

# GENERATION AND CHARACTERIZATION OF EXPRESSED SEQUENCE

## TAGS OF TEA (*Camellia sinensis* L. (O) Kuntze)

THESIS SUBMITTED

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY (SCIENCE) OF

THE UNIVERSITY OF NORTH BENGAL

*AKAN DAS, M.Sc*



DEPARTMENT OF BIOTECHNOLOGY

UNIVERSITY OF NORTH BENGAL

INDIA

2011

31 JAN 2013

245649

TH

633.72

D229g

## Acknowledgement

*This work would not have been possible without the help of various people and organizations who provided support, guidance and encouragement in different ways during the course of this study.*

*I have much pleasure in expressing my deepest sense of gratitude to my supervisor Dr. Dipanwita Saha, Assistant Professor, Department of Biotechnology, University of North Bengal (NBU) for her inspiration, encouragement and guidance.*

*I would like to thank my co-supervisor Dr. Tapan K. Mondal, Assistant Professor, Biotechnology laboratory, Faculty of Horticulture, Uttar Banga Krishi Viswavidyalaya (UBKV), Cooch Behar, West Bengal, who is presently working as a Senior Scientist, National Research Center on DNA Fingerprinting, NBPGR, Pusa, New Delhi, for his untiring guidance, help and involvement throughout the course of my academic endeavors. His incredible availability and pertinent comments largely improved the quality of the work. I never forget the numerous and passionate hours of discussion on and beyond my study topics with him and his confidence in me. I sincerely cannot express in words, my gratefulness to his tremendous support and able guidance throughout the course of this research work.*

*I sincerely express my deep sense of gratitude to Dr. Sudripta Das, Incharge, Department of Biotechnology, Tocklai Experimental Station (TES), TRA, Jorhat, Assam for his kind co-operation, help and constructive suggestions during part of my research works at Molecular Genetics Lab of TES. I extended my gratitude to Dr. Swapan Kumar Ghosh, Dean, Faculty of Horticulture and Dean of students' welfare, UBKV, for his help in various ways and providing me hostel accommodation at UBKV during the course of my research works at Biotechnology Laboratory, UBKV, Coochbehar, West Bengal, India.*

*I am indebted to Vice-chancellor, UBKV, Coochbehar, West Bengal for appointing me as a Junior Research Fellow and subsequently as Senior Research Fellow in the Biotechnology laboratory, Faculty of Horticulture, UBKV. I am also indebted to*

Director, TES, TRA, Assam for his kind permission to work at their Biotechnology labs and to stay at the scholars' hostel inside the campus during part of my research works. It is extended to the Director, NBPGR, New Delhi for allowing me to access the library facilities.

I am thankful to the Head, Dept. of Biotechnology, University of North Bengal, Siliguri, West Bengal for his help in the process of registration as a Ph.D student of the Dept. and to the Head, Dept. of PCP, UBKV for permitting me to access the available facilities in the Department. It is extended to the Incharge and staff of Central Instrumentation Centre, UBKV for allowing me to operate various instruments in the centre.

I sincerely express my special thanks and regard to Dr. Aniruddha Saha, Associate Professor, Dept. of Botany, NBU and to Dr. Mohan C. Kalita, Professor, Dept. of Biotechnology, Gauhati University for their valuable suggestions and encouragement during the course of this study.

I would like to express my deep sense of gratitude to all the scientists and staff of Plant Improvement Division, TES, TRA, Jorhat and the staff of Electron microscopy Lab, Sophisticated Analytical Instrument Facility, North Eastern Hill University, Shillong, Meghalaya for their help in ultrastructural studies.

I appreciate the help and support of my senior colleagues namely Dr. Pranay Bantawa, Mrs. Olivia Saha Roy, Mainaak Mukhapadhaya and Pratap Subba.

