

Ancient signals: comparative genomics of green plant CDPKs

Louis-Philippe Hamel¹, Jen Sheen², and Armand Séguin³

¹ Département de Biologie, Université de Sherbrooke, Sherbrooke, QC J1K 2R1, Canada

² Department of Molecular Biology and Center for Computational and Integrative Biology, Massachusetts General Hospital, Department of Genetics, Harvard Medical School, Boston, MA 02114, USA

³ Natural Resources Canada, Canadian Forest Service, Laurentian Forestry Centre, 1055 rue du P.E.P.S., P.O. Box 10380, Stn. Sainte-Foy, QC G1V 4C7, Canada

Calcium-dependent protein kinases (CDPKs) are multifunctional proteins that combine calcium-binding and signaling capabilities within a single gene product. This unique versatility enables multiple plant biological processes to be controlled, including developmental programs and stress responses. The genome of flowering plants typically encodes around 30 CDPK homologs that cluster in four conserved clades. In this review, we take advantage of the recent availability of genome sequences from green algae and early land plants to examine how well the previously described CDPK family from angiosperms compares to the broader evolutionary states associated with early diverging green plant lineages. Our analysis suggests that the current architecture of the CDPK family was shaped during the colonization of the land by plants, whereas CDPKs from ancestor green algae have continued to evolve independently.

Signaling through plant calcium sensors

The calcium (Ca^{2+}) ion has long been recognized as a versatile secondary messenger that induces signaling in all eukaryotes. Plants, for instance, use Ca^{2+} signaling to regulate development, as well as responses to biotic and abiotic challenges [1]. Following perception of stimuli, plant cells initiate various signaling pathways that relay information to specific primary targets, including Ca^{2+} channels and transporters [2]. Transport from major Ca^{2+} stores (i.e., apoplasts and organelles) results in local Ca^{2+} fluxes at the plasma membrane or phospholipidic membrane of intracellular organelles. This phenomenon, conceptually referred to as the alteration of Ca^{2+} signatures, is constantly monitored by several sensor proteins, which convert upstream Ca^{2+} signals into appropriate downstream cellular responses [3,4].

Plants possess several types of Ca^{2+} sensors, many of which have the EF-hand motif. This helix–loop–helix

structural domain coordinates a single Ca^{2+} ion, providing direct Ca^{2+} -binding activity to the sensor and resulting in Ca^{2+} -dependent conformational changes that promote signaling. The Ca^{2+} sensing apparatus from plants comprises complex protein families that are represented by the calmodulin (CaM) and calmodulin-like (CML) families [5,6], the family of calcineurin B-like (CBL) proteins [7–9], and the CDPKs [10–12].

Among plant Ca^{2+} sensors, CDPKs are unique because they unite Ca^{2+} -binding and signaling capabilities within a single gene product. This combination probably arose following the early fusion of an upstream protein kinase (PK) gene and a downstream CaM gene [13] to enable immediate and efficient translation of input Ca^{2+} signals into appropriate output phosphorylation events. The aim of this review is to trace the evolution of the CDPK family within species that represent key milestones in the evolution of green plants (Box 1; and see Table S1 in the supplementary material online). We first provide a brief description of the various structural features found in higher-plant CDPKs and then assess the conservation of these features in CDPK homologs from descendants of ancestral green plant lineages.

Structure and activation of prototypical CDPKs

Consistent with their function as Ca^{2+} sensors, CDPKs harbor a C-terminal calmodulin-like domain (CLD) that typically comprises four EF-hand motifs (Figure 1). Structural analyses indicate that EF-hands function in pairs, creating two lobes characterized by differential Ca^{2+} -binding affinities [14,15]. The C-terminal lobe has high Ca^{2+} affinity and binds to the ion despite the low concentration that prevails under basal conditions. Within this state, the C-terminal lobe works as a stabilizing structure by promoting intramolecular bonding between the autoinhibitory region (AIR) and the Ser/Thr protein kinase domain (PKD) (Figure 1). As its designation suggests, the AIR keeps the whole enzyme in an inactive mode by acting as a pseudosubstrate that restricts access to the kinase catalytic center [16,17]. Following the perception of a stimulus, the intracellular concentration of Ca^{2+} increases and EF-hand motifs from the N-terminal lobe (Figure 1) bind to the ion despite low Ca^{2+} affinity [18]. The whole protein then undergoes conformational changes that release auto-inhibition and promote phosphorylation of

Corresponding authors: Hamel, L.-P. (Louis-Philippe.Hamel@USherbrooke.ca); Séguin, A. (Armand.Séguin@RNCAN-NRCan.gc.ca).

Keywords: calcium signaling; CDPKs; green plant lineages; evolution.

1360-1385/\$ – see front matter.

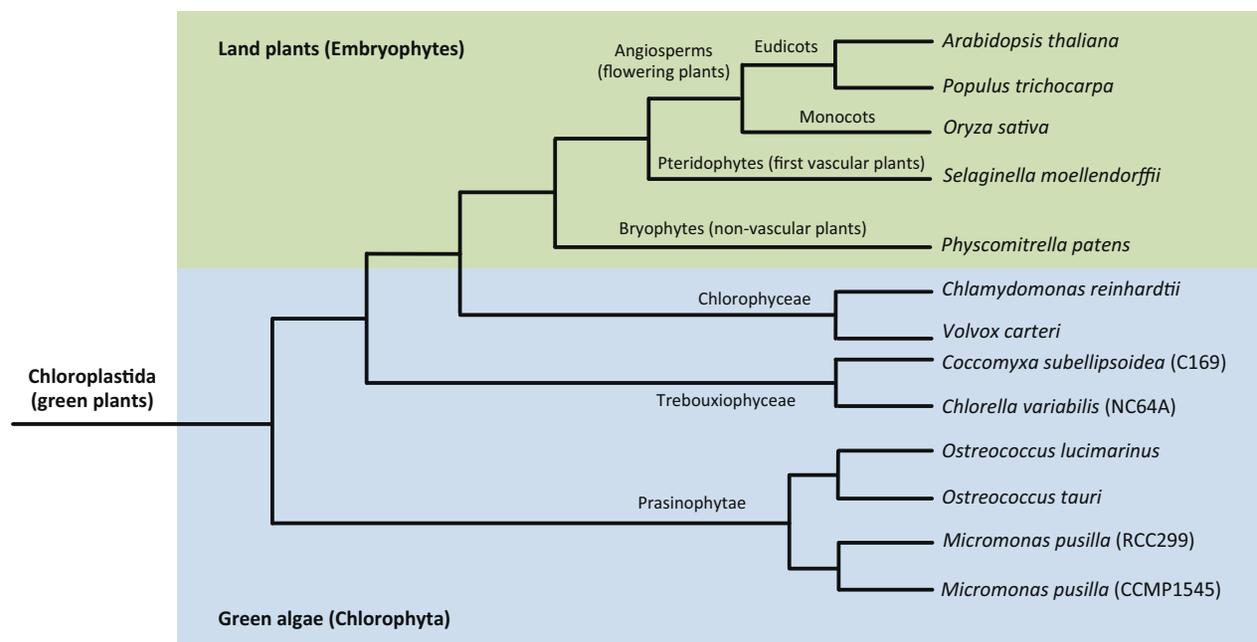
Crown Copyright © 2013 Published by Elsevier Ltd. All rights reserved. <http://dx.doi.org/10.1016/j.tplants.2013.10.009>



