Genome-wide Identification and Analysis of the *CNGC* Gene Family in maize

Lidong Hao[1], Xiuli Qiao[2]*

[1]College of Agriculture and Hydraulic Engineering, Suihua University, 18 Huanghe Road, Suihua, Heilongjiang Province, 152000, PR China

[2]College of Food and Pharmaceutical Engineering, Suihua University, 18 Huanghe Road, Suihua, Heilongjiang Province, 152000, PR China

*Corresponding author:

Xiuli Qiao

Email: shxynxhld@163.com

Abstract

As one of the non-selective cation channel gene family, the cyclic nucleotide-gated channel (CNGC) genes play vital roles in plant physiological processes which are related to signal pathways, plant development, and environment stresses. However, genome-wide identification and analysis of the *CNGC* gene family in maize has not yet been conducted. In this study, 12 *ZmCNGC* genes were identified in the maize genome, which were unevenly distributed on chromosomes 1, 2, 4, 5, 6, 7 and 8. They were classified into five major groups, including Group I, II, III, IVa and IVb. Phylogenetic analysis showed that Group IV CNGC genes emerged the earliest while Group I and II CNGCs appeared later. Prediction of cis-acting regulatory elements showed that 137 putative cis-elements which were related to hormones-response, abiotic stress and organ development. Furthermore, the co-expression network analysis of *ZmCNGC* genes may establish the importance of better understanding ZmCNGC transduction pathways in maize. Additionally, expression profiles of *ZmCNGC* genes were shown to express in a tissue-specific pattern. Our results provide valuable information to increase our understanding of the *CNGC* gene family in maize.

INTRODUCTION

In the process of organism evolution, the formation of complex nutrient absorption and transport system includes ion channels, ion pumps and carriers, and it has been previously shown that those systems respond to endogenous and abiotic stimuli (Saand et al. 2015b). The cyclic nucleotide-gated channel (CNGC) is a Ca^{2+}-permeable cation transport channel, which is suggested to have one of the fundamental mechanisms for organismal systems (Yuen and Christopher 2013; Nawaz et al. 2014). As a molecular switch, secondary messengers such as cyclic nucleotide monophosphates (cNMPs; 3',5'-cAMP and 3',5'-cGMP) and Ca^{2+}/calmodulin (CaM) can regulate CNGCs, those messengers are activated by directing binding of cyclic nucleotides as well as are inhibited by binding of CaM to the CaM binding domain (Saand et al. 2015b; Borsics et al. 2007; Defalco et al. 2016; Kaplan et al. 2007).

In plant CNGCs, they are composed of six transmembrane (TM) domains and one pore region between the fifth and sixth TM domains. Among plants, the CNBD is highly conserved region and carries a PBC and a hinge region (Saand et al. 2015b). The first plant CNGC was identified in *Hordeum vulgare* and named as *HvCBT1* (Schuurink et al. 1998). With the application of bioinformatics tools, identification of *CNGC* gene family in *Arabidopsis*, rice and other plants were carried out. So far, *CNGC* genes have been identified in many plants, with 20 for *Arabidopsis*, 16 for rice, 18 for tomato and 26 in *B. oleracea* (Bridges et al. 2005; Nawaz et al. 2014; Ward et al. 2009; Zelman et al. 2012; Saand et al. 2015a; Chen et al. 2015; Zelman et al. 2013; Guo et al. 2017; Kakar et al. 2017). The identified 20 *Arabidopsis* CNGCs were classified into five groups, including group I, II, III, IVa and IVb (Mäser et al. 2001).

