Chapter III

Comparative Analysis of Transcriptome of Water Stress Tolerant and Susceptible Cultivars Using Suppression Subtractive Hybridization
3.1 Introduction

Tea plant is one of the most important commercial cash crops of China, India, Sri Lanka and other countries. Being a perennial rain fed tree species, its successful cultivation is challenged by a number of biotic and abiotic stresses throughout the year. Among abiotic stresses, drought is one of the major constrains in tea growing areas that reduce yield every year (Singh and Handique, 1993). Since, irrigation is hardly used in tea cultivation; the rain water is the only source that determines the yield potential of tea to a great extent. Therefore, any variation in rainfall and temperature is expected to affect yield and quality of tea. In view of continuous depletion of water resources and global climate change, it has become necessary to study the drought responsive mechanisms in tea at transcriptome level for its improvement of drought tolerance to maintain good yield and quality in water stress environment.

There are several studies in plants generating large gene expression data in response to drought and other abiotic stresses and most of these data has been generated using the model plant Arabidopsis (Seki et al., 2002; Thomashow, 1999; Cheong et al., 2002). Several studies are reported in tea involving effect of water stress on growth and yield (Ng’etich, 1997), root depth, weight and vertical distribution (Nagarajah and Ratnasuriya, 1981). Correlation of drought tolerance of tea with age (Nixon et al., 2001), polyphenol content (Cheruiyot et al., 2007), accumulation of antioxidants (Upadhyay et al., 2008) and clonal variation (Chakraborty et al., 2002) has also been reported. cDNA libraries in tea has been constructed from various tissues to identify transcripts mainly related to secondary metabolism (Matsumoto et al., 1994; Park et al., 2004) as well as other functional categories (Chen et al., 2005) and dormancy (Krishnaraj et al., 2011). Recently, deep sequencing (34.5 million ESTs) of tea transcriptome (Camellia sinensis) has been identified several candidate genes for major metabolic pathways (Shi et al., 2011). There are only a few reports on studies related to drought tolerance in tea at molecular level. For example, Sharma and Kumar (2005) identified three drought responsive ESTs using differential display method which is far behind to well represent the drought-
responsive transcriptome of tea plant. Similar type of study was made by Qi et al. (2010) in which they identified three PEG6000 (Polyethylene glycol) mediated drought responsive fragments. Therefore, a comprehensive investigation is required to study the water stress responses of tea plant in order to understand the molecular basis of drought responses which would contribute towards improvement of tea plant for water stress tolerance. One of the ways to study the drought responsive mechanisms in tea is to identify the novel drought responsive genes and their products involved in various biochemical, physiological and signal transduction pathways by comparing the transcriptome profile of a drought tolerant tea cultivar with a susceptible one. There are several examples of such studies in other plants to compare expression profile of stressed versus non-stress plants resulting into the identification of several abiotic stress related genes (Reddy et al., 2002; Kushwaha et al., 2009).

In the present study, we have compared the drought responsive transcriptome of a tolerant tea cultivar (TV23) with a susceptible one (S.3A/3) by SSH (Suppression Subtractive Hybridization) approach (Diatchenko et al., 1996) to identify the differentially expressed drought-responsive transcripts in leaves at different degrees of drought stress. These two cultivars were found to have contrasting drought tolerance behaviour under field as well as in controlled environmental conditions (Chapter II). Therefore, a comparative analysis of drought responsive transcriptome of these two contrasting cultivars will unravel many genes that may play role in drought tolerance mechanisms in tea.

3.2 Materials and methods

3.2.1 Plant Materials

Young tea leaves (a bud along with 1st and 2nd leaf) of TV23 and S.3A/3 collected at two stages of drought (before wilting stage, BWS; and wilting stage, WS) during the induced water stress experiment (Chapter-II) were used in the present study. These samples were designated as TV23BWS, TV23WS, S.3A/3BWS and S.3A/3WS.
3.2.2 Total RNA Isolation and Quality

Total RNA was isolated from all the four drought induced samples using RNAqueous kit (Ambion, Cat. No. AM1912). The good quality of RNA was ensured by running a denaturing agarose gel. The quality was also checked using a Spectrophotometer (BioPhotometer, Eppendorf, Germany).

3.2.2.1 mRNA Isolation

mRNA was isolated from the high quality total RNA using PolyATtract® mRNA isolation systems III (Promega Corporation, USA, Cat. No. Z5300) following the protocol mentioned in the user manual. The purity and concentration of isolated mRNA was checked using a spectrophotometer.

3.2.3 Subtractive cDNA Library Construction

Four subtractive libraries (Table 3.1) were constructed using PCR-Select cDNA Subtraction Kit (Clontech, USA, Cat. No. 637401) following the protocol described in the user manual. Two subtractive libraries were constructed taking TV23BWS as tester and S.3A/3BWS as driver and vice versa. The forward library is designated as BWE3E4 while reverse as BWE4E3. Likewise, another two libraries (forward and reverse) were constructed from the WS samples (TV23WS and S.3A/3WS) and designated as WE3E4 and WE4E3 (Table 3.1)

<table>
<thead>
<tr>
<th>Tester sample</th>
<th>Driver sample</th>
<th>Library name</th>
<th>Objective to identify differentially expressed genes in</th>
</tr>
</thead>
<tbody>
<tr>
<td>TV23BWS</td>
<td>S.3A/3BWS</td>
<td>BWE3E4</td>
<td>TV23 at before wilting stage</td>
</tr>
<tr>
<td>S.3A/3BWS</td>
<td>TV23BWS</td>
<td>BWE4E3</td>
<td>S.3A/3 at before wilting stage</td>
</tr>
<tr>
<td>TV23WS</td>
<td>S.3A/3WS</td>
<td>WE3E4</td>
<td>TV23 at wilting stage</td>
</tr>
<tr>
<td>S.3A/3WS</td>
<td>TV23WS</td>
<td>WE4E3</td>
<td>S.3A/3 at wilting stage</td>
</tr>
</tbody>
</table>

Table. 3.1  Samples of different stages used in SSH library construction
3.2.3.1 First-strand cDNA Synthesis

The first strand synthesis was carried out from 2 µg (3µl) of Poly A⁺ RNA. The Poly A⁺ RNA was added to a sterile 0.5 ml microcentrifuge tube containing 1 µl of cDNA synthesis primer (10 µM) and 1µl of sterile deionised water. After mixing and brief spin the tube was incubated for 2 min in a thermal cycler (Mastercycler gradient, Eppendorf, Germany) at 70 ºC. This was followed by addition of 2 µl 5X first strand buffer, 1 µl dNTP Mix (10 mM each), 1µl sterile water and 1 µl AMV Reverse Transcriptase (20 units/µl). After mixing and brief spin, tubes were incubated at 42 ºC for 1.5 hr in an air incubator. The reaction was terminated by placing the tube in ice.

