Indian J. Plant Genet. Resour. 30(2): 162-164 (2017)
DOI 10.5958/0976-1926.2017.00023.7

SHORT COMMUNICATION

Text-mining for Identifying Abiotic Stress Candidate Genes to Screen Bread Wheat (*Triticum aestivum*) Germplasm

Monica Jamla, Ambika Baldev Gaikwad, Sherry Rachel Jacob and Sunil Archak
ICAR-National Bureau of Plant Genetic Resources, New Delhi-110012, India

(Received: 30 July 2016; Revised: 26 November 2016; Accepted: 06 December 2016)

An exhaustive text-mining of flowering plants' nucleotide data for the genes known to be associated with abiotic stress tolerance shortlisted 132 putative candidate genes. Primers were designed for orthologous amplification of the gene sequences in two wheat varieties and based on the banding pattern, 24 loci were tested in 13 heat and drought tolerant wheat genotypes along with a susceptible control for their utility in germplasm screening.

Key Words: Abiotic Stress, Bioinformatics, Candidate gene, Text-mining, Wheat

Wheat is the second most produced food-grain worldwide (FAOSTAT, 2016). However, it has been predicted that every ºC of temperature increase will lead to reduction of the global wheat production by 6% (Asseng *et al*., 2015). Consequently, researchers around the world have been trying to develop wheat germplasm lines with tolerance to abiotic stress (Bansal *et al*., 2013). Gene-specific markers have been shown to be highly useful in trait-specific germplasm screening (Singh *et al*., 2015; Archak *et al*., 2016). Abiotic stress is a complex trait and involves signal trafficking of different classes of proteins (Agarwal and Jha, 2010). However, huge amount of sequence data is available in the public domain (Benson *et al*., 2015) which facilitates the application of *in silico* techniques to identify putative genes involved in abiotic stress response pathway in plants. Computational tools (Moreau and Tranchevent, 2012) including simple supervised text-mining (Zhu *et al*., 2013) have been employed in the discovery of candidate genes.

Present study was taken up to text-mine open source gene database to identify candidate genes from plants putatively involved in abiotic stress response, amplify the genes in wheat and test for suitability in screening germplasm. Nucleotide database of the Genbank (www.ncbi.nih.nim.gov.in) was searched using keywords "abiotic stress" and "flowering plants" in December of 2011. The text-mining was supervised to choose only those genes that were (i) published in journals; (ii) with complete coding sequences; and (iii) reported in two or more species. Hypothetical, predicted, putative, partial, unpublished, directly submitted, precursor, exon-only, duplicates of same gene with different sizes and genomic scaffolds were therefore avoided. Sequences from sources like clones, less studied cultivars and naturally interspecific hybrids were also not downloaded from database. The text-mining resulted in 132 putative abiotic stress tolerance genes belonging to 37 plant species. Among the candidates shortlisted, 18 genes were from wheat. The stress-pathway candidate genes included enzymes (e.g. Oxidase and Mitogen activated protein kinase kinase), membrane proteins (e.g. Aquaporins and Copper chaperones), DNA binding proteins (e.g. Dehydration responsive element binding protein and Osmotic stress tolerance protein), ubiquitous proteins (e.g. Defensins and Profilins), and organellar proteins (e.g. chloroplastic Ferritin and mitochondrial ATP-synthase).

Primer 3.0 software tool (fokker.wi.mit.edu/primer3) was used to design heterologous primers for all the 132 genes. Amplification was tested in two wheat genotypes – heat resistant Raj-3765 (IC0532594) and heat susceptible PBW-343 (IC0532751). DNA extraction, PCR amplification and electrophoresis were carried out as described earlier (Archak *et al*., 2003). Amplification patterns are shown in Fig. 1A. Based on clear and single band amplification, 24 STS loci were shortlisted to screen 13 genotypes known to be heat and drought tolerant along with Agra Local, the susceptible

