Comparative Analysis of Small Secreted Peptide Signaling during Defense Response: Insights from Vascular and Non-Vascular Plants

Irina Lyapina[1], Daria Ganaeva[1], Eugene A. Rogozhin[1,2], Ekaterina V. Ryabukhina[1], Dmitry Yu. Ryazantsev[1], Vassili Lazarev[3], Sabina E. Alieva[3], Anna Mamaeva[1] and Igor Fesenko[4*]

[1]Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Miklukho-Maklaya 16/10, 117997 Moscow, Russia

[2]All-Russian Institute for Plant Protection, Pushkin 196608, Russia

[3]Lopukhin Federal Research and Clinical Center of Physical-Chemical Medicine of Federal Medical Biological Agency, Moscow, Russia

[4]No affiliation

*Corresponding author

Abstract

Small secreted peptides (SSPs) play an important role in modulating immune responses in all land plants. However, the evolution of stress peptide signaling in different plant phyla remains poorly understood. Here, we compared the expression of SSP genes in the pathogen-induced transcriptomes of vascular and non-vascular plants. We found 13, 19, 15, and 28 SSP families that were differentially expressed during infection in *Physcomitrium patens*, *Zea mays*, *Brassica napus*, and *Solanum tuberosum*, respectively. A comparative study of peptide motifs and predicted three-dimensional structures confirmed the similarity of SSPs across the examined plant species. In both vascular and non-vascular plants, however, only the RALF peptide family was differentially regulated under infection. We also found that EPFL peptides, which are involved in growth and development processes in angiosperms, were differentially regulated in *P. patens* in response to

pathogen infection. The search for novel immune-specific peptides revealed a family of PSY-like peptides that are differentially regulated during infection in *P. patens*. The treatment with synthetic EPFL, MEG, PSY, and PSY-like peptides validated their roles in the immune response and growth regulation. Thus, our study showed the complex nature of SSP signaling and shed light on the regulation of SSPs in different plant lineages during infection.

1 Introduction

Peptides regulate different processes, including the response to stress conditions (Hou et al., 2021; Lyapina et al., 2019; Olsson et al., 2019). Like animal cytokines (J.-M. Zhang & An, 2007), a group of plant peptides modulating immune response was named "phytocytokines" (Hou et al., 2021). Most of the phytocytokine families were identified in angiosperms. The well-known phytocytokines include PLANT ELICITOR PEPTIDES (PEPs), PAMP-INDUCED PEPTIDES (PIPs), RAPID ALKALINIZATION FACTOR (RALF), PHYTOSULFOKINES (PSKs), PLANT PEPTIDE CONTAINING SULFATED TYROSINE1 (PSY1), SERINE-RICH ENDOGENOUS PEPTIDES (SCOOPs), GOLVENs (GLVs), CLAVATA3/EMBRYO SURROUNDING REGION (CLV3/CLE), CLAVATA3/EMBRYO SURROUNDING REGION (CLV3/CLE), and INFLORESCENCE DEFICIENT IN ABSCISSION (IDA) families (Rzemieniewski & Stegmann, 2022). In addition to the stress response, phytocytokines are shown to be involved in the regulation of plant growth and development (Hou et al., 2021).

Besides phytocytokines, antimicrobial peptides (AMPs) play an important role in plant defense, triggering pathogen cell lysis due to lipid membrane disruption (Moravej et al., 2018). Plant AMPs are small, cysteine-rich proteins (Kulaeva et al., 2020; Tam et al., 2015) that have diverse structures, characteristics, and action mechanisms. The classification of plant AMPs is mostly based on sequence similarity and structures. The most characterized AMPs include thionins, defensins, hevein-like peptides, knottins, hairpin-like peptides, lipid transfer proteins, snakins, and cyclotides (Lima et al., 2022). Antimicrobial peptides were found in most plant taxa.

The phytocytokines, antimicrobial peptides, and plant peptide hormones that control growth and development are often referred to as Small Secreted Peptides (SSPs). The plant SSPs are involved in cell-to-cell communication and share common features: 1) the active peptide sequences are embedded at the C-terminus of 100–250 aa functional or non-functional protein precursors; 2) the protein precursors contain a hydrophobic N-terminal secretion signal; 3) the peptides are released by specific proteases under certain conditions (de Bang et al., 2017; Hu et al., 2021). Nevertheless, precursors of widely recognized phytocytokines, including systemin (SYS) and PEP precursors, do not possess an export signal, which highlights the shortcomings of the SSP classification (Hu et al., 2021; Tavormina et al., 2015). The group of SSPs consists of post-translationally modified (PTM) peptides, cysteine-rich peptides (CRP), and non-Cys-rich and non-PTM peptides. The families of PTM peptides such as PSK, PSY, RGF/GLV/CLEL, CLE, and IDA are widely conserved and were presented in the most recent common ancestor of angiosperms and bryophytes. The cysteine-rich peptides, such as RALF, EPIDERMAL PATTERNING FACTOR (EPF), and EPF-LIKE (EPFL), are shown to be involved in the regulation of many aspects of plant growth and development (Tian et al., 2022). These families were also present in the most recent common ancestor of angiosperms and bryophytes (Furumizu & Shinohara, 2024). The families of SYS and PEP peptides are the most well-known phytocytokines, not belonging to the aforementioned two categories.

During the past decade, thousands of SSPs belonging to more than 70 families have been identified in different plant genomes (Boschiero et al., 2020). In *Triticum aestivum*, 4,981 putative SSPs were identified, among which 1,790 SSPs were grouped into 38 known SSP families (Tian et al., 2022). In *Oryza sativa*, 236 SSPs, including 52 previously unannotated ones, were shown to be involved in the immune response (P. Wang et al., 2020). In *Medicago truncatula*, 4439 genes were predicted to encode SSPs (Boschiero et al., 2020). The bioinformatic search in three bryophyte lineages — mosses, liverworts, and hornworts identified hundreds of SSPs, confirming the important role of peptide signaling in all plant lineages (Lyapina et al., 2021).

The databases of plant SSPs, such as PlantSSP (Ghorbani et al., 2015), OrysSSP (Pan et al., 2013), and MtSSPdb (Boschiero et al., 2020) have been published in recent years.

77 Nevertheless, the catalog of plant SSPs is far from complete, and new peptide families are still
78 being discovered. Recently, a new family of plant phytocytokines, SCREW/CTNIP, was identified
79 in Arabidopsis based on the analysis of stress-induced transcriptomes (Z. Liu et al., 2022; Rhodes
80 et al., 2022). Moreover, plant genomes contain tens of thousands of small open reading frames
81 (smORFs, < 100 aa) that can encode unannotated secreted peptides or be a raw material for the
82 evolution of novel SSPs (Fesenko et al., 2021; Hanada et al., 2013). Such translatable smORFs
83 can be located on transcripts that were annotated as long non-coding RNAs (lncRNAs) (Hsu &
84 Benfey, 2018).

85 The evolution and role of SSPs in land colonization by plants are currently being discussed
86 (Bowman, 2022; Bowman et al., 2017; Ghorbani et al., 2015). The analysis of phytocytokines in
87 different plant groups showed that most of them are family- or species-specific. That probably
88 reflects their fast evolution, which is driven by arms races between plants and microbes. For
89 example, the SCOOP peptide family is specific to *Brassicaceae* family (Yang et al., 2023).
90 Systemin and hydroxyproline-rich systemin were identified only in *Solanaceae* family (Ryan &
91 Pearce, 2003). A species-specific *Z. mays* immune signaling peptide 1 (Zip1) was found in *Zea*
92 *mays* (Ziemann et al., 2018). The PEP family is present in many angiosperm families, but their
93 sequences are highly divergent (Lori et al., 2015; Tanaka & Heil, 2021).

94 Nevertheless, such PTM phytocytokine families as PSK, PSY, RGF/GLV/CLEL, CLE,
95 and IDA were likely present in the most recent common ancestor of angiosperms and bryophytes
96 (Furumizu & Shinohara, 2024). The role of RALF peptides in the immune response was recently
97 demonstrated in both angiosperms and in the moss *P. patens* (Mamaeva et al., 2023). Nevertheless,
98 it is currently unclear whether plant lineages such as bryophytes, lycophytes, ferns, and
99 gymnosperms utilize the same SSP families to modulate immune responses, and how many
100 lineage-specific immune peptide families remain unidentified. Moreover, the understanding of the
101 interactions between several families of phytocytokines in the immune response in diverse plant
102 lineages is currently incomplete.

103 In this study, we compared the patterns of SSP expression under pathogen invasion in three
104 angiosperm species — *Zea mays*, *Brassica napus*, and *Solanum tuberosum* — and one bryophyte,
105 the moss *Physcomitrium patens*. Our analysis revealed different patterns of SSP transcription
106 between vascular plants as well as between vascular and non-vascular plants. In addition, our
107 bioinformatic search identified two novel PSY-like peptide families that are specific to bryophytes.

108 2 Materials and Methods

109 2.1 Pre-processing

110 Raw RNA sequencing data from the following plants under biotic stress were analyzed: the
111 model organism, moss *Physcomitrium patens*, and the angiosperms: *Zea mays*, *Brassica napus*, and
112 *Solanum tuberosum* (accessions: SRP274010, SRP390856, SRP053361, SRP066006, respectively).
113 The following time points were chosen: 4, 8, and 24 hours post inoculation (hpi) for moss *P. patens*
114 infected by *Botrytis cinerea*, 24, 40, 60 hpi for *Z. mays* infected by *Colletotrichum graminicola;* 24,
115 48, 96 hpi for *B. napus* infected by *Sclerotinia sclerotiorum*; 12, 24, 72 hpi for *S. tuberosum* infected
116 by *Pectobacterium carotovorum.*

117 FastQC (Andrews, 2010) was used for quality control, and Trimmomatic (Bolger et al., 2014)
118 was used to clip adapters and nucleotides of low quality. Genomes and genome annotations were
119 accessed via Phytozome (https://phytozome-next.jgi.doe.gov/, Goodstein et al., 2012) and Ensembl
120 Plants (https://plants.ensembl.org/). *P. patens* v3.3, *Z. mays* Ref_Gen_V4, *B. napus*
121 AST_PRJEB5043_v1, and *S. tuberosum* v6.1 genome versions were selected.

122 2.3 Assembly and ORF prediction

123 Genomes were indexed using hisat2-build, reads were mapped to the reference genome using
124 HISAT2 (Kim et al., 2015). Samtools (H. Li et al., 2009) was used for collecting mapping statistics.

125 *De novo* transcript assembly was done using Stringtie (Pertea et al., 2015), Stringtie -e mode
126 was applied for quantification. Prediction of open reading frames (ORFs) from AUG start codons for

127 assembled transcripts, along with further automatic translation to amino acid sequences longer than
128 30 aa, was performed using orfipy (Singh & Wurtele, 2021). The longest ORF per transcript was
129 selected, considering that the longest ORF represents actual coding sequence.

130 2.4 Functional annotation

131 Interproscan (Quevillon et al., 2005, version 5.66-98.0) was used to predict protein domains.
132 Analysis of homology with known families of small secreted peptides (SSPs) was done using the SSP
133 prediction tool (https://mtsspdb.zhaolab.org/prediction/, Boschiero et al., 2020). Plant long non-
134 coding RNAs were accessed via PLncDB (https://www.tobaccodb.org/plncdb/).

135 2.5 DE analysis and filtration

136 Differential gene expression analysis was done using the edgeR package (Robinson et al.,
137 2010). Transcripts with 1 > logFC < -1 and FDR < 0.05 were considered differentially expressed.

