3.0. MATERIALS AND METHODS

Drought stress is one of the major limitations to crop productivity. A basic understanding of plant molecular, physiological and genetic responses to environmental stress conditions, and the development of approaches towards improving stress tolerance crops is need of the hour. Tolerant species may have specific adaptive mechanism and express novel stress response genes. A challenge remains for the identification of stress responsive genes, elucidating their roles in imparting stress tolerance.

In this direction, our parent laboratory has developed a high quality drought stressed subtractive cDNA library from drought tolerant safflower (*Carthamus tinctorius* L.) cultivar A-1 by using subtractive hybridization method (Thippeswamy, 2007). Safflower is considered as one of the important oil seed crops of arid and semi-arid regions. A total 667 ESTs were deposited in the public database (Accession # GT155276-GT155549 and GW584000-GW584264), and library represents the comprehensive stress response transcriptome of safflower. Of the generated ESTs, 73% ESTs have homology with known proteins and 27% ESTs didn’t have homologous from the database (Thippeswamy et al., 2012). The libraries prepared from tissue exposed to stress
conditions were found to be valuable tools to obtain the expressed and stress-regulated portion of the genome. Though a significant portion of these safflower genes code for transporters, osmolytes, chaperones, detoxification enzymes and defense proteins, it is striking that approximately 15% of the transcriptome codes for upstream regulatory genes, including transcription factors, stress responsive kinases and other known signaling molecules. These upstream genes are widely regarded as ideal candidates that might impart stress tolerance to crop plants.

The present study was undertaken to identify EST-SSRs, ILPs, heat shock proteins and analysis of putative candidate genes from drought expressed sequence tags, and expression pattern of abiotic stress responsive WRKY transcription factors from safflower. To full fill these objectives, the following experiments were conducted.

**Sequence resources and ESTs sequences assembly**

A total 667 EST sequences generated in our laboratory (Thippeswamy et al., 2012) was systematically annotated and analyzed using various bioinformatics tools and free online software. All EST sequences of safflower were assembled by CAP3 software (http://genome.cs.mtu.edu/sas.html). The non-redundant unigene sequences from drought responsive ESTs were used to identify microsatellites.

**Identification of EST-SSR and amino acid distribution with trinucleotide repeats in safflower drought responsive EST library**

Online software www.ssrlocator.com was used to search SSRs (motifs ranging from 1 to 10 nucleotides) in order to evaluate the pattern of EST-SSR distribution. The repeat number parameters were as follows: \( \geq 20 \) for mono, \( \geq 10 \) di, \( \geq 7 \) for tri, \( \geq 5 \) for tetra, \( \geq 4 \) for penta and hexa and \( \geq 3 \) for hepta, octa, nona and deca nucleotide respectively. The
space between imperfect SSRs were kept at 5. After predicting the amino acid composition of ESTs containing trinucleotide tandem repeats, the number of amino acid loci and number of amino acid repeats were analyzed in drought EST library.

**Assessment of functional relevance of unigenes having SSRs**

Unigene sequences from safflower drought responsive ESTs containing microsatellites were used for similarity search using Blast2GO piled (Conesa and Gotz, 2008). To identify their putative function, the sequences were run against the non-redundant (nr) protein database of the NCBI (http://www.ncbi.nlm.nih.gov/blast); the obtained hits were compiled (Conesa et al., 2005). Unigene sequences not showing any match were considered as unique to the safflower species.

**Primer designing for EST-SSRs**

All SSR-containing ESTs were individually inspected for suitability for primer design. SSR-containing ESTs that contain sufficient flanking sequences of good quality were selected for primer design. Primers were designed using Primer3 software (http://frodo.wi.mit.edu/primer3/) with an optimum annealing temperature 55–61°C and fragment size 100–300 bp, primer length 15–25bp with 20bp as optimum, and a minimum GC content with 50% being the optimum.

