Australian Journal of Crop Science AJCS
AJCS 9(9):799-806 (2015) ISSN:1835-2707

In-silico expressed sequence tag analysis in identification and characterization of salinity stress responsible genes in *Sorghum bicolor*.

Pavan K. Chaduvula[1], Jyotika Bhati[1], Anil Rai[1], Kishore Gaikwad[2], Soma S. Marla[3], M. Elangovan[4] and Sanjeev Kumar[1*]

[1]Centre for Agricultural Bioinformatics, ICAR- Indian Agricultural Statistics Research Institute, Pusa, New Delhi 110012, India
[2]ICAR-National Research Centre Plant Biotechnology, Pusa, New Delhi 110012, India
[3]DNA Fingerprinting Division, ICAR-National Bureau for Plant Genetic Resources, IARI Campus, New Delhi 110012, India
[4]Plant Variety Protection, ICAR-Directorate of Sorghum Research, Rajendranagar, Hyderabad 500030, India

***Corresponding author: sanjeevk@iasri.res.in**

Abstract

Sorghum (*Sorghum bicolor* L. *Moench*) is the fifth major cereal crop in the world. Sorghum is moderately salt-tolerant crop and different genotypes shows different levels of tolerance. Therefore, the present work emphasizes on reconstruction, validation and annotation of salt stress genes of *Sorghum*. In order to identify genes responsible for salt stress in *Sorghum*, 6749 expressed sequence tags (ESTs) were mined from the web recourses. The downloaded ESTs were clustered and assembled into 823 contigs and biological functions were acquired to the most of the contigs through Gene Ontology (GO). 617 contigs were showed homology to Rice, Maize and Barley protein sequences. The remaining contigs were mapped on *Sorghum* genome and full length gene sequences were obtained by *in-silico* approach. Total 12 candidate genes were identified as salinity stress responsive genes and amplified in nine indigenous Sorghum genotypes. These candidate genes were further validated by means of molecular techniques in nine *Sorghum* cultivars. The study claims the possible involvement of the predicted genes in salt stress mechanism and may be useful in molecular breeding programme in *Sorghum* salinity research.

Keywords: *Sorghum*, salt stress, functional annotation, gene prediction.
Abbreviations: ESTs_Expressed sequence tags; GO_Gene Ontology; cDNA_complimentary DeoxyriboNucleic Acid; ABA_Abscisic acid; JA_Jasmonic acid; SA_Salicylic acid; NaCl_Sodium chloride; NCBI_National Cetre for Biotechnology Information; dbEST_Database of Expressed sequence tags; GC_Guanine:Cytosine; nr_non-redundant; TSS_Transcription Start Site; CDS_Coding Sequence; MAPKK_Mitogen-activated protein kinase kinase; HSP_Heat Shock Protein; BLAST_Basic Local Alignment Search Tool; Asr_Abscisic stress-ripening; mRNA_messenger RiboNucleic acid; GDH_Glutamate dehydrogenase; GS-GOGAT_glutamine synthetase-glutamate synthase; TIGR_The Institute for Genomic Research; CAP_Contig Assembly Program; e-value_expected value; DAG_Directed Acyclic Graph; EC_Enzymatic Codes; KEGG_Kyoto Encyclopedia of Genes and Genomes; CTAB_Cetyl trimethylammonium bromide; PCR_Polymerase Chain Reaction; EtBr_Ethidium bromide; BLAT_BLAST-like alignment tool; dNTPs_Deoxynucleotides; MgCl2_Magnesium chloride.

Introduction

Salinity is a soil condition characterized by a high concentration of soluble salts, a common feature of arid and semi-arid lands, place a major limitation on crop productivity (Rengasamy, 2006; Munns and Tester, 2008). In India nearly 9.38 million ha area is occupied by salt-affected soils out of which 5.5 million ha are saline soils (including coastal) and 3.88 million ha alkali soils (IAB, 2000). It is important to develop salt tolerant varieties to increase crop productivity and extend the area under cultivation even on the saline lands to meet food security of the growing human population. Hence, it is important to study the salt stress mechanism at integrative and whole genome level to identify novel genes, and determine their expression patterns with the aim of improving salt tolerance. The understanding of their functions in stress adaptation will provide the basis of effective engineering strategies to improve crop stress tolerance (Cushman and Bohnert, 2000; Kawasaki et al., 2001).

Sorghum is the third most important food grain crop in India and it is also used as feed (grain and biomass), fuel (ethanol production), fibre (paper), fermentation (methane production) and fertilizer (utilisation of organic by-products). Its small genome (~730 Mb), large genomic variability in salt tolerance and diverse economic importance makes *Sorghum* an attractive model for the functional genomics of Saccharinae and other C4 grasses (Paterson et al., 2009). The availability of a complete genome sequence for *Sorghum* (Paterson et al., 2009) and *Sorghum*Cyc, a metabolic pathways database (http://www.gramene.org/pathway/Sorghumcyc.html) provide a unique opportunity to obtain a comprehensive view of the genes conditioning under abiotic stress in *Sorghum*. *Sorghum*'s response to various abiotic and biotic stresses was examined through cDNA based experiments in recent decade.