3.2.3.2 Second-strand cDNA Synthesis

The second strand cDNA synthesis reaction was performed by adding 48.4 µl sterile water, 16 µl 5X second-strand buffer, 1.6 µl dNTP mix (10 mM) and 4 µl of 20X second-strand enzyme cocktail to the above first strand reaction mixture and incubated at 16 ºC for two hours in a water bath (Pharmacia, USA). The reaction was terminated by adding 4 µl of 20X EDTA/Glycogen mix. The reaction mixture was purified using Phenol: chloroform: isoamylalcohol followed by ammonium acetate precipitation. The pellet was washed, dried and dissolved in 50 µl of sterile water.

3.2.3.3 Preparation of Driver cDNA

The purified second-strand cDNA was digested in a 50 µl reaction mixture containing 15 units of Rsal enzyme (New England BioLabs, UK, Cat. No. R0167S) for 1.5 hrs to produce blunt-ended fragments. The cDNA was phenol-extracted, ethanol-precipitated and then resuspended in 5.5 µl deionised water.

3.2.3.4 Preparation of Tester cDNA

Rsal-digested ds tester cDNA was prepared as described above for the driver. Digested tester cDNA (1 µl) was diluted in 5 µl of deionised water. The diluted tester cDNA (2 µl) was then ligated overnight to 2 µl of adapter 1
(10 µM) and 2 µl adapter 2R (10 µM) in separate ligation reactions in a total volume of 10 µl at 16 ºC using 0.5 units of T4 DNA ligase (New England BioLabs, UK. Cat. No. M0202S) in the buffer supplied by the manufacturer. After ligation, 1 µl of 0.2 M EDTA was added and the sample was heated at 70 ºC for 5 min to inactivate the ligase and stored at -20 ºC for downstream applications.

3.2.3.5 Subtractive Hybridization

To obtain differentially expressed cDNAs, two rounds of hybridizations were performed. The purpose of the first round of hybridization was to equalize and to enrich the differentially expressed sequences. The objective of the second round was to produce double-stranded tester molecules with different adaptors on each end. 1.5 µl of Rsal-digested driver ds cDNA was added to each of two tubes containing 1.5 µl of adapter1- and adapter 2R-ligated tester cDNA. The samples were mixed, ethanol precipitated and resuspended in 1 µl of hybridization buffer (50 mM HEPES, pH 8.3; 500 mM NaCl; 0.02 mM EDTA, pH 8.0; 10 % [w/v] PEG 8000). The solution was overlaid with mineral oil, the DNA was denatured at 98 ºC for 1.5 min and then allowed to anneal for 8 hrs at 68 ºC in a Thermal Cycler (Applied Biosystems, Model 2720). After this first hybridization, the two samples were combined and a fresh portion of heated-denatured driver was added to 1 µl of hybridization buffer according to the procedure specified in the user manual. The sample was left to hybridize for an additional 11 hrs at 68 ºC in the same Thermal Cycler. The final hybridization was then diluted in 200 µl dilution buffer (20 mM HEPES, pH 6.6; 20 mM NaCl; 0.2 mM EDTA, pH 8.0), heated at 72 ºC for 7 min and stored at -20 ºC.

3.2.3.6 PCR Amplification

Two rounds of PCR amplifications were performed for each subtraction. In the first amplification, PCR was suppressed, whereby only differentially expressed sequences can be amplified exponentially. In the second procedure, the background was reduced to enrich the differentially expressed sequences. The primary PCR was conducted in 25 µl containing...
1 µl of diluted subtracted cDNA, 1 µl of PCR primer 1 (10 µM) and 23 µl of PCR master mixture prepared using the 50X Advantage 2 PCR Kit (Clontech, USA, Cat. No.639206). PCR was performed with the following parameters: 75 ºC for 5 min, 27 cycles at (94 ºC for 10 seconds, 68 ºC for 30 sec, 72 ºC for 1.5 min,) and a final extension at 72 ºC for 5 minutes. The amplified products were diluted by a factor of 10 in deionised water. The primary PCR-diluted products (1 µl) were then used as templates in secondary PCR for 10 cycles under the same conditions as for the primary PCR, with the exception that PCR primer 1 was replaced by nested PCR primers 1 and 2R. The PCR products were analyzed on a 2.0% Agarose/EtBr gel. All of the primers (PCR primer 1 and nested PCR primers 1-2R) for the PCR were in the kit.

3.2.3.7 PCR Analysis of Subtraction Efficiency

The efficiency of subtraction was tested by doing a PCR taking equal quantity of subtracted and unsubtracted cDNAs of tester as template. The gene specific primer (5′TCAAATTCCGAAGGTCTAAAG3′ and 5′CGGAAACGGCAAAAGTG3′) of housekeeping gene 26S was used to compare its abundance in the subtracted and unsubtracted cDNAs of tester. The equal abundance of transcripts of 26S in control and experimental samples were ensured.

3.2.4 Cloning of Subtracted cDNA

The subtracted cDNAs, obtained from the second PCR amplification, were ligated to pGEM-T Easy vector (Promega Corporation, USA, Cat. No. A1360) and electroporated into ElectroMax DH10β E. coli cells (Invitrogen, USA, Cat. No.18290-015). The transformed cells were plated onto LB agar culture plates containing Ampicillin (100 µg/ml), IPTG (125 mM), X-Gal (40 µg/ml) and were incubated at 37 ºC overnight. The transformants were selected based on α-complementation. White colonies were randomly picked up and inoculated in 4 ml LB medium containing Ampicillin (100 µg/ml) at 220 rpm at 37 ºC overnight. Next day the cells were harvested and plasmid was isolated using HiPura plasmid miniprep purification kit (Himedia, Cat.
The inserts in the extracted plasmids were verified by digesting with EcoRI (New England BioLabs, UK, Cat. No. R0101S)

### 3.2.5 DNA Sequencing and Homology Search

The inserts in the extracted plasmids were sequenced in GA3130xl Genetic Analyzer (Applied Biosystems) using BigDye Terminator v3.1 Cycle sequencing Kit (Applied Biosystems, Cat. No. 4336917) following the instructions given in the manual. The adapter and vector sequences were removed from the selected good quality (quality value ≥ 20 and length ≥100 nt) sequences using Sequence Scanner software v1.0 (Applied Biosystems) and considered for further analysis. A similarity search of all the ESTs, before contig assembly, was performed using the BlastX programme at the National Centre for Biotechnology Information (http://blast.ncbi.nlm.nih.gov/Blast.cgi/) at default parameters against non-redundant (nr) protein database.

