DNA Research, 2015, 22(6), 471–483
doi: 10.1093/dnares/dsv029
Advance Access Publication Date: 5 November 2015
Full Paper

OXFORD

Full Paper

Phylogeny and evolution of plant cyclic nucleotide-gated ion channel (CNGC) gene family and functional analyses of tomato *CNGCs*

Mumtaz Ali Saand[1], You-Ping Xu[2], Jean-Pierre Munyampundu[1], Wen Li[1], Xuan-Rui Zhang[1], and Xin-Zhong Cai[1,*]

[1]Institute of Biotechnology, College of Agriculture and Biotechnology, Zhejiang University, Hangzhou 310058, China, and [2]Centre of Analysis and Measurement, Zhejiang University, Hangzhou 310058, China

*To whom correspondence should be addressed. Tel. +86 571 88982936. Fax. +86 571 88982936. E-mail: xzhcai@zju.edu.cn

Edited by Dr Kazuo Shinozaki

Received 4 April 2015; Accepted 12 October 2015

Abstract

Cyclic nucleotide-gated ion channels (CNGCs) are calcium-permeable channels that are involved in various biological functions. Nevertheless, phylogeny and function of plant CNGCs are not well understood. In this study, 333 CNGC genes from 15 plant species were identified using comprehensive bioinformatics approaches. Extensive bioinformatics analyses demonstrated that CNGCs of Group IVa were distinct to those of other groups in gene structure and amino acid sequence of cyclic nucleotide-binding domain. A CNGC-specific motif that recognizes all identified plant CNGCs was generated. Phylogenetic analysis indicated that CNGC proteins of flowering plant species formed five groups. However, CNGCs of the non-vascular plant *Physcomitrella patens* clustered only in two groups (IVa and IVb), while those of the vascular non-flowering plant *Selaginella moellendorffii* gathered in four (IVa, IVb, I and II). These data suggest that Group IV CNGCs are most ancient and Group III CNGCs are most recently evolved in flowering plants. Furthermore, silencing analyses revealed that a set of CNGC genes might be involved in disease resistance and abiotic stress responses in tomato and function of *SlCNGC*s does not correlate with the group that they are belonging to. Our results indicate that Group IVa CNGCs are structurally but not functionally unique among plant CNGCs.

Key words: abiotic stress response, cyclic nucleotide-gated channel, gene structure, phylogeny, resistance

1. Introduction

Calcium signal transduction through calcium conducting channels is an indispensable mechanism utilized by plants to sense and respond to internal and environmental stimuli,[1–3] and is involved in various biological processes such as hormone responses,[4] development,[5] light signalling,[6] salt stress[7] and plant–pathogen interaction.[8] The cyclic nucleotide-gated ion channels (CNGCs) are suggested to be one of the important pathways for conducting Ca^{2+} ions in signal transduction.[9] They are ligand-gated, Ca^{2+}-permeable divalent cation-selective channels that are localized in plasma membrane, presumptively are activated by direct binding of cyclic nucleotides and are inhibited by binding of calmodulin (CaM) to the CaM binding (CaMB) domain.[10–15] Plant CNGCs are composed of six transmembrane domains (S1–S6) and a pore region (P) between the fifth and sixth domains, C-terminal CaMB domain and cyclic nucleotide-binding (CNB) domain.[11,16] The CNB domain is highly conserved and carries a plant CNGC-specific motif spanning the phosphate-binding cassette (PBC) and hinge region. This motif identifies CNGCs but no other proteins, hence is recognized as authentic tool to identify plant CNGCs.[16,17] The plant CNGC was first identified in a screen for CaMB proteins in barley,[18] which was followed by identification in Arabidopsis,[19] rice,[20,21] *Populus trichocarpa*,[22] pear,[23] *Selaginella*

moellendorffii,[16] *Physcomitrella patens* and algae.[17,24] The identified 20 Arabidopsis CNGC paralogues were classified into five groups; I, II, III, IVa and IVb.[19] To date, the phylogeny and evolution of CNGC gene family in plant remain unclear.

Prior studies have shown that many members of Arabidopsis CNGC family are implicated in one or several physiological processes, including Ca^{2+} signalling, abiotic stress resistance and defence responses.[25–30] Several members of Arabidopsis CNGC Groups I–III regulate various functions such as plant development and stress tolerance,[25,31] seed germination,[32] plant growth,[33] pollen fertility under stress,[5] pollen tip growth[34] and pathogen defence.[35,36] However, *AtCNGC2* and *AtCNGC4*, members of Group IVb, played their role in disease resistance against various pathogens and in thermotolerance,[29,37–44] while *AtCNGC19* and *AtCNGC20*, members of Group IVa, are involved in abiotic salt stress responses.[19,45–47] Whether plant CNGC genes are functionally distinguished in a group-dependent manner is not clear. Moreover, in our previous study, we found that group IVb *SlCNGC* genes regulate different types of resistances against a wide range of pathogens in tomato, and Group IVa CNGC genes of both Arabidopsis and tomato were distinct to those of all other groups in gene structure and CNGC-specific motifs.[48] Whether this is also the case in other plant species await confirmation.

In this study, we identified CNGC gene family in 15 plant species whose genome has been sequenced at the genome-wide level using comprehensive bioinformatics analyses. Our sequence and phylogenetic data based on 412 CNGC genes from 20 plant species at various positions of evolution revealed for the first time the phylogeny of CNGCs in plant. We also provided evidence that CNGC genes of Group IVa are distinct to those of other groups in gene structure, but function of plant CNGCs is not group dependent.

2. Materials and methods

2.1. Identification of CNGC proteins in plant genomes

BLASTP searches were performed against sequenced genomes of green plants in Phytozome (v9.1) (http://www.phytozome.net/) and in NCBI (http://www.ncbi.nlm.nih.gov/) databases using Arabidopsis and tomato CNGC proteins as queries. Meanwhile, NCBI database searching for sequences containing a plant CNGC-specific motif [LI]-X(2)-[GS]-X-[FYIVS]-X-GX(0,1)-[DE]-LL-X(8,25)-S-X(10)-E-X-F-X-[IL][17] was conducted as well. All retrieved non-redundant sequences were collected and subjected to domain analysis by using the SMART (http://smart.embl-heidelberg.de/) and Conserved Domain Database (CDD) (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi/) programmes. These sequences were compared with Arabidopsis and tomato CNGC proteins using ClustalW2 program (http://www.ebi.ac.uk/Tools/msa/clustalw2/) with default settings and were viewed by GeneDoc program.[49] Those containing both a CNB domain [CNBD, *cNMP* (IPR000595)] and a transmembrane (TM) or ion transport protein (ITP) domain [Ion_trans family (PF00520)] and a plant CNGC-specific motif in the PBC and hinge region within the CNBD were recognized as CNGC proteins. CNGCs in a given species were named in accordance with sequence similarity to Arabidopsis CNGCs in phylogenetic relationship.

2.2. Gene structure, motif and phylogenetic analyses

The exon/intron structure of *CNGC* genes was analysed online using the Gene Structure Display Server (GSDS) (http://gsds.cbi.pku.edu.cn/). The sequence logos were generated online at logo platform (http://weblogo.berkeley.edu/). The CNGC-specific motif in the PBC and hinge region within the CNB domain of all representative plant species were derived after alignment with MEGA 5.0[50] and viewed by GeneDoc.[49] For phylogenetic tree construction, multiple sequence alignments of CNGC proteins from representative plant species were assembled using clustalX 2.01 program.[51] The phylogenetic tree was constructed using MEGA 5.0[50] with maximum likelihood method and bootstrap of 1000.

2.3. Plant material and disease resistance analysis

Tomato (cv. Money maker) plants were grown in growth room at 28° C with16 h light/8 h dark photoperiod. For disease resistance evaluation analyses, tomato plants were inoculated with a variety of pathogens. *Sclerotinia sclerotiorum* (Ss) was grown at 22°C on potato dextrose agar (PDA) medium for 2 days. The PDA plugs of 3 mm at diameter were taken from the colony outside circle that contained most active young mycelia and then were stuck mycelial side down onto the tomato leaves. The inoculated plants were grown under high relative humidity for 24 h. Lesion size was recorded at 30 h post-inoculation. Inoculation and subsequent disease evaluation for bacterial pathogens *Pseudomonas syringae* pv. *tomato* DC3000 (*Pst* DC3000) and *Xanthomonas oryzae* pv. *oryzae* (*Xoo*) were performed as described.[52,53]

2.4. Virus-induced gene silencing analyses

The *SlCNGC* gene members are highly conserved among each other. Therefore, care was taken to ensure the specificity to target the *SlCNGC* genes. The virus-induced gene silencing (VIGS) target fragments of *SlCNGC*1, 6, 7, 8, 11, 14 and 15 were amplified by RT-PCR with gene-specific primers (Supplementary Table S1) and ligated into the TRV VIGS vector pYL156, which was subsequently electroporated into *Agrobacterium tumefaciens* strain GV3101 for VIGS analyses. VIGS analyses were conducted with vacuum infiltration delivery approach as described[54,55] except that recombinant pYL156 with insertion of an eGFP fragment instead of an empty pYL156 was used as control to alleviate viral symptom.[56] At ~3 weeks post agro-infiltration, plants were inoculated with different pathogens and disease was investigated as described above. For each pathogen, at least six silenced plants were examined. The experiments were conducted three times independently.

2.5. Drought and salinity stress tolerance assays of VIGS plants

The plants at 3–4 weeks post silencing were used for both drought and salinity stress tolerance assays. The plants were not watered as drought treatment and the phenotype was observed every day since stop watering. For the salt stress tolerance assay, plants were irrigated with water containing 0.4 M NaCl (200 ml per plant) from the bottoms of the pots. Leaves were sampled after drought and salt treatments to measure relative water content (RWC) and chlorophyll content as described,[57] using the following formulae.

$$\mathrm{RWC}\ (\%) = \left(\frac{\mathrm{FW} - \mathrm{DW}}{\mathrm{TW} - \mathrm{DW}}\right) \times 100$$

$$\mathrm{Chl}\ (\mu\mathrm{g\ ml}^{-1}) = 20.29A_{646} + 8.02A_{663}$$

The plants for both treatments were kept in same condition described above. Ten plants were used for each treatment. Pictures were taken to record the phenotypes. The experiments were conducted three times independently.