These studies included the effects of PEG-induced osmotic stress, exogenous ABA, salt, wounding (by JA), and insects (by SA) (Park et al., 2006; Salzman et al., 2005). Sorghum is known as moderately tolerant to salinity (Igartua et al., 1995). Studies reported relation between salinity stress and gene expression patterns in numerous plant species including Arabidopsis (Kreps et al., 2002; Kilian et al., 2007; Krishnaswamy et al., 2008; Yokotani et al., 2009), Rice (Kawasaki et al., 2001; Walia et al., 2005; Senadheera et al., 2009), and Barley (Zhang et al., 2004 ; Close et al., 2004; Walia et al., 2006). In sorghum, notable phenotypical changes under salinity stress takes place after 4days of growth in 200mM NaCl (Swami et al., 2011). Several sorghum salinity tolerant and sensitive genotypes were reported (Krishnamurty et al., 2007; Hefny et al., 2013). increasing salinity levels decreases the stem yeild and soluble carbohydrate levels in sorghum (Almodares et al., 2008). Studies have been reported transcriptome changes in response to dehydration and high salinity (Buchanan et al., 2005), the evaluation of growth parameters and ion accumulation (Netondo et al., 2004) and soluble carbohydrate contents (Almodares et al., 2008) among various cultivars of sorghum. In present study, single-pass sequences of randomly selected cDNA clones from *Sorghum bicolor* seedlings treated with 150mM NaCl were obtained from NCBI dbEST. In order to identify genes responsible for salt stress in *Sorghum*, ESTs induced under salinity stress condition were mined. The selected ESTs were clustered and assembled into EST contigs. Biological functions were obtained for EST contigs through gene validation and annotation. The contigs, for which GO terms were not assigned, were mapped on to the *Sorghum* genome and full length gene sequences were obtained. These candidate genes were further validated in nine *Sorghum* cultivars by means of molecular techniques.

Results

Assembling of ESTs into contigs

ESTs obtained from salinity induced cDNA library of *Sorghum bicolor* were described here. A total 6749 EST sequences for *Sorghum* bi*color* were collected from GenBank including dbEST. These ESTs were generated from different tissues and stress levels by various workers (http://www.ncbi.nlm.nih.gov/). The average length of theses ESTs is 563 base pairs. Out of these 6749 EST sequences, 59 reads were trimmed and 3 sequences discarded through the sequence cleaning process. Further during masking process 1 hobo-activator belonging to the DNA transposons group was identified and removed. Therefore, total repetitive elements masked 52147 bp i.e. about 1.37% of the total size of query sequence. The primary sequence analysis showed that total GC content of non-redundant EST collection as 53%. The remaining EST sequences for *Sorghum bicolor* were assembled into 823 EST-contigs. Most of these contigs consists of two or three ESTs. These assembled EST account for only 8.68 % of the size of total ESTs. Less abundant or lowly expressed transcripts could not be assembled into larger contigs and remained as 2673 singletons (Table 1).

Functional Categorization of the EST-Contigs

Functional characterization was achieved by comparing 823 assembled and translated EST-contigs against NCBI nr database. Out of 823 EST contigs, 718 EST contigs were selected based on homology search. Blast results showed that more than 9.52% of these EST-contigs were exclusive to

Table 1. Summary of EST analysis performed by EGAssembler.

	Features	Statistics
1	Total number of clean ESTs	6749
2	EST total nucleotides (nt)	3797111
3	Cluster count	3496
4	Singleton Unigenes	2673
5	Multi-member unigenes	823
6	Average EST length (nt)	563
7	Average cluster size (singleton exclude)	3.97
8	Average cluster size (singleton include)	1.8
9	Average GC content (%)	53

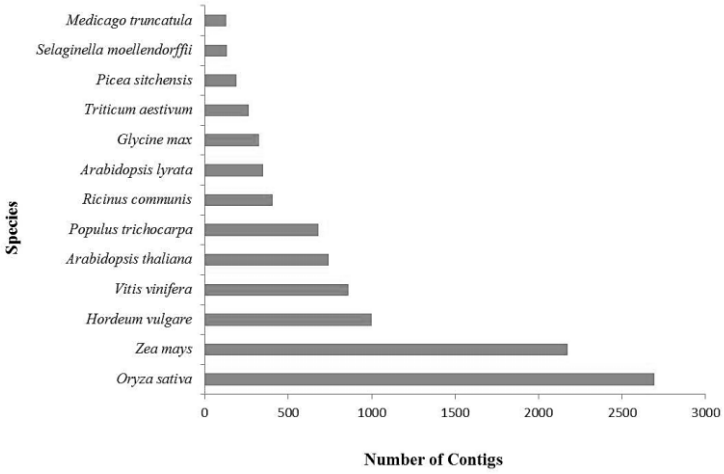

Fig 1. Species distribution of *Sorghum* bi*color* EST-contigs.

Sorghum bi*color* and upto 75% of genes were found as conserved across the higher plant species *Oryza sativa*, *Zea mays*, *Hordeum vulgare*, *Vitis vinifera* and *Arabidopsis thaliana*. The percentage match between sequences of *Sorghum* versus rice, maize and *Arabidopsis* was found to be 20.9%, 16.8% and 5.7% respectively (Fig. 1). These 718 EST contigs were further subjected to GO functional classification. Out of 718 EST contigs GO terms were available only for 681 EST contigs. It has been noticed that overall 3455 GO terms were retrieved which means on an average 5 GO terms per contig was obtained. Maximum 26 GO terms for one contigs and minimum one GO term for 68 contigs were retrieved. However 108 EST contigs having 4 GO terms. The distribution of GO terms per contigs is given in Fig. 2. Again EST-contig sequences were categorized according to the GO vocabularies i.e. Molecular Function (Fig. 3A), Cellular Component (Fig. 3B) and Biological Process (Fig. 3C) with obtained number of GO terms as 1136, 1030 and 1289 respectively.

PCR confirmation of candidate genes in Sorghum cultivars

GO terms obtained only for 681 out of 718 EST-contigs selected. The rest 37 EST contigs were considered as potential for discovery of novel genes. The complete gene structure with TSS, PolyA tails at the extremes and CDS in between were obtained for 12 EST contigs and considered as candidate genes. Remaining 25 contigs were either not aligned on genome or complete gene structure were not obtained. Coding position on chromosome, TSS, polyA tail and length of candidate genes are given in Table 2. These contigs were further used for similarity search against known candidate genes for abiotic stress tolerance. Unigene

Table 2. Coding position of the In-silico validated candidate genes.