### 3.2.6 Functional Annotation of Differentially Expressed ESTs

ESTs of all the four libraries were used for contig assembly by using the web interface (http://mobyle.pasteur.fr/cgi-bin/portal.py#forms::cap3) of the CAP3 software (Huang and Madan, 1999) at default parameters. The unigenes of all the four libraries were individually annotated according to the three main vocabularies of the gene ontology (GO): cellular component, biological process and molecular function. Functional annotations of ESTs were performed using the Blast2GO software v.2.4.8 (Conesa et al., 2005) at default parameters. InterPro Scan (Berardini et al., 2004) was performed to find functional motifs and related GO terms by using the specific tool implemented in the Blast2GO software. Finally, the Augment Annotation by ANNEX' (Zdobnov and Apweiler, 2001) function was used to refine annotations (http://www.goat.no). Enzyme mapping of annotated sequences were done by direct GO to Enzyme annotation and used to query the Kyoto Encyclopaedia of Genes and Genomes (KEGG, http://www.genome.jp/kegg/) (Kanehisa and Goto, 2000; Kanehisa et al.,
2006; Kanehisa et al., 2008) in order to define the main metabolic pathways involved.

Enrichment analysis was performed using Gossip (Bluthgen et al., 2005) package integrated into Blast2GO to identify the functional categories that are over- or under-represented in the set of differentially expressed genes between libraries of different stages at default p value (≤ 0.05). This package employs the Fisher’s exact test and corrects for multiple testing. For this analysis, the involved sequences with their annotations were loaded in the application. Blast2GO returns the GO terms under or over represented at the specified p value.

3.3 Results

3.3.1 RNA Isolation

Good quality of extracted RNA was revealed by two distinct intact bands of 28S and 18S in a denaturing agarose gel (Figure 3.1). The concentration of RNA was found to be 670 ng/µl.

Fig 3.1 Total RNA resolved in 1.2% denaturing Agarose gel
Chapter III

3.3.2 Efficiency of Subtraction

The efficiency of subtraction was checked by performing a PCR reaction. The abundance of 26S rRNA transcripts decreased after subtraction (Figure 3.2). After 23 cycles there was amplification for unsubtracted cDNA while there was no amplification for subtracted cDNA sample. The intensity of band after same number cycles for unsubtracted cDNA is much intense as compared to subtracted sample after 38 cycles. This clearly indicates that abundance of transcripts of 26S rRNA decreased after subtraction and hence proved the efficiency of subtraction.

![Image of gel electrophoresis showing unsubtracted and subtracted samples](image.png)

Fig 3.2 Abundance of 26S rRNA before subtraction and after subtraction

3.3.3 BlastX Analysis

Since, the physiological parameters studied during the induced water stress experiment (Chapter II) as well as under field drought conditions have indicated that TV23 is more drought tolerant than S.3A/3, we were more interested to identify the transcripts that were unique and more abundant in the library BWE3E4. Out of 660 good quality ESTs of BWE3E4, 510 showed hits with a number of known stress responsive, unknown, hypothetical or putative proteins while remaining 150 did not show any hits. From the BlastX results of each EST, we considered that protein hit which is either reported
to be stress responsive or induced during dehydration stress in available literature, although it may not have maximum e-value among the blast hits. The list of selected BlastX results of libraries (BWE3E4 and BWE4E3) can be found in Appendix 3.1 and 3.2. Each of these stress responsive genes were represented by either one, two, three or more ESTs in the libraries. We were more interested in transcripts expressed in TV23 at BWS (i.e., library BWE3E4). Proteins which are known and represented by more than 3 ESTs (49 proteins, Table 3.2) in the library BWE3E4 were considered as highly expressed compared to other known proteins in that library. Presence of these proteins was also compared (Table 3.2) with the BlastX result of the reverse library BWE4E3. Among the most notable are the genes encoding ribosomal RNA protein (17 ESTs), metallothionein (10 ESTs), unnamed protein product (27 ESTs), *ethylene-induced esterase* (7 ESTs), *cinnamoyl-CoA reductase* (7 ESTs) and dehydrin (7 ESTs). However, majority of the proteins were represented by single ESTs in the library (not shown in Table 3.2). Of these 49 genes, 36 are unique to BWE3E4.

<table>
<thead>
<tr>
<th>Sl. No.</th>
<th>Protein name</th>
<th>LIBRARIES</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>BWE3E4</td>
</tr>
<tr>
<td>1</td>
<td>Abscisic stress ripening protein (ASR1)</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>ADP ribosylation factor _1</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td><em>Ascorbate peroxidase</em></td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>APE1</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>Amine oxidase</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>ATP synthase</td>
<td>4</td>
</tr>
<tr>
<td>7</td>
<td>Clp G3PD</td>
<td>3</td>
</tr>
<tr>
<td>8</td>
<td>β-glucanase</td>
<td>3</td>
</tr>
<tr>
<td>9</td>
<td>Calmodulin-related protein</td>
<td>4</td>
</tr>
<tr>
<td>10</td>
<td><em>Glutathione S-transferase</em></td>
<td>3</td>
</tr>
<tr>
<td>11</td>
<td>Chlorophyll a/b binding protein</td>
<td>7</td>
</tr>
<tr>
<td>12</td>
<td>Hexose transporter</td>
<td>7</td>
</tr>
<tr>
<td>13</td>
<td>Photosystem II protein</td>
<td>4</td>
</tr>
<tr>
<td>14</td>
<td>Loss of ET and JA biosynthesis</td>
<td>3</td>
</tr>
<tr>
<td>15</td>
<td><em>Cytochrom C oxidase</em></td>
<td>3</td>
</tr>
<tr>
<td>16</td>
<td>ACC oxidase</td>
<td>4</td>
</tr>
</tbody>
</table>
Chapter III

3.3.4 Contig Assembly

A total of 2365 differentially expressed drought induced ESTs (Gene Bank Acc. No. HS393200 to HS395561, HS395745, HS395830, HS396211 and HS396243) of four libraries were reduced to 1012 unigenes (singletons and contigs) after contig assembly. The details of contig assembly of all the libraries are given in the Table 3.3.
### Table 3.3 Statistics of contig assembly of libraries

