African Journal of Agricultural Research Vol. 7(37), pp. 5252-5258, 25 September, 2012
Available online at http://www.academicjournals.org/AJAR
DOI: 10.5897/AJAR12.078
ISSN 1991-637X ©2012 Academic Journals

Full Length Research Paper

Characteristics and polymorphism of *NAM* gene from *Aegilops* section *sitopsis* species

Xi-Gui Hu[1], Bi-Hua Wu[1]*, Ze-Hong Yan[1], Shou-Fen Dai[1], Lian-Quan Zhang[1], Deng-Cai Liu[1,2] and You-Liang Zheng[1]

[1]Triticeae Research Institute, Sichuan Agricultural University, Wenjiang District, No. 211, Huimin Road, 611130 Chengdu, Sichuan, P. R. China.
[2]Key Laboratory of Adaptation and Evolution of Plateau Biota, Northwest Institute of Plateau Biology, Chinese Academy of Sciences, 810001 Xining, Qinghai, P. R. China.

Accepted 3 September, 2012

In this study, 11 novel *NAM* alleles were cloned from five species of *Aegilops* section *sitopsis*. These nucleotide sequences ranged from 1,540 to 1,555 bp with 3 exons and 2 introns. Their coding proteins contained both N-terminal nitrogen assimilation control (NAC) domain with 5 sub-domains and C-terminal transcriptional activation region (TAR). They have high identities and the closest homology with the *NAM* gene from *Triticum dicoccoides* Körn. and *T. timopheevii* Zhuk. indicating these *NAM* alleles of *A.* sect. *sitopsis* species probably having similar function to the *NAM-B1* gene which control high grain protein content (GPC) in *T. dicoccoides*. Among these 11 alleles, there were 185 variation sites including 143 single nucleotide polymorphism sites (SNPs) and 42 insertion/deletion sites (InDels), despite high identity. Majority of SNPs occurred in coding domain with more non-synonymous mutations, resulting in 53 variation amino acids. These suggested *NAM* alleles in *A.* sect. *sitopsis* species having abundant polymorphism, which mainly existed in the TAR.

Key words: *Aegilops* section *sitopsis*, *NAM* gene, single nucleotide polymorphism, grain protein content.

INTRODUCTION

The genus *Aegilops* is one of greatly important resources to genetic improvement of cultivated wheat because of being involved in the evolution of polyploid wheats; of which, the species *Aegilop tauschii* Coss. (2n = 2x = 14, DD) provides the D genome in *Triticum aestivum* L. (2n = 6x = 42, AABBDD) (McFadden and Sears, 1946; Lafiandra et al., 1992) and the species *A. speltoides* Tausch. (2n = 2x = 14, SS) is widely considered as the donor of B and G genomes of *Triticum* (Shands and Kimber, 1973; Kimber, 1974; Kilian et al., 2007; Salse et al., 2008; Wang et al., 2011) despite having been in controversy (Huang et al., 2002; Liu et al., 2003; Petersen et al., 2006). *A. speltoides* with the four species *A. longissima* Schweinf. and Muschl., *A. sharonensis* Eig., *A. bicornis* (Forssk.) Jaub. and Sp, and *A. searsii* Feld. and Kis belongs to *A.* section *sitopsis*, possessing the genome S (Van Slageren, 1994).

The gene *NAM*, is a member of the nitrogen assimilation control (NAC) family as plant-specific transcriptional regulator, which was originally characterized from consensus sequences from petunia NAM and from *Arabidopsis* ATAF1, ATAF2, and CUC2 (Clark et al., 1996; Waters et al., 2009). Genes with the NAC domain (NAC family genes) are expressed in various developmental stages and tissues, and play many important roles in plant development regulation, as well as adaptation of abiotic and biotic stresses (Olsen et al., 2005; Guo and Gan, 2006; Uauy et al., 2006; Mitsuda et al., 2007; Zhong et al., 2007). The ancestral *NAM-B1* gene was firstly located in a quantitative trait locus (QTL) for grain protein content

*Corresponding author. E-mail: wubihua2001@yahoo.com.cn. Tel: +86 2882650313. Fax: +86 2882650350.

Abbreviations: GPC, Grain protein content; **InDels,** insertion/deletion sites; **QTL,** quantitative trait locus; **SNPs,** single nucleotide polymorphism sites; **TAR,** transcriptional activation region.