As CNGCs are key components for plant development, many previous research studies have been conducted (Wang et al. 2013; Hua et al. 2003). Most *CNGC*s have been characterized by genetic methods and found to be related to plant physiological and molecular functions, including playing vital roles in multiple physiological processes which are involved in signal pathways, plant development, and environmental stresses. For example, *Arabidopsis CNGC7* and *CNGC8* are essential for male reproductive fertility (Tunc-Ozdemir et al. 2013a); *CNGC16* and *CNGC18* participate in the pollen development (Tunc-Ozdemir et al. 2013b; Frietsch et al. 2007); *AtCNGC2*

is involved in jasmonic acid induced apoplastic Ca(2+) influx in epidermal cells (Lu et al. 2015). *Arabidopsis CNGC2*, *CNGC6*, *CNGC19* and *CNGC20* are involved in abiotic stress (Kugler et al. 2009; Gao et al. 2012; Fortuna et al. 2015). Meanwhile, the structure of *Arabidopsis* CNGCs have six transmembrane domains with a pore domain, also have cyclic nucleotide-binding domain and CaM-binding domains in the C-terminus, and these domains have diverse functions (Talke et al. 2003; Chin et al. 2009; Hua et al. 2003; Köhler and Neuhaus 2000). For example, *AtCNGC2* play key roles in stress signaling pathways, including changes the cytosolic free Ca^{2+} in *Arabidopsis* and *CNGC4* is permeable to both K+ and Na+ and activated by both cGMP and cAMP (Balague 2003; Ali et al. 2007).

In recent years, efforts had been made in studying the CNGC gene family in plants, but as one of the most important food crops and source of industrial materials in the world, the maize *CNGC* gene family was rarely reported. In this study, with the benefit from genome-wide sequence information in maize and research information on *Arabidopsis* and rice CNGC families, we conducted genome-wide identification of CNGCs in maize through comprehensive bioinformatics analyses. Furthermore, comprehensive analyses were conducted including multiple alignments, gene structure, conserved motifs and gene duplication of ZmCNGCs, and prediction of cis-acting regulatory elements, expression profiles of *ZmCNGC* genes and a co-expression network between ZmCNGC and other maize genes. This is the first systematic study of CNGC genes in maize and will provide the basis for further research on the *ZmCNGC* gene family.

MATERIALS AND METHODS

Identification of *CNGC* genes in maize genome

To identify the *CNGC* genes in maize, 20 *Arabidopsis* and 16 rice *CNGC* protein sequences were retrieved from TAIR10 database (http://www.arabidopsis.org/index.jsp) and RGAP database (http://rice.plantbiology.msu.edu/), respectively. Then, two method were used to search against the maize protein sequences, one was built a Hidden Markov Model (HMM) to search against maize protein sequences, another was used local BLASTP method with a threshold of e-value <

1e-5. After that, a manual correction was performed to remove the redundancy. To further confirm the ZmCNGC proteins whether have the CNBD domain, those putative ZmCNGC protein sequences were summited to SMART (http://smart.embl-heidelberg.de/) (Letunic and Bork 2018) and NCBI-CDD (https://www.ncbi.nlm.nih.gov/cdd/) (Marchler-Bauer et al. 2017), proteins without the CNBD domains or the amino acids size of below < 200 were removed and the *ZmCNGC*s were confirmed.

The PI (theoretical isoelectric point), MW (molecular weight), and GRAVY of ZmCNGCs were predicted by ExPASy (http://web.expasy.org/protparam/) (Artimo et al. 2012). CELLO v.2.5 (http://cello.life.nctu.edu.tw/) was used to predict the subcellular location of ZmCNGCs. The information of chromosome distribution of ZmCNGCs and the sequences including DNA sequences, CDS, cDNA, up-stream 1500bps of *ZmCNGC*s were obtained from result of BLASTN search in the Ensembl Plant database (http://plants.ensembl.org/index.html) (Bolser et al. 2016).

Multiple alignments, phylogenetic analysis and gene duplication analysis

Multiple sequences alignments were performed using T-COFFEE web tool (http://tcoffee.crg.cat/apps/tcoffee/index.html) (Di Tommaso et al. 2011) and visualized by ESPript with default program (Robert and Gouet 2014). The maximum likelihood (ML) phylogenetic tree was constructed by MEGA 7 with bootstrap 1000 replications and Jones-Taylor-Thornton (JTT) model. To further validate the accuracy of ML tree, the un-root phylogenetic tree was constructed with 1000 bootstrap replications using MEGA 7 (Kumar et al. 2016) based on the full-length protein sequences alignment. Segmental duplication between maize genes as well as the synteny block between maize and *Sorghum*, rice and *Brachypodium* were obtained from the Plant Genome Duplication database (Lee et al. 2013). The substitution rates (Ka/Ks) of duplication events were calculated by using the DnaSP v5 (Librado and Rozas 2009), and the divergence times (Mya) was calculated as a formula $Mya = Ks/2\lambda \times 10^{-6}$, therein $\lambda = 6.5 \times 10^{-9}$ (Lynch and Conery 2000).

