

Genome-wide identification and evolutionary analysis of *CNGC* gene families in sixteen Brassicaceae plant genomes

Akram Ali BALOCH¹ , Kaleem U. KAKAR^{2,*} , Asma ABRO¹ ,
Muhammad MUSHTAQ¹ , Zarqa NAWAZ³ , Abrar HUSSAIN¹ 

¹Department of Biotechnology, Faculty of Life Sciences, Balochistan University of Information Technology, Engineering, and Management Sciences (BUIITEMS), Quetta, Pakistan

²Department of Microbiology, Faculty of Life Sciences, Balochistan University of Information Technology, Engineering and Management Sciences (BUIITEMS), Quetta, Pakistan

³Department of Botany, University of Central Punjab, Rawalpindi, Pakistan

Received: 19.05.2022

Accepted/Published Online: 08.11.2022

Final Version: 23.11.2022

Abstract: Cyclic nucleotide-gated channels (CNGCs) are ligand-gated calcium signaling channels, play vital biological functions. Nonetheless, the *CNGC* gene families have not been well studied in Brassicaceae which includes commercially important and evolutionary model crops. Here, using in silico approaches, we recognized 414 *CNGC* genes in sixteen sequenced genomes of Brassicaceae species, which represent 10 genera of 8 tribes, ranging between 10-51 genes. These CNGCs can be classified into four main groups (I-IV) and two subgroups (IV-A and IV-B), of which, Group-II CNGC is the oldest that has gradually disappeared from the *Arabidopsis thaliana* L. genome. These results present a comprehensive overview of Brassicaceae *CNGC* families and provide a foundation for futuristic genomic studies for assessing and validating functions of CNGCs in plant growth, development, and stress responses.

Key words: Brassicaceae, cyclic nucleotide-gated ion channels, evolutionary model crops

1. Introduction

The Brassicaceae family often known as *Cruciferae* *crucifers* is the most significant model species in plant biology (Inturrisi et al., 2018). Brassicaceae comprises of 338 genera and more than 3700 species of flowering plants distributed throughout the world (Shankar et al., 2019; Sharma et al., 2022). This family includes *Arabidopsis thaliana* (L.) Heynh and well known economically and agronomically important plants e.g., *Brassica* (*B. rapa* L., *B. oleracea* L., *B. napus* L., *B. nigra* L. and *B. juncea* (L.) Czern (Raza et al., 2020; Kaur et al., 2022). *Brassica* species produce vegetable and oilseed crops, accounting for around 10% of global vegetable production and 12% of global edible vegetable oil production (Labana and Gupta, 1993; Beilstein et al., 2006; Klopsch et al., 2017). Brassicaceae members yield unique secondary metabolites that are not only different to the family, but also to the genus and species (Raza et al., 2020). This family consists expanding number of species that are used as research platforms in a variety of fields, including plant genomics, breeding, and evolutionary biology.

Over the past two decades, research on Brassicaceae has changed our understanding in every sector of

* Correspondence: kaleem.ullah3@buitms.edu.pk

plant biology. However, understanding the mode and mechanisms of the evolution of the angiosperm flower is a long-standing and central problem of evolutionary biology (Franzke et al., 2011). Moreover, the systematics and taxonomy of the Brassicaceae family are extremely complicated and need to be addressed with more relevant and trustworthy information (Kiefer et al., 2014; Baloch et al., 2022). As a result, Brassicaceae have recently become a focus of comparative genomes and genetics research.

Cyclic nucleotide gated channels (CNGCs) are calcium-permeable channels that are engaged in a variety of biological functions (Saand et al., 2015a; Baloch et al., 2021). The eukaryotic genomes harbor multiple *CNGC* genes, hence called as *CNGC* gene family. The *CNGC* genes and proteins have been identified and characterized from the genomes of several plants and animals. In animal systems, the *CNGC*-encoded proteins have been well investigated and reported to play important role in sensory transduction (optical and olfactory) (Zagotta and Siegelbaum, 1996; Broillet and Firestein, 1999; Kaupp and Seifert, 2002). In plants, the *CNGC* families have been reported in selected crop genomes, including *A. thaliana* (L.) Heynh (Mäser et al., 2001), *Physcomitrella patens* (Hedw.) Bruch & Schimp

(Gehring, 2013), rice (Nawaz et al., 2014), pear (Chen et al., 2015), tomato (Saand et al., 2015a), Chinese cabbage (Kakar et al., 2017), maize (Hao and Qiao, 2018), tobacco (Nawaz et al., 2019), and turnip (Baloch et al., 2021) respectively. The emerging picture suggests that plant CNGCs are involved in signal transduction pathways (Pandey and Sanyal, 2021) and regulate many key physiological processes such as transport and absorption of toxic or essential cations, Ca²⁺-signaling, development, pollen fertility, geotropism, leaf senescence, innate immunity, and abiotic and biotic stress resistance (Baloch et al., 2019). Using different *in silico* and experimental approaches, studies have unraveled many important features of plant CNGC genes and proteins. For example, most of these proteins reside inside membranes of the membrane-bounded organelles (Yuen and Christopher, 2013; Charpentier et al., 2016), and function in group-specific manner. Phylogenetically, the plant CNGC proteins can be unanimously classified into four major and two subclades. Each CNGC protein comprises a fully conserved transmembrane (TM)/ion-transport domain at N-terminus and cNMP-binding domain (CNBD) at C-terminus with overlapped Calmodulin (CaM) binding (CaMBD) and adjacent isoleucine–glutamine (IQ) domains inside the CNBD that is optional in some CNGCs. Besides, the CNBD contains two well-preserved regions known as PBC (phosphate binding cassettes) motif, which binds the sugar and phosphate moieties of the cNMP ligand, and a “hinge” region adjacent to the PBC, which is believed to contribute to ligand binding efficacy and selectivity (Chin et al., 2009; Saand et al., 2015a; DeFalco et al., 2016; Baloch et al., 2021). The latest criterion for correct identification of CNGC genes are their validation through CNGC-specific motif keys (Baloch et al., 2021; Baloch et al., 2022), which has been used only in selected genome-wide identification-based studies. Despite this progress, our understanding of plant CNGCs in terms of the genomic structural organization, functional and evolutionary mechanism is still not fully understood, primarily due to frequent misidentification and availability of the genomic sequence data for only a few species. Therefore, in the present study, the genome-wide identification, characterization, and evolutionary analysis of CNGC gene families in sixteen species of Brassicaceae was performed by using the recently released reference genomic data in the Brassica database (BRAD). We used *in silico* approaches and utilized various tools and programs for in-depth examination, validation, and phylogenetic analysis of CNGC family members in Brassicaceae.

2. Materials and methods

2.1. Identification of CNGC proteins in Brassicaceae family

Twenty *A. thaliana* (L.) Heynh CNGC family protein sequences taken from TAIR10 (Lamesch et al., 2012) were

used as reference sequences to perform homology-based search (BLASTP program with default parameters) in the BRAD (Cheng et al., 2011). In Blast searches, obtained retrieved sequences showing >75% similarity to the query sequences, were used as input in the domain analysis tools: Simple Modular Architecture Research Tool (SMART) (Letunic et al., 2015), Pfam (Finn et al., 2016), Conserved Domains Database (CDD) (Marchler-Bauer et al., 2015) and Motif search service on GenomeNet., with default cutoff parameters. Obtained sequences containing both CNBD and TM domain, and the presence of “PBC” and “hinge region” within the cNMP-binding domain were recognized as CNGC proteins. The presence of CaMBD and IQ domains inside the CNBD was considered optional in genome-wide identification. Finally, the presence of consensus motif keys “[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]-[X]-[F]-X-[L]) [[L] - X (2) - [G] -X(3)-[G] -X(1,2)-L -L -X -W -X(0,1,2)-[L] -X(7,14)[P] -X(1,5)-S-X(10)-[E] -X -[F] -X -L]” at 90% conservation in aligned proteins was used for validation of candidate CNGCs (Zelman et al., 2013; Baloch et al., 2021; Baloch et al., 2022). The identified CNGC genes were renamed on the basis of their phylogenetic tree positions.

2.2. Multiple sequence alignment and phylogenetic analysis

The multiple sequence alignments of the CNGC peptide/ or protein sequences of selected Brassicaceae species were performed by using the ClustalX 2.0 program (Larkin et al., 2007) and observed by GeneDoc. MEGA 6.0 software (Maximum likelihood, Jones-Taylor-Thornton model, bootstrap of 1000) was utilized to generate phylogenetic trees. Later, another tree (neighbor joining method) was constructed based on multiple sequence alignments of the CNGC proteins from the all selected Brassicaceae species and taking *Amborella trichopoda* Baill. CNGC (AMTR_s00210p00019190) as outgroup (Baloch et al., 2021). For identification purposes, the CNGC proteins of each species were individually aligned with AtCNGCs and phylogenies were performed.

3. Results and discussion

3.1. Genome-wide identification of CNGCs in Brassicaceae

The CNGC gene families have been described in limited important plants (Almoneafy et al., 2014; Chen et al., 2015; Saand et al., 2015b), including *B. oleracea* L. (Kakar et al., 2017), *Nicotiana tabacum* L. (Nawaz et al., 2019), and *B. rapa* L. (Baloch et al., 2021) by our group. However, genome-wide study of the CNGC genes have not been reported in many other important crucifer plants. In present study, we used *in silico* methods to perform de novo identification, characterization and evolutionary

analysis of CNGC gene families in sequenced genomes of Brassicaceae species. For this purpose, homology-based BLAST search was performed by using the available AtCNGCs as queries against selected genomes in BRAD. A total of 469 nonredundant putative gene sequences were retrieved and investigated for the occurrence of a CNBD domain and a TM domain. The CNBD domains are highly conserved among CNGCs, hence recognized as an authentic means to identify plant CNGCs (Lu et al., 2022). Moreover, these sequences were manually analyzed for the occurrence of a CNGC-specific plant motif in PBC and hinge region of the CNBD. Consequently, fifty-five truncated sequences with missing CNGC-specific domains were discarded from analysis (Table S1). Finally, 414 fully annotated CNGC genes were identified in sixteen species that represent 10 genera of 8 tribes of selected Brassicaceae, including our previously identified *BoCNGCs* (Kakar et al., 2017) and *BrCNGCs* (Baloch et al., 2021) (Figure 1). The identified gene families were named

as *AarCNGCs* (*Aethionema arabicum* (L.) Andr. ex DC.), *AICNGC* (*Arabidopsis lyrata* (L.) O’Kane & Al-Shehbaz.), *AalCNGCs* (*Arabis alpina* L.), *BjCNGCs* (*Brassica juncea* (L.) Czern.), *BnCNGCs* (*Brassica napus* L.), *BniCNGCs* (*Brassica nigra* L.), *BrCNGCs* (*Brassica rapa* L.), *CrCNGC* (*Capsella rubella* Reut.), *CsCNGCs* (*Cannabis sativa* L.), *EsCNGCs* (*Eutrema salsugineum* (Pall.) Al-Shehbaz & Warwick.), *LaCNGCs* (*Leavenworthia alabamica* (L.) Rollins.), *SiCNGCs* (*Sisymbrium irio* L.), *SpCNGCs* (*Schrenkiella parvula* (Schrenk) D.A.German & Al-Shehbaz.), *ThCNGCs* (*Thellungiella halophila* (C. A. Meyer) O. E. Schulz.) and *TsCNGCs* (*Thellungiella salsuginea* (Pall.) O.E.Schulz) (Figure S1). The CNGC genes number varied significantly among the plants (Saand et al., 2015a). Similarly, we found that the number of CNGC genes varied among selected Brassicaceae. The *CsCNGC*, *BnCNGC*, and *BjCNGC* were found to be the largest among the sixteen species, containing 51, 50, and 42 genes respectively (Figure 1). The *AalCNGC* family was the shortest among

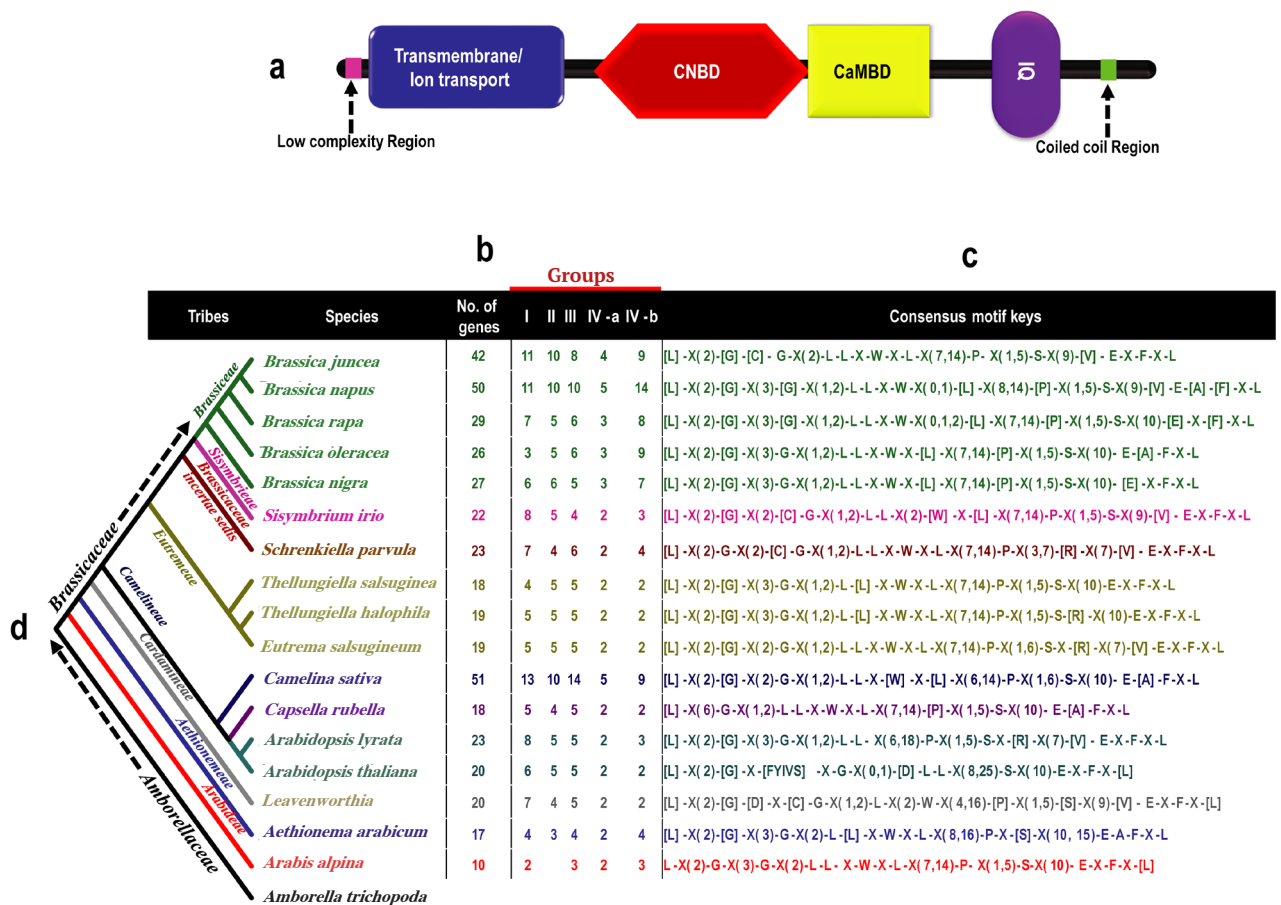


Figure 1. Graphical representation of CNGC genes in Brassicaceae. (a) A model showing the domain architecture of Brassicaceae CNGCs. (b) Total number of CNGCs and their distribution in groups. (c) Consensus motif keys showing conserved regions of CNBD spanning PBD and hinge regions. (d) Phylogenetic relationships of sixteen Brassicaceae species based on neighbor joining tree of AtCNGC1 orthologues using *Am. trichopoda* Baill. CNGC (AMTR_s00210p00019190) as outgroup.

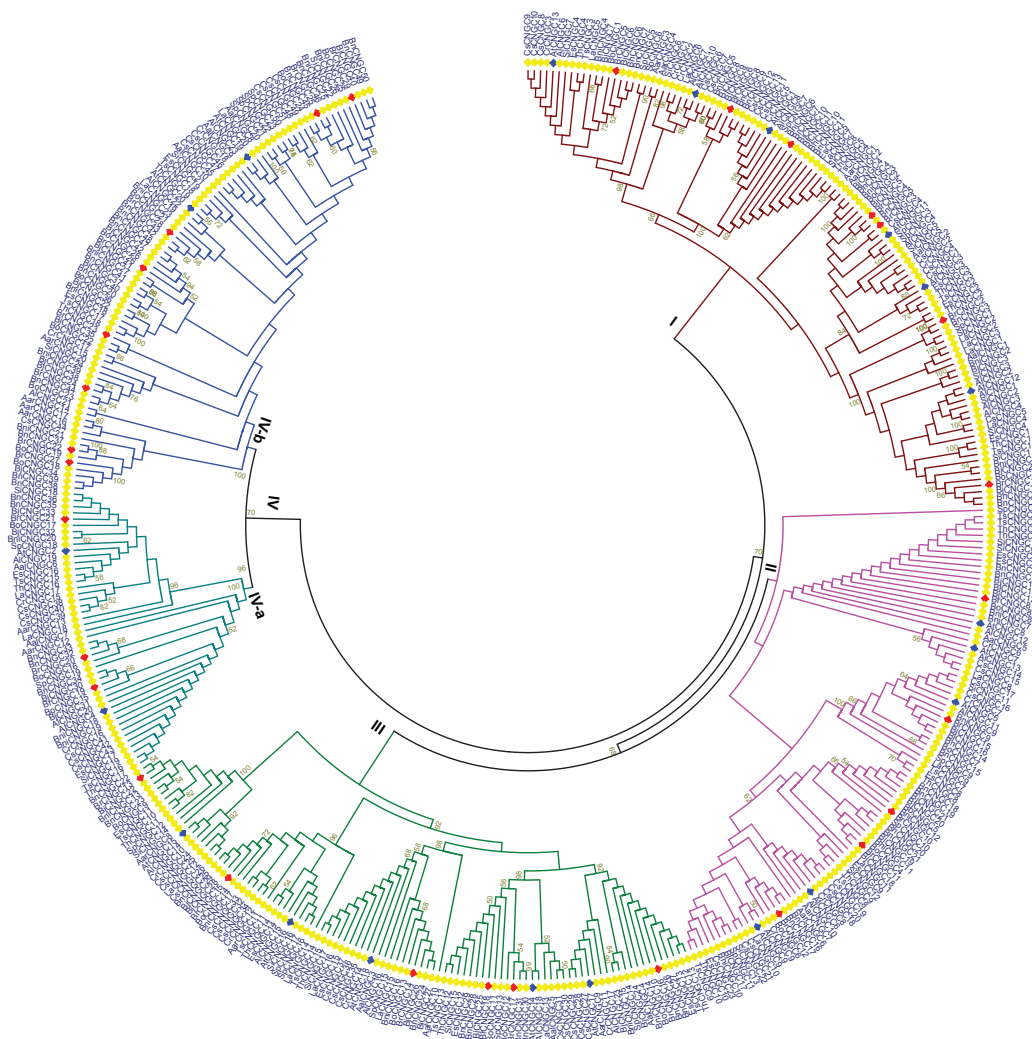


Figure 2. Phylogenetic tree of Brassicaceae CNGC families. The combined tree constructed (maximum likelihood tree with MEGA 6.0, under the Jones-Taylor-Thornton model, bootstrap of 1000) for 434 CNGC proteins, including 414 from sixteen species and 20 from *A. thaliana* (L.) Heynh.

the sixteen plant species, containing only 10 member genes. Vegetable crops (*Brassica*) had more CNGC genes than other crucifers, except *C. sativa* L., implying that CNGC genes expanded after divergence from a common ancestor (Cheng et al., 2014). The comprehensive genes and domain information of the selected species of the Brassicaceae CNGCs are given in Table S2.

3.2. Multiple sequence alignment, conserved domains, and consensus motifs of Brassicaceae CNGCs

The multiple sequence alignment of 434 CNGC proteins, including twenty AtCNGCs, exhibited low to moderate similarity, varying between 50% and 75%. However, each member of the CNGC family contained fully conserved CNBD and TM domains. Most of them harbored overlapped CaMBD and adjacent IQ domains (Figures

1a–1d). The two most conserved regions were recognized inside the CNBD: a PBC motif, which binds the sugar and phosphate moieties of the cNMP ligand, and a “hinge” region that is head-to-head with the PBC and involved in ligand binding effectiveness and choosiness (DeFalco et al., 2016; Baloch et al., 2021). At >90% conservation during alignment, we deduced the CNGC-specific consensus motif keys spanning the PBC and hinge region within the CNBD domain of each crucifer species, which validated all 414 CNGCs identified in this study (Figure 1c and Figure S2). Interestingly, our identified motif contained two variable amino acids (P and V) which were conserved in the hinge region in contrast to the previously described motifs (Nawaz et al., 2014; Saand et al., 2015a). Therefore, the presence of P and V may be

used to identify other crucifer CNGCs. Taking the advice of recent research (Li et al., 2019), the motifs were relaxed enough to avoid miss-identification of new CNGCs and ensured that keys are applicable to identify other plant CNGCs (Zelman et al., 2013; Saand et al., 2015a). Using these keys, four additional CNGCs from the *L. alabamica* (L.) Rollins genome were identified and two accessions (i.e. *Carubv10025546m* and *Carubv10013067m*) from *C. rubella* Reut and a single accession (i.e. *BjuA000233*) from *B. juncea* (L.) Czern were discarded. It may be that these removed accessions are abnormal CNGCs and lost their functions during evolution (Kakar et al., 2017). However, to clearly illustrate this point, more research is required.

