BIOLOGIA PLANTARUM 60 (1): 48-54, 2016 DOI: 10.1007/s10535-015-0565-3

Generation and characterization of expressed sequence tags in young roots of tea (*Camellia assamica*)

A. DAS[1,2], M. MUKHOPADHYAY[1,3], and T.K. MONDAL[1,4]*

Faculty of Horticulture, Uttar Banga Krishi Viswavidyalaya, Pundibari, Cooch Behar-785165, West Bengal, India[1]
Gauhati University, Institute of Science and Technology, Gopinath Bordoloi Nagar, Guwahati-781014, Assam, India[2]
Department of Botany, University of Kalyani, Kalyani, Nadia-741235, West Bengal, India[3]
Division of Genomic Resources, National Bureau of Plant Genetic Resources,Pusa, New Delhi-110012, India[4]

Abstract

Assam tea (*Camellia assamica*) is perennial crop susceptible to moisture stress. We used its tender roots to construct a cDNA library for the identification, functional annotation, and analysis of transcripts. A total of 811 full-length expressed sequence tags were generated. After processing and assembly, 207 unigenes comprising 58 contigs and 149 singletons were registered. Finally, 35.75 % of the unigenes could be assigned to functional categories based on the *Arabidopsis* proteome. There was 43 % of a coding GC content and 1 272 coding DNA sequences found in the unigenes. Codon usage analysis shows leucine as the highest (9.92 %) and tryptophan (2.0 %) as the lowest coded amino acids. Further, a comparative study with drought-induced genes of young roots (reported earlier) reveals that 4.83 % of genes required for normal growth of roots were also induced by a drought stress. Expressions of 10 unigenes under different abiotic stresses, such as drought, cold, and salinity, were further confirmed by RT-qPCR. The sequence tags generated in this study will be valuable resources for functional genomics study of tea and other woody crop plants in future.

Additional key words: cDNA library, contigs, gene expression, unigenes.

Introduction

Tea (*Camellia assamica*) is woody perennial crop grown worldwide. Being grown in upland conditions, tea plants are often subjected to drought stress (Mondal 2002, 2003, Konwar 2004) which severely affects growth of plants, crop yield, and several biochemical processes (Borthakur *et al.*1995, Mondal *et al.* 2001, Deka *et al.* 2006). Thus roots, particularly tender roots, are important organs for continuous supply of water and nutrients (Mukhopadhyay *et al.* 2012) and also primary sites of soil related stress perception. Transcriptional changes can result in successful adaptations leading to stress tolerance by regulating gene expression and signal transduction of regulatory or directly stress protecting proteins (Shinozaki and Yamaguchi-Shinozaki 2007). Although identification of expressed genes can be done in several ways, it mostly depends on the availability of sequence information of a crop species being studied. For a species whose genome sequence is not available, expressed sequence tags (ESTs) are still preferred as a choice for discovering new genes (Adams *et al.* 1991). Expressed sequence tags are short and single-pass sequence reads from an mRNA. They represent a snapshot of genes expressed in a given tissue and/or at a given developmental stage (Mekhedov *et al.* 2000). It also provides a good platform for several other uses such as cloning full length genes, development of genic markers, *etc*. Although construction of a cDNA library and generation and analysis of ESTs have been reported in tea (Park *et al.* 2004, Chen *et al.* 2005, Phukon *et al.* 2012), yet they are not sufficient in number in the absence of a full genome sequence. As ESTs represent the expressed portion of the genome, they are considered to be a more

Submitted 6 March 2015, *last revision* 1 July 2015, *accepted* 27 July 2015.
Abbreviations: ds - double strand; EST - expressed sequence tag; LD-PCR - long distance polymerase chain reaction; RT-qPCR - reverse transcription quantitative PCR.
Acknowledgements: We thank Prof. Swapan Kumar Ghosh from Uttar Banga Krishi Viswavidyalaya for encouraging us in pursuing the work, and the Director, the Tocklai Experimental Station, Tea Research Association, Jorhat, Assam for providing us the sequencing facility. We also thank Mr. Kamal Das for his technical help. The work was funded by the Department of Biotechnology, Govt. of India, New Delhi.
* Corresponding author; e-mail: mondaltk@yahoo.com

effective way for discovering functional genes (Yamamoto and Sasaki 1997, Thanh *et al.* 2011). Identification of new genes, alleles, and functional markers may help in a tea improvement program using molecular breeding techniques. Although conventional breeding and propagation techniques contributed significantly to the improvement of tea for the last several decades, yet due to limitations of conventional breeding and lack of a distinct mutant, application of biotechnology becomes alternative approach (Mondal 2014). In this study, we report the identification and functional characterization of young root specific ESTs of tea.