Table 1. Plant materials used in this study.

Species	Genome	Accession	GenBank No.
Aegilops sharonensis	S^{sh}	CIae32	JQ085412
Aegilops bicornis	S^{b}	CIae47	JQ085413
		CIae70	JQ085414
Aegilops speltoides	S	PI330488	JQ085415
		PI560529	JQ085416
		PI542274	JQ085417
Aegilops searsii	S^{s}	AS2630	JQ085418
		PI599121	JQ085419
		PI599130	JQ085420
Aegilops longissima	S^{l}	PI604106	JQ085421
		PI604109	JQ085422

(GPC) (Joppa et al., 1997), and found in tetraploid wheats *T. dicoccoides* Körn. and *T. dicoccum* Schrank, with genome AABB (Uauy et al., 2006). However, the *NAM-B1* gene in cultivated durum varieties (*T. durum* Desf., AABB) and hexaploid bread wheat (*T. aestivum*) was non-functional, attributed to a frame-shift mutation resulted from 1 bp insertion within it (Uauy et al., 2006; Dubcovsky and Dovrak, 2007). It has been proved that the transgenic cultivated wheat Langdon (*T. durum*) with the *NAM-B1* gene from *T. dicoccoides* expressed high GPC (Uauy et al., 2006). So, this gene is much interest to modern-day breeders because it was considered as a domestication gene affecting GPC in wheat (Dubcovsky and Dovrak, 2007).

It is highly worthwhile to explore further the *NAM* gene resources. The objectives of this study were to isolate *NAM* alleles from *A.* sect. *sitopsis* species, and to analyze their sequence characteristics and relationships with the known *NAM* alleles from tetraploid wheats, for preliminarily investigating the potential value in quality improvement of cultivated wheat and polymorphism of these *NAM* genes.

MATERIALS AND METHODS

Plant materials

In this study, 11 accessions of *A.* sect. *sitopsis*, including the five species *A. speltoides*, *A. bicornis*, *A. longissima*, *A. searsii*, and *A. sharonensis* were used (Table 1), of which, the accession AS2630 was from Triticeae Research Institute, Sichuan Agricultural University, China, and the others were kindly provided by American National Plant Germplasm System (Pullmam, Washington, USA).

DNA extraction, amplification and sequence analysis of *NAM* gene

Genomic DNA was extracted from young leaves by CTAB method (Yan et al., 2002). To obtain the complete nucleotide sequence of *NAM* gene, a set of primers (P1: AGATCTGATGAGGTCCATGGG and P2: ATGCCCGTATGTGGTGTTCATAT) was designed, flanking the coding regions (Figure 1). Polymerase chain reaction (PCR) amplification was performed with Mastercycler proS (eppendorf) in 50 µl volume which consisted of 150 ng of genomic DNA, 0.2 mM of dNTPs, 0.5 µM of each primers, 0.2 mM Mg^{2+}, 5 µl 5× *ExTaq* buffer and 1.5 U of *ExTaq* DNA polymerase with high fidelity (TaKaRa, Dalian, China). The cycling parameters were 95ºC for 5 min to pre-denature, followed by 35 cycles of 95ºC for 50 s, 65ºC for 60 s and 72ºC for 2 min, with a final extension at 72ºC for 7 min. The desired PCR products were recovered from gels and ligated into pMD19-T vector (TaKaRa, Dalian, China) according to the manufacture's instruction. The cloned PCR products were sequenced in both directions by a commercial company (BGI, Shenzhen, China). All sequences were derived from at least five independent clones to exclude sequencing errors.

Data analysis

Sequence multiple alignments were carried out with the software DNAMAN version 6.0.3.48. Polymorphism sites were analyzed with DnaSP 5.10.01. Phylogenetic tree based on complete nucleotide sequence of *NAM* gene was constructed using MEGA 5.05 program by the neighbor joining method (NJ). The bootstrap values were estimated based on 1000 replications.