3.3. Phylogenetic analysis and evolution of Brassicaceae CNGCs

To understand the evolutionary pattern of CNGCs in Brassicaceae, we inferred a rooted tree of single gene orthologs, using *Am. trichopoda* Baill. CNGC as an outgroup. The phylogenetic position of these genes in the sixteen representative genomes was nearly similar, showing that the CNGC gene family first originated in *Arabideae* among eight tribes, which bifurcated into *Aethionemeae* and *Cardamineae* leading to the origination of *Camelineae*, *Eutremeae*, *Brassicaceae incertae sedis*, *Sisymbrieae*, and *Brassicaceae* respectively (Figure 1d).

Phylogenetic analysis showed that CNGC proteins of flowering plant species formed four main classes (I, II, III, Iva, and IVb) (Saand et al., 2015a; Hao and Qiao, 2018). To identify the homology of the 434 CNGCs, a rooted phylogenetic tree of the CNGCs was constructed (Figure 2). Although, the Brassicaceae CNGC families exhibited low similarity between orthologs, their clustering patterns into four main classes (class-I to class-IV) was similar to *A. thaliana* (L.) Heynh (Mäser et al., 2001b), tomato (Saand et al., 2015a), and maize (Hao and Qiao, 2018). Group-IV, which made the largest clade with 127 members, was divided into two subclades, IV-a (43 members) and IV-b (84 members). Phylogenies of individual species with AtCNGCs resulted in similar clustering patterns, except for the AalCNGCs family that did not comprise Group-II (Figure S1). This observation was further supported by

a rooted tree, indicating that during evolution, Group-II CNGCs emerged the earliest amongst entire CNGC groups (I, III, and IV) in crucifer plants (Figure S3) (Saand et al., 2015a). In addition, the gene size of different groups within an individual species and among the groups of different species of Brassicaceae were also distinguishable (Figure 1b). These results are in agreement with previous findings (Saand et al., 2015a; Hao and Qiao, 2018; Mao et al., 2021).

4. Conclusion

In conclusion, this work is the first comprehensive and systematic analysis of the CNGC gene families in sixteen selected Brassicaceae species. This work distinguishes and accomplishes the remaining gaps in the literature, and presents a vibrant representation of plant CNGCs in general, and Brassicaceae in particular.

The outcomes of the current study will certainly provide a foundation, and drive the research forward to the next level, where the researchers can select and clone novel candidate CNGC genes to study signaling pathway mechanisms in detail and make newly improved cultivars through breeding. Additionally, this research will help to establish the roles of expressed candidate CNGC genes in Brassicaceae via gene silencing and cloning in the future.

Acknowledgment

We are grateful to the bioinformaticians who contributed to the development of the BRAD.

Conflict of interests

The authors declare that they have no competing interests.

Funding

The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Availability of data

The sequence datasets generated and analyzed during the current study are available at BRAD database (<http://brassicadb.cn>).

References

- Almoneafy AA, Kakar KU, Nawaz Z, Li B, Chun-lan Y et al. (2014). Tomato plant growth promotion and antibacterial related-mechanisms of four rhizobacterial *Bacillus* strains against *Ralstonia solanacearum*. *Symbiosis* 63 (2):59-70. doi:10.1007/s13199-014-0288-9
- Baloch AA, Kakar KU, Nawaz Z, Ahmed J (2019). Recent trend of genome-wide multigene family analysis and their role in plant drought tolerance. *Annals of Agricultural & Crop Sciences* 4 (2):1046.
- Baloch AA, Kakar KU, Nawaz Z, Mushtaq M, Abro A et al. (2022). Comparative genomics and evolutionary analysis of plant CNGCs. *Biology Methods & Protocols* 7 (1):bpac018. doi:10.1093/biomethods/bpac018
- Baloch AA, Raza AM, Rana SS, Ullah S, Khan S et al. (2021). BrCNGC gene family in field mustard: genome-wide identification, characterization, comparative synteny, evolution and expression profiling. *Scientific reports* 11 (1):1-16. doi:10.1038/s41598-021-03712-y

- Beilstein MA, Al-Shehbaz IA, Kellogg EA (2006). Brassicaceae phylogeny and trichome evolution. *American Journal of Botany* 93 (4):607-619. doi:10.3732/ajb.93.4.607
- Broillet MC, Firestein S (1999). Cyclic Nucleotide-Gated Channels: Molecular Mechanisms of Activation. *Annals of the New York Academy of Sciences* 868 (1):730-740. doi:10.1111/j.1749-6632.1999.tb11352.x
- Charpentier M, Sun J, Martins TV, Radhakrishnan GV, Findlay K et al. (2016). Nuclear-localized cyclic nucleotide-gated channels mediate symbiotic calcium oscillations. *Science* 352 (6289):1102-1105. doi: 10.1126/science.aae0109
- Chen J, Yin H, Gu J, Li L, Liu Z et al. (2015). Genomic characterization, phylogenetic comparison and differential expression of the cyclic nucleotide-gated channels gene family in pear (*Pyrus bretschneideri* Rehd.). *Genomics* 105 (1):39-52. doi:10.1016/j.ygeno.2014.11.006
- Cheng F, Liu S, Wu J, Fang L, Sun S et al. (2011). BRAD, the genetics and genomics database for *Brassica* plants. *BMC Plant Biology* 11 (1):1. doi:10.1186/1471-2229-11-136
- Cheng F, Wu J, Wang X (2014). Genome triplication drove the diversification of *Brassica* plants. *Horticulture Research* 1:14024. doi:10.1038/hortres.2014.24
- Chin K, Moeder W, 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 (7):668-677. doi:10.1139/B08-147
- DeFalco TA, Marshall CB, Munro K, Kang HG, Moeder W et al. (2016). Multiple calmodulin-binding sites positively and negatively regulate Arabidopsis CYCLIC NUCLEOTIDE-GATED CHANNEL12. *The Plant Cell* 28 (7):1738-1751. doi:10.1105/tpc.15.00870
- Finn RD, Coghill P, Eberhardt RY, Eddy SR, Mistry J et al. (2016). The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Research* 44 (D1):D279-D285. doi:10.1093/nar/gkv1344
- Franzke A, Lysak MA, Al-Shehbaz IA, Koch MA, Mummenhoff K (2011). Cabbage family affairs: the evolutionary history of Brassicaceae. *Trends in Plant Science* 16 (2):108-116. doi:10.1016/j.tplants.2010.11.005
- Gehring C (2013). *Cyclic Nucleotide Signaling in Plants*. 1st ed. USA: Humana press.
- Hao L, Qiao X (2018). Genome-wide identification and analysis of the CNGC gene family in maize. *PeerJ* 6:e5816. doi:10.7717/peerj.5816
- Inturrisi FC, Bayer PE, Yang H, Chan C-KK, Edwards D et al. (2018). Genome-wide analysis of NBS-LRR genes in Indian mustard (*Brassica juncea*) and prediction of candidate disease resistance genes. *Phytopathology* 108 (10):111.
- Kakar KU, Nawaz Z, Kakar K, Ali E, Almoneafy AA et al. (2017). Comprehensive genomic analysis of the CNGC gene family in *Brassica oleracea*: novel insights into synteny, structures, and transcript profiles. *BMC Genomics* 18 (1):1-18. doi:10.1186/s12864-017-4244-y
- Kaupp UB, Seifert R (2002). Cyclic nucleotide-gated ion channels. *Physiological Reviews* 82 (3):769-824. doi:10.1152/physrev.00008.2002
- Kaur G, Rajarammohan S, Kumar S, Verma R, Kaur G et al. (2022) Genomic Designing for Biotic Stress Resistance in Rape and Mustard. In: Kole C (editors) *Genomic Designing for Biotic Stress Resistant Oilseed Crops*. 1st ed. Denmark: Springer Cham, pp. 215-264.
- Kiefer M, Schmickl R, German DA, Mandáková T, Lysak MA et al. (2014). BrassiBase: introduction to a novel knowledge database on Brassicaceae evolution. *Plant and Cell Physiology* 55 (1):e3-e3. doi:10.1093/pcp/pt158
- Klopsch R, Witzel K, Börner A, Schreiner M, Hanschen FS (2017). Metabolic profiling of glucosinolates and their hydrolysis products in a germplasm collection of *Brassica rapa* turnips. *Food Research International* 100:392-403. doi:10.1016/j.foodres.2017.04.016
- Labana K, Gupta M (1993). Importance and Origin. In: Labana K, Banga S., Banga S (editors) *Breeding Oilseed Brassicas*. Monographs on Theoretical and Applied Genetics, vol 19. 1st ed. Berlin, Heidelberg: Springer, pp. 1-7.
- Lamesch P, Berardini TZ, Li D, Swarbreck D, Wilks C et al. (2012). The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. *Nucleic Acids Research* 40 (D1):D1202-D1210. doi:10.1093/nar/gkr1090
- Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA et al. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* 23 (21):2947-2948. doi:10.1093/bioinformatics/btm404
- Letunic I, Doerks T, Bork P (2015). SMART: recent updates, new developments and status in 2015. *Nucleic Acids Research* 43 (D1):D257-D260. doi:10.1093/nar/gku949
- Li Q, Yang S, Ren J, Ye X, Liu Z (2019). Genome-wide identification and functional analysis of the cyclic nucleotide-gated channel gene family in Chinese cabbage. *3 Biotech* 9 (3):1-14. doi:10.1007/s13205-019-1647-2
- Lu Z, Yin G, Chai M, Sun L, Wei H et al. (2022). Systematic analysis of CNGCs in cotton and the positive role of GhCNGC32 and GhCNGC35 in salt tolerance. *BMC Genomics* 23 (1):1-13. doi:10.1186/s12864-022-08800-5
- Mao X, Wang C, Lv Q, Tian Y, Wang D et al. (2021). Cyclic nucleotide gated channel genes (CNGCs) in Rosaceae: genome-wide annotation, evolution and the roles on Valsa canker resistance. *Plant Cell Reports* 40 (12):2369-2382. doi:10.1007/s00299-021-02778-2
- Marchler-Bauer A, Anderson JB, Cherukuri PF, DeWeese-Scott C, Geer LY et al. (2005). CDD: a Conserved Domain Database for protein classification. *Nucleic Acids Research* 33 (1):192-196. doi:10.1093/nar/gki069
- Mäser P, Thomine S, Schroeder JI, Ward JM, Hirschi K et al. (2001). Phylogenetic relationships within cation transporter families of Arabidopsis. *Plant Physiology* 126 (4):1646-1667. doi:10.1104/pt.126.4.1646

- Nawaz Z, Kakar KU, Saand MA, 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 (1):1-18. doi:10.1186/1471-2164-15-853
- Nawaz Z, Kakar KU, Ullah R, Yu S, Zhang J et al. (2019). Genome-wide identification, evolution and expression analysis of cyclic nucleotide-gated channels in tobacco (*Nicotiana tabacum* L.). *Genomics* 111 (2):142-158. doi:10.1016/j.ygeno.2018.01.010
- Pandey GK, Sanyal SK (2021) Plant Ligand-Gated Channels 2. In: CNGC Functional Dissection of Calcium Homeostasis and Transport Machinery in Plants. 1st ed. Switzerland: Springer Cham, pp. 63-73.
- Raza A, Hafeez MB, Zahra N, Shaukat K, Umbreen S et al. (2020). The Plant Family Brassicaceae: Introduction, Biology, and Importance. In: Hasanuzzaman M (editors) The Plant Family Brassicaceae. 1st ed. Singapore: Springer, pp. 1-43.
- Saand MA, Xu YP, Muniyampundu JP, Li W, Zhang XR et al. (2015a). Phylogeny and evolution of plant cyclic nucleotide-gated ion channel (CNGC) gene family and functional analyses of tomato CNGCs. *DNA Research* 22 (6):471-483. doi:10.1093/dnares/dsv029
- Saand MA, Xu YP, Li W, Wang JP, Cai XZ (2015b). Cyclic nucleotide gated channel gene family in tomato: genome-wide identification and functional analyses in disease resistance. *Frontiers in Plant Science* 6:303. doi:10.3389/fpls.2015.00303
- Shankar S, Segaran G, Sundar RDV, Settu S, Sathivelu M (2019). Brassicaceae-A classical review on its pharmacological activities. *International Journal of Pharmaceutical Sciences Review and Research* 55 (1):107-113.
- Sharma HK, Kumar A, Singh VV, Meena HS, Meena BL et al. (2022). Genetic Resources of Brassicas. In: Priyadarshan, P, Jain SM (editors) *Cash Crops*. 1st ed. Denmark: Springer Cham, pp. 285-337.
- Yuen CC, Christopher DA (2013). The group IV-A cyclic nucleotide-gated channels, CNGC19 and CNGC20, localize to the vacuole membrane in *Arabidopsis thaliana*. *AoB Plants* 5. doi:10.1093/aobpla/plt012
- Zagotta WN, Siegelbaum SA (1996). Structure and function of cyclic nucleotide-gated channels. *Annual Review of Neuroscience* 19 (1):235-263.
- Zelman AK, Dawe A, Berkowitz GA (2013) Identification of Cyclic Nucleotide Gated Channels Using Regular Expressions. In: Gehring C (editors) *Cyclic Nucleotide Signaling in Plants: Methods and Protocols*. 1st ed. USA: Humana Press, pp. 207-224.

Table S1. List of truncated gene accessions discarded during preliminary investigation.

Species	Accession	Genome location	Start	Stop	Strand	Reason
<i>Aethionema arabicum</i>	AA_scaffold1130_86	AA_scaffold4130	431199	437168	-	Additional irrelevant domains
	AA_scaffold5078_104	AA_scaffold5078	527339	533754	+	Additional irrelevant domains
	AA_scaffold1200_14	AA_scaffold1200	79941	82626	+	Additional irrelevant domains
	AA_scaffold277_3	AA_scaffold277	23575	26340	-	No cNBD domain
	AA_scaffold6631_1	AA_scaffold6631	6013	7624	-	No cNBD domain
	AA_scaffold277_2	AA_scaffold277	8588	11213	-	No cNBD domain
<i>Arabidopsis lyrata</i>	AL_scaffold_0006_3422	c06	21817546	21818993	-	No cNBD domain
<i>Arabis alpina</i>	KFK37412.1	c04	information not given in database	alternatively spliced variant 2 of Aa1CNGC1		truncated
	KFK37410.1	c04	information not given in database	no		truncated
	KFK34623.1	c05	information not given in database	alternatively spliced variant 2 of Aa1CNGC8		truncated
<i>Brassica juncea</i>	BjuA000233	information not given in database				No PBC/hinge
<i>Brassica napus</i>	GSBRNA2T00002433001	chrC03	18993736	18995295	+	No cNBD domain
	GSBRNA2T00054579001	chrC02_random	448541	452088	-	No cNBD domain
	GSBRNA2T00081174001	chrCnn_random	11619683	11621115	+	truncated
	GSBRNA2T00105787001	chrC06	19271982	19275841	-	No cNBD domain
	GSBRNA2T00105787001	chrC06	19271982	19275841	-	No cNBD domain
	GSBRNA2T00105799001	chrC06	19337432	19339987	-	No cNBD domain
	GSBRNA2T00135561001	chrC09	44128700	44131819	+	Two cNBD domains
	Bo1g001000	C1	1449	1625	-	truncated
	Bo3g005110	C3	1956031	1956587	+	truncated
	Bo3g052670	C3	20802976	20804536	+	No cNBD domain
<i>Brassica oleracea</i>	Bo5g001380	C5	206348	207695	+	truncated
Bo5g104990	C5	33415496	33416054	-	truncated	
Bo6g074480	C6	21615284	21619156	-	truncated	
Bo8g027170	C8	7780841	7781396	+	No cNBD domain	
Bo9g001000	C9	728	2075	-	truncated	

Table S1. Continued).

<i>Capsella rubella</i>	Carubv10009174m	Sca1	16258438	16260568	-	No cNBD domain
	Carubv10026259m	information not given in database				No cNBD domain
	Carubv10025546m	Sca4	14370693	14373872	+	No PBC/hinge region
	Carubv10013067m	Sca3	6162568	6165860	+	No PBC/hinge region
<i>Camelina sativa</i>	XP_010505715.1					No cNBD domain/truncated
	XP_010435405.1					No cNBD domain/truncated
	XP_010435423.1					No cNBD domain/truncated
	XP_010447699.1					No cNBD domain/truncated
	XP_010456193.1					No cNBD domain/truncated
	XP_010509562.1					No cNBD domain/truncated
	XP_010509561.1					No cNBD domain/truncated
	XP_010505099.1					No cNBD domain/truncated
	XP_010509558.1					No cNBD domain/truncated
	XP_010505102.1					No cNBD domain/truncated
	XP_010497561.1					No cNBD domain/truncated
	XP_010418823.1					No cNBD domain/truncated
	XP_010516784.1					No cNBD domain/truncated
	XP_010497423.1					No cNBD domain/truncated
<i>Leavenworthia alabamica</i>	LA_scaffold2193_1	LA_scaffold2193	867	4939	-	Additional irrelevant domains
	LA_scaffold1983_6	LA_scaffold1983	18811	24703	+	Additional irrelevant domains
	LA_scaffold1126_72	LA_scaffold1126	216780	223893	+	Additional irrelevant domains
	LA_scaffold241_12	LA_scaffold241	53259	57590	-	No cNBD domain
<i>Sisymbrium irio</i>	SI_scaffold2657_27	SI_scaffold2657	151847	156404	-	Additional irrelevant domains
	c0001_01246	information not given in database				Additional irrelevant domains
<i>Schrenkiella parvula</i>	c0006_00128	Sp7	5422915	5424740	-	No cNBD domain
	Thhalv10002457m	scaffold_4	4667687	4673358	-	No cNBD domain
<i>Thellungiella halophila</i>	Tsa3g16140	ch3-04	187530	190864	+	No Domain
	Tsa3g25770	ch3-14	117343	121067	+	No cNBD domain
<i>Thellungiella salsuginea</i>	Tsa4g31880	ch4-82	2896179	2898483	-	No Ion transport domain

Table S2. List and details of CNGC family genes identified in Brassicaceae species.

Specie	Gene	Gene Information				Protein Information							Group
		Accession	chromosome	Start	Stop	Strand	Primary domain	Secondary domain					
<i>Aethionema arabicum</i>	AarCNGC1	AA_scaffold2791_62	AA_scaffold2791	250126	253871	+	cNMP	TM	CaMBD	IQ			
	AarCNGC2	AA_scaffold3427_5	AA_scaffold3427	47360	49494	+	cNMP	TM	CaMBD				G 1
	AarCNGC3	AA_scaffold6229_5	AA_scaffold6229	10072	13040	-	cNMP	TM	CaMBD	IQ			
	AarCNGC4	AA_scaffold6417_2	AA_scaffold6417	13276	17215	+	cNMP	TM	CaMBD	IQ			
	AarCNGC5	AA_scaffold652_46	AA_scaffold652	186497	189373	-	cNMP	TM	CaMBD	IQ			
	AarCNGC6	AA_scaffold4887_8	AA_scaffold4887	48567	53873	+	cNMP	TM	CaMBD	IQ			G 2
	AarCNGC7	AA_scaffold1970_34	AA_scaffold1970	211216	215495	-	cNMP	TM	CaMBD	IQ			
	AarCNGC8	AA_scaffold4918_54	AA_scaffold4918	230013	232476	-	cNMP	TM	CaMBD	IQ			
	AarCNGC9	AA_scaffold4393_9	AA_scaffold4393	58637	64832	-	cNMP	TM	CaMBD	IQ			G 3
	AarCNGC10	AA_scaffold5136_2	AA_scaffold5136	2002	4906	-	cNMP	TM	CaMBD	IQ			
	AarCNGC11	AA_scaffold5525_47	AA_scaffold5525	238125	244875	+	cNMP	TM	CaMBD	IQ			
	AarCNGC12	AA_scaffold3842_72	AA_scaffold3842	350510	356702	+	cNMP	TM	CaMBD	IQ			
	AarCNGC13	AA_scaffold2718_201	AA_scaffold2718	937192	940986	+	cNMP	TM	CaMBD	IQ			G4a
	AarCNGC14	AA_scaffold1748_49	AA_scaffold1748	218458	221593	-	cNMP	TM	CaMBD	IQ			
	AarCNGC15	AA_scaffold1748_48	AA_scaffold1748	214790	218026	-	cNMP	TM	CaMBD	IQ			G4b
	AarCNGC16	AA_scaffold1748_47	AA_scaffold1748	23575	26340	-	cNMP	TM	CaMBD	IQ			
	AarCNGC17	AA_scaffold1748_46	AA_scaffold1748	205863	208731	-	cNMP	TM	CaMBD	IQ			
<i>Arabidopsis lyrata</i>	AICNGC1	fgenes1_pm.C_scaffold_2000217	c02	2007800	2010303	+	cNMP	TM	CaMBD	IQ			
	AICNGC2	scaffold_403600.1	c04	22532508	22537650	+	cNMP	TM	CaMBD	IQ			
	AICNGC3	fgenes1_pm.C_scaffold_1003675	c01	32007771	32010291	-	cNMP	TM	CaMBD				G 1
	AICNGC4	scaffold_105572.1	c01	31715424	31717448	-	cNMP	TM	CaMBD				
	AICNGC5	scaffold_403599.1	c04	24539220	24541943	+	cNMP	TM	CaMBD	IQ			
	AICNGC6	fgenes1_pm.C_scaffold_6002898	c06	24539220	24541943	+	cNMP	TM	CaMBD	IQ			
	AICNGC7	fgenes2_kg.1_95_AT1G01340.1	c01	377570	380538	+	cNMP	TM	CaMBD	IQ			
	AICNGC8	scaffold_801783.1	c08	15510210	15512917	-	cNMP	TM	CaMBD	IQ			
	AICNGC9	fgenes1_pm.C_scaffold_7000887	c07	4696482	4699204	+	cNMP	TM	CaMBD	IQ			
	AICNGC10	fgenes1_pm.C_scaffold_4000263	c04	2255120	2257856	-	cNMP	TM	CaMBD	IQ			
	AICNGC11	fgenes2_kg.8_1710_AT5G57940.2	c08	18389681	18393605	+	cNMP	TM	CaMBD	IQ			G 2
	AICNGC12	fgenes1_pg.C_scaffold_1001793	c01	8386642	8389342	-	cNMP	TM	CaMBD	IQ			
	AICNGC13	AL_scaffold_0001_1661	c01	6756780	6759385	-	cNMP	TM	CaMBD	IQ			

Table S2. Continued).

