total RNA was isolated using the Nucleospin RNA Plant Kit (Macherey-Nagel) with the extraction buffer supplemented with 2% PVP-40. RNA integrity was confirmed using the RNA 6000 nano kit and a Bioanalyzer (Agilent). The cDNA synthesis was carried out using the M-MLV (Moloney Murine Leukemia Virus Reverse Transcriptase) reverse transcriptase (Promega) and the absence of DNA was checked by PCR using EF-α primers flanking an intron, as described in Gilliland et al. (1990).

2.4.3 | RNA-Seq analyses

Libraries preparation and sequencing were performed using the Illumina NovaSeq 6000 S4 PE100 reads technology (Génome Québec). From 21.5 to 60 million read pairs were generated per sample. Reads were mapped using Salmon software (v1.10.2; Patro et al., 2017) onto the reference transcriptional units from GDDH13 v1.1 improved with newly identified phytocytokine genes (see above). The mapping rates were between 77.5% and 84%. The original sequencing datasets have been deposited in the Gene Expression Omnibus (GEO) with the accession numbers GSE262535 (apple/*Ea*) and GSE262538 (apple/ASM). Differentially expressed genes (DEGs) were defined with DESeq2, including a Benjamini–Hochberg procedure to control the false discovery rate through the AnaDiff v4.3 pipeline (Pelletier, 2022). The criteria for identifying DEGs were as followed: log2 fold change $|\log FC| \geq 0.5$ and p-value/BH ≤ 0.05. Results are means of the three independent experiments (one biological replicate per experiment/modality/timepoint). For the comparisons with *Arabidopsis*, we used available RNA-seq data extracted from Genevestigator (Zimmermann et al., 2005) obtained from *Arabidopsis* leaves challenged with *Botrytis cinerea*, *Phytophthora cinnamomi*, *P. syringae* pv. *Maculicola*, and Flg22 elicitor (GEO/SRA references are GSE66290, SRP253869, ERS1023182 and GSE63603 respectively).

2.4.4 | Reverse transcription quantitative real-time PCR (RT-qPCR) analyses

The primers used for quantitative real-time PCR measurements are listed in Table S1. PCR preparation and calculation of relative changes in gene expression (log2 ratios) were performed as described in Gaucher, Heintz et al. (2022). Results are means of the three independent experiments (one biological replicate per experiment/modality/timepoint) and three technical replicates were run for each biological replicate. Significant differences between inoculated or treated versus respective controls were determined by using the nonparametric rank-based statistical tests Wilcoxon–Mann–Whitney ($p = 0.05$).

3 | RESULTS

3.1 | Identification and annotation of five phytocytokines families in apple: *MdPSK*, *MdIDA*, *MdPIP*, *MdPSY*, and *MdCEP*

Genome-wide similarity search in the apple GDDH13 v1.1 genome using the known *Arabidopsis* PSK, IDA/IDL, PIP, PSY, and CEP proteins allowed the identification of a total of 75 genes: 12 *MdPSK*, 13 *MdIDA/IDL*, 13 *MdPIP*, 18 *MdPSY*, and 19 *MdCEP*, leading to the establishment of a comprehensive structural annotation for these five gene families in apple (Figure 1, Table 1, and Table S2).

Only 37 PTMP genes out of 75 were detected by automatic annotation within the apple genome, with large differences

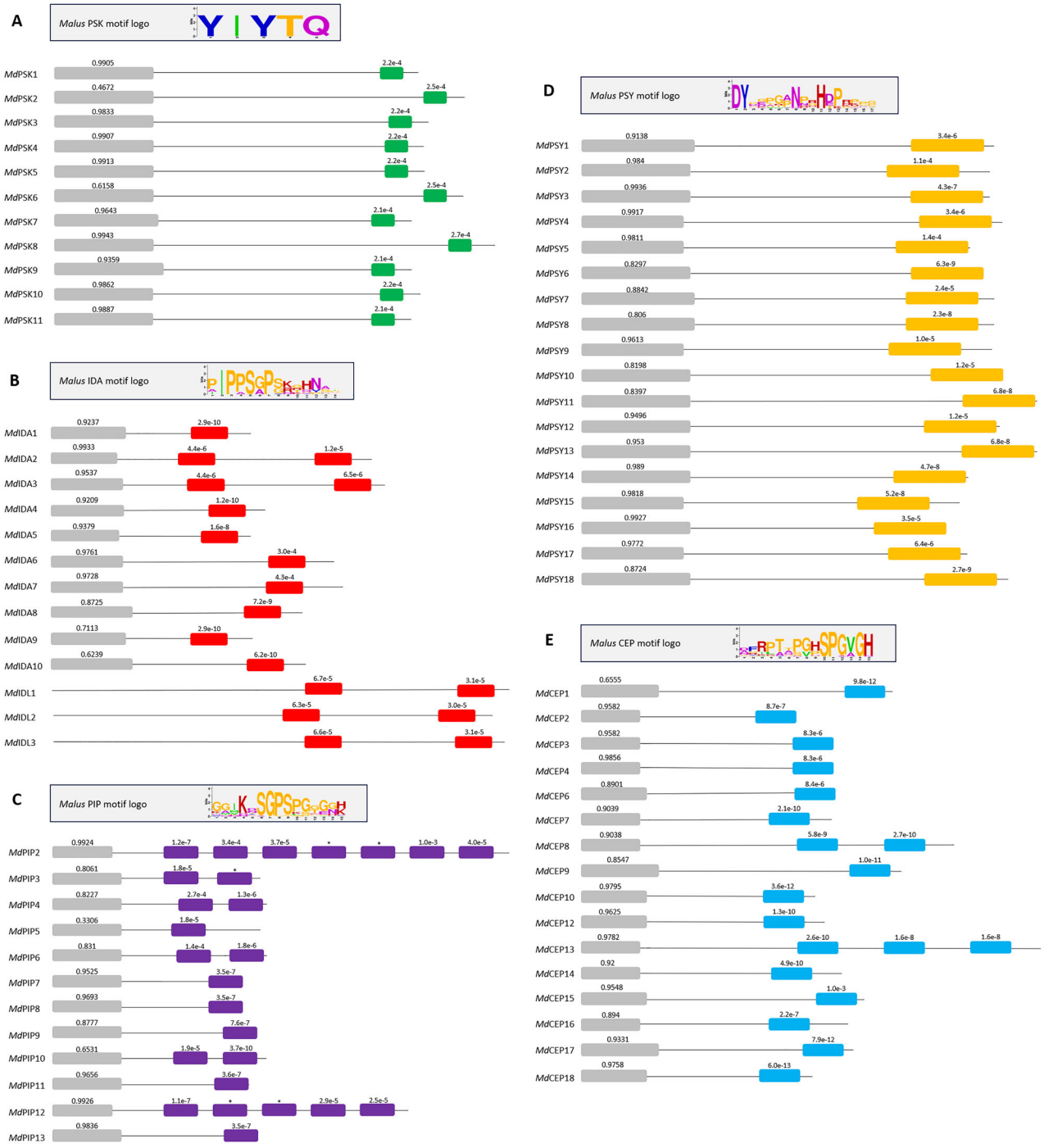

FIGURE 1 Conserved motifs identified in five *Malus* phytocytokine families. (A) *Md*PSK, (B) *Md*IDA, (C) *Md*PIP, (D) *Md*PSY, and (E) *Md*CEP family proteins. The SignalP scores for the presence of signal peptide (gray boxes) and motif location corresponding to putative mature secreted peptides (green, red, purple, yellow, and blue boxes) are shown with respective *e*-values from HMMsearch. The asterisks correspond to motifs with e-values above the cut-off but considered as true positive after manual checking. Pseudogenes, lacking coherent coding sequences, are not presented here. IDA, inflorescence deficient in abscission; CEP, C-terminally encoded peptide; PIP, PAMP-induced secreted peptide; PSK, phytosulfokine; PSY, plant peptide containing sulfated tyrosine.

TABLE 1 Characteristics of *MdPSK*, *MdIDA*, *MdPIP*, *MdPSY* and *MdCEP* gene families in *Malus*.