Materials and methods

Plants: Two-year-old vegetatively propagated well-rooted tea [*Camellia assamica* (Masters) Hung T. Chang cv. TV-23] plants (~ 91 cm in height) were planted in earthen pots (~ 30 cm in diameter) and grown under controlled conditions with a 14-h photoperiod, an irradiance of 300 µmol m^{-2} s^{-1}, a temperature of 25 ± 2 °C, and a relative humidity of 65 - 70 % in a greenhouse at Cooch Behar (26º 41″ latitude and 89º 31″ longitude), West Bengal, India. For two months, the plants were allowed to establish new growth. Thereafter, healthy tender roots were harvested and immediately frozen in liquid N_2.

RNA isolation and cDNA library construction: Total RNA was isolated from collected root samples (100 mg) by using a sodium dodecyl sulphate-LiCl based protocol **(**Das *et al.* 2013**)**. The messenger RNA was isolated and purified using a *PolyATract®* mRNA isolation system (*Promega*, Madison, USA). Double-strand (ds) cDNA was synthesized using long distance polymerase chain reaction (LD-PCR) as described in the manual of the *SMART™* cDNA library construction kit (*Clontech*, Palo Alto, USA). The dscDNA product was digested by proteinase K and *SfiI* (*Streptomyces fimbriatus I*) (*New England BioLabs*, Massachusetts, USA). The digested cDNA was further size fractionated through a *Chroma-spin 400* column (*Clontech*) in order to filter short cDNA fragments and oligonucleotides. The cDNAs in the first four fractions were precipitated using 3 M sodium acetate (pH 4.8), glycogen (20 mg cm^{-3}), and 2.5 volumes of 95% (v/v) ethanol. The ligation mixture was then transformed into competent cells of *Escherichia coli* strain XL1-Blue, and an aliquot was plated on a Luria-Bertani (LB) medium (supplied with 50 µg cm^{-3} ampicillin) containing 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal) and Isopropyl beta-D-thiogalactopyranoside (IPTG) in a final concentration of 40 ug cm^{-3} and 0.1 mM, respectively, to determine the recombinant percentage of the clones. The original library was amplified by an additional culture in 15 cm^3 of an LB/$MgSO_4$/maltose medium. The amplified library was split into 1 cm^3 aliquots by adding 7 % (v/v) dimethyl sulfoxide and stored at -80 °C for further use. The cloned lambda phage was transduced to a competent *E. coli* BM25.8 strain and plated on LB agar (supplied with 100 µg cm^{-3} carbenicillin) and incubated at 31 °C overnight. The white colonies were picked, and plasmids were isolated for sequencing the insert.

Generation, analysis, and functional annotation of ESTs: Transformed bacteria were randomly selected, and plasmid DNAs were extracted using a modified alkaline lysis method **(**Sambrook *et al.* 1989**)**. Inserted cDNAs were sequenced from the 5′ end by using a ready reaction mix of a *Big-dye* termination kit 3.1 and 5′ sequencing primer (5′-CGCCTGGAGACGCCATCC-3′) of a *SMART™* cDNA library construction kit (*Clontech*) in a *3130xl* DNA analyzer *(Applied Biosystems,* Foster City, USA). The raw ESTs were cleaned by removing the vector and adaptor sequences present at both 5′ and 3′ ends using *Sequence scanner v. 3.1* (*Applied Biosystems*). Trimmed sequences with more than 100 bp in length were assembled using the dirty data algorithm considering ambiguous base calls for poor matches to the exact base with the following criteria: *1*) gap optimization for small inserts and double-called bases through the *ReAligner* algorithm, *2*) prefer 3′ gap replacement off, *3*) a minimum overlap of 20 bp, and *4*) a minimum of an 85 % match using *Sequencher 4.1* (*Gene Codes Corporation*, Ann Arbor, USA). Both assembled contigs and singletons were considered as unigenes. *BLASTx* analysis of the unigenes was conducted against the non-redundant protein sequence databases of the National Center for Biotechnology Information **(***NCBI*; www.blast.ncbi.nlm. nih.gov/Blast.cgi), at an e-value threshold of $1E^{-05}$. The percentage of GC in the unigene set was calculated using the *GEECEE* program and codon usage. A functional GC percentage was analysed through the *CUSP* program of the *Pasteur lab* (http://mobyle.pasteur.fr/cgi-bin/ portal.py).

Various functional categories of unigenes were assigned on the basis of *GOslim* categories of *TAIR* (www.arabidopsis.org) based on annotations to terms in *GOslim* category/total annotations to terms in this ontology × 100. The locus hits of unigenes against the *TAIR 10* transcripts database (www.arabidopsis.org/index.jsp) of the model plant *Arabidopsis* (due to the unavailability of the tea genome) were selected at an e-value threshold of $1E^{-10}$ using *WU-BLAST*. For analysis of common genes in roots under drought-stress and normal conditions, our reported drought responsive unigene set **(**Das *et al.* 2012**)** was used as query for searching the homolog genes against the unigene set of roots under a normal growth by using the *BLASTn* program of *NCBI* (Stephen *et al.* 1997) at a threshold identity percentage of minimum 85.