<i>Arabidopsis lyrata</i>	AICNGC14	fgenesH1_pm.C_scaffold_4000624	c04	11908441	11911144	-	cNMP	TM	CaMBD	IQ	G 3
	AICNGC15	fgenesH1_pg.C_scaffold_5001286	c05	13572971	13575665	-	cNMP	TM	CaMBD	IQ	
	AICNGC16	fgenesH2_kg.6_1464_AT5G14870.1	c06	6035223	6037976	-	cNMP	TM	CaMBD	IQ	
	AICNGC17	fgenesH2_kg.7_1149_AT4G30360.1	c07	4782035	4785185	+	cNMP	TM	CaMBD	IQ	G 4a
	AICNGC18	fgenesH2_kg.4_379_AT2G24610.1	c04	3050291	3053715	+	cNMP	TM	CaMBD	IQ	
	AICNGC19	fgenesH2_kg.6_1521_AT5G15410	c06	6281062	6284561	-	cNMP	TM	CaMBD	IQ	G 4b
	AICNGC20	fgenesH2_kg.8_1326_AT5G54250.1	c08	16420408	16424890	-	cNMP	TM	CaMBD	IQ	
	AICNGC21	scaffold_302089.1	c03	7555627	7559037	+	cNMP	TM	CaMBD	IQ	G 1
	AICNGC22	scaffold_302087.1	c03	7543581	7546944	+	cNMP	TM	CaMBD	IQ	
	AICNGC23	fgenesH1_pg.C_scaffold_3001667	c03	7539617	7542783	+	cNMP	TM	CaMBD	IQ	G 3
AaICNGC1	AALP_AA4G253500	4	22227113	22230207	+	cNMP	TM	CaMBD	IQ		
AaICNGC2	AALP_AA4G253100	4	22196843	22200027	+	cNMP	TM	CaMBD	IQ	G 3	
AaICNGC3	AALP_AA8G477900	8	38750808	38753727	-	cNMP	TM	CaMBD	IQ		
AaICNGC4	AALP_AA7G156300	7	18864054	18867338	-	cNMP	TM	CaMBD	IQ	G 4a	
AaICNGC5	AALP_AA6G271300	6	28323545	28326037	+	cNMP	TM	CaMBD	IQ		
AaICNGC6	AALP_AA8G160300	8	10387770	10390974	-	cNMP	TM	CaMBD	IQ	G 4b	
AaICNGC7	AALP_AA8G326400	8	28448919	28453954	-	cNMP	TM	CaMBD	IQ		
AaICNGC8	AALP_AA5G169900	5	14392187	14395596	+	cNMP	TM	CaMBD	IQ	G 1	
AaICNGC9	AALP_AA3G198800	3	11647115	11650842	-	cNMP	TM	CaMBD	IQ		
AaICNGC10	AALP_AA3G200700	3	11832287	11835578	-	cNMP	TM	CaMBD	IQ	G 1	
BjCNGC01	BjuA016240	J04	21906685	21909239	+	cNMP	TM	CaMBD	IQ		
BjCNGC02	BjuB024217	J15	7828833	7831079	+	cNMP	TM	CaMBD	IQ	G 1	
BjCNGC03	BjuB022365	J14	20787546	20789900	-	cNMP	TM	CaMBD	IQ		
BjCNGC04	BjuB019851	J14	1174101	1176691	-	cNMP	TM	CaMBD	IQ	G 1	
BjCNGC05	BjuB019812	J14	913825	916804	+	cNMP	TM	CaMBD	IQ		
BjCNGC06	BjuA011392	J03	18965057	18968996	-	cNMP	TM	CaMBD	IQ	G 1	
BjCNGC07	BjuB017237	J13	26254017	26257911	-	cNMP	TM	CaMBD	IQ		
BjCNGC08	BjuA015292	J04	9197004	9199658	-	cNMP	TM	CaMBD	IQ	G 1	
BjCNGC09	BjuA038153	J10	7202514	7205504	-	cNMP	TM	CaMBD	IQ		
BjCNGC10	BjuA006962	J02	4674944	4677504	+	cNMP	TM	CaMBD	IQ	G 1	
BjCNGC11	BjuB012152	J12	41593077	41595648	-	cNMP	TM	CaMBD	IQ		

Table S2. Continued).

BnCNGC01	GSBRNA2T00058294001	chrA04	12665265	12668409	+	cNMP	TM	CaMBD	IQ
BnCNGC02	GSBRNA2T00127590001	chrC04	39436436	39439178	+	cNMP	TM	CaMBD	IQ
BnCNGC03	GSBRNA2T00098910001	chrA07	14705083	14708986	-	cNMP	TM		IQ
BnCNGC04	GSBRNA2T00074064001	chrC04	984082	987459	+	cNMP	TM	CaMBD	IQ
BnCNGC05	GSBRNA2T00132896001	chrA05	808631	812078	-	cNMP	TM	CaMBD	IQ
BnCNGC06	GSBRNA2T00112178001	chrA03	13180480	13184496	-	cNMP	TM	CaMBD	IQ
BnCNGC07	GSBRNA2T00153491001	chrC03	19518056	19522572	-	cNMP	TM	CaMBD	IQ
BnCNGC08	GSBRNA2T00009392001	chrC09	31884192	31886408	-	cNMP	TM	CaMBD	IQ
BnCNGC09	GSBRNA2T00093047001	chrA10	4931547	4934085	-	cNMP	TM	CaMBD	IQ
BnCNGC10	GSBRNA2T00036836001	chrA02	5360323	5362868	+	cNMP	TM	CaMBD	IQ
BnCNGC11	GSBRNA2T00102267001	chrC02	10092252	10095260	+	cNMP	TM	CaMBD	IQ
BnCNGC12	GSBRNA2T00057927001	chrA06	5629756	5632689	-	cNMP	TM	CaMBD	IQ
BnCNGC13	GSBRNA2T00054207001	chrC05	7083786	7086776	-	cNMP	TM	CaMBD	IQ
BnCNGC14	GSBRNA2T00052790001	chrA02	3784851	3788310	-	cNMP	TM	CaMBD	IQ
BnCNGC15	GSBRNA2T00114952001	chrC02	6416508	6419623	-	cNMP	TM	CaMBD	IQ
BnCNGC16	GSBRNA2T00067432001	chrA03	26087831	26090699	-	cNMP	TM	CaMBD	IQ
BnCNGC17	GSBRNA2T00099720001	chrC07	42093754	42097033	-	cNMP	TM	CaMBD	IQ
BnCNGC18	GSBRNA2T00012172001	chrC04	38310006	38313393	+	cNMP	TM	CaMBD	IQ
BnCNGC19	GSBRNA2T00068273001	chrA04	11848036	11851336	-	cNMP	TM	CaMBD	IQ
BnCNGC20	GSBRNA2T00046585001	chrCnn_random	46519629	46522244	+	cNMP	TM	CaMBD	IQ
BnCNGC21	GSBRNA2T00057361001	chrA09	29111335	29114828	+	cNMP	TM	CaMBD	IQ
BnCNGC22	GSBRNA2T00118729001	chrA07	12144138	12147369	+	cNMP	TM	CaMBD	IQ
BnCNGC23	GSBRNA2T00148160001	chrC04	14008149	14011414	-	cNMP	TM	CaMBD	IQ
BnCNGC24	GSBRNA2T00031646001	chrC09	44128700	44131819	+	cNMP	TM	CaMBD	IQ
BnCNGC25	GSBRNA2T00063043001	chrA06	9550763	9553801	+	cNMP	TM	CaMBD	IQ
BnCNGC26	GSBRNA2T00039899001	chrA01	3095672	3099840	+	cNMP	TM	CaMBD	IQ
BnCNGC27	GSBRNA2T00130317001	chrC01	4234002	4238184	+	cNMP	TM	CaMBD	IQ
BnCNGC28	GSBRNA2T00012272001	chrC04	37794024	37797281	-	cNMP	TM	CaMBD	IQ
BnCNGC29	GSBRNA2T00082312001	chrA04	12210464	12213711	+	cNMP	TM	CaMBD	IQ
BnCNGC30	GSBRNA2T00057304001	chrA09	28879650	28883074	-	cNMP	TM	CaMBD	IQ
BnCNGC31	GSBRNA2T00127302001	chrC08	32189874	32193318	-	cNMP	TM	CaMBD	IQ

Brassica napus (rape)

Table S2. Continued).

<i>Brassica napus</i> (rape)	BnCNGC32	GSBRNA2T00002884001	chrC09	33625801	33632938	-	cNMP	TM	CaMBD	IQ	G4a
	BnCNGC33	GSBRNA2T00078933001	chrA10	5759951	5765207	-	cNMP	TM	CaMBD	IQ	
	BnCNGC34	GSBRNA2T00041436001	chrA02	4902667	4907022	+	cNMP	TM	CaMBD	IQ	
	BnCNGC35	GSBRNA2T00135525001	chrA10	13494368	13497571	+	cNMP	TM	CaMBD	IQ	
	BnCNGC36	GSBRNA2T00035212001	chrC09	44006544	44009743	+	cNMP	TM	CaMBD	IQ	G4b
	BnCNGC37	GSBRNA2T00137471001	chrA03	16872203	16875542	+	cNMP	TM			
	BnCNGC38	GSBRNA2T00043747001	chrCnn_random	44483706	44487129	-	cNMP	TM			
	BnCNGC39	GSBRNA2T00137474001	chrA03	16863304	16866599	+	cNMP	TM			
	BnCNGC40	GSBRNA2T00101781001	chrC05	35081609	35084991	-	cNMP	TM	CaMBD	IQ	
	BnCNGC41	GSBRNA2T00153851001	chrA05	17157632	17161030	-	cNMP	TM	CaMBD	IQ	
	BnCNGC42	GSBRNA2T00048016001	chrCnn_random	47575565	47579101	-	cNMP	TM	CaMBD	IQ	
	BnCNGC43	GSBRNA2T00086350001	chrA01	19425000	19428539	+	cNMP	TM	CaMBD	IQ	
	BnCNGC44	GSBRNA2T00017118001	chrC01_random	1762249	1767846	+	cNMP	TM	CaMBD	IQ	
	BnCNGC45	GSBRNA2T00048015001	chrCnn_random	47560829	47564174	+	cNMP	TM	CaMBD	IQ	
	BnCNGC46	GSBRNA2T00086351001	chrA10	13494368	13497571	+	cNMP	TM	CaMBD	IQ	
	BnCNGC47	GSBRNA2T00101782001	chrC05	35087166	35090279	-	cNMP	TM	CaMBD	IQ	
	BnCNGC48	GSBRNA2T00153850001	chrA05	17162980	17166118	-	cNMP	TM	CaMBD	IQ	
	BnCNGC49	GSBRNA2T00101783001	chrC05	35092931	35096102	-	cNMP	TM	CaMBD	IQ	
	BnCNGC50	GSBRNA2T00153849001	chrA05	17167881	17171529	-	cNMP	TM	CaMBD	IQ	
	BniCNGC01	BniB038821	B5	26176034	26178286	-	cNMP	TM	CaMBD		G-1
BniCNGC02	BniB019484	B4	25373513	25375867	+	cNMP	TM		IQ		
BniCNGC03	BniB044122	B1	28375468	28378346	-	cNMP	TM		IQ		
BniCNGC04	BniB005559	B3	30460377	30464930	+	cNMP	TM	CaMBD	IQ		
BniCNGC05	BniB006017	scaffold_12.1	235834	238475	-	cNMP	TM	CaMBD	IQ		
BniCNGC06	BniB006616	B2	26878414	26881005	+	cNMP	TM	CaMBD	IQ		
BniCNGC07	BniB022693	B4	23635856	23638799	+	cNMP	TM	CaMBD	IQ		
BniCNGC08	BniB043299	B4	14672413	14675075	-	cNMP	TM	CaMBD	IQ		
BniCNGC09	BniB027956	B2	23436315	23439081	-	cNMP	TM	CaMBD	IQ		
BniCNGC10	BniB002567	scaffold_10.1	981272	983914	+	cNMP	TM	CaMBD	IQ		
BniCNGC11	BniB017483	B8	11634581	11637281	-	cNMP	TM	CaMBD	IQ		
BniCNGC12	BniB028064	B5	24929718	24932397	-	cNMP	TM	CaMBD	IQ	G-2	

Table S2. Continued).

<i>Brassica nigra</i>	BniCNGC13	BniB003614	scaffold_106.1	466463	469251	-	cNMP	TM	CaMBD	IQ	G-3
	BmiCNGC14	BmiB018653	B2	40782618	40785437	-	cNMP	TM	CaMBD	IQ	
	BniCNGC15	BniB002576	scaffold_10.1	1052094	1054831	+	cNMP	TM	CaMBD	IQ	
	BniCNGC16	BniB007639	B8	26292138	26295364	+	cNMP	TM	CaMBD	IQ	G-4a
	BmiCNGC17	BmiB042067	B3	13684849	13687796	-	cNMP	TM	CaMBD	IQ	
	BniCNGC18	BniB007763	B2	26309605	26314078	-	cNMP	TM	CaMBD	IQ	
	BniCNGC19	BniB005958	scaffold_12.1	1632641	1637241	-	cNMP	TM	CaMBD	IQ	G-4b
	BniCNGC20	BniB007622	B8	26069033	26072004	+	cNMP	TM	CaMBD	IQ	
	BniCNGC21	BniB011666	scaffold_1658.1	6066	10973	-	cNMP	TM			
	BniCNGC22	BniB048831	B5	6244942	6248272	+	cNMP	TM	CaMBD	IQ	G1
	BniCNGC23	BniB034825	B1	26131236	26134582	+	cNMP	TM	CaMBD	IQ	
	BniCNGC24	BniB031080	B3	37228781	37233216	-	cNMP	TM	CaMBD	IQ	
	BmiCNGC25	BmiB035297	B1	26136630	26139878	-	cNMP	TM	CaMBD	IQ	G2
	BniCNGC26	BniB048828	B5	6217722	6220995	+	cNMP	TM	CaMBD	IQ	
	BniCNGC27	BniB048829	B5	6225911	6228914	+	cNMP	TM			
	BoCNGC1	Bo3g054400	3	21344613	21348748	-	cNMP	TM	CaMBD	IQ	G3
	BoCNGC2	Bo4g009240	4	1033437	1036412	-	cNMP	TM	CaMBD	IQ	
	BoCNGC3	Bo4g158880	4	43301442	43304004	+	cNMP	TM	CaMBD	IQ	
	BoCNGC4	Bo8g100190	8	34690267	34693456	+	cNMP	TM	CaMBD	IQ	G4a
	BoCNGC5	Bo4g154120	4	42081151	42,084,249 r	-	cNMP	TM	CaMBD	IQ	
	BoCNGC6	Bo7g114580	7	45555801	45558581	-	cNMP	TM	CaMBD	IQ	
	BoCNGC7	Bo2g028960	2	8713570	8716224	-	cNMP	TM	CaMBD	IQ	G4a
	BoCNGC8	Bo5g021460	5	7641448	7644249	-	cNMP	TM	CaMBD	IQ	
	BoCNGC9	Bo4g071760	4	15338332	15341436	-	cNMP	TM	CaMBD	IQ	
	BoCNGC10	Bo1g013250	1	4524027	4527775	+	cNMP	TM	CaMBD	IQ	G4a
	BoCNGC11	Bo4g154790	4	42536095	42539364	+	cNMP	TM	CaMBD	IQ	
	BoCNGC12	Bo8g099660	8	34411186	34414631	-	cNMP	TM	CaMBD	IQ	
BoCNGC13	Bo9g165600	9	48846499	48849628	+	cNMP	TM	CaMBD	IQ	G4a	
BoCNGC14	Bo8g076590	8	24659628	24662612	-	cNMP	TM	CaMBD	IQ		
BoCNGC15	Bo2g042020	2	11542927	11548523	-	cNMP	TM	CaMBD	IQ		
BoCNGC16	Bo9g116160	9	36922938	36930370	-	cNMP	TM	CaMBD	IQ	G4a	
BoCNGC17	Bo9g164130	9	48537709	48540641	+	cNMP	TM	CaMBD	IQ		

Table S2. Continued).

<i>Brassica oleracea</i>	BoCNGC18	Bo3g070140	3	27392670	27395852	+	cNMP	TM			
	BoCNGC19	Bo3g070160	3	27410345	27412966	+	cNMP	TM			
	BoCNGC20	Bo5g122720	5	37315175	37,318,430 r	-	cNMP	TM	CaMBD		IQ
	BoCNGC21	Bo1g119310	1	35942625	35945970	+	cNMP	TM	CaMBD		IQ
	BoCNGC22	Bo1g119340	1	35975275	35980560	-	cNMP	TM	CaMBD		IQ
	BoCNGC23	Bo1g079060	1	23608397	23614384	+	cNMP	TM	CaMBD		IQ
	BoCNGC24	Bo1g119320	1	35957633	35965016	-	cNMP	TM	CaMBD		IQ
	BoCNGC25	Bo5g122750	5	37334396	37337680	-	cNMP	TM	CaMBD		IQ
	BoCNGC26	Bo5g122740	5	37329472	37332599	-	cNMP	TM	CaMBD		IQ
	CrCNGC01	Carubv10008436m	Sca1	312824	316422	+	cNMP	TM	CaMBD		IQ
	CrCNGC02	Carubv10008437m	Sca1	312609	316422	+	cNMP	TM	CaMBD		IQ
	CrCNGC03	Carubv10000339m	Sca6	16178825	16182334	-	cNMP	TM	CaMBD		IQ
	CrCNGC04	Carubv10022736m	Sca4	14365086	14368588	+	cNMP	TM	CaMBD		IQ
	CrCNGC05	Carubv10025973m	Sca8	7618145	7622674	-	cNMP	TM	CaMBD		IQ
	CrCNGC06	Carubv10008459m	Sca1	5562756	5565247	-	cNMP	TM	CaMBD		IQ
	CrCNGC07	Carubv10027632m	Sca8	9635751	9639172	+	cNMP	TM	CaMBD		IQ
	CrCNGC08	Carubv10022706m	Sca4	1367088	1370176	-	cNMP	TM	CaMBD		IQ
	CrCNGC09	Carubv10007740m	Sca7	3795493	3798246	+	cNMP	TM	CaMBD		IQ
	CrCNGC10	Carubv10022774m	Sca4	6572160	6575069	+	cNMP	TM	CaMBD		IQ
	CrCNGC11	Carubv10007769m	Sca7	3871140	3874169	+	cNMP	TM	CaMBD		IQ
	CrCNGC12	Carubv10022720m	Sca4	1844204	1847258	+	cNMP	TM	CaMBD		IQ
	CrCNGC13	Carubv10018573m	Sca5	7773848	7776521	-	cNMP	TM	CaMBD		IQ
	CrCNGC14	Carubv10000335m	Sca6	4843174	4846281	-	cNMP	TM	CaMBD		IQ
	CrCNGC15	Carubv10000276m	Sca6	5052588	5056384	-	cNMP	TM	CaMBD		IQ
	CrCNGC16	Carubv10026031m	Sca8	8125959	8130725	-	cNMP	TM	CaMBD		IQ
	CrCNGC17	Carubv10015585m	Sca3	6154056	6157358	+	cNMP	TM	CaMBD		IQ
CrCNGC18	Carubv10013007m	Sca3	6158154	6162396	+	cNMP	TM	CaMBD		IQ	
<i>Capsella rubella</i>											

Table S2. Continued).