<table>
<thead>
<tr>
<th>Characters</th>
<th>LIBRARIES</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>BWE3E4</td>
</tr>
<tr>
<td>Total number of reads before contig assembly (≥100nt)</td>
<td>660</td>
</tr>
<tr>
<td>Total length of reads before contig assembly (nt)</td>
<td>252326</td>
</tr>
<tr>
<td>Total number of reads after contig assembly (≥100nt)</td>
<td>505</td>
</tr>
<tr>
<td>Average length of reads after contig assembly (nt)</td>
<td>307</td>
</tr>
<tr>
<td>Total length of reads after contig assembly (nt)</td>
<td>202736</td>
</tr>
<tr>
<td>Total number of singletons</td>
<td>443</td>
</tr>
<tr>
<td>Total length of singletons (nt)</td>
<td>174732</td>
</tr>
<tr>
<td>Largest singleton (nt)</td>
<td>766</td>
</tr>
<tr>
<td>Shortest singleton (nt)</td>
<td>100</td>
</tr>
<tr>
<td>Average length of singletons (nt)</td>
<td>394</td>
</tr>
<tr>
<td>Total number of contigs</td>
<td>62</td>
</tr>
<tr>
<td>Total length of contigs (nt)</td>
<td>28004</td>
</tr>
<tr>
<td>Largest contig (nt)</td>
<td>968</td>
</tr>
<tr>
<td>Shortest contig (nt)</td>
<td>141</td>
</tr>
<tr>
<td>Average length of contig (nt)</td>
<td>451</td>
</tr>
<tr>
<td>Highest number of ESTs in a contig</td>
<td>26</td>
</tr>
<tr>
<td>Total number of unigenes considered for Blast2GO analysis</td>
<td>505</td>
</tr>
</tbody>
</table>
3.3.5 Functional Annotation of ESTs

The individual library sequences after contig assembly were blasted in Blast2GO suit using BlastX programme at default parameters against NCBI’s nr protein database. The blast results of all the libraries and their dbEST accession numbers can be found in Appendix 3.1 to 3.4. The top blast hit species in all the library was found to be *Vitis vinifera*, followed by *Populus trichocarpa* in BWE4E3, WE3E4, WE4E3 and *Ricinus communis* in BWE3E4 (Figure 3.3a – 3.3d). The first thirteen top-hit species of all the libraries are shown in Figure 3.3a – 3.3d. Mapping of all the libraries have shown that majority of the GO terms were retrieved from the UniProtKB/TrEMBL (82% - 88%) database (Table 3.4). Merging of InterPro Scan results and ANNEX augmentation significantly reduced the total number of successfully annotated GOs (Table 3.4). Different steps of annotation process of four libraries can be found in Appendix 3.5 to 3.8.

<table>
<thead>
<tr>
<th>Databases</th>
<th>LIBRARIES</th>
<th>BWE3E4</th>
<th>BWE4E3</th>
<th>WE3E4</th>
<th>WE4E3</th>
</tr>
</thead>
<tbody>
<tr>
<td>UniProtKB/TrEMBL</td>
<td>8589</td>
<td>5670</td>
<td>1733</td>
<td>1399</td>
<td></td>
</tr>
<tr>
<td>UniProtKB/Swiss-Prot</td>
<td>1241</td>
<td>771</td>
<td>280</td>
<td>206</td>
<td></td>
</tr>
<tr>
<td>TAIR</td>
<td>926</td>
<td>319</td>
<td>248</td>
<td>145</td>
<td></td>
</tr>
<tr>
<td>GR_protein</td>
<td>190</td>
<td>95</td>
<td>39</td>
<td>29</td>
<td></td>
</tr>
<tr>
<td>UniProt</td>
<td>79</td>
<td>36</td>
<td>14</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>SGN</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>MGI</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>WB</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Tair</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Total GO Terms retrieved</td>
<td>11038</td>
<td>6891</td>
<td>2317</td>
<td>1789</td>
<td></td>
</tr>
<tr>
<td>Total GO terms annotated</td>
<td>1189</td>
<td>570</td>
<td>221</td>
<td>158</td>
<td></td>
</tr>
<tr>
<td>No of unigenes annotated</td>
<td>221</td>
<td>111</td>
<td>47</td>
<td>36</td>
<td></td>
</tr>
</tbody>
</table>
### Chapter III

**Figure 3.3a** Top-hit species distribution of BWE3E4

- *Selaginella moellendorffii*
- *Oryza sativa*
- *Actinidia deliciosa*
- *unknown*
- *Jatropha curcas*
- *Arabidopsis thaliana*
- *Arabidopsis lyrata*
- *Nicotiana tabacum*
- *Camellia sinensis*
- *Glycine max*
- *Populus trichocarpa*
- *Ricinus communis*
- *Vitis vinifera*

![Graph](image1)

**Figure 3.3b** Top-hit species distribution of BWE4E3

- *Magnetospirillum*
- *Panax ginseng*
- *Solanum tuberosum*
- *Oryza sativa*
- *Sorghum bicolor*
- *Agave tequilana*
- *Nicotiana tabacum*
- *Medicago truncatula*
- *Glycine max*
- *Camellia sinensis*
- *Ricinus communis*
- *Populus trichocarpa*
- *Vitis vinifera*

![Graph](image2)

**Figure 3.3c** Top-hit species distribution of WE3E4

- *Alopecurus aequalis*
- *Triticum aestivum*
- *Camellia sinensis*
- *Selaginella moellendorffii*
- *Vigna unguiculata*
- *Nematostella vectensis*
- *Litchi chinensis*
- *Camellia oleifera*
- *Solanum lycopersicum*
- *Glycine max*
- *Ricinus communis*
- *Populus trichocarpa*
- *Vitis vinifera*

![Graph](image3)

**Figure 3.3d** Top-hit species distribution of WE4E3

- *Pinus koraiensis*
- *Vigna unguiculata*
- *Nematostella vectensis*
- *Psilotus crispum*
- *Arabidopsis lyrata*
- *Daucus carota*
- *Solanum tuberosum*
- *Solanum lycopersicum*
- *Camellia sinensis*
- *Glycine max*
- *Ricinus communis*
- *Populus trichocarpa*
- *Vitis vinifera*

![Graph](image4)
Chapter III

Among the four libraries the highest diversity of annotated unigenes was observed in BWE3E4 (221 unigenes) followed by BWE4E3 (111 unigenes), WE3E4 (47 unigenes) and WE4E3 (36 unigenes). These annotated sequences were further categorised according to the three main vocabularies of gene ontology as shown in Figure 3.4a - 3.4l.