Box 1. Evolution of green plants

Chloroplastida (also known as green plants) is one of the three lineages derived from the single primary endosymbiosis of a cyanobacterium and a eukaryotic host cell [72,73]. Relying on their outstanding plasticity with respect to biochemistry, cell architecture, morphology, and general lifestyle, green plants have gradually managed to conquer almost every ecosystem on Earth. The Chloroplastida comprises all species of green algae, as well as embryophytes (also known as land plants) [73]. Despite their diversity, these organisms form a monophyletic group divided into two divisions: the Chlorophyta and the Charophyta, which includes the Streptophyta [65,73–76]. The Chlorophyta comprise a large array of marine and freshwater green algae [65,66,73], including the Prasinophytae (e.g., *Ostreococcus* and *Micromonas*), the Trebouxiophyceae (*Coccomyxa* and *Chlorella*), and the Chlorophyceae (*Chlamydomonas* and *Volvox*) (see Figure 1 and Table S1 in the supplementary material online). About 725 million years ago (Mya)–1200 Mya, a green alga species diverged from the Chlorophyta to become the first in the Streptophyta lineage [77–80]. Descendants from the first Streptophyta species nowadays constitute a small but diverse group of freshwater green algae known as the streptophyte

(or charophycean) algae. The Chlorophyta–Streptophyta split was crucial because ancient streptophyte algae evolved new characters that eventually facilitated colonization of terrestrial habitats around 450–500 Mya [81–83]. The first land plants did not possess vascular tissues, and descendants of these ancient lineages today consist of the liverworts, hornworts, and mosses (collectively referred to as the bryophytes or non-vascular plants) (Figure 1). Ancient bryophytes progressively colonized drier habitats, a trend that promoted differentiation of water-conducting tissues known as the phloem and xylem. Land plants with vascular tissues are referred to as higher plants, and the pteridophytes correspond to the most ancestral lineage within this group (Figure 1). Bryophytes and pteridophytes both rely on spores for reproduction, and the differentiation of seeds by spermatophytes (seed plants) was another key breakthrough that accelerated land plant dispersal. Nowadays, spermatophytes comprise five distinct groups, including the gymnosperms and the angiosperms (Figure 1). Expansion of embryophytes was one of the most important steps for the evolution of life because these organisms forever changed the atmospheric oxygen concentration and contributed significantly to the shaping of terrestrial habitats.



TRENDS in Plant Science

Figure 1. Relationships among green plant lineages used to study the evolution of calcium-dependent protein kinases (CDPKs). Histones are among the most highly conserved genes and were therefore used to reflect the evolutionary path of green plant lineages used in this study. The genome assembly of each of the indicated species was searched using the predicted amino acid sequence of the *Arabidopsis* histone H4 gene (AT1G07660) as a query. Full-length coding sequence (CDS) of histone H4 homologs were next aligned with ClustalW using the following alignment parameters: for pairwise alignment, gap opening 10.0 and gap extension 0.1; and for multiple alignment, gap opening 10.0 and gap extension 0.20. Resulting alignments were submitted to the Molecular Evolutionary Genetics Analysis 4 (MEGA4) software to generate a neighbor-joining tree derived from 5000 replicates. Because the tree only depicts topology among the various taxa, branch size is not proportional to the extent of primary sequence divergence. The blue section comprises species of the Chlorophyta division, a taxonomically diverse group of green algae from which Streptophyta diverged 725–1200 Mya. Streptophyta includes a number of green algae collectively known as the streptophyte algae (not shown here), as well as embryophytes (also referred to as land plants). The green section comprises various species of land plants that represent some of the most important evolutionary milestones characterizing the evolution of terrestrial plants: colonization of the land by bryophytes; differentiation of vascular tissues in pteridophytes; and seeds enclosed in fruits for protection in angiosperms.

downstream substrates, including metabolic enzymes, ion channels, and transcription factors [11,19,20].

At the N terminus, CDPK proteins also contain the variable N-terminal domain (VNTD) (Figure 1), which varies in amino acid sequence and length [21]. Although only limited information is available about the function of the VNTD, recent studies suggest that this region is involved in specificity of substrate recognition [22,23]. The VNTD of most angiosperm CDPKs also comprises

predicted *N*-myristoylation and *S*-palmitoylation sites, two features that promote protein targeting to lipid membranes [11,21,24–27] (Figure 1). Experimentally, membrane-anchoring of a CDPK was first demonstrated in rice (*Oryza sativa*) [28]; however, this phenomenon has since been shown for several *Arabidopsis* (*Arabidopsis thaliana*) candidates, as well as for CDPKs from other flowering plant species [11,12]. Importantly, membrane-targeted CDPKs can also move away from membranes in