Analyses of genes structure and conserved motifs

The gene structure (exon-intron) of *ZmCNGC* genes were performed by the Gene Structure Display Server (GSDS, http://gsds.cbi.pku.edu.cn/) (Hu et al. 2015) using the CDS and genome sequence of *ZmCNGC* genes. The conserved motifs of ZmCNGC proteins were using the MEME Suite web server (http://meme-suite.org/index.html) (Bailey et al. 2015) with the maximum number of motif sets at 10 and optimum width of motifs from 6 to 200 amino acids.

Cis-acting regulatory elements of ZmCNGCs

The up-stream 1500bp DNA sequences of *ZmCNGC* genes were used to find cis-acting regulatory elements by 'Signal Scan Search' programs in the NEW PLACE database (https://sogo.dna.affrc.go.jp/cgi-bin/sogo.cgi?lang=en) (Higo et al. 1999).

Expression profiles of *ZmCNGC* genes and network interaction analysis

For understanding the expression of *ZmCNGC* genes in different tissues, two high throughput datasets of maize were obtained from the Expression Atlas datasets (https://www.ebi.ac.uk/gxa/home/) under accession number E-MTAB-3826 and E-MTAB-439. These data were used to analyze the expression of *ZmCNGC* genes in six different tissues (i.e. ear, embryo, endosperm, pollen, root and tassel) and different development stages in embryo, endosperm and seed, respectively. The FPKM values were used to calculate for each *ZmCNGC* genes. The interaction network was constructed on the base on the orthologs between maize and *Arabidopsis* using the AraNet v2 (Lee and Lee 2017) and visualized by the Cytoscape v3.4.0 (Shannon et al. 2003).

RESULTS

Identification of *CNGC* genes in maize

To identify a complete overview of *CNGC* genes in maize, we firstly used 20 *Arabidopsis* and 16 rice CNGC protein sequences align blast against maize protein sequences. After BLAST

Group II contained two maize CNGC genes (ZmCNGC4 and ZmCNGC5), five in *Arabidopsis* and three in rice. Similarly, Group III contained three maize *CNGC* genes (ZmCNGC6, ZmCNGC7, and ZmCNGC8). Group IV embraced five in rice and four in *Arabidopsis*, included one in Group IVa (ZmCNGC9) and three in Group IVb (ZmCNGC10, ZmCNGC11, and ZmCNGC12). Based on the phylogenetic tree among ZmCNGCs, AtCNGCs and OsCNGCs, maize CNGC genes were grouped into five groups (Figure 2A). Consistent with other plant CNGC genes, Group IVa contained only one or two gene members which formed the smallest group (Saand et al. 2015b). Based on the phylogenetic tree, results suggested that during evolution of CNGCs, Group IV CNGC genes emerged the earliest while Group I and II CNGCs appeared later. Furthermore, tree topology resulting from neighbor joining (NJ) analyses was the same as ML tree in Figure 1, all of the groups being retrieved (Supplemental File 1).

In the process of gene evolution, gene duplication play significant roles in generating new members and creating novel gene functions (Magadum et al. 2013). In this study, only one segmental duplication gene pair, ZmCNGC10-ZmCNGC11, was formed in maize genome (Table 2). To further investigated the evolutionary process between maize CNGCs and other gramineae plants, the genome synteny among *Sorghum*, rice and *Brachypodium* were also explored. Results showed that 2, 2 and 1 ZmCNGCs had homologous genes in *Sorghum*, rice and *Brachypodium*, respectively (Table 2). In addition, the substitution rate (Ka/Ks) is used to evaluate the specific positions under positive selection pressure after duplication (Mayrose et al. 2007), Ka/Ks =1, <1 or >1 indicates neutral, purifying and positive selection (Lynch and Conery 2000). As result shown in Table 2, the Ka/Ks of each gene pair was calculated and Ka/Ks of all gene pairs were less than 1, suggested that the selection pressure after duplication was strongly purifying selection. Moreover, the divergence time were also calculated, results indicated that the divergence time between maize CNGCs and other gramineae plants were unevenly divergence. Davidson et al. (2013) had showed that the gene duplication was a significant origin to generate novel genes, thus, these results implied that the duplication events gave principal role in gene evolution.