3. Results

3.1. Identification of CNGC protein sequences in 15 flowering plant species

Fifteen flowering plant species whose genome has been completely sequenced and locate at different evolutional positions were selected for CNGC gene family identification analyses. These species included three monocots (*Brachypodium distachyon*, *Zea mays* and *Sorghum bicolor*) and 12 eudicots (*Aquilegia coerulea*, *Solanum tuberosum*, *Eucalyptus grandis*, *Citrus sinensis*, *Gossypium raimondii*, *Brassica rapa*, *Malus domestica*, *Cucumis sativus*, *Glycine max*, *P. trichocarpa*, *Linum usitatissimum* and *Ricinus communis*) (Fig. 1). To identify CNGC protein sequences in these species, the amino acid (aa) sequence of 20 Arabidopsis and 18 tomato CNGCs were used as the query to BLASTP search the plant genome databases at Phytozome (v9.1) and NCBI. Additionally, NCBI database was also searched for sequences carrying a plant CNGC-specific motif [LI]-X(2)-[GS]-X-[FYIVS]-X-GX(0,1)-[DE]-LL-X(8,25)-S-X(10)-E-X-F-X-[IL].[17] Consequently, 428 non-redundant putative CNGC gene sequences were retrieved, which were further analysed using SMART and CDD programs to confirm the presence of a cyclic nucleotide-binding (CNB) domain [cyclic nucleotide-monophosphate binding (cNMP) domain in SMART, or Cap family effector (CAP_ED) domain in CDD] and a transmembrane (TM) domain (TM domain in SMART or ITP domain in CDD). Subsequently, the candidate sequences were examined for existence of plant CNGC-specific motif in the PBC and hinge region within CNB domain.[16,17] Finally, of a total of 428 retrieved sequences, 329 were full-length sequences containing both required domains and a CNGC-specific motif and thus were recognized as plant CNGCs. Ninety-one sequences had either no CNGC-specific motif or truncated and hence were discarded. The remaining four pair sequences, including two pairs of *Z. mays* sequences GRMZM2G169788_T01 and GRMZM2G129375_T01, GRMZM2G161800_T01 and GRMZM2G141642_T01, one pair of *C. sinensis* sequences orange1.1g045215 m and orange1.1g045637 m and one pair of *C. sativus* sequences Cucsa.026900.1 and Cucsa.194600.1 were incorrectly separated during annotation and thus were corrected into four complete CNGC genes, which hereafter were designated as *ZmCNGC1*, *ZmCNGC8*, *CsCNGC15* and *CusCNGC16* (Supplementary Table S1). Together, 333 CNGC genes were identified in the listed 15 flowering plant species (Fig. 1; Supplementary Table S2). We assigned names of all individual CNGC members in given species in ascending order in accordance with group numbers on the basis of phylogenetic relationship with 20 Arabidopsis CNGCs (Supplementary Table S2).

Composition of CNGC genes in different plant species varied significantly. All three investigated monocots contained <16 CNGC genes, which is similar to what was reported for monocot rice.[20,21] However, 10 out of total 12 analysed dicots comprised at least 18 CNGC genes and half of them consisted of over 28 CNGC genes. Among them, *M. domestica*, *G. max* and *E. grandis* bore the most CNGC genes, 44, 35 and 31, respectively (Fig. 1). Additionally, size of different groups within an individual species was distinguishable. Group IVa was the smallest, mostly containing only one or two gene

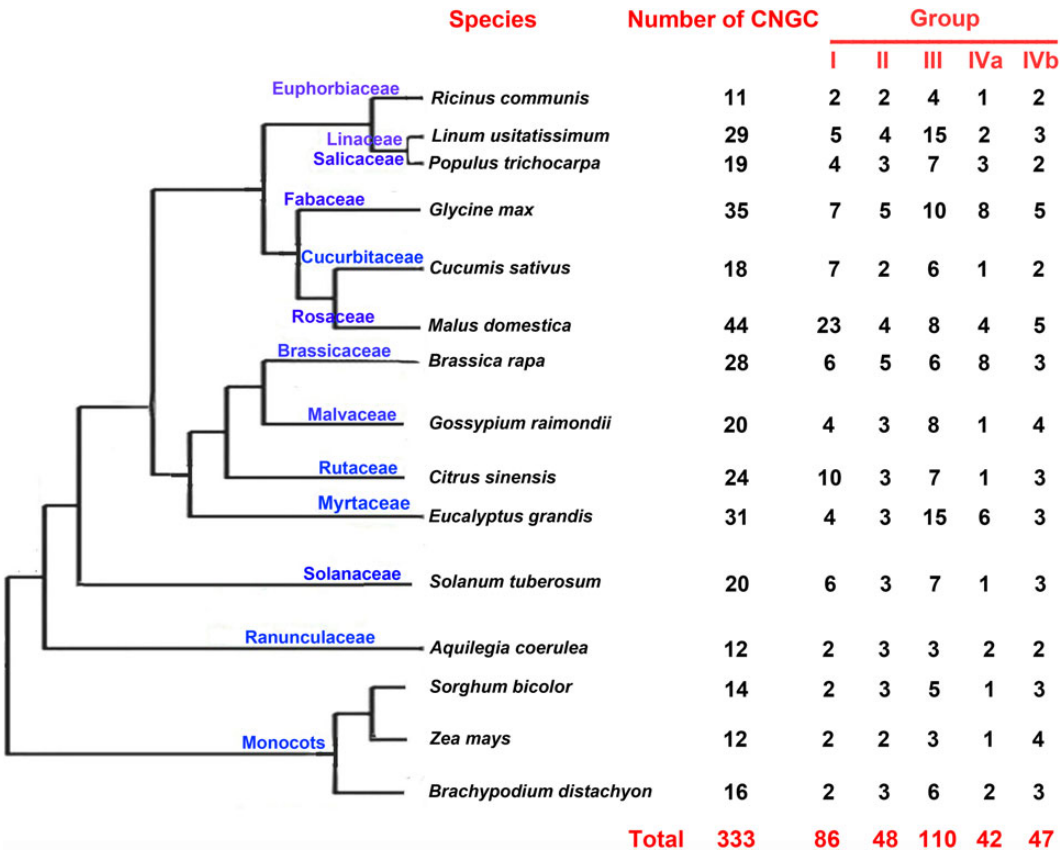

Figure 1. Phylogenetic relationships between the 15 plant species investigated in this study. The total number of CNGC proteins and that of each groups identified in each plant genome is indicated on the right. The phylogenetic tree is modified from Phytozome (http://www.phytozome.net/). This figure is available in black and white in print and in colour at *DNA Research* online.

members, whereas Group III and/or Group I were the largest. Uneven group expansion was evident especially in some species. For example, 23 of total 44 (52%) CNGC genes in *M. domestica* belonged to Group I, while 15 of total 29 (52%) and 15 of total 31 (48%) CNGC genes in *L. usitatissimum* and *E. grandis* belonged to Group III. Similarly, *G. max* and *E. grandis* contained extraordinarily more (8) group IVa CNGC genes (Fig. 1).

3.2. Domain composition and conserved motif at the PBC and hinge region of CNB domain of CNGC proteins in plants

Plant CNGCs comprise a TM/ITP domain and a CNB domain with an overlapped calmodulin-binding (CaMB) domain.[11] However, our domain composition analyses using databases SMART and NCBI-CDD revealed that putative plant CNGC proteins were not only characterized by these well-known domains but also contained other domains in some plant CNGC proteins (Supplementary Table S2). For instance, the CDD database search unveiled that *M. domestica* CNGCs MdCNGC7 and MdCNGC24, two Group I CNGCs, contained DUF2213 (cl19842) and DUF616 (pfam04765) domains, whose function is still unknown. MdCNGC27, a Group II CNGC, carried a LPLATA_GPAT-like domain (cd07989), which is involved in glycerophospholipid biosynthesis. MdCNGC41, a Group IVb CNGC, possessed two other domains, a PTEN_C2 (C2 domain of PTEN tumour-suppressor protein) domain (pfam10409) and a protein tyrosine phosphatase (PTP) domain (cl21483), whereas RcCNGC3, a Group II CNGC, carried a STKc_MAPKKK (Serine/Threonine Kinase, Mitogen-Activated Protein Kinase) domain (cl06606) (Supplementary Table S2). Due to existence of these extra domains, the size of these CNGC proteins was larger than 1000 aa. Nevertheless, whether these domains are indeed functional awaits further experimental confirmation.

It has been suggested that plant CNGC proteins carry a CNGC-specific motif in the PBC and hinge region within CNB domain.[16,17] To confirm whether this motif generally exists in CNGCs of all plant species, we aligned the PBC and hinge region of total 412 CNGC proteins, including 333 from 15 flowering plant species identified in this study and 18 from tomato[48] as well as 61 from Arabidopsis, rice, *S. moellendorffii* and *P. patens* that were previously identified by other labs.[16,17,19] Based on conservation of amino acids at >90%, the motif was deduced as: [L]-X(0,1,2)-[G]-X(3)-G-X(0,1,2)-[E]-L-[L]-X-[W]-X-[L]-X(7,37)-[S]-X(10,11)-[E]-[A]-[F]-X-[L], which was derived from the all-possibility motif: [LIMV0]-X(2)-[GSANCR]-X-[FVIYASCL]-X-G-X(0,1)-X(0,1)-[EDAQGH]-L-[LIVFA]-X-[WRCMLS0]-X-[LMSIQAFT0] -X(7,37)-[SAC]-X(9)-[VTIALMS]-X(0,1)-[EQDN]-[AGSVT]-[FYL]-X-[LIVF], which recognized all 412 plant CNGC proteins identified so far. Notably, this motif contained two invariable amino acids G and L in the PBC. Comparison of this motif with the one suggested earlier, [LI]-X(2)-[GS]-X-[VFIYS]-X-G-X(0,1)-[DE]-L-[LI]-X-[WN]-X(6,32)-[SA]-X(9)-[VTI]-[EN]-[AG]-F-X-[LI],[17] demonstrated that the motif suggested earlier for several plant species[17] is not applicable to all identified plant CNGCs. Except the two invariable amino acids G and L in the PBC, all the remaining positions of the motif had more possibilities than suggested, including the position suggested as invariable F in the hinge region (Supplementary Fig. S1). This F amino acid was replaced by Y in EgCNGC8/9 of Group III and CsCNGC23 of Group IVb while by L in GmCNGC28 of Group IVa (Supplementary Fig. S1). These results demonstrate that it is necessary to relax the motif suggested earlier[17] to the one identified in this study when it is used for identification of plant CNGC.

Furthermore, the CNGC-specific motif for each group was generated. They were [LIV0]-X(2)-[GSNR]-X-[FIYVCL]-X-G-X(0,1)-[EDNKA0]-L-X(2)-[WRCS0]-X(7,22)-[PS]-X-[SC]-X(9)-[VIM]-[EQD]-X-F-X[LIVF] for group I; L-[KQVR]-[ED]-X-[DSAE]-FCG-X-ELLTWALDP-[KR]-X(5)-P-X-S-[TS]-RT-[VA]-X-A-X(2)-EVE-[ASG]-F-[AS]-L for group II; [LIM]-X(2)-[GSCN]-X-F-[CAS]-G-[ED]-ELL-X-W-[AVCTS]-[LI]-X-(7,13)-P-X-S-[TS]-[RQCS]-[TS]-[VALG]-X(4)-[ET]-[VLT]-X(0,1)-[ASG]-[FY]-X-[LF] for group III; L-X-[EDK]-[GR]-X(2)-[CYF]-GEEL-[LIF]-X-W-X-[LFQ0]-[ED]-X(1,2)-[SPA]-[SAL]-X(2,20)-[DGV]-X(11,16)-R-X-[VI]-X-C-X-[TS]-NV-[ED]-[AS]-[FL]-X-[LI] for group IVa, and [LI]-X(2)-G-X-[FY]-X-GDELLSWCLR-[RKQ]-[PS]-F-X(2)-R-[LR]-P-X-S-[STA]-X-[TGS]-X(4)-[ED]-X(2)-[EQ]-[AVG]-[FY]-X-L for Group IVb (Fig. 2). The group-wise PBC and hinge motif analysis revealed that the conservativeness of CNGC sequences was obviously different among the groups. Those of the Groups II and IVb were most conserved while those of the Group I least (Fig. 2). Notably, CNGCs of the Group IVa possessed more amino acids with an average of 52 than those of all other groups, which contained averagely 42 (Fig. 2). The difference existed in the middle portion between PBC and hinge region, which were averagely 26 aa in Group IVa while only 15–17 aa in other groups (Supplementary Fig. S2). To further show the conservativeness and relative frequency of amino acid in each position of the PBC and hinge region of CNGCs belonging to each group, we performed group-wise sequence logo motif analysis. The deduced sequence logos displayed that CNGCs of Groups I, II and III were conserved among each other with differences mainly in the positions 2, 3, 17, 19, 21, 22, 26 and 32. While CNGCs of Groups IVa and IVb were much less conserved with those of the other three groups. CNGCs of group IVa contained 10 extra amino acids in the middle in comparison with the others. Besides, there was also some difference in other positions. Significantly, CNGCs of Group IVb variated in 22 positions of a total of 42 amino acids, including positions 2, 3, 5, 7, 9, 13, 15, 17–23, 29, 31–33, 35–37 and 41, compared with those of other groups (Fig. 3). These data reveal that CNGCs of Groups IVa and IVb are apparently discrete from those of other groups.

