Understanding the molecular basis of differential grain protein accumulation in wheat (*Triticum aestivum* L.) through expression profiling of transcription factors related to seed nutrients storage

Alok Kumar¹ · Jai Prakash Jaiswal² · Netrapal Sharma¹ · Supriya Gupta¹ · Anil Kumar¹

Received: 8 April 2017 / Accepted: 11 January 2018 / Published online: 5 February 2018
© Springer-Verlag GmbH Germany, part of Springer Nature 2018

Abstract

Increasing nutritional value of cereals is one of the important research and breeding objectives to overcome malnutrition in developing countries. The synthesis of grain seed proteins during grain filling is controlled by several mechanisms including transcriptional and posttranscriptional modifications. In the current investigation, transcript abundance analysis of three allelic variants of seed storage protein activator (*Spa A*, *Spa B* and *Spa D*) and *NAM-B1* affecting seed nutrient concentration was carried out in two genotypes (UP 2672 and HS 540) of bread wheat differing in grain protein content. Expression profiling of transcription factor genes was performed using quantitative real time PCR (qRT-PCR). Positive correlation and significant *p* value > 0.05 was observed among the fold expression in developing stages of both the genotypes. Maximum expression of *Spa* genes was observed at S3 stage and maximum fold expression was observed for *Spa B* gene in case of UP 2672, the genotype with high protein content. The transcript profiling of *NAM-B1* gene revealed threefold higher expression in UP 2672 than that of HS 490 at S4 stage. The findings revealed the role of transcriptional regulation in differential grain protein accumulation through varied expression and existence of their allelic variants in wheat genotypes.

Keywords Wheat · Seed storage protein · qRT-PCR · Transcription factors · *NAM-B1*

Wheat (*Triticum aestivum* L., AABBDD, $2n = 42$) is a globally important crop, occupying 17% of all the cultivated land and accounting for 20% of the calories consumed by humans and is, therefore, of great economic importance (Gill et al. 2004; Brenchle et al. 2012). It has been described as the ‘King of cereals’ because of the acreage it occupies, high productivity and the prominent position it holds in the international food grain trade (Kumar et al. 2014). India is one of the major producers (16%) of wheat and occupies second

place after China in terms of area and production among wheat growing countries of the world (ICAR-IIWBR 2017). Wheat constitutes important source of carbohydrates, proteins and also an important source for micronutrients such as Zn and Fe (Distelfeld et al. 2007; Kumar et al. 2014). The storage-protein genes are expressed during seed development stages, which accounts for about 50% of the total protein in mature cereal grains and have important impacts on their nutritional quality for humans and livestock and on their functional properties in food processing (Shewry and Halford 2002). The synthesis of Grain Seed Proteins (GSPs) during grain filling is controlled by several mechanisms, including transcriptional and post transcriptional modifications.

Transcription factors (TFs) regulate gene expression by binding to specific *cis*-acting promoter elements, thereby activating or repressing the transcriptional rates of GSP genes (Verdier and Thompson 2008). In plants, totally more than 50 different families of TFs have been identified based on bioinformatics analysis so far (Hu et al. 2010). Several TFs have been reported to specifically interact with the Endosperm box (EB) and the AACA motif. Three

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s13205-018-1114-5>) contains supplementary material, which is available to authorized users.

✉ Jai Prakash Jaiswal
jpj.gbpu@gmail.com

¹ Department of Molecular Biology & Genetic Engineering, G.B. Pant University of Agriculture & Technology, Pantnagar 263 145, India

² Department of Genetics & Plant Breeding, G.B. Pant University of Agriculture & Technology, Pantnagar 263 145, India

important *cis*-motifs are conserved in GSP promoters from wheat, barley (*Hordeum vulgare*), maize (*Zea mays*), and rice (*Oryza sativa*). These are the GCN4-like motif (GLM) and the prolamin box (PB), which together constitutes the endosperm box (EB), and the AACA motif (Zheng et al. 1993). The GLM motif is recognized by basic Leu zipper (bZIP) TFs of the Opaque2 subfamily, called SPA in wheat (Albani et al. 1997).