**Primer designing for Identified putative genes**

Primers were designed using Primer3 software (http://frodo.wi.mit.edu/primer3/) with an optimum annealing temperature 55–61°C and primer length 15–25bp with 20bp as optimum, and a minimum GC content with 50% being the optimum. The sequences of the primers sets were shown in the following table.
**List of putative genes primer sets**

<table>
<thead>
<tr>
<th>Putative genes</th>
<th>Forward Primer</th>
<th>Reverse primer</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAMs</td>
<td>ACCTCAGAGTCAGTCAACGA</td>
<td>CAGTCTCCCAGGTGAAATCG</td>
</tr>
<tr>
<td>MIPs</td>
<td>TGGATGGGGTGGAAACAAATG</td>
<td>AAAGTTTGAGGGGCAGACAG</td>
</tr>
</tbody>
</table>

**In silico gene cloning**

*In silico* cloning was carried out by readily available online software, snap gene [www.snapgene.com/](http://www.snapgene.com/). For identification of open reading frame (ORF), protein molecular weight (MW) and isoelectric point (pI) of the nucleotide and amino acid sequences online free tools ([www.bioinformatics.org/sms/prot](http://www.bioinformatics.org/sms/prot), [web.expasy.org/compute_pi](http://web.expasy.org/compute_pi)) were used.

**Development of Intron length polymorphism EST Primers**

We have used two model plants *Arabidopsis*, *Medicago*, two cereal crops plants rice, sorghum and one oil yielding crop soybean for the development of Intron-flanking EST-primers from the safflower. The genome sequences (with intron and UTRs) of *Arabidopsis*, rice, sorghum, soybean and *Medicago* reference databases were searched to find out the homolog sequences corresponding to the drought responsive ESTs sequences of *Carthamus tinctorius* using BLASTN program. We used a high e-value ($10^{-20}$) for the BLASTN to remove prologues. Intron–flanking EST-primers were designed based on the “exon/exon” junction site information, inferred from the pair-wise alignments between the ESTs of the *Carthamus* and their homolog exonic sequences from all the crops. As per this pair-wise alignment strategy the predicted “exon/exon” junctions from
safflower ESTs are essentially corresponded to “exon/intron” junctions (splice sites) in all genomic sequences.

The forward and reverse primers were designed using Primer3 (Rozen and Skaletsky, 2000) with default setting based on the EST sequences of safflower that flanks at least two exon/exon junctions. These primers were expected to span the predicted safflower intron regions in a PCR. The identified primer sets were calculated as total number of primers identified to total number of sequence used for the identification of conserved primers and expressed as percentage.

**In silico Polymerase Chain Reaction of ILPs markers**

We tested the designed primers by electronic PCR (e-PCR). Further we have validated our results by using online PCR (http://primer-2.0/) genomic sequences for *Arabidopsis* and rice. PCR product size >300bp, optimum annealing temperature 55–61°C with 60°C as the optimum, and a minimum GC content of 30%, with 50% being the optimum were used as default parameters for PCR. The ILP containing EST sequences of safflower were BLAST-searched against the genome sequences at Phytozome (http://www.phytozome.net/) to find each ILPs marker’s chromosome position on rice, sorghum, soybean and *Medicago*.

**Identification and characterization of putative heat shock proteins/chaperons from drought responsive EST library**

Six hundred and sixty seven drought responsive EST of safflower were used for similarity search using Blast2GO (Conesa et al., 2005) to identify their putative function. These sequences were run against the non-redundant (nr) protein database of the NCBI (http://www.ncbi.nlm.nih.gov/blast). The obtained hits were compiled (Conesa and Gotz,
2008). The sequence annotation function data was manually screened for the identification of putative heat shock proteins/chaperons. Seven sequences containing putative Hsps/chaperons were again validated with by tBLASTx, tBlast DB (nr) with threshold e-value of 1.e-15. We have predicted the localization and other molecular function of the identified HSPs from the data compiled from GO IDs. Sequences not showing any match were considered as unique to the safflower species.

**Motif analysis**

The sequences containing putative Hsps/chaperons were searched for triplet motifs by using SSR locator software at www.ssrlocator.com (Da Maia et al., 2008). SSR Locator is a tool for detecting and characterizing micro-and minisatellites in DNA sequences. In order to avoid bias identification of triplet motif, the parameter used as mono 5, di 3, tri 1, tetra, penta and hexa as 2. With these search criteria the maximum number of triplets can be identified by giving more importance to triplet motif.