Gene name	Gene ID	Exon number	Protein length (aa)	Motif number	Isoelectric point	Molecular weight (kDa)	Remarks
MdPSK1	MD02G1189200	2	84	1	4.85	9.34	
MdPSK2	MD03G1134700	3	94	1	4.61	10.45	
MdPSK3	MD05G1105500	2	86	1	4.97	9.35	
MdPSK4	MD08G1079100	2	85	1	7.84	9.51	
MdPSK5	MD10G1110000	2	85	1	4.68	9.26	
MdPSK6	MD11G1157500	3	94	1	5.87	10.60	
MdPSK7	MD12G1018200	2	82	1	6.71	9.58	
MdPSK8	MD13G1053700	2	101	1	4.48	10.78	
MdPSK9	MD14G1016200	2	82	1	5.19	9.43	
MdPSK10	MD15G1066300	2	84	1	5.82	9.26	
MdPSK11	MD15G1300000	2	82	1	4.85	8.93	
MdPSK12	MD16G1052250	nd	nd	nd	nd	nd	Pseudogene
MdIDA1	MD01G1010200	1	82	1	12.08	9.26	
MdIDA2	MD04G1118430	1	131	2	9.41	13.88	
MdIDA3	MD04G1118450	1	137	2	9.35	14.60	
MdIDA4	MD05G1020400	1	87	1	11.67	9.48	
MdIDA5	MD10G1021350	1	81	1	9.45	8.76	
MdIDA6	MD12G1133850	1	116	1	7.28	12.25	
MdIDA7	MD12G1134050	1	120	1	8.83	12.41	
MdIDA8	MD13G1048650	1	103	1	8.62	20.15	
MdIDA9	MD15G1339250	1	82	1	9.34	19.25	
MdIDA10	MD16G1049950	1	104	1	8.62	20.05	
MdIDL1	MD12G1134080	1	188	2	10.24	11.34	No signal peptide
MdIDL2	MD12G1134310	1	180	2	11.86	9.28	No signal peptide
MdIDL3	MD12G1134390	1	187	2	9.41	11.42	No signal peptide
MdPIP1	MD02G1179250	nd	nd	nd	nd	nd	Pseudogene
MdPIP2	MD02G1179300	1	204	7	9.49	20.565	
MdPIP3	MD05G1090550	1	92	2	10.51	10.03	
MdPIP4	MD05G1090650	1	95	2	10.18	10.02	
MdPIP5	MD05G1090710	1	92	1	9.71	9.91	
MdPIP6	MD05G1090750	1	95	2	10.18	9.88	
MdPIP7	MD08G1091750	1	84	1	5.4	8.85	

(Continues)

TABLE 1 (Continued)

Gene name	Gene ID	Exon number	Protein length (aa)	Motif number	Isoelectric point	Molecular weight (kDa)	Remarks
MdPIP8	MD08G1091770	1	84	1	5.36	8.78	
MdPIP9	MD09G1036050	1	91	1	10.8	10.01	
MdPIP10	MD10G1098120	1	95	2	10.11	10.06	
MdPIP11	MD15G1076000	1	87	1	5.37	8.96	
MdPIP12	MD15G1289400	1	159	5	8.31	16.49	
MdPIP13	MD17G1037650	1	91	1	9.34	9.79	
MdPSY1	MD01G1223300	3	85	1	9.33	9.55	
MdPSY2	MD03G1067000	2	80	1	7.48	8.43	
MdPSY3	MD06G1112150	3	82	1	4.74	9.04	
MdPSY4	MD06G1112200	3	88	1	6.02	9.39	
MdPSY5	MD06G1112300	3	76	1	8.93	8.46	
MdPSY6	MD06G1194400	3	83	1	9.22	9.02	
MdPSY7	MD07G1293950	3	85	1	10.45	9.66	
MdPSY8	MD07G1294000	3	86	1	9.22	9.78	
MdPSY9	MD11G1071800	2	80	1	6.49	8.47	
MdPSY10	MD12G1012300	3	88	1	4.48	9.76	
MdPSY11	MD12G1012400	3	94	1	4.64	10.29	
MdPSY12	MD14G1010100	2	86	1	5.35	9.37	
MdPSY13	MD14G1010200	3	94	1	4.41	10.13	
MdPSY14	MD14G1133550	3	74	1	8.88	8.25	
MdPSY15	MD14G1133600	3	74	1	7.22	8.01	
MdPSY16	MD14G1133700	3	70	1	6.02	7.79	
MdPSY17	MD14G1133800	3	77	1	8.66	8.49	
MdPSY18	MD14G1201500	3	91	1	6.86	10.02	
*MdCEP1**	MD02G1169450	1	119	1	6.99	12.68	
*MdCEP2**	MD07G1232650	1	81	1	8.87	9.18	
*MdCEP3**	MD00G1009200	1	95	1	9.42	10.08	Identical to *MdCEP4*
MdCEP4	MD00G1221100	1	95	1	9.42	10.08	Identical to *MdCEP3*
MdCEP5	MD01G1165350	nd	nd	nd	nd	nd	Pseudogene
*MdCEP6**	MD10G1275500	1	96	1	9.13	10.38	
*MdCEP7**	MD15G1086950	1	96	1	5.6	10.38	
*MdCEP8**	MD15G1281900	1	141	2	7.28	15.10	
*MdCEP9**	MD15G1281950	1	121	1	7.76	13.05	

(Continues)

TABLE 1 (Continued)

Gene name	Gene ID	Exon number	Protein length (aa)	Motif number	Isoelectric point	Molecular weight (kDa)	Remarks
MdCEP10	MD02G1169350	1	88	1	6.65	9.53	
*MdCEP11**	MD08G1106270	nd	nd	nd	nd	nd	Pseudogene
*MdCEP12**	MD06G1011550	1	92	1	10.3	9.75	
MdCEP13	MD02G1169400	1	174	3	5.37	18.23	
MdCEP14	MD08G1106250	1	98	1	5.27	10.48	
MdCEP15	MD11G1264950	1	107	1	6.26	11.54	
MdCEP16	MD15G1086940	1	101	1	5.09	10.62	
MdCEP17	MD15G1281750	1	103	1	8.2	11.35	
MdCEP18	MD15G1281800	1	88	1	7.13	9.59	
MdCEP19	MD16G1079450	nd	nd	nd	nd	nd	Pseudogene

Note: Genes for which ID ends in a multiple of 10 (and not 100) are not present in the automatic GDDH13 v1.1 genome annotation and have been identified as part of this work. The asterisks correspond to *MdCEP* genes previously described by Yu et al. (2019) with former names kept.

between gene families. In the worst cases, only two out of 13 *MdIDA/IDL* genes and three out of 13 PIP genes were detected in the last annotation (v1.1). These results illustrate the necessity of the curation work and the genome-wide search independently of gene prediction software. The new genes have been added to the GDDH13 annotation (with new gene ID inserted in the official nomenclature) and the intron–exon structures of coding sequence (CDS) are available in Table S2. This fine annotation also allowed the identification of genes for which CDS may have been interrupted by frameshifts, stop codon, and/or deletion. We cannot exclude that these pseudogenes in GDDH13 (Table 1) could be functional in other genotypes.

The apple PSK family contains 12 genes encoding precursors ending with the 100% conserved YIYTQ motif (Figure 1A) corresponding to the functional peptide in which tyrosine residues are expected to be sulfated (Matsubayashi, 2011). These genes have been named *MdPSK1* to *MdPSK12*, the last one being a pseudogene. The protein precursors are 87 aa long on average, with a mean isoelectric point of 5.44, indicating that these proteins are mainly composed of acidic aa (Table 1).

We identified in apple 13 single exon genes belonging to the IDA family. They encode proteins containing one or two IDA motifs. Ten genes encode precursors with well predicted N-terminal secretion signal (named *MdIDA1* to *MdIDA10*), but surprisingly, three tandemly arrayed genes encode longer proteins without the secretion signal despite the presence of canonical IDA motifs (Figure 1B). These genes have been considered IDA-like and named *MdIDL1* to *MdIDL3*. We have checked the absence of alternative translation starting site in frame that could lead to shorter or longer proteins carrying a secretion signal. Protein lengths range from 81 to 188 aa and the average isoelectric point around 9.66 suggests that the proteins are highly basic.