Expression analysis of unigenes under abiotic stress: For validation, 10 ESTs were randomly chosen and their expression was studied with the following treatments: *1*) drought treatment was imposed by withholding water for 15 d in three young plants grown in polythene sleeves; well-watered plants of the same age were used as control; *2*) cold treatment was induced by keeping three young plants at 4 °C in a growth chamber for 24 h; seedlings kept under an ambient temperature were used as control; *3*) for salinity treatment, three young plantlets were kept in 150 mM NaCl for 24 h; control plantlets were fed only with clear water for the same period. Total root RNAs were isolated from the controls and different stress induced tissues (100 mg) and used for reverse transcription quantitative polymerase chain reaction (RT-qPCR) analysis. The first-strand cDNA was synthesized from 1 μg of the total RNA using *Leukemia virus* reverse transcriptase (*Invitrogen*, Carlsbad, USA) according to the manufacturer's protocol. To determine expression of candidate genes, RT-qPCR was performed with 2 mm^3 of the first-strand cDNA template and gene-specific primer pairs. Gene-specific RT-qPCR primers were designed by *Primer 3.0* (http://frodo.wi.mit.edu/primer3/) and synthesized commercially (*Sigma-Aldrich*, St. Louis, USA). The tea 26S ribosomal RNA (rRNA) was used as inner control (GenBank ID: AY283368) for RT-qPCR analysis (Mukhopadya *et al.* 2012). All primers for the candidate genes and 26S rRNA are listed in Table 1 Suppl. The resulting cDNA samples were diluted 20 times (1:19) in RNase-free water, and 2 mm^3 of the diluted cDNA was used in a total reaction volume of 25 mm^3 for determining the relative expression of the genes using a *QuantiFast SYBR Green* PCR master mix (*Qiagen*, New Delhi, India). Real-time PCR analysis was performed in a 96-well plate using a *Roche 454* qPCR system (*Roche*, Mumbai, India). The thermal cycling conditions of 95 °C for 5 min followed by 45 cycles of 95 °C for 15 s, 60 °C or temperatures as mentioned in Table 1 Suppl. for 30 s, and 72 °C for 30 s were used. The experiment was performed with at least three independent biological replicates and two technical replicates for each biological replicate. The specificity of the PCR reactions was confirmed by melting curve analysis of the amplicons. The comparative $2^{-\Delta\Delta C_T}$ method was used to calculate the relative quantity of each transcript in the samples (Schmittgen and Livak 2008). Statistical analyses were conducted using the *SAS* software of *JMP Genomics* (*SAS Institute*, Chicago, USA).

Results

The ratio of absorbances A_{260}/A_{280} of the isolated RNA was measured by *Biophotometer Plus* (*Eppendorf India*, Chennai, India) and found 1.88 which confirmed the good quality of the isolated RNA. Denaturing agarose gel electrophoresis of the isolated RNA samples showed distinct bands of 23S and 16S corresponding to molecular masses of 5.0 and 1.8 kb, respectively (data not shown). The synthesized ds cDNA appeared as 0.1 - 4 kb smear on the gel, which confirmed the successful synthesis of ds cDNA (Fig. 1 Suppl.**)**.

The titre of the original and amplified libraries were 3.15×10^5 and 1.82×10^9 pfu cm^{-3}, respectively, with a recombinant rate of 87.8 %. The amplification by PCR shows that the inserted cDNA fragments ranged from 0.3 to 2.0 kb. A total of 1 000 clones were randomly picked, cultured, and subsequently sequenced. Sequencing the cloned inserts produced a total of 811 ESTs which were deposited in *NCBI* (Genbank accessions: GH623575-GH624058; HS389643-HS389969). The read length of the sequences was found to be minimum 100 bp to maximum 700 bp. After removal of vector sequences, there were 586 ESTs with more than 100 bp in length for further analysis. The length of the vector trimmed sequences ranged from 101 to 691 bp with an average length of 376 bp. Finally, 346 sequences passed through the quality parameters, which included 46 % of sequences with no significant hits in the database (160 out of 346), 8.38 % of unknown proteins (29 out of 346), 22.54 % of predicted proteins (78 out of 346), 1.73 % of chaperones and heat shock proteins (6 out of 346), and 21.10 % of enzymes as well as other functional proteins. The sequences were clustered into 58 contigs comprising 2 to 28 individual ESTs in a single contig and 149 singlets producing a total of 207 putative unigenes (Fig. 1 and

Table 1. Assembled clusters that contained more than four full-length expressed sequence tags (ESTs) and E-value less than $1E^{-10}$.