Contig	Gene	Chromosome No/ Position	Length	TSS	CDS*	CDS Start	CDS Stop	Poly A
Contig 60	Abscisic stress ripening protein 2 (ASR2)	6:45323249-45324790	1542	1289	CDSl	339	545	257
					CDSf	623	1225	
Contig 379	Amino acid transporter	4:6076894-6079962	3069	2913	CDSl	321	453	145
					CDSi	571	1036	
					CDSi	1120	1323	
					CDSi	1406	1558	
					CDSf	2054	2465	
Contig 171	Dehydrin	9:46037861-46039248	1388	125	CDSf	284	514	1310
					CDSl	647	874	
Contig 159	Glutamate dehydrogenase	6:53263037-53267060	4024	620	CDSf	883	990	3767
					CDSi	1080	1193	
					CDSi	1956	2177	
					CDSi	2457	2573	
					CDSi	2754	3004	
					CDSi	3207	3282	
					CDSi	3365	3451	
					CDSl	3542	3727	
Contig 281	HSP70	8:49116775-49121517	4743	473	CDSf	611	824	4575
					CDSl	2707	4442	
Contig 452	HSP90	2:63248022-63252131	4110	3720	CDSl	233	1783	39
					CDSi	2747	3169	
					CDSf	3406	3531	
Contig 287	MAPKK1	4:64242735-64247700	4966	300	CDSf	607	1856	4605
					CDSi	2589	2673	
					CDSi	2801	2863	
					CDSi	3523	3582	
					CDSi	3924	4028	
					CDSl	4364	4543	
Contig 738	Phytoene synthetase	10:60688676-60692539	3864	3773	CDSl	343	516	289
					CDSi	617	809	
					CDSi	913	1148	
					CDSi	1803	2002	
					CDSi	2588	2638	
					CDSf	2731	3163	
Contig 169	Proline dehydrogenase	1:51739904-51742879	2976	2618	CDSl	176	226	28
					CDSi	287	1243	
					CDSf	1647	2360	
Contig 759	Serine/threonine kinase	5:15061067-15067115	6049	1263	CDSl	315	338	59
					CDSi	476	574	
					CDSf	639	776	
				5973	CDSl	3149	3328	1817
					CDSi	4029	4127	
					CDSi	4853	4912	
					CDSi	5407	5669	
					CDSf	5785	5941	
Contig 64	Sorbitol transporter	8:44316372-44318810	2439	2130	CDSl	239	1198	14
					CDSi	1299	1645	
					CDSf	1726	1867	
Contig 17	Trehalose 6-p-phosphatase	2:58283549-58289015	5467	107	CDSf	241	291	5244
					CDSi	899	1130	
					CDSi	1215	1554	
					CDSi	1969	4002	
					CDSi	4150	4426	
					CDSi	4548	4864	
					CDSl	4978	5074	

*CDSf-First coding segment, CDSi-Internal Coding Segment, CDSl- Last Coding Segment.

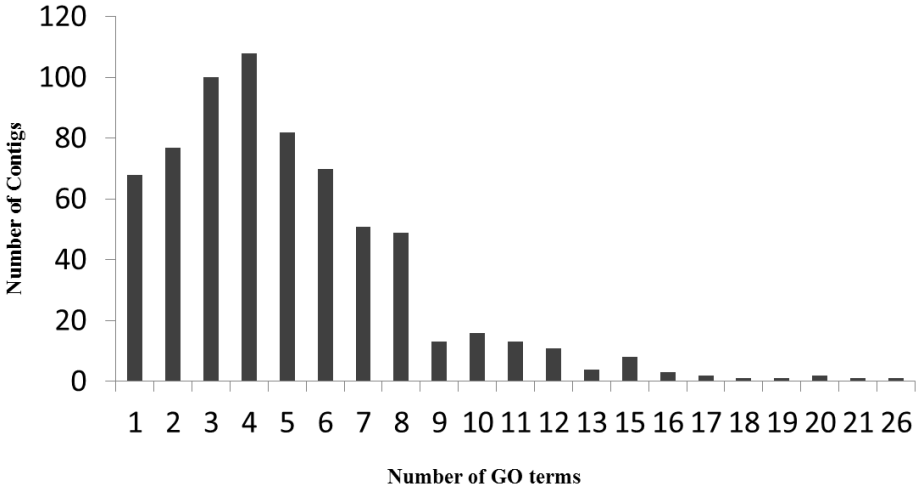

Fig 2. Distribution of number of EST-contigs vs number of GO terms obtained.

sequences showing significant match with a candidate gene were selected and used for primer designing using primer3 software (Table 3). All seven abiotic stress responsive candidate genes were partially amplified in nine diverse *Sorghum* cultivars using gene specific primer pairs. About ~678 bp long Abscisic stress-ripening protein 2 genes was amplified in all cultivars except ERN-11 (Fig. 4A). Glutamate dehydrogenase homologue was amplified using Glutamate dehydrogenase specific primer pair designed considering its predicted gene sequence in *Sorghum*. All nine cultivars showed approximate amplicon size of Glutamate dehydrogenase was ~585 bp (Fig. 4B). Amino acid transporter gene homologue, about 525 bp long, was isolated from eight *Sorghum* genotypes except ERN-9 using primers designed for showing match with predicted candidate gene (Fig. 4C). Amplification of MAPKK1 homologue yielded a product of ~356 bp (Fig. 4D). Except ERN-249 all the cultivars have presence of MAPKK1 gene. Phytoene synthetase gene homologue (about 563 bp long) was isolated from all nine *Sorghum* genotypes using primers designed for showing match with predicted candidate gene (Fig. 4E). Approximate amplicon size of serine/threonine kinase gene was ~ 690bp (Fig. 4F). Three *Sorghum* cultivars viz., ERN-11, ERN-9, ERN-21 showed amplifications of this gene. Approximate amplicon size of the HSP90 gene was ~793 bp in all nine cultivars (Fig. 4G).