The apple PIP family comprises 13 single exon genes encoding proteins containing N-terminal secretion signal followed by one to seven PIP motifs (Figure 1C). These genes have been named *MdPIP1* to *MdPIP13*. Note that *MdPIP1*, with a CDS disrupted by a frameshift, is a pseudogene in GDDH13. This family contains more diverse proteins either mainly acidic or mainly basic, with molecular weights ranging from 8.76 to 20.05 kDa.

The apple PSY family contains 18 homologous genes (*MdPSY1* to *MdPSY18*) encoding proteins with one PSY motif in C-terminal end (Figure 1D). This canonical motif starts with the ubiquitous DY (aspartic acid−Tyrosine [Asp-Tyr]) dipeptide in which the tyrosine residue should be sulfated in the mature peptide (Tost et al., 2021). Protein lengths averaged around 80 aa, either acidic or basic.

After curation, our screening of the apple genome identified 19 genes belonging to the CEP family (Figure 1E). They

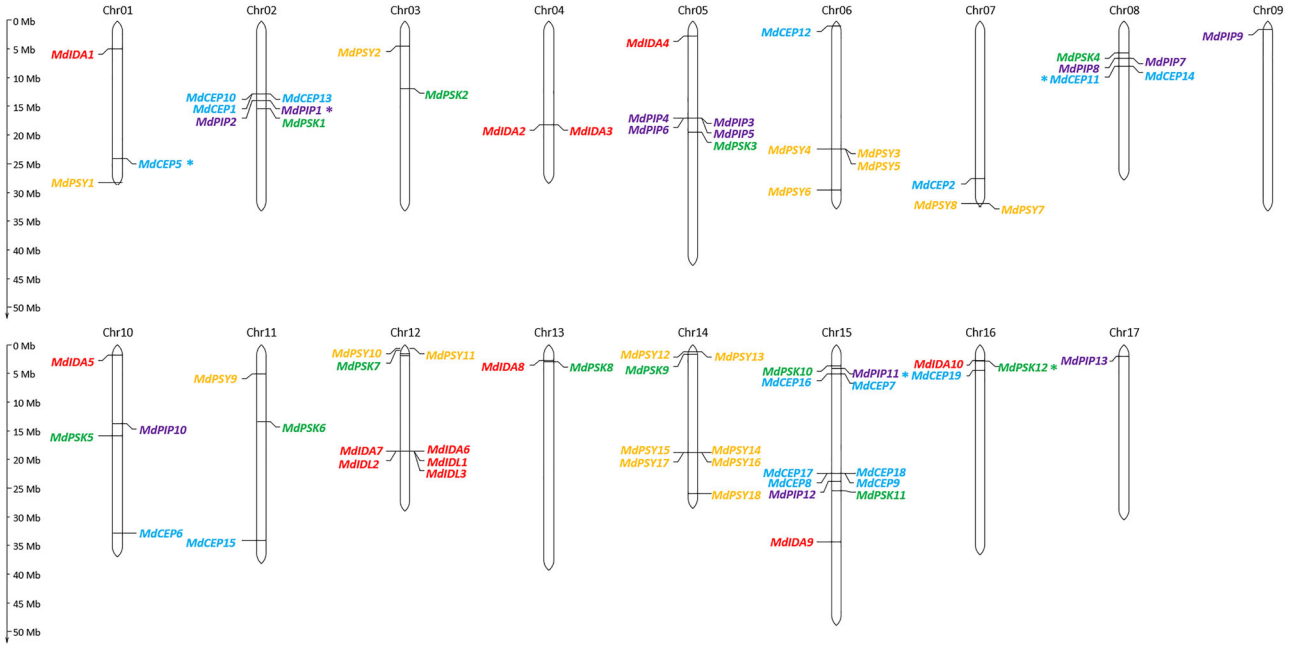

FIGURE 2 Chromosomal organization of the five phytocytokine families (GDDH13 genome). *MdCEP3* and *MdCEP4* genes located on unmapped contigs (Chr00) are not represented here. The pseudogenes with coding sequence (CDS) disrupted by stop codon, frameshift, and/or deletion are tagged by asterisks.

have been named *MdCEP1* to *MdCEP19*. Unlike previous families, genes have not been named according to their physical chromosomal locations (Table 1). Indeed, this *MdCEP* nomenclature takes into account a previous work (Yu et al., 2019), which had identified nine out of the 19 *MdCEP* genes resulting from our annotation (see Section 4 for details). *MdCEP* genes encode proteins containing one to three CEP motifs, three of them (*MdCEP5*, *MdCEP11*, and *MdCEP19*) being pseudogenes in GDDH13. The family harbors more heterogeneous proteins, with size varying from 81 to 174 acidic as well as basic aa.

The chromosomal location of the phytocytokine gene families (Figure 2) indicates that half of the 75 genes are organized in small tandemly arrayed gene clusters (TAGs), ranging from two to five genes. For instance, *MdPSY* and *MdCEP* genes were found in clusters of four genes, while we found a cluster of five *MdIDA*. Only the *MdPSK* family has all its genes dispersed across the genome.

3.2 | Expression analysis of peptide genes in leaves in response to *Ea* and ASM

To decipher the transcriptional pattern of the genes belonging to the five phytocytokine families, we performed RNA-seq analyses of *Ea*-infected leaves from Evereste (resistant) and MM106 (susceptible) genotypes (Durel et al., 2009) and in ASM-treated leaves from Golden Delicious seedlings (susceptible to fire blight and responsive to ASM; Dugé de Bernonville et al., 2014). Differential expression analyses were performed by comparing infected or treated samples with respective controls (Figure 3).

The high level of induction of a gene encoding a polyphenol oxidase (*MdPPO16*) indicates the success of the infection in both the susceptible and resistant genotypes, as early as 6 hours post inoculation (hpi), in accordance with Gaucher, Righetti et al. (2022). Observation of confluent necrosis in the *Ea*-infiltrated areas at 24 hpi confirmed also the infection success, in accordance with Venisse et al. (2002). Similarly, the strong induction of a gene encoding an agglutinin (*MdAGG10*) in ASM-treated leaves demonstrates that the seedlings had effectively detected the treatment and were responding strongly 3 days after, in accordance with Chavonet et al. (2022).

Among the differentially expressed PTMP genes, 26 were selected for RT-qPCR validation. Their differential expression profiles corresponded between the two methods in 82% of the tested comparisons and besides. Similar profiles were observed for an additional 11.5%, with however a lack of statistical significance in one of the two methods (Figure S1). These RT-qPCR results allowed us to interpret the RNA-seq data with complete confidence.

Overall, *Ea* infection and ASM treatment resulted in the upregulation of *MdPIP* family members, the downregulation of *MdPSY* members, and both up- or downregulation of *MdPSK*, *MdIDA/IDL*, and *MdCEP* members (Figure 3). The majority of the genes of the *MdPIP* family (8/13) displayed notable induction levels as early as 6 hpi by *Ea*,