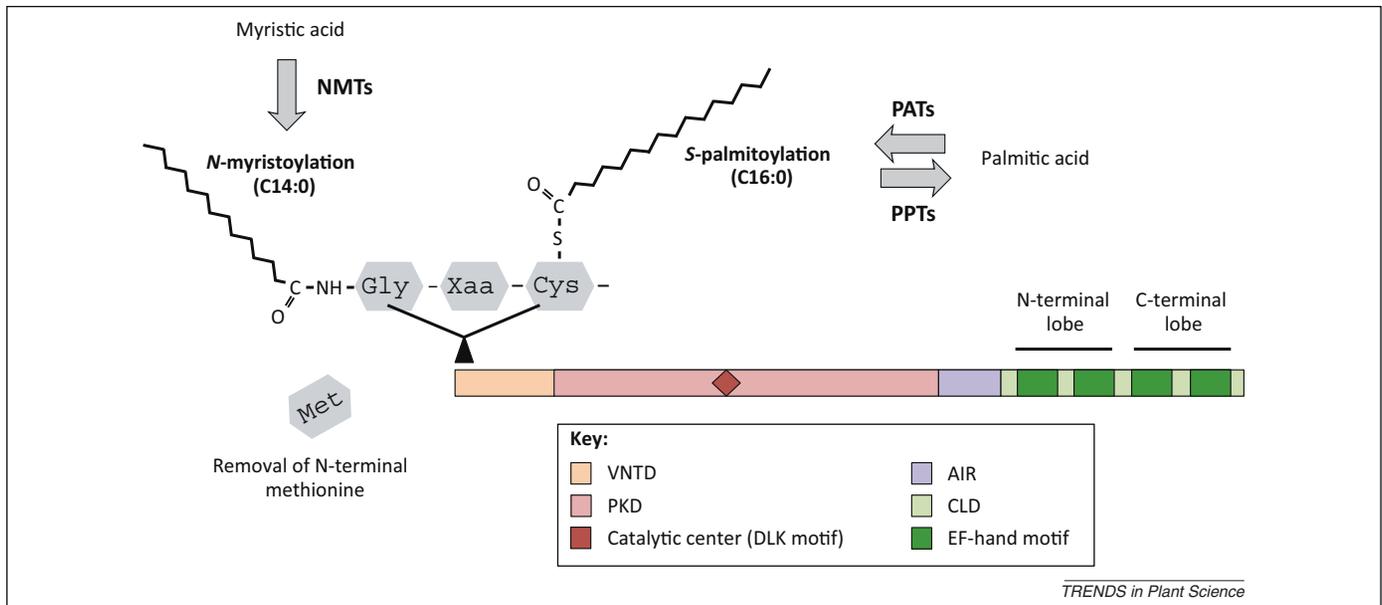


Figure 1. Topology of a prototypical calcium-dependent protein kinase (CDPK) protein. Schematic view of a prototypical CDPK and organization of its conserved domains (see color-coded legend). The N terminus of CDPKs is often targeted for *N*-acylation. *N*-myristoyl transferases (NMTs) catalyze the covalent attachment of myristic acid (a C14:0 fatty acid) to the N-terminal Gly residue (irreversible *N*-myristoylation). Stable anchoring to membranes also requires *S*-palmitoylation, which is performed by palmitoyl acyl transferases (PATs) and palmitoyl protein thioesterases (PPTs). These enzymes catalyze the reversible attachment of palmitic acid (a C16:0 fatty acid) to Cys residues. Abbreviations: AIR, autoinhibitory region; CLD, calmodulin-like domain; PKD, protein kinase domain; VNTD, variable N-terminal domain; Xaa, any amino acid.

response to stress signals [20]. It is likely that this feature enables CDPKs to shuttle between various subcellular compartments as a way to accomplish a broader array of cellular functions.

Evolution of green plant CDPKs

Biochemical evidence showing that CDPKs are calmodulin-independent was first provided more than 25 years ago following the purification and characterization of the first CDPK enzyme from soybean (*Glycine max*) [29]. Since then, CDPK homologs have been identified throughout the plant kingdom, including within species of green algae. Long thought to be plant-specific, CDPK homologs have also been found in two groups of unicellular protists: ciliates and apicomplexan parasites [13,30,31]. The emergence of the first CDPK gene therefore occurred before the basal split between green plants and alveolate protists [13].

The genome of *Arabidopsis* encodes 34 CDPKs that cluster in four well-resolved clades [11,21,32]. Genome sequences from other angiosperms confirm that the basic architecture of the CDPK family is conserved in other flowering plants, including monocot and eudicot species [24–27]. In a previous comparative genomics study, CDPKs from green algae were found to cluster with their flowering plant homologs, whereas CDPKs from protists grouped in a distinct protein subset [31]. The recent completion of genome sequences from early land plants [33,34] and green algae [35–41] now makes it possible to examine the whole CDPK signaling repertoires from descendants of ancient green plant lineages (Box 1 and Table S1). Our analysis provides improved phylogenetic classification of CDPK homologs from green algae. Comparison of the newly identified sequences and the rich CDPK complements from angiosperms also suggests that the basic

architecture of the CDPK family is shared between all land plants, whereas homologs from green algae constitute distinct clades that had not been previously identified.

Green alga CDPKs

Numbers and clustering

Searches of publicly available databases [42,43] have identified 48 putative green alga CDPKs (Table S2 in the supplementary material online). These proteins belong to eight Chlorophyta species, a diverse group of green algae from which Streptophyta (i.e., land plants and precursor streptophyte algae) diverged before colonization of the land by plants (Box 1). Most of the identified CDPK candidates have not yet been characterized, suggesting that our survey would be a valuable resource for those who wish to investigate CDPK functions in green algae.

Following homology-based identification, full-length protein sequences were aligned and phylogenetic relationships were established (Figure 2). Our analysis defines four major clades of green alga CDPKs, two of which have at least one representative member from each of the eight Chlorophyta species examined (Group I and Group IV). Green algae of the Mamiellales order (e.g., the marine species *Micromonas pusilla* and *Ostreococcus* spp.) have particularly small CDPK complements (Table S2) and lack a candidate in Clades II and III (Figure 2). When compared to the freshwater algae *Chlamydomonas reinhardtii* and *Volvox carteri*, the marine species *M. pusilla* and *Ostreococcus* spp. are characterized by extensive gene loss that is associated with reduced genome size [37,39,41]. Gene loss is likely to explain the remarkably small number of CDPK candidates in these species.

With slightly larger CDPK complements, green algae of the Trebouxiophyceae class (*Coccomyxa subellipsoidea* and *Chlorella variabilis*) retain CDPK homologs in Groups I,

correspond to an intermediate evolutionary state that lies between Clades II and III.

Domain organization and membrane targeting

In terms of biochemical properties, CDPKs from green algae are somewhat similar to their angiosperm counterparts (see Tables S2 and S3). Shared features include comparable amino acid length, molecular weight, and isoelectric point (pI). At the structural level, three distinct prediction tools (PROSITE: <http://prosite.expasy.org>, SMART: <http://smart.embl-heidelberg.de>, and Pfam: <http://pfam.sanger.ac.uk>) confirm that green algae possess prototypical CDPKs with four EF-hand motifs (Figure 1), as well as atypical candidates that lack one or more Ca²⁺ binding site(s) at the C terminus (Table S2). This observation confirms that atypical CDPKs are not only found in angiosperms that hold large CDPK families but also in green algae with markedly undersized CDPK complements (e.g., *M. pusilla* and *Ostreococcus* spp.). Recently, a survey has shown that *Arabidopsis* CDPKs have highly variable Ca²⁺ dependences for their kinase activity [44]. Interestingly, low Ca²⁺ sensitivity correlates with structural alterations in the predicted EF-hand motifs, although only one atypical candidate (CPK25) completely lost the ability to bind Ca²⁺. Conservation of atypical CDPKs throughout the green plant lineage suggests that these proteins accomplish vital functions and that better understanding of the functions associated with their unconventional CLD is a key issue in plant signaling [11,44].

Another important feature of CDPKs is the presence of *N*-acylation sites controlling the anchoring of proteins to membranes (Figure 1). Like their flowering plant homologs, several green alga CDPKs harbor predicted *N*-myristoylation and *S*-palmitoylation sites at the N terminus (Table S2). However, the prevalence of these sites seems to be less common than in angiosperms, in which most CDPK candidates carry predicted *N*-acylation sites (Table S3 in the supplementary material online). Because *N*-acylation of cellular proteins is still poorly characterized in Chlorophyta species, we searched genomic data to identify green alga genes that are relevant to this process. Irreversible *N*-myristoylation is performed by *N*-myristoyl transferases (NMTs), whereas reversible *S*-palmitoylation is mediated by palmitoyl acyl transferases (PATs) and palmitoyl protein thioesterases (PPTs) (Figure 1). With the exception of the two *Ostreococcus* species for which no PPT homologs could be identified, it seems that green algae retain the minimal enzymatic machinery necessary to support *N*-acylation of target proteins (Table S4 in the supplementary material online). It can thus be assumed that at least some green alga CDPKs are targeted to lipid membranes.