I gratefully acknowledge the unconditional help and support that I received from Partha, Mayukh, Aditya, Himadri, Swami, Johnson, Saurabhi, Madhu and all members of M. N. Saha Hall of UBKV, Coochbehar as well as from Sangeeta ba, Swarnalee ba, Raju, Tirtha, Bornali, Sushmita, Mayuri, Niraj, Prasenjit, Parveen, Nilakshi, Dhruva da, Debojit, Traloikya during my stay at TES, TRA, Assam.

I am always thankful to some personalities for their valuable time, unconditional help and support that I received in various ways namely Kamal, Mantu, Diganta da, Susheel ji, Himadri Shankar, Chittaranjan and Jyotsna.

*I also acknowledge to the Department of Biotechnology, Govt. of India for their financial support for pursuing this work which would not have been possible, otherwise.*

*It would not have been possible for me to accomplish my PhD work without the constant love, affection and encouragement of my beloved parents, elder brothers, sisters and all of my relatives. They stood by my side throughout the course of this study.*

Dated: 11.04.11

  
(Akan Das)

## LIST OF ABBREVIATIONS

(-)	approximate
1 <sup>o</sup> PCR	primary polymerase chain reaction
2 <sup>o</sup> PCR	secondary polymerase chain reaction
' (prime)	denotes a truncated gene at the indicated side
A <sub>260</sub>	absorption at 260 nm
A <sub>280</sub>	absorption at 280 nm
A <sub>320</sub>	absorption at 320 nm
A <sub>580</sub>	absorption at 580 nm
AA	ascorbic acid
ABA	abscisic acid
ABF	abscisic acid factor
ABRE	abscisic acid responsive element
ACC	1-aminocyclopropane-1-carboxylic acid
AFLP	amplified fragment length polymorphism
AMV	avian myeloblastosis virus reverse transcriptase
AP2	apetala2
APX	ascorbate peroxidase
ATP	adenosine triphosphate
B.C	before christ
BEEM	better equipment for electron microscopy
BHT	butylated hydroxytoluene
BLAST	basic local alignment search tool
BMAA	B-methylamino-L-alanine
bp	base pair
CAT	catalase
CBF	C-repeat binding factor
CDD	conserved domain database
cDNA	complementary deoxyribonucleic acid
cDNA-AFLP	cDNA-amplified fragment length polymorphism
CDP-ME	4-Diphosphocytidyl-2C-methyl-D- erythritol
cds	coding sequence

cfu	colony forming unit
chl	chlorophyll
CI	chloroform isoamylalcohol
cm	centimetre
CoA	coenzyme A
CTAB	cetyl trimethyl ammonium bromide
CUL	cullin
d	days
dbEST	database of expressed sequence tags
DCL1	dicer-like 1
DEPC	diethyl pyrocarbonate
dia	diameter
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DNase I	deoxyribonuclease I
DNPH	2,4-dinitrophenylhydrazine
DNS	3, 5-dinitro-salicylic acid
dNTP	deoxynucleotide triphosphate
DRE	dehydration-responsive element
DREB	dehydration responsive element binding
dsDNA	double-stranded deoxyribonucleic acid
DTT	1,4-dithiothreitol
DW	dry weight
EC	electrical conductivity
ECG	epicatechin gallate
ECQ	epicatechin quinone
EDTA	ethylenediaminetetraacetate
EGCG	epigallocatechin gallate
EGCGQ	epigallocatechin gallate quinone
ERD	early-responsive to dehydration
ERF	ethylene response factor
ESTs	expressed sequence tags
EST-SSRs	expressed sequence tags-simple sequence repeats
et. al	<i>et alia</i>