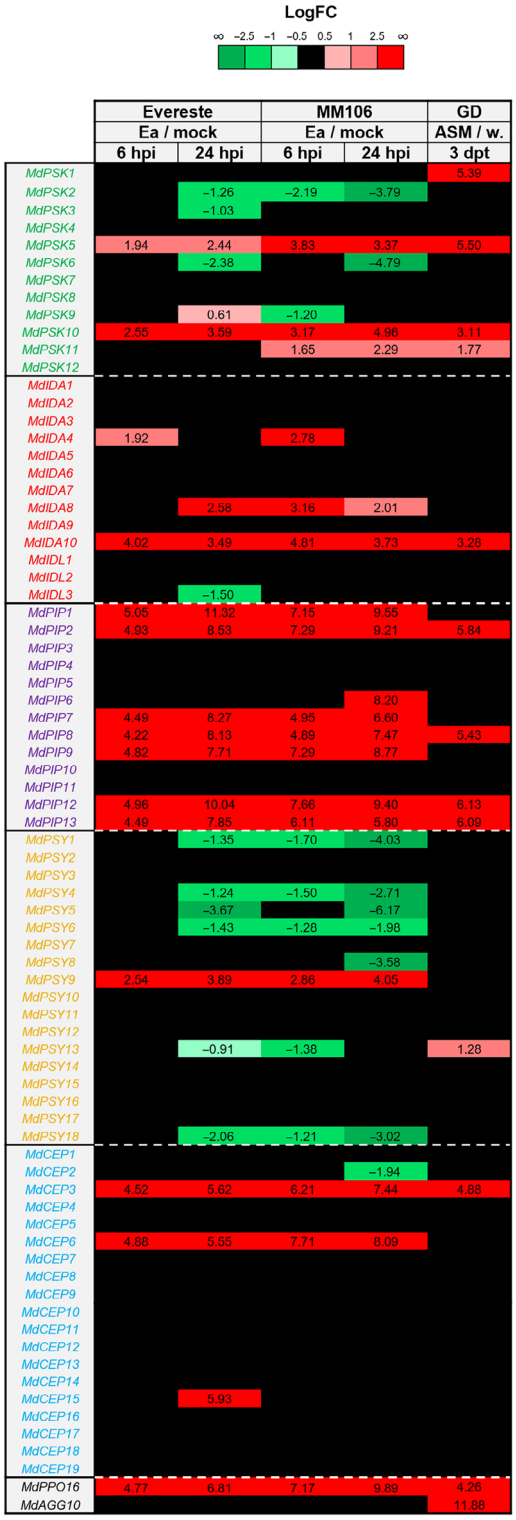

FIGURE 3 Expression analysis of *MdPSK*, *MdIDA*, *MdPIP*, *MdPSY*, and *MdCEP* genes families in apple leaves in response to *Ea* (*Erwinia amylovora*) or ASM (acibenzolar-S-methyl). Heatmap of RNA-seq data representing log2 ratios (|logFC|≥ 0.5 and *p*-value/BH ≤ 0.05) of the expression between Ea- and mock-infiltrated plants (at 6 and 24 hours post inoculation [hpi]) or between ASM- and water (w)-treated plants at 3 days post treatment (dpt). Data were obtained from three biological replicates per modality in three independent

(Continues)

FIGURE 3 (Continued)
experiments ($n = 3$). *MdPPO16* and *MdAGG10* are markers of *Ea*-infection and ASM-treatment respectively. GD, golden delicious seedlings.

irrespective of the genotype (except *MdPIP6*, only induced in MM106 at 24 hpi). Half of them were also induced by ASM. In contrast, *MdPSY* genes were mostly downregulated in response to *Ea* (7/18), particularly at 24 hpi in both genotypes with the exception of *MdPSY9* induced in both genotypes from 6 hpi (Figure 3). *MdPSY13* was the only gene whose expression was specifically modulated (induced) by ASM. Some *MdPSK* genes were either induced (*MdPSK5*, *10* and *11*) or repressed (*MdPSK2* and *6*) by *Ea* in both genotypes depending on the time point, and two others showed differential modulation between genotypes (*MdPSK9*, either induced or repressed and *MdPSK3* only repressed in Evereste at 24 hpi). The two genes (*MdPSK5* and *MdPSK10*) induced by *Ea* in both genotypes were also induced by ASM, which also led to the specific induction of *MdPSK1* (Figure 3). The expression of most genes of the *MdIDA/IDL* and *MdCEP* families was not impacted by the biotic stress or ASM, except *MdIDA4*, *MdIDA8*, *MdIDA10*, *MdCEP3*, and *MdCEP6*, which expressions were induced by *Ea* in both genotypes (globally from 6 hpi) or by ASM for a few of them (*MdIDA10* and *MdCEP3*) and *MdCEP15* specifically induced by *Ea* in Evereste at 24 hpi (Figure 3). Additionally, the RNA-seq data revealed the induction of other genes encoding small secreted peptides of the CRP type (RALFs MD04G1018100 and MD04G1201100, defensins MD08G1027500 and MD15G1119700, thionin MD15G1417600, thaumatin MD04G1064400, CAPEs MD05G1109100, MD05G1108800, and snakin MD17G1039200) known to be involved in plant stress signaling or for their microbiocide activity (Goyal & Mattoo, 2014).

3.3 | Orthology relationships between phytocytokines from *Malus domestica* and *Arabidopsis thaliana*

To evaluate the evolutionary relationships among the *Md*PSK, *Md*IDA, *Md*PIP, *Md*PSY, and *Md*CEP proteins, phylogenetic analyses were performed between the *Arabidopsis* and *Malus* homologs (Figure 4A–E). *Arabidopsis* peptide precursor sequences of the five families of interest have been mined from TAIR10/Araport11 annotations and literature (Amano et al., 2007; Butenko et al., 2003; Hou et al., 2014; Ohyama et al., 2008; Stührwohldt et al., 2021). For a full screening of the *Arabidopsis* genome, we used these sequences to search for possible unannotated homologs. This enabled us to

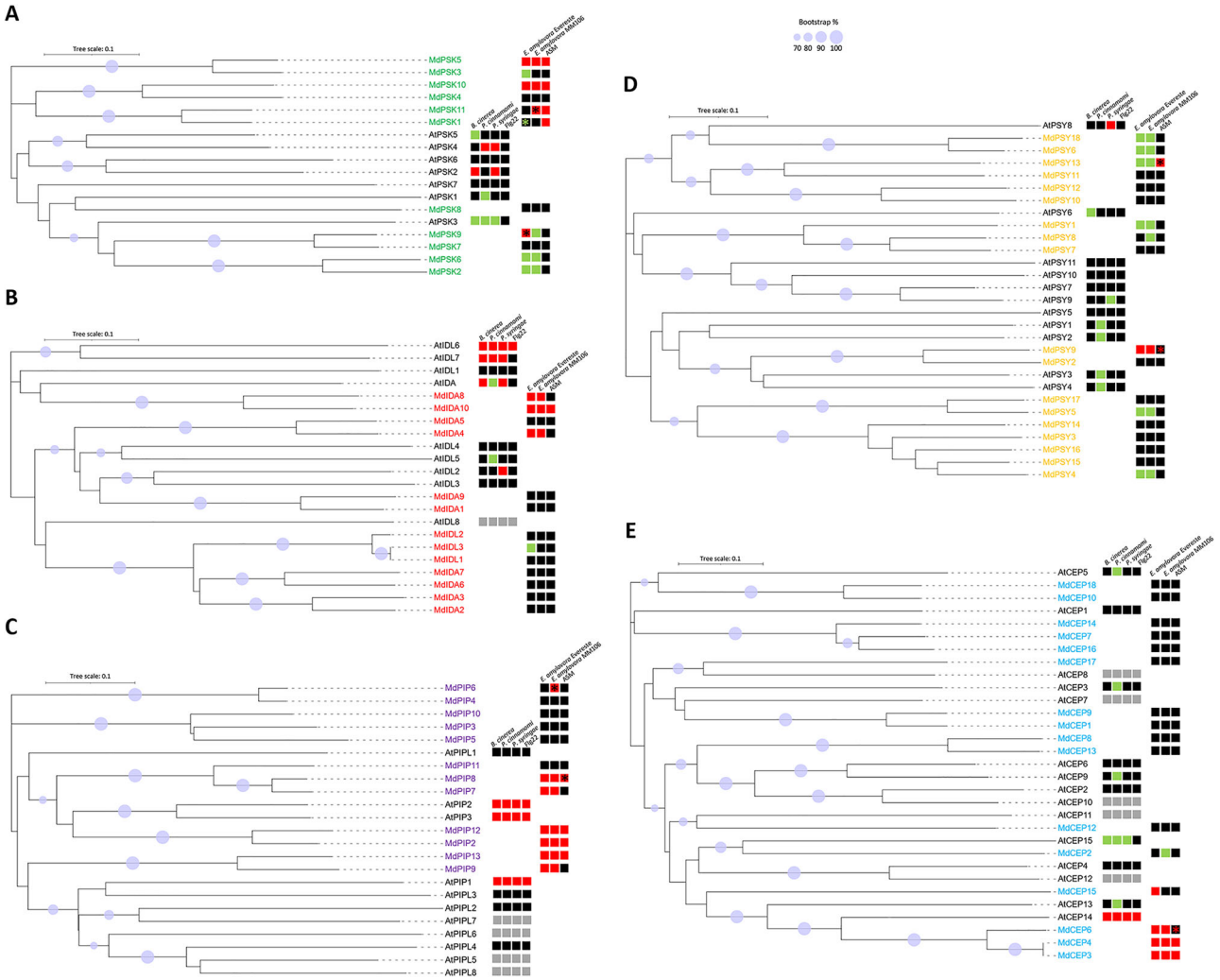

FIGURE 4 Phylogenetic relationships between *Malus* and *Arabidopsis* PSK (phytosulfokine; A), IDA/IDL (inflorescence deficient in abscission/-like; B), PIP (PAMP-induced secreted peptide; C), PSY (plant peptide containing sulfated tyrosine; D), and CEP (C-terminally encoded peptide; E) proteins. The phylogenetic trees were built using the neighbor-joining method with 1000 bootstrap replicates. Bootstrap values of nodes are indicated as circles of different sizes referring to percent intervals from 70% to 100%. Colored squares summarize differential gene expression analyses: red, green, and black squares for induced, repressed, and nonsignificantly differentially expressed genes, respectively. Gray squares mean that no data are available for these genes. *Malus* RNA-seq data are generated in this work and *Arabidopsis* RNA-seq data are extracted from Genevestigator/GEO (see the Results section for descriptions). A square is colored in red or green if the gene is differently expressed in at least one point of the time course analyzed with each pathogen/elicitor. For *Malus*, the asterisks specify the results obtained by reverse transcription quantitative real-time PCR (RT-qPCR) in the few situations where they diverge from the RNA-seq data, using the same color code as for squares.