Domains of the unknown

Several studies have shown that domain organization is highly conserved among CDPK proteins, including more distant homologs from alveolate protists [13,31]. Nonetheless, several unusual CDPKs from apicomplexan parasites possess EF-hand motifs and a pleckstrin homology domain (PHD) at the N terminus [30]. These features are absent in CDPK homologs from green algae, an observation that is consistent with their closer homology with flowering plant

CDPKs [31]. However, further examination of green alga protein sequences has revealed that Cre17.g705000 (*C. reinhardtii*) and Vca20010397m (*V. carteri*) each carry an unusually long VNTD that comprises a C2 domain (see asterisks in Figure 2 and Table S2). With approximately 130 residues, the C2 domain was originally identified as one of two conserved regulatory regions (C1 and C2) within protein kinase C (PKC) [45]. This domain was later found in various signal-transduction and membrane-trafficking proteins, in which it usually functions as a Ca²⁺-dependent module that binds to phospholipids. Given that Cre17.g705000 and Vca20010397 both lack predicted *N*-acylation sites (Table S2), the C2 domain is likely to offer an alternative strategy to guarantee membrane-targeting of these unusual CDPKs.

Although stand-alone C2 domain proteins are encoded by the genome of each Chlorophyta species investigated here (data not shown), CDPKs harboring a C2 domain seem to be uniquely found in green algae of the Chlamydomonadales order (Figure 2 and Table S2). This unexpected combination is likely to result from the fusion of two genes encoding a CDPK and a C2 domain protein, respectively. It remains to be established whether this fusion is lineage-specific, or whether it is a general feature that has been lost in other taxa. A search of the NCBI database (<http://www.ncbi.nlm.nih.gov/>) could only identify two additional CDPKs harboring a C2 domain: that is, DtCPK1 from *Dunaliella tertiolecta* (accession number AAF21062) [46] and CCK1 from *Chlamydomonas moewusii* (CAA89202). These green algae also belong to the Chlamydomonadales order, arguing in favor of a lineage-specific fusion. Overall, our search of genomic data from Chlorophyta species indicates that CDPKs from green algae cluster in four distinct clades and that these PKs have structural features that are generally similar to those observed for angiosperm homologs.

Land plant CDPKs

Numbers and clustering

Angiosperms are the most diverse subset of spermatophytes (plants that produce seeds) and represent the more advanced groups of land plants (Box 1). Although CDPKs from angiosperms have been widely studied [10–12], little is known about the CDPK homologs from more ancient embryophyte lineages, including gymnosperms, pteridophytes, and bryophytes (Box 1). The genome sequences of the bryophyte *Physcomitrella patens* [34] and the pteridophyte *Selaginella moellendorffii* [33] now offer the exciting opportunity to explore CDPK signaling within the oldest land plant lineages.

Searches of genomic data from *P. patens* and *S. moellendorffii* [42,43] have identified 35 novel CDPKs that were compared to 93 homologs from three angiosperm species: that is, the monocot rice and the eudicots poplar (*Populus trichocarpa*) and *Arabidopsis* (Figure S1 and Table S3 in the supplementary material online). Analysis of the phylogenetic relationships has revealed that CDPK homologs from early land plants cluster in four distinct groups (Figure S1), which correspond to the clades formerly identified in various flowering plants [11,21,24–27,32]. Taken as a whole, these evolutionary data indicate that the basal

architecture of the CDPK family is conserved among all land plants: from a primitive bryophyte species to the more sophisticated angiosperms. These results also suggest that large CDPK complements found in angiosperms evolved from the duplication of genes that are common to all embryophytes, rather than from the diversification of lineage-specific candidates that could have fashioned distinct clades of CDPKs.

Expansion through recent gene duplication

In angiosperms, the CDPK family is characterized by the presence of numerous paralogs that share high levels of homology. These closely related genes most probably emerged following recent duplication events and, therefore, have not yet diversified markedly (Figure S1). In some cases, duplication seems to pre-date the monocot–eudicot split (e.g., *OsaCPK17/27*, *AthCPK1/2*, and *PtiCPK1/2*), whereas in others, duplication seems to be specific to a particular lineage (e.g., *OsaCPK12* and *AthCPK29* versus *PtiCPK29-1/29-2*). Duplicated genes arise either by regional genomic events or genome-wide events (also called polyploidization). Whereas polyploidization seems to be the most common mechanism of gene duplication in plants, regional events might still be responsible for a substantial number of duplicated genes. Regional events include local duplication of a single gene (or tandem duplication) as well as transfer of several genes or entire chromosomal segments to more distant chromosomal locations (large segmental duplications). Examination of gene chromosomal distribution indicates that within angiosperms, few CDPK genes occur as tandem duplicates (data not shown). Notable exceptions are *CPK20-1/20-3* and *CPK20-2/20-4* in poplar [27], as well as a cluster of five Group II CDPKs in *Arabidopsis* (*CPK21/22/23/27/31*) [21,32]. As previously shown for mitogen-activated protein kinases (MAPKs) [47], the expansion of the CDPK family from angiosperms thus mainly relied on large-scale DNA rearrangements, namely whole-genome or large segmental duplications.

Gene duplication also seems to be a prominent feature of the CDPK family in *P. patens*. Although this moss is characterized by a relatively simple lifestyle, with 25 distinct members, it retains a CDPK family almost as large as the more complex angiosperms (Table S3). Within each CDPK clade, homologs from *P. patens* cluster tightly (Figure S1), indicating that these genes probably arose following recent duplications. It has been reported that the genome of *P. patens* is characterized by numerous large-scale rearrangements that contributed to the expansion of several gene families, including PKs [34]. A search of the *P. patens* genome could only define four pairs of tandemly duplicated CDPK genes (data not shown), suggesting that large-scale duplications are primarily responsible for the expansion of this protein family. Given that only one bryophyte has been sequenced to date, it is difficult to predict whether recent large-scale duplications contributed to the expansion of the CDPK family in other non-vascular plants. This question is particularly intriguing because support for gene duplication is much less obvious for the CDPK family of *S. moellendorffii* (Figure S1 and Table S3). This species of pteridophyte is characterized by the absence of whole-genome duplication [33], a rare

feature that is likely to explain why this apparently more complex species paradoxically encodes far fewer CDPK homologs (Table S3).

Domain organization and membrane targeting

In terms of biochemical properties, CDPKs from early land plants are highly comparable to their angiosperm homologs (Table S3). For instance, CDPK proteins tend to be slightly acidic, with a pI ranging between five and six. However, a few candidates do not follow this trend and have a basic pI of eight or more. In angiosperms, basic CDPKs are mainly found within Group IV, a feature that is strictly conserved in *P. patens* and *S. moellendorffii* (Table S3). A basic pI might correlate with a specific subcellular localization or function of Group IV CDPKs, which accordingly correspond to the most divergent homologs within this PK family (Figure S1).