Storage protein activator (SPA) is a key regulator of the transcription of wheat grain storage protein genes. Other TF-NAC-domain (NAM; ATAF1, 2; CUC2) TF that has a functional copy in DIC is encoded by GPC-B1 also called (*NAM-B1*) gene. SPA is a key regulator of the transcription of wheat GSP genes and belongs to the Opaque2 (O2) TF subfamily (Albani et al. 1997). Changes in the expression of SPA might allow a shift in grain nitrogen allocation and therefore modify the GSP composition independently of grain protein concentration (Ravel et al. 2006). *NAM-B1* gene affects grain nutrient concentrations in wheat and it is the first identified wheat gene responsible for variation in GPC. A major effect of the *NAM* genes is an increased efflux of nutrients from the vegetative tissues and a higher partitioning of nutrients to grain (Waters et al. 2009). The expression profiles of the TFs selected showed three distinct patterns, associated with gene expression predicted to be in the endosperm, embryo or pericarp (Wan et al. 2008). The genotype is one of the important factors in determining the final protein composition (Bonfil and Kafkafi 1993). In Table 1, two different genotypes of wheat possessing contrasting characters like UP 2672 (> 14%, high grain protein content) and HS 490 (< 11%, low grain protein content) were taken in the present study to investigate the role of some potential TFs affecting grain protein content. In the

present study, expression profiling of *Spa A* and *Spa B*, *Spa D* and *NAM-B1* TF genes was done to reconstruct transcriptional regulatory networks and define its functional role in grain protein concentration as a strategy for the crop improvement and bio-fortification.

The seeds of two wheat genotypes were sown at N. E. Borlaug Crop Research Centre, G.B. Pant University of Agriculture and Technology, Pantnagar, India. Five different developmental stages of the spike, booting (S_1), anthesis (S_2), milk development (S_3), dough development (S_4) and ripening (S_5) were used for qPCR analysis (Supplementary Fig. S1; Supplementary Table 1). Spikes samples were collected from the one-third portion from the tip. All these samples were collected in the forenoon (between 0800 and 0900 hours). At least three independent biological replicates of each tissue sample were harvested and immediately frozen in liquid nitrogen and stored at -80°C till further use.

Total RNA was isolated from five stages of developing spikes (S_1 , S_2 , S_3 , S_4 and S_5) of wheat genotypes (UP 2672 and HS 490) using IRIS total RNA isolation kit (from IHB, Palampur). DNaseI treated total RNA (2 μg) was used to synthesize first strand cDNA using oligo (dT)₁₈ primer. Real-time PCR (qPCR) was carried out using the 5' Real Master Mix SYBR ROX (Eppendorf India Limited, Chennai, India) according to the manufacturer's instructions. The list of gene specific primer used for expression analysis is given in Supplementary Table 2. Actin gene primer was used as internal control. The normalized transcript expression was quantified using $2^{-\Delta\Delta\text{CT}}$ method (Livak and Schmittgen 2001) and presented as fold change over control (calibrator). The correlation coefficient and covariance analysis was performed using the Rx64-3.1.2 Software package. Analysis of variance (ANOVA) was performed using SPSS 21.0 (SPSS

Table 1 Characteristics of the wheat genotypes, UP 2672 and HS 490 used in the study

Particulars	Genotypes	
	UP 2672	HS 490
Pedigree	UP 2382/DLRSN-7	HS 364/HPW 114/HS 240/HS 346
Trait value	Registered with ICAR-NBPGR, New Delhi in 2013 for high protein content (> 14.0%). Evaluated over 40 locations across the year	Released by CVRC for cultivation in NHZ. It possesses low protein content < 10% and very good biscuit making quality
Developed by	G.B. Pant University of Agri. & Tech., Pantnagar	IARI Regional Station, Shimla
Registration No.	IC0597682; INGR13053	Notification No. & date: 449(E), 11.02.2009
Agronomic features	Days to heading: 80	Days to heading: 100
	Days to maturity: 135	Days to maturity: 144
	Plant height: 82 cm	Plant height: 87 cm
Quality attributes	Actual protein content: 14.1%	Actual protein content: 9.6%
	Test weight: 80.8 kg/hl	Test weight: 82.6 kg/hl
	1000 grain wt.: 47.1 g	1000 grain wt.: 44 g
	Sedimentation value: 51.5 ml	Sedimentation value: 33 ml
	Grain appearance score: 5.7	Grain appearance score: 6.6