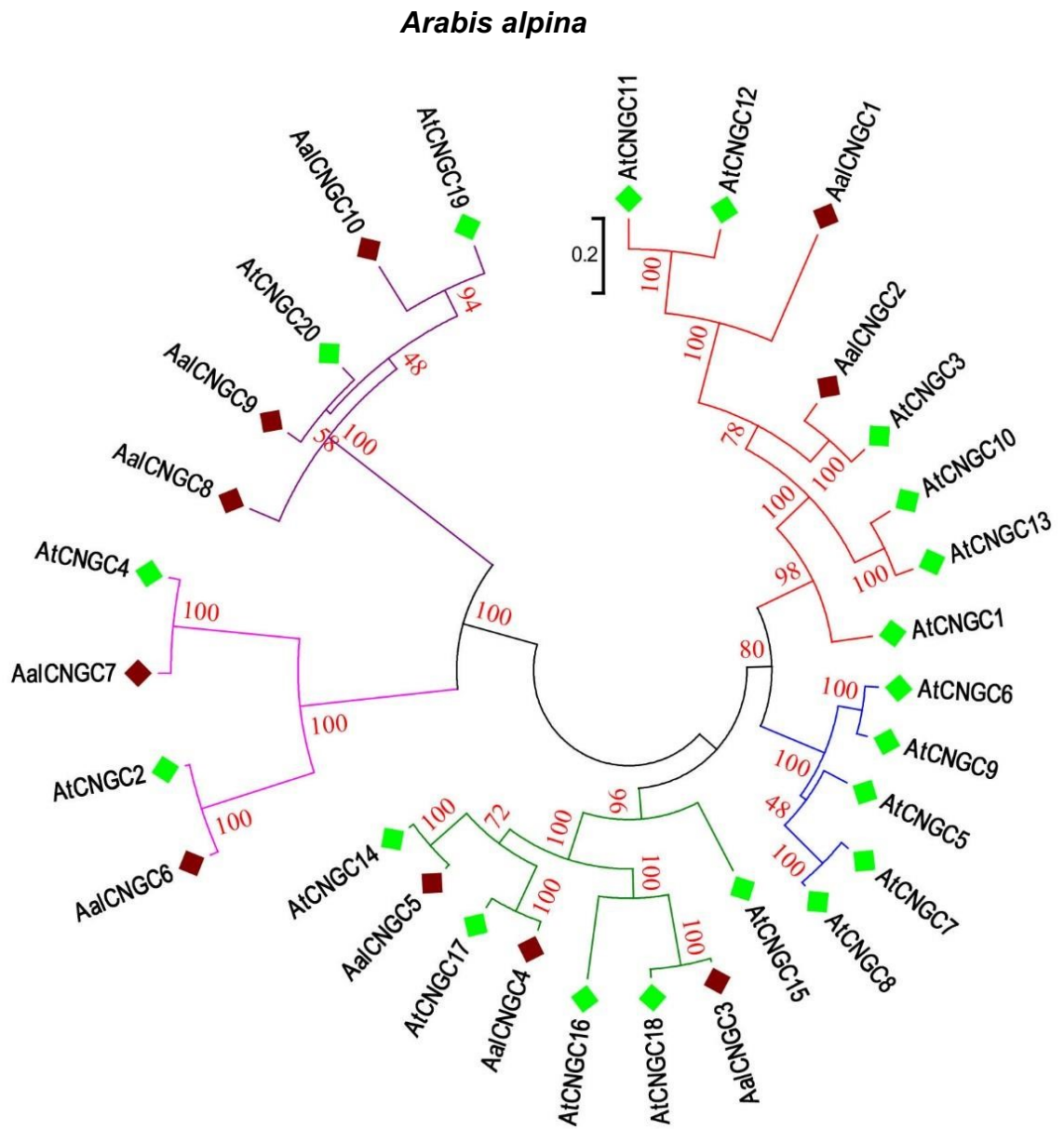
SiCNGC14	SI_scaffold507_5	SI_scaffold507	40042	42658	-	cNMP	TM	CaMBD	IQ	
SiCNGC15	SI_scaffold712_7	SI_scaffold712	37747	40972	+	cNMP	TM	CaMBD	IQ	G3
SiCNGC16	SI_scaffold1205_3	SI_scaffold1205	4209	7024	+	cNMP	TM	CaMBD	IQ	
SiCNGC17	SI_scaffold2067_5	SI_scaffold2067	34816	37383	+	cNMP	TM	CaMBD	IQ	
SiCNGC18	SI_scaffold843_284	SI_scaffold843	1184981	1188195	+	cNMP	TM	CaMBD	IQ	G4a
SiCNGC19	SI_scaffold480_5	SI_scaffold480	30332	35062	+	cNMP	TM	CaMBD	IQ	
SiCNGC20	SI_scaffold2615_4	SI_scaffold2615	15391	18968	+	cNMP	TM	CaMBD	IQ	
SiCNGC21	SI_scaffold2036_23	SI_scaffold2036	103485	106861	+	cNMP	TM	CaMBD	IQ	G4b
SiCNGC22	SI_scaffold2036_22	SI_scaffold2036	98647	102000	+	cNMP	TM	CaMBD	IQ	
SpCNGC1	c0456_00003	Scaffold456	6725	9185	-	cNMP	TM	CaMBD	IQ	
SpCNGC2	c0009_00191	Sp4	2427440	2429900	+	cNMP	TM	CaMBD	IQ	
SpCNGC3	c0046_00029	Scaffold46	108187	111026	-	cNMP	TM	CaMBD	IQ	
SpCNGC4	c0004_00160	Sp4	6579579	6582418	-	cNMP	TM	CaMBD	IQ	G1
SpCNGC5	c0059_00018	Scaffold59	55894	58729	+	cNMP	TM	CaMBD	IQ	
SpCNGC6	c0011_00105	Sp6	402183	405018	+	cNMP	TM	CaMBD	IQ	
SpCNGC7	c0002_00247	Sp6	8172991	8175697	-	cNMP	TM	CaMBD	IQ	
SpCNGC8	c0013_00631	Sp5	15785438	15788253	+	cNMP	TM	CaMBD	IQ	
SpCNGC9	c0002_00648	Sp6	10133985	10137438	+	cNMP	TM	CaMBD	IQ	G2
SpCNGC10	c0009_00480	Sp4	4167071	4169846	+	cNMP	TM	CaMBD	IQ	
SpCNGC11	c0014_00262	Sp7	13528716	13531437	-	cNMP	TM	CaMBD	IQ	
SpCNGC12	c0004_01854	Sp4	13625024	13627715	-	cNMP	TM	CaMBD	IQ	
SpCNGC13	c0009_00408	Sp4	3805721	3808996	-	cNMP	TM	CaMBD	IQ	G3
SpCNGC14	c0014_00241	Sp7	13460292	13463330	-	cNMP	TM	CaMBD	IQ	
SpCNGC15	c0002_01513	Sp6	13777855	13780888	+	cNMP	TM	CaMBD	IQ	
SpCNGC16	c0050_00017	Scaffold50	89997	92872	+	cNMP	TM	CaMBD	IQ	
SpCNGC17	c0007_00264	Sp5	1443823	1446698	-	cNMP	TM	CaMBD	IQ	G4a
SpCNGC18	c0002_01468	Sp6	13586209	13589996	+	cNMP	TM	CaMBD	IQ	
SpCNGC19	c0002_00340	Sp6	8681092	8685447	-	cNMP	TM	CaMBD	IQ	
SpCNGC20	c0005_01561	Sp3	5555883	5559240	+	cNMP	TM	CaMBD	IQ	G4b
SpCNGC21	c0108_00013	Scaffold108	37984	41341	+	cNMP	TM	CaMBD	IQ	
SpCNGC22	c0005_01560	Sp3	5551448	5554858	+	cNMP	TM	CaMBD	IQ	
SpCNGC23	c0108_00012	Scaffold108	33549	36959	+	cNMP	TM	CaMBD	IQ	
ThCNGC1	Thhalv10018258m	scaffold_9	7016987	7019617	-	cNMP	TM	CaMBD	IQ	
ThCNGC2	Thhalv10001345m	scaffold_22	1217304	1220405	+	cNMP	TM	CaMBD	IQ	
ThCNGC3	Thhalv10001335m	scaffold_22	1213484	1217048	+	cNMP	TM	CaMBD	IQ	G1
ThCNGC4	Thhalv10028473m	scaffold_3	474543	477850	+	cNMP	TM	CaMBD	IQ	
ThCNGC5	Thhalv10012801m	scaffold_2	12004747	12007748	+	cNMP	TM	CaMBD	IQ	
<i>Sisymbrium trio (London rocket)</i>										
<i>Thellungiella halophila</i>										

Table S2. Continued).

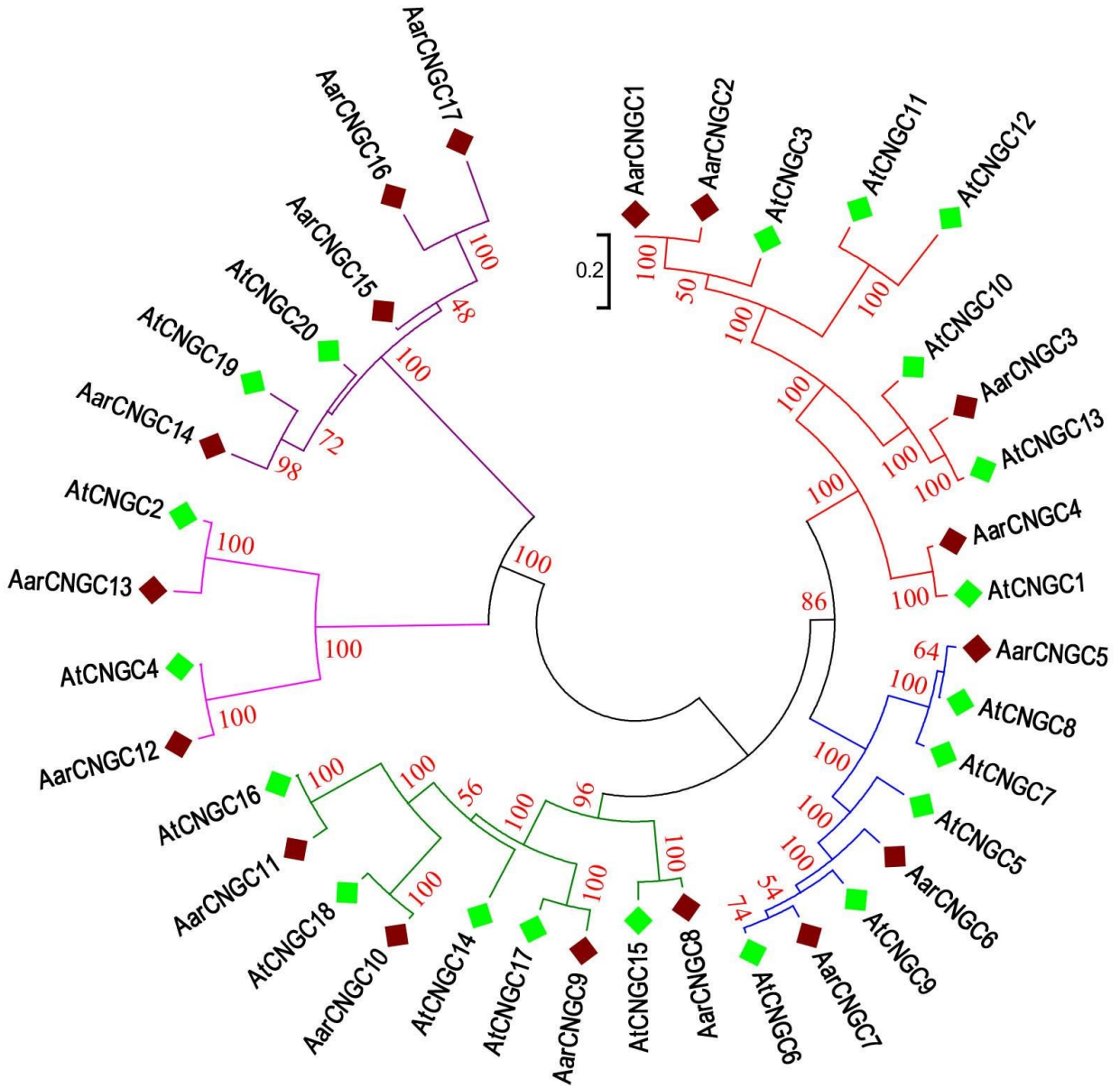
BrCNGC1	Bra034281	A04	11980216	11982791	+	cNMP	TM	CaMBD	IQ	G1
BrCNGC2	Bra003323	A07	15879616	15883454	-	cNMP	TM		IQ	
BrCNGC3	Bra004537	A05	687357	690331	-	cNMP	TM		IQ	
BrCNGC4	Bra031515	A01	16651616	16656087	+	cNMP	TM		IQ	
BrCNGC5	Bra000937	A03	14054247	14058116	-	cNMP	TM	CaMBD	IQ	G2
BrCNGC6	Bra003081	A10	5414086	5416746	-	cNMP	TM	CaMBD	IQ	
BrCNGC7	Bra022632	A02	7390572	7393211	+	cNMP	TM	CaMBD	IQ	
BrCNGC8	Bra026086	A06	5904523	5907153	-	cNMP	TM	CaMBD	IQ	
BrCNGC9	Bra020402	A02	5537255	5540170	-	cNMP	TM	CaMBD	IQ	G3
BrCNGC10	Bra032132	A04	11074762	11077889	-	cNMP	TM	CaMBD	IQ	
BrCNGC11	Bra039221	A09	32929402	32932962	+	cNMP	TM	CaMBD	IQ	
BrCNGC12	Bra024067	A03	27904482	27907069	-	cNMP	TM	CaMBD	IQ	
BrCNGC13	Bra011963	A07	13141306	13144346	+	cNMP	TM	CaMBD	IQ	G4a
BrCNGC14	Bra008733	A10	12426314	12429518	+	cNMP	TM	CaMBD	IQ	
BrCNGC15	Bra018089	A06	9846882	9849809	+	cNMP	TM	CaMBD	IQ	
BrCNGC16	Bra011186	A01	3422819	3426535	+	cNMP	TM	CaMBD	IQ	
BrCNGC17	Bra007839	A09	32710513	32713938	-	cNMP	TM	CaMBD	IQ	G4b
BrCNGC18	Bra032081	A04	11383157	11386389	+	cNMP	TM	CaMBD	IQ	
BrCNGC19	Bra022702	A02	6903420	6907955	+	cNMP	TM	CaMBD	IQ	
BrCNGC20	Bra003001	A10	6203509	6208673	-	cNMP	TM	CaMBD	IQ	
BrCNGC21	Bra008699	A10	12252329	12255245	+	cNMP	TM	CaMBD	IQ	G4b
BrCNGC22	Bra001678	A03	17843897	17852047	+	cNMP	TM			
BrCNGC23	Bra031529	A01	16542495	16546796	-	cNMP	TM			
BrCNGC24	Bra029958	A01	14741686	14746906	+	cNMP	TM	CaMBD	IQ	
BrCNGC25	Bra021265	A01	22108307	22111729	-	cNMP	TM	CaMBD	IQ	G4b
BrCNGC26	Bra022233	A05	19633895	19637029	-	cNMP	TM	CaMBD	IQ	
BrCNGC27	Bra001676	A03	17833555	17836622	+	cNMP	TM			
BrCNGC28	Bra021266	A01	22102703	22106050	+	cNMP	TM	CaMBD	IQ	
BrCNGC29	Bra022232	A05	19638792	19642102	-	cNMP	TM	CaMBD	IQ	

Brassica rapa

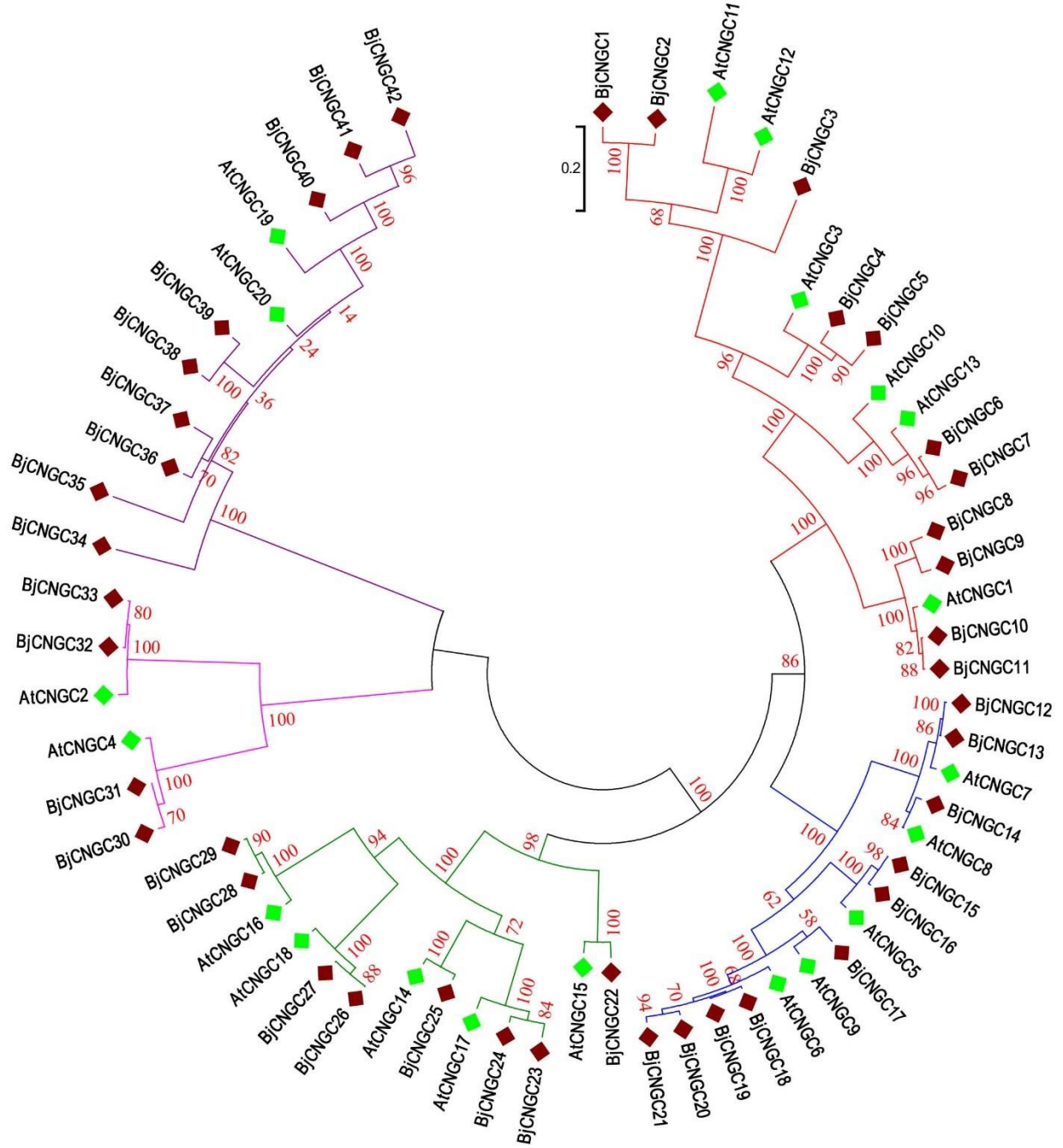
Figure S1. Phylogenies of CNGC families from individual crucifer species with 20 AtCNGCs