alignment, a total of 18 putative *ZmCNGC* genes were identified in the maize genome. Then, to confirm these 18 putative *ZmCNGC* genes, we used the SMART and NCBI CDD to find whether they contained the CNGC-specific domains (CNBD and transmembrane). After removing redundancy genes, a total of 12 *ZmCNGC* genes were detected, less than rice and *Arabidopsis* CNGC genes, the reason for it may be the gene duplication occur in rice and *Arabidopsis* (Paterson et al. 2004; Yu et al. 2005). To better understanding of the following analysis, the predicted ZmCNGC genes were designated as ZmCNGC1 to ZmCNGC12 based on the family classification (Table 1). To further access the existence of ZmCNGCs we identified, all the expressed sequence tags (EST) which aligned to *ZmCNGC* genes using the BLASTN program form NCBI, results demonstrated that only ZmCNGC3 showed no EST hits, other ZmCNGCs had more than 13 representative matches to ESTs. As shown in Table 1, five of them were located in chromosome 5, others were unevenly located in chromosomes 1, 2, 4, 6, 7, and 8. The characteristic features of these 12 *ZmCNGC* genes are listed in Table 1. The ZmCNGCs lengths ranged from 326 to 745 aa with average of 612 aa. The molecular weight of these proteins ranged from 38.63 kDa (ZmCNGC2) to 85.52 kDa (ZmCNGC8) and the pI value ranged from 8.92 (ZmCNGC4) to 9.75 (ZmCNGC12). Subcellular localization analysis showed that all of ZmCNGCs localized in the plasma membrane except for ZmCNGC3 localized in the nuclear fraction, this result is consistent with *Arabidopsis* (Lemtiri-Chlieh and Berkowitz 2004).

Phylogenetic and duplication analyses of ZmCNGCs

To better understanding the evolutionary relationship of CNGC proteins, a maximum likelihood (ML) phylogenetic tree was carried out based on the full-length protein alignments of 12 *ZmCNGCs*, 19 *AtCNGCs* and 16 *OsCNGCs* (Nawaz et al. 2014; Maser et al. 2001). As shown in the phylogenetic tree, the 47 CNGC proteins could be classified into five groups with high support values (Figure 1), including Group I, II, III, IVa and IVb, consistent with what was reported in flowering plant CNGCs (Saand et al. 2015b). Of those Groups, Group I contained three maize *CNGC* genes (ZmCNGC1, ZmCNGC2, and ZmCNGC3), five in *Arabidopsis* and two in rice.