Discussion

Sorghum is considered as moderately salt tolerant field crop. Several studies have been reported on the large scale screening of *Sorghum* varieties for salt tolerance (Krishnamurthy et al., 2007), ion accumulation (Netondo et al., 2004), soluble carbohydrate contents (Almodares et al., 2008a) and transcriptome changes in response to dehydration and high salinity (Johnson et al., 2014; Buchanan et al., 2005). These studies provide valuable information on the effect of salt stress on plant growth and the accumulation of soluble sugars which may function as osmoprotectants. However, there is need to identify specific stress-regulated genes that contribute to *Sorghum*'s salt tolerance mechanisms. An important genomic approach to identify salt stress related genes is based on ESTs generated from different cDNA libraries representing stress treated tissues collected at various stages of development. The clustering of EST sequences generated from salt stress-treated cDNA libraries provides information on gene number, gene content and possible number of gene families involved in stress responses. Putative functions are assigned to such stress-responsive genes by BLASTX comparison to the protein database. This type of analysis provides a valuable resource of information regarding genes associated with stress-responsive mechanism. In this study assembled EST contigs were assigned GO categories for molecular function, biological activity and cellular component (Fig 3). Abscisic stress-ripening protein 2 gene was amplified in all the *Sorghum* cultivars undertaken in this study except ERN-11. It is noted that in *Solanum lycopersicum* Asr1 and Asr2 expression levels increased in dehydrated leaves (whole organ), but only Asr2 levels rose in roots (Maskin et al., 2001). Maize Asr as genes linked to drought resistance (de-Vienne et al., 1999), and the divergence of tomato Asr paralogs due to the adaptation to a hostile environment (Frankel et al., 2006) strongly suggest the role of these proteins in the plant response to environmental signals (Saumonneau et al., 2008). Constitutive expression of MpAsr in *Arabidopsis* helped the plant to reduce cell membrane damage and increase cytoplasmic osmoprotectants in response to stress (Dai et al., 2011). In rice, Asr5 mRNA is up-regulated by cold, drought, high-salinity and ABA (Rabbani et al., 2003). Similarly partial amplification of Asr2 in *Sorghum bicolor* confirms the expression under salinity stress. Glutamate dehydrogenase homologue was amplified using Glutamate dehydrogenase specific primer pair designed considering its predicted gene sequence in *Sorghum*. All nine cultivars showed approximate amplicon size of Glutamate dehydrogenase was ~585 bp (Fig. 4B). Glutamate dehydrogenase present in plants serves as a link between carbon and nitrogen metabolism due to the ability to assimilate ammonium into glutamate or deaminate glutamate into 2-oxoglutarate and ammonium (Forde and Lea, 2007). Excesses of Na^+ and Cl^- also might change the pathway of NH_4^+ assimilation in old leaves of salt stressed rice plants, weaken GOGAT/GS pathway and elevate GDH pathway (Wang et al., 2012). About 525 bp long amino acid transporter gene homologue was isolated from eight *Sorghum* genotypes except ERN-9 using primers designed for showing match with predicted candidate gene (Fig. 4C). Amplification of MAPKK1 homologue yielded a product of ~356 bp (Fig. 4D). Except ERN-249 all the cultivars have presence of

Table 3. List of primers used to amplify *Sorghum* candidate genes related to salinity stress response.

S.No.	Gene	Primer Sequence (5′-3’)	Fragment Size (bp)
1.	ABA Induced protein	F: AGT AGC GAA GAG CAG CAG; R: ACT ACG ACG GTG GGTACA	678
2.	Glutamate dehydrogenase	F: GAT CAC TGG GTA GGG ATG; R: GAG AGA GGC CAA TGA AAC	585
3.	Amino acid transporter	F: GTC AAT TCG AGG TCA AGG; R: TCG AGA CCA CAC CAA AGT	525
4.	MAPKK1	F: ATG TGC ACG CAC AAG TAG; R: GAT GGA GGT GAT CAA AGC	356
5.	Phytoene synthetase	F: GAA GCG AGA GGC AAT AAG; R: CCT GTT GCA GGA AGA ATC	563
6.	Serine/Threonine kinase	F: GAA CCG TAT GTC CCA GTT; R: AAC AAA CCT ACC GGT GAC	690
7.	HSP90	F: CCT TCT ACT CCA ACA AGG AG; R: CTA TCA CAG CTG CAG ATC AC	793

