



# A Comprehensive Phylogenetic Analysis of the MAP4K Family in the Green Lineage

Lixia Pan<sup>1,2†</sup>, Cassio Flavio Fonseca De Lima<sup>1,2†</sup>, Lam Dai Vu<sup>1,2</sup> and Ive De Smet<sup>1,2\*</sup>

<sup>1</sup> Department of Plant Biotechnology and Bioinformatics, Ghent University, Ghent, Belgium, <sup>2</sup> VIB Center for Plant Systems Biology, Ghent, Belgium

## OPEN ACCESS

### Edited by:

Angelica Cibrian-Jaramillo,  
Centro de Investigaciones y Estudios  
Avanzados, Instituto Politécnico  
Nacional de México (CINVESTAV),  
Mexico

### Reviewed by:

Luis Delaye,  
Centro de Investigaciones y Estudios  
Avanzados, Instituto Politécnico  
Nacional de México (CINVESTAV),  
Mexico

Jan de Vries,  
University of Göttingen, Germany  
Adrian Reyes-Prieto,  
University of New Brunswick  
Fredericton, Canada

### \*Correspondence:

Ive De Smet  
Ive.DeSmet@psb.vib-ugent.be

† These authors have contributed  
equally to this work

### Specialty section:

This article was submitted to  
Plant Systematics and Evolution,  
a section of the journal  
Frontiers in Plant Science

Received: 27 January 2021

Accepted: 22 July 2021

Published: 13 August 2021

### Citation:

Pan L, Fonseca De Lima CF,  
Vu LD and De Smet I (2021) A  
Comprehensive Phylogenetic Analysis  
of the MAP4K Family in the Green  
Lineage. *Front. Plant Sci.* 12:650171.  
doi: 10.3389/fpls.2021.650171

The kinase-mediated phosphorylation impacts every basic cellular process. While mitogen-activated protein kinase technology kinase kinases (MAP4Ks) are evolutionarily conserved, there is no comprehensive overview of the MAP4K family in the green lineage (Viridiplantae). In this study, we identified putative MAP4K members from representative species of the two core groups in the green lineage: Chlorophyta, which is a diverse group of green algae, and Streptophyta, which is mostly freshwater green algae and land plants. From that, we inferred the evolutionary relationships of MAP4K proteins through a phylogenetic reconstruction. Furthermore, we provided a classification of the MAP4Ks in the green lineage into three distinct.

**Keywords:** signaling, MAP4K, phylogenetic analysis, evolution, motifs

## INTRODUCTION

Post-translational modifications (PTMs) are vital for plants to sense and respond to environmental changes (Hashiguchi and Komatsu, 2016; Vu et al., 2018). Among these, reversible protein phosphorylation is one of the most widespread and pivotal PTMs that affect every basic cellular process (Ardito et al., 2017; Vu et al., 2018). Phosphorylation is catalyzed by protein kinases, which predominantly phosphorylate substrate proteins on serine, threonine, or tyrosine residues, and protein phosphatases, which mediate the reverse reaction. One of the well-explored, highly conserved kinase subfamilies are components of the mitogen-activated protein kinase (MAPK) cascades, which act through linear sequential serine/threonine and/or tyrosine phosphorylation (Xu and Zhang, 2015). An MAPK signaling module is composed of an MAPK kinase kinase (MAP3K), a MAPK kinase (MAP2K), and a MAPK (Jagodzik et al., 2018). Typically, a MAPK cascade regulates signal transduction through a MAP3K that phosphorylates a MAP2K, which then activates a MAPK by phosphorylation (Xu and Zhang, 2015; Krysan and Colcombet, 2018). Plant MAPK cascades act downstream of receptor-like protein kinases or G-proteins and play important roles in immunity, abiotic stress, and plant growth and development (Cho et al., 2008; Cheng et al., 2015; Xu and Zhang, 2015; Liu and Zhou, 2018; Yan et al., 2018; Zhu et al., 2019).

However, the yeast Ste20 acts as a MAPK kinase kinase (MAP4K) that activates Ste11, a MAP3K, which, in turn, activates downstream components in the mating signaling pathway (Leberer et al., 1992). Interestingly, MAP4Ks are conserved not only in yeast but also in mammals and plants (Dan et al., 2001; Champion et al., 2004; Chuang et al., 2016). The MAP4Ks in plants were first identified in *Brassica* (Leprince et al., 1999), but, so far, plant MAP4Ks have not been extensively characterized (Pan and De Smet, 2020). In *Arabidopsis thaliana*, 10 MAP4Ks were described (Jonak et al., 2002), but not all MAP4Ks have been functionally characterized:

BLUE LIGHT SIGNALING 1 (BLUS1)/MAP4K10 is essential for stomatal opening in response to blue light (Takemiya et al., 2013; Schnabel et al., 2018), SALT INDUCIBLE KINASE 1 (SIK1)/MAP4K3 is critical for cell proliferation and expansion during organ growth and development (Xiong et al., 2016), and both SIK1 and MAP4K4 regulate the flg22-triggered immunity response (Jiang et al., 2019). In addition, a TARGET OF TEMPERATURE 3 (TOT3)/MAP4K4 interacts with TOT3-INTERACTING PROTEIN 4 (TOI4)/MAP4K6 and TOI5/MAP5K5 and controls thermomorphogenesis in Arabidopsis and wheat (Vu et al., 2021).

Although that MAP4Ks are so conserved throughout multiple clades, there is no comprehensive overview of the MAP4K family in the green lineage (Viridiplantae). Viridiplantae is a monophyletic clade of photosynthetic eukaryotic organisms that play important roles in both terrestrial and aquatic ecologies (Leliaert et al., 2012). The group separated around 1 billion years ago into Chlorophyta and Streptophyta, the latter including Embryophyta (the clade of the land plants) (Morris et al., 2018; Fürst-Jansen et al., 2020). One major evolutionary milestone was the conquest of land by the Embryophyta, which is from a freshwater streptophyte algal common ancestor, followed by an expansion in biodiversity (Becker and Marin, 2009). The habitat shift was accompanied by the modification of a series of traits that granted success in land colonization (e.g., the development of specialized organs and tissues such as stomata, roots, leaves, seeds, and flowers), but also a set of molecular innovations (Pires and Dolan, 2012; Ishizaki, 2017).

In this study, we identify putative MAP4K members from representative species of Chlorophyta and Streptophyta, the two clades in the Viridiplantae, and infer the phylogenetic relationships in the green lineage underpinning MAP4K evolution and diversification.

## MATERIALS AND METHODS

### Genome-Wide Identification of (Putative) MAP4Ks

To identify and reconstruct the evolution of putative MAP4Ks throughout the green lineage (Viridiplantae), 6 Chlorophytes, and 29 Streptophytes (including 24 Embryophytes) were chosen depending on the genome availability and annotation status (**Supplementary Table 1**). The MAP4K sequences were identified using a hidden Markov model (HMM) profile generated and calibrated using the HMMER software v3.3 (Eddy, 1998) that was trained based on an alignment of the 10 MAP4Ks previously identified in *A. thaliana* (**Supplementary Figure 1**; Jonak et al., 2002). Local searches were performed on primary protein libraries (Leebens-Mack et al., 2019). To maximize the balance between sensitivity and selectivity, only hits with both per-sequence and per-domain *E*-values lower than  $2e-100$  and primary protein isoforms were considered for downstream analysis. The sequences retrieved were then further confirmed for the presence of the protein kinases domain (PF00069) in the Pfam database (El-Gebali et al., 2019) and InterPro Scan v5 (Jones et al., 2014) and inspected considering the annotation status for

some of the libraries; the 5% longest and shortest were reviewed using a homology-based method (Korf et al., 2001). For this process, the canonical genomic structure [(e.g., evolutionarily conserved exons, introns, and untranslated regions (UTRs))] of the best reciprocal blast hits among *A. thaliana* (Streptophyta-dicotyledon), *Oryza sativa* (Streptophyta-monocotyledon), and *Chlamydomonas reinhardtii* (Chlorophyta) were compared for validation (Altschul et al., 1990, 1997; Moreno-Hagelsieb and Latimer, 2008).