Close examination of the protein topology has not yielded any unexpected domains that would have distinguished CDPK homologs from early land plants. It can thus be assumed that the classical organization of CDPKs is conserved among all land plants and that no striking protein innovation has marked this family since the conquest of the land by plants. The genomes of *P. patens* and *S. moellendorffii* were also found to encode prototypical CDPKs with four EF-hand motifs, as well as atypical candidates with only one to three Ca²⁺-binding site(s) at the C terminus (Table S3). Interestingly, atypical homologs account for 34% of all of the CDPKs found in early land plants, whereas in green algae and angiosperms, atypical homologs account for 48% and 21% of all CPDKs, respectively (these values are based on the number of EF-hand motifs as predicted by PROSITE, but similar results have been obtained with SMART or Pfam prediction tools, see Table S3). Atypical CDPKs are therefore conserved throughout the green plant lineage; however, their overall proportion compared to prototypical CDPKs tends to decrease as plants gain in complexity.

The VNTD of early land plant CDPKs has also been scrutinized, and *N*-acylation sites were defined in several candidates (Table S3). Because genes involved in the *N*-acylation process are conserved in *P. patens* and *S. moellendorffii* (Table S4), it is likely that at least some early land plant CDPKs are targeted to membranes. Interestingly, CDPKs lacking *N*-acylation sites mainly cluster within Group I, a feature shared among all land plants (Table S3). We also noticed that 50% of green alga CDPKs are predicted to lack *N*-acylation sites, compared to 31% in early land plants and only 14% in angiosperms. As observed for the ratio of atypical CDPKs (see the previous section), the proportion of non-acylated CDPKs tends to decrease as plant complexity increases.

Given that CDPKs are PKs, we finally compared the predicted catalytic domain of green alga and embryophyte CDPKs. Within each domain examined, 12 characteristic PK subdomains could be defined, and very few amino acid insertions interrupt the catalytic core sections (data not shown). Within PK subdomains, strictly conserved features include residues that help to anchor and orient ATP: that is, the ATP-binding pocket in subdomain I (GXGXXG motif), the invariant Lys in subdomain II,

and the invariant Asp in subdomain VII (DFG motif). Kinase activity also requires a catalytic base for efficient phosphotransfer reaction, a function mediated by the invariant Asp in subdomain VIB (DLK motif). As for above-mentioned features, we found that the DLK motif is conserved in all CDPKs examined. Overall, our analysis indicates that no particular feature is associated with the PKD of the analyzed CDPKs, suggesting that this protein region is well conserved among all green plant lineages.

Clustering and nomenclature of green plant CDPKs

Given that green algae (Figure 2) and embryophytes (Figure S1) each have four distinct clades of CDPKs, it was tempting to hypothesize that these groups would overlap when compared together because the basic architecture of the CDPK family is conserved throughout the green plant lineage. Surprisingly, phylogenetic analysis of the whole-sequence data set revealed that CDPK homologs from green algae do not cluster with counterpart proteins from land plants (Figure 3). As a result, eight CDPK clades were resolved: the first four being specific to Chlorophyta species and the second four being specific to embryophytes. Among the four groups of land plant CDPKs, Group IV seems to be more closely related to homologs from green algae. The ancestral gene that gave rise to this clade in land plants might therefore serve a general function in growth or development. When compared to CDPKs from alveolate protists, green alga CDPKs were shown to cluster with their land plant homologs [31]. However, our analysis emphasizes that fundamental differences distinguish the CDPK families from Chlorophyta and embryophytes (Figure 3).

In light of their differential clustering, we did not attempt to expand the nomenclature of *Arabidopsis* CDPKs [21] to homologs from Chlorophyta species. The authors of a recent comparative genomics survey reached a similar conclusion regarding the nomenclature of green plant MAPKs [48]. By contrast, CDPK homologs from embryophytes form conserved clades (Figure 3) that might rationalize the establishment of a uniform nomenclature. To further support this initiative, we made use of various web-based tools [43,49] to investigate orthologous relationships between the CDPK homologs from land plants.

With the exception of candidates from poplar and *Arabidopsis*, none of the bioinformatics tools was able to predict reliably orthologous relations between embryophyte CDPKs. As a result, we do not propose an expanded nomenclature for CDPK homologs from early land plants and, hence, we used the previously published nomenclature of rice CDPKs [24]. In the case of poplar and *Arabidopsis*, the prediction tools were generally able to yield clear one-to-one affiliations that matched the phylogenetic classification (Figure S1). Based on this pairing, we named poplar CDPKs in a way that reflects sequence relatedness with the closest homologs from *Arabidopsis* (e.g., PtiCPK1 is most similar to AthCPK1). In some cases, additional numbering was included to discriminate between clearly paralogous forms of a given kinase (e.g., PtiCPK3-1 and PtiCPK3-2 versus AtCPK3) [47,50]. We recognize that such a naming and numbering protocol is strictly driven by predicted evolutionary relationships and that

conservation of biological functions within these relationships remains to be established. Nonetheless, systematic nomenclature should facilitate species-to-species comparisons and help the functional characterization of CDPK homologs showing clear orthologous relationships with candidates from the model plant species *Arabidopsis*.

Concluding remarks and future outlook

Our analysis has revealed that the basic architecture of the CDPK family is conserved among all land plants (Figure S1). This includes a bryophyte species that is related to the first green plant lineage that successfully colonized terrestrial habitats (Box 1). By contrast, CDPK homologs from Chlorophyta form alga-specific clades that apparently kept evolving independently (Figure 3). Taken as a whole, these findings suggest that the Chlorophyta–Streptophyta split (Box 1) played a key role in the evolution of CDPKs. However, major architectural changes that affected this protein family were essentially completed after the colonization of the land by plants (Figure S1). As a result, we propose that plant adaptation to terrestrial habitats has worked as a driving force that provided a basis for modern CDPK complements of embryophytes.

Plant lineages that have progressively developed the capacity to survive in dry habitats required a wide range of structural and physiological adaptations to tolerate the novel stresses associated with life on land. These include enhanced tolerance to desiccation and adaptation to direct UV radiation. Interestingly, abiotic stress responses are among the most frequently reported functions of CDPK homologs in angiosperms [10–12]. Notably, *Arabidopsis* CDPKs from Group I (CPK4/6/11), Group II (CPK3/21/23), and Group III (CPK10) have been shown genetically and/or biochemically to be involved in a complex regulatory network controlling stomatal movements in response to drought and abscisic acid [51–56]. Given that stomata control the essential balance between water loss and CO₂ uptake, this useful structural feature is conserved in all land plants, with the notable exception of liverworts, one of the three bryophyte subgroups (Box 1). Interestingly, recent data confirm that stomata from the bryophyte *P. patens* and the pteridophyte *Selaginella uncinata* respond to environmental signals in a comparable way to those of flowering plants [57,58]. These studies also provide strong evidence that intracellular signaling pathways controlling stomatal aperture are similar in both ancestral and modern land plant lineages. Regulation of stomatal aperture by CDPKs could therefore correspond to one of the key functions acquired during the conquest of the land by plants.