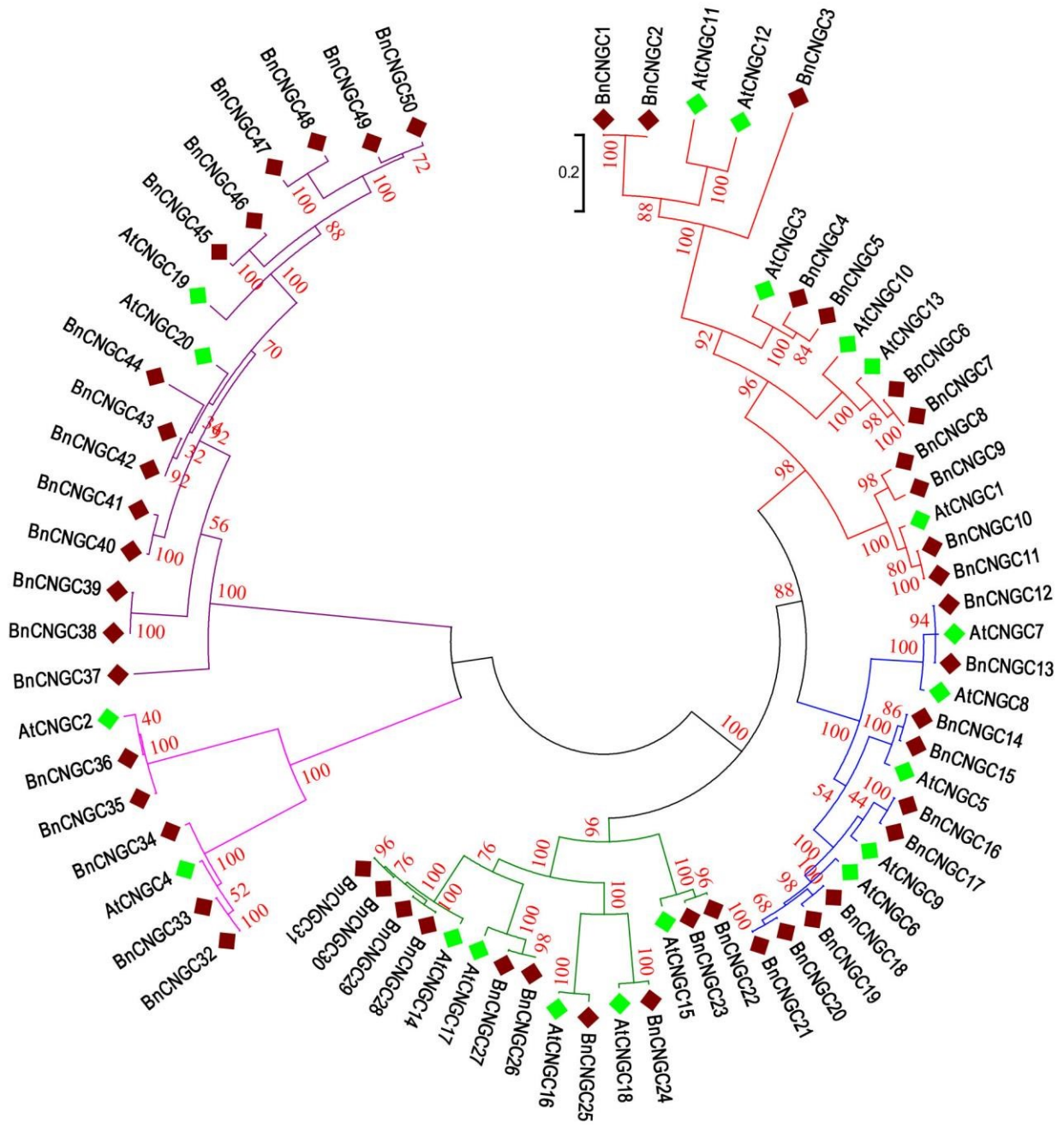
Aethionema arabicum



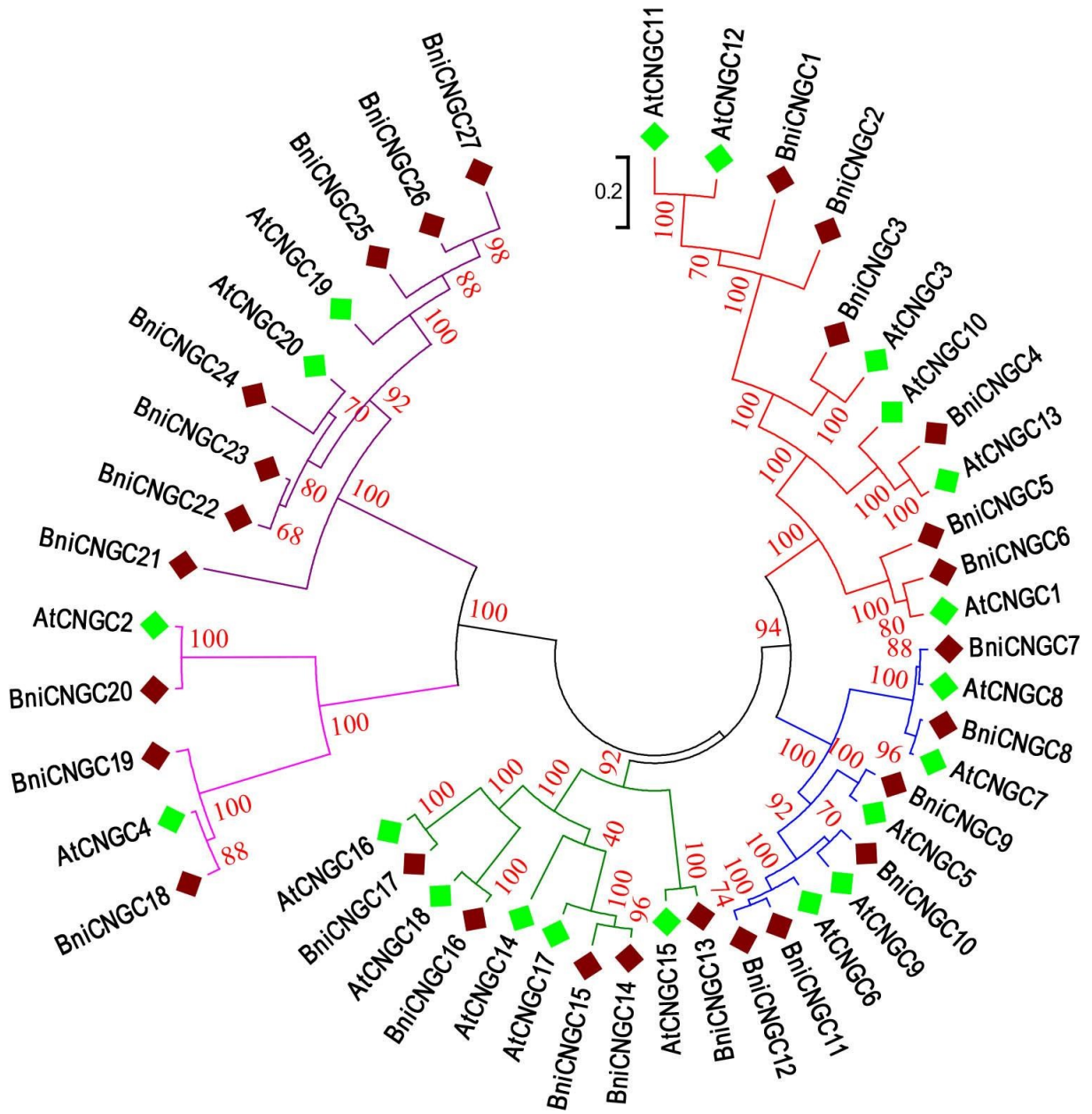
Brassica juncea



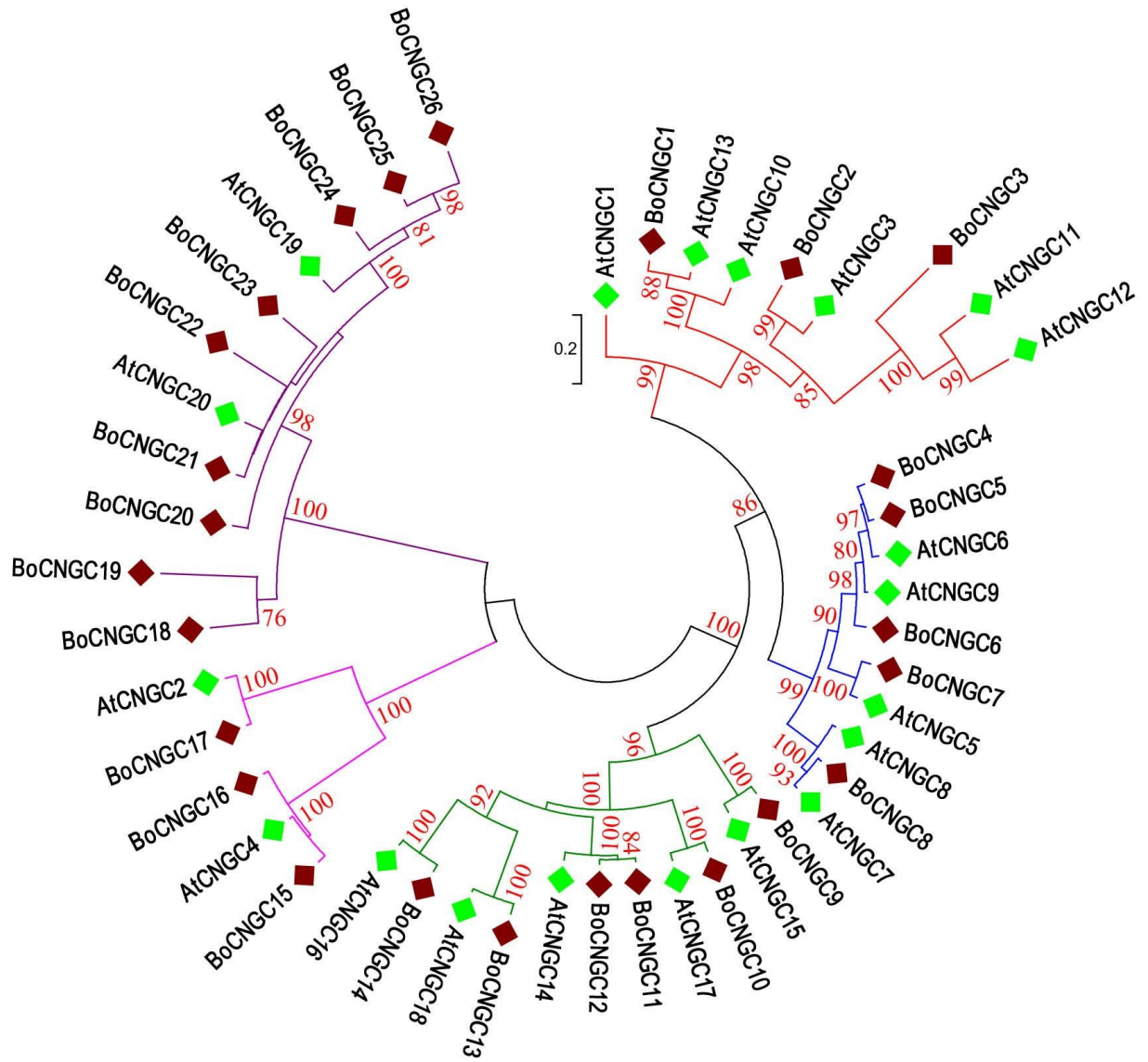
Brassica napus



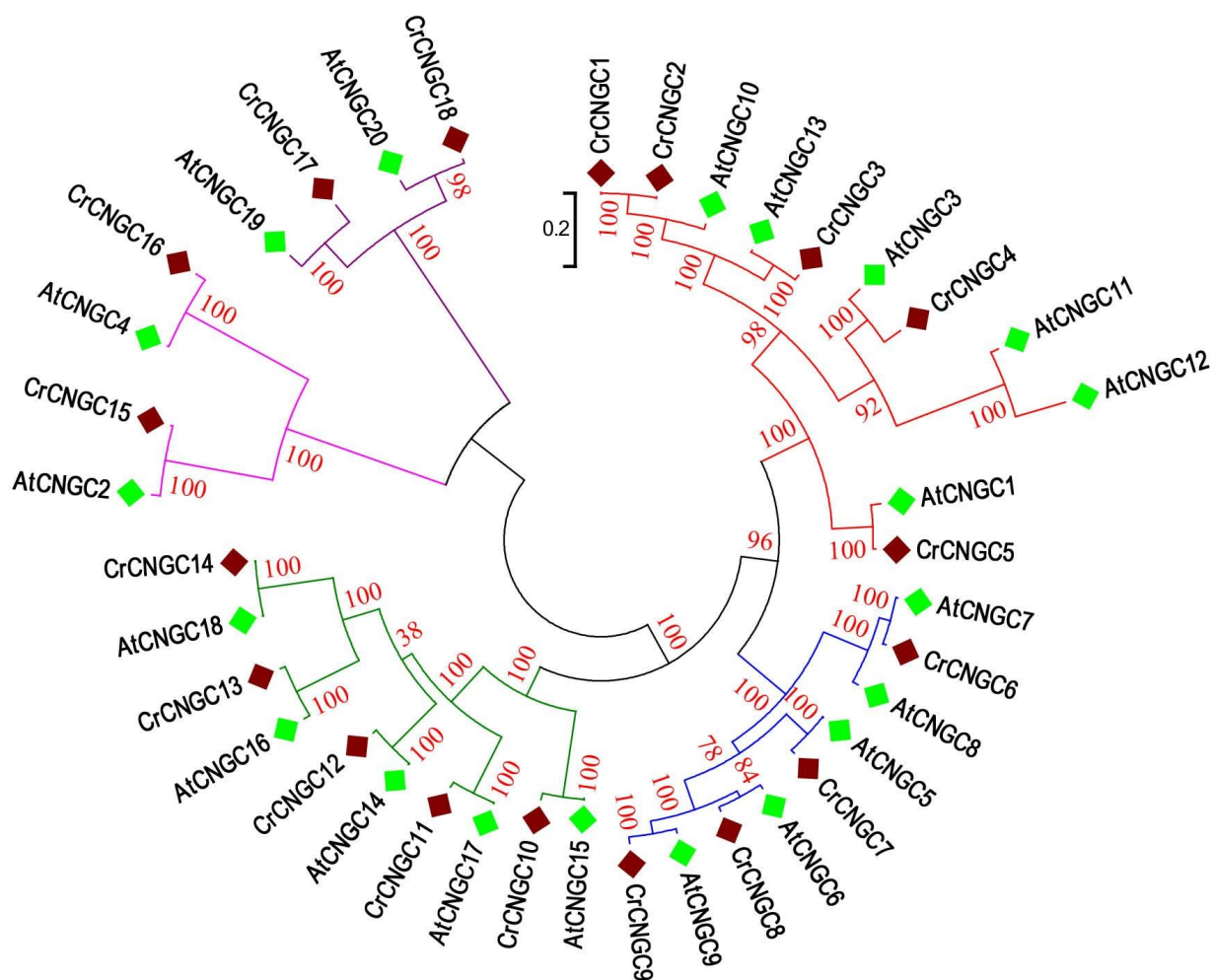
Brassica nigra



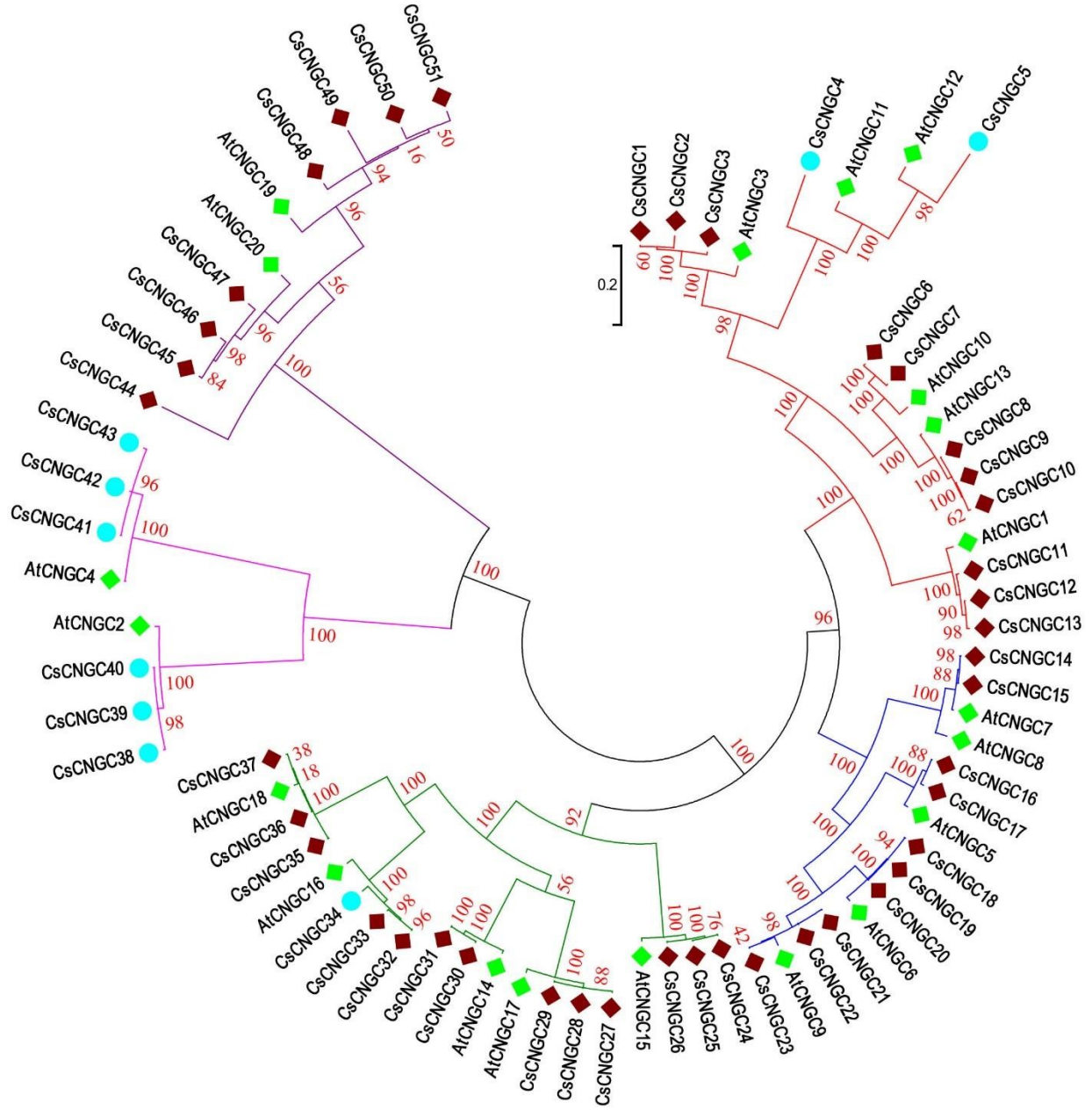
Brassica oleracea



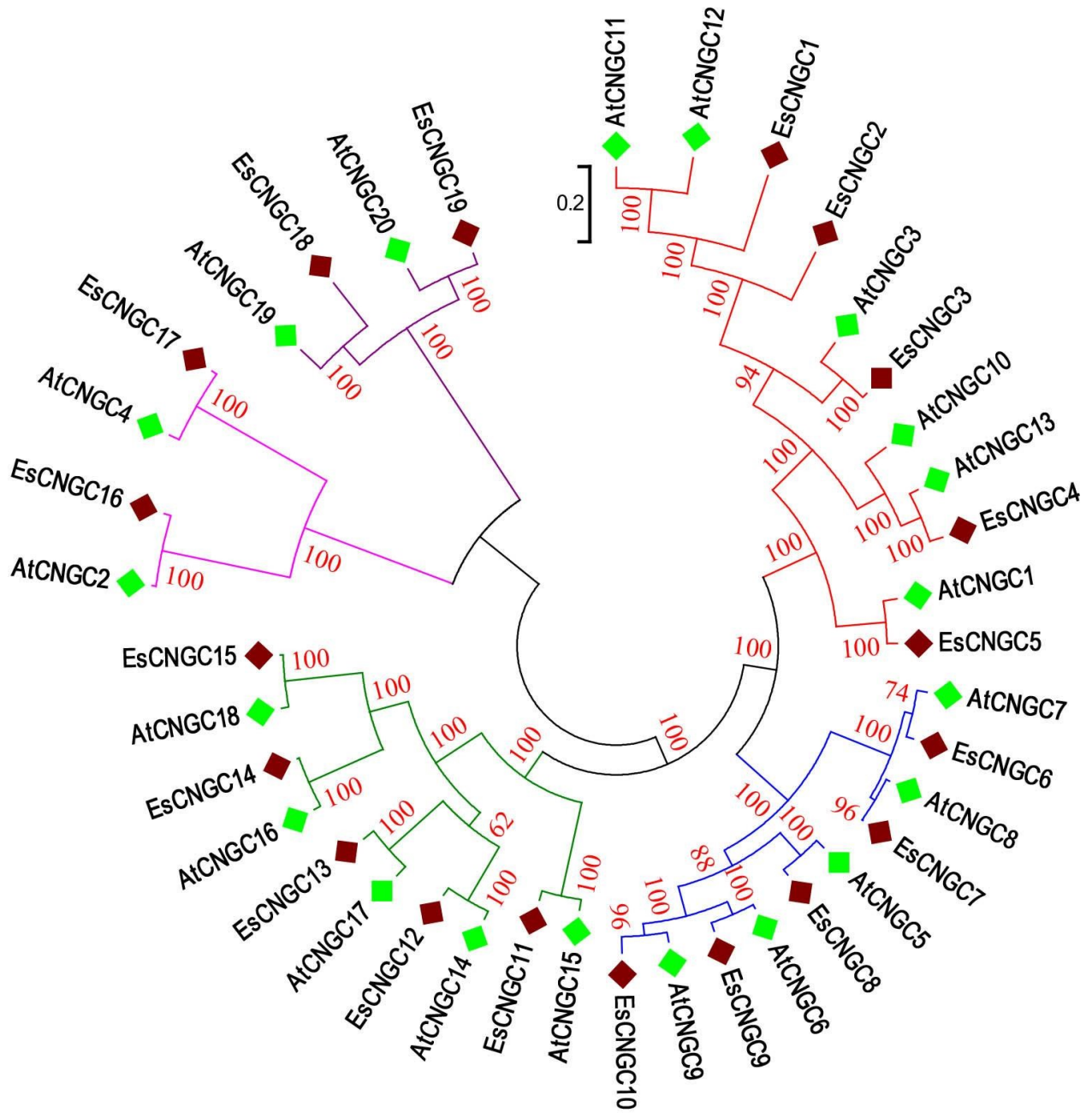
Capsella rubella



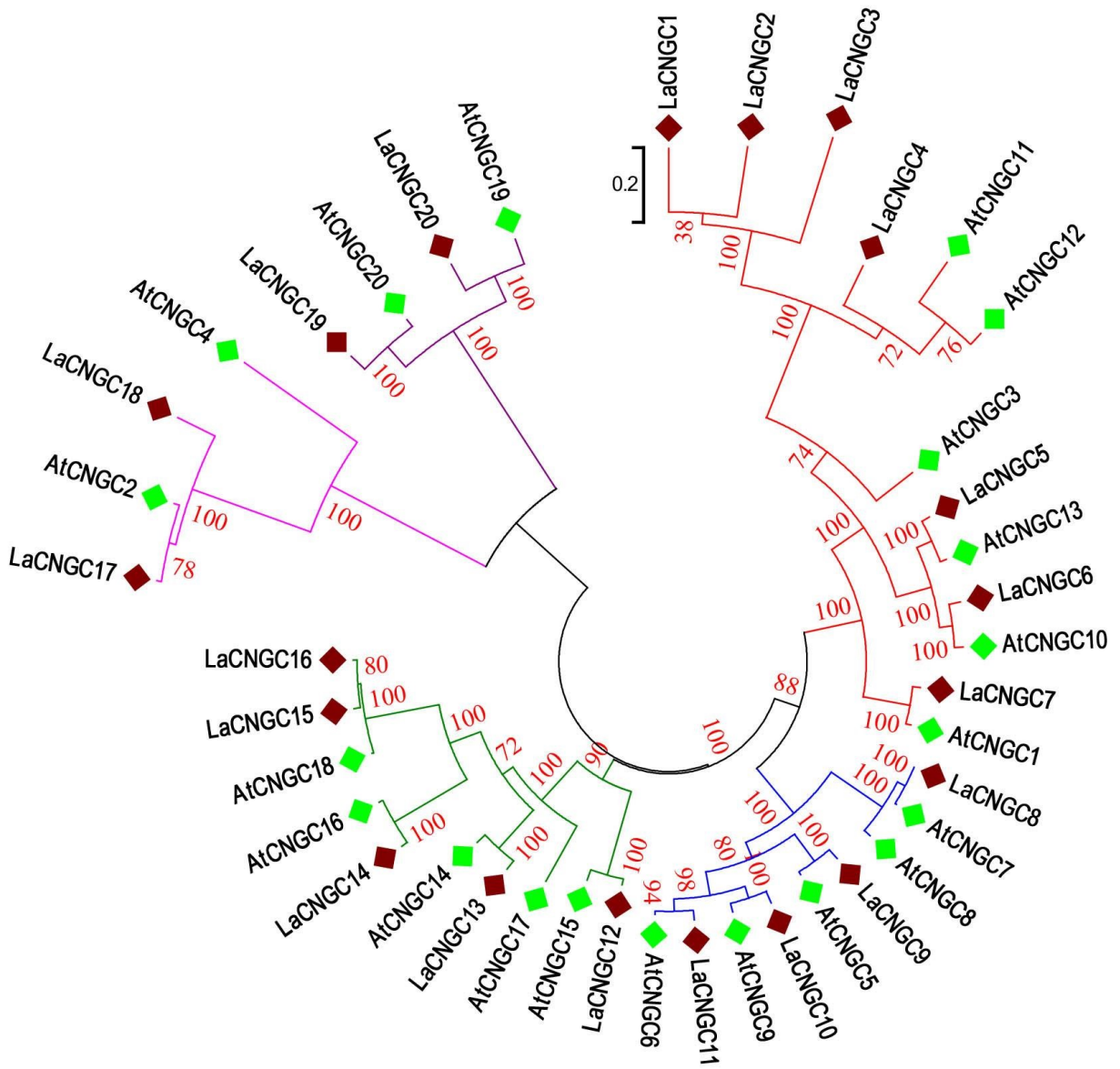
Camelina sativa



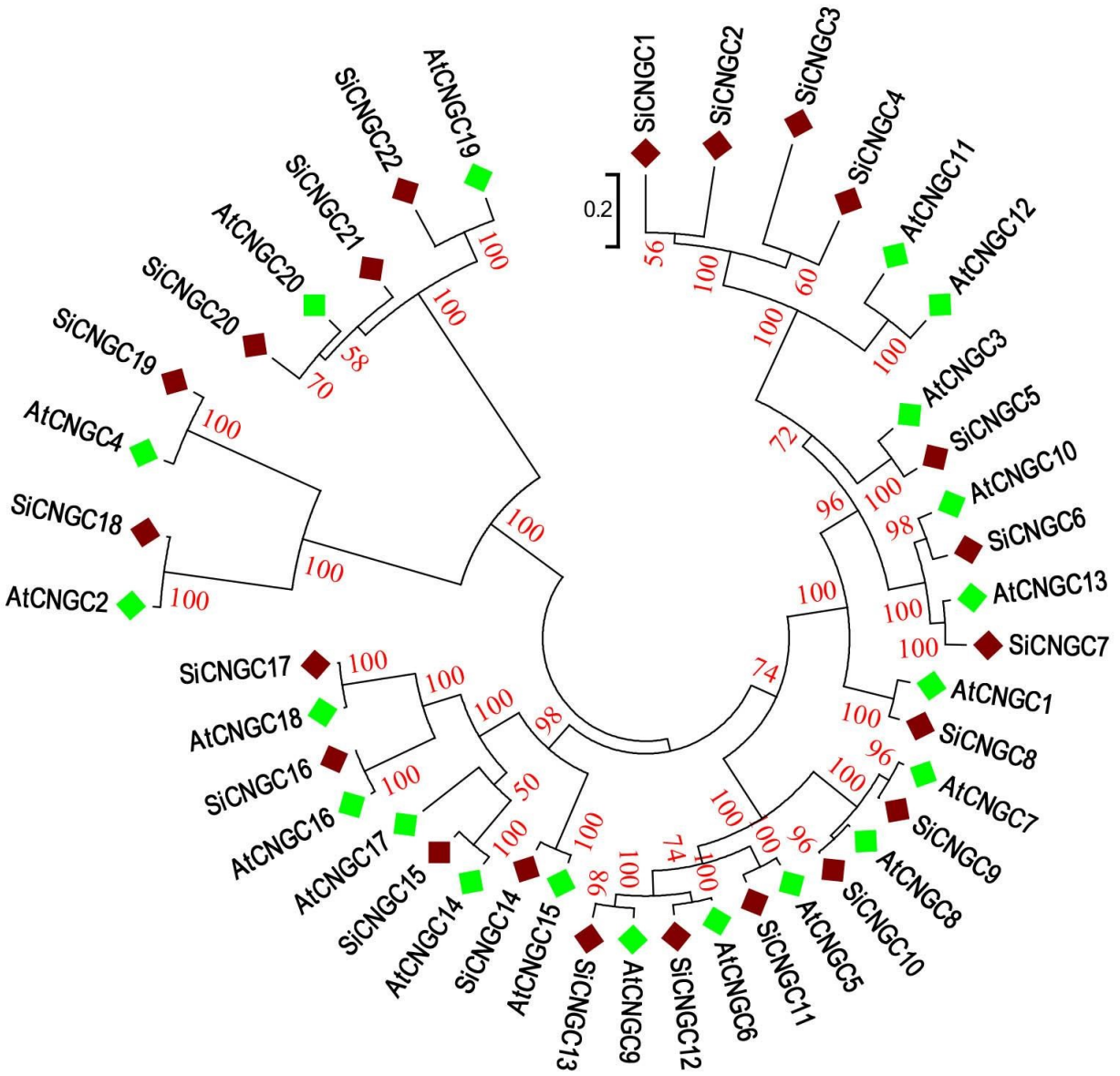
Eutrema salsaugineum



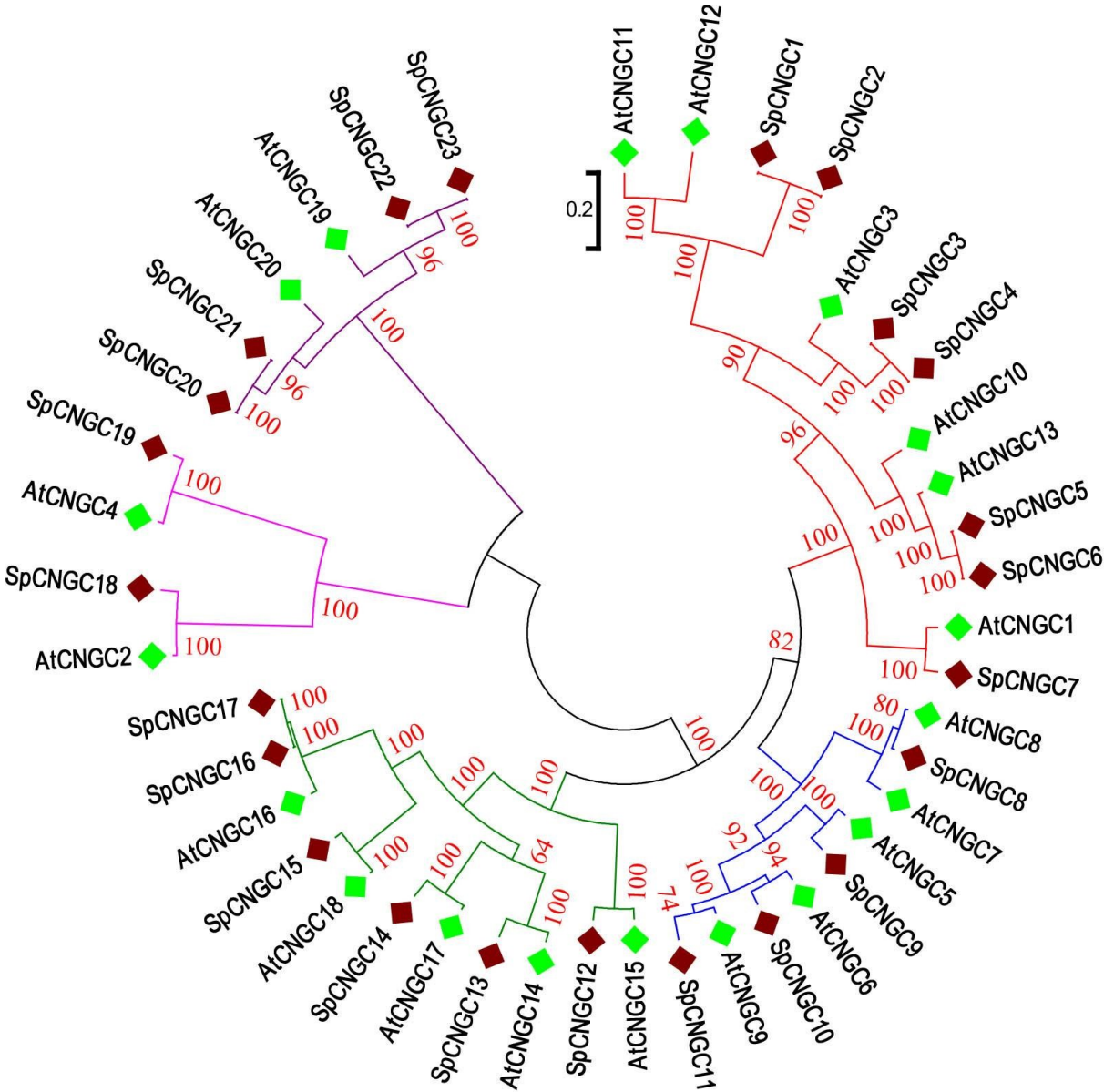
Leavenworthia alabamica



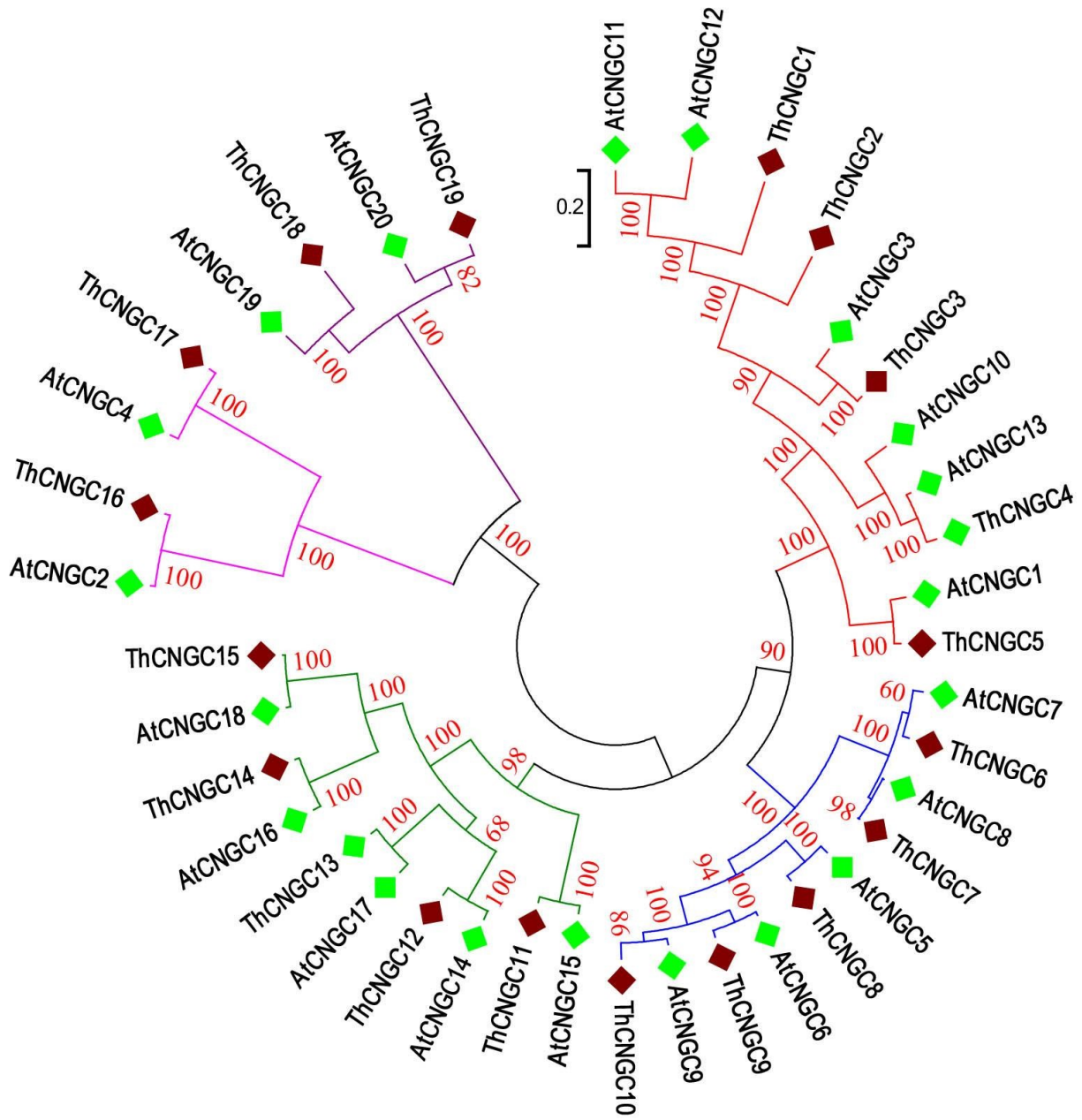
Sisymbrium irio



Schrenkiella parvula



Thellungiella halophila



Thellungiella salsuginea

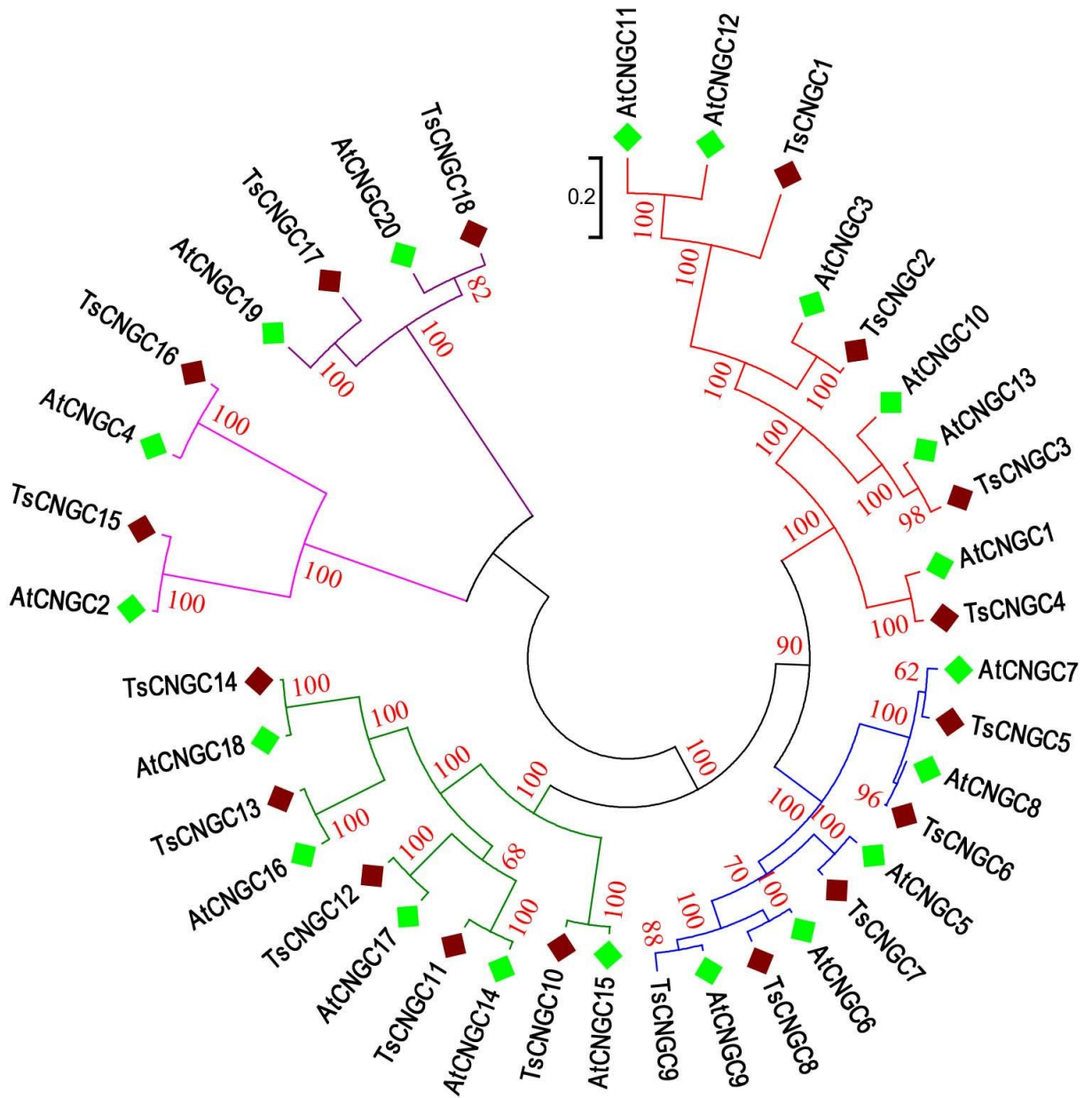
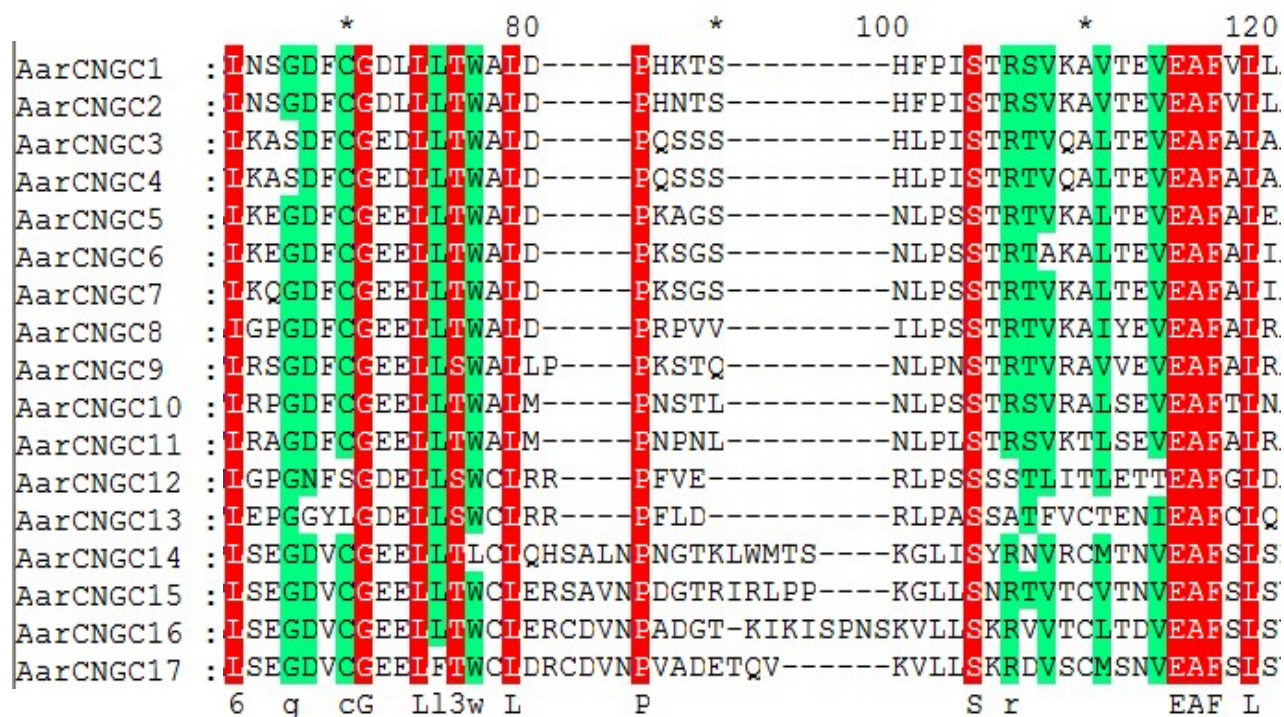


Figure S2. Alignments showing conserved consensus motifs of PBC and hinge regions of CNBD domains from each Brassicaceae species




```

      *           80           *           100           *           120
AlCNGC1 : LQAGDIYGE-LIFN-----GSRLPT---STRTIMTLTEVEGFILLP:
AlCNGC2 : LEAGDICGE-LIFN-----GSRLPT---STRVTMTMTEVEGFILLP:
AlCNGC3 : LKAGDICGE-LIFN-----SSSLPT---STRTIITLTEVEGFITLSP:
AlCNGC4 : LKAGDICGQ-LIFN-----SSCLPT---STRTIITLTEVEGFITLSP:
AlCNGC5 : LVAGDFCGD-LITWALDPL-----SSNFPI---SSRTVQALTEVEGFVLSA:
AlCNGC6 : LKPSDFCGEDLITWALDPQS-----SSHFPPI---STRTVQALTEVEAFALIAA:
AlCNGC7 : LKASDFCGEDLIPWALDPQS-----SSHFPPI---STRTVQALTEVEAFALTA:
AlCNGC8 : LGAGDFCGEELITWALDPHT-----SSNLPPI---STRTVRAIMEVEAFALKA:
AlCNGC9 : LKEGDFCGEELITWALDPKS-----GSNLPS---STRTAKALTEVEAFALIA:
AlCNGC10 : LKEGDFCGDELLITWALDPKS-----GSNLPS---STRTVKALTEVEAFALIA:
AlCNGC11 : LKEGEFCGEELITWALDPKS-----GVNLPS---STRTVKALTEVEAFALTS:
AlCNGC12 : LKEGDFCGEELITWALDPKA-----GSNLPS---STRTVKALTEVEAFALIA:
AlCNGC13 : LKEGDFCGEELITWALDPKA-----GSNLPS---STRTVKALTEVEAFALIA:
AlCNGC14 : LKPGDFCGEELITWALDPRP-----VVILPS---STRTVKALCEVEAFALKA:
AlCNGC15 : LKPGDFCGEELITWALVPNI-----NHNLPL---STRTVRTLSEVEAFALRA:
AlCNGC16 : LKPGDFCGEELITWALMPNS-----TLNLPS---STRSVRALSEVEAFALSA:
AlCNGC17 : LKPGDFCGEELISWALLPKS-----TLNLPS---STRTVRALVEVEAFALRA:
AlCNGC18 : LKPGDFCGEELIAWALLPKS-----TVNLPS---STRTVRALEEVEAFALQA:
AlCNGC19 : LEPGGYLGDELLSWCLRRPF-----LDRLPP---SSATFVCLDNIIEAFSLGC:
AlCNGC20 : LGPGNFSGDELLSWCLRRPF-----VERLPP---SSSTLVTLTETTEAFGLIA:
AlCNGC21 : LSEGDVCGEELITWCLKRFVSNPDGTRIKMPSKGLVSNRNVRCVTNVEAFSLSI:
AlCNGC22 : LYEGDVCGEELITWCLERSVSNPDGTRIRMPKGLLSNRNVRCVTNVEAFSLSV:
AlCNGC23 : LSEGDVCGEELITWCLERSINPDGTRIKMPSKGLVSNRNVRCVTNVEAFSLSV:
      6  g  G  LL                P  S  r  6  E  F  L