e-value	expected value
Exp.	experimental
FAO	food and agricultural organisation
FASTA	fast all
FCR	folin-ciocalteu reagent
FSL	forward subtracted library
FW	fresh weight
g	gram
GA	gibberellic acid
GC	guanine-cytosine
GLR	glutamate receptor
GO	gene ontology
GR	glutathione reductase
GSSG	glutathione disulfide
GSH	$\gamma$ -glutamylcysteinylglycine (glutathione)
GST	glutathione S-transferase
h	hours
ha	hectars
HCl	hydrochloric acid
HMM	hidden markov models
HPLC	high pressure liquid chromatography
HSFs	heat shock factors
HSPs	heat shock proteins
HYL1	hyponastic leaves 1
IAA	indole acetic acid
ID	Identity
IPTG	Isopropyl $\beta$ -D-1-thiogalactopyranoside
JA	jasmonic acids
JTT	Jones, Taylor & Thornton
kb	kilobase
kcal	kilocalorie
kDa	kilodalton
kg	kilogram
kPa	kilopascal

kV	kilovolt
LB	luria bertani
LEA	late embryogenesis abundant proteins
LiCl	lithium chloride
LOR	length of read
M	molar
MAP	mitogen activated protein
max.	maximum
MDA	malondialdehyde (or 3,4-methylenedioxyphenylisopropylamine)
MEGA	molecular evolutionary genetics analysis
mg	milligram
µg	microgram
µl	microlitre
µm	micrometer
µmol	micromole
µM	micromolar
m	meter
min	minutes
MFE	minimal free energy
MIPS	major intrinsic proteins
miRNAs	microRNAs
ml	millilitre
mM	millimolar
mm	millimeter
mmol	millimole
mRNA	messenger ribonucleic acid
mV	millivolt
NAC	nitrogen assimilation control
NADPH	nicotinamide adenine dinucleotide phosphate-oxidase
NAM	no apical meristem
NBT	nitroblue tetrazolium
NCBI	national center for biotechnology information
ng	nanogram
nm	nanometer

nmol	nanomole
nt	nucleotides
O <sub>2</sub> <sup>·-</sup>	superoxide radical
OD	optical density
OH <sup>·</sup>	hydroxyl radical
ORF	open reading frame
p	probability
PAL	phenylalanine ammonia-lyase
PCI	phenol chloroform isoamylalcohol
PCR	polymerase chain reaction
pfu	plaque-forming unit
pg	pico gram
PIP	plasma membrane intrinsic protein
pmol	picomole
POD	polyphenol oxidase
POP7	performance optimized polymer 7
POX	peroxidase
PR	pathogen related
psi	pound per square inch
PVP	polyvinylpyrrolidone
QACs	quaternary ammonium compounds
QV	quality value
RACE	rapid amplification of complementary deoxyribonucleic acid ends
RAPD	random amplification of polymorphic DNA
RFLPs	restriction fragment length polymorphisms
RING	really interesting new gene
RISC	RNA-induced silencing complex
RNA	ribonucleic acid
RNase	ribonuclease
ROS	reactive oxygen species
rpm	revolution per minute
rRNA	ribosomal RNA
RS	reducing sugar
RT-PCR	reverse transcriptase-PCR

RWC	relative water content
S	seconds
SAGE	serial analysis of gene expression
SA-PMPs	streptavidin-paramagnetic particles
SCF	skp, cullin & f-box
SDS	sodium dodecyl-sulphate
siRNAs	small interfering RNAs
SMART	switching mechanism at 5' end of RNA template
SOC	super optimal broth with catabolic repressor
SOD	superoxide dismutase
SOS	salt overly sensitive
ss cDNAs	single-strand complementary deoxyribonucleic acids
SSC	saline-sodium citrate
SSH	suppression subtractive hybridisation
SSR	simple sequence repeat
STMS	sequence tag microsatellite site
TAE	tris-acetate- ethylenediaminetetraacetate
TAIR	the <i>Arabidopsis</i> information resource
TBA	thiobarbituric acid
TC	tentative consensus
TCA	trichloroacetic acid
TE	tris-hydrochloric acid and ethylenediaminetetraacetate
TEM	transmission electron microscope
TIF	transcription initiation factor
Tris	tris (hydroxymethyl) amino methane
Tris-HCl	tris (hydroxymethyl) amino methane-hydrochloric acid
TSS	total soluble sugar
TV	tocklai variety
TW	turgid weight
u	unit
ub	ubiquitin
UDL	unigene of drought stressed leaves
UEST	unique expressed sequence tags of drought stressed roots
UESTM	unique expressed sequence tags markers