CDPK homologs from various plant species have also been frequently involved in the control of biotic stress responses [11,12]. In particular, six *Arabidopsis* CDPKs from Group I have been shown to have a role in innate immunity. CPK1 and CPK2 control the onset of the hypersensitive response (HR) and, along with CPK4 and CPK11, these PKs phosphorylate NADPH oxidases for reactive oxygen species (ROS) production [20]. Treatments with a bacterial elicitor result in transient activation of CPK4/5/6/11, which work as transcriptional regulators of the convergent signaling network downstream of diverse pattern recognition receptors [59]. The prolonged activation of

CPK4/5/6/11 by various bacterial type-III effectors and their cognate NLR (nucleotide-binding domain leucine-rich repeat) intracellular immune sensors leads to the phosphorylation of specific WRKY transcription factors and enhanced plant protection against pathogenic bacteria [20]. Interestingly, some of these defense-related functions (and downstream substrates) seem to be well conserved among Group I CDPKs from various plant species. StCDPK4 and StCDPK5 have for instance been shown to phosphorylate RBOHB (respiratory burst oxidase homolog B), an NADPH oxidase from potato (*Solanum tuberosum*) [60]. Accordingly, overexpression of an active variant of StCDPK5 results in transgenic plants that have increased ROS production and HR-like symptoms [61]. Recent studies also suggest that the *Arabidopsis* CPK5–NADPH oxidase activation circuit has a role in rapid defense signaling propagation [62]. In monocots, the rice OsCPK13 (also referred to as CDPK-7) induces rapid cell death, accumulation of pathogenesis-related (PR) proteins, and upregulation of defense genes when ectopically expressed in *Sorghum bicolor* [63]. Given that every embryophyte lineage experiences invasions by numerous pathogens, the wide conservation of cellular functions within CDPK homologs from Group I suggests that immune signaling through CDPKs also evolved as a result of the new constraint imposed by colonization of the land by plants.

During the course of evolution, CDPK signaling repertoires of land plants have significantly expanded through gene duplications. It is likely that this phenomenon generated novel CDPK homologs that evolved to achieve highly specialized functions, helping plant species to meet a broader array of lineage-specific requirements. Experimental support for this idea also comes from the study of *Arabidopsis*, in which ten CDPK genes (Group I, CPK20/25; Group II, CPK17/19/22/34; Group III, CPK14/24; and Group IV, CPK16/18) show weak or no expression in mesophyll cells [59]. Interestingly, most of these CDPKs are highly and specifically expressed in pollen (data not shown), a situation that is consistent with a recent report showing that CPK17 and CPK34 regulate polarized tip growth in pollen tubes [64]. Taken as a whole, these findings imply that about one-third of *Arabidopsis* CDPKs have specialized reproductive functions instead of more general vegetative growth functions that could be shared with lower land plants and green algae. Each CDPK signaling repertoire thus probably holds a certain level of speciation that will not necessarily be captured by the sole investigation of CDPK homologs from model plants.

In light of the exceptional diversity among land plants, our survey included members of the bryophyte, pteridophyte, and angiosperm lineages (Table S1). These plant species were rationally selected because they represent some of the most important milestones characterizing the evolution of embryophytes (Box 1). Land plants belong to the Streptophyta, early members of which diverged from green algae of the Chlorophyta division (Box 1). Streptophyta not only comprise land plants but also several species of green algae collectively referred to as the streptophyte algae. These organisms are classified into five distinct classes [65,66], and on the basis of different

strategies, several studies have suggested that some of these classes are the sister clade of all land plants [67–71]. Although the true sister clade of land plants is still a matter of debate, it is now widely accepted that characteristics enabling the successful colonization of terrestrial habitats by the first embryophytes have been inherited from ancient streptophyte algae [65]. Descendants of these lineages therefore occupy a key position with regard to our understanding of the evolution of gene families that facilitated colonization of the land by plants. Unfortunately, the study of CDPK homologs from streptophyte algae has been hindered because the genome sequences from these organisms are not yet available. Once processed, this information should help to better define the transition steps that gradually lead to a remodeling of the CDPK family from green algae, eventually leading to the contemporary CDPK signaling repertoires of embryophytes.

Acknowledgments

L.-P.H. was the recipient of a postdoctoral fellowship from the Fonds Québécois de la Recherche sur la Nature et les Technologies (FQRNT). This work was also supported by a grant from the Natural Sciences and Engineering Research Council (NSERC) of Canada to A.S., and a US National Institutes of Health grant (R01 GM70567) to J.S. We apologize to our colleagues whose work could not be cited here owing to space limitations.

Disclosure statement

The authors declare no conflict of interest.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.tplants.2013.10.009](https://doi.org/10.1016/j.tplants.2013.10.009).

References

- Kudla, J. *et al.* (2010) Calcium signals: the lead currency of plant information processing. *Plant Cell* 22, 541–563
- Wheeler, G.L. and Brownlee, C. (2008) Ca²⁺ signalling in plants and green algae-changing channels. *Trends Plant Sci.* 13, 506–514
- Batistič, O. and Kudla, J. (2012) Analysis of calcium signaling pathways in plants. *Biochim. Biophys. Acta* 1820, 1283–1293
- Hashimoto, K. and Kudla, J. (2011) Calcium decoding mechanisms in plants. *Biochimie* 93, 2054–2059
- McCormack, E. *et al.* (2005) Handling calcium signaling: *Arabidopsis* CaMs and CMLs. *Trends Plant Sci.* 10, 383–389
- Perochon, A. *et al.* (2011) Calmodulin and calmodulin-like proteins in plant calcium signaling. *Biochimie* 93, 2048–2053
- Batistič, O. and Kudla, J. (2009) Plant calcineurin B-like proteins and their interacting protein kinases. *Biochim. Biophys. Acta* 1793, 985–992
- Luan, S. (2009) The CBL-CIPK network in plant calcium signaling. *Trends Plant Sci.* 14, 37–42
- Weinl, S. and Kudla, J. (2009) The CBL-CIPK Ca²⁺-decoding signaling network: function and perspectives. *New Phytol.* 184, 517–528
- Asano, T. *et al.* (2012) CDPK-mediated abiotic stress signaling. *Plant Signal. Behav.* 7, 817–821
- Boudsocq, M. and Sheen, J. (2013) CDPKs in immune and stress signaling. *Trends Plant Sci.* 18, 30–40
- Liese, A. and Romeis, T. (2013) Biochemical regulation of *in vivo* function of plant calcium-dependent protein kinases (CDPK). *Biochim. Biophys. Acta* 1833, 1582–1589
- Zhang, X.S. and Choi, J.H. (2001) Molecular evolution of calmodulin-like domain protein kinases (CDPKs) in plants and protists. *J. Mol. Evol.* 53, 214–224
- Christodoulou, J. *et al.* (2004) Evidence for differing roles for each lobe of the calmodulin-like domain in a calcium-dependent protein kinase. *J. Biol. Chem.* 279, 29092–29100