Contig IDs	Homology	Number of ESTs	Plant species	Gene IDs	Score	E-value
Contig45	heat shock protein J	9	*Daucus carota*	AAG24642.1	253	1.00E-65
Contig38	metallothionein-like protein	9	*Camellia sinensis*	ABD97882	93.6	3.00E-17
Contig40	helicase, putative	5	*Ricinus communis*	XP_002520439.1	101	8.00E-20
Contig36	xylem sap protein	4	*Solanum lycopersicum*	ADQ57297.1	173	1.00E-41
Contig39	predicted protein	4	*Populus trichocarpa*	XP_002316984.1	125	1.00E-27
Contig2	proline-rich protein	4	*Gossypium hirsutum*	ABM05952.1	139	2.00E-31
Contig22	acyl-CoA binding protein (ACBP)	4	*Digitalis lanata*	CAB56693.1	89	2.00E-16
Contig30	conserved hypothetical protein	4	*Ricinus communis*	XP_002531502.1	97.4	5.00E-19

Table 2 Suppl.). An average length of the unigenes was 401 bp where consensus sequences of the contigs ranged from 103 to 685 bp and the singlets from 101 to 670 bp. The contigs containing more than four ESTs and an E-value less than $1E^{-10}$ are depicted in Table 1.

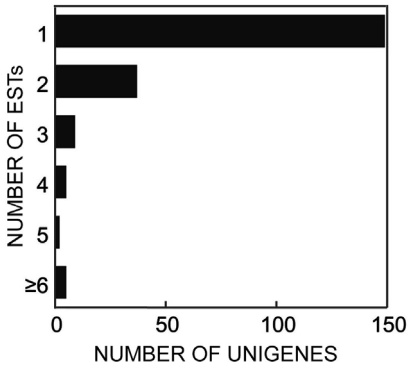

Fig. 1. Distribution and numbers of assembled standard full-length expressed sequence tags (ESTs).

The functional annotation of the full-length unigene set reveals that there were 42 % of overall GC content and 1 272 numbers of coding DNA sequences containing 43 % of coding GC. The codon usage analysis shows that leucine was the most often coded (9.92 %) amino acid followed by serine (8.02 %), and glycine (6.68 %). Besides, the least coded amino acid was tryptophan (2.0 %) followed by methionine (2.03 %) (Fig. 2 Suppl.). A total of 74 out of 207 (35.75 %) unigenes were assigned to functional categories, *i.e.*, 'cellular components', 'biological processes', and 'molecular functions' as defined in the *Arabidopsis* proteome (Fig. 3 Suppl.). In the 'cellular component' category, 'other intracellular component' related genes were found to be the largest group (20.24 %) followed by 'other cytoplasmic component' (14.88 %) and 'unknown cellular component' (13.31 %) related genes. In this category, 'Golgi apparatus', 'extracellular' and 'endo-plasmic reticulum' related genes were found to be the least group (0.6 % in each group) followed by 'mitochondrion' and 'cell wall' related genes (2.38 % in each group). In the 'biological processes' category, 'other cellular process' related genes were found to be the largest group (26.74 %) followed by 'other metabolic process' (19.79 %) and 'protein metabolism' (13.90 %) related genes. In this category, the least number was for 'signal transduction' and 'transcription' related genes (0.54 %) followed by 'DNA or RNA' metabolism (1.07 %) and 'transport' (2.14 %) related genes. In the 'molecular function' category, the highest number of genes was related to 'enzyme activities' (23 %) followed by 'protein binding' and 'unknown molecular functions' related genes (13 %). In this category, the least number of genes were related to 'transporter activity' and 'transcription factor activity' (1 % in each group) followed by kinase and nucleic acid binding factors (2 %). Interestingly, 6.95 % of genes related to abiotic or biotic stimuli were also found under normal growth conditions. Here, the definition of gene functional categories and their sub-categories were followed based on the Gene Ontology Consortium (http:// geneontology.org/).

Homolog analysis of unigenes between previously reported drought stress (246 unigenes) and control libraries (207 unigenes) of young roots of tea revealed 10 common genes. This result shows that probably 4.23 % of drought induced genes were also available under the normal growth of roots, on the contrary, 4.83 % of genes required for normal growth of roots were also induced by the drought stress. The identified homolog genes also coded for heat shock protein, profilin, lipid binding protein, 60S ribosomal protein L31, *etc.* (Fig. 2).