Figure 3.4a Molecular function (BWE3E4) (GO level 3)

Figure 3.4b Molecular function (BWE4E3) (GO level 3)
Chapter III

![Pie chart showing molecular function categories and their percentages.]

**Figure 3.4c Molecular function (WE3E4) (GO level 3)**

- Hydrolase activity: 13%
- Transferase activity: 11%
- Structural constituent of ribosome: 9%
- Protein binding: 11%
- Oxidoreductase activity: 14%
- Nucleic acid binding: 18%
- Structural constituent of ribosome: 7%
- Protein binding: 9%
- Oxidoreductase activity: 11%
- Lyase activity: 7%
- Nucleotide binding: 7%
- Ion binding: 15%
- Transmembrane transporter activity: 5%
- Substrate-specific transporter activity: 5%
- Cofactor binding: 6%

**Figure 3.4d Molecular function (WE4E3) (GO level 3)**

- Hydrolase activity: 13%
- Transferase activity: 7%
- Nucleic acid binding: 18%
- Protein binding: 9%
- Oxidoreductase activity: 11%
- Lyase activity: 7%
- Nucleotide binding: 7%
- Ion binding: 15%
- Transmembrane transporter activity: 5%
- Substrate-specific transporter activity: 5%
- Cofactor binding: 6%
Figure 3.4e Biological Process (BWE3E4) (GO level 3)

Figure 3.4f Biological Process (BWE4E3) (GO level 3)
Chapter III

Figure 3.4g Biological Process (WE3E4) (GO level 3)

Figure 3.4h Biological Process (WE4E3) (GO level 3)
Figure 3.4i Cellular Component (BWE3E4) (GO level 6)

Figure 3.4j Cellular Component (BWE4E3) (GO level 6)
Although the main focus of our study was the identification and functional annotation of the differentially expressed transcripts in TV23 at BWS (i.e., library BWE3E4), we also made an attempt to identify and annotate the transcripts in other libraries. The functional categorizations of ESTs of all the libraries were considered at GO level 3 for molecular function and biological process. For cellular component the GO level considered for categorization was 6 as because there were only few GO subcategories at GO level 3. It is evident that the diversity of drought responsive transcripts involved in molecular function at specified GO level is more in case of TV23 (16 GOs in BWE3E4 and 11 GOs in WE3E4) as compared to S.3A/3 (10 GOs in BE4E3 and 10 GOs in WE4E3) at both the stages of drought induction (BWS and WS, Figure 3.4a, 3.4c and 3.4b, 3.4d). Similar trend was also observed for biological process (26 GOs in BE3E4, Figure 3.4e) and cellular component (17 GOs each in BE3E4 and WE3E4) at GO level 6. The number of GO subcategories also varies between libraries of TV23 at BWS (BWE3E4) and WS (WE3E4). The diversity of GO classes was found to be more in BWS of TV23 compared to S.3A/3, represented by 16 and 26 GOs for molecular function and biological process respectively, while there was no significant differences for cellular component. Similar type of trend was also observed between BWS and WS libraries of S.3A/3 (i.e., BWE3E4 and WE4E3) for biological process and cellular component (represented by 17 and 13 GOs respectively). But, there were no significant differences for molecular function category.

Comparison of GO term distribution between BWE3E4 and BWE4E3 libraries for molecular function and biological process at GO level 3 revealed several library specific GOs. We were interested with those GOs that were unique to BWE3E4. We have identified a total of 12 GO terms unique to BWE3E4 (Appendix 3.9) which includes 80 different proteins (Appendix 3.10). Some of the most abundantly occurring proteins (Appendix 3.10) within this unique GOs are dehydrin (5 times), multidrug resistance protein abc transporter family (4 times), coronatine insensitive 1 (4 times) , ATP synthase beta subunit (3 times), ATP synthase cf1 alpha subunit protein (3 times), ubiquinol-cytochrome-C reductase (3 times), glycosyltransferase
family 35 protein (3 times), stress enhance protein 1 (3 times), NAC domain protein (3 times), ERD15 protein (3 times) and eukaryotic translation initiation factor 5A4 (3 times). The comparison of GO term distribution for cellular component at GO level 3 did not show any significant differences.

3.3.6 Comparative Analysis of Drought Responsive Pathways

We were interested to identify the different drought responsive metabolic pathways in TV23 that were absent in S.3A/3 at BWS. Querying the KEGG retrieved a total of 85 pathway maps involving 130 enzymes (Appendix 3.11) for BWE3E4 and 36 pathway maps and 64 enzymes for BWE4E3 (Appendix 3.12). The most dominant pathways in BWE3E4 were found to be related to the amino acid metabolism (13 pathways), followed by lipid metabolism (11 pathways) and carbohydrate metabolism (10 pathways). The dominant pathways for BWE4E3 were found to be carbohydrate metabolism (4 pathways) and energy metabolism (4 pathways).

The comparative analysis of different metabolic pathways between BWE3E4 and BWE4E3 revealed several pathways that are unique to the former. Those unique pathways that are represented by at least three or more enzymes in the library BWE3E4 are listed in Table 3.5. Besides, there are pathways represented by two or one enzymes in the libraries as shown in Appendix 3.11 and 3.12.