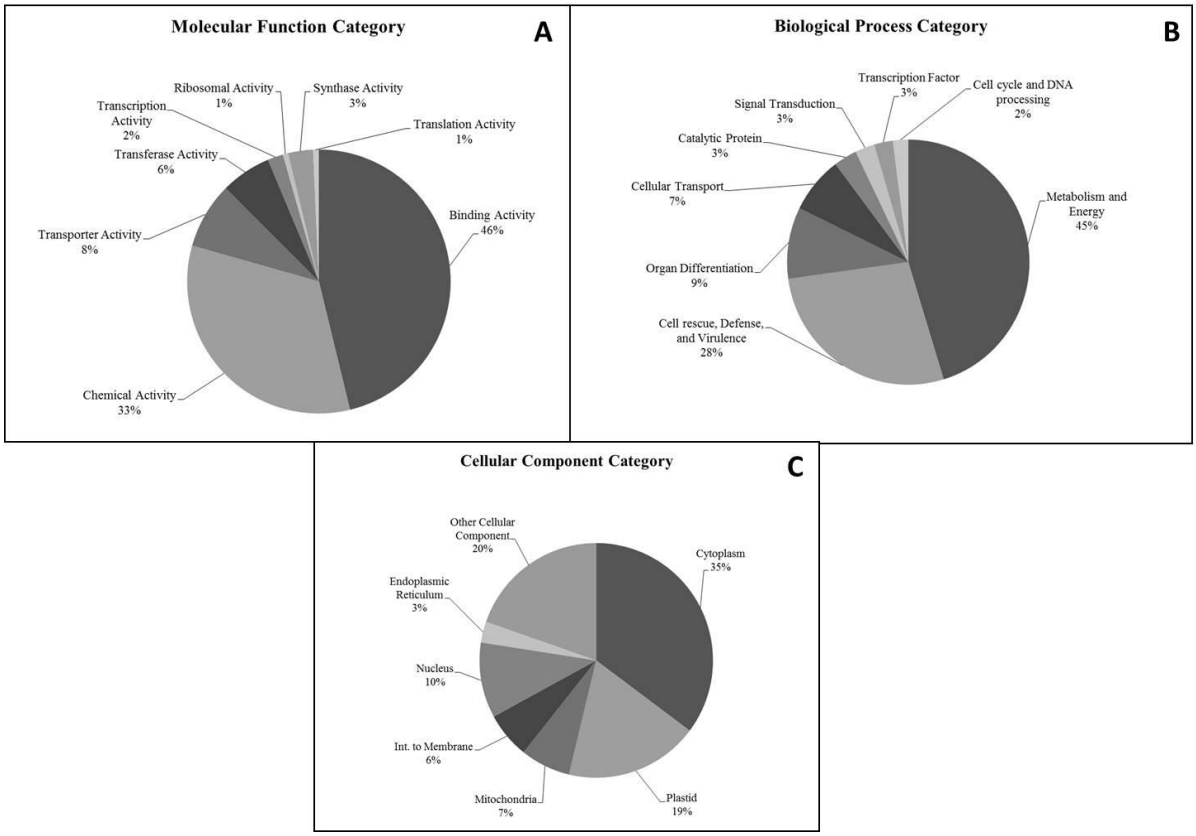

Fig 3. Distribution of GO terms in the (A) Molecular Function, (B) Biological Process and (C) Cellular Component category.

Table4. Indian sorghum cultivars used for identification of salt responsive genes.

S.No	Cultivar Accession	Location
1	E-179	Hyderabad
2	ERN-11	Hyderabad and Udaipur
3	E-250	Hyderabad and Udaipur
4	E-122	Hyderabad and Udaipur
5	ERN-12	Hyderabad, Udaipur and Mauranipur
6	ERN-9	Hyderabad and Udaipur
7	ERN-21	Hyderabad and Udaipur
8	ERN-124	Hyderabad, Udaipur and Mauranipur
9	ERN-249	Hyderabad and Udaipur

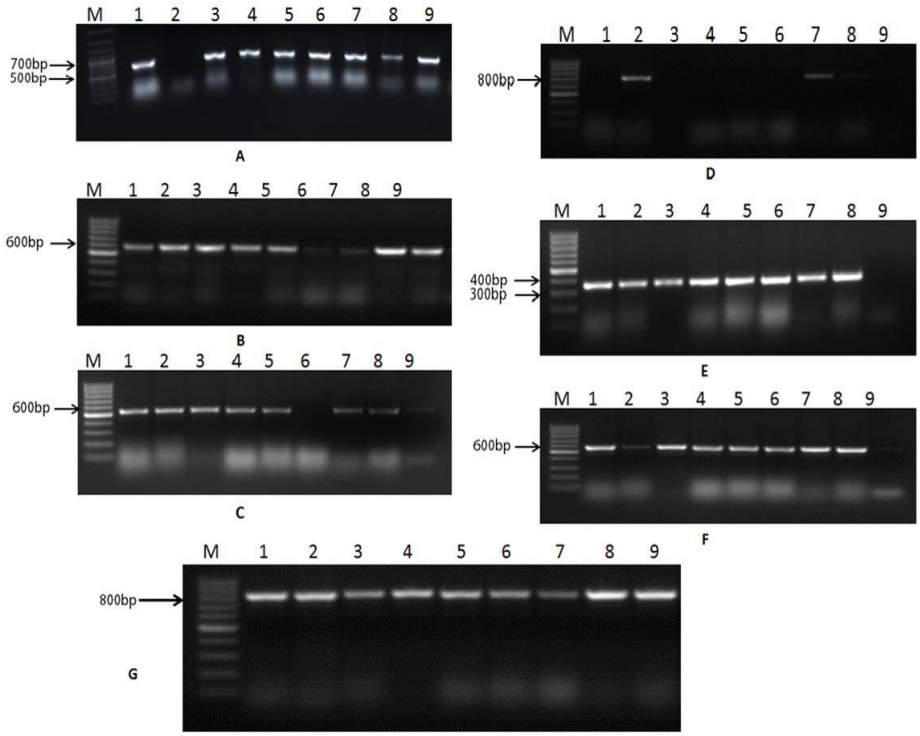

Fig 4. PCR amplification of seven salinity stress responsive genes from nine diverse *Sorghum* genotypes **A.** *ABA-Stress-ripening-2 (ASR2)* **B**. *Glutamate dehydrogenase* **C.** *Amino acid transporter* **D**. *MAPKK1* **E**. *Phytoene synthetase* **F**. *Serine /Threonine Kinase* **G.** *HSP90: Lane no: 1: E-179 2: ERN-11 3: E-250; 4: E-122; 5: ERN-12; 6:ERN-9; 7: ERN-21; 8:ERN-124; 9: ERN-249.*