- 15 Rutschmann, F. *et al.* (2002) *LeCPK1*, a calcium-dependent protein kinase from tomato. Plasma membrane targeting and biochemical characterization. *Plant Physiol.* 129, 156–168
- 16 Weljie, A.M. *et al.* (2003) Conformational changes in the Ca²⁺-regulatory region from soybean calcium-dependent protein kinase- α : fluorescence resonance energy transfer studies. *J. Biol. Chem.* 278, 43764–43769
- 17 Weljie, A.M. and Vogel, H.J. (2004) Unexpected structure of the Ca²⁺-regulatory region from soybean calcium-dependent protein kinase- α . *J. Biol. Chem.* 279, 35494–35502
- 18 Harper, J.F. *et al.* (2004) Decoding Ca²⁺ signals through plant protein kinases. *Annu. Rev. Plant Biol.* 55, 263–288
- 19 Curran, A. *et al.* (2011) Calcium-dependent protein kinases from *Arabidopsis* show substrate specificity differences in an analysis of 103 substrates. *Front. Plant Sci.* 2, 36
- 20 Gao, X. *et al.* (2013) Bifurcation of *Arabidopsis* NLR immune signaling via Ca²⁺-dependent protein kinases. *PLoS Pathog.* 9, e1003127
- 21 Cheng, S.H. *et al.* (2002) Calcium signaling through protein kinases. The *Arabidopsis* calcium-dependent protein kinase gene family. *Plant Physiol.* 129, 469–485
- 22 Asai, S. *et al.* (2013) The variable domain of a plant calcium-dependent protein kinase (CDPK) confers subcellular localization and substrate recognition for NADPH oxidase. *J. Biol. Chem.* 288, 14332–14340
- 23 Ito, T. *et al.* (2010) Alteration of substrate specificity: the variable N-terminal domain of tobacco Ca²⁺-dependent protein kinase is important for substrate recognition. *Plant Cell* 22, 1592–1604
- 24 Asano, T. *et al.* (2005) Genome-wide identification of the rice calcium-dependent protein kinase and its closely related kinase gene families: comprehensive analysis of the CDPKs gene family in rice. *Plant Cell Physiol.* 46, 356–366
- 25 Li, A.L. *et al.* (2008) Evolutionary and functional study of the CDPK gene family in wheat (*Triticum aestivum* L.). *Plant Mol. Biol.* 66, 429–443
- 26 Ma, P. *et al.* (2013) Genome-wide identification of the maize calcium-dependent protein kinase gene family. *Appl. Biochem. Biotechnol.* 169, 2111–2125
- 27 Zuo, R. *et al.* (2013) Genome-wide identification, classification, and expression analysis of CDPK and its closely related gene families in poplar (*Populus trichocarpa*). *Mol. Biol. Rep.* 40, 2645–2662
- 28 Martín, M.L. and Busconi, L. (2000) Membrane localization of a rice calcium-dependent protein kinase (CDPK) is mediated by myristoylation and palmitoylation. *Plant J.* 24, 429–435
- 29 Harmon, A.C. *et al.* (1987) A calcium-dependent but calmodulin-independent protein kinase from soybean. *Plant Physiol.* 83, 830–837
- 30 Billker, O. *et al.* (2009) Calcium-dependent signaling and kinases in apicomplexan parasites. *Cell Host Microbe* 5, 612–622
- 31 Harper, J.F. and Harmon, A. (2005) Plants, symbiosis and parasites: a calcium signalling connection. *Nat. Rev. Mol. Cell Biol.* 6, 555–566
- 32 Hrabak, E.M. *et al.* (2003) The *Arabidopsis* CDPK-SnRK superfamily of protein kinases. *Plant Physiol.* 132, 666–680
- 33 Banks, J.A. *et al.* (2011) The *Selaginella* genome identifies genetic changes associated with the evolution of vascular plants. *Science* 332, 960–963
- 34 Rensing, S.A. *et al.* (2008) The *Physcomitrella* genome reveals evolutionary insights into the conquest of land by plants. *Science* 319, 64–69
- 35 Blanc, G. *et al.* (2012) The genome of the polar eukaryotic microalga *Coccomyxa subellipsoidea* reveals traits of cold adaptation. *Genome Biol.* 13, R39
- 36 Blanc, G. *et al.* (2010) The *Chlorella variabilis* NC64A genome reveals adaptation to photosymbiosis, coevolution with viruses, and cryptic sex. *Plant Cell* 22, 2943–2955
- 37 Derelle, E. *et al.* (2006) Genome analysis of the smallest free-living eukaryote *Ostreococcus tauri* unveils many unique features. *Proc. Natl. Acad. Sci. U.S.A.* 103, 11647–11652
- 38 Merchant, S.S. *et al.* (2007) The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. *Science* 318, 245–250
- 39 Palenik, B. *et al.* (2007) The tiny eukaryote *Ostreococcus* provides genomic insights into the paradox of plankton speciation. *Proc. Natl. Acad. Sci. U.S.A.* 104, 7705–7710
- 40 Prochnik, S.E. *et al.* (2010) Genomic analysis of organismal complexity in the multicellular green alga *Volvox carteri*. *Science* 329, 223–226
- 41 Worden, A.Z. *et al.* (2009) Green evolution and dynamic adaptations revealed by genomes of the marine picoeukaryotes *Micromonas*. *Science* 324, 268–272
- 42 Goodstein, D.M. *et al.* (2012) Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res.* 40, 1178–1186
- 43 Van Bel, M. *et al.* (2012) Dissecting plant genomes with the PLAZA comparative genomics platform. *Plant Physiol.* 158, 590–600
- 44 Boudsocq, M. *et al.* (2012) Characterization of *Arabidopsis* calcium-dependent protein kinases: activated or not by calcium? *Biochem. J.* 447, 291–299
- 45 Cho, W. and Stahelin, R.V. (2006) Membrane binding and subcellular targeting of C2 domains. *Biochim. Biophys. Acta* 1761, 838–849
- 46 Pinontoan, R. *et al.* (2000) Cloning of a cDNA encoding a 66-kDa Ca²⁺-dependent protein kinase (CDPK) from *Dunaliella tertiolecta* (Chlorophyta). *J. Phycol.* 36, 545–552
- 47 Hamel, L.P. *et al.* (2006) Ancient signals: comparative genomics of plant MAPK and MAPKK gene families. *Trends Plant Sci.* 11, 192–198
- 48 Dóczy, R. *et al.* (2012) Exploring the evolutionary path of plant MAPK networks. *Trends Plant Sci.* 17, 518–525
- 49 Östlund, G. *et al.* (2010) InParanoid 7: new algorithms and tools for eukaryotic orthology analysis. *Nucleic Acids Res.* 38, D196–D203
- 50 Hamel, L.P. *et al.* (2012) Mitogen-activated protein kinase signaling in plant-interacting fungi: distinct messages from conserved messengers. *Plant Cell* 24, 1327–1351
- 51 Brandt, B. *et al.* (2012) Reconstitution of abscisic acid activation of SLAC1 anion channel by CPK6 and OST1 kinases and branched ABI1 PP2C phosphatase action. *Proc. Natl. Acad. Sci. U.S.A.* 109, 10593–10598
- 52 Geiger, D. *et al.* (2011) Stomatal closure by fast abscisic acid signaling is mediated by the guard cell anion channel SLAH3 and the receptor RCAR1. *Sci. Signal.* 4, ra32
- 53 Geiger, D. *et al.* (2010) Guard cell anion channel SLAC1 is regulated by CDPK protein kinases with distinct Ca²⁺ affinities. *Proc. Natl. Acad. Sci. U.S.A.* 107, 8023–8028
- 54 Mori, I.C. *et al.* (2006) CDPKs CPK6 and CPK3 function in ABA regulation of guard cell S-type anion- and Ca²⁺-permeable channels and stomatal closure. *PLoS Biol.* 4, e327
- 55 Zhu, S.Y. *et al.* (2007) Two calcium-dependent protein kinases, CPK4 and CPK11, regulate abscisic acid signal transduction in *Arabidopsis*. *Plant Cell* 19, 3019–3036
- 56 Zou, J.J. *et al.* (2010) *Arabidopsis* calcium-dependent protein kinase CPK10 functions in abscisic acid- and Ca²⁺-mediated stomatal regulation in response to drought stress. *Plant Physiol.* 154, 1232–1243
- 57 Ruszala, E.M. *et al.* (2011) Land plants acquired active stomatal control early in their evolutionary history. *Curr. Biol.* 21, 1030–1035
- 58 Chater, C. *et al.* (2011) Regulatory mechanism controlling stomatal behavior conserved across 400 million years of land plant evolution. *Curr. Biol.* 21, 1025–1029
- 59 Boudsocq, M. *et al.* (2010) Differential innate immune signalling via Ca²⁺ sensor protein kinases. *Nature* 464, 418–422
- 60 Kobayashi, M. *et al.* (2007) Calcium-dependent protein kinases regulate the production of reactive oxygen species by potato NADPH oxidase. *Plant Cell* 19, 1065–1080
- 61 Kobayashi, M. *et al.* (2012) StCDPK5 confers resistance to late blight pathogen but increases susceptibility to early blight pathogen in potato via reactive oxygen species burst. *New Phytol.* 196, 223–237
- 62 Dubiella, U. *et al.* (2013) Calcium-dependent protein kinase/NADPH oxidase activation circuit is required for rapid defense signal propagation. *Proc. Natl. Acad. Sci. U.S.A.* 110, 8744–8749
- 63 Mall, T.K. *et al.* (2011) Expression of the rice CDPK-7 in *Sorghum*: molecular and phenotypic analyses. *Plant Mol. Biol.* 75, 467–479
- 64 Myers, C. *et al.* (2009) Calcium-dependent protein kinases regulate polarized tip growth in pollen tubes. *Plant J.* 59, 528–539
- 65 Becker, B. (2013) Snow ball earth and the split of Streptophyta and Chlorophyta. *Trends Plant Sci.* 18, 180–183
- 66 Guiry, M.D. (2012) How many species of algae are there? *J. Phycol.* 48, 1057–1063
- 67 Laurin-Lemay, S. *et al.* (2012) Origin of land plants revisited in the light of sequence contamination and missing data. *Curr. Biol.* 22, R593–R594
- 68 Timme, R.E. *et al.* (2012) Broad phylogenomic sampling and the sister lineage of land plants. *PLoS ONE* 7, e29696