```

```

      *           80           *           100           *
AalCNGC1 : LGECAICGDHLLSWVLDPHSSS-----RLPT---SDRTIRTVSNVEGFIFL:
AalCNGC2 : LSAGGFCDGLLTWALDPHAS-----HFPI---SSRTVQAQTEVEGFVLS:
AalCNGC3 : LKPGDFCGEELITWALMPNSTL-----NLPS---STRSVRALSEVEAFALS:
AalCNGC4 : LKPGDFCGEELISWALLPKSTL-----NLPS---STRTVRALVEVEAFALH:
AalCNGC5 : LKPGDFCGEELIAWALLPKSTV-----NLPS---STRTVRALEEVEAFALQ:
AalCNGC6 : LEPGGYLGDELLSWCLRRPFID-----RLPP---SSATFVCLDNIIEAFSLC:
AalCNGC7 : LGPGNFSGDELLSWCLRRPFVE-----RLPP---SSSTLVTLTETTEAFGLI:
AalCNGC8 : LSEGDVCGEELITWCLERSAVNPHGTTRMPKRLLSNRNVRCVTNVEAFSLS:
AalCNGC9 : LSEGDVCGEELITWCLERSVNSDGTTRIRMPKGLLSNRNVRCVTNVEAFSLC:
AalCNGC10 : LSEGDVCGEELITWCLERSINPDGRRIKMLSKRLVSNRNVRCVTNVEAFSLS:
      L  G  cG eLL W L                p  S  r                EaF l

```



```

      *           80           *           100           *           1
BnCNGC1 : TQAGDICG-DLIFWVLDPHSSS-----SLPT----SDRSVLTITDVEGFIL
BnCNGC2 : TQAGDICG-DLIFWVLDPHSSS-----SLPT----SDRSVLTITDVEGFIL
BnCNGC3 : LMKGDV-WEDLIFWALDPHTSP-----SLPI----SNGTVTTLTIVVEGLTL
BnCNGC4 : LGVGQFCG-DLLTWALDPNTS-----HFPI----STSTVQAQTEVEGFVL
BnCNGC5 : LEAGDFCG-DLLTWALDPNTS-----HLPI----STRTVQAETEVEGFVL
BnCNGC6 : LKASDFCGEDLLTWALDPQSSS-----HFPI----SPRTVQALTEVEAFAL
BnCNGC7 : LNASDFCGEDLLTWALDPQSSS-----HFPI----SPRTVQALTEVEAFAL
BnCNGC8 : LGAGDFCGEELLTWALDPHTST-----NLPI----STRTVQALVEVEAFAL
BnCNGC9 : LGAGDFCGEELLTWALDPHTST-----NLPI----STRTVQALVEVEAFAL
BnCNGC10 : LGAGDFCGEELLTWALDPHTSS-----NLPI----STRTVRALVEVEAFAL
BnCNGC11 : LGAGDFCGEELLTWALDPHTSS-----NLPI----STRTVRALVEVEAFAL
BnCNGC12 : LKEGDFCGEELLTWALDPKAGS-----NLPS----STRTVKALTEVEAFAL
BnCNGC13 : LKEGDFCGEELLTWALDPKAGS-----NLPS----STRTVKALTEVEAFAL
BnCNGC14 : LKEGEFCGEELLTWALDPKSGA-----NLPS----STRTVKALTEVEAFAL
BnCNGC15 : LKEGEFCGEELLTWALDPKSGV-----NLPS----STRTVKALTEVEAFAL
BnCNGC16 : LKEGDFCGEELLTWALDPKSGS-----NLPS----STRTAKALTEVEAFAL
BnCNGC17 : LKEGDFCGEELLTWALDPKSGS-----NLPS----STRTAKALTEVEAFAL
BnCNGC18 : LKEGDFCGDELLTWALDPKSGS-----NLPS----STRTVKALTEVEAFAL
BnCNGC19 : LKEGDFCGDELLTWALDPKSGS-----NLPS----STRTVKALTEVEAFAL
BnCNGC20 : LKEGDFCGDELLTWALDPKSGS-----NLPS----STRTVKALTEVEAFAL
BnCNGC21 : LKEGDFCGDELLTWALDPKSGS-----NLPS----STRTVKALTEVEAFAL
BnCNGC22 : LGPGDFCGEELLTWALDPRPVV-----ILPS----STRTVKAIYEVEAFAL
BnCNGC23 : LGPGDFCGEELLTWALDPRPVV-----ILPS----STRTVKAIYEVEAFAL
BnCNGC24 : LRPGDFCGEELLTWALMPNSTL-----NFPS----STRSVRALSEVEAFAL
BnCNGC25 : LRPGDFCGEELLTWALMPNINQ-----NLPL----STRTVRTLSEVEAFAL
BnCNGC26 : LRPGDFCGEELLSWALLPKSTL-----NLPS----STRTVRALVEVEAFAL
BnCNGC27 : LRPGDFCGEELLSWALLPKSTL-----NLPS----STRTVRALVEVEAFAL
BnCNGC28 : LKPGEFCGEELLSWALLPKSKV-----NLPS----STRTVRALEEVEAFAL
BnCNGC29 : LKPGEFCGEELLSWALLPKSKV-----NLPS----STRTVRALEEVEAFAL
BnCNGC30 : LKPGEFCGEELLSWALLPKSKV-----NLPS----STRTVRALEEVEAFAL
BnCNGC31 : LKPGEFCGEELLSWALLPKSKV-----NLPS----STRTVRALEEVEAFAL
BnCNGC32 : LGPGNFSGDELLSWCLRRPFVE-----RLPP----STSTLVTLTETTEAFGL
BnCNGC33 : LGPGNFSGDELLSWCLRRPFVE-----RLPP----STSTLVTLTETTEAFGL
BnCNGC34 : LGPGNFSGDELLSWCLRRPFVE-----RLPP----SSSTLVTLTETTEAFGL
BnCNGC35 : LEPGGYLGDELLSWCLRRPFID-----RLPP----SSATFVCLDNIEAFSL
BnCNGC36 : LEPGGYLGDELLSWCLRRPFID-----RLPP----SSATFVCLDNIEAFSL
BnCNGC37 : LSKRDVYGEELLTW-YERSSLNPDGTRIRMPKSKGLLSNRDVKCVTNVEAFSL
BnCNGC38 : LSEGDCVCGEELLTWQQRSSVNPDGTMIRMPKSKGLLSNRDVCVTVNVEAFSL
BnCNGC39 : LSEGDCVCGEELLTWQQRSSVNPDGTMIRMPKSKGLLSNRDVCVTVNVEAFSL
BnCNGC40 : LSEGDCVCGEELLTWCLERSVNPDGTRIRMPKSKGLLSNRNVRCVTVNVEAFSL
BnCNGC41 : LSEGDCVCGEELLTWCLERSVNPDGTRIRMPKSKGPLSNRNKVCVTVNVEAFSL
BnCNGC42 : LSEGDCVCGEELLTWCLERSVNPDGTRIRIPKSKGLLSYRNKVCVTVNVEAFSL
BnCNGC43 : LSEGDCVCGEELLTWCLERSVNPDGTRIRIPKSKGLLSYRNKVCVTVNVEAFSL
BnCNGC44 : LSEGDCVCGEELLTWCLERSAVNLDG---RMLSKGLLSNRNVKCVTVNVEAFSL
BnCNGC45 : LSEGDFVGEELLTWCLERSVNPDGRSIKLPLKGLVSNRSVRCVTVNVEAFSL
BnCNGC46 : LSEGDCVCGEELLTWCLERSVNPDGRSIKLPLKGLVSNRSVRCVTVNVEAFSL
BnCNGC47 : LSEGDCVCGEELLTWCLERSVDPDGTRIKMPTKGLVSNRSVRCVTVNVEAFSL
BnCNGC48 : LSEGDCVCGEELLTWCLERSVDPDGTRIKMPTKGLVSNRSVRCVTVNVEAFSL
BnCNGC49 : LSEGDCVCGEELLTWCLERSVNPDGTRIKMPTKGLVSNRNVCVTVNVEAFSL
BnCNGC50 : LSEGDCVCGEELLTWCLERSVNPDGTRIKMPTKGLVSNRNVCVTVNVEAFSL
      6  q   q  LL W l           p   S           Eaf L

```



```

      *           80           *           100           *
BniCNGC1 : :IQAGDICGD-LIFWVLDPHSSS-----SLPT---SDRTVLTITDVEGFIL:
BniCNGC2 : :IKAGDICGD-LIFWVLDPHSP-----RLPT---SSITVMTITDVEGFIL:
BniCNGC3 : :IGAGQFCGD-LITWALDPNTS-----HFPI---STSTVQAQTEVEGFVL:
BniCNGC4 : :INANDFCGEDLLTWALDPQSSS-----HFPI---SPRTVQALTEVEAFAL:
BniCNGC5 : :IGAGDFCGEELLTWALDPHTST-----NLPI---STRTVQALMEVEAFAL:
BniCNGC6 : :IGAGDFCGEELLTWALDPHTSS-----NLPI---STRTVRALVEVEAFAL:
BniCNGC7 : :LKEGDFCGEELLTWALDPKAGS-----NLPS---STRTVKALTEVEAFAL:
BniCNGC8 : :LKEGDFCGEELLTWALDPKAGS-----NLPS---STRTVKALTEVEAFAL:
BniCNGC9 : :LKEGEFCGEELLTWALDPKSGV-----NLPS---STRTVKALTEVEAFAL:
BniCNGC10 : :LKEGDFCGEELLTWALDPKSGS-----NLPS---STRTAKALTEVEAFAL:
BniCNGC11 : :LKEGDFCGDELLTWALDPKSGS-----NLPS---STRTVKALTEVEAFAL:
BniCNGC12 : :LKEGDFCGDELLTWALDPKSGS-----NLPS---STRTVKALTEVEAFAL:
BniCNGC13 : :IGPGDFCGEELLTWALDPRPVV-----ILPS---STRTVKAIYEVEAFAL:
BniCNGC14 : :LRPGDFCGEELLWALLPKSTL-----NLPS---STRTVRALVEVEAFAL:
BniCNGC15 : :LRPGDFCGEELLWALLPKSTL-----NLPS---STRTVKASVEVEAFAL:
BniCNGC16 : :LRPGDFCGEELLTWALMPNSTL-----NFPS---STRSVRALSEVEAFAL:
BniCNGC17 : :LRPGDFCGEELLTWALMPNINL-----NLPL---STRTVRTLSEVEAFAL:
BniCNGC18 : :IGPGNFSGDELLSWCLRRPFVE-----RLPP---SSSTLVTLTETTEAFGL:
BniCNGC19 : :IGPGNFSGDELLSWCLRRPFVE-----RLPP---SSSTLVTLTETTEAFGL:
BniCNGC20 : :LEPGGYLGDDELLSWCLRRPFID-----RLPP---SSATFVCLDNIEAFSL:
BniCNGC21 : :LSEGDVCGEELLTWQERSSVNPDGTRIRMPNSGLRSNRDVKCVTDVAAFSL:
BniCNGC22 : :LSEGDVCGEELLTWCLERSSVNPDGTRIKMPTKGLASYRNVTCVTNVEAFSL:
BniCNGC23 : :LSEGDVCGEELLTWCLERSSVNPDGTRIKMPSKGLLSNRNVKCVTNVEAFSL:
BniCNGC24 : :LSEGDVCGEELLTWCLERSAVNPDG---RMPSKGLLSNRNVKCVTNVEAFSL:
BniCNGC25 : :LSEGDVCGEELLTWCLERSSVNPDGRSIKLPLKGLVSNRSVICVTNVEAFSL:
BniCNGC26 : :LSEGDVCGEELLTWCLDRSSVNPDGTRIKMPAKGLVSNRRVRCVTNVEAFSL:
BniCNGC27 : :LSEGDVCGEELLTWCLDRSSVNRDGTTRKMQTKGLVSNRSVRCVTNVEAFSL:
      6  q  cG  LL  W  l           p  S           eaF  L

```

```

      *           80           *           100           *
BoCNGC1 : TNASDFCGEDLITWALDPQS-----SSHFPFI----SPRTVQALTEVEAFAL
BoCNGC2 : TGVGQFCG-DLITWALDPNT-----SHFPFI----STSTVQAQTEVEGFVL
BoCNGC3 : TQAGDICG-DLIFWVLDPHS-----SSSLPT----SDRSVLTITDVEGFIL
BoCNGC4 : TKEGDFCGDELITWALDPKS-----GSNLPS----STRTVKALTEVEAFAL
BoCNGC5 : TKEGDFCGDELITWALDPKS-----GSNLPS----STRTVKALTEVEAFAL
BoCNGC6 : TKEGDFCGEELITWALDPKS-----GSNLPS----STRTAKALTEVEAFAL
BoCNGC7 : TKEGEFCGEELITWALDPKS-----GVNLPS----STRTVKALTEVEAFAL
BoCNGC8 : TKEGDFCGEELITWALDPKA-----GSNLPS----STRTVKALTEVEAFAL
BoCNGC9 : TGP GDFCGEELITWALDPRP-----VVILPS----STRTVKAIYEVEAFAL
BoCNGC10 : TRPGDFCGEELISWALLPKS-----TLNLPS----STRTVRALVEVEAFAL
BoCNGC11 : TKPGEFCGEELIAWALLPKS-----KVNLP S----STRTVRALEEVEAFAL
BoCNGC12 : TKPGEFCGEELIVWALLPKS-----KVNLP S----STRTVRALEEVEAFAL
BoCNGC13 : TRPGDFCGEELITWALMPNS-----TLNFP S----STRSVRALSEVEAFAL
BoCNGC14 : TRPGDFCGEELITWALMPNI-----NQNLPL----STRTVRTLSEVEAFAL
BoCNGC15 : TGP GNFSGDELISWCLRRPF-----VERLPP----SSSTLVLTLETTEAFGL
BoCNGC16 : TGP GNFSGDELISWCLRRPF-----VERLPP----STSTLVLTLETTEAFGL
BoCNGC17 : TEPGGYLGDELISWCLRRPF-----IDRLPP----SSATFVCLDSIEAFSL
BoCNGC18 : TSEG DVCGEELITWCQERS SVNPDGTMIRMPSKGLLSSRDVWCVTNVEAFSL
BoCNGC19 : TSEEFVCGEELITWCQERS SVNPDGTMIRMPSKGLISSRAVWCVTNVEAFSL
BoCNGC20 : TSEG DVCGEELITWCLERS SVNPDGTRIRMPSKGLLSSNRNVRCVTNVEAFSL
BoCNGC21 : TSEG DVCGEELITWCLERS SVNPDGTRIRIPSKGLLSSYRNVRCVTNVEAFSL
BoCNGC22 : TSEG DVCGEELIAWCLGRSSVNPDGTRIRMPSKGLVSSNRNVRCVTDVEAFSL
BoCNGC23 : TSEG DVCGEELITWCLERSAVNL---DGRMLSKGLLSSSRNVKCVTNVEAFSL
BoCNGC24 : TSEG DVFGEELITWCLERS SVNPDGRS IKLPLKGLVSSNRSVRCVTNVEAFSL
BoCNGC25 : TSEG DVCGEELITWCLERS SVNPDGTRIKMPTKGLVSSNRNVRCVTNVEAFSL
BoCNGC26 : TSEQ DVCGEELITWCLERS SVDPDGTRIKMPTKGLVSSNRSVRCVTNVEAFSL
6   g   G eLl W l           p   S           EaF L

```


		*		840	*		860	*					
BrCNGC1	:	LQAGDICG	-DLIFWV	LDPHSS	-----	SLPT	----	SDRSVLTLDVEGFIL					
BrCNGC2	:	LMKGDVW	-EDLIWA	LDPHTSP	-----	SLPI	----	SNGTVTTLTYVEGLTL					
BrCNGC3	:	LEAGDFCG	-DLITWA	LDPNTS	-----	HLPI	----	STRTVQAETEVEGFVL					
BrCNGC4	:	RSVGEFCGEELI	SWALLDPHSS	-----	CVPI	----	SSMIIKARTEVEGFVL						
BrCNGC5	:	LKASDFCGEDLI	TWALDPQSSS	-----	HFPI	----	SPRTVQALTEVEAFAL						
BrCNGC6	:	LGAGDFCGEELI	TWALDPHTST	-----	NLPI	----	STRTVQALVEVEAFAL						
BrCNGC7	:	LGAGDFCGEELI	TWALDPHTSS	-----	NLPI	----	STRTVRALVEVEAFAL						
BrCNGC8	:	LKEGDFCGEELI	TWALDPKAGS	-----	NLPS	----	STRTVKALTEVEAFAL						
BrCNGC9	:	LKEGEFCGEELI	TWALDPKSGV	-----	NLPS	----	STRTVKALTEVEAFAL						
BrCNGC10	:	LKEGDFCGDELI	TWALDPKSGS	-----	NLPS	----	STRTVKALTEVEAFAL						
BrCNGC11	:	LKEGDFCGDELI	TWALDPKSGS	-----	NLPS	----	STRTVKALTEVEAFAL						
BrCNGC12	:	LKEGDFCGEELI	TWALDPKSGS	-----	NLPS	----	STRTAKALTEVEAFAL						
BrCNGC13	:	IGPGDFCGEELI	TWALDPRPVV	-----	ILPS	----	STRTVKAIYEVEAFAL						
BrCNGC14	:	LRPGDFCGEELI	TWALMPNSTL	-----	NFPS	----	STRSVRALSEVEAFAL						
BrCNGC15	:	LRPGDFCGEELI	TWALMPNINQ	-----	NLPL	----	STRTVRTLSEVEAFAL						
BrCNGC16	:	LRPGDFCGEELI	SWALLPKSTL	-----	NLPS	----	STRTVRALVEVEAFAL						
BrCNGC17	:	LKPGDFCGEELI	AWALLPKSKV	-----	NLPS	----	STRTVRALEEVEAFAL						
BrCNGC18	:	LKPGDFCGEELI	AWALLPKSKV	-----	NLPS	----	STRTVRALEEVEAFAL						
BrCNGC19	:	LGPGNFSGDELI	SWCLRRPFVE	-----	RLPP	----	SSSTLVLTLETTEAFGL						
BrCNGC20	:	LGPGNFSGDELI	SWCLRRPFVE	-----	RLPP	----	STSTLVLTLETTEAFGL						
BrCNGC21	:	LEPGGYLGDELI	SWCLRRPFID	-----	RLPP	----	SSATFVCLDNIEAFSL						
BrCNGC22	:	LSKRDVYGEELI	TW--YERSSLNPDGTRIRMP	SKGLLSNRDVKCVTNVEAFSL									
BrCNGC23	:	LSEGDFVGEELI	SWCLERAASNTDGTRIWIKRKGLL	SYRSVRCVTNVDAFSL									
BrCNGC24	:	LSEGDFVGEELI	TWCLERSAVNLDG---	RMLSKGLLSSRNVKCVTNVEAFSL									
BrCNGC25	:	LSEGDFVGEELI	TWCLERSSVNPdGRSIKLP	LKGLVSNRSVRCVTNVDAFSL									
BrCNGC26	:	LSEGDFVGEELI	TWCLERSSVNPdGTRIKMPT	TKGLVSNRSVRCVTNVDAFSL									
BrCNGC27	:	LSEGDFVGEELI	TWCLERSSVNPdGTIIRMP	SKGLLSSRDVWCVTNVDAFSL									
BrCNGC28	:	LSEGDFVGEELI	TWCLERSSVNPdGTRIRIP	SKGLLSYRNVKCVTNVEAFSL									
BrCNGC29	:	LSEGDFVGEELI	TWCLERSSVNPdGTRIKMPT	TKGLVSNRNVRCVTNVDAFSL									
		g	g	LL	W	l		p	S		e	f	L

```

          *           80           *           100           *
CrCNGC1  : MKASEFCGEDLLPWALDPQSS-----SHFP----ISTRRTVQALTEVEAFAL
CrCNGC2  : MKASEFCGEDLLPWALDPQSS-----SHFP----ISTRRTVQALTEVEAFAL
CrCNGC3  : MKASDFCGEDLLTWALDPQSS-----SHFP----ISTRRTVQALTEVEAFAL
CrCNGC4  : LVAGDFCG-DLLTWALDPLSS-----HLP----ISSRTVQAKTEVEGFVL
CrCNGC5  : LGAGDFCGEELLTWALDPHTS-----SNLP----ISTRRTVRAALMEVEAFAL
CrCNGC6  : LKEGDFCGEELLTWALDPKAG-----SNLP----SSTRRTVKALTEVEAFAL
CrCNGC7  : LKEGEFCGEELLTWALDPKSG-----INLP----SSTRRTVKALTEVEAFAL
CrCNGC8  : LKEGDFCGDELLTWALDPKSG-----SNLP----SSTRRTVKALTEVEAFAL
CrCNGC9  : LKEGDFCGEELLTWALDPKSG-----SNLP----SSTRRTAKALTEVEAFAL
CrCNGC10 : LGPGDFCGEELLTWALDPRPV-----VILP----SSTRRTVKAICEVEAFAL
CrCNGC11 : LRPGDFCGEELLVWALLPKST-----LNLP----SSTRRTVRTLVEVEAFAL
CrCNGC12 : LRPGDFCGEELLVWALLPKST-----VNLP----SSTRRTVRALEEVEAFAL
CrCNGC13 : LRPGDFCGEELLTWALMPNIT-----VNLP----LSTRRTVKTLSEVEAFAL
CrCNGC14 : LRPGDFCGEELLTWALMPNST-----LNLP----SSTRSVRALSEVEAFAL
CrCNGC15 : LEPGGYIGDELLVWCLRRPFI-----DRLP----PSSATFVCLDNIEAFSL
CrCNGC16 : LGPGNFSGDELLVWCLRRPFV-----ERLP----PSSSTLVTLTETTEAFGL
CrCNGC17 : LSEGDFCGEELLTWCLERSVINPDGTRIKMQPKGLVSNRNVRCMTNVEAFAL
CrCNGC18 : LSEGDFCGEELLTWCLERSVNPDPGTRIRMPKGLVSNRNVRCVTNVEAFSL
          6       cG   LL W L           p       S r           EaF L