138 ORFs predicted as known SSPs by the SSP prediction tool were assigned to the corresponding
139 SSP family based on homology to SSPs from other plants. The search for potential novel candidates
140 of SSPs was conducted in ORFs shorter than 200 aa, corresponding to differentially regulated
141 transcripts, with predicted signal sequences, no transmembrane domains, no known protein domains,
142 and no homology to known SSPs from other plants.

143 2.6 Sequence alignment and structure prediction

144 Multiple sequence alignments (MSAs) were created using the high precision L-INS-I
145 algorithm of MAFFT v7 (Katoh & Standley, 2013). Alignments were visualized using JalView
146 v2.11.2.2, Clustalx color (Waterhouse et al., 2009).

147 Prediction of three-dimensional structures was done using AlphaFold2 (Jumper et al., 2021).
148 Structural comparison was performed using TM align (Y. Zhang & Skolnick, 2005). Such metrics as

149 RMSD and TM-score were considered. Structural search against PDB and AlphaFold database
150 (AFDB) was performed using Foldseek (van Kempen et al., 2024).

151 2.7 Search for peptide homologs

152 A search for paralogs of novel immune peptide candidates was performed using BLASTP
153 against all major ORFs assigned to each transcript. Hits with E-value < 0.00001 were considered
154 significant. The search for orthologs and paralogs for novel candidates in other databases was done
155 using NCBI BLASTP against the non-redundant protein sequences database, organism: *Viridiplantae.*
156 Sequences with no matches or with matches to unannotated proteins ("putative", "unknown",
157 "uncharacterized", etc) were selected. TBLASTN against 1000 plant transcriptomes (1KP,
158 https://db.cngb.org/datamart/plant/DATApla4/) was run to search for homologs of novel candidates
159 among translated nucleotide sequences of other plants published in 1KP. The following default
160 parameters were customized: a maximum of 250 target sequences and a word size of 3. Hits with E-
161 value < 0.01 were considered significant.

162 2.8 Cloning and recombinant protein expression

163 The pET32b+ vector ("Novagen", USA) was used for recombinant plasmids construction.
164 When the obtained recombinant plasmid is expressed, a hybrid protein is created that contains
165 thioredoxin, polyhistidine domain, enteropeptidase cleavage site (DDDDR) and the peptide
166 sequences. Previously, an assembly of synthetic genes encoding MEG and EPFL peptides was carried
167 out by PCR method. This procedure was made according to Shevchuk et al. (2004): the long
168 oligonucleotides (Supplementary Table 1) (100 ng of each) and all required components were added
169 to PCR mixture (the Phusion HS II DNA polymerase was used). The PCR amplification was carried
170 out using the following program: 94°C – 60 s; 94°C – 20 s, 60°C – 20 s, 72°C – 20 s – 10 cycles;
171 72°C – 60 s. Then, 2.5 µl of the obtained mixture was added to a novel amplification mixture
172 containing flanking oligonucleotides for the corresponding peptide (0.5 µM of each), and the
173 amplification was carried out according to the same program (20 cycles). The PCR product was
174 purified from the agarose gel and cloned into the pET32b+ vector using *HihdIII/KpnI* restriction

175 endonucleases. The obtained construction was transformed into *E. coli* strain Origami. The induction
176 and purification of the recombinant protein were carried out under native conditions according to the
177 QIAexpressionist manual ("Qiagen", Germany). For the purpose of fusion protein purification from
178 bacterial cells 1000 ml LB growth medium was inoculated with 25 ml of overnight *E. coli* culture.
179 The induction was carried out for 4 h in the presence of 1 mM IPTG, then the cells were centrifuged,
180 and the precipitate was resuspended in 50 ml of buffer A (50 mM NaH_2PO_4, 300 mM NaCl, 10 mM
181 imidazole, pH 8.0). The *E. coli* cells disruption was realized for 30 min on ice with the addition of
182 lysozyme (1 mg/ml) followed by sonication. The purified cell lysate was applied to Ni-NTA agarose
183 column (10 ml) pre-equilibrated with buffer A, then the column was washed with buffer A containing
184 20 mM imidazole. The fusion protein was eluted with buffer B (50 mM NaH_2PO_4, 300 mM NaCl,
185 250 mM imidazole, pH 8.0). The purity and size of recombinant hybrid proteins were evaluated by
186 SDS-PAGE electrophoresis.

187 ## 2.9 Peptide chemical synthesis

188 *P.patens* PSY1, PSYL1, and PSYL2 peptides (Table 1) were chemically synthesized at the
189 genetic engineering laboratory of the FRCC PCM (Moscow, Russia). The peptides were synthesized
190 using solid-phase synthesis on Rink resin. The synthesis was carried out with microwave irradiation
191 of the reaction mixture, according to the Fmoc-strategy. The purity of the lyophilized peptides was
192 > 95%, their molecular weight was confirmed by mass spectrometric analysis. The synthesized
193 peptides were dissolved in sterile water and stored at − 80°C.

194 **Table 1.** Sequences of synthesized peptides

Peptide	**Sequence**
PSY1	DYEKPCANKKHDPSVCKNG
PSYL1	DYHDPEFNHIHDLPKR
PSYL2	DYDDSQANERLHFSPFT

195 ## 2.10 Plant growth

196 Protonemata of *Physcomitrium patens* subsp. patens Gransden 2004 (Freiburg, Germany)
197 were grown in 100 ml liquid Knop medium containing 500 mg/l ammonium tartrate for ROS
198 accumulation analysis or in 600 µl solid Knop medium with 450 µl sterile water containing a
199 corresponding peptide, ammonium tartrate and glucose for cell length analysis, under white light with
200 a photon flux of 61 µmol/m2 s under a 16-hour photoperiod at 24°C. Seven-day-old protonema was
201 used for analysis.

202 2.11 ROS detection and cell length analysis

203 For intracellular ROS detection, we used the fluorescent dye 2′,7′-Dichlorofluorescin
204 Diacetate (DCFH-DA; Sigma-Aldrich, USA). *P. patens* protonemata were treated with 5 µM
205 synthesized PpEPFL or PpMEG for 10 min. Samples were incubated with 10 µM DCFH-DA for 15
206 min. The detection was performed on the fluorescent microscope Axio Imager M2 (Zeiss) with an
207 AxioCam 506 mono digital camera and filter units. The No. 44 filter (λex BP 475 nm/40 nm; λem BP
208 530 nm/50 nm) was used for DCFH-DA fluorescence detection. Data on the fluorescence intensity
209 were obtained from the related Zeiss software Zen.

210 For cell length analysis, protonema was grown with 1 nM - 1 µM of each peptide for 7 days.
211 Cell walls were stained with 10 µg/ml propidium iodide for 5 min. Filter unit 20 Rhodamin (λex BP
212 546/12 nm; λem BP 575–640 nm) was used.

213 2.12 Statistics

214 Statistical analysis and visualization were made in Python v. 3.7.4 (Van Rossum and Drake,
215 1995) using modules scipy 1.5.2 (Virtanen et al., 2020), seaborn 0.11.1 (Waskom, 2021), numpy
216 1.20.1, pandas 1.2.3 (McKinney, 2012). For determining statistically significant differences two- or
217 more-way analysis of variance (ANOVA), followed by Tukey's honestly significant difference (HSD)
218 test, or Mann-Whitney U test were applied. Differences were considered significant at a P-value <
219 0.05.

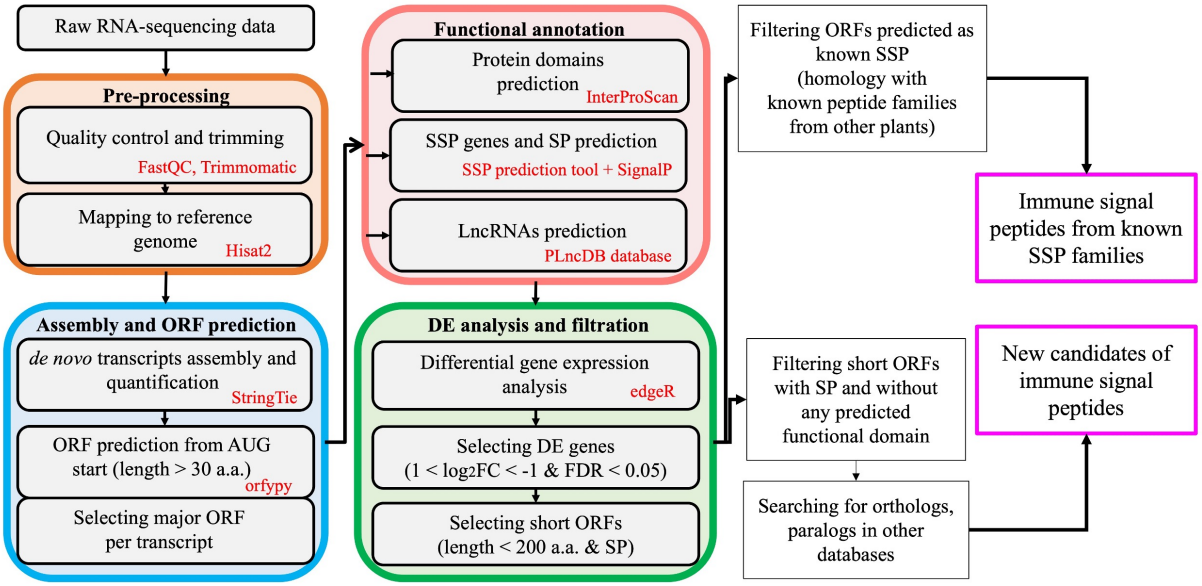
3 Results

3.1 Identification of SSPs in pathogen-induced transcriptomes

In order to analyze the expression patterns of known and previously unannotated SSPs during immune response, we re-analyzed four published transcriptomes of phylogenetically distant species: the moss *Physcomitrium patens* (Physcomitrella), *Zea mays*, *Brassica napus*, and *Solanum tuberosum* (Kwenda et al., 2016; Reboledo et al., 2021; D. Wang et al., 2022; Xia et al., 2023; Figure 1A). These transcriptomes were selected based on the following criteria: 1) the plants were infected with either necrotrophic or hemibiotrophic pathogens; 2) the quality of raw data (see Methods); and 3) RNA-seq analysis was performed at different time points after pathogen inoculation.

The RNA-seq reads were remapped to the corresponding genomes with the following alignment rate: 91% for *P. patens*; 96% for *Z. mays*; 83% for *B. napus*; and 85% for *S. tuberosum*. Overall, we identified from 28,234 to 67,590 transcripts at different hpi (hours post infection) across all selected species (Figure 1B). The following number of transcripts were assigned to annotated genes: 23,843 for *P. patens*; 15,269 for *S. tuberosum*; 15,138 for *Z. mays* and 19,875 for *B. napus*. To annotate all assembled transcripts, open reading frames (ORFs) larger than 30 amino acids (aa) that start with AUG codons were predicted using the orfipy tool (Singh & Wurtele, 2021). For annotation, the major ORF was selected for each transcript, resulting in approximately 45,000 predicted proteins in *P. patens* and *S. tuberosum*; approximately 40,000 and 60,500 predicted proteins in *Z. mays* and *B. napus*, respectively.