RESULTS AND DISCUSSION

Molecular characterization and comparison analysis

Using the primers P1 and P2, one desired DNA band was detected in genomic amplifications from each accession of *A.* sect. *sitopsis*, and 11 sequences were obtained with the GenBank No. JQ085412-JQ085422 (Table 1). Their full lengths were ranged from 1,540 to 1,555 bp, with 3 exons and 2 introns (Figure 2), similar to the results in *Triticum* and *Hordeum* (Uauy et al., 2006; Jamar et al., 2010). They had high identities and the closest homology with *NAM-B1* gene from *T. dicoccoides* (Uauy et al., 2006) and *NAM-G1* gene (HQ843865-HQ843876) from *T. timopheevii* Zhuk. (GenBank, www.ncbi.nlm.nih.gov)

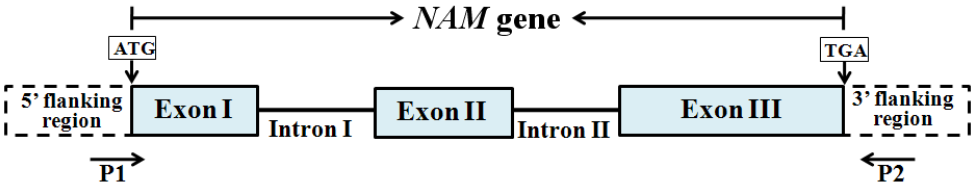

Figure 1. Schematic structure of *NAM* gene. A set of primer P1 and P2 was located in its flanking regions.