3.3. Gene structural analysis reveals the uniqueness of Group IVa CNGC genes

To understand the possible gene structural relationship among CNGC orthologues and paralogues, the exon/intron structure of individual CNGC genes identified from 15 flowering plant species were analysed using GSDS software.[58] Intron numbers of these genes varied from 0 to 15, with wide divergence. Notably, the Group IVa CNGCs had distinct gene structures to those of all other groups, with more introns at different phases and lengths. The number of introns in Group IVa CNGCs ranged from 9 (*ZmCNGC8*, *BrCNGC22* and *MdCNGC38*) to 15 (*MdCNGC39*) and dominated with 11 (57%) or 10 (24%), which were obviously different from that in the other groups. Group I CNGCs contained introns ranged from 3 (*ZmCNGC1*) to 12 (*LuCNGC1* and *LuCNGC2*) and dominated with 6 (49%) or 7 (22%); Group II CNGCs carried introns ranged from 0 (6 genes) to 13 (*MdCNGC24*) with the majority being 6 (60%); Group III CNGC genes had introns of 1 (*PtCNGC10*) to 11 (*MdCNGC30*) with mostly being 6 (49%) or 5 (32%), while Group IVb CNGC genes constituted introns of 3 (*ZmCNGC11*) to 9 (*MdCNGC41*) mainly of 6 (51%) or 7 (30%) (Supplementary Fig. S3 and Tables S2 and S3). To better understand the phylogeny of gene structure of plant CNGCs, we also analysed the gene structure for CNGC genes from a vascular non-flowering plant (lycophyte) *S. moellendorffii*

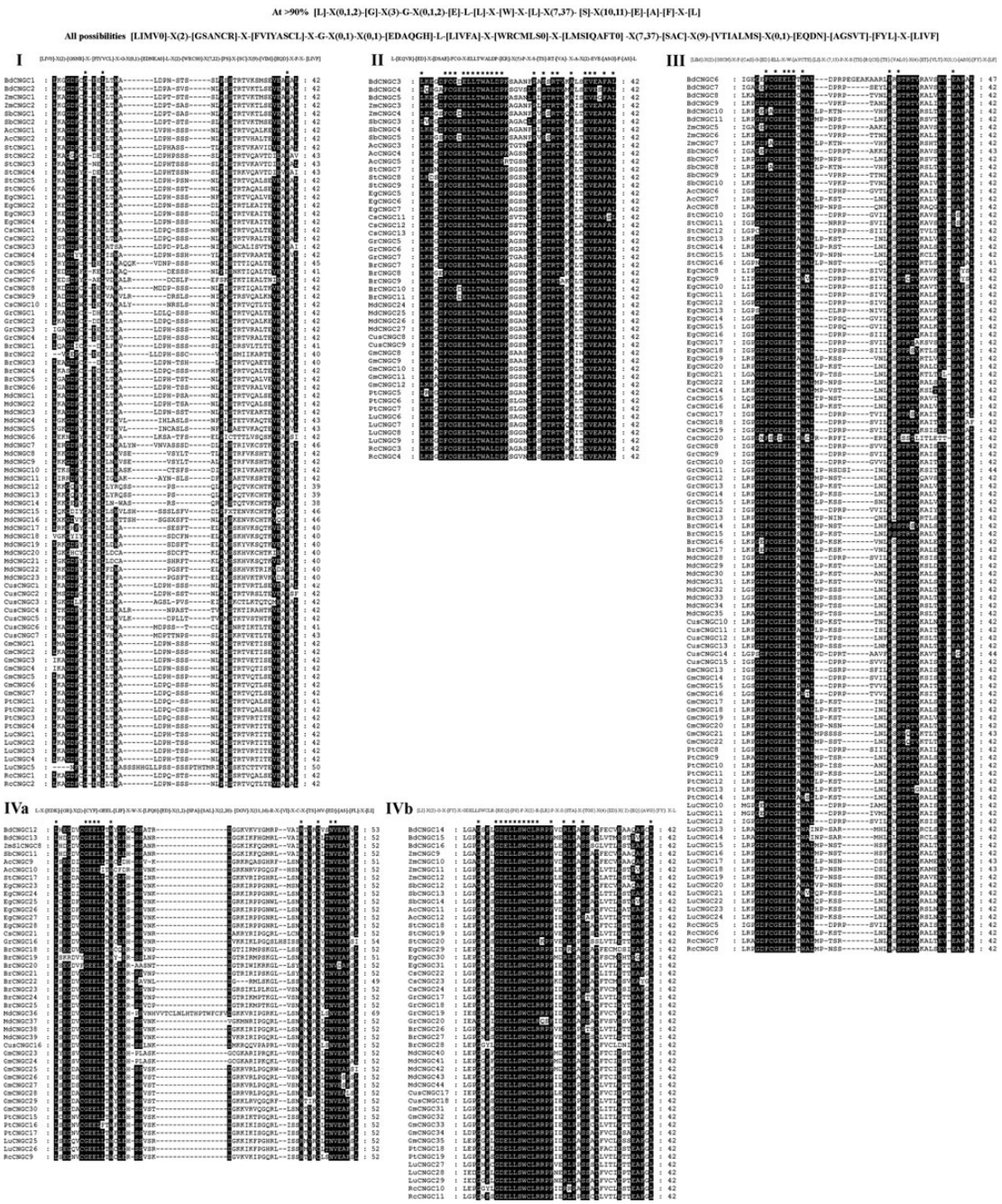

Figure 2. The group-wise alignment of CNGC-specific motif spanning the PBC and hinge region within CNB domain of plant CNGC proteins identified in this study. The CNGC-specific motifs both at >90% conservation and for all possibilities for all identified plant CNGCs are shown at top of the alignments. The motifs for each group are also shown at top of the alignments for each group. The square bracket '[ ]' indicates the amino acids allowed in this position of motif; A 'X' represents any amino acid, while the round bracket '()' denotes the number of amino acids. The names of groups and CNGC genes within groups are indicated to the left of the alignments. The length of the PBC and hinge region fragment is indicated to the right. Residues conserved at >90% were highlighted in black, while those were invariable were marked with asterisks above the alignments.

and a non-vascular plant (moss) *P. patens*.[12,13] Remarkably, the moss only contained Group IVa and IVb CNGCs and the lycophyte carried Groups I and II besides Groups IVa and IVb CNGCs (Refer to Section 3.4. for detail, Fig. 4). Intriguingly, one moss Group IVa CNGC genes, *PpCNGCg*, consisted of 12 introns, and all two lycophyte Group IVa CNGC genes (TRD8QYM9 SELML, TRD8RB40 SELML) contained 12 and 11 introns, respectively, as observed for other Group IVa genes in 15 flowering plant species (Supplementary Fig. S3). However, the remaining genes of both moss and *Selaginella* had only 4 to 8 introns (Supplementary Table S3). Moreover, introns of Group IVa CNGC genes had distinct phase profile to other group genes. Except *MdCNGC36*, all Group IVa genes possessed all of three intron phases; 2, 1 and 0 while introns of the majority of other group genes had no phase 1. The intron of the majority Group IVa CNGC genes had the

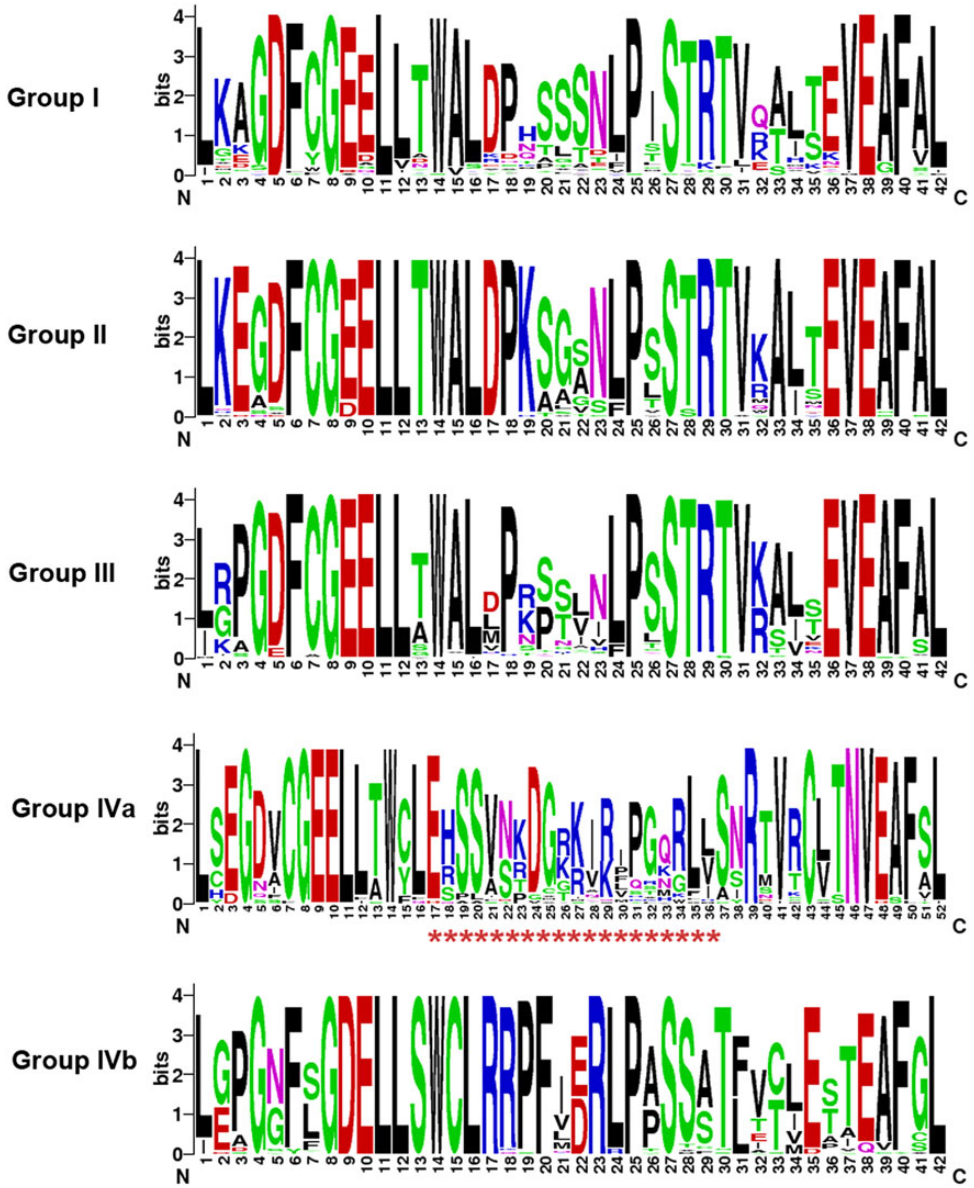

Figure 3. The group-wise sequence logos of the CNGC-specific motifs spanning the PBC and hinge region within CNB domain of plant CNGC proteins. The bit score for each position in the sequence and the group names are indicated to the left. The red asterisks below the logo for Group IVa indicate the distinct sequence uniquely existing in this group.