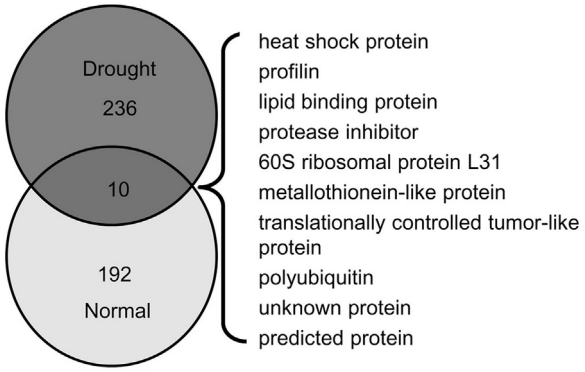

Fig. 2. Venn diagram of genes that expressed in roots under drought and normal conditions. The genes were put in three groups: genes expressed in normal or drought conditions and in both drought and normal conditions.

To validate the expression of the sequenced ESTs, we selected 10 random contigs to understand their expression pattern by RT-qPCR in root tissues (Fig. 3). The study reveals that, in general, the expression depended on gene regulation as well as stress. It was found that cyclin dependent kinase (Cacdk), which is regulatory enzyme, was upregulated under three different stresses compared to respective control plants. On the other hand, cystatin (Cacys) was highly expressed under the salinity stress in comparison to the drought or cold stresses. Similarly, an acetyl-co-A binding protein (Caacy) was up-regulated under the salinity. Ubiquitin activating enzyme E1 (Cauae), an important enzyme of ubiquitine-ligase pathways, showed a higher expression under cold and salinity than under drought, however, under all stresses higher than in the controls. Interestingly, ubiquitin-conjugating enzyme (Cauce) showed very different expression patterns. Under the drought, it was up-regulated whereas under the cold and the salinity, it was either similar to the control or down regulated. Eukaryotic translation initiation factor 5A4 (Cauti) was up-regulated under the drought and the cold but remained unchanged under the salinity stress. Late embryogenesis abundant protein 3L-1 (Calea) was highly expressed under the salinity, and moderately up-regulated under the

cold and drought stresses in comparison to the control plants. Another ubiquitin pathway enzyme, ubiquitin-conjugating enzyme 2 (Cauce2) remained unchanged except under the salinity where it was found to be up-regulated. Metallothionein (Camet; a transporter) did not show any response to the three stresses.

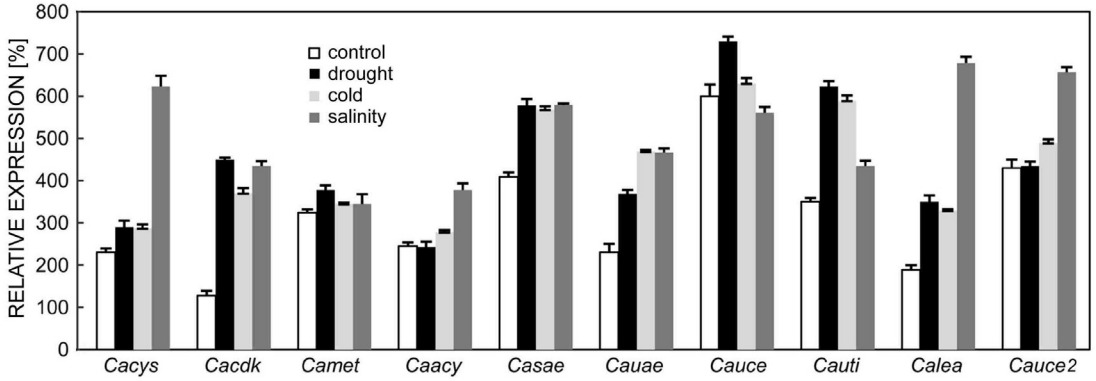

Fig. 3. Relative expression of different genes in *Camellia assamica* (*Ca*) roots as revealed by RT-qPCR analysis using three biological and two technical replicates; *cys* - cystatin, *cdk* - cyclin-dependent kinases, *met* - metallothionein, *acy* - acyl-CoA binding protein, *sae* - stress associated endoplasmatic reticulum protein, *uae* - ubiquitin activating enzyme E1, *uce* - ubiquitin conjugating enzyme, *uti* - eukaryotic translation initiation factor, *lea* - late embryogenesis abundant protein, *uce2* - ubiquitin conjugating enzyme E2.

Discussion

Tea plants remain under the vegetative stage for nearly 60 years. It requires continuous supply of nutrients and water. Therefore, roots of tea plants are an important organ for tea production and productivity. Between two different types of roots, younger roots are more active than the older ones for absorption of water and nutrients (Konwar 2004, Das *et al.* 2012). Additionally, the root is a primary organ which perceives any kinds of soil related stresses and thus emits the first layer of defence by activating several metabolic processes. Hence, understanding the molecular events underlying root physiology holds a promise to identify and exploit potentially endogenous defence mechanisms for various abiotic stresses. The present investigation is a novel endeavour to generate root specific ESTs which will be useful for identification of new genes.