(A)



(B)

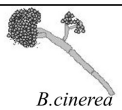
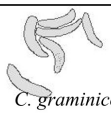
Plant	*Physcomitrium patens*			*Zea mays*			*Solanum tuberosum*			*Brassica napus*		
Pathogen	 *B.cinerea*			 *C. graminicola*			*P. carotovorum*			*S. sclerotiorum*		
Time Point, h	**4**	**8**	**24**	**24**	**40**	**60**	**12**	**24**	**72**	**24**	**48**	**96**
Transcripts	43780	43861	43978	42208	41277	28234	48590	49887	49434	65613	66116	67590
DE transcripts (% lncRNA)	251 (0.4)	6256 (0.21)	13190 (0.17)	18129 (0.41)	8572 (0.36)	9685 (0.41)	2266 (0.75)	704 (0.85)	1900 (1.05)	7938 (0.38)	10523 (0.29)	15143 (0.32)
DE short ORF + SP (% lncRNA)	2 (0)	172 (0.6)	348 (0.57)	1184 (0.34)	698 (0.14)	987 (0.61)	296 (1.35)	109 (3.7)	240 (1.25)	633 (0.32)	1121 (0.27)	1535 (0.52)
DE SSP	2	27	38	53	35	48	42	14	20	70	99	140
All SSP	129	129	129	89	94	92	174	174	157	314	289	250

Figure 1. (A) - General scheme of work pipeline; **(B) -** transcripts assembly statistics.

242 The predicted proteins were then annotated using the InterProScan tool (Quevillon et al.,
243 2005). This resulted in assigning to known protein domains about 57% of the assembled transcripts
244 in *P. patens* and *B. napus*, 53% in *S. tuberosum* and 70% in *Z. mays*. Among the most represented
245 domains, such as chloroplast- or ribosome-related domains, a significant portion of the predicted
246 ORFs belonged to uncharacterized domains or were unannotated (Supplementary Table 2).

247 We next used the online SSP prediction tool (https://mtsspdb.zhaolab.org/prediction/,
248 Boschiero et al., 2020) to determine the transcripts encoding putative precursors of known SSP
249 families (Figure 1B; Supplementary Table 3). Overall, 129, 92, 168, and 284 transcripts encoding
250 potential SSP precursors were identified in the assembled transcriptomes of *P. patens*, *Z. mays*, *S.*
251 *tuberosum*, and *B. napus*, respectively (Figure 1B; Supplementary Table 3).

252 For each species, we next identified differentially expressed (DE) transcripts ($-1 \leq \log2$
253 fold change ≥ 1, $P\text{adj} < 0.05$) at each time point (Figure 1B; Supplementary Table 2). In *P. patens*
254 and *Z. mays*, the highest number of DE transcripts was found at 24 hpi - 30% and 43% of all
255 transcripts, respectively (Figure 1B; Supplementary Table 2). For *S. tuberosum* the largest number
256 of DE transcripts was found at 12 hpi (4.7% of overall transcripts), and for *B. napus* - at 96 hpi
257 (22.4% of overall transcripts) (Figure 1B; Supplementary Table 2).

258 We additionally identified lncRNAs in assembled transcriptomes that could potentially
259 encode for unannotated peptide precursors (Hsu & Benfey, 2018). To estimate the proportion of
260 lncRNAs in differentially regulated transcripts, we selected the assembled transcripts that 1) do
261 not code for annotated proteins; 2) do not assign to known protein domains and SSP families; 3)
262 contain smORFs below 100 codons; and 4) do not overlap known genes on the same strand. We
263 annotated such transcripts using the lncRNA database PLncDB
264 (https://www.tobaccodb.org/plncdb/). At each time point, lncRNA transcripts accounted for
265 approximately 0.17–1.05% of differentially regulated transcripts in each species, with *S.*
266 *tuberosum* having a higher percentage and *P. patens* having a lower one (Figure 1B;
267 Supplementary Table 2). We suggest that it can be the result of both biological and technical
268 reasons.

269 For further analysis, we selected 13, 19, 15, and 28 DE SSP families from *P. patens*, *Z.*
270 *mays*, *S. tuberosum,* and *B. napus*, respectively (22, 45, 25 and 103 DE transcripts, respectively;
271 Figure 1B; Supplementary Table 3). Moreover, we additionally selected DE unannotated proteins
272 (≤ 200 aa), containing N-terminal signals (Figure 1B; Supplementary Table 2). This set of DE
273 proteins was used for the identification of novel, previously unannotated peptide families. Some
274 of them were located on predicted long noncoding RNAs (Figure 1B; Supplementary Table 2).

275 3.2 Transcriptional patterns of SSP genes in different plant 276 lineages

277 We next analyzed the transcriptional patterns of different groups of SSPs in response to
278 phytopathogen infection. The first group consists of selected known phytocytokines. The second
279 group included "peptide hormones," which have been demonstrated to control growth and
280 development but not stress responses. The third group consists of known families of AMPs.

281 3.2.1 The transcriptional regulation of phytocytokines

282 The most common differentially regulated phytocytokines in all examined species were
283 from conserved phytocytokine families such as PSK, CLE, PSY, and RALF. The transcripts that
284 encode RALF peptides were identified in all examined species at three time points. The members
285 of this family were differentially expressed under infection in *P. patens*, *B. napus* and *Z. mays*. In
286 *P. patens*, the PpRALF1 (Pp3c3_15280) gene was significantly upregulated under infection (Table
287 2; Supplementary Table 3). Some members of the RALF family were significantly downregulated
288 under infection in *B. napus* and *Z. mays* (e.g., CDX78035 in *B. napus* and Zm00001d033941_T001
289 in *Z. mays*; Table 2).

290 **Table 2.** Groups of differentially expressed gene families predicted as known SSPs involved in
291 immunity

SSP families	Up-regulated	Down-regulated

Phytocytokines	***Negative regulators***	**RALF**	*P. patens* *B. napus*	*Z. mays* *B. napus*
		PSK	*Z. mays* *B. napus* *S. tuberosum*	*Z. mays* *B. napus*
		PSY	*Z. mays* *B. napus*	*Z. mays* *B. napus* *S. tuberosum*
	SGP-rich peptides	**CLE**	*B. napus* *P. patens*	*B. napus*
		PIP	*B. napus*	
		CEP	*B. napus*	*B. napus*
		IDA	*B. napus*	
		SCOOP	*B. napus*	*B. napus*
		GLV		*B. napus* *S. tuberosum*
Defense peptides	***From functional precursor***	**CAPE**	*P. patens* *Z. mays* *B. napus*	*P. patens*
	AMPs	**nsLTP**	*Z. mays* *B. napus* *S. tuberosum* *P. patens*	*Z. mays* *B. napus* *S. tuberosum* *P. patens*
		Hevein-like	*Z. mays* *B. napus* *S. tuberosum*	
		Thionin	*B. napus* *S. tuberosum*	
Unclear functions in immunity		**TAX**	*P. patens* *Z. mays* *B. napus*	*S. tuberosum*
		MEG	*Z. mays*	*P. patens* *B. napus* *Z. mays*
		EPFL		*P. patens*

Previously, PSK peptides were shown to regulate the immune response to biotrophic and necrotrophic pathogens in an antagonistic manner in Arabidopsis (Y. Li et al., 2024). The genes belonging to the PSK family were differentially regulated in vascular plants but not in *P. patens*. In Z. *mays*, PHYTOSULFOKINES 6-RELATED gene (Zm00001d023385_T001) was upregulated during infection. In addition, another phytosulfokine gene (Zm00001d023384_T001) was up- and downregulated at different hpi, and Zm00001d048006_T001 was downregulated. In

298 *B. napus*, we identified 22 upregulated transcripts of PSK genes, and two genes, CDX68551 and
299 CDY36518, were downregulated. One PSK gene (Soltu.DM.02G027440.2) was upregulated in *S.*
300 *tuberosum* (Table 2; Supplementary Table 3).

301 The PSY family was recently shown to play a role in regulating the trade-off between plant
302 growth and stress response in vascular plants (Ogawa-Ohnishi et al., 2022). PSY receptors
303 (PSYRs) activate the expression of various genes encoding stress response transcription factors
304 upon depletion of the ligands (Ogawa-Ohnishi et al., 2022). We found that PSY transcripts were
305 significantly downregulated in all vascular plants under infection: MSTRG.18847.1 in *S.*
306 *tuberosum*; CDY71288, CDX80347, and MSTRG.42712.1 in *B. napus*; 14 transcripts in *Z. mays*
307 (Table 2; Supplementary Table 3). Some PSY genes in *Z. mays* and *B. napus* were also
308 upregulated: Zm00001d038523_T001, Zm00001d020443_T001, MSTRG.18595.2 and
309 MSTRG.17209.2 in *Z. mays*; CDY64688, CDY35659, CDY69338 and MSTRG.43937.2 in *B.*
310 *napus* (Table 2; Supplementary Table 3). However, our analysis did not confirm the expression of
311 the predicted PSY gene in P. *patens*.

312 The analysis of phytocytokines specific to vascular plants identified PIP (7 upregulated
313 transcripts), CEP (12 upregulated transcripts), IDA (CDY08563), and SCOOP (three upregulated
314 and one downregulated) peptides in *B. napus* (Table 2; Supplementary Table 3). Transcripts of
315 IDA peptide precursors were detected in *S. tuberosum* plants, but their transcription did not show
316 significant changes under infection.

317 Transcripts of PEP precursors (PROPEPs) were not detected in our analysis across the
318 examined plant species, likely due to their low transcription levels.

319 Thus, our analysis showed that only the RALF peptide family was differentially regulated
320 under infection in vascular and non-vascular plants.

321 ### 3.2.2 The transcriptional patterns of known peptide growth hormones

322 The number of SSP families that were shown to regulate only growth and development
323 processes is limited. For example, cysteine-rich EPIDERMAL PATTERNING FACTOR (EPF)
324 and EPF-LIKE (EPFL) peptide families play a role in regulating plant growth and development
325 (Y. He et al., 2023; Xiong et al., 2022). Recently, the EPFL family was found to be significantly
326 downregulated upon *F. oxisporum* infection and elicitor treatments in tomato (Slezina et al., 2021).
327 Our analysis revealed transcripts encoding members of the EPFL family in all angiosperms,
328 although their transcription levels remained unchanged during infection. However, we found that
329 some members of the EPFL family were significantly downregulated in the moss *P. patens*.
330 According to our analysis, 14 transcripts belonging to the EPFL peptide family were identified in
331 *P. patens*. From them, nine transcripts were significantly downregulated (six annotated genes:
332 Pp3c6_12270, Pp3c23_5720, Pp3c24_9860, Pp3c23_11350, Pp3c1_26030, Pp3c6_27020; Table
333 2; Supplementary Table 3). The function of this peptide family in bryophytes is currently
334 unknown. Thus, these results suggest that moss EPFL peptides may be involved in the regulation
335 of the immune response, as has been recently shown in tomato (Slezina et al., 2021).

336 The genes encoding members of the cysteine-rich TAX (TAXIMIN) family peptide were
337 significantly upregulated in *Z. mays* (four transcripts, three isoforms of Zm00001d008923_T001
338 and one not annotated MSTRG.18593.1), *B. napus* (seven transcripts, two annotated: CDY31955
339 and CDY19064), *P. patens* (Pp3c4_17790) and downregulated in *S. tuberosum* (MSTRG.21268.1;
340 Table 2; Supplementary Table 3). This family was suggested to play a role in nutrient-status-
341 related signaling (de Bang et al., 2017).

342 Another peptide family of cysteine-rich peptides, MEG (Maternally Expressed Gene), was
343 previously shown to be downregulated in tomato plants under infection (Slezina et al., 2021). We
344 found that members of this family were significantly downregulated in *P. patens* (two isoforms of
345 Pp3c5_6640) and *B. napus* (MSTRG.22520.1) and some members were significantly up- and
346 down-regulated in *Z. mays* (three up-regulated transcripts: two isoforms of
347 Zm00001d005037_T001 and Zm00001d052375_T002; two down-regulated transcripts:
348 Zm00001d052375_T002 and Zm00001d042365_T001; Table 2; Supplementary Table 3).