192

193 **Multiple alignments, gene structure and conserved motif of ZmCNGCs**

194 The CNBD domain is a gene structural feature element in plant CNGCs which contain the PBC
195 and the hinge region (Diller et al. 2001). However, many ion transporters other than CNGCs also
196 have CNBD in C-terminal and a hexa-transmembrane (TM) in N-terminal, for example, potassium
197 AKT/KAT channels (Shaker type) also contain both a CNBD and TM domain. All AKT/ KAT-
198 type channels consist of six transmembrane (TM) regions with one P region (Su et al. 2001).
199 Therefore, we aligned 14 AKT/KAT proteins of maize from NCBI and 12 ZmCNGC protein
200 sequences (Supplemental File 2), results showed that these 14 AKT/KAT and 12 ZmCNGCs were
201 highly conserved and all of them composed six transmembrane domains (S1-S6) and a pore region.
202 We also paid attention to the CNBD region in AKT/KAT proteins and CNGCs, but they were not
203 highly conserved in PBC and hinge domain. The CNBD identified in CNGCs but no other proteins,
204 hence was recognized as a tool to identify plant CNGCs (Zelman et al. 2012). As shown in Figure
205 3, after aligning the CNBD region of maize CNGCs, the putative PBC (phosphate binding cassette)
206 and hinge domain were also identified, which were consistent with rice and *Arabidopsis* CNGCs
207 (Nawaz et al. 2014). Within the maize PBCs, a conserved phenylalanine (F), a stabilizing glycine
208 (G) and an acidic residue (D or E), followed by two aliphatic leucines (L), they were 100%
209 conserved inside the PBCs. As well, aromatic phenylalanine (F) and leucine (L) were the 100%
210 conserved within the hinge region. These two conserved regions occurred in between the CNBD
211 and CaMBD regions. On the basis of the corresponding alignment, a stringent motif (L-X(2)-G-
212 [ED]-ELL-[TSG]-W-[ACY]-L-X(10,20)-[SA]-X-T-X(7)-[EQ]-[AG]-F-X-L) that recognized in
213 those 12 maize CNGCs, which included the PBC and hinge domain, consistent with other plant
214 species (Saand et al. 2015b; Nawaz et al. 2014). The alignment of maize, rice and Arabidopsis
215 CNGCs also performed, results showed that no positions specific to the maize CNGC consensus,
216 suggested the PBCs and hinge were highly conserved among plants (Supplemental File 3).

217

218 Meanwhile, gene structure analysis could add better understanding to the gene function and

evolution. As a whole, the number of introns ranged from 0 to 7 (Figure 2B), different from rice and Arabidopsis CNGCs. In rice, OsCNGC ranged from 1 to 11 introns, while *Arabidopsis* CNGCs ranged from 4 to 10 introns (Nawaz et al. 2014). Notably, the Group IVa and IVb ZmCNGCs had distinct gene structures to those of all other groups, with more introns at different phases and lengths, consistent with most flowering plant species Group IV CNGC genes (Saand et al. 2015b).

Furthermore, motif-based recognition of proteins give understanding the evolution history (Seoighe and Gehring 2004). Ten putative motifs were characterized and named as motif1 to motif10 in ZmCNGCs. The relative positions of motif in the five groups were found to have various patterns (Figure 2C). All of those ZmCNGCs harbor motif1, motif2, motif3 and motif4, representing they were the typical ZmCNGC domain. The motif3 (QWRTWAA[CV]FIQ[AL]AW[RH]RY) were identified as the IQ domain among those 12 ZmCNGCs. Fischer et al. (2013) have showed that IQ as a functional motif within CaMBD and downstream of the CNBD domain, also conserves among plant CNGCs and enhances the changeable of Ca^{2+}-dependent channel control mechanisms in plant. In this study, all CNGC proteins contained IQ motif, suggesting that those proteins bind CaM in a Ca^{2+}-dependent manner. Notably, motif4 as sequence logo of CNBD domain, which most conserved in plants and animals (Jackson et al. 2007). In addition, 10 ZmCNGCs except ZmCNGC2 and ZmCNGC3 possessed motif 5 and 6 which associated with ion transport (Nawaz et al. 2014). Although other motifs had not been reported in plants or animals, it hold an important position in the function of organism.

Prediction of cis-acting regulatory elements of ZmCNGC proteins

Cis-acting regulatory elements are important molecular witches which associated with the transcriptional regulatory of genes when encounter environment stresses (Nakashima et al. 2009). To better understand the possible biological processes of these ZmCNGCs involved, 1.5 kb upstream of *ZmCNGC* genes genomic sequences were used to identify cis-regulatory elements and these were submitted to the NEW PLACE web tool. 137 different putative cis-elements were found