Isolation of high-quality RNA is an initial but critical step in several molecular biology experiments. Similarly, the integrity of the RNA is also of a supreme importance for a number of downstream applications (Yadav *et al.* 2014). The synthesized ds cDNA appears as a 0.1 to 4 kb smear on the gel corroborating a successful synthesis of ds cDNA. A higher number of clones having inserts above 1 kb is one of the major challenges in cDNA library construction (Yadav *et al.* 2014). The PCR amplification shows that the inserted cDNA fragments ranged from 0.3 to 2 kb in a smear which ascertains the desired results of containing long cDNA fragments.

Construction and analysis of a cDNA library is essential tool for functional analysis of a genome of an organism. Simultaneously, amplification of the primary cDNA library is also necessary due to its instability and a limited volume. In order to screen out the a low abundance mRNA with 99 % probability, a well-constructed cDNA library should have 1.7×10^5 independent clones, and a desirable titer of the amplified cDNA should not be less than 1×10^9 pfu cm^{-3} (Thanh *et al.* 2011). The titre of the original and amplified libraries of the tea roots were 3.15×10^5 and 1.823×10^9 pfu cm^{-3}, respectively, with a recombinant rate of 87.8 %. Earlier reports suggested that the titers of unamplified and amplified libraries of shoots of *C. sinensis* are 6.8×10^5 and 7.2×10^9 pfu cm^{-3}, respectively (Chen *et al.* 2005), and 1.4×10^6 and 5.27×10^8 pfu cm^{-3}, respectively (Phukon *et al.* 2012), which are analogous results to our study.

The information provided by ESTs of randomly isolated gene transcripts generated under specific abiotic stress conditions provides an opportunity for gene discovery in addition to identification of underlying diverse processes of a plant that offers physiological responses against stresses (Wong *et al.* 2005). Genomic approaches to scrutinize genetic responses of plants for stress tolerance may help in understanding the mechanism of tolerance at molecular level (Chaves *et al.* 2003). In the present study, it was found that 157 ESTs, corresponding to 45.36 % of the total ESTs, had a significant homology with registered genes with known functions, which is comparable to that observed in sugarcane (38 %), grape (39.56 %), and citrus (34 %) (Phukon *et al.* 2012). The rest of unknown ESTs can become very crucial while discovering new genes with their actual functions.

In our expression analysis, we found four different categories of genes when tested under three different stresses such as cold, drought, and salinity. In the first category, we noticed that four genes, namely *Cacys*,

Caacy, *Calea,* and *Cauce* were more upregulated under the salinity compared to other two stresses, which indicates that these alleles might be salinity responsive in tea. It has also been reported that a cysteine protease inhibitor (cystatin) is salinity responsive in halophytes *Cakile maritima* (Megdiche *et al.* 2009) and *Amaranthus hypochondriacus* (Valdés-Rodríguez *et al.* 2007), which concur with our findings. Du *et al.* (2013) reported a role of acyl-CoA-binding protein in the regulation of PHOSPHOLIPASE Dα1 (*PLDα*) expression. *PLDα* has a role in the biosynthesis of the ABA through the regulation of expression of lipid messengers, which indicates that modulation of cellular lipid profiles is essential for the tolerance of abiotic stresses involving ABA signaling (Lu *et al.* 2013). Similarly, a late embryogenesis abundant protein is well-documented to confer the tolerance of salinity stress on rice (Xu *et al*. 1996) and mulberry (Lal *et al.* 2008). A ubiquitin-conjugating enzyme also confers osmotic stress tolerance on *Arabidopsis* (Chung *et al.* 2013).

In the second category, the expressions of *Camet* and *Caucp* remained unchanged under all the three different stresses. In contrary to our findings, Kumar *et al.* (2012) reported that metallothionein is upregulated under salinity stress in rice. Although the reasons are not known yet, differences in the system as well as allelic variations may contribute to the expression difference of the same gene in different plants. However, the ubiquitin-conjugating enzyme was found to be upregulated under drought stress (Zhou *et al.* 2010). In the third category of expressed genes, *Cacdk* and *Casae* were up-regulated almost uniformly under the stresses compared to the control plants, which indicates that these two genes might be involved in the cross-talk. An increased expression of cyclin-dependent protein kinases under salinity stress was also reported in rice (Huang *et al.* 2008), which supports our present finding. Moreover, *Cauae* and *Cauti* were also up-regulated, but their regulation was stress specific. Although the former gene was up-regulated in the cold and the salinity, the latter was up-regulated in the drought and the cold. Zhou *et al.* (2010) reported that an ectopic expression of the ubiquitin activating enzyme confers salinity tolerance. Sing *et al.* (2013) also reported in wheat that eukaryotic a translation initiation factor is upregulated under drought stress, which supports our present finding.

