*Indian J. Genet.*, 76(3): 233-245 (2016)
DOI: 10.5958/0975-6906.2016.00038.9

# Cloning and characterization of drought stress-induced NAC transcription factors from *Brassica juncea* and *Sinapis alba*

**Dharitree Phukan, Bhupendra Singh P., Indu Ravi[1], Amitha S. V. Mithra, Devendra K. Yadava[2], Viswanathan Chinnusamy[2] and Trilochan Mohapatra[4]***

ICAR-National Research Center on Plant Biotechnology, New Delhi 110 012; [1]IGNOU Regional Centre, Jaipur 302 020; [2]Division of Seed Science and Technology; [3]Division of Plant Physiology, ICAR-Indian Agricultural Research Institute, New Delhi 110 012; [4]ICAR, Krishi Bhawan, New Delhi 110 001

(Received: February 2016; Revised: July 2016; Accepted: August 2016)

**Abstract**

**The plant specific NAC [for NAM (no apical meristem), ATAF, CUC (cup-shaped cotyledon)] TFs are one of the largest plant TF families that play important roles in plant development and stress tolerance. Suppression subtractive hybridization (SSH) analysis with using drought stressed plants of *S. alba* lead to the identification of several stress responsive ESTs. Two of them homologous to *Arabidopsis* NAC14 and NAC19 were selected for cloning of full length CDS and expression analysis in *Brassica* and related species with contrasting drought tolerance. *NAC14* and *NAC19* genes were cloned from drought tolerant *Sinapis alba* and *Brassica juncea* cvs. RGN73 and Varuna, and drought sensitive *B. juncea* cvs. RLM619, BEC144 and BioYSR. Sequencing of genomic region coding for these NACs revealed that both *NAC14* and *NAC19* contain 3 exons and 2 introns each. *In silico* analysis of protein structure led to development of 3D models of these stress responsive NAC TFs. Although both proteins have 7 strands of β sheets, the NAC14 had 5 β sheets of type A and 2 β sheets of type B, while NAC14 have all 7 β sheets of type A. These proteins also differed in helix content, β turns and g turns. This suggest their functional diversity under abiotic stresses. Real-time RT-PCR expression analysis revealed that both the genes were up-regulated under drought stress in the leaves of *B. juncea* genotypes Varuna and BioYSR. In addition, *NAC14* was up-regulated in the leaves of RLM619, while *NAC19* was up-regulated in the leaves of *S. alba* and BEC144 under drought stress as compared to control conditions. Interestingly, drought stress did not up-regulate these genes in RGN73. This study revealed genotypic variation in the drought regulation of *NAC* TFs in *B. juncea* and *S. alba*.**

**Key words:** Drought, mustard, NAC proteins, *S. alba,* structural analysis

## Introduction

The projected demand for oilseeds in India alone is around 34 million tonnes by 2020, out of which nearly 14 million tonnes (41%) need to be met by mustard (Yadava and Singh 1999). Hence, improvement in the productivity and abiotic stress tolerance of Indian mustard is critical for meeting edible oil demand. NAC (NAM, ATAF and CUC) transcription factor (TF) family represents one of the largest families of plant-specific TFs. NAC TFs regulate developmental processes such as maintenance of shoot apical meristem, flower development, leaf senescence, embryo development, lateral root formation, secondary wall thickening, and abiotic and biotic stress responses (Nakashima et al. 2012; Nuruzzaman et al. 2013; Pereira-Santana et al. 2015; Shao et al. 2015; Kim et al. 2016a). Transgenic Arabidopsis overexpressing abiotic stress inducible *ANAC019*, *ANAC055*, or *ANAC072* genes resulted in enhanced expression of several stress responsive genes and conferred enhanced drought tolerance (Tran et al. 2004). Similarly, transgenic Arabidopsis plants overexpressing *ZmSNAC1* (Lu et al. 2012), *TaNAC2* (Mao et al. 2012) and *AhNAC2* (Liu et al. 2011) showed tolerance to different abiotic stresses. Rice transgenic plants overexpressing SNAC1, SNAC2, SNAC3, OsNAC4, OsNAC5, OsNAC6, OsNAC10, ONAC022, ONAC045, ONAC106 showed tolerance to different abiotic stresses by modulating leaf senescence, reactive oxygen species, enhanced ABA sensitivity, redox homeostasis, proteolytic degradation and

*Corresponding author's e-mail: tmnrcpb@gmail.com; dharitree19@gmail.com
Published by the Indian Society of Genetics & Plant Breeding, F2, First Floor, NASC Complex, PB#11312, IARI, New Delhi 110 012
Online management by indianjournals.com

enhanced expression of several stress responsive genes (Hu et al. 2006; Zheng et al. 2009; Nakashima et al. 2007; Jeong et al. 2010; Song et al. 2011; You et al. 2014; Fang et al. 2015; Sakuraba et al. 2015; Hong et al. 2016). Abiotic stress inducible wheat gene TaNAC69 overexpression in transgenic wheat resulted in increased drought tolerance (Xue et al. 2011). Wheat TaNAC67 conferred enhanced tolerance to drought, salt and freezing stresses, as supported by enhanced expression of multiple abiotic stress responsive genes and improved physiological traits, *viz*., cell membrane stability, retention of improved chlorophyll contents, Na+ efflux rates, better photosynthetic potential and elevated water retention capability.

Arabidopsis and rice NAC family consists of >100 members each. In *Brassica rapa*, about 96 members have been predicted (Cheng et al. 2011; http://brassicadb.org/brad/geneFamily.php), while 65 members have been reported in B. napus as given in Plant Transcription Factor Database v3.0 (http://planttfdb.cbi.pku.edu.cn/family.php?sp=Bna&fam=NAC). However, only few NAC TFs have been characterized from *Brassica* species. *B. napus* NAC14, NAC19, NAC485, BnaNAC82 and NAC103 have been found to be regulated by different abiotic stresses. (Hegedus et al. 2003; Zhong et al. 2012; Niu et al. 2014; Ying et al. 2014; Wang et al. 2015). BnNAC103 transcription factor gene when overexpressed resulted in higher reactive oxygen species (ROS) accumulation and cell death in plants (Niu et al. 2014). Similarly, Overexpression of *BnaNAC19* and *BnaNAC82* resulted in accumulation of ROS and hypersensitive response like cell death when in transgenic tobacco (Wang et al. 2015). BnNAC14 was identified from subtractive expressed sequence tag analysis and screening of cDNA libraries of *B. napus* (Hegedus et al. 2003). BnNAC19 gene was identified from canola through a systematical analysis and mining of expressed sequence tags (Wang et al. 2015).

Earlier by employing suppression subtractive hybridization (SSH) analysis, we identified several stress responsive ESTs from *S. alba* including partial ESTs of NAC14 and NAC19 (Palit et al. 2014). We screened 38 genotypes of mustard and *Sinapis alba*, and identified genotypes with contrasting drought tolerance and identified *S. abla*, RGN73 and Varuna as tolerant whereas RLM619, BEC144 and BioYSR as susceptible (Phukan et al. 2016). Hence these genotypes were selected for cloning and characterization of *NAC14* and *NAC19*.

## Materials and methods

### *Cloning of NAC14 and NAC19*

The primers (NAC14_F 5'CAATAAGAAGAAGAAGAAAAAGTGG3', NAC14_R 5'CCGGTTCAGCATAGTGGATT3', NAC19_F 5'ATGGGTATCCAAGAAACTGACCCGT3' and NAC19_R 5'TCACATAAACCCAAACCCACCA3') were designed to amplify the full length *NAC14* (945bp) and *NAC19* (1164bp) from 6 genotypes: five of *B. juncea* namely RGN73, Varuna, RLM619, BEC144 and BioYSR and one of *Sinapis alba*. The resulting PCR products were cloned in the pDrive Cloning Vector (QIAGEN) and sequenced by ABI 3730XL DNA analyzer (Applied Biosytems, Hitachi, USA). The sequences were submitted in NCBI GenBank [GenBank accession numbers: *S. alba* NAC14 (KT281870), RGN73 NAC14 (KT281871), Varuna NAC14 (KT281872), RLM619 NAC14 (KT281873), BEC144 NAC14 (KT281874), BioYSR NAC14 (KT281875), *S. alba* NAC19 (KT281876), RGN73 NAC19 (KT281877), Varuna NAC19 (KT281878), RLM619 NAC19 (KT281879), BEC144 NAC19 (KT281880), BioYSR NAC19 (KT281881)].

### *Sequence analysis of NAC14 and NAC19 proteins*

Protein sequences of NAC14 and NAC19 from *S. alba,* RGN73, Varuna, RLM619, BEC144 and BioYSR cloned in this study were used to identify closely related gene using NCBI blast module blastp. User friendly BLAST output visualization tool Circos was used to visualize sequence similarity (Krzywinski et al. 2009; Darzentas 2010). Homologues of NAC genes from different species having top blast hits of >77% identity were selected and used for phylogenetic analysis.

### *Multiple sequence alignment and phylogenetic analysis*

Top quality hits generated by blastp were aligned with MAFFT v.7 (Katoh et al. 2002; Katoh and Standley 2013) and figured using ESPript (http://espript.ibcp.fr/ESPript/cgi-bin/ESPript.cgi) (Gouet et al. 1999, 2003, Robert and Gouet 2014). A phylogenetic tree was constructed using the maximum-likelihood method implemented in the PhyML program version 3.1 aLRT (Anisimova and Gascuel 2006) on the Phylogeny.fr platform (Dereeper et al. 2008) with the LG substitution model and default settings. Reliability for internal branch was assessed using the bootstrapping method with 100 bootstrap replicates. Graphical representation and edition of the phylogenetic tree were performed with Itol: Interactive Tree Of Life (Letunic and Bork 2007).

### *Physiochemical analysis of NAC14 and NAC19*

Physiochemical analysis of NAC14 and NAC19 were predicted by online software ProtParam (http://www.expasy.org/tools/protparam.html) (Gasteiger et al. 2005), which predicts molecular mass, theoretical isoelectric point (pI), amino acid composition, atomic composition, instability index as well as grand average of hydropathicity (GRAVY).

### *Structure analysis of NAC14 and NAC19*

Secondary structure of deduced amino acid sequence of NAC14 and NAC19 was analyzed by GOR secondary structure prediction method version IV. The structure was predicted and compared (npsa-pbil.ibcp.fr/cgi-bin/npsaautomat.pl?page=npsagor4.html). The folding state of NAC14 and NAC19 was predicted by Fold Index program (bioportal.weizmann.ac.il/fldbin/findex). 3D structure prediction of NAC14 and NAC19 was done by iterative threading assembly refinement in I-TASSER server (Roy et al. 2010). The predicted five different models of *NAC14* and *NAC19* were generated and the model showing overall best stereo chemical quality was selected for further quality assessment.

### *Quality assessment of predicted structures*

Modeling of loops in protein was performed with ModLoop (http://modbase.compbio.ucsf.edu/modloop/) (Foster 2002; Fiser and Sali 2003). Energy minimization of the constructed protein structure was done by YASARA Energy Minimization Server (http://www.yasara.org/minimizationserver.htm). Ramachandran plot analysis was done with RAMPAGE (http://mordred.bioc.cam.ac.uk/~rapper/rampage.php), where all the stereo chemical properties of the proteins were analyzed. The structures were further verified with SAVES (http://nihserver.mbi.ucla.edu/SAVES/). The 3D model of NAC14 and NAC19 was subjected to Pymol Molecular Graphic System (http://pymol.org/ep) to obtain the final structure. The PDB files of modeled NAC14 and NAC19 were subjected to PDBsum server (http://www.ebi.ac.uk/thorntonsrv/databases/pdbsum/generate.html) for structural motif analysis and to ProSA-web server for getting the reliable values for the model generated (http://prosa.services.came.sbg.ac.at/prosa.php).

### *Expression analysis of NAC14 and NAC19*

The seedlings of *S. alba,* RGN73*,* Varuna*,* RLM619*,* BEC144 and BioYSR were grown in soil in the culture room. Fifteen days after germination, drought stress was imposed by withholding irrigation to the plants for eight days. Leaf samples of control and drought stress treated plants were collected for expression analysis. Expression of *NAC14* and *NAC19* genes was analyzed by quantitative RT-PCR, using gene-specific primers: NAC14-qF TGGACGATTGGGTGTTGTGTCGTA, NAC14-qR ATCACATGACCGTTCGCTACCTCA, NAC19-qF TGGGTATCCAAGAAACTGACCCGT, NAC19-qR TAAGCTCTTCGTCGGTCGGGTAAA. Expression of the *NAC* genes in various samples was normalized with actin (GenBank: KM881428.1) reference gene as an internal control to quantify the expression level of the transcripts. Total RNA was isolated from controlled and stressed plants by using GeneJET$^{TM}$ Plant RNA Purification mini kit (Fermentas, EU) and cDNA was synthesized using SuperScript$^{TM}$III First-Strand Synthesis System for RT-PCR (Invitrogen, USA) following the manufacturer's instructions. The resulted cDNA samples were diluted 5 times (1:5) in RNase-free water and 1µl of the diluted cDNA was used as template in total reaction volume of 25µl using Power SYBR Green PCR Master Mix (Applied Biosystem, Life Technology, USA). Real time PCR analysis was performed using Stratagene M X 3005P Q-PCR system. The Q-PCR amplification was performed with triplicates. The specificity of the PCR reactions was confirmed by melting curve analysis of the amplicons. The comparative $2^{-\Delta\Delta CT}$ method was used to calculate the normalized fold change of each transcript in the samples.