### Alignment and Phylogenetic Analyses

To build a phylogenetic reconstruction for the MAP4K group, a MAPK family member (MPK3, AT3G45640.1) was first added to the list of putative MAP4Ks as a possible outgroup. Following this, full-length sequences were aligned using the GISMO software, which applies a top-down sequence alignment and is optimized for protein sampling (Neuwald and Altschul, 2016) (**Supplementary Data Sheet 1**). Subsequently, we inferred the maximum likelihood phylogenetic tree using the concatenated alignment under the best fitting model JTT + I + G4 in IQ-TREE v2.0.6 (Neuwald and Altschul, 2016), executing 1,000 ultrafast replicates. The resulting Newick consensus tree was visualized using the Interactive Tree Of Life version 4 (iTOL) (Letunic and Bork, 2019). The taxonomy tree was generated with the phyloT online tool (phylo.t.biobyte.de). In addition, a Bayesian Markov chain Monte Carlo (MCMC) sampling method for phylogenetic reconstruction was used. The alignment was submitted to the best fitting model JTT + G4 implemented in PhyloBayes with 5,000 generations and, after the convergence of the two chains, a burn-in of 25% was applied. The resulting consensus tree (data not shown) was inspected and compared with the previous tree, revealing a high conservation of the three clades (100%, 100%, and 99.41% for clade I, II, and II, respectively) and high posterior probability support (1.0.83, and 1 for Clade I, II, and III, respectively), but several unresolved polytomies.

### Feature Search and Characterization

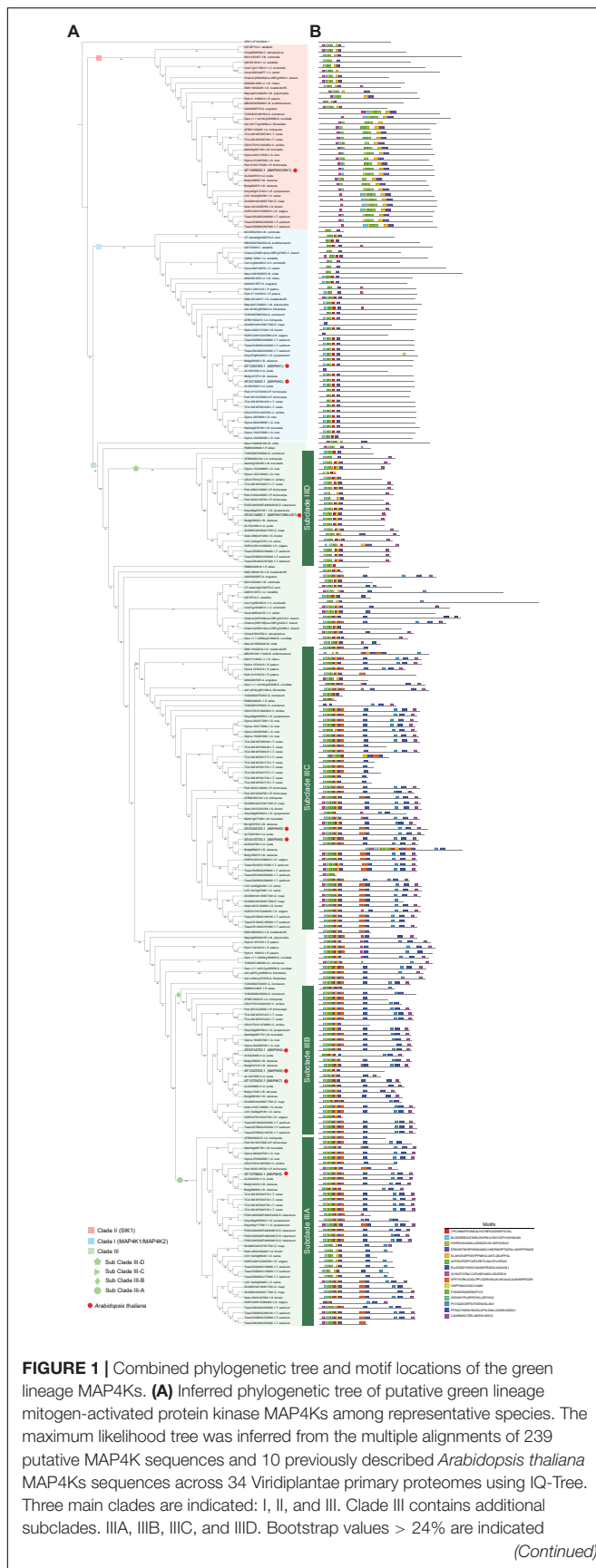
To gain insight into the conserved motifs among plant MAP4K sequences, the MEME suite 5.3.3<sup>1</sup> (Bailey and Gribskov, 1998) was used. The number of discoverable motifs was set to 15 and the width of motifs was set to an interval of a minimum of 3 and maximum of 50 amino acids. To create sequence logos, WEBLOGO<sup>2</sup> was used with the default settings (Crooks et al., 2004). The analysis of disorder was performed by the web-based IUPred2A tool<sup>3</sup> (Mészáros et al., 2018) utilizing the *A. thaliana* MAP4K protein sequences. The resulting disorder probability was used to categorize each residue as either ordered (<0.4), intermediate (0.4–0.6), or disordered (>0.6). Analyses of the biotic and abiotic influences on *A. thaliana* MAP4K expression was done through analyses of available transcriptome data on the Genevestigator platform<sup>4</sup> (Zimmermann et al., 2004,

<sup>1</sup><http://meme-suite.org/tools/meme>

<sup>2</sup><https://weblogo.berkeley.edu/logo.cgi>

<sup>3</sup><https://iupred2a.elte.hu/>

<sup>4</sup><https://genevestigator.com/gv/start/start.jsp>

**FIGURE 2 |** Continued

in the branches. Red dots indicate *Arabidopsis* MAP4Ks. The *Arabidopsis* MAP4Ks included in the (sub)clades are MAP4K1 (I), MAP4K2 (II), MAP4K3/SIK1 (II), MAP4K4 (IIB), MAP4K5 (IIC), MAP4K6 (IIC), MAP4K7 (IIB), MAP4K8 (IIA), MAP4K9 (IIB), and MAP4K10/BLUS1 (IID). **(B)** Motif distribution along the putative MAP4K sequences identified by MEME tool. The height of motif “block” is proportional to  $-\log(p\text{-value})$ , the taller the block, and the lower the probability of a wrong match. The motif sequence is indicated in the figure.

2005; Hruz et al., 2008) on August 25, 2020. The eFP browser<sup>5</sup> (Hayes et al., 2008) with standard settings was used for the cell, tissue, and organ MAP4K expression analyses.