In conclusion, the present study generated the root specific genomic resources, many of which are trait specific. A challenge remains to utilize them to develop stress tolerant plants, or alternatively, if they can be used to develop useful markers for resistant breeding in the tea improvement program. These ESTs identified in this study may provide a useful genomic resource for biologists and plant breeders in developing new strategies for improving the stress-tolerant tea cultivar.

References

Adams, M.D., Kelley, J.M., Gocayne, J.D., Dubnick, M., Polymeropoulos, M.H., Xiao, H., Merril, C.R., Wu, A., Olde, B., Moreno, R.F., Kerlavage, A.R., McCombie, W.R., Venter, J.C.: Complementary DNA sequencing: expressed sequence tags and human genome project. - Science **252**: 1651-1656, 1991.

Borthakur, S., Mondal, T.K., Borthakur, A., Deka, P.C.: Variation in peroxidase and esterase isoenzymes in tea leaves. -Two and a Bud **42**: 20-23, 1995.

Chaves, M.M., Maroco, J.P., Pereira, J.S.: Understanding plant responses to drought: from genes to the whole plant. - Funct. Plant Biol. **30**: 239-264, 2003.

Chen, L., Zhao, L.P., Gao, Q.K.: Generation and analysis of expressed sequence tags from tender shoot cDNA library of tea plant (*Camellia sinensis*). - Plant Sci. **168**: 359-363, 2005.

Chung, E., Cho, C.W., So, H.A., Kang, J.S., Chung, Y.S., Lee, J.H.: Overexpression of VrUBC1, a mung bean E2 ubiquitin-conjugating enzyme, enhances osmotic stress tolerance in *Arabidopsis*. - PLoS ONE **8**: e66056, 2013.

Das, A., Das, S., Mondal, T.K.: Identification of differentially expressed gene profiles in young roots of tea [*Camellia sinensis* (L.) O. Kuntze] subjected to drought stress using suppression subtractive hybridization. - Plant mol. Biol. Rep. **30**: 1088-1101, 2012.

Das, A., Saha, D., Mondal, T.K.: An optimized method for extraction of RNA from tea roots for functional genomics analysis. - Indian J. Biotechnol. **12**: 129-132, 2012.

Deka, A., Deka, P.C., Mondal, T.K.: Tea. - In: Parthasarathy, V.A., Chattopadhyay, P.K., Bose, T.K. (ed.): Plantation Crops. Pp. 1-147. Naya Udyog, Kolkata 2006.

Du, Z.Y., Chen, M.X., Chen, Q.F., Xiao, S., Chye, M.L.: *Arabidopsis* acyl-CoA-binding protein ACBP1 participates in the regulation of seed germination and seedling development. - Plant J. **74**: 294-309, 2013.

Huang, Y.W., Tsay, W.S., Chen, C.C., Lin, C.W., Huang, H.J.: Increased expression of the rice C-type cyclin-dependent protein kinase gene, *Orysa; CDKC; 1*, in response to salt stress. - Plant Physiol. Biochem. **46**: 71-81, 2008.

Konwar, B.K.: Evaluation and development of planting material. - In: Goswami, B.K. (ed.): Notes on Tea Management. Pp. 1-6. Tea Research Association, Assam 2004.

Kumar, G., Kushwaha, H.R., Panjabi-Sabharwal, V., Kumari, S., Joshi, R., Karan, R., Mittal, S., Pareek, S.L., Pareek, A.: Clustered metallothionein genes are co-regulated in rice and ectopic expression of *OsMT1e-P* confers multiple abiotic stress tolerance in tobacco *via* ROS scavenging. - BMC Plant Biol. **12**: 107, 2012.

Lal, S., Gulyani, V., Khurana, P.: Overexpression of *HVA1* gene from barley generates tolerance to salinity and water stress in transgenic mulberry (*Morus indica*). - Transgenic Res. **7**: 651-663, 2008.

Lu, S., Bahn, S.C., Qu, G., Qin, H., Hong, Y., Xu, Q., Zhou, Y., Hong, Y., Wang, X.: Increased expression of phospholipase Dα1 in guard cells decreases water loss with improved seed production under drought in *Brassica napus*. - Plant Biotechnol. J. **11**: 380-389, 2013.

Megdiche, W., Passaquet, C., Zourrig, W, Zuily, F.Y., Abdelly, C.: Molecular cloning and characterization of novel cystatin

gene in leaves *Cakile maritime*, halophyte. - J. Plant Physiol. **166**: 739-749, 2009.

Mekhedov, S., Martinez, O., Ohlrogge, J.: Towards a functional catalogue of the plant genome: a survey of genes for lipid biosynthesis. - Plant Physiol. **122**: 389-401, 2000.

Mondal, T.K.: *Camellia* biotechnology: a bibliographic search. - Int. J. Tea Sci. **1**: 28-37, 2002.