349 TPD 1 (TAPETUM DETERMINANT 1) is a secreted cysteine-rich small protein that was
350 shown to control anther cell differentiation in Arabidopsis (Huang et al., 2016). We identified two
351 TPD transcripts in the moss *P. patens* (STRG.20746.1, STRG.20746.2) and three of the five
352 transcripts encoded TPD-like peptides in *Z. mays* (two of them already annotated
353 (Zm00001d049823_T001, Zm00001d006677_T001) were significantly up-regulated
354 (Supplementary Table 3).

355 Thus, our analysis shows that cysteine-rich peptides, which were not considered to play a
356 role in regulating immune responses, were actively regulated during infection in all plant lineages.
357 Our findings show that the widely conserved RALF and EPFL peptide families may play a role in
358 the immune response in both vascular and non-vascular plants.

359 3.2.3 The transcriptional regulation of antimicrobial SSPs

360 We identified several different SSPs classified as antimicrobial and protease inhibitors in
361 the analyzed species, such as Gibberellic acid stimulated in Arabidopsis (GASA) from Snakin
362 family, ARACIN, Potato type II proteinase inhibitor, SubIn and others. The most known
363 antimicrobial peptides, hevein-like and thionins, were identified in the transcriptomes of all
364 examined species (Supplementary Table 3). However, hevein-like transcripts were significantly
365 upregulated during infection only in vascular plants (Table 2; Supplementary Table 3). Thionins
366 were significantly upregulated in *B. napus* and *S. tuberosum*, but we did not identify differential
367 expression in other species.

368 The non-specific lipid transfer proteins (nsLTPs) are a plant-specific superfamily of
369 cysteine-rich AMPs (Santos-Silva et al., 2023). The transcripts of nsLTP genes were identified in
370 all analyzed plant species and were the most highly represented group of transcripts, both up- and
371 down-regulated (Supplementary Table 3). Our findings showed that nsLTPs have similar patterns
372 of regulation in response to pathogen attack in vascular and non-vascular plants (Table 2).

373 3.3 Comparative analysis of SSP motifs

374 We next compared the peptide motifs of bryophyte and angiosperm SSPs differentially
375 regulated in the transcriptomes of the examined species. The SSPs were selected based on the
376 following criteria: 1) differentially expressed in *P. patens* and in at least one angiosperm species;
377 or 2) differentially expressed in *P. patens*. We used only functional peptide sequences for this
378 analysis (Figure 2A; Supplementary Figure 1).

379 The EPFL peptides identified in *P. patens*, *S. tuberosum*, and *B. napus* revealed a typical
380 cysteine pattern as well as a conserved GS region (Figure 2A, Supplementary Figure 1A), which
381 is similar to the previous studies (Takata et al., 2013; Xiong et al., 2022).

382 The YISY motif, which is shown to play a role in receptor binding (R. Zhang et al., 2023),
383 was identified in most RALF peptides from *P. patens*, *S. tuberosum,* and *B. napus*. *Z. mays*
384 peptides contain an ISY motif (Figure 2A, Supplementary Figure 1B). The RGC region was found
385 in all studied plants, suggesting that RALF peptides share known motifs in all examined species
386 (R. Zhang et al., 2023).

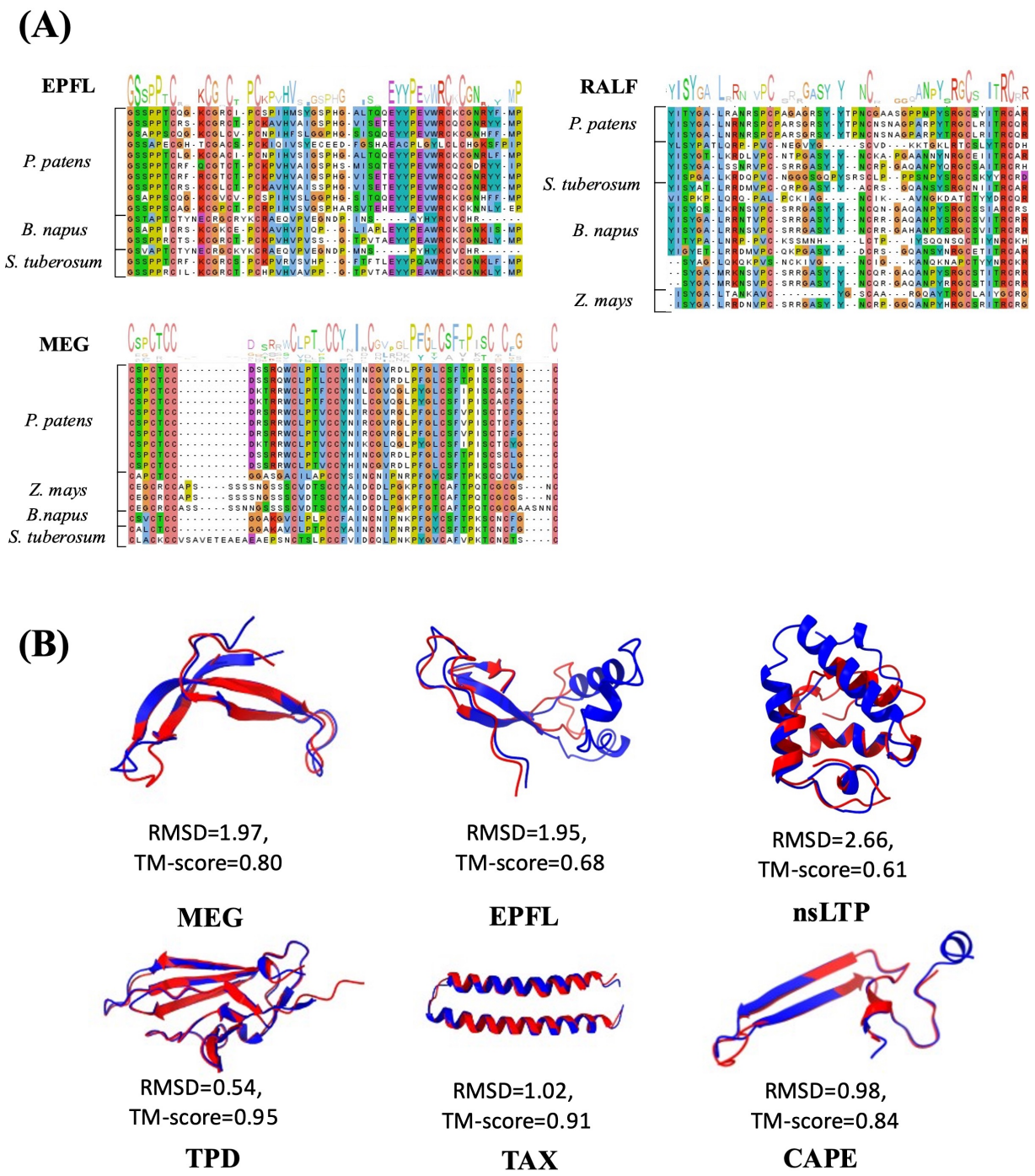


Figure 2. (A) - Sequence similarity analysis for peptide families in *Physcomitrium patens* and angiosperms; (**B) -** Structural comparison of angiosperm and *Physcomitrium patens* peptide predicted structures.

391 Previously, MEG peptides with 8 and 12 cysteine residues were identified in angiosperms
392 (Gutiérrez-Marcos et al., 2004; Slezina et al., 2021). We found that MEG peptides from *P.*
393 *patens*,*Z. mays*, *S. tuberosum,* and *B. napus* contain 12 conserved cysteine residues, including two
394 CC regions (Figure 2A, Supplementary Figure 1C). In contrast to the cysteine patterns, the MEG
395 amino acid sequences showed significant variation among the examined sequences.

396 Thus, our analysis showed that EPFL peptides are the most conserved among all the plants
397 analyzed. The largest difference in peptide motifs was observed in MEG peptides, including the
398 number of cysteine residues and overall amino acid composition.

399 3.4 Structure-based comparison of the identified SSPs

400 3.4.1 The prediction of peptide structures using AlphaFold2

401 The accuracy of sequence-based methods for predicting peptide structures has been
402 demonstrated in recent years (McDonald et al., 2023). Moreover, current methods for structure
403 comparison allow the detection of distant orthologs (Monzon et al., 2022). Therefore, we next
404 compared predicted structures of EPFL, RALF, CLE, MEG, nsLTP, CAPE, TPD, and TAX
405 peptides from different species. For this, 135 peptide structures were predicted by AlphaFold2
406 (Jumper et al., 2021), and structures with an average pLDDT > 70 were selected for further
407 analysis. Peptide structures of *P. patens* were compared with those of angiosperms. Structural
408 alignment was performed using the TM-align tool, and metrics including sequence identity
409 (Seq_Id), Root Mean Square Deviation (RMSD), and TM-score were calculated (Supplementary
410 Table 4). Because clear secondary structures were not predicted for RALF and CLE peptide, they
411 were excluded from further analysis. Overall, 21 peptide structures of MEG, EPFL, CAPE, TPD,
412 TAX, and nsLTP peptides with the highest average pLDDT per family for each plant species were
413 selected for structural comparison (Supplementary Table 4, Supplementary Materials).

414 The predicted structures of the MEG peptides consist of multiple beta-sheets. The highest
415 similarity to *P. patens* MEG structures was observed in the case of *Z. mays* MEG (TM-score 0.80;

416 Figure 2B), which is in line with our motif comparison. The lowest structural similarity was found
417 between *P. patens* and *S. tuberosum* MEG structures (TM-score 0.44, Supplementary Table 4).
418 The MEG structure of *Z. mays* contains an additional loop composed of alpha-helices, which is
419 absent in the predicted MEG structure of *P. patens*.

420 The structure of EPFL peptides comprises two antiparallel beta-sheets connected by a loop.
421 The structures of EPFL peptides from *Z. mays* and *P. patens* showed the highest similarity (TM-
422 score 0.68; Figure 2A) compared to other analyzed plants (TM-scores 0.55 for *B. napus* and *S.*
423 *tuberosum;* Supplementary Table 4).

424 The predicted structure of non-specific lipid transport proteins (nsLTPs) contains four
425 alpha-helices. *B. napus* nsLTP peptide structure revealed the highest similarity to *P. patens* ndLTP
426 structures (Figure 2B). TM-scores for *Z. mays* and *S. tuberosum* structures compared to *P. patens*
427 were 0.50 and 0.56, respectively (Supplementary Table 4).

428 The predicted structures of the TPD and TAX peptides showed high structural similarity
429 across all examined species (Supplementary Table 4). The highest similarity between the predicted
430 structures of *P. patens* and *B. napus* peptides was observed (Figure 2B).

431 The predicted structures of the CAPE peptides contain two beta-sheets and an alpha-helix.
432 The highest similarity to *P. patens* CAPE was observed between P. *patens* and *B. napus* structures
433 (Figure 2B).

434 Therefore, the cysteine-rich peptides (MEG, EPFL) from *P. patens* showed the highest
435 similarity to *Z. mays* homologs. The structures of CAPE, nsLTP, TAX, and TPD peptides from *P.*
436 *patens* were the most similar to the predicted structures from *B. napus*. The TPD peptide has the
437 most conserved structure among *P. patens, Z. mays,* and *B. napus* species.

438 3.4.2 The search for structural homologs of identified SSPs

439 Then, we searched for structural homologs of *P. patens* peptide families MEG, EPFL,
440 CAPE, nsLTP, TAX, and TPD against AlphaFold Protein Structure Database

441 (https://alphafold.ebi.ac.uk) and PDB (https://www.rcsb.org) databases by Foldseek (van Kempen
442 et al., 2024).