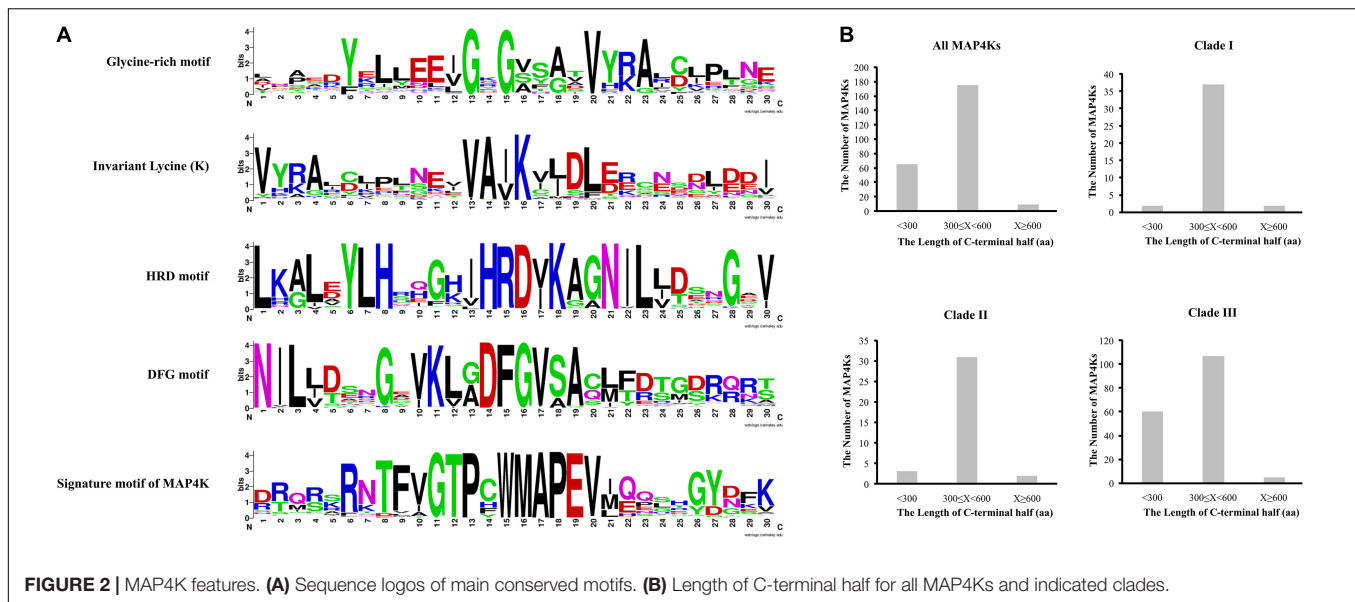
## RESULTS AND DISCUSSION

### The MAP4K Family Members in *A. thaliana*

In *A. thaliana*, the MAP4K family consists of 10 members, namely, MAP4K1 (At1G53165), MAP4K2 (At3G15220), MAP4K3/SIK1 (At1G69220), MAP4K4/TOT3 (At5G14720), MAP4K5/TO15 (At4G24100), MAP4K6/TO14 (At4G10730), MAP4K7 (At1G70430), MAP4K8 (At1G79640), MAP4K9 (At1G23700), and MAP4K10/BLUS1 (At4G14480) (Pan and De Smet, 2020; Vu et al., 2021). Most *A. thaliana* MAP4K family members share common features, such as a MAP4K signature motif and several conserved residues in the kinase domain, and display a C-terminal half from the end of the kinase domain to the stop codon that is largely intrinsically disordered (**Supplementary Figure 1**). However, some differences can be observed. For example, SIK1 contains a long intrinsically disordered N-terminal part before the kinase domain, which is important for interaction with MOB1 to control cell expansion (Xiong et al., 2016; **Supplementary Figure 1**). Furthermore, MAP4K9 lacks the HRD and DFG motifs in the kinase domain that are vital for catalytic activity, does not contain an intrinsically disordered C-terminal part, and does not display the conserved MAP4K signature and GXGXXG/A motifs (**Supplementary Figure 1**). In *Homo sapiens*, MAP4Ks contain a characteristic, highly conserved C-terminal citron-homology domain (CNH) in all seven known members of the group (Chuang et al., 2016; Seo et al., 2020), which is also present in MAP4K orthologs in *Drosophila melanogaster* and *Caenorhabditis elegans* (Fiedler et al., 2014), which is important for protein-protein interactions (Poinat et al., 2002; Machida et al., 2004; Taira et al., 2004). Strikingly, a fully conserved CNH domain is absent in *A. thaliana* MAP4Ks.

To gain insight into the expression patterns of the *A. thaliana* MAP4K family during development, the MAP4K expression data were compiled and visualized from online repositories, namely, eFP browser (Winter et al., 2007) and Genevestigator v3 (Hruz et al., 2008). These *in silico* expression patterns suggested that most MAP4Ks are broadly expressed throughout the plant, with some MAP4Ks expressed stronger or specifically

<sup>5</sup><http://www.bar.utoronto.ca/efp/cgi-bin/efpWeb.cgi>



in one organ, such as in pollen (Supplementary Figure 2). This indicates distinct roles for some of these kinases. Furthermore, the *MAP4K* expression was hardly affected by hormones or nutrients and was mainly regulated by abiotic and biotic stresses (Supplementary Table 2).

## Identification of Putative MAP4K Family Members in the Green Lineage

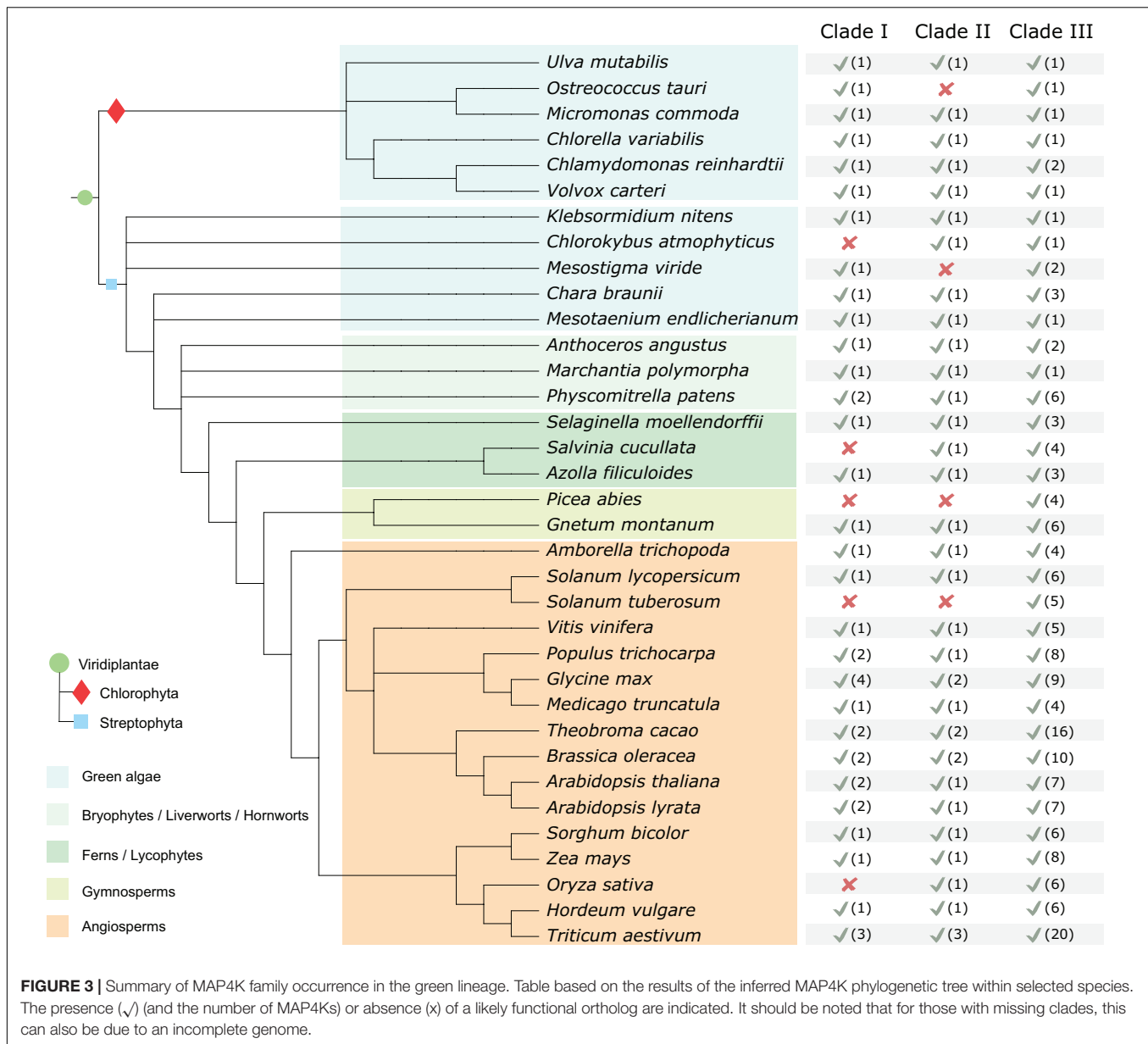
While there is some information on the MAP4K family members in a few plant species, such as *A. thaliana*, *Brassica napus*, *Zea mays*, *Solanum chacoense*, and *Vitis vinifera* (Leprince et al., 1999; Llompert et al., 2003; Major et al., 2009; Cakir and Kılıçkaya, 2015; Pan and De Smet, 2020), a comprehensive overview is lacking. Therefore, to reconstruct the evolutionary relationships of the MAP4K family in the green lineage, we analyzed 34 representative species with available genome or protein information (Tuskan et al., 2006; Jaillon et al., 2007; Merchant et al., 2007; Ouyang et al., 2007; Palenik et al., 2007; Worden et al., 2009; Blanc et al., 2010; Prochnik et al., 2010; Schmutz et al., 2010; Banks et al., 2011; Hu et al., 2011; Young et al., 2011; Lamesch et al., 2012; Sato et al., 2012; Albert et al., 2013; Motamayor et al., 2013; Nystedt et al., 2013; Sharma et al., 2013; Blanc-Mathieu et al., 2014; Hori et al., 2014; Liu et al., 2014; Hirsch et al., 2016; Beier et al., 2017; Bowman et al., 2017; Appels et al., 2018; De Clerck et al., 2018; Lang et al., 2018; Li et al., 2018; McCormick et al., 2018; Nishiyama et al., 2018; Wan et al., 2018; Cheng et al., 2019; Wang et al., 2020; Zhang et al., 2020; Supplementary Table 1). In this analysis, we included representative green algae from Chlorophyta [*Chlamydomonas reinhardtii* and *Volvox carteri* (Chlorophyceae), *Micromonas pusilla* (Mamiellophyceae), *Ostreococcus tauri* and *Chlorella variabilis* (Trebouxiophyceae), and *Ulva mutabilis* (Ulvophyceae)] and Streptophyta [*Klebsormidium nitens* (Klebsormidiophyceae), *Chlorokybus atmophyticus* (Chlorokybophyceae), *Mesostigma*