unique phase pattern 0-0-0-0-0-0-2-2-0-1-2 (52%) for 11-intron genes and 0-0-0-0-0-0-2-2-0-1 (14%) for 10-intron genes, which were mainly present in *B. rapa*. The only exception is those in *M. domestica*. Additionally, although CNGC genes of Group IVb and those of the other groups contain similar number of introns, their phase patterns differed. For 6-intron genes, those from Group IVb especially from eudicots exhibited a phase pattern of 0-0-2-0-0-0 (34%) while those from the Groups I, II and III had a 2-0-0-2-0-0 phase pattern (30, 56 and 36%, respectively). While for 7-intron genes, those from Group IVb showed a 2-0-0-2-0-0-0 phase pattern (19%) while those from Group I displayed a 2-0-0-2-0-0-2 phase pattern (9%) (Supplementary Table S3). In addition, most of Group IVa genes had larger gene size probably due to containing more introns in comparison with genes of other groups within the same species (Supplementary Fig. S3 and Table S3). These data clearly showed that the CNGC genes of Groups IVa and IVb especially Group IVa are distinct in gene structure to those of other groups.

3.4. The phylogeny of plant CNGC proteins

To elucidate the evolutionary history of the plant CNGC proteins, we comprehensively analysed the phylogeny of all identified 412 CNGCs in plant to date. We constructed maximum likelihood (ML) phylogenetic tree based on the alignment of the CNGC proteins, which included 333 from 15 flowering plant species identified in this study, 20 in Arabidopsis,[19] 18 in tomato,[48] 28 in rice, 8 in *P. patens* and 5 in *S. moellendorffii*.[16,17] Based on the phylogenetic tree, plant CNGC proteins clustered into five groups, named as Groups I, II, III, IVa and IVb with significant bootstrap values (Fig. 4) which is in accordance with what was reported for Arabidopsis CNGCs.[19] CNGCs of the

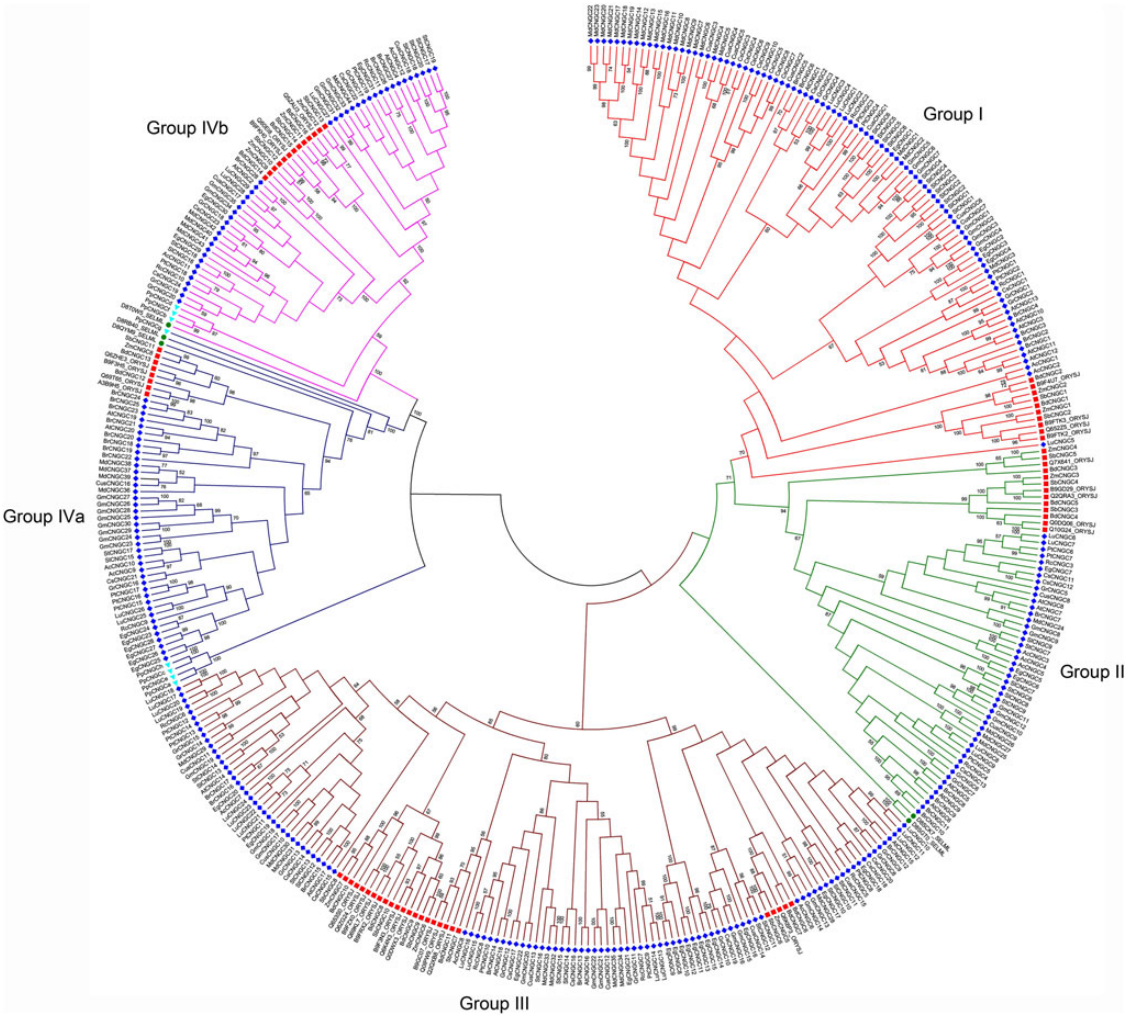

Figure 4. Phylogenetic tree of the plant CNGC families. The tree was constructed for 412 CNGC proteins including 333 from 15 plant species identified in this study, 20 in Arabidopsis,[19] 18 in tomato,[48] 28 in rice, 5 in *S. moellendorffii* and 8 in *P. patens*.[16,17] The tree was created using ClustalX program by maximum likelihood (ML) method with bootstrap of 1000 in MEGA5. Groups were indicated. The plant lineages are shown in different shapes and colours; moss in aqua triangle, lycophyte in green circle, monocots in red square and eudicots in blue diamond.

18 flowering higher plant species including both monocots and eudicots separated into five groups. However, markedly, all 8 CNGCs from the non-vascular land plant *P. patens* clustered only with Groups IVa and IVb. Among the five CNGCs of lower vascular (non-flowering) plant *S. moellendorffii*, two and one gathered in the Group IVa and IVb, respectively, while the remaining two formed a sister clade with Groups I and II CNGCs of flowering plant species (Fig. 4). This result suggests that during evolution Group IV CNGC genes emerged the earliest among all CNGCs in green plants; Group I and II CNGCs appeared later in vascular plants and Group III CNGCs are likely the most recently evolved in flowering plants.

Additionally, the groups were remarkably of unequal size. Group III is the largest one with 132 genes, including 106 from eudicots and 26 from monocots. Group I was the second largest group. It was composed of 102 genes including 92 from eudicots and 10 from monocots. In contrast, Groups II, IVa and IVb were only constituted of 63, 56 and 59 genes, respectively, which was significantly less than Groups III and I (Fig. 4). These data demonstrate that plant CNGC gene family expansion event occurred unequally in a group-dependent manner during evolution.

3.5. Silencing of the unique tomato Group IVa *CNGC* gene *SlCNGC15* reduces drought resistance but does not affect resistance to a variety of pathogens in tomato

As described above, the CNGC genes of Groups IVa are ancient and distinct in gene structure as well as CNGC motif in PBC and hinge region within CNB domain to those of other groups. To understand whether genes of this group are also functionally distinguished from those of the other groups, we analysed the function of the unique Group IVa gene in tomato, *SlCNGC15*, in biotic and abiotic stress resistance using VIGS technique and compared it with that of Group IVb *SlCNGC* genes, which have been known to regulate a wide range of resistance in tomato including negatively regulating resistance to fungal pathogens *Pythium aphanidermatum* and *S. sclerotiorum* and positively regulating resistance to viral pathogen *Tobacco rattle virus*, flg22-triggered PTI (PAMP-triggered immunity) and Pep1-triggered DPI (DAMP-triggered immunity).[48] A 331 bp fragment of the *SlCNGC15* gene, which is specific in this gene was amplified through RT-PCR, and cloned into the TRV silencing vector pYL156 for silencing analyses, while non-silenced eGFP fragment-inserted recombinant pYL156 vector was used as a negative control.[56] RT-PCR