**Author for Correspondence: Email- sunil.archak@icar.gov.in*

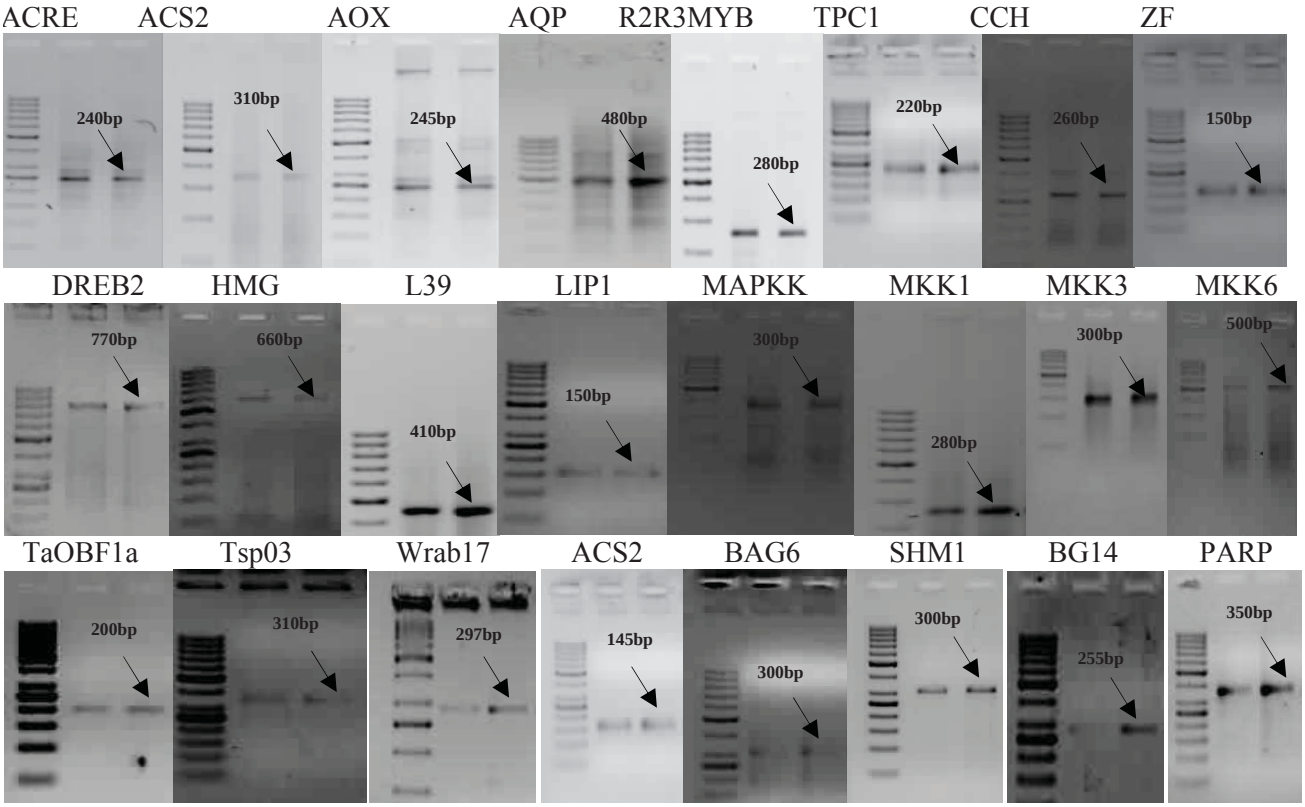

Figure 1A: Amplification profile of 24 STS loci.
Gene names are mentioned on the top of each gel picture. Lane: 1: 50/100 bp ladder, 2: Raj3765, 3: PBW343.

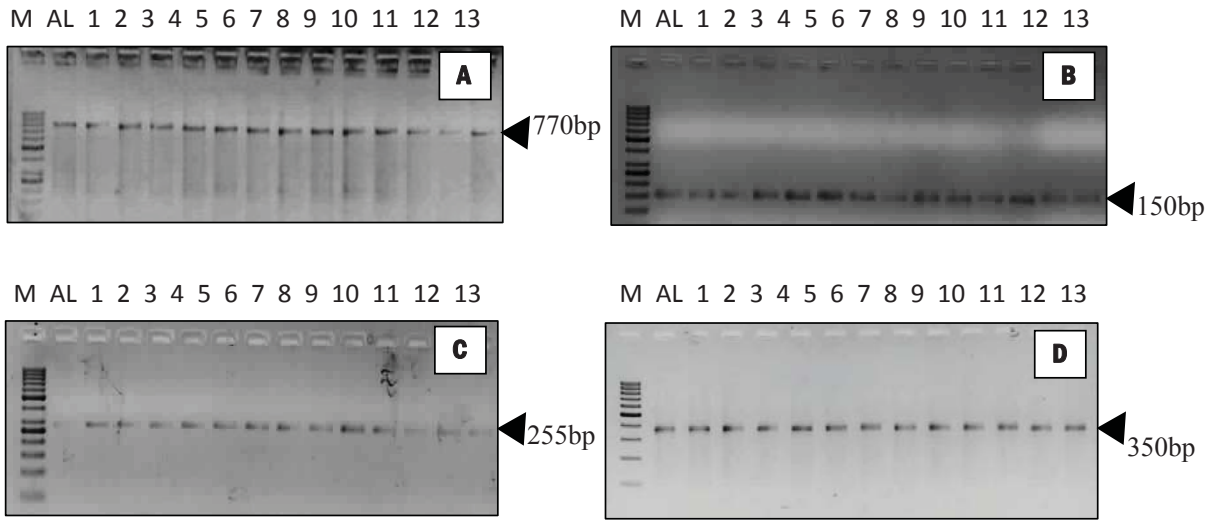

Figure 1B: Amplification profile of four candidate genes in 14 wheat accessions
Candidate genes: (A) LIP1 (B) TaOBF1a (C) Wrab17 (D) SHM1. Lanes: M: 50bp ladder, AL: Agra Local, 1: AKAW-3717, 2: HTW 6, 3: HTW 11, 4: WCF8-HT13, 5: Hindi 62, 6: KRL 99, 7: KRL 3-4, 8: Kharchia Local, 9: Kharchia 65, 10: WCF 12-7, 11: WCF 12-61, 12: WCF 12-19, 13: WCF 12-208.