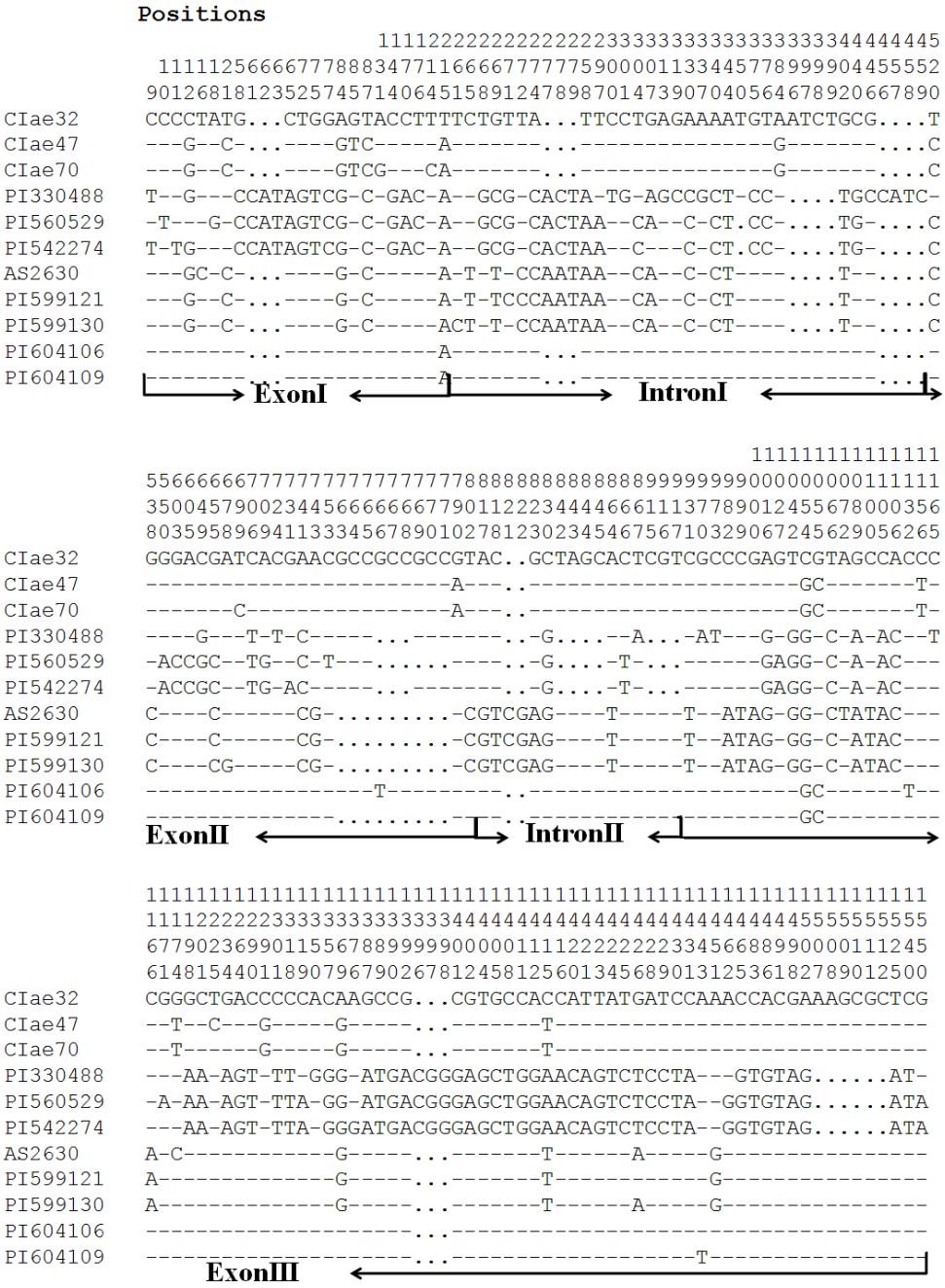

Figure 2. Sequence variations of the 11 *NAM* alleles from *Aegilops* sect. *sitopsis* species. Dashed (-) and Dots (.), respectively indicate identical and deletion nucleotides. The corresponding three exon and two intron domains are indicated.

(Figure 4).

Despite high identity (96.28%), many differences existed among the 11 nucleotide sequences. Among them, there were 185 variation sites containing 143 single nucleotide polymorphism sites (SNPs). Of which, 79 were transition of C-T for 42 and A-G for 37, including 65 in the exons and 14 in the introns. And, 59 were transversions of G-C for 23, A-C for 17, A-T for 11 and G-T for 8, including 43 in the exons and only 16 in the introns. In addition, there were 5 triallelic SNPs involving the two sites 269, 271 in intron I, and the three 1174, 1415 and 1426 in exon III, which were with both transition and transversion including G/C/T for 3, C/T/A and G/T/A for 1, respectively (Figure 2, Table 2). 111 SNPs occurred in coding sequences including 21 in exon I, 18 in exon II and 72 in exon III, and 32 in non-coding sequences including 25 in intron I and 7 in intron II. These results showed that the coding domain had higher proportion of SNPs than the non-coding region in the *NAM* gene of *S*. sect. *sitopsis* species, which was in agreement with the results by Jamar et al. (2010).

There were 42 insertion/deletion sites (InDels) among the 11 allelic sequences, of which, some were characteristic of certain species. For example, the insertions at positions 61 to 63, 277 to 279 and 1396 to 1398, as well as the deletions at positions 842 to 845, 915 to 917 and 1507 to 1512 were only presented in *A. speltoides* and the insertions at positions 277 to 279 and 821 to 822 were only in *A. searsii*. The other InDels sites were discovered in two or more species (Figure 2). These indicated that *NAM* gene of *A.* sect. *sitopsis* species possessed abundance polymorphism in comparison with that of *Hordeum* (Jamar et al., 2010; Madhavan, 2011).

Deduced amino acid analysis

The amino acid sequences deduced from these 11 *NAM* alleles of *A.* sect. *sitopsis* shared the typical structural characteristics of NAC protein (Aida et al., 1997; Ooka et al., 2003). They had both the N-terminal NAC domain containing A, B, C, D and E sub-domains, and the C-terminal TAR, which was similar to the structure of group I of *NAC* genes (Ooka et al., 2003).

Comparing these protein sequences with their coding nucleotide sequences, we found that non-synonymous mutations were more than synonymous mutations in coding region, which were derived from 61 and 50 SNPs, respectively. The 21 InDels at the positions 61 to 63, 763 to 771, 1396 to 1398 and 1507 to 1512 resulted in 7 amino acid InDels. And, 53 sites of amino acid variation were detected in the 11 protein sequences, of which, more variations existed in the C-terminal TAR with 31, than those in the N-terminal NAC domain (Figure 3), according to the characteristics of NAC protein (Ren et al., 2000; Xie et al., 2000; Duval et al., 2002; Ernst et al., 2004). The highly divergent TAR was considered as affecting function of NAC protein (Ooka et al., 2003; Olsen et al., 2005).

It has been proved that the functional *NAM* gene was correlated to high GPC (Uauy et al., 2006; Jamar et al., 2010), since the QTL of GPC was found to co-map with it in wheat and barley (Joppa et al., 1997; Distelfeld and Fahima, 2007; Distelfeld et al., 2008). Recent researches discovered that polymorphism among functional *NAM* alleles seems also to affect the protein content (Jamar et al., 2010; Madhavan, 2011). As a result, the variations in TAR might result in certain differences of GPC among the 11 accessions of *A*. sect. *stopsis*, which needs to be further studied.