MAPKK1 gene. Different MAPK elements are activated under salinity stress such as MAPK4, MAPK6, and MAPKK1 (Taj et al., 2010). AtMAPK1 negatively regulates a putative Na+/H+ antiporter, leading to salinity sensitivity (Chinnusamy et al., 2006). About 563 bp long phytoene synthetase gene homologue was isolated from all nine *Sorghum* genotypes using primers designed for showing match with predicted candidate gene (Fig. 4E). Approximate amplicon size of serine/threonine kinase gene was ~ 690bp (Fig. 4F). Three *Sorghum* cultivars viz., ERN-11, ERN-9, ERN-21 showed amplifications of this gene. Serine/threonine protein kinase, is a positive regulator of plant tolerance to salt stress and well characterized in *Glycine soja* (Sun et al., 2013). Approximate amplicon size of the HSP90 gene was ~793 bp in all nine cultivars (Fig. 4G). Salt stress causes an augmentation of HSP90 family proteins with in tomato roots (Manaa et al., 2011). Expression of HSP90 in *Arabidopsis* is developmentally regulated and is responsive to heat, cold, salinity, heavy metals, phytohormones, and light and dark transitions (Krishna and Gloor, 2001). ERN-11 is most tolerant than all the other genotypes by showing amplification of all the genes which represent salinity stress response. ERN-249 is susceptible to salinity stress by showing non-amplification of major salinity response genes including Amino acid Transporter, Phytoene synthetase and Serine/Threonine Kinase.

Materials and Methods

Data collection and assembly

A total of 6749 ESTs, both 5′ and 3′, were collected from *Sorghum* bi*color* cDNA library derived from NCBI dbEST. The collected ESTs were subjected to trimming process, which include the removal of low quality sequences, poly(A) tails, ribosomal RNA and vector regions. The trimming process was conducted through EGAssembler (Masoudi-Nejad et al., 2006) and vector regions were screened using NCBI's core vector library. The resulted ESTs were scanned for removal of mitochondrial chloroplast sequences and for repeat masking process using TIGR *Sorghum* repeat library. The cleaned ESTs were assembled into EST-contigs with improved length and quality using CAP3 (Huang and Madan, 1999) software with default parameter values (gap penalty factor N>6, segment pair score cut-off N>40, overlap percent identity cut-off N>65 etc). Only EST-contigs with e-value $<10^{-20}$ were considered for further analysis. GC percentage, average length and length range of contigs were found out using a custom developed perl script. Number of clustered sequences present in different contigs was obtained through analysis of CAP3 assembly files (.ace).

Functional annotation

In silico functional characterization of the assembled EST-contigs was carried out by using Blast2GO v.2.5.0 software (Conesa and Gotz, 2008). At the first step of annotations, all the contigs were compared to the Gene Bank nr sequence database (REF) by BLAST (BLASTX 2.2.26+) (Altschul et al., 1997) with BLOSUM62 scoring matrix. The parameters for BLASTX *i.e.* high-scoring segment pair (HSP) cut-off and e-value threshold were set at 33 and 1e-3 respectively. Based on BLASTX results, the GO terms for each EST-contigs were extracted at level 3 of the DAG. EST-contigs were assigned the molecular function, biological process and cellular localization. At the second step conserved domains present in the EST-contigs were identified by screening with InterProscan (Quevillon et al., 2005), which is a database that contains information about domains, motifs and conserved regions in protein families. At the last step of annotation, ECs were obtained based on gene ontology of the EST-contigs and

their associated metabolic pathways were mapped using KEGG database.

Plant material and DNA extraction

Plant materials for present study were comprised of sorghum genotypes from Hyderabad, Udaipur, and Mauranipur of India. Nine diverse *Sorghum* genotypes namely E-122, E-179, E-250, ERN-9, ERN-11, ERN-12, ERN-21, ERN-124 and ERN-249 obtained from Directorate of *Sorghum*, Hyderabad (Table 4). Plants were grown in pots under controlled conditions, used for the amplification of salinity stress responsive candidate genes. Total genomic DNA was extracted from 2-week old seedlings using a modified CTAB method (Doyel and Doyel, 1987). The DNA concentration was measured by Thermo scientific Nano Drop 8000 Spectrophotometer and adjusted 50ng/µl and stored at 4 ºC to be used as PCR templates for PCR amplification. Integrity and quality of DNA was checked by electrophoresis on 0.8%(w/v) subsequent EtBr staining.

Identification of salinity stress responsive candidate genes

The EST-contigs, for which GO terms were not assigned, were aligned on genome of *Sorghum bicolor* using BLAT (Kent, 2002). The length of the aligned EST-contigs on *Sorghum* genome was further extended 1 kb up and down stream in order to predict the complete gene structure. Computational gene prediction program, FGENESH (Salamov and Solovyev, 2000), was used for prediction of complete gene structure with TSS, PolyA tails at the extremes and CDS in between. These contigs were then used for similarity search against known candidate genes for abiotic stress tolerance. Unigene sequences showing significant match with a candidate gene were selected and used for primer designing using primer3 software (Table 3). PCR amplification for all the genes was standardized and carried out for nine diverse *Sorghum* cultivars in TECHNO gradient thermal cycler in 15 µl reaction containing 10 ng template DNA, 0.2 mM dNTPs, 2 mM $MgCl_2$, 10 pmol of each of forward and reverse primer, 0.5 U of Taq DNA polymerase (vivantis) and 1X PCR buffer. The amplification profile included: initial denaturation for 5 min at 94°C followed by 40 cycles of denaturation for 40 s at 94°C, annealing at temperature specific for each target gene for 40 s and extension at 72°C for 1 min 20 s. Final extension was allowed for 10 min at 72°C and storage at 4°C. Amplified PCR products were resolved on 1.2% agarose gels.