identify a new undescribed *PSK* gene (AT2G22942, named *At*PSK7) encoding a protein with two YIYTQ motifs and two PSY genes (AT3G49270 named *AtPSY10* and AT3G49307 named *AtPSY11*). These sequences were added to the previously published *Arabidopsis* members of PSK, IDA, PIP, PSY, and CLE families for building phylogenetic trees with *Malus* homologs.

In order to identify orthologs between *Malus* and *Arabidopsis*, the homologous relationships between peptide precursors of the two species were also compared based on gene expression profile under biotic stress conditions (expressologs, Das et al., 2016). We used available RNA-seq data obtained from *Arabidopsis* leaves challenged with oomycete, fungus, gram-negative bacteria, and Flg22 elicitor (Figure 4). In some interesting situations, sequence similarities and transcriptional behaviors converge in the clustering of *Malus* and *Arabidopsis* genes, identifying them as candidate orthologs. For instance, in the CEP family, the three *MdCEP6*, four and three genes, induced upon *Ea* infection and ASM application, were clustered with *AtCEP14*, the only *Arabidopsis* CEP gene induced in presence of *B. cinerea*, *P. cinnamomi*, *P. syringae*, and flg22 (Figure 4E). A similar observation was made in the PIP family where *MdPIP8*, *MdPIP7*, *MdPIP12*, and *MdPIP2* are clustered with *AtPIP2* and *AtPIP3*, all of them significantly induced by the tested pathogens (Figure 4C). In PSK family, three genes repressed after biotic stresses are

gathered in a same monophyletic group with *AtPSK3* while other groups of *MdPSK* showed induced expression profiles for which no orthologs were found in *Arabidopsis* in response to biotic stresses (Figure 4A).

4 | DISCUSSION

This work presents a comprehensive characterization and an accurate nomenclature of five phytocytokine gene families in apple, namely, *MdPSK*, *MdIDA*, *MdPIP*, *MdPSY*, and *MdCEP*, and provides the basis for determining their functional roles in apple defense against biotic stresses.

Our bioinformatics pipeline has led to the identification of 12 *MdPSK*, 13 *MdIDA/IDL*, 13 *MdPIP*, 18 *MdPSY*, and 19 *MdCEP*, whereas the *Arabidopsis* genome contains seven *AtPSK*, nine *AtIDA/IDL*, 11 *AtPIP*, 11 *AtPSY*, and 15 *AtCEP* genes. Out of these 75 phytocytokine genes, 38 were not detected in the apple genome annotation (v1.1). Indeed, due to their short size, low conservation level and quite specific condition of expression (leading to overall low transcript coverage), genes encoding PTMPs are known to be poorly detected by regular gene prediction tools (Takahashi et al., 2019). The *Md*PSK and *Md*PSY proteins all featured one C-terminal PSK and PSY motif, comprising five and 17 aa, respectively. Conversely, *Md*IDA/*Md*IDL, *Md*PIP, and *Md*CEP proteins exhibited IDA, PIP, and CEP conserved motif with a canonical size of 14–15 residues. Some of the *MdPIP* and *MdCEP* genes encode precursors with duplicated motifs and therefore seem to be able to produce several distinct peptides during their maturation process. The most striking situation concerns *Md*PIP2, which contains seven repeated motifs. A few similar cases have been previously described in *Arabidopsis* in the PIP, SCOOP, and CEP phytocytokine families with two to five repeats (Delay et al., 2013; Guillou et al., 2022; Vie et al., 2015).

The prediction of multiple secreted peptides in apple might have resulted from the whole genome duplication event that occurred approximately 27 million years ago (Lallemand et al., 2023). This is supported by the many instances of homology found between several pairs of *Malus* genes. In addition, local intragenic and intergenic duplication events may have occurred. For instance, combined to the chromosomal distribution, the analysis revealed that several TAG clusters were observed for *MdPSK*, *MdIDA*, *MdPIP*, *MdPSY*, and *MdCEP* genes suggesting local duplication events across the chromosomes. Their phylogenetic relationships suggest that some of these duplications were recent, as supported by the 100% sequence identity between *MdCEP3* and *MdCEP4* in nucleic acid sequences. Among the TAG cluster on chromosome 12, *Md*IDL proteins do not feature signal peptide. This loss could have occurred recently in the history of the family, then followed by two local duplications on the genome-wide scale leading to the occurrence of the three *MdIDL* genes. *MdPIP2* and *MdPIP12* were not located on the same chromosome but their sequence featured seven and five PIP motifs, respectively, and they are closely related by phylogenetic analysis, suggesting first intragenic duplications followed by a distant duplication event. Interestingly, these repeated PIP motifs are separated by a quite conserved spacer of seven residues that could be involved in the proteolytic cleavage of the precursors to release the distinct mature peptides. The numerous duplicated genes and the ability to generate several peptides from one single precursor raise the question of redundancy and diversity of biological functions among these multigenic families.

The CEP gene family plays crucial roles in a wide range of cellular processes involved in growth and development and in response to abiotic stresses (Ogilvie et al., 2014). To date, numerous gene members of the CEP family have previously been identified across various plant genomes such as the 22 CEPs identified in *Glycine max* (Sin et al., 2022), 24 in *Solanum lycopersicum* (Liu et al., 2022), 15 in *Oryza sativa* (Aggarwal et al., 2020), and 21 in *Nicotiana tabacum* (Pan et al., 2024). In apple, two previous publications focused on the CEP family but were based on a first draft version of a heterozygous *M. domestica* genome (Velasco et al., 2010). Li et al. (2018) had predicted 27 CEP genes, but 19 of them are clearly false positives and encode other proteins (notably kinesins and PIP precursors). Another bioinformatics work describes 12 *MdCEP* genes (Yu et al., 2019). Their annotations and our results converge for nine genes and we have therefore taken into account their name in the nomenclature of the family that we propose (Table 1). The three false positives correspond to a group of genes encoding PIP precursors and can be explained by probable common origin (Vie et al., 2015). Our approach enabled us to discriminate between these two families and identified a total of 19 *MdCEP* genes (see specific consensus motifs in Figure 1).

Our RNA-seq analysis revealed a similar up- or down-regulation of most of the PTMP genes between a resistant and a susceptible genotype challenged with *Ea*. These similar modulations, recorded in the five PTMP families, were significant as early as 6 hpi for many members (particularly for PIP genes that were all upregulated), suggesting an early signaling role in the *Ea*/apple interaction irrespective of the outcome of the interaction (compatible vs. incompatible). This could indicate that downstream events are not involved in resistance of Evereste to *Ea*, or that effective downstream defenses are lacking in the susceptible genotype. In Golden Delicious seedlings, ASM only triggered upregulation of PTMP genes. Several members of the five PTMP families were concerned, albeit in smaller numbers than those triggered by *Ea*, and were common to the latter, with the exception of *MdPSK1* and *MdPSY13*. Interestingly, none of the two genes specifically induced by *Ea* in the resistant

genotype (*MdPSK9* and *MdCEP15*) were induced by ASM in Golden Delicious seedlings but these results remain to be confirmed with comparable biological samples. Moreover, none of the PTMP genes described in this study colocalize with the strong *Ea* resistance quantitative trait locus (QTL) from Evereste, located on chromosome 12. Candidate genes under this QTL would rather be nucleotidebinding site and leucine-rich repeat protein or serine/threonine kinase genes (Parravicini et al., 2011). These results could suggest different resistance mechanisms between the Evereste resistance and the induced resistance in Golden Delicious. However, the role of these PTMP in apple immunity remains to be investigated.