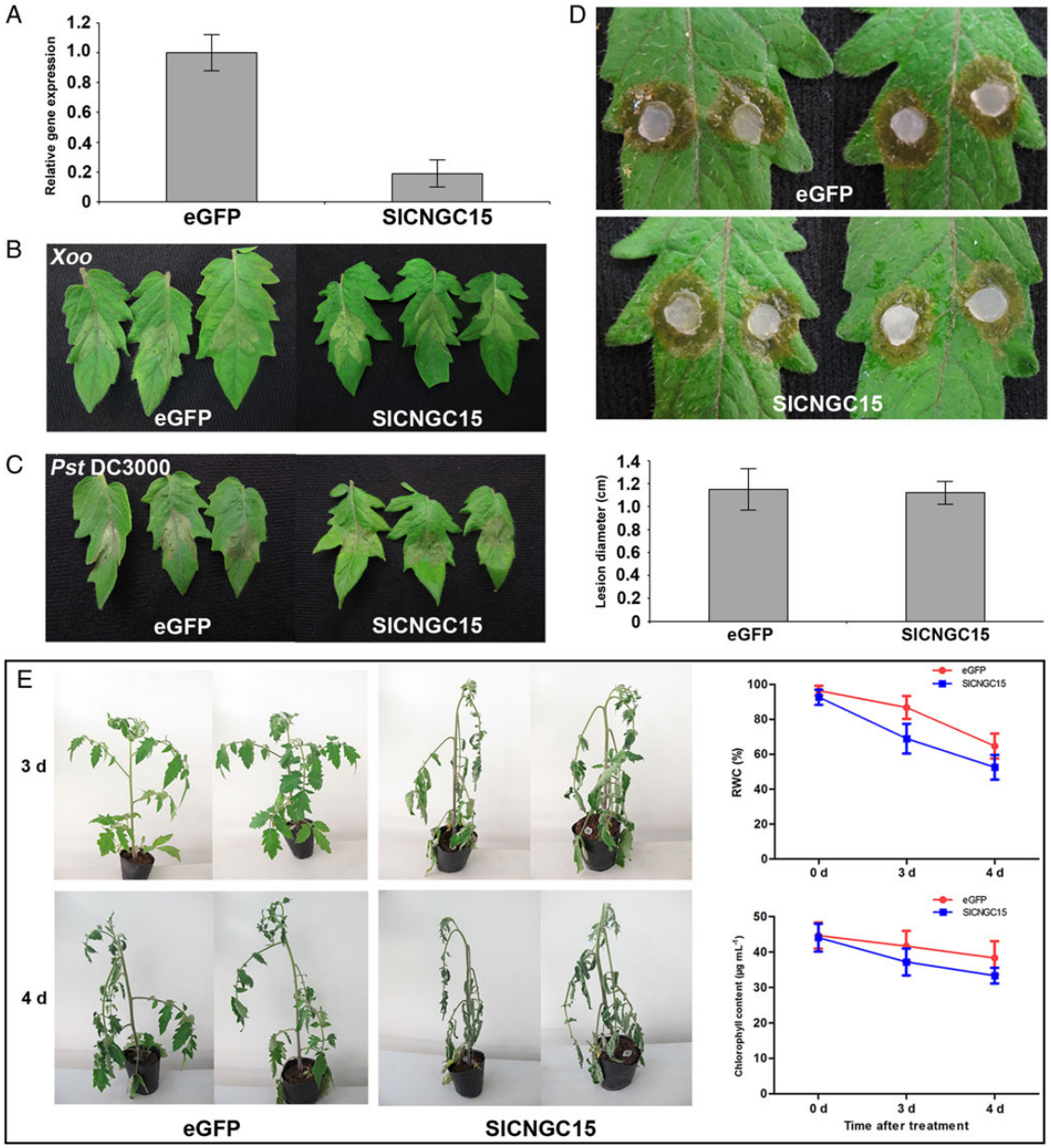

Figure 5. Functions of the Group IVa gene *SlCNGC15* in disease resistance and abiotic stress responses revealed by VIGS analyses. (A) Silencing efficiency analysis. Plants infiltrated with *Agrobacterium* suspensions carrying an eGFP-control vector served as control plants. Accumulation level of *SlCNGC15* transcript in VIGS-treated plants and the eGFP-control plants was detected by qRT-PCR analyses. (B) *Xoo*-induced hypersensitive response. Photographs were taken at 14 hpi. (C) *Pst* DC3000 caused necrosis symptoms. Photographs were taken at 48 hpi. (D) *Sclerotinia sclerotiorum* caused necrosis symptoms. Photographs were taken at 30 hpi. Lesion diameter was measured and statistically analysed for all plants. Significant difference between lesion diameter of the silenced plants and that of the eGFP-control plants is indicated as small letters ($P < 0.05$, DMRT). (E) Drought stress response. (Left) Phenotypes of the *SlCNGC15*-silenced and eGFP-control plants at 3 and 4 days after initiation of drought assay by withholding watering. (Right) Comparisons of leaf RWC and chlorophyll content of the *SlCNGC15*-silcenced and eGFP-control plants at 0, 3 and 4 days after drought treatment. Ten plants were used for each treatment. The experiments were conducted three times independently. Values represent mean ± SD ($n = 3$).

results showed that transcripts of *SlCNGC15* accumulated only 19% of those of the eGFP-control (Fig. 5A), indicating that *SlCNGC15* was efficiently silenced.

To probe function of *SlCNGC15* in disease resistance, the silenced tomato plants were inoculated with a set of different types of pathogens as biotic stresses. We inspected various non-host and host pathogens, including bacterial pathogens *P. syringae* pv. *tomato* DC3000 (*Pst* DC3000) and *X. oryzae* pv. *oryzae* (*Xoo*) and a fungal pathogen *S. sclerotiorum*. Resistance to both non-host pathogen *Xoo* and host pathogen *Pst* DC3000 in the *SlCNGC15*-silenced tomato plants was similar to that in the eGFP-control plants, as both plants developed clear hypersensitive response and necrosis symptoms with similar severity in infiltrated areas at 14 and 48 hpi, respectively (Fig. 5B and C). Resistance to host pathogen *S. sclerotiorum* in the *SlCNGC15*-silenced tomato plants was also similar to that in the eGFP-control plants. Necrotic symptoms of the leaves of the

SlCNGC15-silenced plants and the eGFP-control plants were similar (Fig. 5D). The size of lesions in the *SlCNGC15*-silenced plants was 1.12 cm in diameter on average, while that in the eGFP-control plants was 1.15 cm in diameter at 30 hpi (Fig. 5D). This is in contrast with the function of Group IVb genes *SlCNGC16*, *SlCNGC17* and *SlCNGC18*, which negatively regulate tomato resistance to this necrotrophic pathogen.[48]

The observation that *SlCNGC15* does not play a role in resistance to a variety of pathogens prompts us to examine whether this gene is involved in abiotic stress tolerance. To this aim, role of this gene in drought and salt tolerance was investigated. As shown in Fig. 5E, the leaves of the *SlCNGC15*-silenced plants showed obvious desiccation symptoms at 3 days after stopping watering, including leaf rolling and wilting, while the eGFP-control plants did not show any obvious desiccation symptom. At 4 days after drought treatment, the *SlCNGC15*-silenced plants only exhibited slight wilting symptoms in upper leaves, while the whole eGFP-control plants had completely wilted. To verify this stress tolerance phenotype in the silenced plants, we further comparatively measured the RWC and chlorophyll content of the *SlCNGC15*-silenced and the eGFP-control plants under drought stress conditions. The RWC in the eGFP-control plants decreased to ~86 and 64% at 3 and 4 days post drought treatment, respectively. However, that in the *SlCNGC15*-silenced plants significantly dropped to ~68 and 52% at 3 and 4 days post drought treatment, respectively (Fig. 5E), demonstrating that the *SlCNGC15*-silenced plants lost more water than the eGFP-control plants. Similarly, higher reduction of chlorophyll was observed in the *SlCNGC15*-silenced plants than in the eGFP-control plants after drought treatment (Fig. 5E). Taken together, these results indicate that the Group IVa gene *SlCNGC15* plays a role in drought tolerance in tomato. Similar analyses were conducted to probe the role of *SlCNGC15* in salinity tolerance. The *SlCNGC15*-silenced and eGFP-control plants displayed similar severity of symptoms after salinity treatment (Supplementary Fig. S4), indicating that *SlCNGC15* might be not required for salinity tolerance in tomato.

3.6. Silencing analyses of other group *SlCNGC* genes do not reveal group-specific function in disease resistance and abiotic stress response in tomato

The above finding that Group IVa and Group IVb *SlCNGC* genes play different roles in biotic and abiotic stress responses prompts us to investigate whether CNGCs function in a group-dependent manner. To address this hypothesis, two members per group, *SlCNGC1* and *SlCNGC6* of Group I, *SlCNGC7* and *SlCNGC8* of Group II and *SlCNGC11* and *SlCNGC14* of Group III, were selected for comparative functional analyses by VIGS. RT-PCR results showed that transcripts of the six SlCNGC genes accumulated only lower than

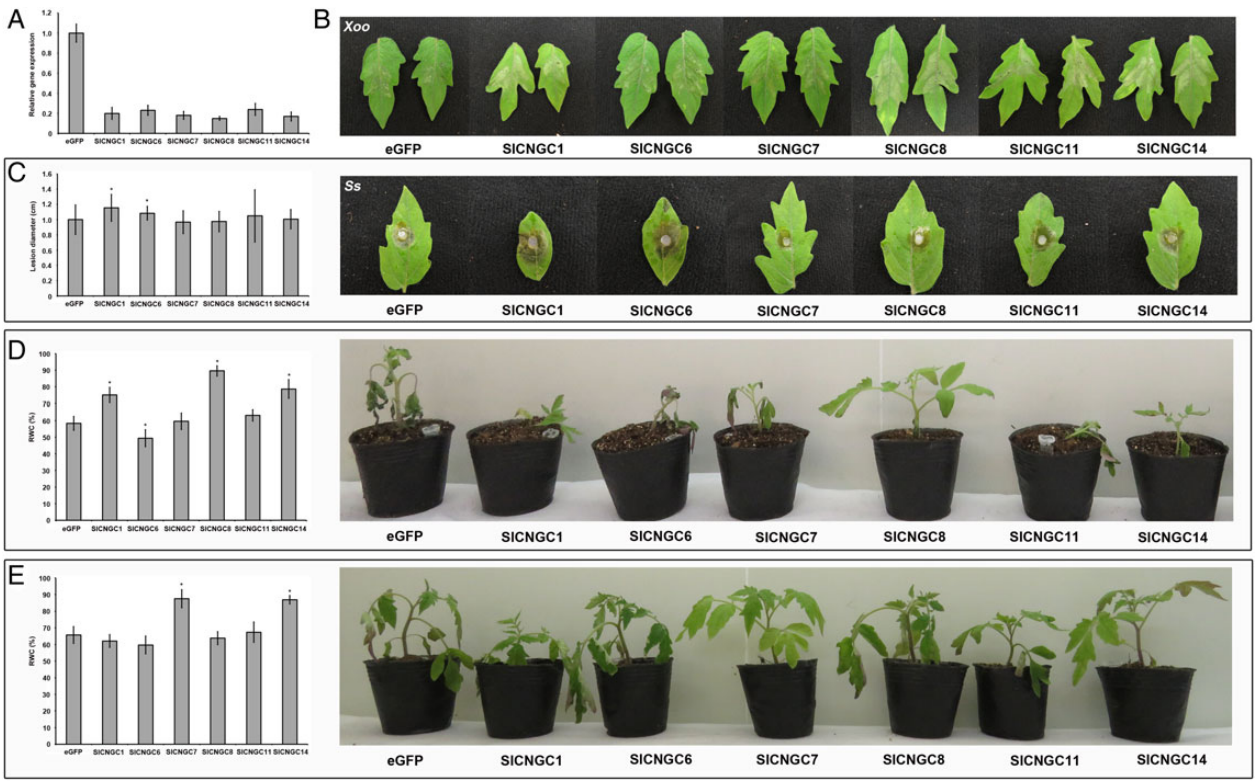

Figure 6. Functions of six Group I–III *SlCNGC* genes in disease resistance and abiotic stress responses revealed by VIGS analyses. (A) Silencing efficiency analysis. Plants infiltrated with *Agrobacterium* suspensions carrying an eGFP-control vector served as control plants. Accumulation level of *SlCNGC* transcript in VIGS-treated plants and the eGFP-control plants was detected by qRT-PCR analyses. (B) *Xoo*-induced hypersensitive response. Photographs were taken at 14 hpi. (C) *Sclerotinia sclerotiorum* caused necrosis symptoms. Photographs were taken at 24 hpi. Lesion diameter was measured and statistically analysed for all plants. Significant difference between lesion diameter of the silenced plants and that of the eGFP-control plants is indicated as small letters ($P < 0.05$, DMRT). (D) Salt stress response. (Right) Phenotypes of the *SlCNGC*-silenced and eGFP-control plants at 3 days after salt treatment. (Left) Comparison of leaf RWC of the *SlCNGC*-silenced and eGFP-control plants at 3 days after salt treatment. (E) Drought stress response. (Right) Phenotypes of the *SlCNGC*-silenced and eGFP-control plants at 8 days after initiation of drought assay by withholding watering. (Left) Comparison of leaf RWC of the *SlCNGC*-silcenced and eGFP-control plants at 8 days after drought treatment. Ten plants were used for each treatment. The experiments were conducted three times independently. Values represent mean ± SD ($n = 3$).