UN	united nations
UPASI	united planters' association of southern India
US	united states
USL	unigenes of standard leaves
USR	unigenes of standard roots
USSR	union of soviet socialist republics
UV	ultraviolet
UV-VIS	ultraviolet-visible
UWL	unigenes of winter dormancy stressed leaves
v	version
V	voltage
v/v	volume by volume
vol	volume
W	watt
WU-BLAST	washington university-basic local alignment search tool
w/v	weight by volume
x-gal	5-bromo-4-chloro-3-indolyl-beta-D-galacto-pyranoside

## LIST OF TABLES

- Table 2.1: Drought stress responses of major antioxidants in various crop plants.
- Table 2.2: Effect of drought stress on RWC, photosynthesis rate, transpiration rate, stomatal conductance and pigment content of tea leaves.
- Table 2.3: Effect of drought stress on ROS, lipid peroxidation and level of membrane stability.
- Table 2.4: Effect of drought stress on different biochemicals.
- Table 3.1: A snapshot of EST analysis in woody plants.
- Table 3.2: RNA yield and quality by different methods.
- Table 3.3: Assembled clusters that contain more than 4 drought induced ESTs and e-value less than  $1E-10$ .
- Table 3.4: Drought induced unigenes containing protein domain, signal peptide and signal anchor.
- Table 3.5: Identified ESTs of tea roots which were known to be associated with drought stress.
- Table 3.6: Detection of EST-SSRs and putative functions of the SSR-containing drought induced unigenes.
- Table 3.7: Assembled clusters that contain more than 4 full-length ESTs and e-value less than  $1E-10$ .
- Table 3.8: Details of the ESTs under normal growth, drought and winter dormancy stresses of tea leaves.
- Table 4.1: Identification of miRNAs in important crop plants.
- Table 4.2: Predicted miRNAs of *Camellia sp.*
- Table 4.3: Potential targets of the candidate miRNA families.

## LIST OF FIGURES

- Fig. 2.1: Flowchart of main antioxidants and redox signaling processes in higher plants.
- Fig. 2.2: Drought induction experiment of tea seedlings.
- Fig. 2.3: Photographs of photosynthesis rate, transpiration rate and stomatal conductance measurement using a photosynthesis system, an uprooted plant, root length measurement of control and drought induced plants of TV-23 and roots for various experiments.
- Fig. 2.4: Light microscopic images of stomata and transverse section of petioles.
- Fig. 2.5: Ultra-structural images of TEM.
- Fig. 2.6: Estimation of different antioxidative enzymes activity of leaves and roots of drought stressed and control tea plants.
- Fig. 3.1: Functions of drought associated genes.
- Fig. 3.2: Transcriptional regulatory networks of drought as well as other abiotic stress signals.
- Fig. 3.3: Agarose gel electrophoresis of total RNA isolated from tea roots by different reagents and methods.
- Fig. 3.4: Evaluation of *Rsa*I restriction digestion of dsDNAs.
- Fig. 3.5: Adaptor ligation efficiency analyses of 'tester' dsDNAs.
- Fig. 3.6: Analyses of primary and secondary PCR products.
- Fig. 3.7: Subtraction efficiency analyses of SSH libraries.
- Fig. 3.8: Blue white screening of transformed DH10B *E. coli* cells.
- Fig. 3.9: Extracted plasmids of transformed cells containing the DNAs of interest.
- Fig. 3.10: Distribution and numbers of assembled drought induced ESTs.
- Fig. 3.11: Codon usage analyses of drought induced unigenes.
- Fig. 3.12: Phylogenetic tree of drought associated conserved protein domain sequences.
- Fig. 3.13: Gene ontology classification of drought induced unigenes as defined for *Arabidopsis* proteome.