## Results and discussion

### *Cloning of NAC14 and NAC19*

Full length coding sequences of *NAC14* and *NAC19* were cloned and sequenced from six genotypes namely *S.alba,* RGN73, Varuna, RLM619, BEC144 and BioYSR, and these sequences were submitted to NCBI GenBank (Supplementary Fig. 1: available online at http://www.isgpb.co.in). Comparison of sequences of the amplicon from genomic DNA with that of cDNA revealed that *NAC14* gene has 3 exons (1 to 169, 245 to 519, 607 to945) and 2 introns (170 to 244 and 520 to 606). It encodes a protein with 260 amino acids. NAC domain of *NAC14* ranged from 9$^{th}$ to 139$^{th}$ amino acid positions. The NAC domain was subdivided into five sub-domains (A-E). Sub-domains A and B extended from amino acid 9$^{th}$ to 30$^{th}$, and 41$^{st}$ to 55$^{th}$ positions, respectively, which contained a high proportion of acidic amino acids as described previously (Puranik et al. 2012). Sub-domains C and

D ranged from 63rd to 99th, and 105th to 136th positions, respectively and were rich in basic amino acids. The N-terminal sub-domains (A and B) exhibited a net negative charge, while the others were positively charged as shown by Xie et al. 1999. Putative nuclear localization signals (NLS) have been detected in sub-domains C and D as shown in earlier study (Kikuchi et al. 2000). The DBD contained within a 60 amino acid region was located within sub-domains D and E which extended up to 160th position of amino acid sequence. Besides, sub-domains A, C and D were distinctly conserved, whereas sub-domains B and E were relatively less conserved (Ooka et al. 2003). The highly conserved sub-domains C and D might be responsible for binding to DNA, and sub-domain A might be involved in homo- and hetero-dimerization, whereas the divergent sub-domains B and E might be implicated in the functional diversity of the NAC proteins (Jensen et al. 2010; Chen et al. 2011). Four amino acid substitutions in the coding region of *NAC14* gene were identified among the six genotypes of *Brassica*.

*NAC19* also found to contain 3 exons and 2 introns. *NAC19* cloned from *S. alba* encoded a putative protein with 309 amino acids whereas RGN73 and RLM619 encoded 315 and Varuna, BEC144 and BioYSR encoded 314 amino acids. NAC domain of *NAC19* ranged from 14th to 139th position of amino acid sequence. There were eleven amino acid substitutions in the coding region of *NAC19* gene among the six genotypes of *Brassica* due to nucleotide substitutions and deletions. Deletions observed in the NAC domain of *S. alba* could be due its genetic distance being a wild relative of the cultivated *B. juncea*. In previous studies on wild relatives of *Brassica* crops, *S. alba* has been reported to exhibit greater tolerance to drought (Warwick 1993; Brown et al. 1997; Phukan et al. 2016). It would be interesting to experimentally establish relationship between the observed sequence level differences with trait variation among different genotypes of *Brassica* and its wild relatives.

### *Multiple sequence alignment and Phylogenetic Analysis of NAC14 and NAC19*

Multiple sequence alignments of NAC14 and NAC19 proteins cloned from RGN73, Varuna, RLM619, BEC144, BioYSR and *S. alba* were used for comparison with of NAC TFs from different species. It depicted that most of the regions of the NACs are conserved showing similar evolutionary relationship among different species (Supplementary Fig. 2: available online at http://www.isgpb.co.in). It showed high sequence conservation in the NAC domain present in the N-terminal of the protein (Supplementary Fig. 3: available online at http://www.isgpb.co.in). Most of the conserved motifs found within the N-terminal NAC domain indicated that these motifs may be essential for the function of NAC proteins (Liu et al. 2014). The C-terminal is known as the transcription regulatory region of NAC proteins (Puranik et al. 2012). This region was found highly diverged between NAC14 and NAC19 suggesting that these proteins might regulate different regulons involved in stress tolerance in *Brassica* species (Supplementary Fig. 3: available online at http://www.isgpb.co.in).

Circos drawn for NAC14 and NAC19 is shown in Figs. 1a and 1b, which provide an essential first glimpse at sequence relationships. NAC14 cloned from RGN73 and BEC144 showed 100% similarity with NAC14 of *B. rapa* and *S. alba*, and 97.69% similarity with *B. napus*, NAC14 cloned from Varuna showed 99.62% similarity with that of *B. rapa* and *S. alba*, while it showed 98.08% similarity with NAC14 *B. napus*. NAC14 from RLM619 had 98.85% similarity with that of *B. rapa* and *S. alba*, and 97.31% similarity with *B. napus*. NAC14 cloned from BioYSR showed 99.62% similarity with NAC14 of *B. rapa* and *S. alba*, and 97.31% with that of *B. napus*. NAC14 cloned in this study showed 80-86% similarity with *E. salsugineum*, *A. thaliana*, *A. lyrata*, *A. alpina*, *C. rubella* and *C. sativa*. NAC19 cloned from RGN73, RLM619, Varuna and BioYSR showed about 98.41 to 99.37% similarity with that of *B. rapa* and *B. napus*, while NAC19 of BEC144 showed about 95% similarity with that of *B. rapa* and *B. napus*. NAC19 from RGN73, RLM619, Varuna and BioYSR showed 92% similarity with that of *S. alba*, while NAC19 of BEC144 showed 89.52% similarity with that of *S. alba*. The NAC19 cloned in this study showed 83.28 to 88.61% similarity with that of *E. salsugineum*, *A. lyrata*, *A. thaliana*, *C. sativa* and *A. alpine*. Sequence similarity scores can be used as the base for the phylogenetic analysis. The NAC14 cloned from *S. alba* and *B. juncea* genotypes showed high similarity in the N terminal DBD (DNA-binding domain) region with respective domain in proteins from different species, while the C-terminal TRD (transcription regulatory domain) showed large variation between *Brassica* species as compared with *E. salsugineum*, *A. lyrata*, *A. thaliana*, *C. sativa* and *A. alpine*. This suggested that the transcription activation function is diverged during the evaluation (Tran et al. 2004; Hu et al. 2006; Kim et al. 2007b). However, NAC19 showed high similarity in both DBD and TRD

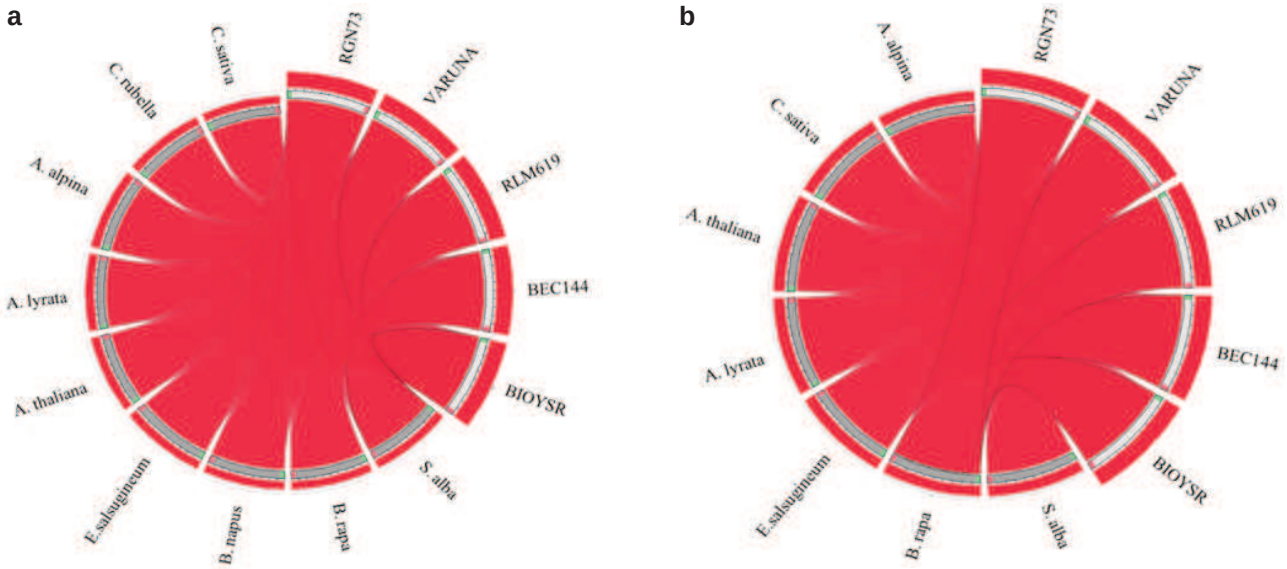



**Fig. 1. Circoletto showing similarity of NAC14 (a) and NAC19 (b) with NACs of different plant species. NAC14 cloned from *B. juncea* and *S. alba* were compared with NAC TFs of BraNAC2 (XP_009128229.1), BnaNAC14 (AAP35055.1), EsEUTSA (XP_006390089.1), AtANAC032 (NP_177869.1), AlANAC032 (XP_002887684.1), AaAALP (KFK42139.1), CrCARUB (XP_006302722.1), and CsNAC2 (XP_010428823.1). NAC19 cloned from *B. juncea* and *S. alba* were compared with BraNAC19 (XP_009147649.1), BnaNAC19 (AHN60135.1), EsEUTSA (XP_006392855.1), AtNAC19 (NP_175697.1), AlANAC019 (XP_002894409.1), CsNAC19 (XP_010462142.1) and AaAALP (KFK35760.1)**

domains across species, suggesting that NAC19 function appears to be conserved across the species studied here.

Sequence conservation across different species is an important indicator of functionality and evolution. To study the evolutionary relationship among the NAC TFs in different plant species, NAC14 cloned from six genotypes in this study were compared with NAC transcription factors from different species, BraNAC2, BnaNAC14, EsEUTSA, AtANAC032, AlANAC032, AaAALP, CrCARUB, and CsNAC2 to study their evolution. On the other hand NAC19 sequences were compared with BraNAC19, BnaNAC19, EsEUTSA, AtNAC19, AlANAC019, CsNAC19 and AaAALP (Supplementary Fig. 4: available online at http://www.isgpb.co.in). The tree generated with NAC homologues suggested that they might have been evolved from same ancestors of the NAC family. Since the *B. juncea* (AABB) and *B. napus* (AACC) were derived from ancestral genome of *B. rapa* (AA), it is quite obvious that five genotypes of *B. juncea* i.e. RGN73, Varuna, RLM619, BEC144 and BioYSR showed close relation with *B. rapa* and *B. napus*. *S. alba* (genome SS, 2n = 24) which belongs to the *nigra* lineage (Agerbirk et al. 2008), was also found phylogenetically closer to *B. juncea*.

### *Physiochemical analysis of NAC14 and NAC19*

The average molecular weight of NAC14 and NAC19 proteins is 29621.9 and 34612.7 g $mol^{-1}$, respectively. Considerable difference in isoelectric point (pI) was observed between these two NACs as the pI of NAC14 was 9.11, while that of NAC19 was 6.19. The computed isoelectric point (PI) of NAC14 was above 7 indicating that the proteins are useful for developing buffers for purification by isoelectric focusing method (Reehana et al. 2013). The total number of negatively charged residues (Asp+ Glu) was 32 and positively charged residues (Arg+ Lys) was 39 in NAC14, where as for NAC19 (Asp+ Glu) was 36 and (Arg+ Lys) was 33. Consistent with the fact that the NAC domain is rich in positive charges and might involved in DNA binding (Ernst et al. 2004; Puranik et al. 2012, Le et al. 2011), NAC14 and NAC19 cloned in this study also showed richness of positively charged amino acids in the NAC domain, suggesting its function in DNA binding activity. Considerable differences were observed in extinction coefficient and instability index between these two NACs studied here (Table 1). The differences in the extinction coefficient of NACs appear to be due to the concentration of Cys residues.