The comparative analysis of phytocytokine gene expression profiles in apple and *Arabidopsis* under biotic stresses conditions provides further support for the orthologous relationships between the two species. Two monophyletic clusters in PIP and CEP families were particularly interesting as combining the two species with similar overexpression following various biotic stresses. Therefore, these consistent transcription profiles might suggest that these gene clusters have similar biological functions in both species. Knowing their role in plant immunity (Delay et al., 2013; Hou et al., 2014), it would be interesting to decipher their biological function and signaling pathway triggered in apple in response to biotic stresses.

This study describes five major families of phytocytokines in apple and establish a reference nomenclature. It provides a starting point to further explore their role in apple, particularly in plant's resistance against infections. It would be worthwhile to study whether the peptide/receptor pairs are conserved across species as well as to decipher the biological meaning of these multigenic families and their complex proteolytic processes. These phytocytokines open new avenues for potential applications in fruit trees toward improved resistance against biotic stresses, or development of new biocontrol methods.

AUTHOR CONTRIBUTIONS

Marie-Charlotte Guillou: Data curation; formal analysis; investigation; visualization; writing—original draft; writing—review and editing. **Matthieu Gaucher**: Formal analysis; investigation; validation; visualization; writing—original draft; writing—review and editing. **Emilie Vergne**: Investigation; writing—review and editing. **Jean-Pierre Renou**: Conceptualization; writing—review and editing. **Marie-Noëlle Brisset**: Conceptualization; writing—review and editing. **Sébastien Aubourg**: Conceptualization; data curation; investigation; supervision; validation; writing—original draft; writing—review and editing.

ACKNOWLEDGMENTS

The authors thank the University of Angers and INRAE for fundings. The authors are grateful to Jean-Marc Celton for improving the manuscript, Sylvain Gaillard for updating JBrowse, Sandra Pelletier for the advice on using AnaDiff, the ANAN (SFR QuaSaV, Angers, France), and the Génome Québec facilities (Montreal, Canada) for RNA sequencing, and to the PHENOTIC plant facility (https://doi.org/10.17180/YKBZ-2V85) for plant cultivation and experiments.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

All data supporting the findings of this study are available within the paper and within its Supporting Information published online. Biological materials are available from the corresponding authors upon request. RNA-seq data were deposited at Gene Expression Omnibus (NCBI) under the accession numbers GSE262535 (apple/*Ea*) and GSE262538 (apple/ASM).

ORCID

Marie-Charlotte Guillou https://orcid.org/0000-0002-9700-4635
Sébastien Aubourg https://orcid.org/0000-0002-0695-4767

REFERENCES

Aćimović, S. G., Zeng, Q., McGhee, G. C., Sundin, G. W., & Wise, J. C. (2015). Control of fire blight (*Erwinia amylovora*) on apple trees with trunk-injected plant resistance inducers and antibiotics and assessment of induction of pathogenesis-related protein genes. *Frontiers in Plant Science*, *10*, Article 16.

Aggarwal, S., Kumar, A., Jain, M., Sudan, J., Singh, K., Kumari, S., & Mustafiz, A. (2020). C-terminally encoded peptides (CEPs) are potential mediators of abiotic stress response in plants. *Physiology and Molecular Biology of Plants*, *26*, 2019–2033. https://doi.org/10.1007/s12298-020-00881-4

Amano, Y., Tsubouchi, H., Shinohara, H., Ogawa, M., & Matsubayashi, Y. (2007). Tyrosine-sulfated glycopeptide involved in cellular proliferation and expansion in *Arabidopsis*. *Proceedings of the National Academy of Sciences*, *104*, 18333–18338. https://doi.org/10.1073/pnas.0706403104

Brisset, M.-N., Cesbron, S., Thomson, S. V., & Paulin, J. P. (2000). Acibenzolar-S-methyl induces the accumulation of defense-related enzymes in apple and protects from fire blight. *European Journal of Plant Pathology*, *106*, 529–536. https://doi.org/10.1023/A:1008728119087

Buscaill, P., & Rivas, S. (2014). Transcriptional control of plant defence responses. *Current Opinion in Plant Biology*, *20*, 35–46. https://doi.org/10.1016/j.pbi.2014.04.004

Butenko, M. A., Patterson, S. E., Grini, P. E., Stenvik, G. E., Amundsen, S. S., Mandal, A., & Aalen, R. B. (2003). Inflorescence deficient in abscission controls floral organ abscission in *Arabidopsis* and identifies a novel family of putative ligands in plants. *The Plant Cell*, *15*, 2296–2307. https://doi.org/10.1105/tpc.014365

Chao, J., Li, Z., Sun, Y., Aluko, O. O., Wu, X., Wang, Q., & Liu, G. (2021). MG2C: A user-friendly online tool for drawing genetic maps. *Molecular Horticulture*, *1*, 1–4. https://doi.org/10.1186/s43897-021-00020-x

Chavonet, E., Gaucher, M., Warneys, R., Bodelot, A., Heintz, C., Juillard, A., Cournol, R., Widmalm, G., Bowen, J. K., Hamiaux, C., Brisset, M. N., & Degrave, A. (2022). Search for host defense markers uncovers an apple agglutination factor corresponding with fire blight resistance. *Plant Physiology*, *188*, 1350–1368. https://doi.org/10.1093/plphys/kiab542

Couto, D., & Zipfel, C. (2016). Regulation of pattern recognition receptor signalling in plants. *Nature Reviews Immunology*, *16*, 537–552. https://doi.org/10.1038/nri.2016.77

Crooks, G. E., Hon, G., Chandonia, J. M., & Brenner, S. E. (2004). WebLogo: A sequence logo generator. *Genome Research*, *14*, 1188–1190. https://doi.org/10.1101/gr.849004

Daccord, N., Celton, J.-M., Linsmith, G., Becker, C., Choisne, N., Schijlen, E., Van De Geest, H., Bianco, L., Micheletti, D., Velasco, R., Di Pierro, E. A., Gouzy, J., Rees, D. J. G., Guérif, P., Muranty, H., Durel, C.-E., Laurens, F., Lespinasse, Y., Gaillard, S., … Bucher, E. (2017). High-quality *de novo* assembly of the apple genome and methylome dynamics of early fruit development. *Nature Genetics*, *49*, 1099–1106. https://doi.org/10.1038/ng.3886

Das, M., Haberer, G., Panda, A., Das Laha, S., Ghosh, T. C., & Schäffner, A. R. (2016). Expression pattern similarities support the prediction of orthologs retaining common functions after gene duplication events. *Plant Physiology*, *171*, 2343–2357. https://doi.org/10.1104/pp.15.01207

Delay, C., Imin, N., & Djordjevic, M. A. (2013). CEP genes regulate root and shoot development in response to environmental cues and are specific to seed plants. *Journal of Experimental Botany*, *64*, 5383–5394. https://doi.org/10.1093/jxb/ert332

Depotter, J. R., Misas Villamil, J. C., & Doehlemann, G. (2022). Maize immune signalling peptide ZIP1 evolved de novo from a retrotransposon. *BioRxiv*. https://doi.org/10.1101/2022.05.18.492421

De Smet, I., Voß, U., Jürgens, G., & Beeckman, T. (2009). Receptor-like kinases shape the plant. *Nature Cell Biology*, *11*, 1166–1173. https://doi.org/10.1038/ncb1009-1166

Dugé de Bernonville, T., Gaucher, M., Flors, V., Gaillard, S., Paulin, J. P., Dat, J. F., & Brisset, M. N. (2012). T3SS-dependent differential modulations of the jasmonic acid pathway in susceptible and resistant genotypes of *Malus* spp. challenged with *Erwinia amylovora*. *Plant Science*, *188*, 1–9. https://doi.org/10.1016/j.plantsci.2012.02.009