to be associated with identified *ZmCNGC* genes and only 12, including CACTFTPPCA1, EBOXBNNAPA, DOFCOREZM, MYCCONSENSUSAT, CAATBOX1, GTGANTG10, WRKY71OS, GT1CONSENSUS, ROOTMOTIFTAPOX1, POLLEN1LELAT52, MYBCORE, and OSE2ROOTNODULE, were apparently in the promoter region of all *ZmCNGC* genes (Supplemental File 4: Table S1) and highly consistent with rice CNGCs, maybe these elements in the upstream region were conserved (Nawaz et al. 2014). Additionally, five cis-elements were gene-specific, such as ACGTCBOX, TATABOX3, HDZIP2ATATHB2, CTRMCAMV35S, -300CORE, ABREMOTIFAOSOSEM and MRNA3ENDTAH3 were unique to ZmCNGC6, ZmCNGC7, ZmCNGC8, ZmCNGC11, respectively. Also, some cis-elements were involved in different abiotic/biotic stimuli, including those such as hormone-response (i.e. abscisic acid, auxin, ethylene, etc.), stress-related (i.e. drought, temperatures, disease, etc.) and development-related (i.e. mesophyll specific, tissue specific, etc.), indicating that these *ZmCNGC* genes may be involved in regulating diverse stress responses. For example, study showed that the CACTFTPPCA1 motif is for mesophyll-specific gene expression in the C4 plant (Gowik et al. 2004); MYCCONSENSUSAT is MYC recognition site which related to abiotic stress signaling (Liu et al. 2015); and WRKY71OS is reported as binding site of WRKY71 which involved in gibberellin signaling pathway (Zhang et al. 2004). Different cis-elements presenting in *ZmCNGC* genes indicated that they may relate to different regulatory networks.

Co-expression network between ZmCNGC and other maize genes

To get the detailed information about the interaction relationship between *ZmCNGC* genes and other maize genes, the co-expression network based on the orthology-based predictions following the network in *Arabidopsis* were constructed. As shown in and Figure 4 and Supplemental File 4: Table S2, a total of 5 *ZmCNGC* genes including ZmCNGC5, ZmCNGC6, ZmCNGC7, ZmCNGC9 and ZmCNGC10 with 76 gene pairs of network interactions were identified. GO annotations of interacting genes were also performed. Some symbols such as SOS1, BPM2, SKOR and BPM4, which play an essential role in regulation of organ development and osmotic stress response were

shown. The co-expression network analysis of *ZmCNGC* genes may provide comprehensive information for understanding ZmCNGC genes transduction pathways in maize.

Expression profiles of *ZmCNGC* genes in different tissues

We performed transcriptome sequencing to evaluate tissue-specific expression levels of *ZmCNGC* genes in different tissues based on previous study (see method). As shown in Figure 5A and Supplemental File 4: Table S3, the expression levels among *ZmCNGC* genes were tissue-specific in their expression. For example, *ZmCNGC5* was highly expressed in pollen and all Group IVb genes (*ZmCNGC10*, *ZmCNGC11* and *ZmCNGC12*) were mainly expressed in embryo, while *ZmCNGC2* and *ZmCNGC6* were expressed at low levels in embryo, pollen and endosperm. We also evaluated some *ZmCNGC* genes in embryo, endosperm and seed expression in some days after pollination. As shown in Figure 5A, B, C and D, ZmCNGC3, ZmCNGC5 and ZmCNGC7 were not detected or had no expression in all tissues, except ZmCNGC5 only expressed in pollen. The embryo specific-expression gene *ZmCNGC10* is gradually went up with time in embryo (Figure 5B). ZmCNGC8 was highly expressed in embryo, endosperm and seed after pollination. Moreover, the IVb gene also had a similar expression pattern.

DISCUSSION

Based on the phylogenetic, gene structure and conserved motifs analysis, we classified 12 ZmCNGCs into five groups with high support values. Among these five groups, Group IV ZmCNGCs emerged the earliest and had more introns than other group genes, followed by Group I and II CNGCs genes during evolution of plants. Previous researches have demonstrated that duplication genes often obtain new functions to enhance plant adapt to the environment (Dias et al. 2003). Interestingly, we found that all Group IVb genes had participated in gene duplication. Compared to other *ZmCNGC* genes, the gene expression of Group IV genes were more intense, suggested that these three *ZmCNGC* genes obtained gene functions from other plants after duplication. As the ancestor of CNGC genes, Group IVb genes mainly expressed in embryo and

embryo after pollination, suggested that Group IVb genes play a significant role in embryo development. These results further illustrated that Group IVb gene play important role in maize gene duplication, evolution and expression. Meanwhile, due to the less duplication event occurred in maize CNGCs, the less of the gene numbers.