Of the 22 variation sites in the N-terminal NAC domain, 18 occurred in the outer of the 5 sub-domains, only 4 in the three sub-domains A, B and D, while none was in the two sub-domains C and E (Figure 3). The NAC domain was associated with DNA-binding ability (Ooka et al., 2003; Olsen et al., 2005), and the three sub-domains B, C and E were responsible for maintaining the correct structure of NAC domain in dimerization or DNA attachment (Duval et al., 2002; Xue et al., 2006). It was obvious that the N-terminal NAC domain of the present *NAM* gene of *A.* sect. *sitopsis* species was highly conserved, especially the C and E sub-domains. This suggested that *NAM* alleles from *A.* sect. *sitopsis* species might correctly express their functions.

Phylogenetic analysis

In this study, the *NAM-H1* gene (EU368851) from *Hordeum vulgare* ssp. *vulgare* (Lewis) (Distelfeld et al., 2008) was used as out group, and a phylogenetic tree was constructed based on the complete nucleotide sequences from these 11 cloned *NAM* alleles of *A.* sect. *sitopsis* species, 2 *NAM-B1* alleles (DQ871219, DQ869673) from *T. dicoccoides* (Uauy et al., 2006), and 12 *NAM-G1* alleles (HQ843865-HQ843876) from *T. timopheevii* (Figure 4). The result showed that the *NAM* alleles from tetraploid wheats of both *T. dicoccoides* and *T. timopheevii* were more closely related to those from *A. speltoides*, than those from *A. sharonensis*, *A. searsii*, *A. bicornis* and *A. longissima* in *A.* sect. *sitopsis*, showing the closer relationships among the B, G and S genomes (Kilian et al., 2007; Wang et al., 2011). This further supported the results that the B and G genomes could be related to *A. speltoides* but not to other species in *A.* sect. *sitopsis* (Shands and Kimber, 1973; Kimber, 1974; Salina et al., 2006; Kilian et al., 2007; Salse et al., 2008; Wang et al., 2011). These implied that the novel *NAM* alleles in *A.* sect. *sitopsis* species might have similar function to the *NAM-B1* gene affecting GPC in *T. dicoccoides* (Uauy et al., 2006).