NHZ Northern hills zone

Inc., Chicago, IL, USA) to detect differences in expression among genotypes and developing stages of spikes (Kumar et al. 2015).

The quantitative expression of *SPA* genes (*Spa A*, *Spa B*, *Spa D*) and *NAM-B1* gene was studied to understand the correlation between the expression pattern of these gene with seed protein accumulation in developing spikes of wheat genotypes (UP 2672 and HS 490) differing in grain protein content. Expression profiling of *Spa A* gene transcript showed marked difference at different stages of developing spikes in both the studied genotypes. S_4 stage of the HS490 genotype was taken as a reference stage for finding the relative expression of all the other stages in both the genotype. Irrespective of the genotypes, the transcript level was in decreasing order starting from the S_3 stage to S_5 stage, i.e. as the maturation of the seed continues the transcript level decreases. As calculated, about 151.27- and 4.72-fold lower expression was reported for HS 490 and UP 2672, respectively at ripening stage as compared to the milking stage. The maximum transcript accumulation was at S_3 stage of about 29.65 and 51.98 for low and high protein genotype, respectively.

Higher expression in UP 2672 might define the role of *Spa A* TF gene during protein storage in wheat seed (Fig. 1a). The results suggest that *Spa A* gene have relatively less or no role during the booting (S_1) and anthesis (S_2) stages as no increase in transcript accumulation have been observed in studied wheat genotypes. *f* test and *t*-test were also conducted to compare two variances and mean *f*-test has produced *p* value = 0.303, ratio of variances is 0.3263554 and *t* test *p* value = 0.2927 with mean of HS 490: 6.169393 and mean of UP 2672: 19.504760. Analysis of variance (ANOVA) conclude that significant difference exist in expression among genotypes and developing stages of spikes as the *p* value obtained is about 0.1104 and 0.0621, respectively, which is greater than 0.05.

Transcript abundance of *Spa B* gene depicts expression initiates from the S_3 stage, i.e. milk stage that decrease from S_3 (32.4467) to S_5 (0.273573) stage for both the genotypes. This decrease was 118- and 13-fold from milk stage to ripening stage for HS 490 and UP 2672, respectively. Similar to *Spa A*, no transcript accumulation has been observed for *Spa B* gene at S_1 and S_2 stage. The transcript accumulation

Fig. 1 Relative expression analysis of **a** *Spa A*, **b** *Spa B*, **c** *Spa D*, and **d** *NAM-B1* gene by quantitative RT-PCR in different stages (S1, S2, S3, S4 and S5) of developing spikes in two genotypes (HS490 and UP2672) of wheat differing in their seed protein content

was highest at S_3 stage which was more than twofold higher for UP 2672 as compared to HS 490 (Fig. 1b). This further suggests that higher expression of *Spa B* TF gene might play role during protein storage in high protein wheat genotype since most of the protein stored at S_3 stage. *F* test has produced *p* value = 0.2686, ratio of variances 0.2986256 and *t*-test produced *p* value = 0.5428 with mean for HS 490: 6.169393 and mean for UP 2672: 19.504760. ANOVA results showed significant difference in expression among genotypes and developing stages of spikes as the *p* value obtained is about 2.432785 and 5.308955, respectively, which is greater than 0.05.