**Table 1.** Physiochemical analysis of *NAC14* and *NAC19*

| Physicochemical parameters | *NAC14* | *NAC19* |
|---|---|---|
| Length (total no. of amino acids) | 260 | 308 |
| Isoelectric point (PI) | 9.11 | 6.19 |
| Molecular weight (MW) | 29621.9 | 34612.7 |
| Negatively charged residues (Asp+ Glu) | 32 | 36 |
| Positively charged residues (Arg+ Lys) | 39 | 33 |
| Extinction coefficient (EC) | 55140 | 42985 |
| Instability index (II) | 40.15 | 26.84 |
| Aliphatic index (AI) | 73.88 | 68.02 |
| Grand average of hydropathicity (GRAVY) | –0.506 | –0.492 |

The ProtParam server predicted an instability index of 40.15 and 26.84 for *NAC14* and *NAC19* proteins, respectively. As proteins with instability index larger than 40 are unstable (Guruprasad et al. 1990), NAC14 would be unstable whereas NAC19 would be stable in solution. The aliphatic index for the NAC14 is 73.88 and for NAC19 is 68.02, which is regarded as a positive factor for the increase of thermal stability of globular proteins (Ikai 1980) which will help to maintain their stability and activities at high temperature. The GRAVY value for both ranged from - 0.506 to -0.492 indicating that these proteins will interact favorably with water (Table 1).

### *Structural analysis of NAC14 and NAC19*

Structural analyses are precious sources of information on shapes and domain structure, protein classification, function prediction and interaction with other macromolecules. The structure provides the first framework to understand the interactions of NAC TFs with DNA at the molecular level. The secondary structure helps in determining the exact structure of the gene. The secondary structure of the NAC14 protein is dominated by random coil (64.23%), whereas alpha helices and extended strand contributed to 13.46 and 22.31 %, respectively (Fig. 2a). The secondary structure of NAC19 showed 59.09, 17.53 and 23.38 % random coil, alpha helices and extended strand, respectively (Fig. 2 b). The secondary structures are more conserved than the nucleotide sequences, which help in understanding molecular evolution (Reehana et al. 2013).

Further 260 residues of NAC14 revealed 0.058 unfoldability, 0.027 charge and 0.444 phobic values. Predicted disordered region of NAC14 was between 54-155 amino acids (score: –0.15 ± 0.08) having 120 disordered residues (Fig. 3a). 309 residues of NAC19 revealed 0.079 unfoldability, 0.010 charge and 0.445 phobic values. There were 5 predicted disordered regions of NAC19 having 175 disordered residues. Predicted disorder segment were [76]-[87] length: 12 score: -0.06 ± 0.03, [89]-[126] length: 38 score: -0.09 ± 0.06, [129]-[166] length: 38 score: –0.09 ± 0.04, [173]-[230] length: 58 score: -0.13 ± 0.05 (longest disordered region) and [248]-[276] length: 29 score: –0.06 ± 0.04 (Fig. 3b).

Pymol Molecular Graphic System was used to generate the final 3D structure of NAC14 and NAC19 (Fig. 5), and surface morphology of NAC14 and NAC19 protein was modelled in a mesh style (Suplementary Fig. 5: available online at http://www.isgpb.co.in). Molecular modelling methods are now routinely used to analyze the structure, dynamics, surface properties and thermodynamics of inorganic, biological and polymeric systems. Ramachandran plot assessment of NAC14 revealed 96.1% of residues in most favored zone, 3.9% in allowed region and no amino acid in disallowed/outlier region. Ramachandran plot assessment of NAC19 revealed 95.8% of residues in most favored zone, 3.9% in allowed region and 0.3% amino acid in disallowed region (Fig. 4). The statistics in the favored and allowed region and relative low percentage in the outlier suggest Ramachandran plots for NAC14 and NAC19 are acceptable. The Goodness factor (G-factor) from the PROCHECK results showed relevant information between covalent and overall bond-angle distances. Analysis of G-Factor of the modeled NAC14 was 0.00 and NAC19 was 0.04, which revealed the quality of the predicted model is very good. The overall Ramachandran plot attributes and the G-factor assured the quality of NAC14 and NAC19 structures (Table 2). ERRAT analysis of NAC14 showed overall quality factor of 81.988 and NAC19 revealed overall quality factor of 78.302.

The PDBsum server facilitated to derive and verify the secondary structure and topology of NAC proteins. NAC14 consist of 5 antiparallel strands of β sheets A and 2 antiparallel strands of β sheets B surruounded by helices. This results were similar with the fact that NAC domain fold consist of a twisted â sheets surruounded by helical element (Ernst et al. 2004). One helice is of G type (Leu55 to Met58)

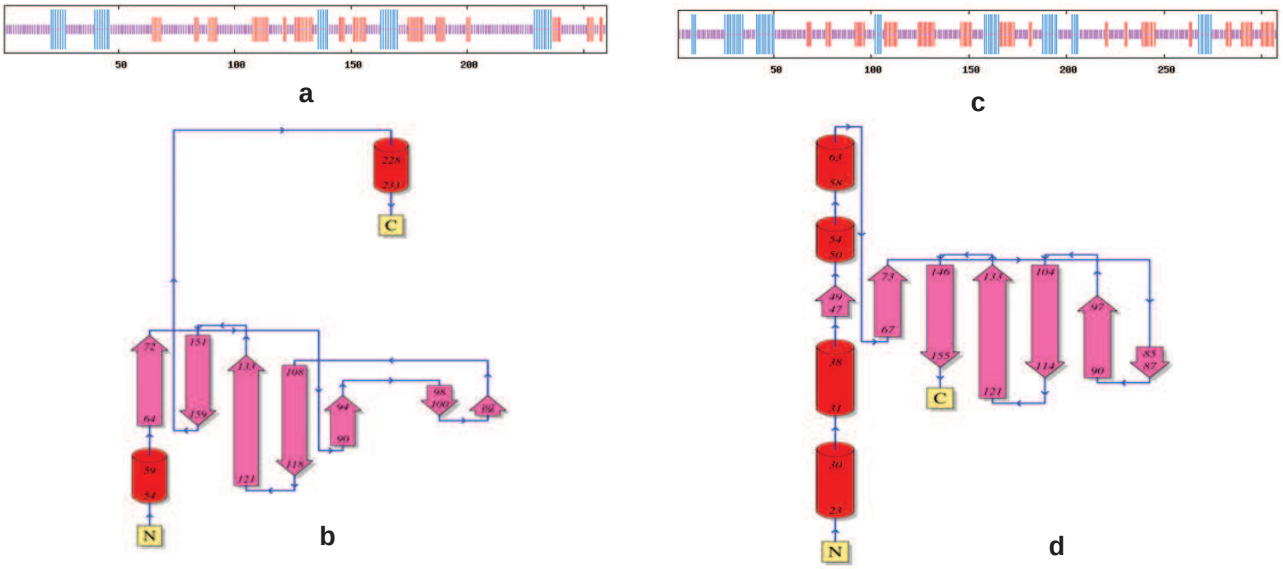



**Fig. 2. Secondary Structure and topology diagram of NAC14 (a, b) and NAC19 (c, d)**

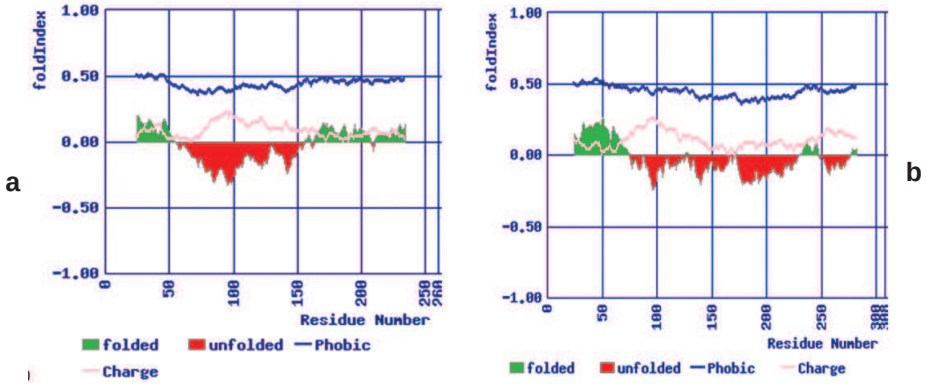



**Fig. 3. Predicted folding state of NAC14 (a) and NAC19 (b). Amino acids suggested in ordered and non-ordered regions are shown towards positive and negative numbers, respectively**

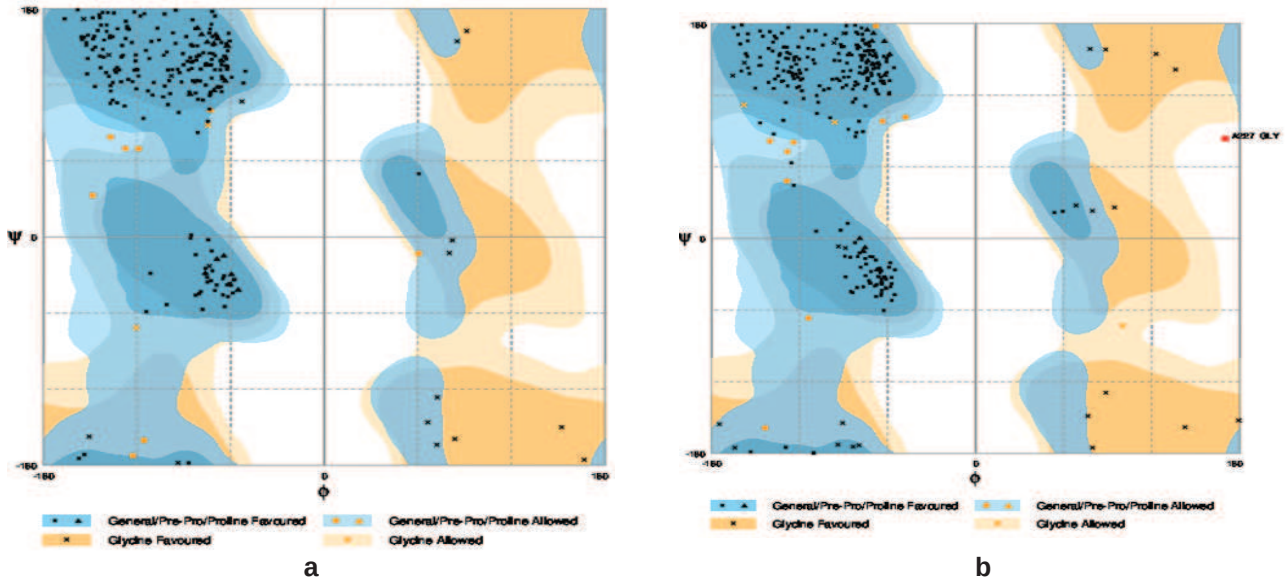



**Fig. 4. Ramachandran Plot Assessment of NAC14 (a) and NAC19 (b). The model obtained for NAC14 and NAC19 showed 96.1 and 95.8% residues in the denser core region, respectively that accounted for reliable and consistent structure**