control (Table 1). DNA extraction, PCR amplification and electrophoresis were carried out as described earlier (Archak *et al.*, 2003). It was observed that amplicons did not exhibit discernible length polymorphism (Fig 1B). Amplicons of BG14 (Non-specific lipid transfer protein), SHM1 (Serine hydroxyl methyltransferase 1), LIP1 (GDSL-lipase protein), TaOBF1a (Basic leucine zipper protein) and Wrab17 (Late embryogenesis

Table 1. Details of wheat genotypes employed in marker screening

S.No.	Accesion No.	Name	Biological Status	Unique feature
1	IC582907	AKAW-3717	Genetic Stock	Early heat and drought tolerance
2	IC29007A	HTW 6	Others	Terminal heat tolerance
3	IC35117	HTW 11	Others	Terminal heat tolerance
4	IC443622	WCF8-HT13	Genetic Stock	Drought and lodging resistance; WUE
5	IC296681	Hindi 62	Others	Heat and drought tolerance
6	IC546936	KRL 99	Genetic Stock	Salinity and sodicity tolerance
7	IC408331	KRL 3-4	Genetic Stock	Salt tolerance greater than Kharchia 65
8	IC296742	Kharchia Local	Others	Salt tolerance
9	IC335540	Kharchia 65	Released cultivar	Salt tolerance
10	IC594376	WCF12-7	Genetic Stock	Drought tolerance
11	IC594377	WCF12-61	Genetic Stock	Drought tolerance
12	IC594378	WCF12-19	Genetic Stock	Drought tolerance
13	IC594379	WCF12-208	Genetic Stock	Drought tolerance
14	IC138332	Agra Local	Landrace	Susceptible control

abundant protein) were gel-extracted using QIAquick columns and were sequenced (single pass) to find out the existence of sequence polymorphism. The chromatogram analysis was done using Finch TV (www.geospiza.com/Products/finchtv.shtml) and sequences were aligned using BIOEDIT (www.mbio.ncsu.edu/bioedit/bioedit.html). No sequence polymorphism was observed.

Present study could list 132 candidate genes involved in various pathways linked to response to heat and drought in plants. Primers were designed and successful amplification of these candidate genes in wheat was demonstrated. Neither fragment length nor sequence polymorphism was observed in selected 24 loci. Considering the indispensable nature of each of the gene shortlisted, this result was not unexpected. Based on the limited data generated by the study, it was concluded that the candidate genes could only be shortlisted based on experiments to detect differential expression. However, the study opens up possibility of employing candidate genes for validating germplasm accessions identified as tolerant to heat or moisture stress in field experiments.

Acknowledgements

The research work was funded by sponsored grants of the National Innovations in Climate Resilient Agriculture (NICRA) project of the Indian Council of Agricultural Research. Sunil Archak is supported by the ICAR-National Fellowship.

References

Agarwal KP and B Jha (2010) Transcription factors in plants and ABA dependent and independent abiotic stress signaling. *Biol. Plantarum* **54**: 201-212.

Archak S, AB Gaikwad, D Gautam, EVVB Rao, KRM Swamy and JL Karihaloo (2003) Comparative assessment of DNA fingerprinting techniques (RAPD, ISSR and AFLP) for genetic analysis of cashew (*Anacardium occidentale* L.) accessions of India. *Genome* **46:** 362-369.

Asseng *et al.* (2015). Rising temperatures reduce global wheat production. *Nat. Clim. Chang.* **5:** 43–147.

Bansal KC, RK Tyagi, M Dutta, BS Phogat, and S Archak (2013). Strategies for enhanced utilization of *ex situ* wheat genetic resources to boost wheat productivity: Indian initiatives. In Proceedings of the Regional Consultation on Improving Wheat Productivity in Asia, 48 p.

Benson DA, K Clark, I Karsch-Mizrachi, DJ Lipman, J Ostell, and EW Sayers (2015) *GenBank. Nucleic Acids Res*. **43**(Database issue), D30.

Moreau Y and LC Tranchevent (2012) Computational tools for prioritizing candidate genes: boosting disease gene discovery. *Nat. Rev. Genet.* **13:** 523-536.

Singh AK, PK Singh, M Arya, NK Singh and US Singh (2015). Molecular screening of blast resistance genes in rice using SSR markers. *Plant Pathol. J.* **31:** 2-24.

Archak S, JC Rana, P Singh and AB Gaikwad (2016). Potential of gene-specific sequence-tagged-sites (STS) as trait specific markers in buckwheat (*Fagopyrum* spp.). *J. Plant Biochem. Biot*. (in press).

Zhu F, P Patumcharoenpol, C Zhang, Y Yang, J Chan, A Meechai, W Vongsangnak, and B Shen (2013). Biomedical text mining and its applications in cancer research. *J. Biomed. inform.* **46:** 200-211.