The presence of CNBD domain in C-terminal and a hexa-transmembrane (TM) in N-terminal are the characteristic in plant CNGCs (Saand et al. 2015b). However, many ion transporters other than CNGCs also have these domain. For example, as homologs of CNGCs, potassium AKT /KAT channels (Shaker type) also contain both a CNBD domain and a TM domain. All AKT/ KAT-type channels consist of six transmembrane (TM) regions with one P region (Su et al. 2001). Previous studies showed that CNGC-specific motif with PBCs and hinge domain only exist in plant CNGCs rather than other ion transporters, for example, in rice and *Arabidopsis*. In this study, 12 ZmCNGCs were identified in maize and found that maize contained PBCs and hinge in CNBD domain after aligning, further confirmed the previous hypothesis (Saand et al. 2015b). A stringent motif (L-X(2)-G-[ED]-ELL-[TSG]-W-[ACY]-L-X(10,20)-[SA]-X-T-X(7)-[EQ]-[AG]-F-X-L) that included the PBC and hinge domain, consistent with other plant species (Saand et al. 2015b; Nawaz et al. 2014). In animals, the hinge occurs within the CNBD itself was the difference between plants and animals (Jackson et al. 2007). However, no positions specific to the maize CNGC consensus, the PBCs and hinge were highly conserved in plants. It is noteworthy that the existence of CNBD is sufficient but not necessary for identifying a CNGC protein.

Meanwhile, most researches showed that cyclic nucleotide-gated channels (CNGCs) genes have been related to pollen development and in response to environmental stimulus. For example, *Arabidopsis CNGC16* is critical for pollen fertility under conditions of heat stress and drought stress (Tunc-Ozdemir et al. 2013b), and *CNGC18* has been shown to function in pollen tube tip growth (Frietsch et al. 2007). In rice, *OsCNGC13* promotes seed-setting rate by facilitating pollen tube growth in stylar tissues (Xu et al. 2017). *ZmCNGC1* and *ZmCNGC5*, two homologous gene of *CNGC16* and *CNGC18* in our study are mainly expressed in pollen, indicated they mainly involved in pollen development. Previous study showed that the *AtCNGC3* expressed in the

327 embryo, leaves and roots, the expression level of *ZmCNGC4* was consistent with *AtCNGC3* which
328 highly expressed in plant development except pollen (Kaplan et al. 2007).
329
330 **CONCLUSION**
331 A total of 12 CNGC genes were identified in maize using bioinformatics method based on the
332 presence plant CNGC-specific motif spanning the PBCs and hinge region with the CNBD of
333 CNGC proteins. Phylogenetic analyses Group IV ZmCNGCs emerged earliest and had more
334 introns than other group ZmCNGCs, while Group I and II seemed to have evolved later.
335 Significantly, ZmCNGCs genes were diverse in genes structure, protein length and size. We
336 modified a maize stringent motif (L-X(2)-G-[ED]-ELL-[TSG]-W-[ACY]-L-X(10,20)-[SA]-X-T-
337 X(7)-[EQ]-[AG]-F-X-L) that included the PBC and hinge domain. Expression profiles of
338 *ZmCNGC* genes were tissue-specific expressed and related to pollen development. Our results
339 provide a reference for plant CNGCs during gene evolution.
340
341 **ACKNOWLEDGEMENTS**
342 We extend our gratitude to Dr. Lihe Yu for providing valuable suggestions.
343
344 **ADDITIONAL INFORMATION AND DECLARATIONS**
345 **Grant Disclosures**
346 The following grant information was disclosed by the authors:
347 Natural Science Foundation of Heilongjiang Province, China: B2016010.
348 **Supplemental Information**
349 Supplemental File 1
350 The un-root neighbor joining (NJ) tree based on the maize, rice and *Arabidopsis* CNGCs protein
351 sequences by using MEGA7.
352 Supplemental File 2
353 Conserved regions and multiple sequences alignments between maize AKT/KAT channels gene