*viride* (Mesostigmatophyceae), *Mesotaenium endlicherianum* (Zygnematophyceae), and *Chara braunii* (Charophyceae)]. In addition, several land plants from the Streptophyta were included to trace the expansion of MAP4Ks in land plants: the bryophyte *Physcomitrium patens*, the liverwort *Marchantia polymorpha*, the hornwort *Anthoceros agrestis*, the lycophyte *Selaginella moellendorffii*, the ferns *Azolla filiculoides* and *Salvinia cucullata*, the gymnosperms *Picea abies* and *Gnetum montanum*, and 16 angiosperms (including *Amborella trichopoda*, 10 eudicots, and 5 monocots). Through in-depth analyses of protein-containing databases (see section “Materials and Methods”), we generated a set of 249 putative MAP4K family members with a minimum and maximum amino acid sequence length of 225 and 2,251, respectively (Supplementary Table 3 and Supplementary Data Sheet 2). These results showed that putative members of the MAP4K family are present in both Chlorophyta and Streptophyta and all representative species used in this analysis (Figure 1 and Supplementary Tables 1, 3). The resulting phylogenetic consensus tree revealed three major clades (Figure 1). With respect to clade III, four subclades could be distinguished in angiosperms: one subclade with AtMAP4K8 (Subclade IIIA), a second subclade with AtMAP4K4, AtMAP4K7, and AtMAP4K9 (Subclade IIIB), the third subclade with AtMAP4K5 and AtMAP4K6 (Subclade IIIC), and the fourth clade with AtMAP4K10 (Subclade IIID) (Figure 1 and Supplementary Table 3).

## Conserved Features of the MAP4Ks in Plants

As proteins are generally composed of one or more functional regions or domains that can provide insight into their function and evolutionary relationships (Lee et al., 2007), we explored to what level the key domains and motifs are conserved in the putative MAP4K family members (Figures 1, 2A, Supplementary Figure 3, and Supplementary Table 3).





First, we assessed to what extent the identified sequences have a likely functional kinase domain based on some key features, such as a glycine-rich loop motif (GxGxxG) that is essential for nucleotide binding and is an integral part of the ATP-binding site. In general, in about 20% of kinases, the third G is substituted by A or S, which disrupts kinase activity (Hanks and Quinn, 1991; Hemmer et al., 1997; Torkamani et al., 2008; Chmielecki et al., 2010; Steinberg, 2018; Zhang et al., 2018). We also observed this in several MAP4Ks (**Supplementary Table 3** and **Supplementary Figure 3**). In addition, we analyzed HRD and DFG motifs and an invariant lysine (K), all of which are pivotal for catalytic activity (Hanks et al., 1988; Carrera et al., 1993; Hanks and Hunter, 1995; Xu et al., 2000; Roskoski, 2004; Kannan and Neuwald, 2005; Zhang et al., 2015; Modi and Dunbrack, 2019; Pan and De Smet, 2020; **Supplementary Table 3** and **Supplementary Figure 3**). Due to changes in these amino acid triads, the putative MAP4Ks

that lack all or some of the above-mentioned kinase features are possibly kinase-inactive MAP4Ks (**Supplementary Table 3**; Paul and Srinivasan, 2020).

Second, we evaluated the presence of the previously proposed signature motif for (mammalian) MAP4Ks (GTPyWMAPEv, with Y and V being less conserved) located in the kinase subdomain VIII (Sells and Chernoff, 1997; Dan et al., 2001). The majority (94%) of the putative MAP4Ks in the green lineage share this well-conserved signature motif (**Supplementary Table 3** and **Supplementary Figure 3**). Interestingly, two *A. thaliana* and *A. lyrata* MAP4K9s from subclade IIIB do not contain this conserved signature motif and show severe alterations in the HRD, DFG, and glycine-rich loop motifs (GxGxxG) (**Supplementary Table 3**), in addition, which raises the question if these should still be considered as true (functional) MAP4Ks.

Third, we investigated the length of the C-terminal half of the MAP4Ks. Although the disordered C-terminal half is essential for the interaction with substrates, it is largely not conserved across sequences (Lee et al., 2017). For almost 70% of the putative MAP4Ks, the length of the C-terminal half was between 300 and 600 amino acids (Figure 2B, Supplementary Figure 4, and Supplementary Table 3). Putative MAP4Ks that (partially) lack such a C-terminal half, such as SMO169G0133.1 in *Selaginella*, TraesCS6A02G353400.1 in wheat, and Bo9g089620.1 in *Brassica*, will likely be affected in the interaction with substrates (Zhang et al., 2018; Jiang et al., 2019).

Strikingly, none out of the 249 sequences contained a fully conserved CNH domain, suggesting that this characteristic element for several non-plant MAP4Ks might not have been present in the ancestor of the Viridiplantae. However, we identified conserved but clade-specific motifs at a similar position where the CNH domain is found in other eukaryotes (Figure 1, Supplementary Figures 1–5, and Supplementary Table 3). Furthermore, an experimental analysis will be required to reveal whether these motifs are involved in similar functions as the CNH domain, such as protein–protein interaction.

Finally, the phototropin-mediated phosphorylation of BLUS1 Ser-348 alleviates the autoinhibitory activity of the C-terminal part on its kinase activity (Hosotani et al., 2021). This serine residue belongs to a conserved RRISGWNF consensus motif (Takemiya et al., 2013). Therefore, we checked for the presence of such a motif in all the putative MAP4Ks. This revealed that an S(G/A)WNF motif is absent in Clade I and Clade II of MAP4Ks, but present in 64% of Clade III MAP4Ks (Supplementary Table 3 and Supplementary Figure 3).

## Evolutionary Insight in the MAP4K Family in the Green Lineage

Based on the tree topology of the inferred phylogenetic reconstruction of the MAP4K family, we divided the group into three major clades (I, II, and III). In addition, based on the protein motif structure and expansion in flowering plants, we defined four subclades in clade III (IIIA, IIIB, IIIC, and IIID) (Figure 1). While MAP4Ks are present in all the plant species we investigated, the absolute number of group members increased in land plants, especially in Clade III, and some species seemed to lack members for a particular clade (Figure 3). For example, *O. tauri* (Chlorophyta), *Mesostigma viride* (Streptophyta), *Picea abies*, and *Solanum tuberosum* lacked a clade II member, and clade I was absent in *C. atmophyticus* (Streptophyta), *Salvinia cucullate* (Fern), *P. abies*, *S. tuberosum*, and *O. sativa* (Figures 1, 3). Furthermore, Subclade IIIA (which contained angiosperm-exclusive MAP4Ks) appeared as a sister group of Subclade IIIB (which was exclusive to seed plant MAP4Ks) and orthologs for both clades were identified in ferns. Likely, the increase in the number of MAP4Ks in vascular plants (ferns, gymnosperms, and angiosperms) is associated with the increased rate of whole-genome duplications or polyploidization largely observed in land plants; thus, it is often invoked as one of the main causal agents for diversification and land colonization (Bowers et al., 2003; Qiao et al., 2019; Zhang et al., 2019). In addition, Subclade IIID (mainly

defined by the presence of BLUS1) contained only members of gymnosperms and angiosperms (Figure 1), extending this beyond the angiosperm lineage (Harris et al., 2020). However, the single *Gnetum montanum* (TnS000821935t04) MAP4K sequence identified as belonging to this clade has retained only some of the described angiosperm motifs and lacks the LQQRMSLTELQKEKLKNVQ motif in the disordered domain, which is likely essential for protein–protein interactions (Figure 1). Therefore, whether the gymnosperm MAP4K in subclade IIID retains a similar function as the other members of the clade remains unknown.