```


	*	80	*	100	*
CsCNGC1	: LVAGDFCG	-DLITWALDPLS	-----SHFF	I-----	SNRTVQAQTEVEEGFVL:
CsCNGC2	: LVAGDFCG	-DLITWALDPLS	-----SHFF	I-----	SSRTVQAQTEVEEGFVL:
CsCNGC3	: LVAGDFCG	-DLITWALDPLS	-----SHFF	I-----	SNRTVQAQTEVEEGFVL:
CsCNGC4	: LVTGDFCG	HDLIPWALDPLS	-----SHFF	I-----	SNRTVRAQTEVEEGFL:
CsCNGC5	: LKAGDICC	-ELIFN	-----GSRL	PT-----	STRTIMTLTQVEEGFVL:
CsCNGC6	: LKASDFCG	EDLIPWALDPQS	-----SSHFF	I-----	STRTVQALTEVEAFAL:
CsCNGC7	: LKASDFCG	EDLIPWALDPQS	-----SSHFF	I-----	STRTVQALTEVEAFAL:
CsCNGC8	: LKASDFCG	EDLITWALDPQS	-----SSHFF	I-----	STRTVQALTEVEAFAL:
CsCNGC9	: LKASDFCG	EDLITWALDPQS	-----SSHFF	I-----	STRTVQALTEVEAFAL:
CsCNGC10	: LKASDFCG	EDLITWALDPQS	-----SSHFF	I-----	STRTVQALTEVEAFAL:
CsCNGC11	: LGAGDFCG	EELITWALDPHT	-----SSNL	PI-----	STRTVRALMEVEAFAL:
CsCNGC12	: LGAGDFCG	EELITWALDPHT	-----SSNL	PI-----	STRTVRALMEVEAFAL:
CsCNGC13	: LGAGDFCG	EELITWALDPHT	-----SSNL	PI-----	STRTVRALMEVEAFAL:
CsCNGC14	: LKEGDFCG	EELITWALDPKA	-----GSNL	PS-----	STRTVKALTEVEAFAL:
CsCNGC15	: LKEGDFCG	EELITWALDPKA	-----GSNL	PS-----	STRTVKALTEVEAFAL:
CsCNGC16	: LKEGEFCG	EELITWALDPKS	-----GINL	PS-----	STRTVKALTEVEAFAL:
CsCNGC17	: LKEGEFCG	EELITWALDPKS	-----GINL	PS-----	STRTVKALTEVEAFAL:
CsCNGC18	: LKEGDFCG	DELITWALDPKS	-----GANL	PS-----	STRTVKALTEVEAFAL:
CsCNGC19	: LKEGDFCG	DELITWALDPKS	-----GANL	PS-----	STRTVKALTEVEAFAL:
CsCNGC20	: LKEGDFCG	DELITWALDPKS	-----GANL	PS-----	STRTVKALTEVEAFAL:
CsCNGC21	: LKEGDFCG	EELITWALDPKS	-----GSNL	PS-----	STRTAKALTEVEAFAL:
CsCNGC22	: LKEGDFCG	EELITWALDPKS	-----GSNL	PS-----	STRTAKALTEVEAFAL:
CsCNGC23	: LKEGDFCG	EELITWALDPKS	-----GSNL	PS-----	STRTAKALTEVEAFAL:
CsCNGC24	: IGPDFCG	EELITWALDPRP	-----VVTL	PS-----	STRTVKAI CEVEAFAL:
CsCNGC25	: IGPDFCG	EELITWALDPRP	-----VVIL	PS-----	STRTVKAI CEVEAFAL:
CsCNGC26	: IGPDFCG	EELITWALDPRP	-----VVIL	PS-----	STRTVKAI CEVEAFAL:
CsCNGC27	: LRPDFCG	EELISWALLPKS	-----TLNL	PS-----	STRTVRALVEVEAFAL:
CsCNGC28	: LRPDFCG	EELISWALLPKS	-----TLNL	PS-----	STRTVRALVEVEAFAL:
CsCNGC29	: LRPDFCG	EELISWALLPKS	-----TLNL	PS-----	STRTVRALVEVEAFAL:
CsCNGC30	: LRPDFCG	EELIAWALLPKS	-----TVNL	PS-----	STRTVRALEEVEAFAL:
CsCNGC31	: LRPDFCG	EELIAWALLPKS	-----TVNL	PS-----	STRTVRALEEVEAFAL:
CsCNGC32	: LRPDFCG	EELITWALMPVI	-----TVNL	PL-----	STRTVKTLSEVEAFAL:
CsCNGC33	: LRPDFCG	EELITWALMPVI	-----TVNL	PL-----	STRTVKTLSEVEAFAL:
CsCNGC34	: LRPDFCG	EELITWALMPVI	-----TVNL	PL-----	STRTVKTLSEVEAFAL:
CsCNGC35	: LRPDFCG	EELITWALMPNS	-----TLNL	PS-----	STRSVRALSEVEAFAL:
CsCNGC36	: LRPDFCG	EELITWALMPNS	-----TLNL	PS-----	STRSVRALSEVEAFAL:
CsCNGC37	: LRPDFCG	EELITWALMPNS	-----TLNL	PS-----	STRSVRALSEVEAFAL:
CsCNGC38	: LEPGGYLC	DELLSWCLRRPF	-----TDRL	LP-----	SSATFVCLDNIEAFSL:
CsCNGC39	: LEPGGYLC	DELLSWCLRRPF	-----TDRL	LP-----	SSATFVCLDNIEAFSL:
CsCNGC40	: LEPGGYLC	DELLSWCLRRPF	-----TDRL	LP-----	SSATFVCLDNIEAFSL:
CsCNGC41	: LGPGNFSC	DELLSWCLRRPF	-----VERL	LP-----	SSSTLVTLLETTEAFGL:
CsCNGC42	: LGPGNFSC	DELLSWCLRRPF	-----VERL	LP-----	SSSTLVTLLETTEAFGL:
CsCNGC43	: LGPGNFSC	DELLSWCLRRPF	-----VERL	LP-----	SSSTLVTLLETTEAFGL:
CsCNGC44	: LSEGDFCG	EELITLCTKRSG	----GTTIKM	PKGLVI	SNRDVTCITNVEAFWL:
CsCNGC45	: LSEGDFCG	EELITWCLERSVTPDGTRIRM	PAKGLL	-----	SNRNVRCVTNVEAFSL:
CsCNGC46	: LSEGDFCG	EELITWCLERSVTPDGTRIRM	PAKGLL	-----	SNRNVRCVTNVEAFSL:
CsCNGC47	: LSEGDFCG	EELITWCLERSVTPDGTRIRM	PAKGLH	-----	SNRNVRCVTNVEAFSL:
CsCNGC48	: LSEGDFCG	EELISWCLERSINPDGTRIKM	PKGLV	-----	SNRNVRCVTNVEAFSL:
CsCNGC49	: LSEGDFCG	EELITWCLERSINPDGTRITM	PKGLV	-----	SNRNVRCVTNVEAFSL:
CsCNGC50	: LSEGDFCG	EELITWCLERSINPDGTRIKM	PKGLV	-----	SNRNVRCVTNVEAFSL:
CsCNGC51	: LSEGDFCG	EELITWCLERSINPDGTRIKM	PKGLV	-----	SNRNVRCVTNVEAFSL:

6 g G LL w l P S Eaf L


```

          *           80           *           100           *
EsCNGC1 : LQAGDIFG-DLIFWVLDPHP-----SSGLP----TSNRTILMTDVEGFILH
EsCNGC2 : LGAGDFCG-ELLFWALDPNA-----SHFP----ISSRTVQALTEVEGFLLS
EsCNGC3 : LNAGDFCG-DLITWALDPNA-----SHFP----ISSRTVQALTEVEGFLLS
EsCNGC4 : LKASDFCGEDLLITWALDPQS-----SSHFP----ISTRTVQALTEVEAFALH
EsCNGC5 : LGAGDFCGEELLITWALDPHT-----SSNLP----ISTRTVRALVEVEAFALH
EsCNGC6 : LKEGDFCGEELLITWALDPKT-----GSNLP----SSTRTVKALTEVEAFALH
EsCNGC7 : LKEGDFCGEELLITWALDPKA-----GSNLP----SSTRTVKALTEVEAFALH
EsCNGC8 : LKEGDFCGEELLITWALDPKS-----GVNLP----SSTRTVKALTEVEAFALH
EsCNGC9 : LKEGDFCGDELLITWALDPKS-----GVNLP----SSTRTVKALTEVEAFALH
EsCNGC10 : LKEGDFCGEELLITWALDPKS-----GSNLP----SSTRTAKALTEVEAFALH
EsCNGC11 : IGP GDFCGEELLITWALDPRP-----VVILP----SSTRTVKAVYEVEAFALH
EsCNGC12 : LRPGEFCGEELLAWALLPKS-----KVNLP----SSTRTVRALEEVEAFALH
EsCNGC13 : LRP GDFCGEELLISWALLPKS-----TLNLP----SSTRTVRALVEVEAFALH
EsCNGC14 : LRP GDFCGEELLITWALMPNV-----NLNLP----LSTRTVRTLSEVEAFALH
EsCNGC15 : LRP GDFCGEELLITWALMPNS-----TLNLP----SSTRSVRALSEVEAFALH
EsCNGC16 : LEQGGYLGDELLSWCIRRRPF-----IDRLP----PSSATFVCIDNIEAFSLH
EsCNGC17 : LGPGNFSGDELLSWCIRRRPF-----VERLP----PSSSTLVTLLETTEAFGLH
EsCNGC18 : LSEGDFCGEELLITWCLERSINPDGTRIKMPPKGLVSSRNVRQVTNVEAFSLH
EsCNGC19 : LTEGDFCGEELLITWCLERSSVNPDGTRIRMPKGLLSNRNVRQVTNVEAFSLH
          6  q  G  LL  W  L           P  S  r  6  E  F  L

```

```

      *           80           *           100           *
LaCNGC1 :IKYGDVCGD-LHFWIPNPTPSL-----HLPI----SRTVMTIEDVEGFIL
LaCNGC2 :LQKGDICGD-LHFWIFNPTHST-----QIPT----FNKTVNTVTDVEGFIL
LaCNGC3 :LKKGDICGD-LHLWFFNPNPSS-----HLRF----SKKTVTTVNNVEGFII
LaCNGC4 :LTTGDICGELLFC-----GS-----RLLT----SERTVMTLTQVEGFIL
LaCNGC5 :LKASDFCGEDLLTWALDPQSSS-----HFPI----STRTVQALTEVEAFAL
LaCNGC6 :LKASDFCGEDLLPWALDPQSSS-----HLPI----STRTVQALTEVEAFAL
LaCNGC7 :LGAGDFCGEELLTWALDPHTSS-----NLPI----STRTVRALVEVEAFAL
LaCNGC8 :LKEGDFCGEELLTWALDPKAGS-----NLPS----STRTVKALTEVEAFAL
LaCNGC9 :LKEGDFCGEELLTWALDPKSGV-----NLPS----STRTVKALTEVEAFAL
LaCNGC10 :LKEGDFCGEELLTWALDPKSGS-----NLPS----STRTAKALTEVEAFAL
LaCNGC11 :LKEGDFCGDELLTWALDPKSGS-----NLPS----STRTVKALTEVEAFAL
LaCNGC12 :LGGDFCGEELLTWALDPRPVV-----ILPS----STRTVKALICEVEAFAL
LaCNGC13 :LRPGDFCGEELLAWALLPKSTV-----NLPS----STRTVRALLEEVEAFAL
LaCNGC14 :LRPGDFCGEELLTWALMPNINL-----NLPL----STRTVRTLSEVEAFAL
LaCNGC15 :LRPGDFCGEELLTWALMPNSTL-----NFPS----STRSVRALSEVEAFAL
LaCNGC16 :LRPGDFCGEELLTWALMPNSTL-----NLPS----STRSVRALSEVEAFAL
LaCNGC17 :LEPGGYLGDELLSWCLRRPFVD-----RLPP----SSATFVSLDNIEAFSL
LaCNGC18 :VEPGGYLGDELLFSWCLRRPFVK-----RLPP----SSATFVSVDNIEAFSL
LaCNGC19 :LSEGDFCGEELLTWCLERSNVHDGTRIRIPSRGLLSNRNVRCVTNVEAFSL
LaCNGC20 :LSEGDFCGEELLTWCLDRSSINPDGTRIKMPSKGLVSNRSVRCVTNVETESL
      6  gd  cG  L  w           p  s           6  6E  F  6

```



```

      80          *          100          *          120
SiCNGC1 : IQAGDICG-DLLFWILDPHPSS-----GLPT---SNKTIILTLTDVEGFIL:
SiCNGC2 : IQEGDICG-DLIF---NPNPSS-----HLPT---SSETIMSLTDVEGFIL:
SiCNGC3 : IGAGDICG-DLIFLIINTSSSS-----GLPT---SNRTVITDTHVEGFIL:
SiCNGC4 : IEGGDICG-ELLFWMLAPYPSS-----RLPT---SNKTIMTLTDVEGFIL:
SiCNGC5 : IGAGEFCG-DLLTWALDPNTS-----QFPI---SSRTVQALTEVEGFVL:
SiCNGC6 : IKASEFCGEDLLTWALDPQSSS-----NFPI---STRTVQAI TEVEAFAL:
SiCNGC7 : IKASDFCGEDLLPWALDPQSSS-----HFPI---STRTVQALTEVEAFAL:
SiCNGC8 : IGAGDFCGEELLTWALDPHTTS-----NLFI---STRTVRALVEVEAFAL:
SiCNGC9 : IKEGDFCGEELLTWALDPKAGS-----NLPS---STRTVKALTEVEAFAL:
SiCNGC10 : IKEGDFCGEELLTWALDPKAGS-----NLPS---STRTVKALTEVEAFAL:
SiCNGC11 : IKEGEFCGEELLTWALDPKSGV-----NLPS---STRTVKALTEVEAFAL:
SiCNGC12 : IKEGDFCGDELLTWALDPKSGS-----NLPS---STRTVKALTEVEAFAL:
SiCNGC13 : IKEGDFCGEELLTWALDPKSGS-----NLPS---STRTAKALTEVEAFAL:
SiCNGC14 : IGPGDFCGEELLTWALDPRPVV-----ILPS---STRTVKAMYVEVEAFAL:
SiCNGC15 : IRRPGEFCGEELLIAWALLPKSKV-----NLPS---STRTVRALEEVEAFAL:
SiCNGC16 : IRRAGDFCGEELLIWALMPNVNL-----NLPL---STRTVRTLSEVEAFAL:
SiCNGC17 : IRRPGDFCGEELLIWALMPNSTL-----NLPS---STRSVRALSEVEAFAL:
SiCNGC18 : IEPGGYLGDELLSWCLRRPFID-----RLPP---SSATFVCLDNIEAFSL:
SiCNGC19 : IGPGNFSGDELLSWCLRRPFVE-----RLPP---SSSTLVTLTETEAFFGL:
SiCNGC20 : ISEGDVCGEELLIWCLERSAVNPDGTTIRIPSKGLLSRNVRVCTNVEAFSL:
SiCNGC21 : ISEGDVCGEELLIWCLERTSVNPDGTKIRLPSRGLLSNRNVRCVTNVEAFSL:
SiCNGC22 : ISEGDVCGEELLIWCLERSTVNPDPGTRIKMPPKGLVSNRSVRCVTNVEAFSL:
      6  g  cG  LL  w  l          P          S          E  F  L

```



```

          *           80           *           100           *
ThCNGC1 : LQAGDIFGD-LLFWVIDPHP-----SSGLP----TSNRTILTM TDVEGFIL
ThCNGC2 : LGAGDFCGETLFWAIDPNA-----SHFP----ISSRTVQAITEVEGFLL
ThCNGC3 : LNAGDFCGD-LLTWAIDPNA-----SHFP----ISSRTVQAITEVEGFLL
ThCNGC4 : LKASDFCGEDLLTWAIDPQS-----SSHFP----ISTRTVQAITEVEAFAL
ThCNGC5 : LGAGDFCGEELLTWAIDPHT-----SSNLP----ISTRTVRAIVEVEAFAL
ThCNGC6 : LKEGDFCGEELLTWAIDPKT-----GSNLP----SSTRTVKAITVEVEAFAL
ThCNGC7 : LKEGDFCGEELLTWAIDPKA-----GSNLP----SSTRTVKAITVEVEAFAL
ThCNGC8 : LKEGDFCGEELLTWAIDPKS-----GVNLP----SSTRTVKAITVEVEAFAL
ThCNGC9 : LKEGDFCGDELLTWAIDPKS-----GVNLP----SSTRTVKAITVEVEAFAL
ThCNGC10 : LKEGDFCGEELLTWAIDPKS-----GSNLP----SSTRAKAITVEVEAFAL
ThCNGC11 : IGPDFCGEELLTWAIDPRP-----VVILP----SSTRTVKAVYEVEAFAL
ThCNGC12 : LRPGEFCGEELLAWAILPKS-----KVNLP----SSTRTVRAIEEVEAFAL
ThCNGC13 : LRPGEFCGEELLSWAILPKS-----TLNLP----SSTRTVRAIVEVEAFAL
ThCNGC14 : LRPGEFCGEELLTWAIDMPNV-----NLNLP----ISTRTVRTLSEVEAFAL
ThCNGC15 : LRPGEFCGEELLTWAIDMPNS-----TLNLP----SSTRSVRAISEVEAFAL
ThCNGC16 : LEQGGYLCDELLSWCI RRRPF-----IDRLP----PSSATFVCI DNI EAFSL
ThCNGC17 : LGPGNFSGDELLSWCI RRRPF-----VERLP----PSSSTLVTL ETT EAFGL
ThCNGC18 : LSEGDFCGEELLTWCI ERSSINPDGTRIKM PPKGLVSSRNVR CVTNVEAFSL
ThCNGC19 : LTEGDFCGEELLTWCI ERSSVNPDGTRIRM PPKGLVSSRNVR CVTNVEAFSL
          6  g  G  Ll W L           P           S  r           6  E  F  L

```

```

          *           80           *           100           *
TsCNGC1 : LQAGDIFG-DLLFWVIDPHF-----SSGLP----TSNRTILTMTDVEGFIL
TsCNGC2 : LNAGDFCG-DLLTWALDPNA-----SHFP----ISSRTVQALTEVEGFIL
TsCNGC3 : LKASDFCGEDLLTWALDPQS-----SSHFP----ISTRVQALTEVEAFAL
TsCNGC4 : LGAGDFCGEELLTWALDPHT-----SSNLP----ISTRVRAIVEVEAFAL
TsCNGC5 : LKEGDFCGEELLTWALDPKT-----GSNLP----SSTRVKAITEVEAFAL
TsCNGC6 : LKEGDFCGEELLTWALDPKA-----GSNLP----SSTRVKAITEVEAFAL
TsCNGC7 : LKEGDFCGEELLTWALDPKS-----GVNLP----SSTRVKAITEVEAFAL
TsCNGC8 : LKEGDFCGDELLTWALDPKS-----GVNLP----SSTRVKAITEVEAFAL
TsCNGC9 : LKEGDFCGEELLTWALDPKS-----GSNLP----SSTRAKAITEVEAFAL
TsCNGC10 : LGPGDFCGEELLTWALDPRP-----VVILP----SSTRVKAVYEVEAFAL
TsCNGC11 : LRPGDFCGEELLTWALMPNV-----NLNLP----ISTRVRTLSEVEAFAL
TsCNGC12 : LRPGDFCGEELLTWALPKS-----TLNLP----SSTRVRAIVEVEAFAL
TsCNGC13 : LRPGEFCGEELLAWALPKS-----KVNLP----SSTRVRAIEEVEAFAL
TsCNGC14 : LRPGDFCGEELLTWALMPNS-----TLNLP----SSTRSVRAISEVEAFAL
TsCNGC15 : LEQGGYLGDELLSWCI RRPF-----IDRLP----PSSATFVCLDNI EAFSL
TsCNGC16 : LGPGNFSCDELLSWCI RRPF-----VERLP----PSSSTLVTLTET EAFGL
TsCNGC17 : LSEGDFCGEELLTWCI ERSSINPDGTRIKM PKGLVSSRNVR CVTNVEAFSL
TsCNGC18 : LTEGDFCGEELLTWCI ERSSVNPDGTRIRM P SKGLLSNRNVR CVTNVEAFSL
          6  g  G  LL  W  L           P           S  r           6  EaF  L

```