The enrichment analysis of the libraries BWE3E4 (test) and WE3E4 (reference) resulted into several classes of GO terms that are over-represented either in test group or reference group (Figure 3.5) while rest of the GO terms are unique. There were four GO terms being up-regulated in test group as compared to reference group. These are nucleic acid binding (GO:0005488) represented by 130 sequences, cellular nitrogen compound metabolic process (GO:0034641) represented by 51 sequences, nitrogen compound metabolic process (GO:0006807) represented by 52 sequences, nucleobase, nucleoside, nucleotide and nucleic acid metabolic process (GO:0006139) represented by 24 sequences.
Table 3.5 Pathways (those represented by three or more enzymes) unique to BWE3E4

<table>
<thead>
<tr>
<th>Enzymes</th>
<th>E. C. code</th>
<th>Metabolic pathways*</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Alcohol dehydrogenase</em></td>
<td>ec:1.1.1.1</td>
<td>F I B T1 D A T2 T3</td>
</tr>
<tr>
<td><em>Enoyl-CoA hydratase</em></td>
<td>ec:4.2.1.17</td>
<td></td>
</tr>
<tr>
<td><em>3-hydroxybutyryl-CoA epimerase</em></td>
<td>ec:5.1.2.3</td>
<td></td>
</tr>
<tr>
<td><em>3-hydroxyacyl-CoA dehydrogenase</em></td>
<td>ec:1.1.1.35</td>
<td></td>
</tr>
<tr>
<td><em>Dodecenoyl-CoA isomerase</em></td>
<td>ec:5.3.3.8</td>
<td></td>
</tr>
<tr>
<td><em>Monoamine oxidase</em></td>
<td>ec:1.4.3.4</td>
<td></td>
</tr>
<tr>
<td><em>Salutaridinol 7-O-acetyltransferase</em></td>
<td>ec:2.3.1.150</td>
<td></td>
</tr>
<tr>
<td><em>Aspartate transaminase</em></td>
<td>ec:2.6.1.1</td>
<td></td>
</tr>
<tr>
<td><em>Flavin-containing monooxygenase</em></td>
<td>ec:1.14.13.8</td>
<td></td>
</tr>
<tr>
<td><em>Pyrroline-5-carboxylate reductase</em></td>
<td>ec:1.5.1.2</td>
<td></td>
</tr>
<tr>
<td><em>Spermidine synthase</em></td>
<td>ec:2.5.1.16</td>
<td></td>
</tr>
<tr>
<td><em>Dimethylallyltransferase</em></td>
<td>ec:2.5.1.1</td>
<td></td>
</tr>
<tr>
<td><em>Farnesyltransf erase</em></td>
<td>ec:2.5.1.29</td>
<td></td>
</tr>
<tr>
<td><em>1-deoxy-D-xylulose-5-phosphate synthase</em></td>
<td>ec:2.2.1.7</td>
<td></td>
</tr>
<tr>
<td><em>(2E,6E)-farnesyl diphosphate synthase</em></td>
<td>ec:2.5.1.10</td>
<td></td>
</tr>
<tr>
<td><em>Putrescine N-methyltransferase</em></td>
<td>ec:2.1.1.53</td>
<td></td>
</tr>
<tr>
<td><em>Carboxylesterase</em></td>
<td>ec:3.1.1.1</td>
<td></td>
</tr>
</tbody>
</table>

*F- Fatty acid metabolism (map00071); I-Isoquinoline alkaloid biosynthesis (map00950); B- Butanoate metabolism (map00650); T1-Tyrosine metabolism (map00350); D – Drug metabolism-cytochrome (map00982) p450; A – Arginine and proline metabolism (map00330); T2 – Terpenoid backbone biosynthesis (map00900); T3 – Tropane, piperidine and pyridine alkaloid biosynthesis (map00960).
There are several GO terms which are being over-represented in the reference group. These are membrane part (GO:0044425), cytoplasmic part (GO:0044444), transport (GO:0006810) etc. represented by 10, 29 and 9 sequences respectively. Besides, there are four classes of GO terms that are unique to reference group. For example, identical protein binding (GO:0042802), UDP-glycosyltransferase activity (GO:0008194), lipid transport (GO:0006869) and response to reactive oxygen species (GO:0000302) represented by 2 sequences each. There was only one GO term (GO:0046483) unique to the test group representing heterocycle metabolic process by 19 sequences. These group of sequences mainly includes metabolic enzymes like cinnamoyl-CoA reductase, RNA polymerase beta subunit protein, pyrroline-5-carboxylate reductase, aspartate aminotransferase, spermidine synthase, 1-deoxy-d-xylulose 5-phosphate synthase, ethylene-induced esterase, Desacetoxyvindoline 4-hydroxylase, folylpolyglutamate synthase, ATP synthase beta subunit, ATP synthase CF₅ subunit IV protein, ATP synthase cf1 alpha subunit protein.

Besides, there were non enzymatic proteins like meiosis 5, magnesium-protoporphyrin ix monomethyl ester, senescence-inducible chloroplast stay-
green protein and tetrafunctional protein of glyoxysomal fatty acid betaoxidation.

3.4 Discussion

Drought has become one of the major abiotic constrains for crop production due to continuous depletion of water resources and increase of global temperature. Tea being a rain fed tree species, its production is highly dependent on availability of rain water. Keeping in mind the rapid expansion of water stress areas and increase of global temperature it has become important to study the drought tolerance mechanism in tea in order to devise future strategies to combat with increasing global drought scenario. Drought tolerance being a polygenic trait, improvement through conventional breeding techniques is difficult. One of the alternative ways is to identify and study the drought responsive genes and their regulatory pathways by comparative gene expression profiling of a drought tolerant tea cultivar with a susceptible one. In the present study, an attempt was made to identify the differentially expressed transcripts in a drought tolerant cultivar TV23 as compared to a drought susceptible cultivar S.3A/3 using SSH technique.

Four subtractive libraries were constructed at two stages of drought induction in order to identify the differentially expressed transcripts. Comparative transcript profiling between libraries of different stages of drought could identify several transcripts. As TV23 was found to be tolerant compared to S.3A/3 (Chapter II), our main focus was to identify the drought responsive transcripts that are highly represented and/or unique in TV23 at BWS (i.e., BWE3E4). Comparative analysis of BlastX results of BWE3E4 with BWE4E3 identified 36 proteins unique to BWE3E4 represented by 3 or more ESTs.