25% of those of eGFP-control (Fig. 6A), indicating that the SlCNGC genes were efficiently silenced. *SlCNGC1*- and *SlCNGC11*-silenced plants displayed retarded growth, while plants in which the other four SlCNGC genes were silenced showed normal growth as the eGFP-control plants (Supplementary Fig. S5).

To compare function of these *SlCNGC* genes in disease resistance, the silenced tomato plants were inoculated with bacterial pathogen *Xoo* and fungal pathogen *S. sclerotiorum*. *SlCNGC1*-, *SlCNGC8*-, *SlCNGC11*- and *SlCNGC14*-silenced plants developed much stronger hypersensitive necrosis at 14 hpi compared with the eGFP-control plants, while *SlCNGC6*- and *SlCNGC7*-silenced plants formed similar hypersensitive necrosis to the eGFP-control plants (Fig. 6B). This result indicated that *SlCNGC1*, *SlCNGC8*, *SlCNGC11* and *SlCNGC14* may play a negative role in tomato non-host resistance to Xoo. *SlCNGC1*- and *SlCNGC6*-silenced plants displayed more severe necrotic symptoms with statistically significant larger lesions compared with the eGFP-control plants, while plants in which the other four SlCNGC genes were silenced did not show statistically significant difference in lesion size compared with the eGFP-control plants (Fig. 6C). This result demonstrated that *SlCNGC1* and *SlCNGC6* might play a positive role in tomato resistance to *S. sclerotiorum*. Together, these results suggest that function of *SlCNGC* genes in disease resistance is not correlated with the group that they are belonging to.

To compare function of these *SlCNGC* genes in abiotic stress responses, the silenced tomato plants were examined for their role in salt and drought tolerance. Regarding salt tolerance, compared with the eGFP-control plants, *SlCNGC6*-silenced plants exhibited more severe symptoms, *SlCNGC1*-, *SlCNGC8*- and *SlCNGC14*-silenced plants developed less severe symptoms, while *SlCNGC7*- and *SlCNGC11*-silenced plants showed similar symptoms. At 3 days after 0.4 M NaCl supply, *SlCNGC8*-silenced plants kept normal growth; *SlCNGC1*- and *SlCNGC14*-silenced plants only displayed slight wilt in hypocotyls or leaves, *SlCNGC7*- and *SlCNGC11*-silenced plants severely wilted, while *SlCNGC6*-silenced plants completely wilted with all leaves desiccated (Fig. 6D). Moreover, the RWC in *SlCNGC8*-, *SlCNGC1*- and *SlCNGC14*-silenced plant leaves (89.6, 75.1 and 78.7%, respectively) was high; that in *SlCNGC7*- and *SlCNGC11*-silenced plant leaves (59.6 and 63.1%) was medium and similar to control (58.3%), while that in *SlCNGC6*-silenced plant leaves (49.4) was obviously lower than in control (Fig. 6D). This RWC data correlated well with the severity of wilting symptoms in the silenced plants. These results demonstrated that *SlCNGC6*-silenced plants are more sensitive, while *SlCNGC1*-, *SlCNGC8*- and *SlCNGC14*-silenced plants are more tolerant to high concentration of salt stress, and indicated that *SlCNGC6* may play a positive role, while *SlCNGC1*, *SlCNGC8* and *SlCNGC14* may play a negative role in salt tolerance in tomato. As for the drought tolerance, at 8 days after drought treatment, the *SlCNGC7*- and *SlCNGC14*-silenced plants remained normal growth without showing obvious wilting symptoms, while *SlCNGC1*-, *SlCNGC6*-, *SlCNGC8*- and *SlCNGC11*-silenced plants and the eGFP-control plants had severely wilted. The RWC content showed similar trend to the severity of wilting symptoms (Fig. 6E), indicating that *SlCNGC7*- and *SlCNGC14*-silenced plants are more tolerant to drought stress and, thus, that these two SlCNGC genes may play a negative role in drought tolerance in tomato.

Collectively, these results imply that *SlCNGC* genes of the same group may play different role while those of different groups may function similarly in biotic and abiotic stress responses, and thus, *SlCNGC* genes do not function in a group-dependent manner.

4. Discussion

4.1. CNGC family in plant

Plant CNGCs are characterized by the presence of a C-terminal CNB domain as well as an N-terminal hexa-transmembrane (TM) or ITP domain. However, it is noteworthy that the existence of these domains is only a necessary but not a sufficient condition to judge a CNGC protein, since many ion transporters other than CNGCs contain these domains as well. For example, potassium EAG/ERG/KAT channels (Shaker type) contain both a CNB domain and a TM domain.[59,60] To address this issue, a plant CNGC-specific motif in the PBC and hinge region within CNB domain of CNGC proteins is proposed.[16,17] It has been claimed that this motif only exists in plant CNGCs rather not other ion transporters. However, it is validated only in some plant species such as Arabidopsis, rice, a moss *P. patens* and a lycophyte *S. moellendorffii*.[16,17] Whether this is generally correct in other plant species remains unproved. In this study, we identified 333 CNGCs from 15 flowering plant species. We aligned the PBC and hinge region of 412 CNGC proteins, including these 333 sequences and 18 from tomato[48] as well as 61 from Arabidopsis, rice, *S. moellendorffii* and *P. patens* that were previously identified by other labs,[16,17,19] and consequently generate a motif [LIMV0]-X(2)-[GSANCR]-X-[FVIYASCL]-X-G-X(0,1)-X(0,1)-[EDAQGH]-L-[LIVFA]-X-[WRCMLS0]-X-[LMSIQAFT0]-X(7,37)-[SAC]-X(9)-[VTIALMS]-X(0,1)-[EQDN]-[AGSVT]-[FYL]-X-[LIVF], which recognizes all identified 412 plant CNGC proteins so far. Comparison of our motif with the one suggested for plant CNGCs, [LI]-X(2)-[GS]-X-[VFIYS]-X-G-X(0,1)-[DE]-L-[LI]-X-[WN]-X(6,32)-[SA]-X(9)-[VTI]-[EN]-[AG]-F-X-[LI] by Zelman et al.[17] demonstrated that the suggested motif is not applicable to all identified plant CNGCs. Almost all the positions of the motif had more possibilities than suggested, including the position suggested as invariable F in the hinge region (Supplementary Fig. S1). The only exception is the two invariable amino acids G and L in the PBC (Supplementary Fig. S1). Our results provide a more applicable CNGC motif that facilitates identification of plant CNGC.

Another finding of this study for plant CNGCs is that the size of CNGC family in different plant species varied significantly and independent of genome size. We observed that generally CNGC family in monocots is smaller than in dicots. All three analysed monocots in this study contain <16 CNGC genes. However, 10 out of 12 dicots under analysis comprise at least 18 CNGC genes and half of them consist of over 28 CNGC genes (Fig. 1). This result indicates independent expansion of CNGC gene members through duplications in distinct dicots lineages since their divergence from monocots. Further, it is obvious that independent genome duplications in Fabaceae, Solanaceae, Brassicaceae and Populus have contributed the increase in dicot CNGC genes compared with monocots, as well as variation in gene copy number across different dicot species. In addition, we scrutinized the structure of all identified plant CNGC genes and found that a total of six CNGC genes do not contain any intron. All these genes belong to Group II. Four of them are from monocots; one is from *A. coerulea*, a species very close to monocots in evolution (Fig. 1), while the remaining one is from dicot species *B. rapa* (Supplementary Fig. S3). Moreover, all three monocots under this study contain 1 to 2 intronless gene(s), while only 2 out of 12 dicots under this study carry one such a gene. This observation suggests that gain of introns by monocot intronless CNGC genes may have led to gene recombination and subsequent gene family expansion in dicots after their separation from monocots.

4.2. Phylogeny and evolution of CNGC family in plant

One of the aims of this study is to address the evolution of CNGC family in plant. To this end, we first identified CNGCs in single cellular algal species. Our BLASP searches using Arabidopsis and tomato CNGC proteins retrieved 67 sequences from six alga species, including *Ostreococcus lucimarinus*, *Micromonas pusilla* RCC299 and CCMP1545, *Coccomyxa subellipsoidea* C-169, *Volvox carteri* and *Chlamydomonas reinhardtii*. Domain analysis revealed that 49 of them contained both CNB and TM/ITP domains (Supplementary Table S5). However, analysis of CNGC-specific motif in the PBC and hinge region of CNB domain demonstrated that none of these sequences carried the CNGC-specific motif (Supplementary Table S5). Three CNGC genes were reported in green alga *C. reinhardtii*.[24] However, none of these sequences has a plant-specific CNGC motif as well (Ref. 16 and Supplementary Table S5, this study). Collectively, all retrieved and reported putative CNGC sequences from six algal species do not fit the plant CNGC motif criterion, although they could meet the domain criterion to be identified as CNGC genes. Therefore, probably canonical CNGC does not exist in the single cell species algae. Nevertheless, we successfully identified 333 CNGC genes in 15 higher flowering plant species including 3 monocots and 12 eudicots (Fig. 1; Supplementary Table S2). Together with those identified previously in Arabidopsis,[19] tomato,[48] rice, *P. patens* and *S. moellendorffii*,[16,17] a total of 412 CNGCs from 20 plant species at various evolutional nodes provide a good platform to analyse the phylogeny and evolution of CNGC family in plant. The ML phylogenetic tree for the 412 CNGCs clearly shows that plant CNGCs cluster into five groups. Strikingly, although CNGC family of each flowering plant species that are under this analysis contain members of all five groups, the non-vascular land plant *P. patens* only carries members of Groups IVa and IVb, while the lower vascular non-flowering plant *S. moellendorffii* possesses CNGCs of Groups IVa and IVb and a group of its own sharing a node with Groups I and II of flowering plants (Fig. 4). In other words, *P. patens* lacks CNGC of Groups I, II and III while *S. moellendorffii* lacks CNGC of Groups III. This indicates that Group III CNGCs are only present in higher flowering plant species. Collectively, these data reveals that Group IV CNGC genes are the most ancient CNGCs in green plants, and Group I and II CNGCs seem to have followed during evolution of vascular plants, while Group III CNGCs are likely the most recently evolved and their emergence must have been essential during the appearance of flowering plants.