## CONCLUSION

In this study, we identified putative MAP4K members from representative plant species and inferred the phylogenetic relationships in the green lineage underpinning MAP4K evolution and diversification. As a whole, the presence and absence of typical sequence motifs in (putative) MAP4K sequences likely resulted in a functional diversification within the MAPK4 family. The next important and necessary step is thus to evaluate the functionality and activity of these MAP4Ks, not only in *A. thaliana*, but also in other members of the green lineage. Importantly, this also relates to some of the motifs that were identified and the large intrinsically disordered C-terminal half. Some of these analyses will be facilitated by the fewer MAP4K family members identified in this study in, for example, *Marchantia*, which is a valuable model for genetic studies (Ishizaki et al., 2016; Poveda, 2020).

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article and Supplementary Material. Further inquiries can be directed to the corresponding author.

## AUTHOR CONTRIBUTIONS

LP, CF, and LDV performed analyses. LP, CF, LDV, and ID interpreted the results and wrote the manuscript. ID coordinated the study. All authors contributed to the article and approved the submitted version.

## FUNDING

This work was supported by the Research Foundation, Flanders (FWO.OPR.2019.0009.01). LP is a recipient of a Ph.D. grant from the Chinese Scholarship Council. LDV was supported by the post-doctoral fellowship of the Special Research Fund (Bijzonder Onderzoeksfonds, Ghent University).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.650171/full#supplementary-material>

## REFERENCES

- Albert, V. A., Barbazuk, W. B., Depamphilis, C. W., Der, J. P., Leebens-Mack, J., Ma, H., et al. (2013). The Amborella genome and the evolution of flowering plants. *Science* 342:1241089. doi: 10.1126/science.1241089
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi: 10.1016/S0022-2836(05)80360-2
- Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402. doi: 10.1093/nar/25.17.3389
- Appels, R., Eversole, K., Stein, N., Feuillet, C., Keller, B., Rogers, J., et al. (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science* 361:eaar7191. doi: 10.1126/science.aar7191
- Ardito, F., Giuliani, M., Perrone, D., Troiano, G., and Lo Muzio, L. (2017). The crucial role of protein phosphorylation in cell signaling and its use as targeted therapy. *Int. J. Mol. Med.* 40, 271–280. doi: 10.3892/ijmm.2017.3036
- Bailey, T. L., and Gribskov, M. (1998). Combining evidence using p-values: application to sequence homology searches. *Bioinformatics (Oxford, England)* 14, 48–54. doi: 10.1093/bioinformatics/14.1.48
- Banks, J. A., Nishiyama, T., Hasebe, M., Bowman, J. L., Gribskov, M., DePamphilis, C., et al. (2011). The Selaginella genome identifies genetic changes associated with the evolution of vascular plants. *Science* 332, 960–963. doi: 10.1126/science.1203810
- Becker, B., and Marin, B. (2009). Streptophyte algae and the origin of embryophytes. *Ann. Bot.* 103, 999–1004. doi: 10.1093/aob/mcp044
- Beier, S., Himmelbach, A., Colmsee, C., Zhang, X.-Q., Barrero, R. A., Zhang, Q., et al. (2017). Construction of a map-based reference genome sequence for barley, *Hordeum vulgare* L. *Sci. Data* 4, 1–24. doi: 10.1038/sdata.2017.44
- Blanc, G., Duncan, G., Agarkova, I., Borodovsky, M., Gurnon, J., Kuo, A., et al. (2010). The *Chlorella variabilis* NC64A genome reveals adaptation to photosymbiosis, coevolution with viruses, and cryptic sex. *Plant Cell* 22, 2943–2955. doi: 10.1105/tpc.110.076406
- Blanc-Mathieu, R., Verhelst, B., Derelle, E., Rombauts, S., Bouget, F.-Y., Carré, I., et al. (2014). An improved genome of the model marine alga *Ostreococcus tauri* unfolds by assessing Illumina de novo assemblies. *BMC Genomics* 15:1103. doi: 10.1186/1471-2164-15-1103
- Bowers, J. E., Chapman, B. A., Rong, J., and Paterson, A. H. (2003). Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. *Nature* 422, 433–438. doi: 10.1038/nature01521
- Bowman, J. L., Kohchi, T., Yamato, K. T., Jenkins, J., Shu, S., Ishizaki, K., et al. (2017). Insights into land plant evolution garnered from the *Marchantia polymorpha* genome. *Cell* 171, 287.e5–304.e5. doi: 10.1016/j.cell.2017.09.030
- Cakir, B., and Kılıçkaya, O. (2015). Mitogen-activated protein kinase cascades in *Vitis vinifera*. *Front. Plant Sci.* 6:556. doi: 10.3389/fpls.2015.00556
- Carrera, A. C., Alexandrov, K., and Roberts, T. M. (1993). The conserved lysine of the catalytic domain of protein kinases is actively involved in the phosphotransfer reaction and not required for anchoring ATP. *Proc. Natl. Acad. Sci. U.S.A.* 90, 442–446. doi: 10.1073/pnas.90.2.442
- Champion, A., Picaud, A., and Henry, Y. (2004). Reassessing the MAP3K and MAP4K relationships. *Trends Plant Sci.* 9, 123–129. doi: 10.1016/j.tplants.2004.01.005
- Cheng, S., Xian, W., Fu, Y., Marin, B., Keller, J., Wu, T., et al. (2019). Genomes of subaerial Zygnematophyceae provide insights into land plant evolution. *Cell* 179, 1057.e4–1067.e4. doi: 10.1016/j.cell.2019.10.019