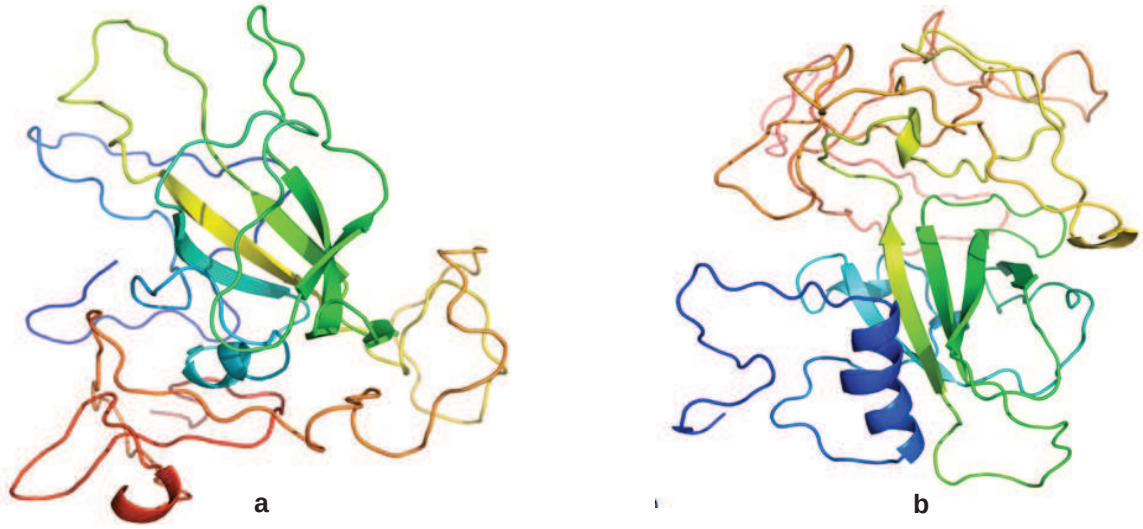



**Fig. 5. Predicted 3-D structure NAC14 (a) and NAC19 (b)**

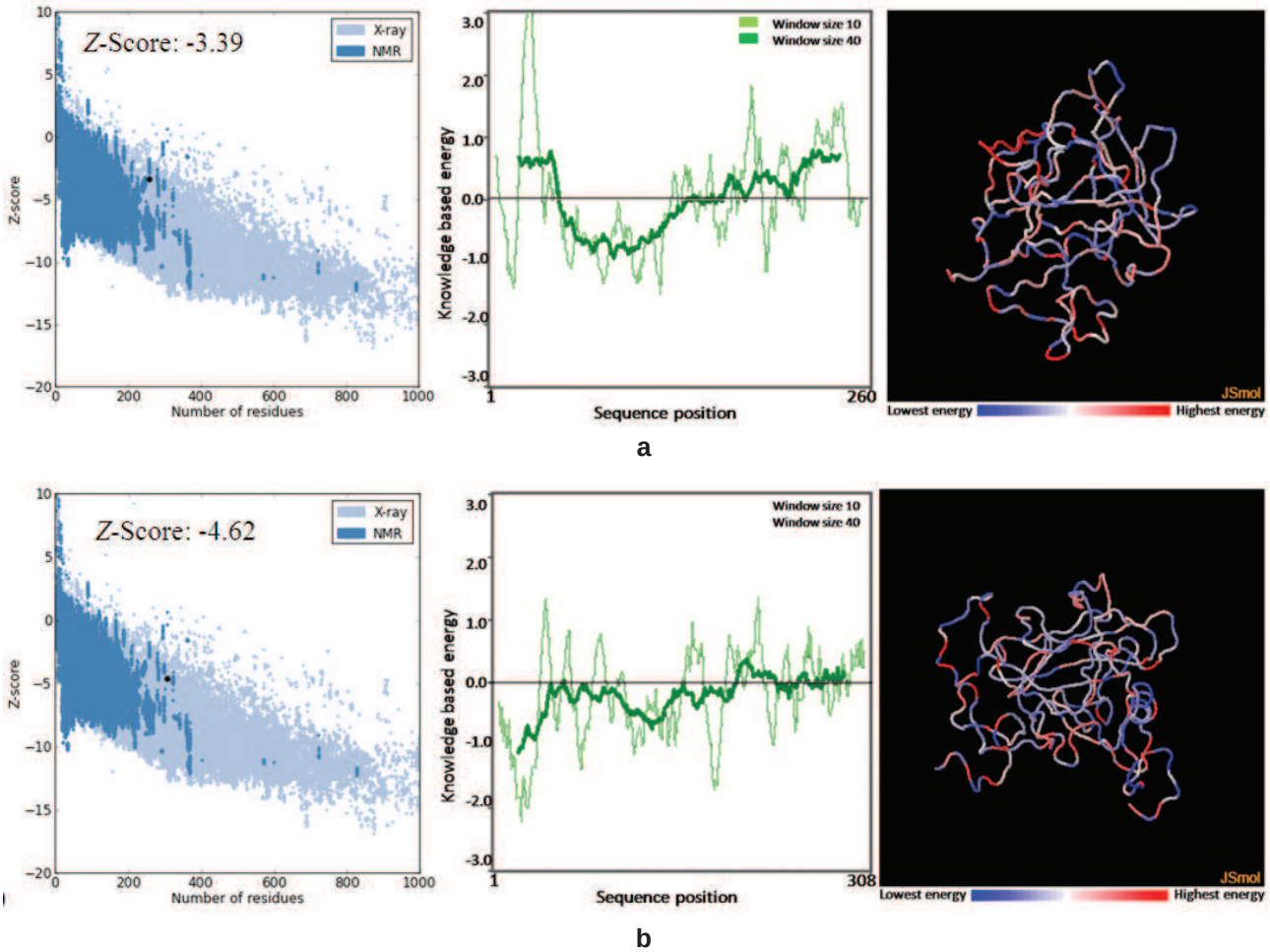



**Fig. 6. Investigation of the structure of NAC14 (a) and NAC19 (b) using the ProSA-web service. ProSA-web z-scores of all protein chains in PDB were obtained by X-ray crystallography (light blue) or NMR spectroscopy (dark blue) with reference to their length. The plot shows only chains with less than 1000 residues and a z-score d 10. The z-score is highlighted by a black dot. Screenshot of C-α trace with JSmol visualization. Residues are colored from blue to red in the order of increasing residue energy**

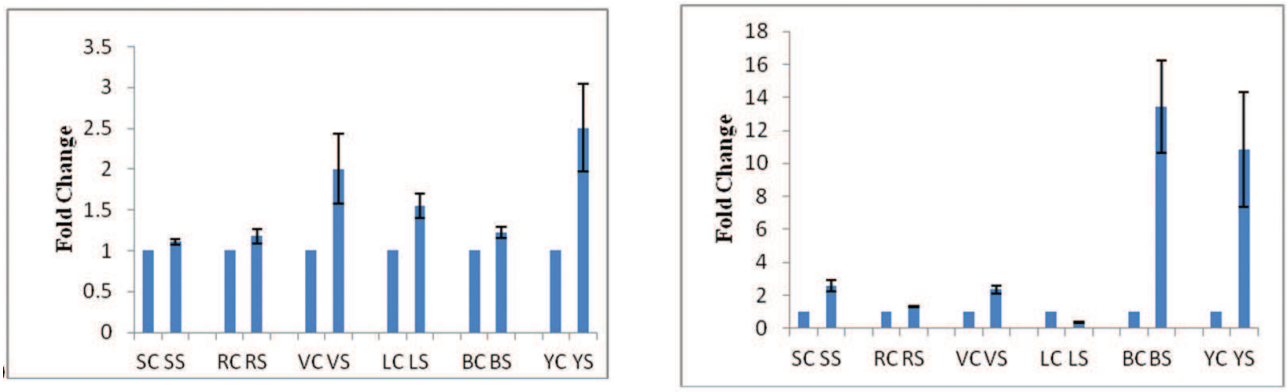



**Fig. 7. Relative quantification of *NAC14* (a) and *NAC19* (b) expression under control and drought stress conditions in *Brassica* genotypes. Actin was used as reference gene for normalization. The means are generated from three independent measurements and the bars indicate standard errors. SC-*S. alba* control, SS-*S. alba* stressed, RC-RGN73 control, RS-RGN73 stressed, VC-Varuna control, VS-Varuna stressed, LC-RLM619 control, LS-RLM619 stressed, BC-BEC144 control, BS-BEC144 stressed, YC-BioYSR control and YS-BioYSR stressed**

**Table 2.** PROCHECK statistics of predicted 3-D model of NAC14 and NAC19

| | NAC14 | | NAC19 | |
|---|---|---|---|---|
| Ramachandran plot statistics | No. of residues | % of residues | No. of residues | % of residues |
| Residues in most favored regions | 248 | 96.1% | 293 | 95.8% |
| Residues in additional allowed regions | 10 | 3.9% | 12 | 3.9% |
| Residues in disallowed regions | 0 | 0.0% | 0 | 0.3% |
| Total number of residues | 260 | | 308 | |
| G – Factors | Score | Average score | Score | Average score |
| Dihedral angles | | | | |
| Phi-Psi distribution | -0.53 | | -0.44 | |
| Chi1-Chi2 distribution | 0.30 | | 0.37 | |
| Chi1 only | 0.25 | | 0.27 | |
| Chi3 & Chi4 | 0.53 | | 0.63 | |
| Omega | -0.62 | | -0.68 | |
| Average of dihedral angles | | -0.22 | | -0.20 |
| Main-chain covalent forces | | | | |
| Main-chain bond length | 0.61 | | 0.64 | |
| Main-chain bond angles | 0.11 | 0.32 | 0.21 | 0.39 |
| Average of G-Factors | | 0.00 | | 0.04 |

towards the N-terminal and other is of H types (Asp229 to Ala232) towards the C-terminal. Concisely, the protein motif of NAC14 consist of 2 helices, 7 strands of β sheets (5A and 2B), 3 β hairpins, 2 β bulges, 24 β turns and 2 g turns [Supplementary Table 1: available online at http://www.isgpb.co.in]. NAC19 consists of 7 strands of β sheets A and 4 helices towards the N-terminal forming 1 helix-helix interaction. Out of 4 helices, 2 are of H type (Asp24 to Val29 and Leu32 to Ala37) and remaining 2 helices are of G type (Leu51 to Lys53 and Leu59 to Lys62). Overall the protein of NAC19 consists of 4 helices, 1 helix-helix interaction, 7 strands of β sheets A, 4 β hairpins, 1 β bulge, 26 β turns and 5 g turns (Supplementary Table 2: available

online at http://www.isgpb.co.in). The NAC14 and NAC19 showed considerable variation in their protein structure. Although both had 7 strands of β sheets, the NAC14 had 5 β sheets of type A and 2 β sheets of type B, while NAC14 had all 7 β sheets of type A. These proteins also differed in helix content, β turns and g turns.

The ProSA server displayed plots containing the z-score (-3.39) of NAC14 and (-4.62) of NAC19 model with comparable protein chains of PDB, indicating the reliability of the structure. The residues with the negative energies further confirmed the consistency of this predicted model (Figs. 6a and 6b). This plot showed local model quality by plotting energies as a function of amino acid sequence position (Kulkarni and Devarmath 2013).

### *Expression analysis of NAC14 and NAC19 by Q-RTPCR*

The expression pattern of *NACs* in leaves of the seedlings of *B. juncea* (RGN73*,* Varuna*,* RLM619*,* BEC144 and BioYSR) and *S. alba* were analyzed to understand the regulation by drought stress*. NAC14* showed >1.5 fold up-regulation in Varuna, RLM619 and BioYSR under drought stress (Fig. 7a). *NAC19* showed >2 fold up-regulation in the leaves of *S. alba*, Varuna, BEC144 and BioYSR under drought stress as compared with control conditions (Fig. 7b). Previous study showed that *BnaNAC14* expression is induced by mechanical wounding but not by dehydration (Hegedus et al. 2003). In this study, we found that drought stress induces the expression of *NAC14* in some *B. juncea* genotypes. Previously it was shown that *BnaNAC19* is upregulated by Paraquat (10 μM), ABA (50 μM), dehydration, NaCl (200 mM), heat (37°C) and cold (4°C) treatments (Wang et al. 2015) in *B. napus*. In Arabidopsis also, *ANAC019* was found to be upregulated by drought, salt and ABA (Jensen et al. 2010). Hence, drought induction of *NAC19* gene in *S. alba* and *B. juncea* found in this study is similar to that reported previously in *B. napus* and *Arabidopsis*. The drought inducible expression of *NAC14* was found in three out of five *B. juncea* genotypes tested. In this study also we did not observe drought induced expression of *NAC14* in RGN73, BEC144 and *S. alba* (Fig. 7a) similar to the response of *B. napus* reported previously (Hegedus et al. 2003). These results revealed genotypic differences in drought induced expression of NAC genes. Rice NAC19 was shown to be induced by ABA, MeJ and Blast fungus, suggesting its potential role in biotic stress as well (Lin et al. 2007). Thus, the NAC19 cloned in this study may play important roles in both biotic and abiotic stress tolerance.

In this study we cloned and sequenced two drought-inducible NAC TF genes (*NAC14* and *NAC19*) from five *B. juncea* cultivars and one accession of *S. alba*. *In-silico* analysis enabled us to structurally characterize the NAC protein. The cloned *NAC* genes have high sequence homology and evolutionary relationship with NAC proteins. We predicted the secondary and tertiary structures of NAC14 and NAC19, and derived important information regarding the DNA binding domains and their expression under environmental stress. The 3D structures of these TFs were predicted and validated. The results revealed that the 3D models developed were highly accurate, and will be useful for functional analysis of different domains of NAC transcription factors. The structure analysis also suggested that these two stress-inducible NAC TFs may regulate different regulons involved in stress responses of *Brassica* species. Real-time RT-PCR expression analysis revealed genotype specific regulation of *NAC14* and *NAC19*, which was unknown previously. Functional validation of these genes will help understand their role in drought tolerance of oilseed *Brassica*.

### Authors' contribution

Conceptualization of research (DP, VC, TM); Designing of the experiments (DP, VC, TM); Contribution of experimental materials (ASVM, DKY); Execution of field/lab experiments and data collection (DP); Analysis of data and interpretation (DP, BSP); Preparation of manuscript (DP, VC, TM, IR).

### Declaration

The authors declare no conflict of interest.

### Acknowledgment

The work presented in the manuscript was funded under the "Network Project on Transgenics in Crops" (NPTC) of the Indian Council of Agricultural Research (ICAR).

### References

Anisimova M. and Gascuel O. 2006. Approximate likelihood-ratio test for branches: a fast, accurate, and powerful alternative. Systematic Biol*.,* **55**: 539-552.