443 The analysis of MEG peptides yielded three significant hits (TM-score > 0.5, Probability
444 > 0.8) against the AlphaFold Protein Structure Database (AFDB) with CRP7-Cysteine-rich family
445 protein from *Zea mays* (Supplementary Figure 2).

446 For EPFL peptides, structural similarity to ~70 EPFL proteins from *Oryza sativa, Glycine*
447 *max, Arabidopsis thaliana, Nicotiana tabacum, etc.* was identified. Hits to structures from the PDB
448 database were also obtained, such as EPFL and plant receptor ERL1-TMM in complex with EPF1
449 from *Arabidopsis thaliana* (Accessions in PDB database: 2liy_A, 5xkj_F, 5xjo_E; Supplementary
450 Figure 2). We also found hits to the structures of nsLTP, TPD, TAX, and CAPE peptides in the
451 AFDB and PDB databases (Supplementary Figure 2).

452 Thus, our analysis revealed high structural similarity among MEG, EPFL, CAPE, nsLTP,
453 TAX, and TPD peptides from different plant taxonomic groups. For CAPE and EPFL precursors,
454 hits to PDB structures were observed. Therefore, our results confirmed that searches by Foldseek
455 (van Kempen et al., 2024) can be used to find peptide homologs at the structural level.

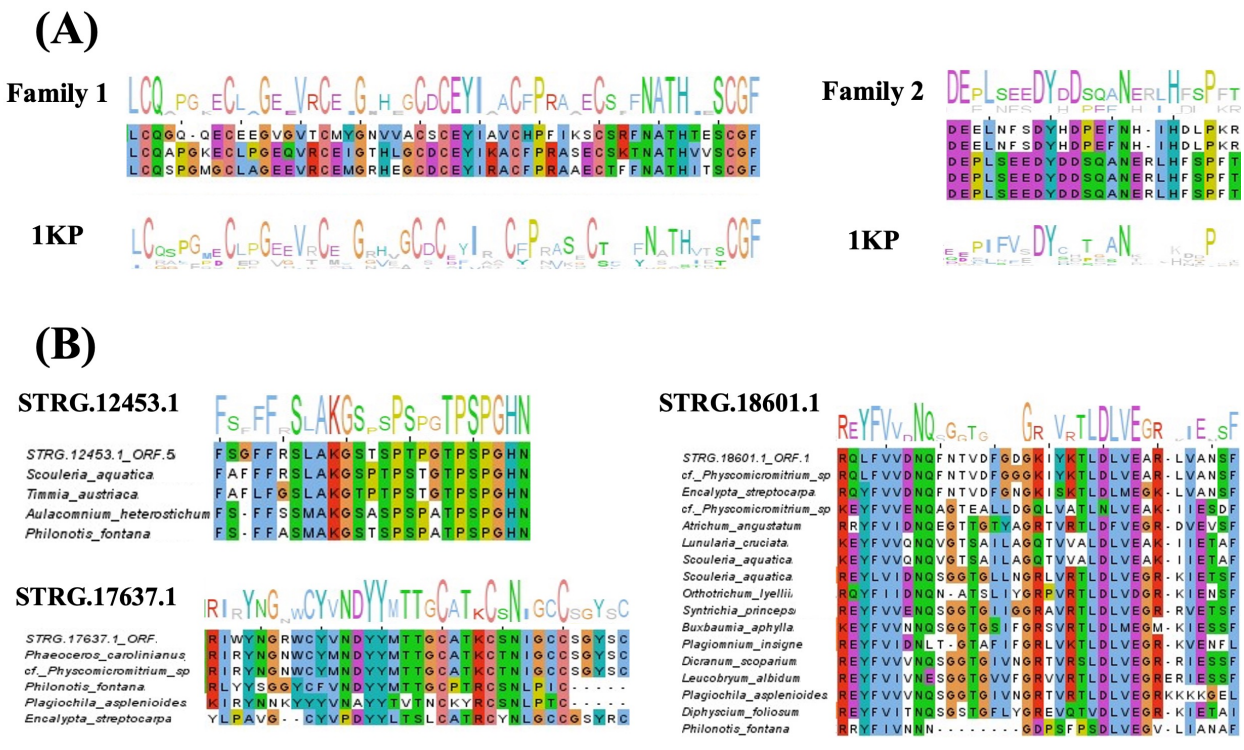
456 3.5 Identification of novel families of SSPs in the moss *P. patens*

457 Novel peptide families have emerged during the evolution of land plants (Furumizu &
458 Shinohara, 2024). However, the identification of taxon-specific peptides and novel peptide
459 families is a challenging task. Moreover, lineage-specific peptide families have been identified
460 only in angiosperms to date. Therefore, we next sought to search for novel SSPs involved in the
461 immune response in the non-vascular plant, *P. patens,* using our custom pipeline.

462 At first, we selected 356 ORFs below 200 aa located on DE transcripts that contained a
463 predicted N-terminus signal sequence but did not contain a transmembrane domain or known
464 peptide motif (Figure 1). We additionally filtered out "candidate" SSPs that have protein
465 annotation or have hits in the NCBI non-redundant protein sequence database (BLASTP, E-value

< 0.01, organism: *Viridiplantae*). We next searched for possible orthologs of 58 "candidate" SSPs in the 1000 plant transcriptome (1KP) database (https://db.cngb.org/datamart/plant/DATApla4/) using TBLASTN (E-value < 0.01). Then, we manually checked the presence of possible peptide motifs based on multiple sequence alignments (MSAs).

Based on the analysis of possible paralogs in the *P. patens* genome, we identified two families of putative novel SSPs in *P. patens*. The consensus sequences of the identified motifs are presented in Figure 3A (Supplementary Figure 3). Family 1 comprises three *P. patens* genes (*Pp3c2_4250*, *Pp3c14_17750*, and *Pp3c21_20040*) that have not been functionally annotated yet. One of these genes, *Pp3c21_20040*, was differentially regulated upon *B. cinerea* infection (Figure 3A, Supplementary Figure 3A). The members of Family 1 are cysteine-rich and comprise 8 conservative cysteine residues as well as conserved "NATH" and "SCGF" motifs (Figure 3A). The TBLASTN search in the 1KP database identified possible orthologs in 49 species, including mosses, liverworts, and lycophytes (Figure 3A, Supplementary Figure 4). The median length of predicted ORFs was 119 aa.



481 **Figure 3**. (**A) -** Alignment logos of novel candidate families found in *P. patens* (**Family 1:**
482 STRG.19963.1_ORF.1; STRG.1695.1_ORF.2; STRG.14124.1_ORF.4; alignment logo with 1KP;
483 **Family 2:** STRG.4116.1_ORF.1; STRG.4116.2_ORF.1; STRG.4476.2_ORF.1;
484 STRG.4476.3_ORF.1; STRG.4473.1_ORF.1; alignment logo with 1KP); (**B) -** Alignment logos
485 of novel peptide candidates corresponding to unique *P. patens* ORFs with 1KP
486 (STRG.12453.1_ORF.5; STRG.18601.1_ORF.1; STRG.17637.1_ORF.1).

487 The putative peptide precursors of Family 2 shared two similar peptide motifs:
488 DY*D***N**H**P and DY*D***N***H**P (Figure 3A, Supplementary Figure 3B). These
489 motifs are similar to the known PSY peptide motif DY*(D)(P)**N**H*P (Furumizu et al., 2021),
490 but differ in the position of conserved residues H and P. Previously, one *P. patens* gene
491 (Pp3c8_3200) was identified as a precursor of PSY peptide (Furumizu et al., 2021; Tost et al.,
492 2021). Two predicted protein precursors from Family 2, STRG.4116.1_ORF.1 (117 aa) and
493 STRG.4116.2_ORF.1 (91aa), shared the DY*D***N**H**P motif at the C-end and came from
494 the *Pp3c3_33940V3.3* gene that was erroneously annotated as alternative splicing factor
495 SRp20/9G8 in Phytozome 13. The STRG.4116.1 transcript was differentially upregulated upon
496 infection with *B. cinerea*. The other three predicted precursors contain the motif
497 DY*D***N***H**P. Two of them, STRG.4476.2_ORF.1 (114 aa) and STRG.4476.3_ORF.1
498 (118 aa), belonged to the gene *Pp3c4_6140*. The assembled transcript STRG.4473.1 intersects the
499 *P. patens* gene *Pp3c4_6040* and encodes 118 aa precursor STRG.4473.1_ORF.1. The TBLASTN
500 analysis of the corresponding peptide sequences against 1000 Plant transcriptomes identified 20
501 possible orthologs of Family 2 in mosses and one liverwort species (Supplementary Figure 5). The
502 median length of predicted ORFs was 110 aa which is similar to known PSY genes. Thus, we
503 identified a novel family of immune-responsive PSY-like (PSYL) peptides that are encoded by
504 three *P. patens* loci.

505 Because SSPs can be presented by only one copy in the genomes of non-vascular plants,
506 we additionally analyzed putative SSP precursors without possible paralogs in *P. patens*. We
507 searched for their possible homologs against transcripts from the 1KP database. The resulting
508 alignments of the three most interesting up-regulated peptide candidates are presented in Figure

509 3B. Full alignments can be found in Supplementary Figure 6. These putative peptide precursors
510 contain motifs that are conserved among different plant species. In the case of
511 STRG.12453.1_ORF.5 (140aa), P***TPSPGHN motif is observed, which is similar to CLE and
512 IDA peptide motifs (Furumizu et al., 2021). This candidate was found in four moss species (Figure
513 3B, Supplementary Figure 6A). Candidate STRG.18601.1_ORF.1 (115aa) contains a conserved
514 LDLVE motif, which is observed in 13 other species, including mosses and liverworts (Figure 3A,
515 Supplementary Figure 6B). Finally, *P. patens* candidate STRG.17637.1_ORF.1 (121aa) is
516 cysteine-rich and, notably, contains a conserved DYY region. Hits to four other species were
517 found, including moss, liverwort, and hornwort plants (Figure 3B, Supplementary Figure 6C).
518 Further study is needed to explore the functions of these SSP candidates.

519 Thus, we propose an approach to identify novel, not annotated, plant peptide families. This
520 approach focuses on the identification of putative SSP precursors (based on length, N-terminal
521 signal sequence, no TM-domain) combined with analysis of possible peptide motifs. Overall, we
522 identified 5 potential candidates for novel signaling peptides in *P. patens*. Two of them comprise
523 families based on sequence similarity, and some of them are unique ORFs predicted in *P. patens.*
524 Nevertheless, all potential novel candidates are found in other plant species belonging to non-
525 vascular plants. For the two candidates, distant homology to known peptide families (PSY, IDA,
526 and CLE) is suspected.