**Prediction of amino acids from triplet motif**

The type of amino acids and their distributions in drought responsive ESTs containing putative Hsps/chaperons were predicted from trinucleotide repeats using SSR Locator software at www.ssrlocator.com. Translating the sequences containing putative Hsps/chaperons to their corresponding amino acids provides some clues about distribution of amino acids at the protein level.

**Phylogenetic analysis of the identified Hsps/chaperons**

For analyzing phylogenetic relationships between molecular sequences of safflower and other plant species, free online software is used at http://www.phylogeny.fr/. Blast 2.2.18 program BLASTN (http://blast.ncbi.nlm.nih.gov/)
with threshold e-value of 1e^{-5}, 1e^{-10} and 1e^{-20} were kept as default settings. A fast BLAST inbuilt search on Gigablast in http://www.phylogeny.fr/, quickly explore the sequence neighbours, facilitates the selection of homologous sequences. Based on a 'quick-and-dirty' phylogenetic representation using BLAST results and an estimator of the final multiple alignment length it analyses phylogenetic relationships between molecular sequences.

**Primer design and in silico cloning**

The safflower EST sequence containing putative Hsps/chaperons were in the range of 400-600bp. By using NCBI database, (www.ncbi.nlm.nih.gov) all these putative Hsps/Chaperons was individually BLAST (blast.ncbi.nlm.nih.gov) with tblastx. The full length CDS with criteria of showing more than 60% homology were searched because our sequences are 400-600bp. Primer pair was designed using PRIMER3 software at http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi. The sequence of the primer sets is as follows.

<table>
<thead>
<tr>
<th>Putative gene</th>
<th>Forward Primer</th>
<th>Reverse primer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hsps</td>
<td>CTTGCGTACATGGTGTGGAG</td>
<td>ACGAGGACGAAGGGGATAC</td>
</tr>
</tbody>
</table>

**Expression profiling of WRKY transcription factor genes by qRT-PCR**

Transcription factors (TFs) are key proteins regulating many signal transduction pathways. Within TFs, the WRKY proteins constitute a family involved in the plant abiotic and biotic stress responses. Safflower complete genome information is not available yet. In order to perform expression profiling of WRKY transcription factors
from safflower, we have used orthologs from sunflower (*Helianthus annuus*), belongs to the same family *Asteraceae*.

**Identification of WRKY genes**

Sunflower complete genome sequence is also limited, however the information regarding the huge quantity of expressed sequence tags (ESTs) is available in public databases. Giacomelli et al., (2010) carried out comprehensive, comparative EST analysis and identified 97 putative WRKY genes from sunflower. In addition, WRKY genes available from sunflower transcription factor database, selected orthologs sequences for qRT-PCR based expression profiling.

**Plant growth conditions**

Seeds of safflower (*Carthamus tinctorius* L.) cultivar A-1 were procured from Agricultural Department, Anantapuram. Seeds were surface sterilized with 0.1% (w/v) sodium hypochlorite solution for 5 minutes, thoroughly rinsed with distilled water, germinated in plastic pots containing 2kg of soil: sand (2:1) mixture and allowed to grow for nineteen days. The pots were maintained in the departmental botanical garden under natural photoperiod (10-12 hours; temperature 28 ± 4°C). Plants were maintained at 100% soil moisture level (SML).

**Stress imposition**

Nineteen-day-old safflower plants were subjected to 1) drought stress, by withholding water and samples were collected at the time of wilting 2) salt stress, by adding 250 mM NaCl solution to the pots and collected samples after 72 hours 3) cold stress, transferring plants to a growth chamber set at 10°C and samples were collected after 8 hours 4) dehydration stress was imposed by keeping uprooted plants in ambient
Materials and Methods

conditions for 8 hours. After completion of stress treatments, stressed and unstressed leaf samples were collected, flash frozen in liquid nitrogen and stored at -80°C prior to RNA isolation.