Mondal, T.K.: Micropropagation of tea (*Camellia sinensis*). - In Jain, S.M., Ishii, K. (ed.): Micropropagation of Woody Trees and Fruits. Pp. 671-720. Kluwer Academic Publishers, Dordrecht 2003.

Mondal, T.K. (ed.): Breeding and Biotechnology of Tea and Wild Relatives. - Springer India, New Delhi 2014.

Mondal, T.K., Bhattacharya, A., Sharma, M., Ahuja, P.S.: Induction of *in vivo* somatic embryos from tea (*Camellia sinensis*) cotyledons. - Curr. Sci. **81**: 297-300, 2001.

Mukhopadyay, M., Bantawa, P., Das, A., Sarkar, B., Ghosh, P.D., Mondal, T.K.: Changes of growth, photosynthesis and alteration of leaf antioxidative defence system of tea [*Camellia sinensis* (L.) O. Kuntze] seedling under aluminum stress. - Biometal **25**: 1141-1154, 2012.

Park, J., Kim, J.B., Hahn, B.S., Kim, K.H., Ha, S.H., Kim, J.B., Kim, Y.H.: EST analysis and genes involved in secondary metabolism in *Camellia sinensis* (tea) using suppression subtractive hybridization. - Plant Sci. **166**: 953-961, 2004.

Phukon, M., Namdev, R., Deka, D., Modi, M.K., Sen, P.: Construction of cDNA library and preliminary analysis of expressed sequence tags from tea plant [*Camellia sinensis* (L) O. Kuntze]. - Gene **506**: 202-206, 2012.

Sambrook, J., Fritsch, E.F., Maniatis, T. (ed.): Molecular Cloning: a Laboratory Manual. - Cold Spring Harbor Laboratory Press, Cold Spring Harbor - New York 1989.

Schmittgen, T.D., Livak, K.J.: Analyzing real-time PCR data by the comparative CT method. - Natur. Protocols **3**: 1101-1108, 2008.

Shinozaki, K., Yamaguchi-Shinozaki, K.: Gene networks involved in drought stress response and tolerance. - J. exp. Bot. **58**: 221-227, 2007.

Singh, B., Chauhan, H., Khurana, J.P., Khurana, P., Singh, P.: Evidence for the role of wheat eukaryotic translation initiation factor 3 subunit g (TaeIF3g) in abiotic stress tolerance. - Gene **532**: 177-185, 2013.

Stephen, F.A., Thomas, L.M., Alejandro, A.S., Jinghui, Z., Zheng, Z., Webb, M., David, J.L.: Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. - Nucl. Acids Res. **25**: 3389-3402, 1997.

Thanh, T., Chi, V.T., Abdullah, M.P., Omar, H., Noroozi, M., Ky, H., Napis, S.: Construction of cDNA library and preliminary analysis of expressed sequence tags from green microalga *Ankistrodesmus convolutes corda*. - Mol. Biol. Rep. **38**: 177-182, 2011.

Valdés-Rodríguez, S., Guerrero-Rangel, A., Melgoza-Villagómez, C., Chagolla-López, A., Delgado-Vargas, F., Martínez-Gallardo, N., Sánchez-Hernández, C., Délano-Frier, J.: Cloning of a cDNA encoding a cystatin from grain amaranth (*Amaranthus hypochondriacus*) showing a tissue-specific expression that is modified by germination and abiotic stress. - Plant Physiol. Biochem. **45**: 790-798, 2007.

Wong, C.E., Li, Y., Whitty, B.R., Diaz-Camino, C., Akhter, S.R., Brandle, J.E., Golding, G.B., Weretilnyk, E.A., Moffatt, B.A., Griffith, M.: Expressed sequence tags from the Yukon ecotype of *Thellungiella* reveal that gene expression in response to cold, drought and salinity shows little overlap. - Plant mol. Biol. **58**: 561-574, 2005.

Xu, D., Duan, X., Wang, B., Hong, B., Ho, T., Wu, R.: Expression of a late embryogenesis abundant protein gene, *HVA1*, from barley confers tolerance to water deficit and salt stress in transgenic rice. - Plant Physiol. **110**: 249-257, 1996.

Yadav, R.K., Barbi, F., Ziller, A., Luis, P., Marmeisse, R., Reddy, M.S., Fraissinet-Tachet, L.: Construction of sized eukaryotic cDNA libraries using low input of total environmental meta transcriptomic RNA. - BMC Biotech. **14**: 80-90, 2014.

Yamamoto, K., Sasaki, T.: Large-scale EST sequencing in rice. - Plant mol. Biol. **35**: 135–144, 1997.

Zhou, G.A., Chang, R.Z., Qiu, L.J.: Over expression of soybean ubiquitin-conjugating enzyme gene *GmUBC2* confers enhanced drought and salt tolerance through modulating abiotic stress-responsive gene expression in *Arabidopsis*. - Plant mol. Biol. 72: 357-67, 2010.