- 69 Turmel, M. *et al.* (2006) The chloroplast genome sequence of *Chara vulgaris* sheds new light into the closest green algal relatives of land plants. *Mol. Biol. Evol.* 23, 1324–1338
- 70 Turmel, M. *et al.* (2007) The green algal ancestry of land plants as revealed by the chloroplast genome. *Int. J. Plant Sci.* 23, 679–689
- 71 Wodniok, S. *et al.* (2011) Origin of land plants: do conjugating green algae hold the key? *BMC Evol. Biol.* 11, 104
- 72 Keeling, P.J. (2010) The endosymbiotic origin, diversification and fate of plastids. *Philos. Trans. R. Soc. Lond. B: Biol. Sci.* 365, 729–748
- 73 Adl, S.M. *et al.* (2012) The revised classification of eukaryotes. *J. Eukaryot. Microbiol.* 59, 429–493
- 74 Becker, B. and Marin, B. (2009) Streptophyte algae and the origin of embryophytes. *Ann. Bot.* 103, 999–1004
- 75 Leliaert, F. *et al.* (2012) Phylogeny and molecular evolution of the green algae. *Crit. Rev. Plant Sci.* 31, 1–46
- 76 Lewis, L.A. and McCourt, R.M. (2004) Green algae and the origin of land plants. *Am. J. Bot.* 91, 1535–1556
- 77 Douzery, E.J. *et al.* (2004) The timing of eukaryotic evolution: does a relaxed molecular clock reconcile proteins and fossils? *Proc. Natl. Acad. Sci. U.S.A.* 101, 15386–15391
- 78 Hedges, S.B. *et al.* (2004) A molecular timescale of eukaryote evolution and the rise of complex multicellular life. *BMC Evol. Biol.* 4, 2
- 79 Yoon, H.S. *et al.* (2004) A molecular timeline for the origin of photosynthetic eukaryotes. *Mol. Biol. Evol.* 21, 809–818
- 80 Zimmer, A. *et al.* (2007) Dating the early evolution of plants: detection and molecular clock analyses of orthologs. *Mol. Genet. Genomics* 278, 393–402
- 81 Gensel, P.G. (2008) The earliest land plants. *Annu. Rev. Ecol. Evol. Syst.* 39, 459–477
- 82 Kenrick, P. *et al.* (2012) A timeline for terrestrialization: consequences for the carbon cycle in the Palaeozoic. *Philos. Trans. R. Soc. Lond. B: Biol. Sci.* 367, 519–536
- 83 Sanderson, M.J. *et al.* (2004) Molecular evidence on plant divergence times. *Am. J. Bot.* 91, 1656–1665
- 84 Tamura, K. *et al.* (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* 24, 1596–1599

Plant Science Conferences in 2014

Plant Biology Europe FESPB/EPSO 2014 Congress

Learning from the past, preparing for the future

22–26 June, 2014

Dublin, Ireland

<http://europlantbiology.org/>

XVI International Congress on Molecular Plant-Microbe Interactions

6–10 July, 2014

Rhodes, Greece

<http://www.mpmi2014rhodes-hellas.gr/index.php>

Plant Biology 2014

12–16 July, 2014

Portland, USA

<http://www.aspb.org/meetings>