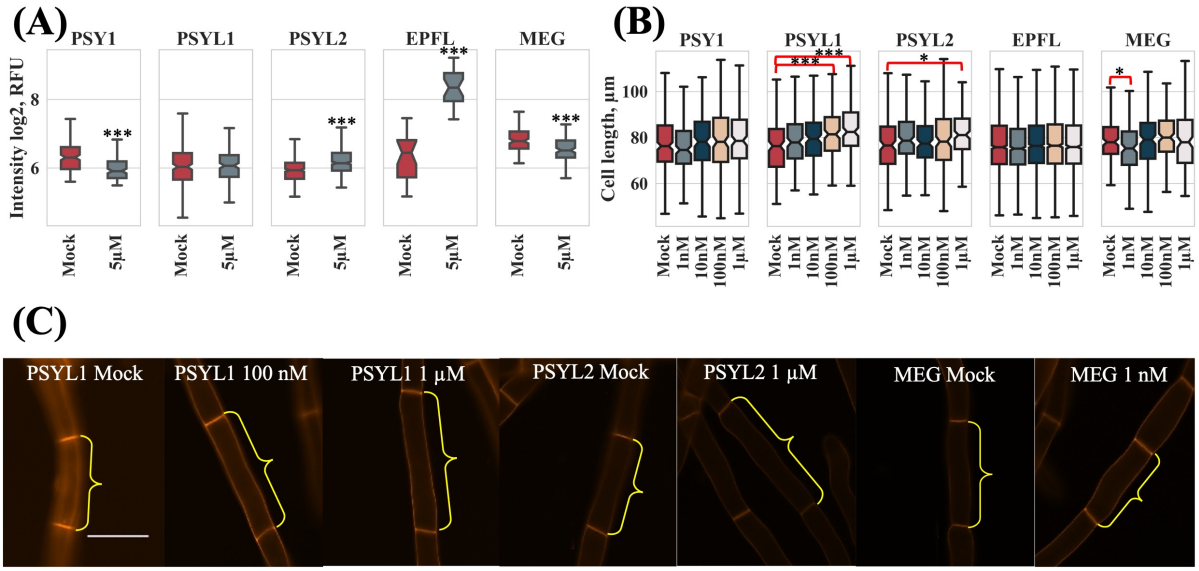
527 3.6 Experimental validation of potential immune peptides

528 We next synthesized and assessed the biological activity of the moss PSY, PSYL1, PSYL2,
529 PpEPFL, and PpMEG peptides. At first, we used reactive oxygen species (ROS) accumulation as
530 a marker of the immune response. Treatment with 5 μM of synthetic PpPSYL2 and PpEPFL
531 induced a significant accumulation of ROS molecules in *P. patens* cells compared to the mock
532 (Mann-Whitney U test: P-value < 0.0005; Figure 4A). In contrast, treatment with syntenic PpPSY1
533 and PpMEG resulted in a significant decrease in the accumulation of ROS in moss cells, which
534 may indicate their role as negative regulators of the immune response (Mann-Whitney U test: P-

535 value < 0.0005; Figure 4A). We did not observe any significant difference in the ROS
536 accumulation between the mock and PSYL1 peptide treatments (Figure 4A).

537 We also explored the effects of peptide treatment on *P. patens* growth. The moss
538 protonemata were grown on a semi-solid medium supplemented with either a corresponding
539 peptide or a mock treatment for one week. A significant difference in cell length was found
540 between mock and treatment with PSYL1 (100 nM and 1 µM; Tukey's test: P-value < 0.001),
541 PSYL2 (1 µM), and MEG (1 nM) peptides (Tukey's test: P-value < 0.05; Figure 4B and C).

542 Thus, our experiments confirmed the biological activity of differentially regulated SSPs,
543 including previously unknown PSY-like peptides.



544

545 **Figure 4.** (**A**) **-** Analysis of the accumulation of intracellular ROS molecules using a
546 fluorescent dye DCFH-DA after treatment with 5 µM PpPSY1, PpPSYL1, PpPAYL2, PpEPFL
547 and PpMEG in the protonema of moss *P. patens* (Mann-Whitney U test: ***P-value < 0.001).
548 Control samples (C) were treated with water. **(B) -** Measurement of the length of protonema cells
549 grown in a semi-solid medium containing PpPSY1, PpPSYL1, PpPSYL2, PpMEG, PpEPFL
550 (Tukey's test: ***P-value < 0.001; *P-value < 0.05). (**C**) **-** Microscope images showing moss *P.*
551 *patens* protonemata cells stained with propidium iodide. Bar = 50 µm. All experiments were
552 performed in at least three replicates.

4 Discussion

Recent studies showed that peptides derived from protein precursors play an important role in plant immunity, modulating the defense response. The plant cell-surface receptors that recognize these peptide ligands contain Leucine-Rich Repeat (LRR) or Malectin ectodomains and have an ancient origin (Ngou et al., 2024). Contrary to immune receptors, the origin and evolution of immune peptides are still poorly understood (Furumizu et al., 2021). One of the reasons is that, due to rapid changes, they appear to quickly enter the "twilight zone", where sequence similarity is no longer detectable. Nevertheless, several databases that include plant Small Secreted Peptides (SSPs) were recently published (Boschiero et al., 2020; Ghorbani et al., 2015; Pan et al., 2013). Sequence similarity, however, does not necessarily mean functional similarity across different plant phyla. In this study, we compared pathogen-induced transcriptomes in vascular and non-vascular plants to understand how SSP families from different plant lineages respond to infection. Along with conserved SSPs associated with immune responses that were found in the transcriptomes of both vascular and non-vascular plants, we observed species-specific regulation of certain SSP families.

Our analysis shows that cysteine-rich peptides were actively regulated during infection in all plant lineages. The EPFL, RALF, TAX, and MEG were differentially regulated in *P. patens* and angiosperm species. The widely conserved family of EPIDERMAL PATTERNING FACTOR (EPF) peptides is involved in regulating plant growth or development in vascular and non-vascular land plants (Caine et al., 2016; Y. He et al., 2023). The nine EPFL genes that are upregulated in the developing sporophyte have been identified in previous bioinformatics studies of the *P. patens* genome (Caine et al., 2016; Lyapina et al., 2021; Takata et al., 2013), but their functions remain unknown. However, the role of EPF and EPF-LIKE (EPFL) peptides as positive or negative regulators of immune response has not been shown to date. Nevertheless, several genes of the EPFL family were shown to be significantly downregulated upon pathogen infection in tomato (Slezina et al., 2021). In this study, we found that the transcriptional level of six EPFL genes was significantly downregulated in *P. patens*. This result can be explained by a specific response to pathogens or by the inhibition of growth processes under stress conditions.

581 Our treatment of moss protonemata with synthetic EPFL peptide resulted in an increase in
582 intracellular reactive oxygen species (ROS) levels, suggesting the activation of downstream
583 signaling. In Arabidopsis, the receptor-like kinase (RLK) ERECTA and its two closely related
584 homologs, ERECTA-LIKE 1 (ERL1) and ERL2 are responsible for perceiving the EPF/EPFL
585 family peptides (Y. He et al., 2023; Tameshige et al., 2017). In the genome of *P. patens,* six
586 potential orthologues of the ERECTA genes were found (Villagarcia et al., 2012). Thus, the EPFL
587 family might be a target for future work on understanding the evolution of phytocytokines in land
588 plants.

589 The RALF peptides were shown to be involved in immune response across different plant
590 lineages (Y.-H. He et al., 2023; Mamaeva et al., 2023; Slezina et al., 2021; Wolf & Höfte, 2014).
591 We found both up-regulated and down-regulated RALF genes in pathogen-induced
592 transcriptomes, which is consistent with previous results showing different roles of different
593 members of RALFs in modulating plant immunity (Y.-H. He et al., 2023; L. Liu et al., 2024;
594 Mamaeva et al., 2023; Stegmann et al., 2017).

595 The poorly characterized cysteine-rich peptides, such as TAX and MEG, were
596 differentially regulated in both vascular and non-vascular plant species. However, how TAX
597 peptides are involved in the immune response is currently unknown. Previously, tomato MEGs
598 were found to be downregulated in response to infection (Slezina et al., 2021). This is similar to
599 our findings and suggests the conservative role of this SSP family in vascular and non-vascular
600 land plants. Nevertheless, the role of MEG peptides in plant growth and defense is poorly
601 understood.

602 In addition to common approaches for the identification of SSPs based on sequence
603 similarity (Boschiero et al., 2020), we developed a new pipeline to identify novel lineage-specific
604 secreted peptides. Our pipeline resulted in the identification of two putative SSP families in *P.*
605 *patens*. One of these families (Family 2) consists of PSY-like peptides, which are similar to the
606 known phytocytokine PSY family in angiosperms. The PSY peptide family was present in the
607 most recent common ancestor of vascular plants and bryophytes (Furumizu & Shinohara, 2024).

608 Previously, one gene potentially encoding the PSY peptide and another encoding the cognate
609 PSYR receptor were identified in *P. patens* (Furumizu et al., 2021; Furumizu & Aalen, 2023).
610 Thus, we significantly expanded the family of PSY-like peptides in bryophytes. The identification
611 of previously unknown PSY-like peptides in bryophyte genomes could expand our understanding
612 of the peptide evolution in land plants.

613 In conclusion, our comprehensive analysis of small secreted peptides in the pathogen-
614 induced transcriptomes of four plant species showed the conservation of cysteine-rich peptides
615 across different plant lineages. We found high motif and structural conservation of these peptides
616 in vascular and non-vascular plants. In addition, we identified previously uncharacterized putative
617 SSP families in bryophytes. One of these bryophyte-specific families shows similarity to
618 previously characterized PSY peptides. Thus, our study provides a new avenue to explore the
619 evolution of known and previously uncharacterized peptide families.

620 Author contributions

621 IF supervised the study. IL, DG conducted all experiments and performed all bioinformatic
622 analyses. AM performed ROS measurement and cell length analysis. EAR, DYR and EVR
623 synthesized recombinant moss *P. patens* peptides. VL and SEA synthesized moss *P. patens* PSY
624 family peptides. IL, DG, and IF wrote the manuscript. All authors contributed to the article and
625 approved the submitted version.

626 Acknowledgment

627 We thank the Center for Precision Genome Editing and Genetic Technologies for Biomedicine,
628 Lopukhin Federal Research and Clinical Center of Physical-Chemical Medicine of Federal
629 Medical Biological Agency for the expertise and guidance in genetic engineering.

630 Funding

This work was supported by Russian Science Foundation (project no. 23-74-10048).

Data availability statement

The data that support the findings were derived from the NCBI Sequence Reads Archive (SRA) available in the public domain with the following accession numbers: SRP274010, SRP390856, SRP053361, SRP066006.

References

Andrews, S. (2010). *FastQC: A Quality Control tool for High Throughput Sequence Data*. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics (Oxford, England)*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Boschiero, C., Dai, X., Lundquist, P. K., Roy, S., Bang, T. C. de, Zhang, S., Zhuang, Z., Torres-Jerez, I., Udvardi, M. K., Scheible, W.-R., & Zhao, P. X. (2020). MtSSPdb: The Medicago truncatula Small Secreted Peptide Database. *Plant Physiology*, *183*(1), 399–413. https://doi.org/10.1104/pp.19.01088

Bowman, J. L. (2022). The origin of a land flora. *Nature Plants*, *8*(12), 1352–1369. https://doi.org/10.1038/s41477-022-01283-y

Bowman, J. L., Kohchi, T., Yamato, K. T., Jenkins, J., Shu, S., Ishizaki, K., Yamaoka, S., Nishihama, R., Nakamura, Y., Berger, F., Adam, C., Aki, S. S., Althoff, F., Araki, T., Arteaga-Vazquez, M. A., Balasubrmanian, S., Barry, K., Bauer, D., Boehm, C. R., … Schmutz, J. (2017). Insights into Land Plant Evolution Garnered from the Marchantia polymorpha Genome. *Cell*, *171*(2), 287-304.e15. https://doi.org/10.1016/j.cell.2017.09.030

Caine, R. S., Chater, C. C., Kamisugi, Y., Cuming, A. C., Beerling, D. J., Gray, J. E., & Fleming, A. J. (2016). An ancestral stomatal patterning module revealed in the non-vascular land plant Physcomitrella patens. *Development*, *143*(18), 3306–3314. https://doi.org/10.1242/dev.135038

de Bang, T. C., Lay, K. S., Scheible, W.-R., & Takahashi, H. (2017). Small peptide signaling pathways modulating macronutrient utilization in plants. *Current Opinion in Plant Biology*, *39*, 31–39. https://doi.org/10.1016/j.pbi.2017.05.005