- Cheng, Z., Li, J.-F., Niu, Y., Zhang, X.-C., Woody, O. Z., Xiong, Y., et al. (2015). Pathogen-secreted proteases activate a novel plant immune pathway. *Nature* 521, 213–216. doi: 10.1038/nature14243
- Chmielecki, J., Peifer, M., Jia, P., Succi, N. D., Hutchinson, K., Viale, A., et al. (2010). Targeted next-generation sequencing of DNA regions proximal to a conserved GXGXXG signaling motif enables systematic discovery of tyrosine kinase fusions in cancer. *Nucleic Acids Res.* 38, 6985–6996. doi: 10.1093/nar/gkq579
- Cho, S. K., Larue, C. T., Chevalier, D., Wang, H., Jinn, T.-L., Zhang, S., et al. (2008). Regulation of floral organ abscission in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U.S.A.* 105, 15629–15634. doi: 10.1073/pnas.0805539105
- Chuang, H. C., Wang, X., and Tan, T. H. (2016). MAP4K family kinases in immunity and inflammation. *Adv. Immunol.* 129, 277–314. doi: 10.1016/bs.ai.2015.09.006
- Crooks, G. E., Hon, G., Chandonia, J.-M., and Brenner, S. E. (2004). WebLogo: a sequence logo generator. *Genome Res.* 14, 1188–1190. doi: 10.1101/gr.849004
- Dan, I., Watanabe, N. M., and Kusumi, A. (2001). The Ste20 group kinases as regulators of MAP kinase cascades. *Trends Cell Biol.* 11, 220–230. doi: 10.1016/S0962-8924(01)01980-8
- De Clerck, O., Kao, S. M., Bogaert, K. A., Blomme, J., Foflonker, F., Kwantes, M., et al. (2018). Insights into the evolution of multicellularity from the *Sea Lettuce* genome. *Curr. Biol. : CB* 28, 2921.e5–2933.e5. doi: 10.1016/j.cub.2018.08.015
- Eddy, S. R. (1998). Profile hidden Markov models. *Bioinformatics (Oxford, England)* 14, 755–763. doi: 10.1093/bioinformatics/14.9.755
- El-Gebali, S., Mistry, J., Bateman, A., Eddy, S. R., Luciani, A., Potter, S. C., et al. (2019). The Pfam protein families database in 2019. *Nucleic Acids Res.* 47, D427–D432. doi: 10.1093/nar/gky995
- Fiedler, L. R., Maifoshie, E., and Schneider, M. D. (2014). Mouse models of heart failure: cell signaling and cell survival. *Curr. Topics Dev. Biol.* 109, 171–247. doi: 10.1016/b978-0-12-397920-9.00002-0
- Fürst-Jansen, J. M., de Vries, S., and de Vries, J. (2020). Evo-physio: on stress responses and the earliest land plants. *J. Exp. Bot.* 71, 3254–3269. doi: 10.1093/jxb/eraa007
- Hanks, S. K., and Hunter, T. (1995). The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification I. *FASEB J.* 9, 576–596. doi: 10.1096/fasebj.9.8.7768349
- Hanks, S. K., and Quinn, A. M. (1991). Protein kinase catalytic domain sequence database: identification of conserved features of primary structure and classification of family members. *Methods Enzymol.* 200, 38–62. doi: 10.1016/0076-6879(91)00126-h
- Hanks, S. K., Quinn, A. M., and Hunter, T. (1988). The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. *Science* 241, 42–52. doi: 10.1126/science.3291115
- Harris, B. J., Harrison, C. J., Hetherington, A. M., and Williams, T. A. (2020). Phylogenomic evidence for the monophyly of Bryophytes and the reductive evolution of stomata. *Curr. Biol. : CB* 30, 2001.e2–2012.e2. doi: 10.1016/j.cub.2020.03.048
- Hashiguchi, A., and Komatsu, S. (2016). Impact of post-translational modifications of crop proteins under abiotic stress. *Proteomes* 4:42. doi: 10.3390/proteomes4040042
- Hayes, J. Jr., Underwood, R. M., and Pettis, J. (2008). A survey of honey bee colony losses in the US, fall 2007 to spring 2008. *PLoS One* 3:e4071. doi: 10.1371/journal.pone.0004071
- Hemmer, W., McGlone, M., Tsigelny, I., and Taylor, S. S. (1997). Role of the glycine triad in the ATP-binding site of cAMP-dependent protein kinase. *J. Biol. Chem.* 272, 16946–16954. doi: 10.1074/jbc.272.27.16946
- Hirsch, C. N., Hirsch, C. D., Brohammer, A. B., Bowman, M. J., Soifer, I., Barad, O., et al. (2016). Draft assembly of elite inbred line PH207 provides insights into genomic and transcriptome diversity in maize. *Plant Cell* 28, 2700–2714. doi: 10.1105/tpc.16.00353
- Hori, K., Maruyama, F., Fujisawa, T., Togashi, T., Yamamoto, N., Seo, M., et al. (2014). *Klebsormidium flaccidum* genome reveals primary factors for plant terrestrial adaptation. *Nat. Commun.* 5:3978. doi: 10.1038/ncomms4978
- Hosotani, S., Yamauchi, S., Kobayashi, H., Fuji, S., Koya, S., Shimazaki, K.-I., et al. (2021). A BLUS1 kinase signal and a decrease in intercellular CO<sub>2</sub> concentration are necessary for stomatal opening in response to blue light. *Plant Cell* 33, 1813–1827. doi: 10.1093/plcell/koab067
- Hruz, T., Laule, O., Szabo, G., Wessendorp, F., Bleuler, S., Oertle, L., et al. (2008). Genevestigator V3: a reference expression database for the meta-analysis of transcriptomes. *Adv. Bioinform.* 2008:420747. doi: 10.1155/2008/420747
- Hu, T. T., Pattyn, P., Bakker, E. G., Cao, J., Cheng, J.-F., Clark, R. M., et al. (2011). The *Arabidopsis lyrata* genome sequence and the basis of rapid genome size change. *Nat. Genet.* 43, 476–481. doi: 10.1038/ng.807
- Ishizaki, K. (2017). Evolution of land plants: insights from molecular studies on basal lineages. *Biosci. Biotechnol. Biochem.* 81, 73–80. doi: 10.1080/09168451.2016.1224641