Conclusion

Present study was undertaken with an objective to analysis of the ESTs with putative functional annotations and amplification of candidate genes in *Sorghum bicolor*. Large numbers of putative salt stress responsive genes were identified in *Sorghum*. Seven salt stress responsive candidate genes validated for their significance in stress responses *Sorghum*. Therefore, present study provided basic information about some salt stress responsive genes in *Sorghum* that can be exploited in overcoming various salt stress related problems limiting its production and subsequently in breeding superior varieties giving better yield under salt stress conditions. Through Results of this study will improve understanding of the molecular mechanisms of salinity tolerance in *Sorghum bicolor*.

Acknowledgement

We acknowledge support from National Agricultural Bioinformatics Grid (NABG), National Agricultural Innovation Project (NAIP), Indian Council for Agricultural Research (ICAR).

References

Almodares A, Hadi MR, Ahmadpour H (2008) *Sorghum* stem yield and soluble carbohydratesunder different salinity levels. Afr J Biotechnol. 7: 4051-4055.

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. Nucl Acids Res. 25: 3389-3402.

Buchanan CD, Lim SY, Salzman RA, Kagiampakis L, Morishige DT, Weers BD, Klein RR, Pratt LH, Cordonnier-Pratt MM, Klein PE, Mullet JE (2005) Sorghum bicolor's transcriptome response to dehydration, high salinity and ABA. Plant Mol Biol. 58(5): 699–720.

Cushman JC, Bohnert HJ (2000) Genomic approaches to plant stress tolerance. Opin Plant Biol. 3(2): 117-124.

Kawasaki S, Borchert C, Deyholos M, Wang H, Brazille S, Kawai K, Galbraith D, Bohnert HJ (2001) Gene expression profiles during the initial phase of salt stress in rice. Plant Cell. 13: 889-905.

Close TJ, Wanamaker SI, Caldo RA, Turner SM, Ashlock DA, Dickerson JA, Wing RA, Muelbauer GJ, Kleinhofs A, Wise RP (2004) A new resource for cereal genomics: 22K barley GeneChip comes of age. Plant Physiol. 134: 960–968.

Conesa A, Gotz S (2008) Blast2GO: A Comprehensive Suite for Functional Analysis in Plant Genomics. Int J Plant Genomics. 1-13.

de-Vienne D, Leonardi A, Damerval C, Zivy M (1999) Genetics of proteome variation for QTL characterization: application to drought-stress responses in maize. J Exp Bot. 50: 303-309.

Chinnusamy V, Zhu J, Zhu JK (2006) Salt stress signalling and mechanisms of plant stress tolerance. Genet Eng. 27: 141-177.

Dai JR, Liu B, Feng DR, Liu H, He YM, Qi KB, Wang HB, Wang JF (2011) MpAsr encodes an intrinsically unstructured protein and enhances osmotic tolerance in transgenic *Arabidopsis*. Plant Cell Rep. 30: 1-12.

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochemical Bulletin. 19: 11-15

Frankel N, Carrari F, Hasson E, Iusem ND (2006) Evolutionary history of the Asr gene family. Gene. 378: 74-83.

Forde BG, Lea PJ (2007) Glutamate in plants: metabolism, regulation, and signalling. J Exp Bot. 58(9): 2339–2358.

Hefny MM, Metwali EMR, Mohamed AI (2013) Assessment of genetic diversity of sorghum (Sorghum bicolor L. Moench) genotypes under saline irrigation water based on some selection indices. Aust J Crop Sci. 7 (12): 1935-1945.

Huang X, Madan A (1999) CAP3: A DNA sequence assembly program. Genome Res. 9: 868-877.

IAB (2000) Indian Agriculture in Brief. (27th edition). Agriculture Statistics Division, Ministry of Agriculture, Govt. of India, New Delhi.

Igartua E, Gracia MP, Lasa JM (1995) Field responses of grain sorghum to a salinity gradient. Field Crops Res. 42: 15–25.

Johnson SM, Lim FL, Finkler A, Fromm H, Slabas AR, Knight MR (2014) Transcriptomic analysis of Sorghum bicolor responding to combined heat and drought stress. BMC Genomics. 15: 456.

Krishnamurthy L, Serraj R, Hash CT, Dakheel AJ, Reddy BVS (2007) Screening *Sorghum* genotypes for salinity tolerant biomass production. Euphytica. 156: 15-24.

Kawasaki S, Borchert C, Deyholos M, Wang H, Brazille S, Kawai K, Galbraith D, Bohnert HJ (2001) Gene expression profiles during the initial phase of salt stress in rice. Plant Cell. 13(4): 889-905.

Kreps JA, Wu Y, Chang HS, Zhu T, Wang X, Harper J (2002) Transcriptome changes for Arabidopsis in response to salt, osmotic, and cold stress. Plant Physiol. 130: 2129–2141.

Kilian J, Whitehead D, Horak J, Wanke D, Weinl S, Batistic O, D'Angelo C, Bornberg-Bauer E, Kudla J, Harter K (2007) The AtGenExpress global stress expression data set: protocols, evaluation and model data analysis of UV-B light, drought and cold stress responses. Plant J. 50: 347–363

Krishna P, Gloor G (2001) The Hsp90 family of proteins in *Arabidopsis thaliana*. Cell Stress Chaperones. 6: 238–246.

Krishnaswamy SS, Srivastava S, Mohammadi M, Rahman MH, Deyholos MK, Kav NN (2008) Transcriptional profiling of pea ABR17 mediated changes in gene expression in Arabidopsis thaliana. BMC Plant Biol. 8: 91.