Dugé de Bernonville, T., Marolleau, B., Staub, J., Gaucher, M., & Brisset, M. N. (2014). Using molecular tools to decipher the complex world of plant resistance inducers: An apple case study. *Journal of Agricultural and Food Chemistry*, *62*, 11403–11411. https://doi.org/10.1021/jf504221x

Durel, C.-E., Denancé, C., & Brisset, M.-N. (2009). Two distinct major QTL for resistance to fire blight co-localize on linkage group 12 in apple genotypes 'Evereste' and *Malus floribunda* clone 821. *Genome*, *52*, 139–147. https://doi.org/10.1139/G08-111

Duvaud, S., Gabella, C., Lisacek, F., Stockinger, H., Ioannidis, V., & Durinx, C. (2021). Expasy, the Swiss bioinformatics resource portal, as designed by its users. *Nucleic Acids Research*, *49*, W216–W227. https://doi.org/10.1093/nar/gkab225

Eddy, S. R. (2011). Accelerated profile HMM searches. *PLoS Computational Biology*, *7*, e1002195. https://doi.org/10.1371/journal.pcbi.1002195

Furumizu, C., & Shinohara, H. (2024). Land plant peptide signaling: What we know—and don't know—about its evolution. *Physiologia Plantarum*, *176*, e14172. https://doi.org/10.1111/ppl.14172

Gaucher, M., Heintz, C., Cournol, R., Juillard, A., Bellevaux, C., Cavaignac, S., Coureau, C., Giraud, M., Le Maguet, J., Bérud, M., Koké, E., Crété, X., Lemarquand, A., Orain, G., & Brisset, M.-N. (2022). The use of potassium phosphonate (KHP) for the control of major apple pests. *Plant Disease*, *106*, 3166–3177. https://doi.org/10.1094/PDIS-01-22-0183-RE

Gaucher, M., Righetti, L., Aubourg, S., Dugé de Bernonville, T., Brisset, M. N., Chevreau, E., & Vergne, E. (2022). An *Erwinia amylovora* inducible promoter for improvement of apple fire blight resistance. *Plant Cell Reports*, *41*, 1499–1513. https://doi.org/10.1007/s00299-022-02869-8

Gilliland, G., Perrin, S., & Bunn, H. F. (1990). Competitive PCR for quantitation of mRNA. In *PCR protocols: A guide to method and applications* (pp. 60–70). Academic Press.

Goyal, R. K., & Mattoo, A. K. (2014). Multitasking antimicrobial peptides in plant development and host defense against biotic/abiotic stress. *Plant Science*, *228*, 135–149. https://doi.org/10.1016/j.plantsci.2014.05.012

Guillou, M.-C., Balliau, T., Vergne, E., Canut, H., Chourré, J., Herrera-León, C., Ramos-Martín, F., Ahmadi-Afzadi, M., D'amelio, N., Ruelland, E., Zivy, M., Renou, J.-P., Jamet, E., & Aubourg, S. (2022). The *PROSCOOP10* gene encodes two extracellular hydroxylated peptides and impacts flowering time in *Arabidopsis*. *Plants*, *11*, 3554. https://doi.org/10.3390/plants11243554

Gully, K., Pelletier, S., Guillou, M.-C., Ferrand, M., Aligon, S., Pokotylo, I., Perrin, A., Vergne, E., Fagard, M., Ruelland, E., Grappin, P., Bucher, E., Renou, J.-P., & Aubourg, S. (2019). The SCOOP12 peptide regulates defense response and root elongation in *Arabidopsis thaliana*. *Journal of Experimental Botany*, *70*, 1349–1365. https://doi.org/10.1093/jxb/ery454

Gust, A. A., Pruitt, R., & Nürnberger, T. (2017). Sensing danger: Key to activating plant immunity. *Trends in Plant Science*, *22*, 779–791. https://doi.org/10.1016/j.tplants.2017.07.005

Hou, S., Liu, D., & He, P. (2021). Phytocytokines function as immunological modulators of plant immunity. *Stress Biology*, *1*, 1–14. https://doi.org/10.1007/s44154-021-00009-y

Hou, S., Wang, X., Chen, D., Yang, X., Wang, M., Turrà, D., Di Pietro, A., & Zhang, W. (2014). The secreted peptide PIP1 amplifies immunity through receptor-like kinase 7. *PLoS Pathogens*, *10*, e1004331. https://doi.org/10.1371/journal.ppat.1004331

King, E. O., Ward, M. K., & Raney, D. E. (1954). Two simple media for the demonstration of pyocyanin and fluorescin. *Journal of Laboratory and Clinical Medicine*, *44*, 301–307.

Lallemand, T., Leduc, M., Desmazières, A., Aubourg, S., Rizzon, C., Landès, C., & Celton, J.-M. (2023). Insights into the evolution of ohnologous sequences and their epigenetic marks post-WGD in *Malus domestica*. *Genome Biology and Evolution*, *15*, evad178. https://doi.org/10.1093/gbe/evad178

Lalun, V. O., Breiden, M., Galindo-Trigo, S., Smakowska-Luzan, E., Simon, R. G. W., & Butenko, M. A. (2023). A dual function of the IDA peptide in regulating cell separation and modulating plant immunity at the molecular level. *Elife*, *12*, RP87912. https://doi.org/10.7554/eLife.87912

Li, R., An, J. P., You, C. X., Shu, J., Wang, X. F., & Hao, Y. J. (2018). Identification and expression of the CEP gene family in apple (*Malus* × *domestica*). *Journal of Integrative Agriculture*, *17*, 348–358. https://doi.org/10.1016/S2095-3119(17)61653-8

Liu, D., Shen, Z., Zhuang, K., Qiu, Z., Deng, H., Ke, Q., Liu, H., & Han, H. (2022). Systematic annotation reveals CEP function in tomato root

development and abiotic stress response. *Cells*, *11*, 2935. https://doi.org/10.3390/cells11192935

Madeira, F., Pearce, M., Tivey, A. R. N., Basutkar, P., Lee, J., Edbali, O., Madhusoodanan, N., Kolesnikov, A., & Lopez, R. (2022). Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic Acids Research*, *50*, W276–W279. https://doi.org/10.1093/nar/gkac240

Marmiroli, N., & Maestri, E. (2014). Plant peptides in defense and signaling. *Peptides*, *56*, 30–44. https://doi.org/10.1016/j.peptides.2014.03.013

Matsubayashi, Y. (2011). Small post-translationally modified peptide signals in *Arabidopsis*. *The Arabidopsis Book*, *9*. https://bioone.org/journals/the-arabidopsis-book/volume-2011/issue-9

Matsubayashi, Y., & Sakagami, Y. (1996). Phytosulfokine, sulfated peptides that induce the proliferation of single mesophyll cells of *Asparagus officinalis* L. *Proceedings of the National Academy of Sciences*, *93*, 7623–7627. https://doi.org/10.1073/pnas.93.15.7623

Mosher, S., Seybold, H., Rodriguez, P., Stahl, M., Davies, K. A., Dayaratne, S., Morillo, S. A., Wierzba, M., Favery, B., Keller, H., Tax, F. E., & Kemmerling, B. (2013). The tyrosine-sulfated peptide receptors PSKR1 and PSY1R modify the immunity of *Arabidopsis* to biotrophic and necrotrophic pathogens in an antagonistic manner. *The Plant Journal*, *73*, 469–482. https://doi.org/10.1111/tpj.12050

Murphy, E., Smith, S., & De Smet, I. (2012). Small signaling peptides in *Arabidopsis* development: How cells communicate over a short distance. *The Plant Cell*, *24*, 3198–3217. https://doi.org/10.1105/tpc.112.099010

Nielsen, H. (2017). Predicting secretory proteins with SignalP. *Methods in Molecular Biology*, *1611*, 59–73. https://doi.org/10.1007/978-1-4939-7015-5_6

Ogilvie, H. A., Imin, N., & Djordjevic, M. A. (2014). Diversification of the C-TERMINALLY ENCODED PEPTIDE (CEP) gene family in angiosperms, and evolution of plant-family specific CEP genes. *BMC Genomics*, *15*, 1–15. https://doi.org/10.1186/1471-2164-15-870

Ohyama, K., Ogawa, M., & Matsubayashi, Y. (2008). Identification of a biologically active, small, secreted peptide in Arabidopsis by in silico gene screening, followed by LC-MS-based structure analysis. *Plant Journal*, *55*, 152–160. https://doi.org/10.1111/j.1365-313X.2008.03464.x