RNA isolation and DNase digestion

Total RNA from leaf samples was isolated using RNeasy Plant Mini Kit (Qiagen, Germany) as per manufacturer’s instructions. DNA was removed by incubating isolated RNA with DNase at 37°C for 30 minutes (TURBO DNase Kit; Ambion, USA). DNase was activated by treating the reaction mixture with inactivation buffer. Finally, the reaction mixture was centrifuged and separated RNA free from DNA.

Quantification of RNA

The purity and quantity of RNA was estimated spectrophotometrically. 2 µl of the RNA extract was taken in 1 ml of DEPC treated water for spectrophotometric quantification. The isolated RNA was measured at 260 and 280 nm in the UV-visible spectrophotometer (Shimadzu 1800, Japan) and quantified by using the formula given below.

\[
\text{Concentration} = \frac{\text{O.D at } 260 \text{ nm} \times 40}{\text{Volume of RNA taken (µl)}}
\]

Purity of RNA solution was checked by taking the \(A_{260}/A_{280}\) ratio.

RNA gel analysis

Formaldehyde agarose gel (1%) was prepared and equal volumes of RNA mixed with RNA tracking dye and were loaded on to the gel slots and electrophoresed at 300 V. After complete run the gel was visualized under UV-trans illuminator (UVI Tec, UK).
First strand cDNA synthesis

The purified RNA was taken (5 µl) in a fresh tube and 1 µl of oligo dT primer (5’GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAAGCTAGTCTCGAGTTTTTTTT TT-3’) and 6 µl of DEPC treated water were added to it. The reaction mixture was kept at 65°C for 5 minutes and immediately chilled on ice. Then 4µl of reverse transcriptase buffer (RT buffer), 2 µl of 10 mM dNTP mix, 2 µl of RT (M-MLV Reverse Transcriptase Fermentas, Germany) enzyme was added, mixed gently and incubated at 42°C for one hour. Then the reaction mixture kept at 70°C for 10 minutes to stop the reaction and immediately chilled on ice.

Preparation of Agarose gel electrophoresis

Nucleotide fragments were separated using horizontal gel electrophoresis. Depending on the requirement 1 to 1.5% agarose solutions were prepared in 1X TAE buffer (40 mM Tris acetate, 2 mM EDTA, pH 8.0) by heating in a microwave oven, cool the agarose solution up to 40-50°C and poured into the boat. 0.5 µg/ml of ethidium bromide was used for staining. After agarose polymerization nucleotide samples were mixed with 0.1 volume loading buffer (0.25% bromophenol blue, 0.25% xylene cyanol and 15% glycerol) and separated electrophoretically. The detection of the DNA fragments was carried out under UV trans-illuminator.

Confirmation of cDNA by agarose gel electrophoresis

Agarose gel (1%) was prepared and 2 µl of cDNA mixed with DNA tracking dye was loaded on to the gel slots and electrophoresed at 300 V. After complete run, the gel was visualized under UV-trans-illuminator. cDNA synthesis efficiency was normalized
by quantitative RT-PCR amplification of transcripts from reference gene coding actin. The sequence of actin primer set as follows:

Actin forward: 5’- AGGGCGGTCTTTTCCAAGTAT-3’
Actin reverse: 5’-ACATAATGGCGGGAACATT-3’

**Primer designing**

In order to monitor PCR amplification with maximum specificity and efficiency, qRT-PCR primers were designed using Primer3 software. Considering a stringent set of criteria such as: melting temperature (Tm) of 60±2, primer length of 20-24 nucleotides, guanine-cystosine (GC) content of 45-55% and PCR amplicon length of 100-150 base pairs (bp). The sequences of the primer sets were shown in the following table.