- Ishizaki, K., Nishihama, R., Yamato, K. T., and Kohchi, T. (2016). Molecular genetic tools and techniques for *Marchantia polymorpha* Research. *Plant Cell Physiol.* 57, 262–270. doi: 10.1093/pcp/pcv097
- Jagodzik, P., Tajdel-Zielinska, M., Ciesla, A., Marczak, M., and Ludwikow, A. (2018). Mitogen-activated protein kinase cascades in plant hormone signaling. *Front. Plant Sci.* 9:1387. doi: 10.3389/fpls.2018.01387
- Jaillon, O., Aury, J.-M., Noel, B., Policriti, A., Clepet, C., Casagrande, A., et al. (2007). The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature* 449, 463–467. doi: 10.1038/nature06148
- Jiang, Y., Han, B., Zhang, H., Mariappan, K. G., Bigeard, J., Colcombet, J., et al. (2019). MAP4K4 associates with BIK1 to regulate plant innate immunity. *EMBO Rep.* 20:e47965. doi: 10.15252/embr.201947965
- Jonak, C., Ökrész, L., Bögre, L., and Hirt, H. (2002). Complexity, cross talk and integration of plant MAP kinase signalling. *Curr. Opin. Plant Biol.* 5, 415–424. doi: 10.1016/s1369-5266(02)00285-6
- Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C., et al. (2014). InterProScan 5: genome-scale protein function classification. *Bioinformatics* 30, 1236–1240. doi: 10.1093/bioinformatics/btu031
- Kannan, N., and Neuwald, A. F. (2005). Did protein kinase regulatory mechanisms evolve through elaboration of a simple structural component? *J. Mol. Biol.* 351, 956–972. doi: 10.1016/j.jmb.2005.06.057
- Korf, I., Flicek, P., Duan, D., and Brent, M. R. (2001). Integrating genomic homology into gene structure prediction. *Bioinformatics* 17, S140–S148. doi: 10.1093/bioinformatics/17.suppl\_1.S140
- Krysan, P. J., and Colcombet, J. (2018). Cellular complexity in MAPK signaling in plants: questions and emerging tools to answer them. *Front. Plant Sci.* 9:1674. doi: 10.3389/fpls.2018.01674
- Lamesch, P., Berardini, T., Li, D., Swarbreck, D., Wilks, C., Sasidharan, R., et al. (2012). A gene expression map of *Arabidopsis thaliana* development. *Nucleic Acids Res.* 40, D1202–D1210. doi: 10.1038/ng1543
- Lang, D., Ullrich, K. K., Murat, F., Fuchs, J., Jenkins, J., Haas, F. B., et al. (2018). The *Physcomitrella patens* chromosome-scale assembly reveals moss genome structure and evolution. *Plant J.* 93, 515–533. doi: 10.1111/tbj.13801
- Leberer, E., Dignard, D., Harcus, D., Thomas, D. Y., and Whiteway, M. (1992). The protein kinase homologue Ste20p is required to link the yeast pheromone response G-protein beta gamma subunits to downstream signalling components. *EMBO J.* 11, 4815–4824. doi: 10.1002/j.1460-2075.1992.tb05587.x
- Lee, D., Redfern, O., and Orengo, C. (2007). Predicting protein function from sequence and structure. *Nat. Rev. Mol. Cell Biol.* 8, 995–1005. doi: 10.1038/nrm2281
- Lee, S., Devamani, T., Song, H. D., Sandhu, M., Larsen, A., Sommese, R., et al. (2017). Distinct structural mechanisms determine substrate affinity and kinase activity of protein kinase Ca. *J. Biol. Chem.* 292, 16300–16309. doi: 10.1074/jbc.M117.804781
- Leebens-Mack, J. H., Barker, M. S., Carpenter, E. J., Deyholos, M. K., Gitzendanner, M. A., Graham, S. W., et al. (2019). One thousand plant transcriptomes and the phylogenomics of green plants. *Nature* 574, 679–685. doi: 10.1038/s41586-019-1693-2
- Leliaert, F., Smith, D. R., Moreau, H., Herron, M. D., Verbruggen, H., Delwiche, C. F., et al. (2012). Phylogeny and molecular evolution of the green algae. *Crit. Rev. Plant Sci.* 31, 1–46. doi: 10.1080/07352689.2011.615705
- Leprince, A.-S., Jouannic, S., Hamal, A., Kreis, M., and Henry, Y. (1999). Molecular characterisation of plant cDNAs BnMAP4K $\alpha$ 1 and BnMAP4K $\alpha$ 2 belonging to the GCK/SPS1 subfamily of MAP kinase kinase kinase. *Biochim. Biophys. Acta (BBA)-Gene Struct. Exp.* 1444, 1–13. doi: 10.1016/S0167-4781(98)00246-2
- Letunic, I., and Bork, P. (2019). Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res.* 47, W256–W259. doi: 10.1093/nar/gkz239
- Li, F.-W., Brouwer, P., Carretero-Paulet, L., Cheng, S., De Vries, J., Delaux, P.-M., et al. (2018). Fern genomes elucidate land plant evolution and cyanobacterial symbioses. *Nat. Plants* 4, 460–472. doi: 10.1038/s41477-018-0188-8
- Liu, S., Liu, Y., Yang, X., Tong, C., Edwards, D., Parkin, I. A., et al. (2014). The *Brassica oleracea* genome reveals the asymmetrical evolution of polyploid genomes. *Nat. Commun.* 5, 1–11. doi: 10.1038/ncomms4930
- Liu, Y., and Zhou, J. (2018). MAPping kinase regulation of ICE1 in freezing tolerance. *Trends Plant Sci.* 23, 91–93. doi: 10.1016/j.tplants.2017.12.002
- Llompart, B., Castells, E., Río, A., Roca, R., Ferrando, A., Stiefel, V., et al. (2003). The direct activation of MIK, a germinal center kinase (GCK)-like kinase, by MARK, a maize atypical receptor kinase, suggests a new mechanism for signaling through kinase-dead receptors. *J. Biol. Chem.* 278, 48105–48111. doi: 10.1074/jbc.M307482200
- Machida, N., Umikawa, M., Takei, K., Sakima, N., Myagmar, B.-E., Taira, K., et al. (2004). Mitogen-activated protein kinase kinase kinase 4 as a putative effector of Rap2 to activate the c-Jun N-terminal kinase. *J. Biol. Chem.* 279, 15711–15714. doi: 10.1074/jbc.C300542200
- Major, G., Daigle, C., Stafford-Richard, T., Tebbji, F., Lafleur, E., Caron, S., et al. (2009). Characterization of ScMAP4K1, a MAP kinase kinase kinase involved in ovule, seed and fruit development in *Solanum chacoense* Bitt. *Plant Biol.* 10, 27–46.
- McCormick, R. F., Truong, S. K., Sreedasyam, A., Jenkins, J., Shu, S., Sims, D., et al. (2018). The *Sorghum bicolor* reference genome: improved assembly, gene annotations, a transcriptome atlas, and signatures of genome organization. *Plant J.* 93, 338–354. doi: 10.1111/tbj.13781
- Merchant, S. S., Prochnik, S. E., Vallon, O., Harris, E. H., Karpowicz, S. J., Witman, G. B., et al. (2007). The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. *Science (New York, NY)* 318, 245–250. doi: 10.1126/science.1143609
- Mészáros, B., Erdős, G., and Dosztányi, Z. (2018). IUPred2A: context-dependent prediction of protein disorder as a function of redox state and protein binding. *Nucleic Acids Res.* 46, W329–W337. doi: 10.1093/nar/gky384
- Modi, V., and Dunbrack, R. L. (2019). Defining a new nomenclature for the structures of active and inactive kinases. *Proc. Natl. Acad. Sci. U.S.A.* 116, 6818–6827. doi: 10.1073/pnas.1814279116
- Moreno-Hagelsieb, G., and Latimer, K. (2008). Choosing BLAST options for better detection of orthologs as reciprocal best hits. *Bioinformatics* 24, 319–324. doi: 10.1093/bioinformatics/btm585
- Morris, J. L., Puttick, M. N., Clark, J. W., Edwards, D., Kenrick, P., Pressel, S., et al. (2018). The timescale of early land plant evolution. *Proc. Natl. Acad. Sci. U.S.A.* 115, E2274–E2283. doi: 10.1073/pnas.1719588115
- Motamayor, J. C., Mockaitis, K., Schmutz, J., Haiminen, N., Livingstone, D., Cornejo, O., et al. (2013). The genome sequence of the most widely cultivated cacao type and its use to identify candidate genes regulating pod color. *Genome Biol.* 14, 1–25. doi: 10.1186/gb-2013-14-6-r53
- Neuwald, A. F., and Altschul, S. F. (2016). Bayesian top-down protein sequence alignment with inferred position-specific gap penalties. *PLoS Comput. Biol.* 12:e1004936. doi: 10.1371/journal.pcbi.1004936
- Nishiyama, T., Sakayama, H., de Vries, J., Buschmann, H., Saint-Marcoux, D., Ullrich, K. K., et al. (2018). The *Chara* genome: secondary complexity and implications for plant terrestrialization. *Cell* 174, 448.e4–464.e4. doi: 10.1016/j.cell.2018.06.033
- Nystedt, B., Street, N. R., Wetterbom, A., Zuccolo, A., Lin, Y.-C., Scofield, D. G., et al. (2013). The Norway spruce genome sequence and conifer genome evolution. *Nature* 497, 579–584. doi: 10.1038/nature12211
- Ouyang, S., Zhu, W., Hamilton, J., Lin, H., Campbell, M., Childs, K., et al. (2007). The TIGR rice genome annotation resource: improvements and new features. *Nucleic Acids Res.* 35, D883–D887. doi: 10.1093/nar/gkl976
- Palenik, B., Grimwood, J., Aerts, A., Rouzé, P., Salamov, A., Putnam, N., et al. (2007). The tiny eukaryote *Ostreococcus* provides genomic insights into the paradox of plankton speciation. *Proc. Natl. Acad. Sci. U.S.A.* 104:7705. doi: 10.1073/pnas.0611046104
- Pan, L., and De Smet, I. (2020). Expanding the mitogen-activated protein kinase (MAPK) universe: an update on MAP4Ks. *Front. Plant Sci.* 11:1220. doi: 10.3389/fpls.2020.01220
- Paul, A., and Srinivasan, N. (2020). Genome-wide and structural analyses of pseudokinases encoded in the genome of *Arabidopsis thaliana* provide functional insights. *Proteins: Struct. Funct. Bioinform.* 88, 1620–1638. doi: 10.1002/prot.25981
- Pires, N. D., and Dolan, L. (2012). Morphological evolution in land plants: new designs with old genes. *Philosop. Trans. R. Soc. B: Biol. Sci.* 367, 508–518. doi: 10.1098/rstb.2011.0252
- Poinat, P., De Arcangelis, A., Sookhareea, S., Zhu, X., Hedgcock, E. M., Labouesse, M., et al. (2002). A conserved interaction between  $\beta$ 1 integrin/PAT-3 and Nck-interacting kinase/MIG-15 that mediates commissural axon navigation in *C. elegans*. *Curr. Biol.* 12, 622–631. doi: 10.1016/s0960-9822(02)00764-9