3.4.1 Comparative Transcriptome Profiling of TV23 and S.3A/3 at Before Wilting Stage

The comparative transcript profiling of the two cultivars at BWS identified many transcripts that are differentially expressed in TV23. These transcripts can be considered as early responsive and probably enabled TV23 to
maintain its physiological activities at low soil water content for a prolonged drought period (as observed in Chapter II) compared to S.3A/3. Water stress results in the imbalance of electron transport rate and thus accumulation of reactive oxygen species (Randy, 1995). We have identified several differentially expressed transcripts in TV23 at BWS (i.e., BWE3E4) which code for enzymes reported to be involved in neutralization of ROS. These enzymes mainly includes glutathione peroxidase, glutathione S-transferase and ascorbate peroxidase. Presence of these detoxifying enzymes (or antioxidants) in BWE3E4 implies that TV23 was more efficient in neutralization of ROS and therefore, more drought tolerant compared to S.3A/3. It has already been reported earlier that drought tolerance of plants is unequivocally related to efficient antioxidant cellular processes (Kranner et al., 2002; Montero-Tavera et al., 2008). Also, there are examples of overproduction of antioxidants in response to drought stress resulting in improved drought tolerance in other plants (Pastori and Foyer, 2002; Sunkar et al., 2006). The presence of scavenging enzymes such as ascorbate peroxidise which has been reported to neutralize the effect of \( \text{H}_2\text{O}_2 \) (Asada, 1992) may helps TV23 to combat negative effect of ROS during water stress. Ascorbate peroxidase gene from Arabidopsis have been over-expressed in tobacco chloroplast to improve drought tolerance (Badawi et al., 2004). Another ROS scavenging enzymes differentially expressed in TV23 at BWS is Glutathione S-transferase which catalyzes glutathione-dependent detoxification reactions and the reduction of hydroperoxides (Galle et al., 2005). Higher expression and enzymatic activity of glutathione S-transferase was observed in tolerant potato genotypes under cold and osmotic stress (Seppanen et al., 2000). Furthermore, this enzyme has been reported to act as binding protein that sequestrates flavonoids in vacuole for protection against environmental stress (Tahkokorpi et al., 2007). The metallothionein which appeared in a big way in both the libraries (BWE3E4 and BWE4E3) also acts as antioxidant during stress (Seki et al., 2001) and scavenge hydroxyl radical HO\(^\cdot\) (Thornalley and Vasak, 1985; Viarengo et al., 2000).
The BWE3E4 library also showed transcripts showing homology with proteins having protective functions i.e., those proteins which help to maintain cell structure and thus its functions. This group includes transcripts showing homology with LEA (late embryogenesis abundant) protein, water channel proteins like aquaporin 1 (MIP family, PIP subfamily), proline rich protein and dehydrin. This is in accordance to earlier studies reporting accumulation of LEA protein during drought stress (Hong et al., 2005; Gosal et al., 2009). It is thought that the LEA proteins act as molecular chaperones and maintain cell membrane structure and ion balance during drought stress (Close, 1997; Browne et al., 2002; Babu et al., 2004). Also, there are reports of engineering LEA genes for enhanced drought tolerance in crop plants (Cheng et al., 2002; Xu et al., 1996). Dehydrin which is also highly represented in BWE3E4 are reported to over express during drought (Cesar et al., 2003) and cold stress (Zhu et al., 2000) in other plants. It is associated with cellular turgor maintenance during drought stress (Cellier et al., 1998). The dehydrin, LEA and proline rich proteins are thought to provide stability to other proteins in osmotic stress (Ingram and Bartels, 1996). The aquaporin identified in our study belongs to the MIP (major intrinsic protein) family and PIPs (plasma membrane intrinsic protein) group. This protein has been reported to regulate water transportation and associated with physiological processes such as stomatal opening and closing, organ movement, cell elongation and stress responses (Johansson et al., 2000). This protein has already been used to develop transgenic Tobacco with improved drought tolerance (Yu et al., 2005). From the above discussion it is evident that the higher expression of antioxidant enzymes and other protective proteins in BWE3E4 as compared to BWE4E3 contribute towards enhance drought tolerance of TV23.

Drought inducible gene expression may be ABA dependent or ABA independent (Yamaguchi-Shinozaki et al., 2005; Shinozaki et al., 1996). We found certain transcripts in BWE3E4 that showed homology with proteins which has been reported to be ABA responsive. Such transcripts includes Asr1 gene, which is represented by 3 ESTs. The expression of Asr1 was reported to be induced by ABA and PEG (polyethylene glycol) mediated
drought in tomato (Kalifa et al., 2004). Transgenic Arabidopsis thaliana over expressing LLA23 gene (an ortholog of Asr1 from Lily) have been reported to show increased tolerance to drought and to alter the expression of ABA and other stress regulated genes (Yang et al., 2005). Thus, Asr1 gene which appeared in our library may play a role in drought tolerance behaviour of TV23 but still further investigation is required to confirm its role in tea plant under water stress. Also, it has been reported that accumulation of ABA activates a number of transcription factors like NAC, AREB/ABF, and MYB (Shinozaki and Yamaguchi-Shinozaki, 1997). Significantly, none of these transcription factors were highly represented in the library BWE3E4. This may be due to less number of sequenced ESTs (660) in the library BWE3E4. However, ERD15 gene does exist in the library whose expression is controlled by transcription factors induced by ABA.

We have identified several genes whose expression are unique to TV23 and represented by one or more ESTs. Although some genes are represented by single ESTs in TV23, it doesnot necessarily mean that they donot have a role in drought tolerance mechanisms and needs to be considered. We have also identified ESTs corresponding to transcription factors like WRKY, TCP transcription factor, ethylene responsive transcription factor which are reported to be involved in biotic and abiotic stress responses (Pandey and Somssich, 2009), regulation of growth and cell division (Li et al., 2005), regulation of transcription by ethylene (Ohta et al., 2000) respectively. Several protein homologues of ribosomal proteins were also identified which may function in restructuring the protein synthesis apparatus in tea during water stress (Kawasaki et al., 2001).

3.4.2 Functional Annotation and Enrichment Analysis

Functional annotation results clearly indicated that diversity of drought responsive transcripts in TV23 is much more complex than S.3A/3 at both the stages of drought induction. The diversity of GO term was also more in case of TV23 for molecular function, biological process and cellular component at specified GO levels. This differential response of TV23 therefore may contribute to its high drought tolerance compared to S.3A/3.
Comparison of BWE3E4 and BWE4E3 showed several proteins unique to TV23 at BWS represented under GOs unique to BWE3E4. Most of these proteins were reported to be induced under stress conditions in other plant species. For example, the gene for the multidrug resistance protein abc transporter family, represented 4 times under different unique GOs have been reported to be involved in guard cell hormonal signalling and water use (Klein et al., 2003). Their study, using Arabidopsis mutant for the gene mrp5-1, clearly demonstrated its involvement in several signalling pathways controlling stomatal movements, which is very vital for a plant to overcome water stress conditions. Among others, the genes for dehydrin (Cesar et al., 2003), aquaporin 1 of MIP family and PIP subfamily (Yu et al., 2005), abscisic stress ripening protein 1 (Kalifa et al., 2004), late embryogenesis-abundant protein (Hong et al., 2005; Gosal et al., 2009), were also reported to be induced during water stress. This group also includes genes encoding for proteins like NAC domain and ethylene responsive transcription factor. The NAC domain protein was found to bind a drought-responsive cis-element in the early responsive to dehydration stress 1 Promoter (Tran et al., 2004). It was observed that over-expressing NAC transcription factor in transgenic Arabidopsis induces the up-regulation of a number of stress inducible genes and the plant showed significantly increased drought tolerance. Hcp70-1 is a member of hsp-70 family and is involved in optimum growth, development, thermotolerance, and regulation of the heat shock response in plants (Sung and Charles, 2003). Alcohol dehydrogenase which appeared in this group is reported to catalyse the oxidation of various toxic aldehydes that accumulated as a result of side reactions of ROS with lipids and proteins to protect cells against the excessive accumulation of ROS (Sunkar et al., 2003). All these genes under GOs unique to TV23 are found to be related to stress directly or indirectly and some of them are reported to give drought tolerance in other plant species. Expression of these genes in TV23 therefore might be responsible for its contrasting drought tolerance.