4.3. Group IVa CNGC genes are structurally distinct to those of other groups

Group IVa CNGC genes are highly conserved in all plant species including moss, lycophyte and higher flowering plant species. Although plant species bear different number of Group IVa CNGC genes, most if not all of them in 12 out of 15 flowering plant species under this study contain 11 introns with a pattern of 0-0-0-0-0-0-2-2-0-1-2 except the apple gene *MdCNGC36* whose intron pattern is 0-0-0-0-2-0-2-2-0-0-0 (Supplementary Fig. S3 and Table S3). The remaining three species, maize, cucumber and *B. rapa*, do not contain any 11-intron Group IVa CNGC gene. BLASTP search in genome sequence of maize and cucumber did not retrieve any full-length Group IVa CNGC gene. However, we found that two successive ORFs coded the N- and C-truncated parts of a putative full-length CNGC protein, respectively. Therefore, we considered the fused sequence of the two truncated ones as a full-length CNGC gene (Supplementary Table S2). Whether these two sequences, *ZmCNGC8* and *CusCNGC19*, are indeed full-length Group IVa CNGC genes awaits further experimental confirmation. *Brassica rapa* is unique in that although it carries the highest number (8) of Group IVa CNGC genes, none of them contains 11 introns; rather they possess 9–10 introns (Supplementary Fig. S3 and Table S3). The reason to cause this remains unclear.

Group IVa CNGCs differs from those of other groups in many aspects. First, Group IVa CNGCs have larger size of protein compared with other group CNGCs. The average size of full length is 50–70 aa larger in Group IVa CNGCs, compared with other group CNGCs (Supplementary Table S6). It is similar for the PBC and hinge region fragment within CNB domain. It is generally 52 aa in those of Group IVa, which is 10 aa larger than those of other groups (Fig. 2; Supplementary Figs S1 and S2). Second, the CNGC-specific motif of Group IVa CNGCs is different from that of other groups at many positions (Fig. 2; Supplementary Fig. S2). Additionally, the exon/intron gene structure and gene length of group IVa CNGCs are distinct to those of other groups. Group IVa CNGC genes generally contain 10–11 introns with intron pattern of 0-0-0-0-0-0-2-2-0-1-2 for 11-intron genes and 0-0-0-0-0-0-2-2-0-1 for 10-intron genes, while the other group CNGC genes mainly comprise 5–7 introns with no Phase 1 intron (Supplementary Fig. S3). The Group IVa CNGC genes generally have a larger size than those of other groups probably due to carrying more introns. As far as we know, this is the first time to systemically compare the structural characteristics between CNGC genes of Group IVa and those of other groups.

4.4. Function of tomato CNGCs in disease resistance and abiotic stress responses

Considering that Group IVa CNGCs are distinct to those of all other groups in terms of structure as described above, we are interested to make clear whether Group IVa CNGCs are also distinguishable from those of other groups in functions in disease resistance and abiotic stress responses. VIGS analyses indicate that the unique Group IVa CNGC gene in tomato, *SlCNGC15*, plays an important role in drought tolerance but does not affect resistance to a variety of pathogens including bacterial non-host pathogen *Xoo*, bacterial host pathogen *Pst* DC3000 and fungal pathogen *S. sclerotiorum* (Fig. 5). This is in contrast with the three Group IVb CNGC genes in tomato, *SlCNGC16/17/18*, which function in resistance to the *S. sclerotiorum*.[48] However, VIGS analyses for six CNGC genes belonging to three other groups demonstrate that function of *SlCNGC* genes is not correlated with the group that they are belonging to. *SlCNGC* genes of the same group may play different role while those of different groups may function similarly in biotic and abiotic stress responses. For instance, *SlCNGC8* of Group II and *SlCNGC14* of Group III may both function in resistance to *Xoo* and tolerance to salt stress, while *SlCNGC7*, which belongs to the same group as *SlCNGC8*, plays no role in these biological processes (Fig. 6). Some SlCNGC genes of other groups such as *SlCNGC7* of Group II and *SlCNGC14* of Group III may be also involved in drought tolerance as *SlCNGC15* of Group IVa (Figs 5 and 6). These results reveal that *SlCNGC* genes do not function in a group-dependent manner.

Our VIGS analyses indicate that the Group IVa gene *SlCNGC15* might be not required for salinity tolerance in tomato (Supplementary Fig. S4). However, *AtCNGC19* and *AtCNGC20*, two Group IVa Arabidopsis CNGC genes, were found to be involved in salinity response and could assist the plant to cope with toxic effects caused by salt stress, probably by contributing to a re-allocation of sodium within the plant.[46] The reason to cause this discrepancy is unclear. It may reflect the difference in function of CNGC homologues of the same group in different plant species. Alternatively, since plant CNGCs

may form heterotetramer as found for animal CNGCs and redundancy may exist among CNGC members, the phenotype of VIGS may be affected by other CNGC genes. Moreover, VIGS only results in knock-down rather not knock-out of *SlCNGC15* in tomato (Fig. 5A). Therefore, results from the VIGS analyses await further confirmation using mutants or stable high-efficient RNAi lines. In addition, AtCNGC20 can bind with CaM at its isoleucine glutamine (IQ) motif locating outside of CNB domain.[61] Whether SlCNGC15 and even other SlCNGCs contain similar function mechanism is worthy of further study.

5. Conclusion

A total of 333 CNGC genes from 15 plant species were identified using bioinformatics approaches based on the presence of a CNB domain and a transmembrane pore-forming domain as well as a plant CNGC-specific motif spanning the PBC and hinge region within CNB domain of CNGC proteins. Eight loci for *ZmCNGC1*, *ZmCNGC8*, *CsCNGC15* and *CusCNGC16* that were misannotated at Phytozome database were corrected. We also modified the CNGC-specific stringent motif as [LIMV0]-X(2)-[GSANCR]-X-[FVIYASCL]-X-G-X(0,1)-X(0,1)-[EDAQGH]-L-[LIVFA]-X-[WRCMLS0]-X-[LMSIQAFT0]-X(7,37)-[SAC]-X(9)-[VTIALMS]-X(0,1)-[EQDN]-[AGSVT]-[FYL]-X-[LIVF] to be applicable to all 412 CNGCs from 20 plant species that were identified in this study and by other labs so far. Phylogenetic analyses revealed that Group IV CNGCs were the first to arise from the last common ancestors of all green plants, while Groups I and II seem to have evolved later along with appearance of vascular plants with subsequent divergence of Group III CNGCs in flowering plants. Expansion and diversification of CNGC genes into various groups most probably occurred independently in distinct plant lineages. Significantly, we discover that CNGCs of Group IVa are distinct to those of other groups in many aspects of structure, including protein size, length and CNGC-specific motif in the PBC and hinge region of CNB domain, exon/intron structure especially intron number and phase pattern as well as gene size. Members of different groups of tomato CNGC genes play distinguishable roles in disease resistance and abiotic stress responses. Our results provide insights into the phylogeny and evolution as well as function of plant CNGCs.

Conflict of interest statement

The authors have declared that no competing interest exists.

Supplementary data

Supplementary data are available at www.dnaresearch.oxfordjournals.org

Funding

This work was financially supported by grants from the Genetically Modified Organisms Breeding Major Projects (no. 2014ZX0800905B), the Special Fund for Agro-scientific Research in the Public Interest (no. 201103016), the SRFDP (no. 20110101110092) and the Program for New Century 151 Talents of Zhejiang Province. Funding to pay the Open Access publication charges for this article was provided by the grant from the Genetically Modified Organisms Breeding Major Projects (no. 2014ZX0800905B).

References

1. Kudla, J., Batistic, O. and Hashimoto, K. 2010, Calcium signals: the lead currency of plant information processing, *Plant Cell*, **22**, 541–63.
2. Jammes, F., Hu, H.C., Villiers, F., Bouten, R. and Kwak, J.M. 2011, Calcium-permeable channels in plant cells, *FEBS J.*, **278**, 4262–76.
3. Reddy, A.S., Ali, G.S., Celesnik, H. and Day, I.S. 2011, Coping with stresses: roles of calcium- and calcium/calmodulin-regulated gene expression, *Plant Cell*, **23**, 2010–32.
4. Munemasa, S., Oda, K., Watanabe-Sugimoto, M., Nakamura, Y., Shimoishi, Y. and Murata, Y. 2007, The coronatine-insensitive 1 mutation reveals the hormonal signaling interaction between abscisic acid and methyl jasmonate in Arabidopsis guard cells. Specific impairment of ion channel activation and second messenger production, *Plant Physiol.*, **143**, 1398–407.
5. Frietsch, S., Wang, Y.F., Sladek, C., et al. 2007, A cyclic nucleotide-gated channel is essential for polarized tip growth of pollen, *Proc. Natl Acad. Sci. USA*, **104**, 14531–36.
6. Harada, A., Sakai, T. and Okada, K. 2003, Phot1 and phot2 mediate blue light-induced transient increases in cytosolic Ca^{2+} differently in Arabidopsis leaves, *Proc. Natl Acad. Sci. USA*, **100**, 8583–8.
7. Tracy, F.E., Gilliham, M., Dodd, A.N., Webb, A.A. and Tester, M. 2008, NaCl-induced changes in cytosolic free Ca^{2+} in *Arabidopsis thaliana* are heterogeneous and modified by external ionic composition, *Plant Cell Environ.*, **31**, 1063–73.
8. Qi, Z., Verma, R., Gehring, C., et al. 2010, Ca^{2+} signaling by plant *Arabidopsis thaliana* Pep peptides depends on AtPepR1, a receptor with guanylyl cyclase activity, and cGMP-activated Ca2^{+} channels, *Proc. Natl Acad. Sci. USA*, **107**, 21193–98.
9. Talke, I.N., Blaudez, D., Maathuis, F.J. and Sanders, D. 2003, CNGCs: prime targets of plant cyclic nucleotide signalling? *Trends Plant Sci.*, **8**, 286–93.
10. Kaupp, U.B. and Seifert, R. 2002, Cyclic nucleotide-gated ion channels, *Physiol. Rev.*, **82**, 769–824.
11. Chin, K., Moeder, W. and Yoshioka, K. 2009, Biological roles of cyclic-nucleotide-gated ion channels in plants: what we know and don't know about this 20 member ion channel family, *Botany*, **87**, 668–77.
12. Ma, W., Qi, Z., Smigel, A., Walker, R.K., Verma, R. and Berkowitz, G.A. 2009, Ca^{2+}, cAMP, and transduction of non-self perception during plant immune responses, *Proc. Natl Acad. Sci. USA*, **106**, 20995–1000.
13. Wang, Y.F., Munemasa, S., Nishimura, N., et al. 2013, Identification of cyclic GMP-activated nonselective Ca^{2+}-permeable cation channels and associated CNGC5 and CNGC6 genes in Arabidopsis guard cells, *Plant Physiol.*, **163**, 578–90.
14. Gao, Q.F., Fei, C.F., Dong, J.Y., Gu, L.L. and Wang, Y.F. 2014, Arabidopsis CNGC18 is a Ca^{2+}-permeable channel, *Mol. Plant*, **7**, 739–43.
15. Zhou, L., Lan, W., Jiang, Y., Fang, W. and Luan, S. 2014, A calcium-dependent protein kinase interacts with and activates a calcium channel to regulate pollen tube growth, *Mol. Plant*, **7**, 369–76.
16. Zelman, A.K., Dawe, A., Gehring, C. and Berkowitz, G.A. 2012, Evolutionary and structural perspectives of plant cyclic nucleotide-gated cation channels, *Front. Plant Sci.*, **3**, 95.
17. Zelman, A.K., Dawe, A. and Berkowitz, G.A. 2013, Identification of cyclic nucleotide gated channels using regular expressions. In: Gehring, C. (ed.), *Cyclic nucleotide signaling in plants*, vol. **1016**. Humana Press: New York, pp.207–24.
18. Schuurink, R.C., Shartzer, S.F., Fath, A. and Jones, R.L. 1998, Characterization of a calmodulin-binding transporter from the plasma membrane of barley aleurone, *Proc. Natl Acad. Sci. USA*, **95**, 1944–9.
19. Maser, P., Thomine, S., Schroeder, J.I., et al. 2001, Phylogenetic relationships within cation transporter families of Arabidopsis, *Plant Physiol.*, **126**, 1646–67.
20. Bridges, D., Fraser, M.E. and Moorhead, G.B. 2005, Cyclic nucleotide binding proteins in the *Arabidopsis thaliana* and *Oryza sativa* genomes, *BMC Bioinformatics*, **6**, 6.
21. Nawaz, Z., Kakar, K.U., Saand, M.A. and Shu, Q.Y. 2014, Cyclic nucleotide-gated ion channel gene family in rice, identification, characterization and experimental analysis of expression response to plant hormones, biotic and abiotic stresses, *BMC Genomics*, **15**, 853.