Kent WJ (2002) BLAT - the BLAST-like alignment tool. Genome Res. 12(4): 656-664.

Munns R, Tester M (2008). Mechanisms of salinity tolerance. Annu Rev Plant Biol. 59: 651–681.

Manaa HB, Ahmed B, Valot JP, Bouchet S, Aschi-Smiti M, Causse M, Faurobet (2011) Salt and genotype impact on plant physiology and root proteome variations in tomato. J Exp Bot. 62: 2797–2813.

Masoudi-Nejad A, Tonomura K, Kawashima S, Moriya Y, Suzuki M, Itoh M, Kanehisa M, Endo T, Goto S (2006) EGassembler: online bioinformatics service for large-scale processing, clustering and assembling ESTs and genomic DNA fragments. Nucl Acids Res. 34: W459-462.

Maskin L, Gudesblat GE, Moreno JE, Carrari FO, Frankel N, Sambade A, Rossi M, Iusem ND (2001) Differential expression of the members of the Asr gene family in tomato (Lycopersicon esculentum). Plant Sci. 161:739-746.

Netondo GW, Onyango JC, Beck E (2004) *Sorghum* and salinity: I. Response of growth, water relations, and ion accumulation to NaCl salinity. Crop Sci. 44: 797-805.

Paterson AH, Bowers JE, Bruggmann R, Dubchak I, Grimwood J, Gundlach H, Haberer G, Hellsten U, Mitros T, Poliakov A (2009) The *Sorghum bicolor* genome and the diversification of grasses. Nature. 457(7229): 551-556.

Park SJ, Huang YH, Ayoubi P (2006) Identification of expression profiles of *Sorghum* genes in response to green bug phloem-feeding using cDNA subtraction and microarray analysis. Planta. 223(5): 932-947.

Quevillon E, Silventoinen V, Pillai S, Harte N, Mulder N, Apweiler R, Lopez R (2005) InterProScan: protein domains identifier. Nucl Acids Res. 33: W116-20.

Rengasamy P (2006) World salinization with emphasis on Australia. J Exp Bot. 57: 1017 – 1023.

Rabbani MA, Maruyama K, Abe H, Khan MA, Katsura K, Ito Y, Yoshiwara K, Seki M, Shinozaki K, Yamaguchi-Shinozaki K (2003) Monitoring expression profiles of rice genes under cold, drought, and high-salinity stresses and abscisic acid application using cDNA microarray and RNA get-blot analyses. Plant Physiol. 133: 1755–1767.

Salzman, RA, Brady JA, Finlayson SA, Buchanan CD, Summer EJ, Sun F, Klein PE, Klein RR, Pratt LH (2005) Transcriptional profiling of *Sorghum* induced by methyl jasmonate, salicylic acid, and aminocyclopropane carboxylic acid reveals cooperative regulation and novel gene responses. Plant Physiol. 138(1): 352-368.

Senadheera P, Singh RK, Maathuis FJM (2009) Differentially expressed membrane transporters in rice roots may contribute to cultivar dependent salt tolerance. J Exp Bot. 60: 2553–2563.

Swami AK, Alam SI, Sengupta N, Sarin R (2011) Differential proteomic analysis of salt stress response in Sorghum bicolor leaves. Environ Exp Bot. 71: 321–332.

Saumonneau A, Agasse A, Bidoyen MT, Lallemand M, Cantereau A, Medici A, Laloi M, Atanassova R (2008) Interaction of grape ASR proteins with a DREB transcription factor in the nucleus. FEBS Lett. 582(23-24): 3281-3287.

Sun XL, Yu QY, Tang LL, Ji W, Bai X, Cai H, Liu XF, Ding XD, Zhu YM (2013) GsSRK, a G-type lectin S-receptor-like serine/threonine protein kinase, is a positive regulator of plant tolerance to salt stress. J Plant Physiol. 170(5): 505-515.

Salamov A, Solovyev V (2000) Ab initio gene finding in *Drosophila* genomic DNA. Genome Res. 10: 516-522.

Taj G, Agarwal P, Grant M, Kumar A (2010) MAPK machinery in plants: Recognition and response to different stresses through multiple signal transduction pathways. Plant Signal Behav. 5: 1370-1378.

Walia H, Wilson C, Condamine P, Liu X, Ismail AM, Zeng L, Wanamaker SI, Mandal J, Xu J, Cui X, Close TJ (2005) Comparative transcriptional profiling of two contrasting rice genotypes under salinity stress during the vegetative growth stage. Plant Physiol. 139(2): 822-835.

Walia H, Wilson C, Wahid A, Condamine P, Cui X, Close T (2006) Expression analysis of barley (Hordeum vulgare L.) during salinity stress. Funct Integr Genomics. 6: 143–156

Wang H, Zhang M, Guo R, Shi D, Liu B, Lin X, Yang C (2012) Effects of salt stress on ion balance and nitrogen metabolism of old and young leaves in rice (*Oryza sativa* L.). BMC Plant Biol. 12: 194.

Yokotani N, Ichikawa T, Kondou Y, Matsui M, Hirochika H, Iwabuchi M, Oda K (2009) Tolerance to various environmental stresses conferred by the salt-responsive rice gene ONAC063 in transgenic Arabidopsis. Planta. 229: 1065–1075.

Zhang H, Sreenivasulu N, Weschke W, Stein N, Rudd S, Radchuk V, Potokina E, Scholz U, Schweizer P, Zierold U, Langridge P, Varshney RK, Wobus U, Graner A (2004) Large-scale analysis of the barley transcriptome based on expressed sequence tags. Plant J. 40: 276–290.