Pan, X., Deng, Z., Wu, R., Yang, Y., Akher, S. A., Li, W., Zhang, Z., & Guo, Y. (2024). Identification of CEP peptides encoded by the tobacco (*Nicotiana tabacum*) genome and characterization of their roles in osmotic and salt stress responses. *Plant Physiology and Biochemistry*, *209*, 108525. https://doi.org/10.1016/j.plaphy.2024.108525

Parravicini, G., Gessler, C., Denancé, C., Lasserre-Zuber, P., Vergne, E., Brisset, M.-N., Patocchi, A., Durel, C.-E., & Broggini, G. A. L. (2011). Identification of serine/threonine kinase and nucleotide-binding site–leucine-rich repeat (NBS-LRR) genes in the fire blight resistance quantitative trait locus of apple cultivar 'Evereste'. *Molecular Plant Pathology*, *12*, 493–505. https://doi.org/10.1111/j.1364-3703.2010.00690.x

Patro, R., Duggal, G., Love, M. I., Irizarry, R. A., & Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. *Nature Methods*, *14*, 417–149. https://doi.org/10.1038/nmeth.4197

Pearce, G., Strydom, D., Johnson, S., & Ryan, C. A. (1991). A polypeptide from tomato leaves induces wound-inducible proteinase inhibitor proteins. *Science*, *253*, 895–897. https://doi.org/10.1126/science.253.5022.895

Pelletier, S. (2022). *AnaDiff: A tool for differential analysis of microarrays and RNAseq* (Version v4. 3) [Computer software]. Zenodo. https://doi.org/10.5281/zenodo.6477918

Pruitt, R. N., Joe, A., Zhang, W., Feng, W., Stewart, V., Schwessinger, B., Dinneny, J. R., & Ronald, P. C. (2017). A microbially derived tyrosine-sulfated peptide mimics a plant peptide hormone. *New Phytologist*, *215*, 725–736. https://doi.org/10.1111/nph.14609

Rzemieniewski, J., Leicher, H., Lee, H. K., Broyart, C., Nayem, S., Wiese, C., Maroschek, J., Camgöz, Z., Lalun, V. O., Djordjevic, M. A., Corina Vlot, A., Hückelhoven, R., Santiago, J., & Stegmann, M. (2022). CEP signaling coordinates plant immunity with nitrogen status. *BioRxiv*. https://doi.org/10.1101/2022.12.20.521212

Sauter, M. (2015). Phytosulfokine peptide signalling. *Journal of Experimental Botany*, *66*, 5161–5169. https://doi.org/10.1093/jxb/erv071

Shen, Y., & Diener, A. C. (2013). *Arabidopsis thaliana* RESISTANCE TO FUSARIUM OXYSPORUM 2 implicates tyrosine-sulfated peptide signaling in susceptibility and resistance to root infection. *PLoS Genetics*, *9*, e1003525. https://doi.org/10.1371/journal.pgen.1003525

Sin, W. C., Lam, H. M., & Ngai, S. M. (2022). Identification of diverse stress-responsive xylem sap peptides in soybean. *International Journal of Molecular Sciences*, *23*, 8641. https://doi.org/10.3390/ijms23158641

Stührwohldt, N., Bühler, E., Sauter, M., & Schaller, A. (2021). Phytosulfokine (PSK) precursor processing by subtilase SBT3. 8 and PSK signaling improve drought stress tolerance in *Arabidopsis*. *Journal of Experimental Botany*, *72*, 3427–3440. https://doi.org/10.1093/jxb/erab017

Takahashi, F., Hanada, K., Kondo, T., & Shinozaki, K. (2019). Hormone-like peptides and small coding genes in plant stress signaling and development. *Current Opinion in Plant Biology*, *51*, 88–95. https://doi.org/10.1016/j.pbi.2019.05.011

Tavormina, P., De Coninck, B., Nikonorova, N., De Smet, I., & Cammue, B. P. (2015). The plant peptidome: An expanding repertoire of structural features and biological functions. *The Plant Cell*, *27*, 2095–2118. https://doi.org/10.1105/tpc.15.00440

Tost, A. S., Kristensen, A., Olsen, L. I., Axelsen, K. B., & Fuglsang, A. T. (2021). The PSY peptide family-expression, modification and physiological implications. *Genes (Basel)*, *12*, 218. https://doi.org/10.3390/genes12020218

Van Canh, L. T., & Aubourg, S. (2024). Bioinformatics methods for prediction of gene families encoding extracellular peptides. *Methods in Molecular Biology*, *2731*, 3–21. https://doi.org/10.1007/978-1-0716-3511-7_1

Vanneste, J. L. (Ed.). (2000). *Fire blight: The disease and its causative agent, Erwinia amylovora*. CABI.

Velasco, R., Zharkikh, A., Affourtit, J., Dhingra, A., Cestaro, A., Kalyanaraman, A., Fontana, P., Bhatnagar, S. K., Troggio, M., Pruss, D., Salvi, S., Pindo, M., Baldi, P., Castelletti, S., Cavaiuolo, M., Coppola, G., Costa, F., Cova, V., Dal Ri, A., … Viola, R. (2010). The genome of the domesticated apple (*Malus* × *domestica* Borkh.). *Nature Genetics*, *42*, 833–839. https://doi.org/10.1038/ng.654

Venisse, J. S., Malnoy, M., Faize, M., Paulin, J. P., & Brisset, M. N. (2002). Modulation of defense responses of *Malus* spp. during compatible and incompatible interactions with *Erwinia amylovora*. *Molecular Plant Microbe Interactions*, *15*, 1204–1212. https://doi.org/10.1094/MPMI.2002.15.12.1204

Vie, A. K., Najafi, J., Liu, B., Winge, P., Butenko, M. A., Hornslien, K. S., Kumpf, R., Aalen, R. B., Bones, A. M., & Brembu, T. (2015). The IDA/IDA-LIKE and PIP/PIP-LIKE gene families in

Arabidopsis: Phylogenetic relationship, expression patterns, and transcriptional effect of the PIPL3 peptide. *Journal of Experimental Botany*, *66*, 5351–5365. https://doi.org/10.1093/jxb/erv285

Warneys, R., Gaucher, M., Robert, P., Aligon, S., Anton, S., Aubourg, S., Barthes, N., Braud, F., Cournol, R., Gadenne, C., Heintz, C., Brisset, M.-N., & Degrave, A. (2018). Acibenzolar-s-methyl reprograms apple transcriptome toward resistance to rosy apple aphid. *Frontiers in Plant Science*, *9*, Article 1795. https://doi.org/10.3389/fpls.2018.01795

Yu, Z., Xu, Y., Liu, L., Guo, Y., Yuan, X., Man, X., Liu, C., Yang, G., Huang, J., Yan, K., Zheng, C., Wu, C., & Zhang, S. (2019). The importance of conserved serine for C-terminally encoded peptides function exertion in apple. *International Journal of Molecular Sciences*, *20*, 775. https://doi.org/10.3390/ijms20030775

Yuan, X., Gdanetz, K., Outwater, C. A., Slack, S. M., & Sundin, G. W. (2023). Evaluation of plant defense inducers and plant growth regulators for fire blight management using transcriptome studies and field assessments. *Phytopathology*, *113*, 2152–2164. https://doi.org/10.1094/PHYTO-04-23-0147-KC

Zhang, T., Li, X., Zhao, Q., Shi, Y., Hao, Y., & You, C. (2022). Genome-wide identification and functional characterization of the MdCLE peptide family in apple (*Malus* × *domestica*). *Horticultural Plant Journal*, *8*, 279–288. https://doi.org/10.1016/j.hpj.2021.12.003

Zimmermann, P., Hennig, L., & Gruissem, W. (2005). Gene-expression analysis and network discovery using Genevestigator. *Trends in Plant Science*, *10*, 407–409. https://doi.org/10.1016/j.tplants.2005.07.003

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Guillou, M.-C., Gaucher, M., Vergne, E., Renou, J.-P., Brisset, M.-N., & Aubourg, S. (2025). Phytocytokine genes newly discovered in *Malus domestica* and their regulation in response to *Erwinia amylovora* and acibenzolar-S-methyl. *The Plant Genome*, *18*, e20540. https://doi.org/10.1002/tpg2.20540