22. Ward, J.M., Maser, P. and Schroeder, J.I. 2009, Plant ion channels: gene families, physiology, and functional genomics analyses, *Annu. Rev. Physiol.*, **71**, 59–82.
23. Chen, J., Yin, H., Gu, J., et al. 2015, Genomic characterization, phylogenetic comparison and differential expression of the cyclic nucleotide-gated channels gene family in pear (*Pyrus bretchneideri* Rehd.), *Genomics*, **105**, 39–52.
24. Verret, F., Wheeler, G., Taylor, A.R., Farnham, G. and Brownlee, C. 2010, Calcium channels in photosynthetic eukaryotes: implications for evolution of calcium-based signalling, *New Phytol.*, **187**, 23–43.
25. Kaplan, B., Sherman, T. and Fromm, H. 2007, Cyclic nucleotide-gated channels in plants, *FEBS Lett.*, **581**, 2237–46.
26. Sherman, T. and Fromm, H. 2009, Physiological roles of cyclic nucleotide gated channels in plants. In: Mancuso, S. and Baluska, F. (eds.), *Signaling in plants*. Springer Press: Heidelberg, pp.91–106.
27. Abdel-Hamid, H., Chin, K., Moeder, W. and Yoshioka, K. 2011, High throughput chemical screening supports the involvement of Ca^{2+} in cyclic nucleotide-gated ion channel-mediated programmed cell death in Arabidopsis, *Plant Signal. Behav.*, **6**, 1817–9.
28. Ma, W. 2011, Roles of Ca^{2+} and cyclic nucleotide gated channel in plant innate immunity, *Plant Sci.*, **181**, 342–6.
29. Ma, W. and Berkowitz, G.A. 2011, Ca^{2+} conduction by plant cyclic nucleotide gated channels and associated signaling components in pathogen defense signal transduction cascades, *New Phytol.*, **190**, 566–72.
30. Moeder, W., Urquhart, W., Ung, H. and Yoshioka, K. 2011, The role of cyclic nucleotide-gated ion channels in plant immunity, *Mol. Plant*, **4**, 442–52.
31. Ma, W., Ali, R. and Berkowitz, G.A. 2006, Characterization of plant phenotypes associated with loss-of-function of *AtCNGC1*, a plant cyclic nucleotide gated cation channel, *Plant Physiol. Biochem.*, **44**, 494–505.
32. Gobert, A., Park, G., Amtmann, A., Sanders, D. and Maathuis, F.J. 2006, *Arabidopsis thaliana* cyclic nucleotide gated channel 3 forms a non-selective ion transporter involved in germination and cation transport, *J. Exp. Bot.*, **57**, 791–800.
33. Borsics, T., Webb, D., Andeme-Ondzighi, C., Staehelin, L.A. and Christopher, D.A. 2007, The cyclic nucleotide-gated calmodulin-binding channel *AtCNGC10* localizes to the plasma membrane and influences numerous growth responses and starch accumulation in *Arabidopsis thaliana*, *Planta*, **225**, 563–73.
34. Tunc-Ozdemir, M., Tang, C., Ishka, M.R., et al. 2013, A cyclic nucleotide-gated channel (CNGC16) in pollen is critical for stress tolerance in pollen reproductive development, *Plant Physiol.*, **161**, 1010–20.
35. Yoshioka, K., Kachroo, P., Tsui, F., Sharma, S.B., Shah, J. and Klessig, D.F. 2001, Environmentally sensitive, SA-dependent defense responses in the cpr22 mutant of Arabidopsis, *Plant J.*, **26**, 447–59.
36. Yoshioka, K., Moeder, W., Kang, H.G., et al. 2006, The chimeric Arabidopsis CYCLIC NUCLEOTIDE-GATED ION CHANNEL11/12 activates multiple pathogen resistance responses, *Plant Cell*, **18**, 747–63.
37. Yu, I.C., Parker, J. and Bent, A.F. 1998, Gene-for-gene disease resistance without the hypersensitive response in Arabidopsis *dnd1* mutant, *Proc. Natl Acad. Sci. USA*, **95**, 7819–24.
38. Clough, S.J., Fengler, K.A., Yu, I.C., Lippok, B., Smith, R.K. Jr. and Bent, A. F. 2000, The Arabidopsis *dnd1* "defense, no death" gene encodes a mutated cyclic nucleotide-gated ion channel, *Proc. Natl Acad. Sci. USA*, **97**, 9323–8.
39. Balague, C., Lin, B., Alcon, C., et al. 2003, *HLM1*, an essential signaling component in the hypersensitive response, is a member of the cyclic nucleotide-gated channel ion channel family, *Plant Cell*, **15**, 365–79.
40. Jurkowski, G.I., Smith, R.K. Jr., Yu, I.C., et al. 2004, Arabidopsis *DND2*, a second cyclic nucleotide-gated ion channel gene for which mutation causes the "defense, no death" phenotype, *Mol. Plant Microb. Interact.*, **17**, 511–20.
41. Ali, R., Ma, W., Lemtiri-Chlieh, F., et al. 2007, Death don't have no mercy and neither does calcium: Arabidopsis cyclic nucleotide gated channel2 and innate immunity, *Plant Cell*, **19**, 1081–95.
42. Finka, A., Cuendet, A.F., Maathuis, F.J., Saidi, Y. and Goloubinoff, P. 2012, Plasma membrane cyclic nucleotide gated calcium channels control land plant thermal sensing and acquired thermotolerance, *Plant Cell*, **24**, 3333–48.
43. Chin, K., DeFalco, T.A., Moeder, W. and Yoshioka, K. 2013, The Arabidopsis cyclic nucleotide-gated ion channels AtCNGC2 and AtCNGC4 work in the same signaling pathway to regulate pathogen defense and floral transition, *Plant Physiol.*, **163**, 611–24.
44. Fortuna, A., Lee, J., Ung, H., Chin, K., Moeder, W. and Yoshioka, K. 2015, Crossroads of stress responses, development and flowering regulation- the multiple roles of Cyclic Nucleotide Gated Ion Channel 2, *Plant Signal. Behav.*, **10**, article no. e989758.
45. Maathuis, F.J. 2006, The role of monovalent cation transporters in plant responses to salinity, *J. Exp. Bot.*, **57**, 1137–47.
46. Kugler, A., Kohler, B., Palme, K., Wolff, P. and Dietrich, P. 2009, Salt-dependent regulation of a CNG channel subfamily in Arabidopsis, *BMC Plant Biol.*, **9**, 140.
47. Yuen, C.Y. and Christopher, D.A. 2010, The role of cyclic nucleotide-gated channels in cation nutrition and abiotic stress. In: Demidchik, V. and Maathuis, F. (eds.), *Ion channels and plant stress responses*, Springer Press: Heidelberg, pp.137–57.
48. Saand, M.A., Xu, Y.P., Li, W., Wang, J.P. and Cai, X.Z. 2015, Cyclic nucleotide gated channel gene family in tomato: genome-wide identification and functional analyses in disease resistance, *Front. Plant Sci.*, **6**, 303.
49. Nicholas, K.B., Nicholas, H. and Deerfield, D. 1997, GeneDoc: analysis and visualization of genetic variation, *Embnet.news*, **4**, 1–4.
50. Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S. 2011, MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods, *Mol. Biol. Evol.*, **28**, 2731–9.
51. Larkin, M.A., Blackshields, G., Brown, N.P., et al. 2007, Clustal W and Clustal X version 2.0, *Bioinformatics*, **23**, 2947–8.
52. Li, W., Xu, Y.P., Zhang, Z.X., et al. 2012, Identification of genes required for nonhost resistance to *Xanthomonas oryzae* pv. *oryzae* reveals novel signaling components, *PLoS ONE*, **7**, e42796.
53. Zhao, Y., Liu, W., Xu, Y.P., Cao, J.Y., Braam, J. and Cai, X.Z. 2013, Genome-wide identification and functional analyses of calmodulin genes in Solanaceous species, *BMC Plant Biol.*, **13**, 70.
54. Wang, C., Cai, X., Wang, X. and Zheng, Z. 2006, Optimisation of tobacco rattle virus-induced gene silencing in Arabidopsis, *Funct. Plant Biol.*, **33**, 347–55.
55. Cai, X.Z., Wang, C., Xu, Y., Xu, Q., Zheng, Z. and Zhou, X. 2007, Efficient gene silencing induction in tomato by a viral satellite DNA vector, *Virus Res.*, **125**, 169–75.
56. Cheng, W.S., Xu, Q.F., Li, F., Xu, Y.P. and Cai, X.Z. 2012, Establishment of a suitable control vector for Tobacco rattle virus-induced gene silencing analysis in *Nicotiana benthamiana*, *J. Zhejiang Uni. Agri. Life Sci.*, **38**, 10–20.
57. Zhu, M., Chen, G., Zhang, J., et al. 2014, The abiotic stress-responsive NAC-type transcription factor SlNAC4 regulates salt and drought tolerance and stress-related genes in tomato (*Solanum lycopersicum*), *Plant Cell Rep.*, **33**, 1851–63.
58. Guo, A.Y., Zhu, Q.H., Chen, X. and Luo, J.C. 2007, GSDS: a gene structure display server, *Yi Chuan*, **29**, 1023–6.
59. Cherel, I. 2004, Regulation of K^+ channel activities in plants: from physiological to molecular aspects, *J. Exp. Bot.*, **55**, 337–51.
60. Su, H., Golldack, D., Katsuhara, M., Zhao, C. and Bohnert, H.J. 2001, Expression and stress-dependent induction of potassium channel transcripts in the common ice plant, *Plant Physiol.*, **125**, 604–14.
61. Fischer, C., Kugler, A., Hoth, S. and Dietrich, P. 2013, An IQ domain mediates the interaction with calmodulin in a plant cyclic nucleotide-gated channel, *Plant Cell Physiol.*, **54**, 573–84.