- Poveda, J. (2020). *Marchantia polymorpha* as a model plant in the evolutionary study of plant-microorganism interactions. *Curr. Plant Biol.* 23:100152. doi: 10.1016/j.cpb.2020.100152
- Prochnik, S. E., Umen, J., Nedelcu, A. M., Hallmann, A., Miller, S. M., Nishii, I., et al. (2010). Genomic analysis of organismal complexity in the multicellular green alga *Volvox carterii*. *Science* 329:223. doi: 10.1126/science.1188800
- Qiao, X., Li, Q., Yin, H., Qi, K., Li, L., Wang, R., et al. (2019). Gene duplication and evolution in recurring polyploidization–diploidization cycles in plants. *Genome Biol.* 20:38. doi: 10.1186/s13059-019-1650-2
- Roskoski, R. Jr. (2004). Src protein–tyrosine kinase structure and regulation. *Biochem. Biophys. Res. Commun.* 324, 1155–1164. doi: 10.1016/j.bbrc.2004.09.171
- Sato, S., Tabata, S., Hirakawa, H., Asamizu, E., Shirasawa, K., Isobe, S., et al. (2012). The tomato genome sequence provides insights into fleshy fruit evolution. *Nature* 485, 635–641. doi: 10.1038/nature11119
- Schmutz, J., Cannon, S. B., Schlueter, J., Ma, J., Mitros, T., Nelson, W., et al. (2010). Genome sequence of the palaeopolyploid soybean. *Nature* 463, 178–183. doi: 10.1038/nature08670
- Schnabel, J., Hombach, P., Waksman, T., Giuriani, G., Petersen, J., and Christie, J. M. (2018). A chemical genetic approach to engineer phototropin kinases for substrate labeling. *J. Biol. Chem.* 293, 5613–5623. doi: 10.1074/jbc.RA118.001834
- Sells, M. A., and Chernoff, J. (1997). Emerging from the Pak: the p21-activated protein kinase family. *Trends Cell Biol.* 7, 162–167. doi: 10.1016/S0962-8924(97)01003-9
- Seo, G., Han, H., Vargas, R. E., Yang, B., Li, X., and Wang, W. (2020). MAP4K interactome reveals STRN4 as a key STRIPAK complex component in Hippo pathway regulation. *Cell Rep.* 32:107860. doi: 10.1016/j.celrep.2020.107860
- Sharma, S. K., Bolser, D., de Boer, J., Sønderkær, M., Amoros, W., Carboni, M. F., et al. (2013). Construction of reference chromosome-scale pseudomolecules for potato: integrating the potato genome with genetic and physical maps. *G3 Genes Genomes Genet.* 3, 2031–2047. doi: 10.1534/g3.113.007153
- Steinberg, S. F. (2018). Post-translational modifications at the ATP-positioning G-loop that regulate protein kinase activity. *Pharmacol. Res.* 135, 181–187. doi: 10.1016/j.phrs.2018.07.009
- Taira, K., Umikawa, M., Takei, K., Myagmar, B.-E., Shinzato, M., Machida, N., et al. (2004). The Traf2 and Nck-interacting kinase as a putative effector of Rap2 to regulate actin cytoskeleton. *J. Biol. Chem.* 279, 49488–49496. doi: 10.1074/jbc.M406370200
- Takemiya, A., Sugiyama, N., Fujimoto, H., Tsutsumi, T., Yamauchi, S., Hiyama, A., et al. (2013). Phosphorylation of BLUS1 kinase by phototropins is a primary step in stomatal opening. *Nat. Commun.* 4, 1–8. doi: 10.1038/ncomms3094
- Torkamani, A., Kannan, N., Taylor, S. S., and Schork, N. J. (2008). Congenital disease SNPs target lineage specific structural elements in protein kinases. *Proc. Natl. Acad. Sci. U.S.A.* 105, 9011–9016. doi: 10.1073/pnas.0802403105
- Tuskan, G. A., DiFazio, S., Jansson, S., Bohlmann, J., Grigoriev, I., Hellsten, U., et al. (2006). The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* 313:1596. doi: 10.1126/science.1128691
- Vu, L. D., Gevaert, K., and De Smet, I. (2018). Protein language: post-translational modifications talking to each other. *Trends Plant Sci.* 23, 1068–1080. doi: 10.1016/j.tplants.2018.09.004
- Vu, L. D., Xu, X., Zhu, T., Pan, L., Van Zanten, M., De Jong, D., et al. (2021). The membrane-localized protein kinase MAP4K4/TOT3 regulates thermomorphogenesis. *Nat. Commun.* 12, 1–14. doi: 10.1038/s41467-021-23112-0
- Wan, T., Liu, Z.-M., Li, L.-F., Leitch, A. R., Leitch, I. J., Lohaus, R., et al. (2018). A genome for gnetophytes and early evolution of seed plants. *Nat. Plants* 4, 82–89. doi: 10.1038/s41477-017-0097-2
- Wang, S., Li, L., Li, H., Sahu, S. K., Wang, H., Xu, Y., et al. (2020). Genomes of early-diverging streptophyte algae shed light on plant terrestrialization. *Nat. Plants* 6, 95–106. doi: 10.1038/s41477-019-0560-3
- Winter, D., Vinegar, B., Nahal, H., Ammar, R., Wilson, G. V., and Provart, N. J. (2007). An “Electronic Fluorescent Pictograph” browser for exploring and analyzing large-scale biological data sets. *PLoS One* 2:e718. doi: 10.1371/journal.pone.0000718
- Worden, A. Z., Lee, J.-H., Mock, T., Rouzé, P., Simmons, M. P., Aerts, A. L., et al. (2009). Green evolution and dynamic adaptations revealed by genomes of the marine picoeukaryotes *Micromonas*. *Science* 324:268. doi: 10.1126/science.1167222
- Xiong, J., Cui, X., Yuan, X., Yu, X., Sun, J., and Gong, Q. (2016). The Hippo/STE20 homolog SIK1 interacts with MOB1 to regulate cell proliferation and cell expansion in Arabidopsis. *J. Exp. Bot.* 67, 1461–1475. doi: 10.1093/jxb/erv538
- Xu, B.-E., English, J. M., Wilsbacher, J. L., Stippes, S., Goldsmith, E. J., and Cobb, M. H. (2000). WNK1, a novel mammalian serine/threonine protein kinase lacking the catalytic lysine in subdomain II. *J. Biol. Chem.* 275, 16795–16801. doi: 10.1074/jbc.275.22.16795
- Xu, J., and Zhang, S. (2015). Mitogen-activated protein kinase cascades in signaling plant growth and development. *Trends Plant Sci.* 20, 56–64. doi: 10.1016/j.tplants.2014.10.001
- Yan, H., Zhao, Y., Shi, H., Li, J., Wang, Y., and Tang, D. (2018). BRASSINOSTEROID-SIGNALING KINASE1 phosphorylates MAPKKK5 to regulate immunity in Arabidopsis. *Plant Physiol.* 176, 2991–3002. doi: 10.1104/pp.17.01757
- Young, N. D., DeBellé, F., Oldroyd, G. E. D., Geurts, R., Cannon, S. B., Udvardi, M. K., et al. (2011). The *Medicago* genome provides insight into the evolution of rhizobial symbioses. *Nature* 480, 520–524. doi: 10.1038/nature10625
- Zhang, J., Fu, X.-X., Li, R.-Q., Zhao, X., Liu, Y., Li, M.-H., et al. (2020). The hornwort genome and early land plant evolution. *Nat. Plants* 6, 107–118. doi: 10.1038/s41477-019-0588-4
- Zhang, K., Wang, X., and Cheng, F. (2019). Plant polyploidy: origin, evolution, and its influence on crop domestication. *Horticult. Plant J.* 5, 231–239. doi: 10.1016/j.hpj.2019.11.003
- Zhang, L., Wang, J.-C., Hou, L., Cao, P.-R., Wu, L., Zhang, Q.-S., et al. (2015). Functional role of histidine in the conserved His-x-Asp motif in the catalytic core of protein kinases. *Sci. Rep.* 5:10115. doi: 10.1016/j.hpj.2019.11.003
- Zhang, M., Chiang, Y.-H., Toruño, T. Y., Lee, D., Ma, M., Liang, X., et al. (2018). The MAP4 kinase SIK1 ensures robust extracellular ROS burst and antibacterial immunity in plants. *Cell Host Microbe* 24, 379.e5–391.e5. doi: 10.1016/j.chom.2018.08.007
- Zhu, Q., Shao, Y., Ge, S., Zhang, M., Zhang, T., Hu, X., et al. (2019). A MAPK cascade downstream of IDA-HAE/HSL2 ligand–receptor pair in lateral root emergence. *Nat. Plants* 5, 414–423. doi: 10.1038/s41477-019-0396-x
- Zimmermann, P., Hennig, L., and Grussem, W. (2005). Gene-expression analysis and network discovery using Genevestigator. *Trends Plant Sci.* 10, 407–409. doi: 10.1016/j.tplants.2005.07.003
- Zimmermann, P., Hirsch-Hoffmann, M., Hennig, L., and Grussem, W. (2004). GENEVESTIGATOR. Arabidopsis microarray database and analysis toolbox. *Plant Physiol.* 136, 2621–2632. doi: 10.1104/pp.104.046367

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Pan, Fonseca De Lima, Vu and De Smet. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.