Transcript level of *Spa D* gene showed marked difference in different stages of both genotypes. In HS 490, 133.43-fold lower expressions were observed at S_5 stage as compared to S_3 stage. The highest accumulation of transcript was at S_3 stage of about 11.08 that further supports its role during protein accumulation. Similarly, in UP 2672, the highest relative expression of *Spa D* gene was at S_3 stage (milk stage) of about 22, which was twofold higher as compared to HS 490 (Fig. 1c). By using the *f*-test and two sample *t*-test, the expression values were analyzed for significance. It was seen that by *f*-test, *p* value = 0.2686, ratio of variances: 0.2986256. By *t*-test statistics, *p* value = 0.5428 which is greater than 0.05 so our values obtained are acceptable and sample estimates mean for HS 490 and UP 2672 is 2.432785 and 5.308955, respectively. Using ANOVA, the *p* value obtained is significant to about 0.1935 and 0.0194, respectively, which is greater than 0.05 suggesting there is significant difference in expression among genotypes and developing stages of spikes.

Expression profiling of *Spa A*, *Spa B* and *Spa D* genes showed similar expression pattern in both the studied wheat genotypes. Irrespective of the genotypes, the expression of studied *Spa* genes was maximum at S_3 stage (milking stage) that further declines at S_4 and minimum at S_5 stage of spikes. Similar findings were reported by Albani et al. (1997) and Ravel et al. (2006) about *SPA* TFs that the three *SPA* homologs have similar patterns of expression during grain development, with a peak in expression around 300 degree days after anthesis.

The expression profiling of *NAM-B1* gene showed no relative expression at S_1 and S_2 stages of developing spikes in both the wheat genotypes (UP 2672 and HS 490). The expression increases from S_3 stage that reaches maximum at S_4 stage up to sevenfold for HS 490 and 11-fold for UP 2672 that further declines to half-fold for both the genotypes at S_5 stage. It showed threefold high expressions in UP 2672 (22.7) than HS 490 (7.91) at S_4 stage (Fig. 1d). This suggests that there might be the low protein genotype (HS490) contains low GPC allele of *NAM-B1*, which is non-functional *NAM-B1* as reported earlier in wheat genotypes. Uauy et al. (2006b) reported that in addition to the functional *NAM-B1*

allele a low GPC allele of *NAM-B1* exists in durum wheat genotypes that affects the grain protein content in wheat. These durum wheat accessions had a 1-bp insertion causing a frame-shift mutation making it non-functional. Thus high protein genotype UP 2672 having functional *NAM-B1* can be explored as a potential material for crossing *NAM-B1* allele into high-performance genetic backgrounds. It was seen that by *f* test *p* value = 0.05062, ratio of variances: 0.1048476. By *t*-test statistics: *p* value = 0.3352, which is greater than 0.05 so our values obtained are acceptable and sample estimates mean for HS 490 is 2.33 and UP 2672 is 6.953439. This shows that there is statistical significant values of means and ratio of variance which are acceptable. ANOVA result supports that there is significant difference in expression among genotypes and developing stages of spikes as the *p* value obtained is significant to about 0.186 and 0.111, respectively, which is greater than 0.05.

In conclusion, in nutshell, three *SPA* homologs have similar patterns of expression and *NAM-B1* has slightly different patterns of expression as compared to *SPA* during grain development in wheat genotypes. *SPA* showed high expression during milk stage (S_3) while *NAM-B1* showed high transcript accumulation in late stage as compared to *SPA* TFs, i.e. at the dough stage (S_4). At the late maturity stage (S_5) the transcript level of all the TFs declines. A strong positive correlation was observed among the *SPA* TFs, while negative correlation was observed among *NAM-B1* and *SPA* TFs. This shows that *SPA* TFs are linked to each other in transcription regulatory network of seed storage proteins (SSPs), which needs to be further investigated. However, detailed molecular insight will allow exploration of correlation among the different *SPA* TFs (*Spa A*, *Spa B*, *Spa D*) gene and *NAM-B1* gene in transcription regulatory network to investigate the role of SSP in wheat.