**List of gene specific primers for qRT-PCR**

<table>
<thead>
<tr>
<th>Gene</th>
<th>Forward</th>
<th>Reverse</th>
</tr>
</thead>
<tbody>
<tr>
<td>SF WRKY 1</td>
<td>5-CGGAGAAAGCACAGAGGT-3</td>
<td>5-TCCGAACCAGACTGCAAAAGA-3</td>
</tr>
<tr>
<td>SF WRKY 8</td>
<td>5-AGACCGCTATCGTAGGAGGAAG-3</td>
<td>5-TCTCGGATAAGGACTGTGGTTT-3</td>
</tr>
<tr>
<td>SF WRKY 12</td>
<td>5-CCCAAGGAGTATACAGATGCA-3</td>
<td>5-CCGATAGTCGATCCACTCTCTTT-3</td>
</tr>
<tr>
<td>SF WRKY 13</td>
<td>5-CACCAATCATGACGACTATT-3</td>
<td>5-TCTGCCACTCTCTCTCTCTACTACT-3</td>
</tr>
<tr>
<td>SF WRKY 21</td>
<td>5-TGGGATCCGGAATGC-3</td>
<td>5-TGGCAATTTCGGGTGGTAT-3</td>
</tr>
<tr>
<td>SF WRKY 25</td>
<td>5-CGGAGAGGGATGGATTCAATT-3</td>
<td>5-CGATTTCGCTTCTTTCAC-3</td>
</tr>
<tr>
<td>SF WRKY 33</td>
<td>5-AGGAGGAGGCGGCTCAG-3</td>
<td>5-TAACTGGGAAGCGGCCACTT-3</td>
</tr>
<tr>
<td>SF WRKY 38</td>
<td>5-CCGATCGGAGAGAAGGGA-3</td>
<td>5-CCGATCGGAGAGAAGGGA-3</td>
</tr>
<tr>
<td>SF WRKY 45</td>
<td>5-GGCAATGGGGAGCGAAG-3</td>
<td>5-TCAACGGCATGGATGACAC-3</td>
</tr>
<tr>
<td>SF WRKY 54</td>
<td>5-TGGTGCAAGAAGCAAGAACA-3</td>
<td>5-CCCTAAACCACATGCCTGCTCTA-3</td>
</tr>
</tbody>
</table>
Materials and Methods

Recipe for qRT-PCR reaction (µl)

Applied Biosystems 2X SYBR GREEN master mix …… 2.5

- Forward primer (1pm/µl) 1
- Reverse primer (1pm/µl) 1
- Template 0.5
- Total volume …… ………………5

RT-PCR conditions followed are given below

1. 58°C …………………… 2.0 min
2. 94°C ……………………10.0 min
3. 94°C …………………… 0:15 min
4. 55°C …………………… 1.0 min

step 3-4 ………………… 40 cycles

Quantitative RT-PCR conditions and analysis

Quantitative RT-PCR (qRT-PCR) reactions were conducted in StepOne Real Time PCR system (Applied Biosystems, USA) essentially described by Caldana et al., (2007). Reactions with a final volume of 5µl contained 0.5µl of template (cDNA), 200 nmole of each gene specific primer (2µl of mixed 0.5mM forward and reverse primers), and 2.5µl of 2x SYBR Green reagents, using an electronic pipette as a master mix of cDNA and 2x SYBR Green reagents, using an electronic pipette (Eppendorf, Germany). Reaction plates were sealed with a transparent adhesive cover before proceeding to quantitative reactions. The following standard thermal profile was used for all qRT-PCR reactions: 50°C for 2 minutes, 95°C for 10 minutes; 40 cycles of 95°C for 15 seconds, and 55°C for 1 minute. After 40 cycles, the specificity of the amplifications was checked.
by heating from 60°C to 95°C with a ramp speed of 1.9°C minute⁻¹, using SDS 2.2.1 Software (Applied Biosystems, USA). Amplification curves were analysed with a normalized reporter (Rn: the ratio of the fluorescence emission intensity of the SYBR Green to the fluoresce signal of the passive reference dye) threshold of 0.2 to obtain the Cₜ values (Threshold cycle). Data were normalized to reference gene Actin; ∆Cₜ(gene) - Cₜ(Actin). Its expression was measured with three replicates in each PCR run, and the average Cₜ was used for relative expression analyses. The fold change value was calculated using the expression 2⁻∆∆Cₜ, where ∆∆Cₜ represents, Cₜ Treatment - Cₜ Control. The results obtained were transformed to log2 scale.

**In gel analysis**

The RT-PCR amplified product was analysed on 2% agarose gel, and images were captured using Vilbour-Lourmat Gel documentation system (France).