Fesenko, I., Shabalina, S. A., Mamaeva, A., Knyazev, A., Glushkevich, A., Lyapina, I., Ziganshin, R., Kovalchuk, S., Kharlampieva, D., Lazarev, V., Taliansky, M., & Koonin, E. V. (2021). A vast pool of lineage-specific microproteins encoded by long non-coding RNAs in plants. *Nucleic Acids Research*, *49*(18), 10328–10346. https://doi.org/10.1093/nar/gkab816

Furumizu, C., & Aalen, R. B. (2023). Peptide signaling through leucine-rich repeat receptor kinases: Insight into land plant evolution. *New Phytologist*, *238*(3), 977–982. https://doi.org/10.1111/nph.18827

Furumizu, C., Krabberød, A. K., Hammerstad, M., Alling, R. M., Wildhagen, M., Sawa, S., & Aalen, R. B. (2021). The sequenced genomes of nonflowering land plants reveal the innovative evolutionary history of peptide signaling. *The Plant Cell*, *33*(9), 2915–2934. https://doi.org/10.1093/plcell/koab173

Furumizu, C., & Shinohara, H. (2024). Land plant peptide signaling: What we know—and don't know—about its evolution. *Physiologia Plantarum*, *176*(1), e14172. https://doi.org/10.1111/ppl.14172

Ghorbani, S., Lin, Y.-C., Parizot, B., Fernandez, A., Njo, M. F., Van de Peer, Y., Beeckman, T., & Hilson, P. (2015). Expanding the repertoire of secretory peptides controlling root development with comparative genome analysis and functional assays. *Journal of Experimental Botany*, *66*(17), 5257–5269. https://doi.org/10.1093/jxb/erv346

Goodstein, D. M., Shu, S., Howson, R., Neupane, R., Hayes, R. D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N., & Rokhsar, D. S. (2012). Phytozome: A comparative platform for green plant genomics. *Nucleic Acids Research*, *40*(Database issue), D1178-1186. https://doi.org/10.1093/nar/gkr944

Gutiérrez-Marcos, J. F., Costa, L. M., Biderre-Petit, C., Khbaya, B., O'Sullivan, D. M., Wormald, M., Perez, P., & Dickinson, H. G. (2004). Maternally expressed gene1 Is a novel maize endosperm transfer cell-specific gene with a maternal parent-of-origin pattern of expression. *The Plant Cell*, *16*(5), 1288–1301. https://doi.org/10.1105/tpc.019778

Hanada, K., Higuchi-Takeuchi, M., Okamoto, M., Yoshizumi, T., Shimizu, M., Nakaminami, K., Nishi, R., Ohashi, C., Iida, K., Tanaka, M., Horii, Y., Kawashima, M., Matsui, K., Toyoda, T., Shinozaki, K., Seki, M., & Matsui, M. (2013). Small open reading frames associated with morphogenesis are hidden in plant genomes. *Proceedings of the National Academy of Sciences*, *110*(6), 2395–2400. https://doi.org/10.1073/pnas.1213958110

He, Y., He, X., Wang, X., Hao, M., Gao, J., Wang, Y., Yang, Z.-N., & Meng, X. (2023). An EPFL peptide signaling pathway promotes stamen elongation via enhancing filament cell proliferation

692 to ensure successful self-pollination in Arabidopsis thaliana. *New Phytologist*, *238*(3), 1045–1058.
693 https://doi.org/10.1111/nph.18806

694 He, Y.-H., Chen, S.-Y., Chen, X.-Y., Xu, Y.-P., Liang, Y., & Cai, X.-Z. (2023). RALF22 promotes
695 plant immunity and amplifies the Pep3 immune signal. *Journal of Integrative Plant Biology*,
696 *65*(11), 2519–2534. https://doi.org/10.1111/jipb.13566

697 Hou, S., Liu, D., & He, P. (2021). Phytocytokines function as immunological modulators of plant
698 immunity. *Stress Biology*, *1*(1), 8. https://doi.org/10.1007/s44154-021-00009-y

699 Hsu, P. Y., & Benfey, P. N. (2018). Small but Mighty: Functional Peptides Encoded by Small
700 ORFs in Plants. *PROTEOMICS*, *18*(10), 1700038. https://doi.org/10.1002/pmic.201700038

701 Hu, X.-L., Lu, H., Hassan, M. M., Zhang, J., Yuan, G., Abraham, P. E., Shrestha, H. K., Villalobos
702 Solis, M. I., Chen, J.-G., Tschaplinski, T. J., Doktycz, M. J., Tuskan, G. A., Cheng, Z.-M. (Max),
703 & Yang, X. (2021). Advances and perspectives in discovery and functional analysis of small
704 secreted proteins in plants. *Horticulture Research*, *8*(1), 1–14. https://doi.org/10.1038/s41438-
705 021-00570-7

706 Huang, J., Zhang, T., Linstroth, L., Tillman, Z., Otegui, M. S., Owen, H. A., & Zhao, D. (2016).
707 Control of Anther Cell Differentiation by the Small Protein Ligand TPD1 and Its Receptor EMS1
708 in Arabidopsis. *PLOS Genetics*, *12*(8), e1006147. https://doi.org/10.1371/journal.pgen.1006147

709 Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool,
710 K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J.,
711 Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., … Hassabis, D. (2021). Highly
712 accurate protein structure prediction with AlphaFold. *Nature*, *596*(7873), 583–589.
713 https://doi.org/10.1038/s41586-021-03819-2

714 Katoh, K., & Standley, D. M. (2013). MAFFT Multiple Sequence Alignment Software Version 7:
715 Improvements in Performance and Usability. *Molecular Biology and Evolution*, *30*(4), 772–780.
716 https://doi.org/10.1093/molbev/mst010

717 Kim, D., Langmead, B., & Salzberg, S. L. (2015). HISAT: A fast spliced aligner with low memory
718 requirements. *Nature Methods*, *12*(4), 357–360. https://doi.org/10.1038/nmeth.3317

719 Kulaeva, O., Kliukova, M., Afonin, A., Sulima, A., Zhukov, V., & Tikhonovich, I. (2020). The
720 role of plant antimicrobial peptides (AMPs) in response to biotic and abiotic environmental factors.
721 *Biological Communications*, *65*(2), Article 2. https://doi.org/10.21638/spbu03.2020.205

722 Kwenda, S., Birch, P. R. J., & Moleleki, L. N. (2016). Genome-wide identification of potato long
723 intergenic noncoding RNAs responsive to Pectobacterium carotovorum subspecies brasiliense
724 infection. *BMC Genomics*, *17*(1), 614. https://doi.org/10.1186/s12864-016-2967-9

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & 1000 Genome Project Data Processing Subgroup. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics (Oxford, England)*, *25*(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Li, Y., Di, Q., Luo, L., & Yu, L. (2024). Phytosulfokine peptides, their receptors, and functions. *Frontiers in Plant Science*, *14*. https://doi.org/10.3389/fpls.2023.1326964

Lima, A. M., Azevedo, M. I. G., Sousa, L. M., Oliveira, N. S., Andrade, C. R., Freitas, C. D. T., & Souza, P. F. N. (2022). Plant antimicrobial peptides: An overview about classification, toxicity and clinical applications. *International Journal of Biological Macromolecules*, *214*, 10–21. https://doi.org/10.1016/j.ijbiomac.2022.06.043

Liu, L., Liu, X., Bai, Z., Tanveer, M., Zhang, Y., Chen, W., Shabala, S., & Huang, L. (2024). Small but powerful: RALF peptides in plant adaptive and developmental responses. *Plant Science*, *343*, 112085. https://doi.org/10.1016/j.plantsci.2024.112085

Liu, Z., Hou, S., Rodrigues, O., Wang, P., Luo, D., Munemasa, S., Lei, J., Liu, J., Ortiz-Morea, F. A., Wang, X., Nomura, K., Yin, C., Wang, H., Zhang, W., Zhu-Salzman, K., He, S. Y., He, P., & Shan, L. (2022). Phytocytokine signalling reopens stomata in plant immunity and water loss. *Nature*, *605*(7909), 332–339. https://doi.org/10.1038/s41586-022-04684-3

Lori, M., van Verk, M. C., Hander, T., Schatowitz, H., Klauser, D., Flury, P., Gehring, C. A., Boller, T., & Bartels, S. (2015). Evolutionary divergence of the plant elicitor peptides (Peps) and their receptors: Interfamily incompatibility of perception but compatibility of downstream signalling. *Journal of Experimental Botany*, *66*(17), 5315–5325. https://doi.org/10.1093/jxb/erv236

Lyapina, I., Filippova, A., & Fesenko, I. (2019). The Role of Peptide Signals Hidden in the Structure of Functional Proteins in Plant Immune Responses. *International Journal of Molecular Sciences*, *20*(18), 4343. https://doi.org/10.3390/ijms20184343

Lyapina, I., Filippova, A., Kovalchuk, S., Ziganshin, R., Mamaeva, A., Lazarev, V., Latsis, I., Mikhalchik, E., Panasenko, O., Ivanov, O., Ivanov, V., & Fesenko, I. (2021). Possible role of small secreted peptides (SSPs) in immune signaling in bryophytes. *Plant Molecular Biology*, *106*(1–2), 123–143. https://doi.org/10.1007/s11103-021-01133-z

Mamaeva, A., Lyapina, I., Knyazev, A., Golub, N., Mollaev, T., Chudinova, E., Elansky, S., Babenko, V. V., Veselovsky, V. A., Klimina, K. M., Gribova, T., Kharlampieva, D., Lazarev, V., & Fesenko, I. (2023). RALF peptides modulate immune response in the moss Physcomitrium patens. *Frontiers in Plant Science*, *14*. https://doi.org/10.3389/fpls.2023.1077301

McDonald, E. F., Jones, T., Plate, L., Meiler, J., & Gulsevin, A. (2023). Benchmarking AlphaFold2 on Peptide Structure Prediction. *Structure (London, England : 1993)*, *31*(1), 111-119.e2. https://doi.org/10.1016/j.str.2022.11.012

McKinney, W. (2012). Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython. *"O'Reilly Media, Inc."*

Monzon, V., Paysan-Lafosse, T., Wood, V., & Bateman, A. (2022). Reciprocal best structure hits: Using AlphaFold models to discover distant homologues. *Bioinformatics Advances*, *2*(1), vbac072. https://doi.org/10.1093/bioadv/vbac072

Moravej, H., Moravej, Z., Yazdanparast, M., Heiat, M., Mirhosseini, A., Moosazadeh Moghaddam, M., & Mirnejad, R. (2018). Antimicrobial Peptides: Features, Action, and Their Resistance Mechanisms in Bacteria. *Microbial Drug Resistance*, *24*(6), 747–767. https://doi.org/10.1089/mdr.2017.0392

Ngou, B. P. M., Wyler, M., Schmid, M. W., Kadota, Y., & Shirasu, K. (2024). Evolutionary trajectory of pattern recognition receptors in plants. *Nature Communications*, *15*(1), 308. https://doi.org/10.1038/s41467-023-44408-3

Ogawa-Ohnishi, M., Yamashita, T., Kakita, M., Nakayama, T., Ohkubo, Y., Hayashi, Y., Yamashita, Y., Nomura, T., Noda, S., Shinohara, H., & Matsubayashi, Y. (2022). Peptide ligand-mediated trade-off between plant growth and stress response. *Science (New York, N.Y.)*, *378*(6616), 175–180. https://doi.org/10.1126/science.abq5735