Brown J., Brown A.P., Davis J.B. and Erickson D. 1997. Intergeneric hybridization between *Sinapis alba* x *Brassica napus*. Euphytica, **93**:163-168.

Chen Q., Wang Q., Xiong L. and Lou Z. 2011. A structural view of the conserved domain of rice stress-responsive NAC1. Protein Cell, **2**: 55-63.

Cheng F, Liu S, Wu J, Fang L, Sun S, Liu B, Li P, Hua W, Wang X. 2011. BRAD, the genetics and genomics database for *Brassica* plants. BMC Plant Biol., **11**: 136.

Darzentas N. 2010. Circoletto: visualizing sequence similarity with Circos. Bioinformatics, **26**: 2620-2621.

Fang Y., Liao K., Du H., Xu Y., Song H., Li X. and Xiong L. 2015. A stress-responsive NAC transcription factor SNAC3 confers heat and drought tolerance through modulation of reactive oxygen species in rice. J. Exp. Bot., **66**: 6803-17.

Fiser A. and Sali A. 2003. ModLoop: automated modeling of loops in protein structures. Bioinformatics, **19**: 2500-2501.

Foster M.J. 2002. Molecular Modelling on Structural Biology. Micron., **33**: 365-384.

Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M. R., Appel R. D., Bairoch A. 2005. *Protein Identification and Analysis Tools on the ExPASy Server;* (In) John M. Walker (ed): The Proteomics Protocols Handbook, Humana Press, 571-607.

Gouet P., Courcelle E., Stuart D. I. and Metoz F. 1999. ESPript: multiple sequence alignments in PostScript. Bioinformatics, **15**: 305-308.

Gouet P., Robert X. and Courcelle E. 2003. ESPript/ ENDscript: extracting and rendering sequence and 3D information from atomic structures of proteins. Nucleic Acids Res., **31**: 3320-3323.

Guruprasad K., Reddy B.V.B. and Pandit M.W. 1990. Correlation between stability of a protein and its dipeptide composition- a novel approach for predicting in vivo stability of a protein from its primary sequence. Protein Eng. Protein Eng., **4**: 155-161.

Hegedus D., Yu M., Baldwin D., Gruber M., Sharpe A., Parkin I., Whitwill S. and Lydiate D. 2003. Molecular characterization of *Brassica napus* NAC domain transcriptional activators induced in response to biotic and abiotic stress. Plant Mol. Biol., **53**: 383-397.

Hong Y., Zhang H., Huang L., Li D., Song F. 2016. Overexpression of a stress-responsive NAC transcription factor gene ONAC022 improves drought and salt tolerance in rice. Front Plant Sci., **22**: 7-4.

Hu H., Dai M., Yao J., Xiao B., Li X., Zhang Q. and Xiong L. 2006. Overexpressing a NAM, ATAF, and CUC (NAC) transcription factor enhances drought resistance and salt tolerance in rice. Proc. Natl. Acad. Sci. U. S. A., **103:** 12987-12992.

Ikai A. 1980. Thermostability and aliphatic index of globular proteins. J. Biochem., **88**: 1895-1898.

Jensen M. K., Kjaersgaard T. and Nielsen M. M. 2010. The *Arabidopsis thaliana* NAC transcription factor family: structure-function relationships and determinants of ANAC019 stress signaling. Biochem. J., **426**: 183-96.

Jeong J. S., Kim Y. S., Baek K. H., Jung H., Ha S. H., Do Choi Y. 2010. Root-specific expression of OsNAC10 improves drought tolerance and grain yield in rice under field drought conditions. Plant Physiol., **153**: 185-197.

Katoh K., Misawa K., Kuma K. and Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res., **30**: 3059-3066.

Katoh and Standley. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol., **30**:772-780.

Kikuchi K., Ueguchi-Tanaka M., Yoshida K. T., Nagato Y., Matsusoka M. and Hirano H. Y. 2000. Molecular analysis of the NAC gene family in rice. Mol. Gen. Genet., **262:** 1047-1051.

Kim H. G., Nam H. G. and Lim P. O. 2016a. Regulatory network of NAC transcription factors in leaf senescence. Curr. Opin. Plant Biol., **33**:48-56.

Kim S. Y., Kim S. G. and Kim Y. S. 2007b. Exploring membrane- associated NAC transcription factors in *Arabidopsis*: implications for membrane biology in genome regulation. Nucleic Acids Res., **35**: 203-213.

Krzywinski M., Schein J., Biroli., Connors J., Gascoyne R., Horsmon D., Jones S. J. and Marra M. A. 2009.Circos: an information aesthetic for comparative genomics. Genome Res., **19**: 1639-1645.

Kulkarni P. A. and Devarumath R. M. 2014. *In silico* 3D-structure prediction of SsMYB2R: A novel MYB transcription factor from *Saccharum spontaneum*. Indian J. Biotech., **13**: 437-447.

Le D.T., Nishiyama R., Watanabe Y., Mochida K., Yamaguchi-Shinozaki K., Shinozaki K. and Tran L. S. 2011. Genome-wide survey and expression analysis of the plant-specific NAC transcription factor family in soybean during development and dehydration stress. DNA Res., **18**: 263-276.

Letunic I. and Bork P. 2007. Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. Bioinformatics, **23**: 127-128.

Lin R., Zhao W., Meng X., Wang M. and Peng Y. 2007. Rice gene OsNAC19 encodes a novel NAC-domain transcription factor and responds to infection by *Magnaporthe grisea*. Plant Science, **172**: 120-130.

Liu X., Hong L., Li X. Y., Yao Y., Hu B. and Li L. 2011. Improved drought and salt tolerance in transgenic Arabidopsis overexpressing a NAC transcriptional factor from *Arachis hypogaea*. Biosci. Biotechnol. Biochem., **75**: 443-50.

Liu T., Song x., Duan W., Huang Z., Liu G. and Hou X. 2014. Genome-wide analysis and expression patterns of NAC transcription factor family under different developmental stages and abiotic stresses in Chinese cabbage. Plant Mol. Biol. Rep., **32**: 1041-1056.

Liu G. Z., Li X. L., Jin S. X., Liu X. Y., Zhu L. F., Nie Y. C., Zhang X. L. 2014. Overexpression of rice NAC Gene SNAC1 improves drought and salt tolerance by enhancing root development and reducing transpiration rate in transgenic cotton. PLoS One, **9**: 86895.

Lu M., Ying S., Zhang D. F., Shi Y. S., Song Y. C., Wang T. Y., et al. 2012. A maize stress-responsive NAC transcription factor, ZmSNAC1, confers enhanced tolerance to dehydration in transgenic *Arabidopsis*. Plant Cell Rep., **31**: 1701-1711.

Mao X., Zhang H., Qian X., Li A., Zhao G., Jing R. 2012. TaNAC2, a NAC-type wheat transcription factor conferring enhanced multiple abiotic stress tolerances in *Arabidopsis*. J. Exp. Bot., **63**: 2933-•–2946.

Nakashima K., Tran L. S., Van Nguyen D., Fujita M., Maruyama K. and Todaka D. 2007. Functional analysis of a NAC-type transcription factor OsNAC6 involved in abiotic and biotic stress-responsive gene expression in rice. Plant J., **51**: 617-630.

Nakashima K., Takasaki H., Mizoi J., Shinozaki K., Yamaguchi-Shinozaki K. 2012. NAC transcription factors in plant abiotic stress responses. Biochem. Biophys. Acta, **1819**: 97-103.

Niu F., Wang B., Wu F., Yan J., Li L., Wang C., Wang Y., Yang B. and Jiang Y.Q. 2014. Canola (*Brassica napus* L.) NAC103 transcription factor gene is a novel player inducing reactive oxygen species accumulation and cell death in plants. Biochem. Biophys. Res. Commun., **454**: 30-5.

Nuruzzaman M., Shaoni A. M. and Kikuchi S. 2013. Roles of NAC transcription factors in the regulation of biotic and abiotic stress responses in plants. Front. Microbiol., **4**: 248.

Ooka H., Satoh K. and Doi K. 2003. Comprehensive analysis of NAC family genes in *Oryza sativa* and *Arabidopsis thaliana.* DNA Res., **10**: 239-247.

Palit P., Nagaich D., Phukan D., Goyal V., Kumar S., Yadava D. K., Bhat S. R. and Mohapatra T. 2012. Identification of moisture stress responsive genes through suppression subtractive hybridization in *Brassica juncea* and *Sinapis alba*. Proc. International Conference on Plant Biotechnology for food security: new frontiers, National Research Centre on Plant Biotechnology, NASC Complex, New Delhi, India.

Pereira-Santana A., Alcaraz L. D., Castano E., Sanchez-Calderon L., Sanchez-Teyer f. and Rodriguez-Zapata L. 2015. Comparative genomics of NAC transcription factors in angiosperms: implications for the adaption and diversification of flowering plants. PLoS One, **10**(11): e141866.

Phukan D., Goyal V., Palit P., Kalia R., Koundal M., Mithra S. V. A., Ravi I., Yadava D. K., Chinnusamy V. and Mohapatra T. 2016. Expression analysis of candidate genes for abiotic stress tolerance in oilseed *Brassica* genotypes with contrasting osmotic stress tolerance. Indian J. Exp. Biol., (Accepted)

Puranik S., Sahu P. P., Srivastava P. S. and Prasad M. 2012. NAC proteins: regulation and role in stress tolerance. Trends Plant Sci., **17**: 369-381.

Reehana N., Ahamed A. P., Ali D. M., Suresh A., Kumar R. A. and Thajuddin N. 2013. Structure based computational analysis and molecular phylogeny of C-Phycocyanin gene from the selected cynobacteria. International J. Biological, Veterinary, Agric. Food Eng., **7**: 47-51.

Robert X. and Gouet P. 2014. Deciphering key features in protein structures with the new ENDscript server. Nucl. Acids Res., **42**: W320-W324.

Roy A., Kucukural A. and Zhang Y. 2010. I-TASSER: a unified platform for automated protein structure and function prediction. Nature Protocol, **5**: 725-738.

Sakuraba Y., Piao W., Lim J. H., Han S. H., Kim Y. S., An G. and Paek N. C. 2015. Rice ONAC106 inhibits leaf senescence and increases salt tolerance and tiller angle. Plant Cell Physiol., **56**: 2325-39.

Satoh K., Shimizu T., Kondoh H., Hiraguri A., Sasaya T. and Choi I. R. 2011. Relationship between symptoms and gene expression induced by the infection of three strains of riced warf virus. PLoS One, **3**: e18094.

Shao H., Wang H. and Tang X. 2015. NAC transcription factors in plant multiple abiotic stress responses: progress and prospects. Front Plant Sci., **29**: 902.

Song S. Y., Chen Y., Chen J., Dai X. Y. and Zhang W. H. 2011. Physiological mechanisms underlying OsNAC5-dependent tolerance of rice plants to abiotic stress. Planta, **234**: 331-345.

Sperotto R. A., Ricachenevsky F. K., Duarte G. L., Boff T., Lopes K. L. and Sperb E. R. 2009. Identification of up-regulated genes in flag leaves during rice grain filling and characterization of OsNAC5, a new ABA-dependent transcription factor. Planta, **230**: 985-1002.

Tran L. S. P., Nakashima K. and Sakuma Y. 2004. Isolation and functional analysis of *Arabidopsis* stress-inducible NAC transcription factors that bind to a drought responsive cis-element in the early responsive to dehydration stress 1 promoter. Plant Cell, **16**: 2481-98.

Wang B., Guo X., Wang C., Ma J., Niu F., Zhang H., Yang B., Liang W., Han F. and Jiang Y. Q. 2015. Identification and characterization of plant-specific NAC gene family in canola (*Brassica napus* L.) reveal

novel members involved in cell death. Plant Mol Biol., **87**: 395-411.

Warwick S. I. 1993. Guide to the wild germplasm of Brassica and allied crops. Part IV. Wild species in the Tribe *Brassiceae* (Cruciferae) as sources of agronomic traits. Agriculture Canada Research Branch, Technical Bulletin 1993-17E, Ottawa, Canada. 19 pp.

Xie Q., Sanz-Burgos A. P., Guo H., García J. A. and Gutiérrez C. 1999. GRAB proteins, novel members of the NAC domain family, isolated by their interaction with a geminivirus protein. Plant Mol. Biol., **39**: 647-656.

Xue G. P., Wa, H. M., Richardson T., Drenth J., Joyce P. A. and McIntyre C. L. 2011. Overexpression of TaNAC69 leads to enhanced transcript levels of stress up-regulated genes and dehydration tolerance in bread wheat. Mole. Plant, **4**: 697-712.

Yadava J. S. and Singh N. B. 1999. Proc. 10th International Rapeseed Congress, Canberra, Australia.