The enrichment analysis between BWE3E4 (test) and WE3E4 (reference) revealed a GO term (GO:0046483; heterocycle metabolic process) unique to BWS of TV23 and includes 19 proteins (Figure 3.5). This unique group of
proteins includes enzyme like *Cinnamoyl-CoA reductase* which catalyzes the conversion of cinnamoyl-CoA esters to their corresponding cinnamaldehydes, in lignin biosynthetic pathways in plants (Bai *et al.*, 2003). Lignification is an adaptive process in plant to avoid different environmental stress like drought and water-logging. Therefore, the expression of *cinnamoyl CoA reductase* in TV23 supports the fact that the plant was under water stress and adapted to neutralize the effect by activating the lignin biosynthetic pathway. Another enzyme, *pyrroline-5-carboxylate reductase*, in this group is reported to be involved in synthesis of proline, an osmoprotectant synthesized by plants to protect cells from damage due to water stress (Shinozaki and Yamaguchi-Shinozaki, 1997). Another enzyme in this group is *spermidine synthase* that synthesizes spermidine, a polyamine, which may play pivotal role in plant defence against environmental stresses by controlling cell division, root formation, and flowering and also in slowing senescence (Tamaoki *et al.*, 2004). The spermidine synthesized by this enzyme may have dual function in plant stress tolerance: as a protectant in ROS scavenging and a membrane-bound protecting compound and as a signalling regulator in stress signalling pathways that leads to the build-up of stress tolerant mechanisms under stress conditions (Kasukabe *et al.*, 2004). Thus, induced expression of *spermidine synthase* in TV23 may play an important role in drought tolerance through membrane protection and/or regulating the stress signalling pathways. The enzyme 1-*deoxy-d-xylulose 5-phosphate synthase* catalyzes the initial step of the plastidic 2C-methyl-D-erythritol 4-phosphate (MEP) pathway that produces isopentenyl diphosphate (IPP), the common precursor for synthesis of a large numbers of plant isoprenoids (Estevez *et al.*, 2001). Plant isoprenoids are a group of compounds having more than 30,000 individual isoprenoid compounds. In addition to their function as pigments, phytohormones, defensive agents against pathogens and constituents of membranes, to name only a few, isoprenoids may also have a significant photoprotective role, which is of particular importance in plant stress responses (Penuelas and Munne-Bosch, 2005). Therefore, presence of this gene encoding this enzyme in TV23 at BWS may be correlated to the activation of MEP pathway for production of isoprenoids for stress tolerance.
in tea. Folypolyglutamate synthase, whose activity was reported in plant development (Mehrshahi et al., 2010), resulted in reduction of primary root elongation in Arabidopsis, due to disruption of the gene. (Srivastava et al., 2011). Therefore, appearance of ESTs of this particular protein in TV23 at BWS can be supported by the fact that plant under drought stress increases root length to cope with the water stress condition. The gene for non-enzymatic chloroplast stay-green protein is induced under water stress environment. Study of mutant for this gene in Soybean has been reported to increase water susceptibility. Therefore, most of the genes that are unique to TV23 (i.e., BWE3E4 library) are induced under stress and some of them were shown to contribute directly or indirectly to drought tolerance in other crop plants. However, the precise function of most of these genes thus identified in the present study is not well known and needs to be studied to understand the drought tolerance mechanisms in tea. We have identified a total of eight pathways unique to BWE3E4 represented by three or more enzymes (Table 3.5). The other enzymes of these pathways, not identified in our study, could have been identified by increasing the number of sequencing reactions. Being unique to BWE3E4, these might be the putative pathways, along with others, that play a role for higher drought tolerance of TV23 compared to S.3A/3. Further investigation is required to understand the role of these pathways during water stress in tea.

3.5 Conclusion

To date, limited number of studies on gene expression in Tea has been reported so far as drought stress is concerned. Although, recently a group from China (Shi et al., 2011) have reported 34.5 million ESTs, but those are not derived from stressed tissues. In the present study, we have identified 2365 drought responsive genes in two stages of drought induction in TV23 and S.3A/3. Comparison of BlastX results identified a large numbers of genes common to both the cultivars and most of them were reported to be common drought and other stress responsive genes in several other studies. We focused on a set of 49 genes that are represented by more than 3 ESTs in TV23 of which 36 are not reported from BWE4E3 library. These 36 genes were considered as unique to TV23 at BWS and therefore might be
responsible for high drought tolerance of TV23 along with others. However, ESTs of these proteins may have appeared in BWE4E3 library if we would have sequenced more numbers of ESTs. To confirm uniqueness of these ESTs and also to identify more unique and/or highly expressed genes, a high throughput sequencing of tea transcriptome under drought stress is required. There are several transcripts that are under-represented in TV23 compared to S.3A/3 and may also potentially contributing for high drought tolerance of TV23 and needs to be studied further to get a more clear picture of drought tolerance in tea. We have noticed a large numbers of ESTs in all the four libraries that did not show any homology with proteins reported in the public databases. This group of ESTs also needs to be considered for further study as they might play a role in drought tolerance.

The GO term comparison and enrichment analysis revealed several GOs that are unique to BWS of TV23. Most of these genes under those GOs were reported to be induced under stress environment and some of them are found to improve drought tolerance when introduced in plants. The unique occurrence of these genes in TV23 therefore may responsible for its higher drought tolerance compared to S.3A/3. However, detail expression and functional analysis of these genes under different degrees of drought stress, which is beyond the scope of this chapter, is necessary to understand the drought response mechanisms and subsequently to improve drought tolerance in tea and other crops.

References