Acknowledgement The logistic support provided by Director, Experiment Station, G.B. Pant University of Agriculture and Technology, Pantnagar is acknowledged.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

- Albani D, Hammond-Kosack MCU, Smith C, Conlan S, Colot V, Holdsworth M, Bevan M (1997) The wheat transcriptional activator SPA: a seed-specific bZIP protein that recognizes the GCN4-like motif in the bifactorial endosperm box of prolamin genes. *Plant Cell* 9:171–184
- Bonfil DJ, Kafkafi U (1993) Effect of soil type and mineral nutrition on glutenin-gene expression in wild and cultivated wheat. VIII Int Wheat Genet Symp Beijing, China 2:1087–1090

- Brenchle R, Spannag M, Pfeifer M, Barker GL, D'Amore R (2012) Analysis of the bread wheat genome using whole-genome shotgun sequencing. *Nature* 491:705–710
- Distelfeld A, Cakmak I, Peleg Z, Ozturk L, Yazici AM, Budak H, Saranga Y, Fahima T (2007) Multiple QTL-effects of wheat Gpc-B1 locus on grain protein and micronutrient concentrations. *Physiol Plant* 129:635–643
- Gill BS, Appels R, Botha-Oberholster AM, Buell CR, Bennetzen JL (2004) A workshop report on wheat genome sequencing. *Int Gen Res Wheat Consort Genet* 168:1087–1096
- Guillaumie S, Charmet G, Linossier L, Torney V, Robert N, Ravel C (2004) Co-location between a gene encoding for the bZip factor SPA and an eQTL for a high-molecular-weight glutenin subunit in wheat (*Triticum aestivum*). *Genome* 47:705–713
- Hu R, Qi G, Kong Y, Kong D, Gao U, Zhou G (2010) Comprehensive analysis of NAC domain transcription factor gene family in *Populus trichocarpa*. *BMC Plant Biol* 10:145–154
- ICAR-IIWBR (2017) Director's report of AICRP of wheat and barley 2016–17. In: Singh GP (ed) ICAR-Indian Institute of Wheat and Barley Research, Karnal, Haryana, India. P. 87.
- Kumar B, Dhananjay Singh BN (2014) Evaluation of genetic divergence in Wheat (*Triticum aestivum* L.) germplasms. *The Bioscan* 9:755–758
- Kumar A, Pathak RK, Gupta SM, Gaur VS, Pandey D (2015) Systems biology for smart crops and agricultural innovation: filling the gaps between genotype and phenotype for complex traits linked with robust agricultural productivity and sustainability. *OMICS A J Integ Biol* 19:581–601
- Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2^{-ΔΔC_T} method. *Methods* 25:402–408
- Ravel C, Nagy JJ, Martre P, Sourdille P, Dardevet M, Balfourier F, Pont C, Giancola S, Praud S, Charmet G (2006) Single nucleotide polymorphism, genetic mapping, and expression of genes coding for the DOF wheat prolamin-box binding factor. *Funct Integr Genom* 6:310–321
- Shewry PR, Halford NG (2002) Cereal seed storage proteins: structures, properties and role in grain utilization. *J Exp Bot* 53:947–958
- Uauy C, Brevis JC, Dubcovsky J (2006a) The high grain protein content gene Gpc-B1 accelerates senescence and has pleiotropic effects on protein content in wheat. *J Exp Bot* 57:2785–2794
- Uauy C, Distelfeld A, Fahima T, Blechl A, Dubcovsky J (2006b) A NAC gene regulating senescence improves grain protein, zinc, and iron content in wheat. *Science* 314:1298–1301
- Verdier J, Thompson RD (2008) Transcriptional regulation of storage protein synthesis during dicotyledon seed filling. *Plant Cell Physiol* 49:1263–1271
- Wan Y, Poole RL, Huttly AK, Toscano-Underwood C, Feeney K, Welham S, Gooding MJ, Mills C, Edwards KJ, Shewry PR (2008) Transcriptome analysis of grain development in hexaploid wheat. *BMC Genom* 9:121–125
- 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
- Zheng Z, Kawagoe Y, Xiao S, Li Z, Okita T, Hau TL, Lin A, Murai M (1993) Distal and proximal cis-acting regulator elements are required for developmental control of a rice seed storage protein glutelin gene. *Plant J* 4:357–366