Ying L., Chen H. and Cai W. 2014. BnNAC485 is involved in abiotic stress responses and flowering time in *Brassica napus*. Plant Physiol. Biochem., **79**: 77-87.

You J., Zong W., Hu H., Li X., Xiao J. and Xiong L. 2014. A STRESS-RESPONSIVE NAC1-regulated protein phosphatase gene rice protein phosphatase18 modulates drought and oxidative stress tolerance through abscisic acid-independent reactive oxygen species scavenging in rice. Plant Physiol., **166**: 2100-14.

Zheng X., Chen B., Lu G. and Han B. 2009. Overexpression of a NAC transcription factor enhances rice drought and salt tolerance. Biochem. Biophys. Res. Commun., **379**: 985-989.

Zhong H., Guo Q. Q., Chen L., Ren F., Wang Q. Q., Zheng Y. and Li X. B. 2012. Two Brassica napus genes encoding NAC transcription factors are involved in response to high-salinity stress. Plant Cell Rep., **31**: 1991-2003.

**Supplementary Table 1.** Secondary structure details of NAC14 (a-e)

(a) Table of helices

| No. | Start | End | Type | No. resid | Length | Unit rise | Residues/ turn | Pitch | Deviation from ideal | Sequence |
|---|---|---|---|---|---|---|---|---|---|---|
| 1. | Leu55 | Met58 | G | 4 | 7.46 | 1.88 | 3.33 | 6.24 | 41.0 | LPDM |
| 2. | Asp229 | Ala232 | H | 4 | 6.31 | 1.40 | 3.89 | 5.46 | 40.7 | DAIA |

| No. | Start | End | Sheet | No. resid | Edge | Sequence |
|---|---|---|---|---|---|---|
| 1. | Glu65 | Arg72 | A | 8 | Yes | EWYFFSPR |
| 2. | Tyr90 | Thr94 | A | 5 | Yes | YWKAT |
| 3. | Pro99 | Ile100 | B | 2 | Yes | PI |
| 4. | Val106 | Gly107 | B | 2 | Yes | VG |
| 5. | Lys109 | Gly117 | A | 9 | No | KKALVFYSG |
| 6. | Gly122 | Tyr132 | A | 11 | No | GEKTNWIMHEY |
| 7. | Asp151 | Tyr158 | A | 8 | No | DWVLCRIY |

(b) Table of beta hairpins

| | Strand 1 | | | Strand 2 | | | Hairpin |
|---|---|---|---|---|---|---|---|
| No. | Start | End | Length | Start | End | Length | Class |
| 1. | Pro99 | Ile100 | 2 | Val106 | Gly107 | 2 | 5:5 |
| 2. | Lys109 | Gly117 | 9 | Gly122 | Tyr132 | 11 | 2:4 |
| 3. | Gly122 | Tyr132 | 11 | Asp151 | Tyr158 | 8 | 21:21 |

(c) Table of beta bulges

| No. | Bulge type | Res X | Res 1 | Res 2 | Res 3 | Res 4 |
|---|---|---|---|---|---|---|
| 1. | Antiparallel classic | Lys98A | Ile108A | Lys109A | | |
| 2. | Antiparallel special | Phe114A | Thr125A | Asn126A | Trp127A | |

(d) Table of beta turns

| No. | Turn | Residue | Seq. | Turn type | Residue phi | i+1 psi | Chi | Residue phi | i+2 psi | i to i+3 CA-dist | H-bond |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. | Pro11-Phe14 | PPGF | II | -65.5 | 139.4 | 27.3 | 82.2 | -2.7 | - | 5.5 | No |
| 2. | Thr19-Glu22 | TDEE | IV | -71.6 | -26.7 | -71.5 | -77.4 | 145.1 | -74.5 | 6.7 | Yes |
| 3. | Tyr27-Arg30 | YLCR | VIII | -82.3 | -45.9 | -68.9 | -151.9 | 128.5 | -172.3 | 6.7 | Yes |
| 4. | Ile37-Pro40 | IAAP | IV | -58.0 | -33.7 | - | -151.3 | 159.0 | - | 6.8 | Yes |
| 5. | Asp51-Asp54 | DAWD | I | -72.1 | -11.1 | - | -82.3 | -14.0 | -85.3 | 5.2 | No |
| 6. | Ala52-Leu55 | AWDL | I | -82.3 | -14.0 | -85.3 | -111.2 | -28.8 | -69.5 | 5.4 | Yes |
| 7. | Pro77-Ser80 | PNGS | II | -62.1 | 132.2 | -167.8 | 80.4 | -12.7 | - | 5.8 | No |
| 8. | Gly79-Pro82 | GSRP | VIII | -80.1 | -44.7 | -73.3 | -130.1 | 122.3 | -66.0 | 6.7 | Yes |
| 9. | Arg102-Pro105 | RPKP | VIII | -68.6 | -17.0 | 31.0 | -100.8 | 145.8 | -54.6 | 6.6 | Yes |
| 10. | Lys118-Arg121 | KPPR | I | -55.1 | -41.1 | -25.3 | -76.1 | -28.0 | 31.9 | 5.7 | Yes |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 11. | Pro119-Gly122 | PPRG | VIII | -76.1 | -28.0 | 31.9 | -135.4 | 121.7 | -73.2 | 6.8 | Yes |
| 12. | Arg142-Asn145 | RKGN | IV | -69.9 | 174.1 | -52.8 | -52.7 | 106.7 | - 6.8 | | Yes |
| 13. | Arg148-Asp151 | RLDD | IV | -59.7 | -38.2 | -76.5 | -77.7 | 107.9 | -62.4 | 6.1 | Yes |
| 14. | Lys161-Ile164 | KGVI | I | -73.6 | -22.9 | - | -62.2 | -37.8 | -176.8 | 5.4 | Yes |
| 15. | Gly162-Glu165 | GVIE | VIII | -62.2 | -37.8 | -176.8 | -127.9 | 103.2 | -51.6 | 6.5 | Yes |
| 16. | Glu165-Arg168 | EKRR | IV | -69.0 | -37.9 | -65.1 | -75.5 | 136.5 | -169.2 | 6.6 | Yes |
| 17. | Leu189-Gly192 | LIGG | IV | -77.9 | -57.1 | -60.7 | -93.6 | -178.0 | - 6.9 | | Yes |
| 18. | Arg200-Ser203 | RVVS | IV | -74.8 | -35.5 | -178.0 | -63.7 | -54.9 | -168.0 | 6.0 | Yes |
| 19. | Ser203-Phe206 | SPEF | IV | -59.4 | -29.0 | -24.8 | -139.4 | 162.0 | -65.8 | 6.6 | Yes |
| 20. | Asn219-Glu222 | NALE | IV | -64.8 | -45.7 | - | -122.6 | -172.4 | -56.9 | 6.7 | Yes |
| 21. | Glu235-Ser238 | EIVS | IV | -100.9 | -53.3 | -53.2 | -141.4 | 141.7 | 177.6 | 5.8 | Yes |
| 22. | Arg239-Gly242 | RLLG | I | -60.7 | -37.5 | -173.1 | -72.5 | -34.5 | -65.0 | 5.4 | Yes |
| 23. | Leu240-Gly243 | LLGG | IV | -72.5 | -34.5 | -65.0 | -120.4 | -71.5 | -6.4 | | Yes |
| 24. | Val254-Gln257 | VVRQ | IV | -85.3 | 151.5 | -179.9 | 60.8 | 49.8 | -63.7 | 5.9 | Yes |

(e) Table of gamma turns

| No. | Start | End | Sequence | Turn type | Residue i+1 | | | i to i+2 |
|---|---|---|---|---|---|---|---|---|
| | | | | | Phi | Psi | Chi1 | |
| 1. | Ile100 | Arg102 | I G R | INVERSE | 74.4 | 87.9 | - | 5.9 |
| 2. | Asn183 | Asp185 | N F D | INVERSE | 81.0 | 82.3 | -161.8 | 5.7 |

**Supplementary Table 2.** Secondary structure details of NAC19 (a-f)

(a) Table of helices

| No. | Start | End | Type | No. resid | Length | Unit rise | Residues/ turn | Pitch | Deviation from ideal | Sequence |
|---|---|---|---|---|---|---|---|---|---|---|
| 1. | Asp24 | Val29 | H | 6 | 9.09 | 1.46 | 3.51 | 5.13 | 6.8 | DEELMV |
| 2. | Leu32 | Ala37 | H | 6 | 9.30 | 1.48 | 3.53 | 5.23 | 4.0 | LCRKAA |
| 3. | Leu51 | Lys53 | G | 3 | - | - | - | - | - | LYK |
| 4. | Leu59 | Lys62 | G | 4 | 8.75 | 2.36 | 2.87 | 6.78 | 37.7 | LPSK |

(b) Beta strands

| No. | Start | End | Sheet | No. resid | Edge | Sequence |
|---|---|---|---|---|---|---|
| 1. | Ala47 | Glu48 | A | 2 | Yes | AE |
| 2. | Val68 | Pro72 | A | 5 | No | VFFCP |
| 3. | Val85 | Ala86 | A | 2 | Yes | VA |
| 4. | Lys91 | Ile96 | A | 6 | No | KLRVRI |
| 5. | Gly105 | Ile114 | A | 10 | No | GIKKALVFYI |
| 6. | Thr121 | Leu132 | A | 12 | No | TKTNWIMHEYRL |
| 7. | Trp147 | Lys154 | A | 8 | No | WVLCRIYK |

(c) Table of beta hairpins

| No. | Strand 1 | | | Strand 2 | | | Hairpin |
|---|---|---|---|---|---|---|---|
| | Start | End | Length | Start | End | Length | Class |
| 1. | Val85 | Ala86 | 2 | Lys91 | Ile96 | 6 | 2:4 |
| 2. | Lys91 | Ile96 | 6 | Gly105 | Ile114 | 10 | 9:11 |
| 3. | Gly105 | Ile114 | 10 | Thr121 | Leu132 | 12 | 6:6 |
| 4. | Thr121 | Leu132 | 12 | Trp147 | Lys154 | 8 | 18:18 |

(d) Table of beta bulges

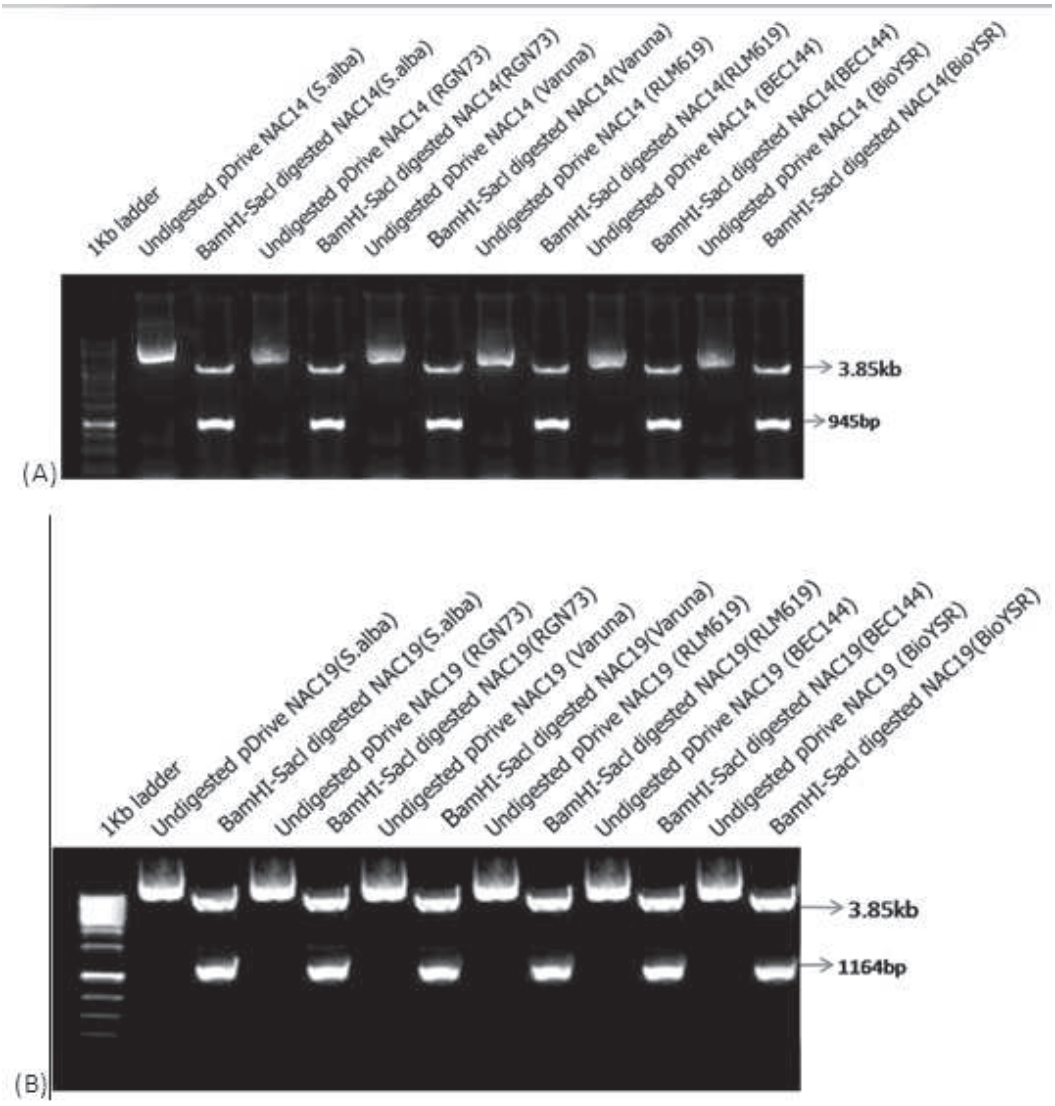
| No. | Bulge type | Res X | Res 1 | Res 2 | Res 3 | Res 4 |
|---|---|---|---|---|---|---|
| 1. | Antiparallel special | Phe112A | Thr123A | Asn124A | Trp125A | |