ACKNOWLEDGEMENTS

We wish to thank the anonymous reviewers for helpful

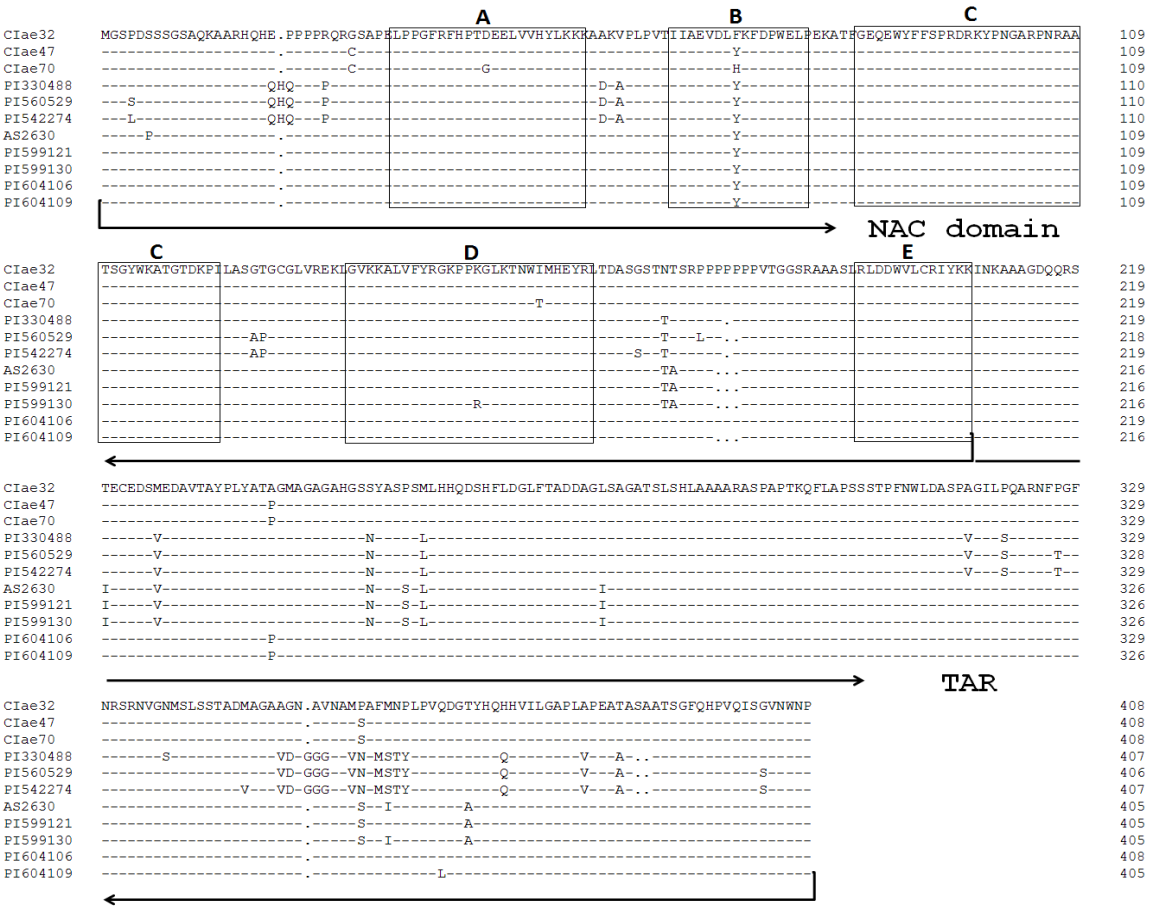

Figure 3. Comparison of the deduced amino acid sequences from the 11 *NAM* alleles from *Aegilops* sect. *sitopsis* species. Dashed (-) and Dots (.) respectively indicate identical and deletion amino acid residues. The NAC domain, transcriptional activation region (TAR), and the five sub-domains A, B, C, D and E are, respectively indicated with arrows and rectangle frames.

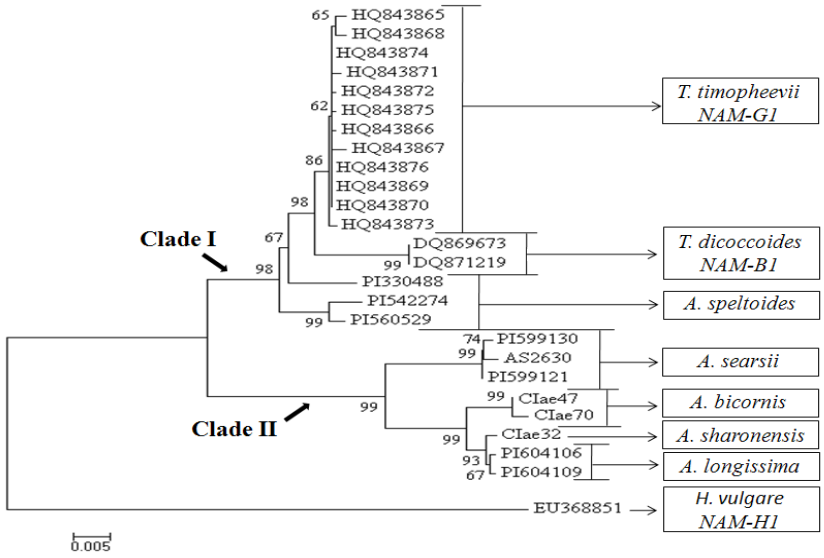

Figure 4. Phylogenetic relationship among 26 *NAM* alleles from *Aegilops* sect. *sitopsis* species, tetraploid wheats *T. dicoccoides* and *T. timopheevii*, and *Hordeum vulgar*e ssp. *vulgare*.

Table 2. Single nucleotide polymorphism sites (SNPs) among 11 new *NAM* alleles from *Aegilops* sect. *sitopsis* species.

Domains	No.	SNPs								
		Transitions		Transversions				Transitions/Transversions		
		T-C	A-G	G-C	A-C	G-T	A-T	G/T/C	C/T/A	G/T/A
Exon I	21	8	3	4	2	3	1	0	0	0
Exon II	18	6	6	4	2	0	0	0	0	0
Exon III	72	20	22	9	10	3	5	1	1	1
Exons	111	34	31	17	14	6	6		3	
Intron I	25	6	4	5	3	2	3	2	0	0
Intron II	7	2	2	1	0	0	2	0	0	0
Introns	32	8	6	6	3	2	5		2	

comments that improved the manuscript. This research was supported by the National Natural Science Foundation of China No. 30571139 and No. 30671271, as well as by both the Personnel Training Foundation and Education Committee Accented Term in Sichuan Province.