Olsson, V., Joos, L., Zhu, S., Gevaert, K., Butenko, M. A., & De Smet, I. (2019). Look Closely, the Beautiful May Be Small: Precursor-Derived Peptides in Plants. *Annual Review of Plant Biology*, *70*, 153–186. https://doi.org/10.1146/annurev-arplant-042817-040413

Pan, B., Sheng, J., Sun, W., Zhao, Y., Hao, P., & Li, X. (2013). OrysPSSP: A comparative Platform for Small Secreted Proteins from rice and other plants. *Nucleic Acids Research*, *41*(D1), D1192–D1198. https://doi.org/10.1093/nar/gks1090

Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T.-C., Mendell, J. T., & Salzberg, S. L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nature Biotechnology*, *33*(3), 290–295. https://doi.org/10.1038/nbt.3122

Quevillon, E., Silventoinen, V., Pillai, S., Harte, N., Mulder, N., Apweiler, R., & Lopez, R. (2005). InterProScan: Protein domains identifier. *Nucleic Acids Research*, *33*(Web Server issue), W116–W120. https://doi.org/10.1093/nar/gki442

Reboledo, G., Agorio, A. D., Vignale, L., Batista-García, R. A., & Ponce De León, I. (2021). Transcriptional profiling reveals conserved and species-specific plant defense responses during

791 the interaction of Physcomitrium patens with Botrytis cinerea. *Plant Molecular Biology*, *107*(4–
792 5), 365–385. https://doi.org/10.1007/s11103-021-01116-0

793 Rhodes, J., Roman, A.-O., Bjornson, M., Brandt, B., Derbyshire, P., Wyler, M., Schmid, M. W.,
794 Menke, F. L., Santiago, J., & Zipfel, C. (2022). Perception of a conserved family of plant signalling
795 peptides by the receptor kinase HSL3. *eLife*, *11*, e74687. https://doi.org/10.7554/eLife.74687

796 Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2010). edgeR: A Bioconductor package for
797 differential expression analysis of digital gene expression data. *Bioinformatics (Oxford, England)*,
798 *26*(1), 139–140. https://doi.org/10.1093/bioinformatics/btp616

799 Ryan, C. A., & Pearce, G. (2003). Systemins: A functionally defined family of peptide signals that
800 regulate defensive genes in Solanaceae species. *Proceedings of the National Academy of Sciences*,
801 *100*(suppl_2), 14577–14580. https://doi.org/10.1073/pnas.1934788100

802 Rzemieniewski, J., & Stegmann, M. (2022). Regulation of pattern-triggered immunity and growth
803 by phytocytokines. *Current Opinion in Plant Biology*, *68*, 102230.
804 https://doi.org/10.1016/j.pbi.2022.102230

805 Santos-Silva, C. A. dos, Ferreira-Neto, J. R. C., Amador, V. C., Bezerra-Neto, J. P., Vilela, L. M.
806 B., Binneck, E., Rêgo, M. de S., da Silva, M. D., Mangueira de Melo, A. L. T., da Silva, R. H., &
807 Benko-Iseppon, A. M. (2023). From Gene to Transcript and Peptide: A Deep Overview on Non-
808 Specific Lipid Transfer Proteins (nsLTPs). *Antibiotics*, *12*(5), Article 5.
809 https://doi.org/10.3390/antibiotics12050939

810 Shevchuk, N. A., Bryksin, A. V., Nusinovich, Y. A., Cabello, F. C., Sutherland, M., & Ladisch, S.
811 (2004). Construction of long DNA molecules using long PCR-based fusion of several fragments
812 simultaneously. *Nucleic Acids Research*, *32*(2), e19. https://doi.org/10.1093/nar/gnh014

813 Singh, U., & Wurtele, E. S. (2021). orfipy: A fast and flexible tool for extracting ORFs.
814 *Bioinformatics*, *37*(18), 3019–3020. https://doi.org/10.1093/bioinformatics/btab090

815 Slezina, M. P., Istomina, E. A., Korostyleva, T. V., Kovtun, A. S., Kasianov, A. S., Konopkin, A.
816 A., Shcherbakova, L. A., & Odintsova, T. I. (2021). Molecular Insights into the Role of Cysteine-
817 Rich Peptides in Induced Resistance to Fusarium oxysporum Infection in Tomato Based on
818 Transcriptome Profiling. *International Journal of Molecular Sciences*, *22*(11), 5741.
819 https://doi.org/10.3390/ijms22115741

820 Stegmann, M., Monaghan, J., Smakowska-Luzan, E., Rovenich, H., Lehner, A., Holton, N.,
821 Belkhadir, Y., & Zipfel, C. (2017). The receptor kinase FER is a RALF-regulated scaffold
822 controlling plant immune signaling. *Science*, *355*(6322), 287–289.
823 https://doi.org/10.1126/science.aal2541

Takata, N., Yokota, K., Ohki, S., Mori, M., Taniguchi, T., & Kurita, M. (2013). Evolutionary Relationship and Structural Characterization of the EPF/EPFL Gene Family. *PLOS ONE*, *8*(6), e65183. https://doi.org/10.1371/journal.pone.0065183

Tam, J. P., Wang, S., Wong, K. H., & Tan, W. L. (2015). Antimicrobial Peptides from Plants. *Pharmaceuticals*, *8*(4), 711–757. https://doi.org/10.3390/ph8040711

Tameshige, T., Ikematsu, S., Torii, K. U., & Uchida, N. (2017). Stem development through vascular tissues: EPFL-ERECTA family signaling that bounces in and out of phloem. *Journal of Experimental Botany*, *68*(1), 45–53. https://doi.org/10.1093/jxb/erw447

Tanaka, K., & Heil, M. (2021). Damage-Associated Molecular Patterns (DAMPs) in Plant Innate Immunity: Applying the Danger Model and Evolutionary Perspectives. *Annual Review of Phytopathology*, *59*(Volume 59, 2021), 53–75. https://doi.org/10.1146/annurev-phyto-082718-100146

Tavormina, P., De Coninck, B., Nikonorova, N., De Smet, I., & Cammue, B. P. A. (2015). The Plant Peptidome: An Expanding Repertoire of Structural Features and Biological Functions. *The Plant Cell*, *27*(8), 2095–2118. https://doi.org/10.1105/tpc.15.00440

Tian, D., Xie, Q., Deng, Z., Xue, J., Li, W., Zhang, Z., Dai, Y., Zheng, B., Lu, T., De Smet, I., & Guo, Y. (2022). Small secreted peptides encoded on the wheat (triticum aestivum L.) genome and their potential roles in stress responses. *Frontiers in Plant Science*, *13*, 1000297. https://doi.org/10.3389/fpls.2022.1000297

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*, *12*(2), Article 2. https://doi.org/10.3390/genes12020218

van Kempen, M., Kim, S. S., Tumescheit, C., Mirdita, M., Lee, J., Gilchrist, C. L. M., Söding, J., & Steinegger, M. (2024). Fast and accurate protein structure search with Foldseek. *Nature Biotechnology*, *42*(2), 243–246. https://doi.org/10.1038/s41587-023-01773-0

Van Rossum, G., and Drake, F. L., Jr (1995). Python tutorial: Centrum voor Wiskunde en Informatica Amsterdam. *The Netherlands*

Villagarcia, H., Morin, A.-C., Shpak, E. D., & Khodakovskaya, M. V. (2012). Modification of tomato growth by expression of truncated ERECTA protein from Arabidopsis thaliana. *Journal of Experimental Botany*, *63*(18), 6493–6504. https://doi.org/10.1093/jxb/ers305

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., et al. (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods, 17*, 261–272. https://doi.org/10.1038/s41592-019-0686-2

Wang, D., Jin, S., Chen, Z., Shan, Y., & Li, L. (2022). Genome-wide identification of the pectin methylesterase inhibitor genes in Brassica napus and expression analysis of selected members. *Frontiers in Plant Science*, *13*, 940284. https://doi.org/10.3389/fpls.2022.940284

Wang, P., Yao, S., Kosami, K., Guo, T., Li, J., Zhang, Y., Fukao, Y., Kaneko-Kawano, T., Zhang, H., She, Y.-M., Wang, P., Xing, W., Hanada, K., Liu, R., & Kawano, Y. (2020). Identification of endogenous small peptides involved in rice immunity through transcriptomics- and proteomics-based screening. *Plant Biotechnology Journal*, *18*(2), 415–428. https://doi.org/10.1111/pbi.13208

Waskom, M. (2021). seaborn: statistical data visualization. *Journal of Open Source Software, 6*(60), 3021. https://doi.org/10.21105/joss.03021

Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M., & Barton, G. J. (2009). Jalview Version 2—A multiple sequence alignment editor and analysis workbench. *Bioinformatics*, *25*(9), 1189–1191. https://doi.org/10.1093/bioinformatics/btp033

Wolf, S., & Höfte, H. (2014). Growth Control: A Saga of Cell Walls, ROS, and Peptide Receptors. *The Plant Cell*, *26*(5), 1848–1856. https://doi.org/10.1105/tpc.114.125518

Xia, X., Wei, Q., Xiao, C., Ye, Y., Li, Z., Marivingt-Mounir, C., Chollet, J.-F., Liu, W., & Wu, H. (2023). Genomic survey of NPF and NRT2 transporter gene families in five inbred maize lines and their responses to pathogens infection. *Genomics*, *115*(2), 110555. https://doi.org/10.1016/j.ygeno.2022.110555

Xiong, L., Huang, Y., Liu, Z., Li, C., Yu, H., Shahid, M. Q., Lin, Y., Qiao, X., Xiao, J., Gray, J. E., & Jin, J. (2022). Small EPIDERMAL PATTERNING FACTOR-LIKE2 peptides regulate awn development in rice. *Plant Physiology*, *190*(1), 516–531. https://doi.org/10.1093/plphys/kiac278

Yang, H., Kim, X., Skłenar, J., Aubourg, S., Sancho-Andrés, G., Stahl, E., Guillou, M.-C., Gigli-Bisceglia, N., Tran Van Canh, L., Bender, K. W., Stintzi, A., Reymond, P., Sánchez-Rodríguez, C., Testerink, C., Renou, J.-P., Menke, F. L. H., Schaller, A., Rhodes, J., & Zipfel, C. (2023). Subtilase-mediated biogenesis of the expanded family of SERINE RICH ENDOGENOUS PEPTIDES. *Nature Plants*, *9*(12), 2085–2094. https://doi.org/10.1038/s41477-023-01583-x

Zhang, J.-M., & An, J. (2007). Cytokines, inflammation, and pain. *International Anesthesiology Clinics*, *45*(2), 27–37. https://doi.org/10.1097/AIA.0b013e318034194e

Zhang, R., Shi, P.-T., Zhou, M., Liu, H.-Z., Xu, X.-J., Liu, W.-T., & Chen, K.-M. (2023). Rapid alkalinization factor: Function, regulation, and potential applications in agriculture. *Stress Biology*, *3*(1), 16. https://doi.org/10.1007/s44154-023-00093-2

Zhang, Y., & Skolnick, J. (2005). TM-align: A protein structure alignment algorithm based on the TM-score. *Nucleic Acids Research*, *33*(7), 2302–2309. https://doi.org/10.1093/nar/gki524

Ziemann, S., van der Linde, K., Lahrmann, U., Acar, B., Kaschani, F., Colby, T., Kaiser, M., Ding, Y., Schmelz, E., Huffaker, A., Holton, N., Zipfel, C., & Doehlemann, G. (2018). An apoplastic peptide activates salicylic acid signalling in maize. *Nature Plants*, *4*(3), Article 3. https://doi.org/10.1038/s41477-018-0116-y