(e) Table of beta turns

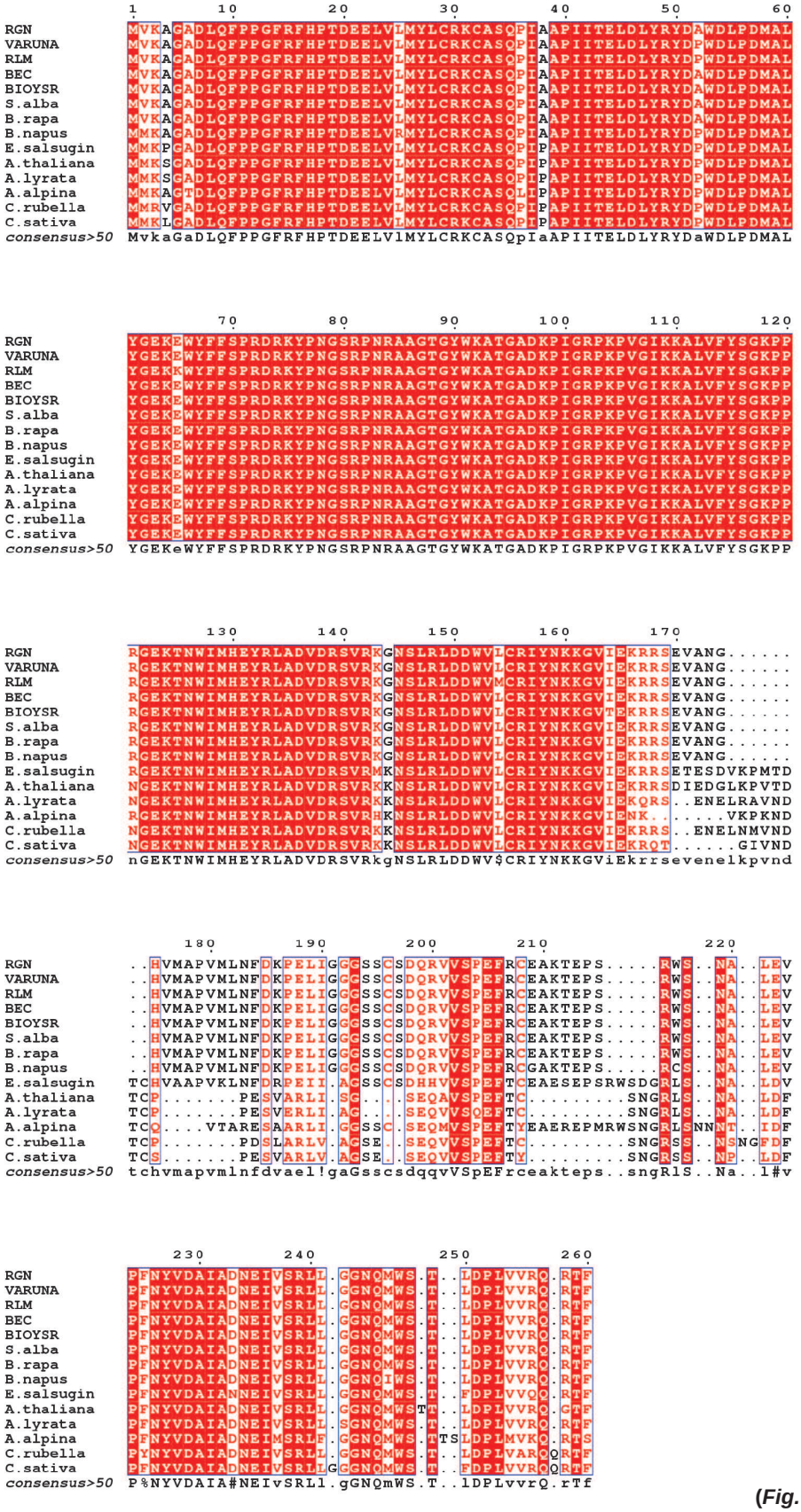
| No. | Turn | Residue | Seq. | Turn type | Residue phi | i+1 psi | Chi | Residue phi | i+2 psi | i to i+3 CA-dist | H-bond |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *1. | Gln11-Leu14 | QLSL | VIII | -64.4 | -37.1 | -73.8 | -123.9 | 80.6 | -58.6 | 6.4 | Yes |
| *2. | Pro15-Phe18 | PPGF | I | -75.3 | -7.1 | 32.7 | -88.6 | -8.0 | - | 6.0 | No |
| *3. | Phe41-Gln44 | FSLQ | VIII | -66.6 | -31.8 | 30.6 | -146.3 | 145.0 | 176.6 | 6.5 | Yes |
| *4. | Gln44-Ala47 | QLIA | VIII | -76.4 | -40.5 | -69.4 | -107.2 | 114.0 | -56.2 | 6.5 | Yes |
| *5. | Asp55-Val58 | DPWV | I | -76.2 | -10.3 | 32.2 | -89.0 | -14.9 | 53.2 | 5.6 | No |
| 6. | Pro56-Leu59 | PWVL | IV | -89.0 | -14.9 | 53.2 | -80.6 | -41.1 | 179.6 | 5.3 | Yes |
| 7. | Leu64-Glu67 | LFGE | II | -60.5 | 150.7 | -71.3 | 94.8 | 26.0 | - | 6.7 | Yes |
| 8. | Gln78-Lys81 | QTVK | IV | -78.2 | -24.1 | 55.2 | -104.1 | 161.9 | -58.2 | 6.9 | Yes |
| 9. | Gly87-Gly90 | GQYG | VIII | -65.6 | -32.1 | -172.7 | -133.1 | 154.8 | -55.6 | 5.3 | Yes |
| 10. | Lys116-Lys119 | KAPK | IV | -57.0 | -51.9 | - | -79.5 | 161.2 | 34.9 | 6.6 | Yes |
| 11. | Ala117-Gly120 | APKG | IV | -79.5 | 161.2 | 34.9 | 59.5 | 22.7 | -48.0 | 6.1 | No |
| 12. | Ser142-Asp145 | SKLD | IV | -63.8 | -43.4 | -176.6 | -113.8 | 173.9 | -59.1 | 6.4 | Yes |
| 13. | Gln162-Glu165 | QAYE | IV | -65.1 | -24.8 | - | -67.6 | -48.6 | -53.8 | 5.8 | Yes |
| 14. | Ala163-His166 | AYEH | I | -67.6 | -48.6 | -53.8 | -74.5 | -21.6 | 62.8 | 4.9 | Yes |
| 15. | Arg172-Ser175 | RELS | IV | -68.5 | -39.7 | -61.5 | -123.8 | 44.5 | -55.4 | 5.3 | No |
| 16. | Val192-Ser195 | VLDS | I | -61.7 | -32.5 | -176.6 | -60.9 | -34.1 | -73.4 | 5.9 | No |
| 17. | Leu193-Leu196 | LDSL | IV | -60.9 | -34.1 | -73.4 | -85.0 | -44.4 | 14.5 | 4.8 | Yes |
| 18. | Leu196-Glu199 | LHHE | VIII | -80.2 | -15.1 | -67.3 | -147.4 | 144.8 | -170.1 | 6.6 | Yes |
| 19. | Ser215-Pro218 | SLRP | VIII | -57.1 | -37.8 | -177.4 | -110.5 | 158.4 | -56.3 | 5.9 | Yes |
| 20. | Phe237-Gly240 | FDWG | I | -70.1 | -15.4 | -71.2 | -79.0 | -12.1 | -68.6 | 5.6 | No |
| 21. | Val246-Asn249 | VEHN | I | -55.0 | -26.3 | -65.5 | -67.9 | -18.1 | -59.3 | 5.8 | No |
| 22. | Pro261-Glu264 | PSLE | VIII | -64.1 | -27.1 | 63.0 | -128.0 | 127.7 | -63.9 | 5.5 | Yes |
| 23. | Tyr269-Met272 | YLKM | VIII | -71.9 | -37.5 | -174.0 | -86.4 | 135.2 | -173.1 | 6.8 | Yes |
| 24. | Asp281-Phe284 | DFGF | IV | -123.6 | 151.6 | -66.2 | 88.9 | -129.1 | - | 6.9 | Yes |
| 25. | Phe282-Ala285 | FGFA | IV | 88.9 | -129.1 | - | -67.5 | 123.0 | -179.2 | 6.3 | Yes |
| 26. | Gly303-Gly306 | GGFG | IV | -93.4 | -6.6 | - | -155.1 | 171.2 | 57.0 | 6.8 | Yes |

(f) Table of gamma turns

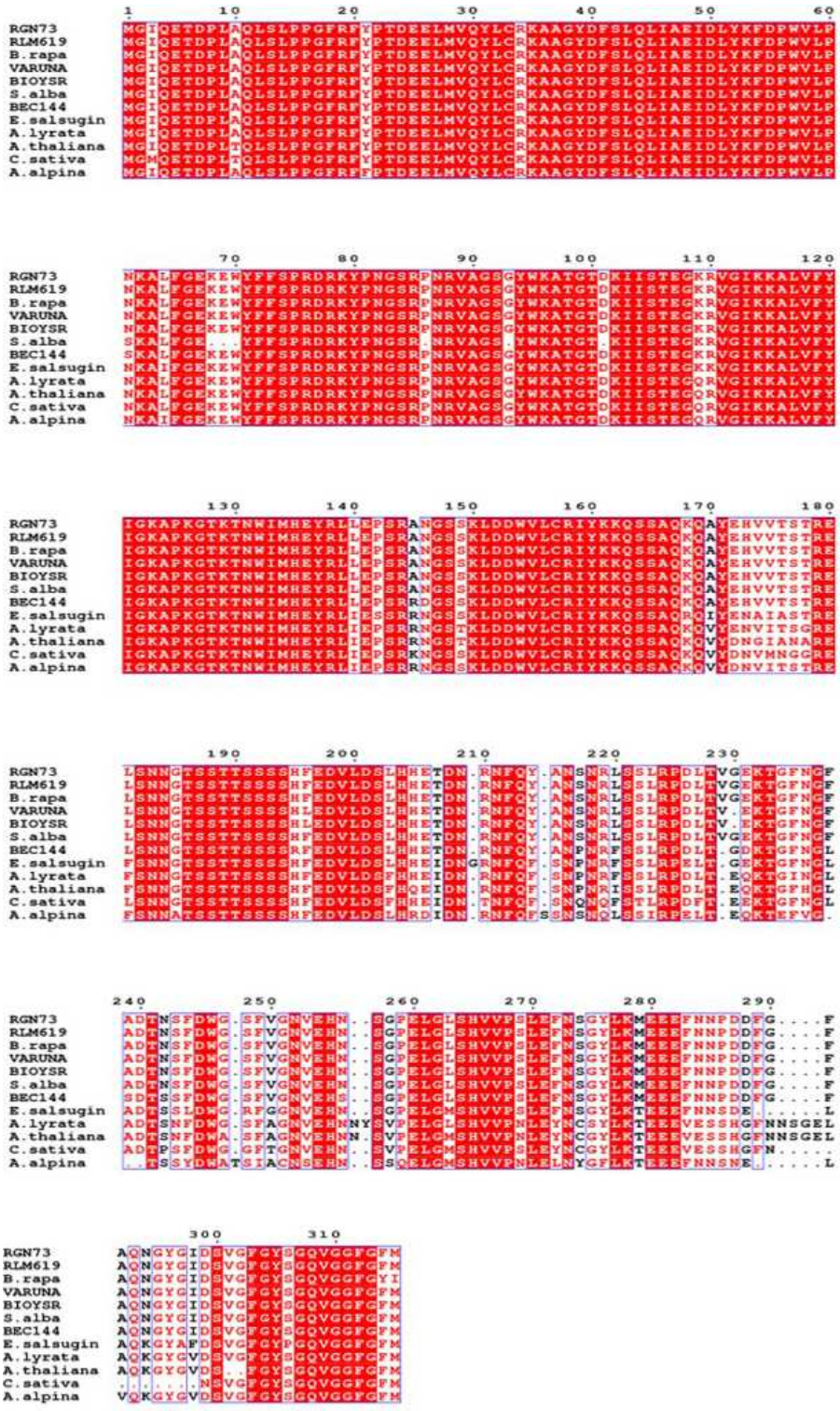
| No. | Start | End | Sequence* | Turn type | Residue i+1 | | | i to i+2 |
|---|---|---|---|---|---|---|---|---|
| | | | | | Phi | Psi | Chi1 | CA-dist |
| 1. | Arg19 | Tyr21 | R F Y | INVERSE | -86.6 | 91.4 | -161.6 | 6.0 |
| 2. | Leu43 | Leu45 | L Q L | INVERSE | -76.7 | 89.5 | -170.5 | 5.7 |
| 3. | Ile49 | Leu51 | I D L | INVERSE | -81.8 | 85.3 | -167.5 | 5.6 |
| 4. | Lys102 | Val104 | K R V | INVERSE | -81.1 | 88.7 | -172.1 | 5.8 |
| 5. | Lys271 | Glu273 | K M E | INVERSE | -79.3 | 98.1 | 179.8 | 5.7 |