REFERENCES

Aida M, Ishida T, Fukaki H, Fujisawa H, Tasaka M (1997). Genes involved in organ separation in *Arabidopsis*: an analysis of the cup-shaped cotyledon mutant. Plant Cell 9:841-857.

Clark SE, Jacobsen SE, Levin JZ, Meyerowitz EM (1996). The *CLAVATA* and shoot meristemless loci competitively regulate meristem activity in *Arabidopsis*. Development 122:1567-1575.

Distelfeld A, Fahima T (2007). Wild emmer wheat as a source for high-grain-protein genes: map-based cloning of *Gpc-B1*. Israel J. Plant Sci. 55:297-306.

Distelfeld A, Korol A, Dubcovsky J, Uauy C, Blake T, Fahima T (2008). Colinearity between the barley grain protein content (GPC) QTL on chromosome arm 6HS and the wheat *Gpc-B1* region. Mol. Breed. 22:25-38.

Dubcovsky J, Dvorak J (2007). Genome plasticity a key factor in the success of polyploidy wheat under domestication. Science 316:1862-1866.

Duval M, Hsieh TF, Kim SY, Thomas TL (2002). Molecular characterization of *AtNAM*: a member of the *Arabidopsis* NAC domain superfamily. Plant Mol. Biol. 50:237-248.

Ernst HA, Olsen AN, Skriver K, Larsen S, Leggio LL (2004). Structure of the conserved domain of ANAC, a member of the NAC family of transcription factors. EMBO Rep. 5:297-303.

Guo YF, Gan SS (2006). AtNAP, a NAC family transcription factor, has an important role in leaf senescence. Plant J. 46:601-612.

Huang S, Sirikhachornkit A, Su X, Faris J, Gill B, Haselkorn R, Gornicki P (2002). Genes encoding plastid acetyl-CoA carboxylase and 3-phosphoglycerate kinase of the *Triticum*/*Aegilops* complex and the evolutionary history of polyploidy wheat. Proc. Natl. Acad. Sci. USA 99:8133-8138.

Jamar C, Loffet F, Frettinger P, Ramsay L, Fauconnier ML, Jardin PD (2010). *NAM-1* gene polymorphism and grain protein content in *Hordeum*. J. Plant Physiol. 167:497-501.

Joppa LR, Du C, Hart GE, Hareland GA (1997). Mapping gene(s) for grain protein in tetraploid wheat (*Triticum turgidum* L.) using a population of recombinant inbred chromosome lines. Crop Sci. 37:1586-1589.

Kilian B, Özkan H, Deusch O, Effgen S, Brandolini A, Kohl J, Martin W, Salamini F (2007). Independent wheat B and G genome origins in outcrossing *Aegilops* progenitor haplotypes. Mol. Biol. Evol. 24:217-227.

Kimber G (1974). A reassessment of the origin of the polyploid wheats. Genetics 78:487-492.

Lafiandra D, Masci S, D'Ovidio R, Tanzarella OA, Porceddu E, Margiotta B (1992). Relationship between the D genome of hexaploid wheats (AABBDD) and *Ae. squarrosa* as deduced by seed storage proteins and molecular marker analyses. Hereditas 116:233-238.

Liu B, Segal G, Rong JK, Feldman M (2003). A chromosome-specific sequence common to the B genome of polyploid wheat and *Aegilops searsii*. Plant Syst. Evol. 241:55-66.

Madhavan S (2011). Screening of *HvNAM-B1* polymorphism, grain nutrient content and seed size in 80 Scandinavian barley cultivars. Dissertation, Sweden Linköpings universitet.

McFadden ES, Sears ER (1946). The origin of *Triticum spelta* and its free-threshing hexaploid relatives. J. Hered. 37:81-89.

Mitsuda N, Iwase A, Yamamoto H, Yoshida M, Seki M, Shinozaki K, Ohme-Takagi M (2007). NAC transcription factors, NST1 and NST3, are key regulators of the formation of secondary walls in woody tissues of *Arabidopsis*. Plant Cell. 19:270-280.