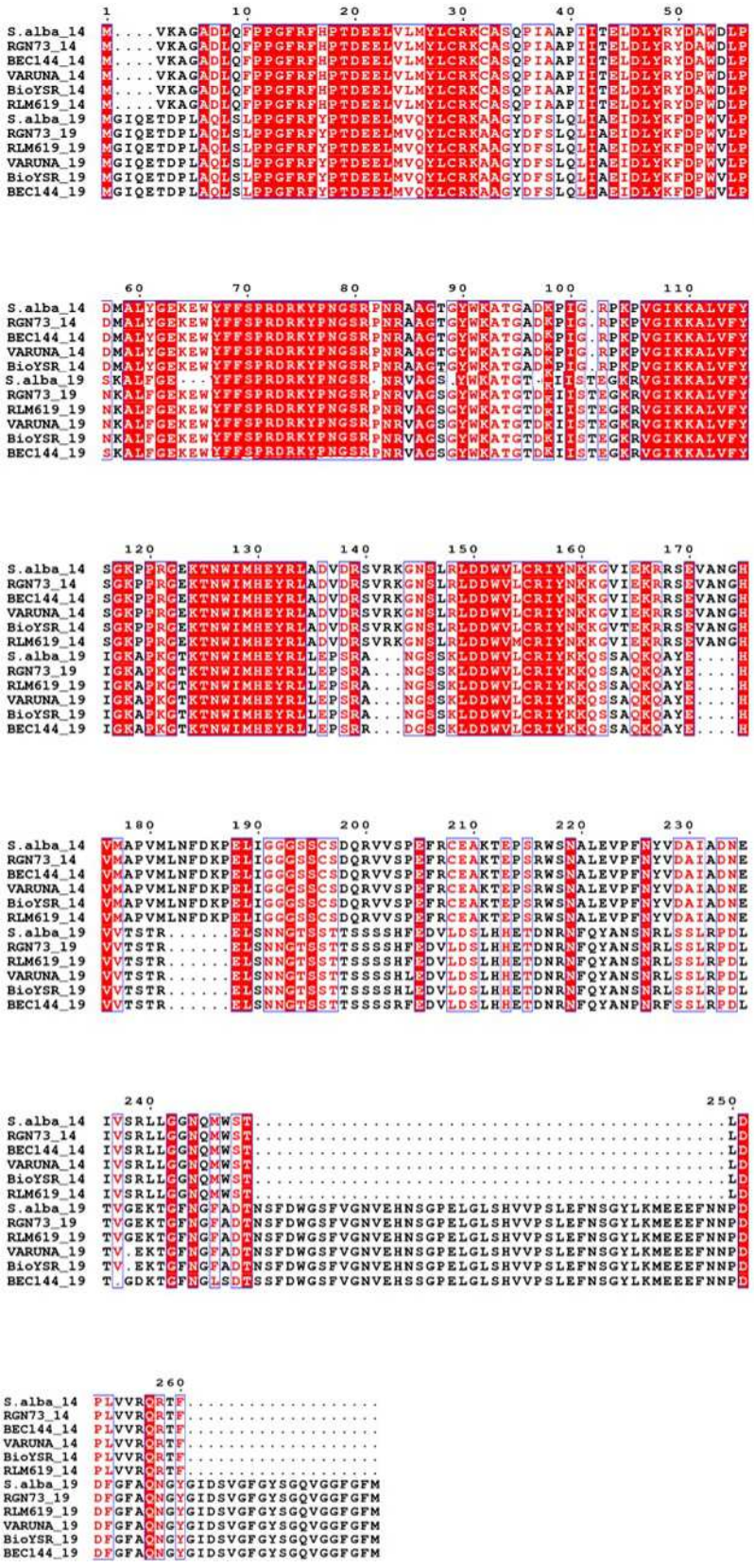

**Supplementary Fig. 1: Full length genes *NAC14* (A) and *NAC19* (B) from *S. alba,* RGN73, Varuna, RLM619, BEC144 and BioYSR. The PCR amplicons were cloned in pDrive vector and the inserts were released by BamHI and SacI double digestion. [GenBank accession numbers: *Sinapis alba* NAC14 (KT281870), RGN73 NAC14 (KT281871), Varuna NAC14 (KT281872), RLM619 NAC14 (KT281873), BEC144 NAC14 (KT281874), BioYSR NAC14 (KT281875), *Sinapis alba* NAC19 (KT281876), RGN73 NAC19 (KT281877), Varuna NAC19 (KT281878), RLM619 NAC19 (KT281879), BEC144 NAC19 (KT281880), BioYSR NAC19 (KT281881)]**



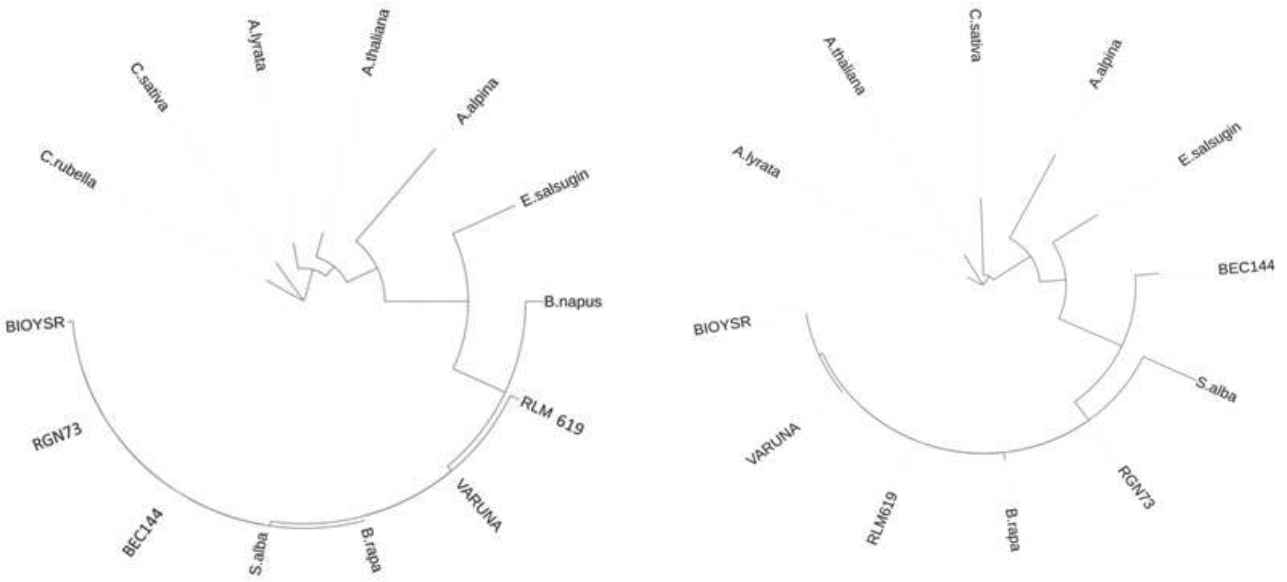

(***Fig. 2 contd. on next page***)



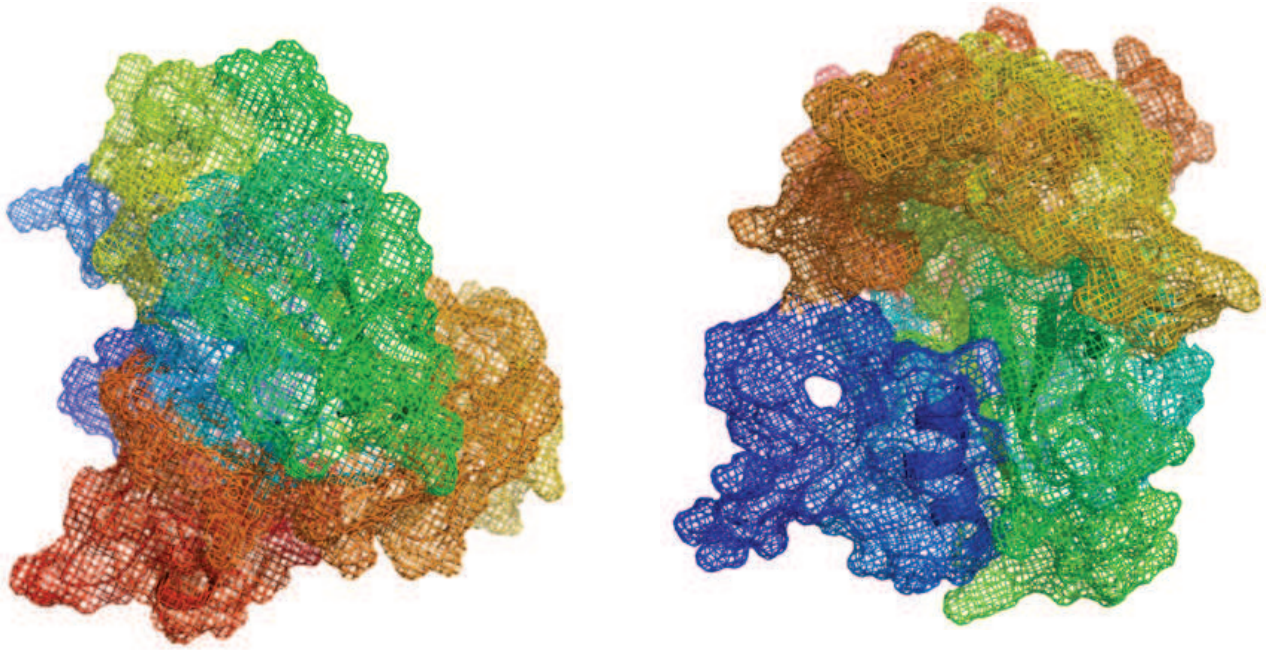

**Supplementary Fig. 2: Multiple sequence alignment of NAC14 (a) and NAC19 with NACs from different species. NAC14 cloned from *B. juncea* and *S. alba* were compared with NAC TFs of BraNAC2 (XP_009128229.1), BnaNAC14 (AAP35055.1), EsEUTSA (XP_006390089.1), AtANAC032 (NP_177869.1), AlANAC032 (XP_002887684.1), AaAALP (KFK42139.1), CrCARUB (XP_006302722.1), and CsNAC2 (XP_010428823.1). NAC19 cloned from *B. juncea* and *S. alba* were compared with BraNAC19 (XP_009147649.1), BnaNAC19 (AHN60135.1), EsEUTSA (XP_006392855.1), AtNAC19 (NP_175697.1), AlANAC019 (XP_002894409.1), CsNAC19 (XP_010462142.1) and AaAALP (KFK35760.1). The protein sequences were aligned with MAFFT v.7 and figured using ESPript. The conserved regions are highlighted by red color**




**Supplementary Fig.3: Multiple sequence alignment of NAC14 and NAC19 cloned from *B. juncea* genotypes and *S. alba***




**Supplementary Fig. 4. Phylogenetic tree of NAC14 (A) and NAC19 (B) from *S. alba*, RGN73, Varuna, RLM619, BEC144 and BIO-YSR with six eudicots namely *Brassiceae*, *Eutremeae*, *Arabideae*, *Arabidopsis*, *Camelina* and *Capsella***



**A** **B**

**Supplementary Fig. 5. Predicted S-D mesh models of NAC14 (A) and NAC 19 (B)**