Olsen AN, Ernst HA, Leggio LL, Skriver K (2005). NAC transcription factors: structurally distinct, functionally diverse. Trends Plant Sci. 10:79-87.

Ooka H, Satoh K, Doi K, Nagata T, Otomo Y, Murakami K, Matsubara K, Osato N, Kawai J, Carninci P, Hayashizaki Y, Suzuki K, Kojima K, Takahara Y, Yamamoto K, Kikuchi S (2003). Comprehensive analysis of NAC family genes in *Oryza sativa* and *Arabidopsis thafiana*. DNA Res. 10:239-247.

Petersen G, Seberg O, Yde M, Berthelsen K (2006). Phylogenetic relationships of *Triticum* and *Aegilops* and evidence for the origin of the A, B, and D genomes of common wheat (*Triticum aestivum*). Mol. Phylogenet. Evol. 39:70-82.

Ren T, Qu F, Morris TJ (2000). *HRT* gene function requires interaction between a NAC protein and viral capsid protein to confer resistance to turnip crinkle virus. Plant Cell 12:1917-1925.

Salina EA, Lim KY, Badaeva ED, Shcherban AB, Adonina IG, Amosova AV, Samatadze TE, Vatolina TY, Zoshchuk SA, Leitch AR (2006). Phylogenetic reconstruction of *Aegilops* section *Sitopsis* and the evolution of tandem repeats in the diploids and derived wheat polyploids. Genome 49:1023-1035.

Salse J, Chagué V, Bolot S, Magdelenat G, Huneau C, Pont C, Belcram H, Couloux A, Gardais S, Evrard A, Segurens B, Charles M, Ravel C, Samain S, Charmet G, Boudet N, Chalhoub B (2008). New insights into the origin of the B genome of hexaploid wheat: evolutionary relationships at the *SPA* genomic region with the S genome of the diploid relative *Aegilops speltoides*. BMC Genomics 9:555.

Shands HL, Kimber G (1973). Reallocation of the genomes of *Triticum timopheevii* Zhuk. In: Sears ER, Sears LMS (eds.) Proc. 4th Int. Wheat Genet. Symp., Missouri Agricultural Experiment Station, University of Missouri, Columbia, MO. pp.101-108.

Uauy C, Distelfeld A, Fahima T, Blechl A, Dubcovsky J (2006). A NAC gene regulating senescence improves grain protein, zinc, and iron content in wheat. Science 314:1298-1301.

Van Slageren MW (1994). Wild wheats: a monograph of *Aegilops* L., and

Amblyopyrum (Jaub & Spach) Eig. (Poaceae). Dissertation, Wageningen Agricultural University.
Wang SL, Li XH, Wang K, Wang XZ, Li SS, Zhang YZ, Guo GF, Zeller FJ, Hsam SLK, Yan YM (2011). Phylogenetic analysis of C, M, N, and U genomes and their relationships with *Triticum* and other related genomes as revealed by LMW-GS genes at *Glu-3* loci. Genome 54:273-284.
Waters BM, Uauy C, Dubcovsky J, Grusak MA (2009). Wheat (*Triticum aestivum*) NAM proteins regulate the translocation of iron, zinc, and nitrogen compounds from vegetative tissues to grain. J. Exp. Bot. 60:4263-4274.
Xie Q, Frugis G, Colgan D, Chua NH (2000). *Arabidopsis* NAC1 transduces auxin signal downstream of TIR1 to promote lateral root development. Gene Dev. 14:3024-3036.
Xue GP, Bower NI, McIntyre L, Riding GA, Kazan K, Shorter R (2006). TaNAC69 from the NAC superfamily of transcription factors is up-regulated by abiotic stresses in wheat and recognises two consensus DNA-binding sequences. Funct. Plant Boil. 33:43-57.
Yan ZH, Wan YF, Liu KF, Zheng YL, Wang DW (2002). Identification of a novel HMW glutenin subunit and comparison of its amino acid sequence with those of homologous subunits. Chin. Sci. Bull. 47:222-227.
Zhong RQ, Richardson EA, Ye ZH (2007). Two NAC domain transcription factors, SND1 and NST1, function redundantly in regulation of secondary wall synthesis in fibers of *Arabidopsis*. Planta 225:1603-1611.