- Fig. 3.14: SCF complex mediated ubiquitin-proteasome degrading pathway associated ESTs of tea roots under drought stress.
- Fig. 3.15: Size ranges of ds cDNA synthesized through LD PCR.
- Fig. 3.16: Size fractionation of cDNAs through chroma spin 400 column.
- Fig. 3.17: Plaque forming units of unamplified and amplified full-length cDNA libraries.
- Fig. 3.18: Distribution and numbers of assembled standard full-length ESTs.
- Fig. 3.19: Codon usage analysis of standard full-length unigenes.
- Fig. 3.20: Gene ontology classification of full-length unigenes as defined for *Arabidopsis* proteome.
- Fig. 3.21: Venn diagram of genes that expressed in roots under drought and normal conditions.
- Fig. 3.22: BLASTx analysis of tea leaf unigenes under normal growth, drought and winter dormancy stresses.
- Fig. 3.23: Functional assignment of unigenes under normal growth, drought and winter dormancy stresses of tea leaves as defined for *Arabidopsis* proteome.
- Fig. 3.24: Venn diagram of drought stress induced genes in roots and leaves.
- Fig. 3.25: Venn diagram of genes that expressed in leaves under drought, winter dormancy and normal conditions as well as those expressed mutually among these stresses.
- Fig. 4.1: Different steps involved in the identification process of conserved miRNAs and their targets.
- Fig. 4.2: Predicted hairpin secondary structures of candidate pre-miRNAs.
- Fig. 4.3: Conservation of tea miRNA (csi-miR408) in diverse plant species.
- Fig. 4.4: Phylogenetic relationship among the members of each miRNA family, namely: miRNA414, miRNA1122 and miRNA408.
- Fig. 4.5: Predicted miRNA targets and their complementary sites with miRNAs.

## PREFACE

Understanding the responses of plants to their environment in terms of adaptability and performance is of paramount importance for selecting and/or developing cultivars that can withstand the unwanted environmental changes. Apart from the interest in factors important for ecological and evolutionary distribution, much of the present day interest in agri-biotechnology is on selecting stress resistant quality cultivars. Modern agriculture is affected by environmental factors such as water-logging, drought, temperature, light, and salt stress. Sustainable agriculture in harsh environment requires an understanding of the ways that plant antioxidants and genes respond to abiotic stresses. In recent years, biochemical and transcriptional analysis of plants under abiotic stress brought to light some of the complex mechanisms of stress tolerance. Understanding of mechanisms that operate in signal perception, transduction and downstream regulation as well as cellular pathways that are involved in abiotic stress responses provide valuable information of plant's tolerance to such stresses.

Drought tolerance is a complex trait, expression of which depends on the action and interaction of different morphological, physiological and biochemical characters. The level of drought stress correlates with the degree of changes in these characters which provides the scope for identifying drought tolerant cultivars and the study of tolerance mechanism at molecular level thereof. Under drought, plants accumulate various compatible solutes and antioxidants for maintaining water homeostasis. For the biosynthesis of these solutes and compounds, numbers of genes and genetic pathways get activated. Hence, a detail knowledge of physiological and biochemical changes under drought provides the scope of correlation with transcriptional analysis, and considerably broadens the possibility of understanding the lying mechanism of tolerance.

Identification and characterization of genes expressed differentially under drought stress has gained much attention in crop plants research. Dramatic improvements in DNA sequencing technology have paved the way for the use of large-scale single-pass cDNA sequencing which has given rise to large EST collections to address many

biological questions. It facilitates to study the genetic mutation, conservation, repeat markers, phylogeny as well as functional and comparative genomics, finally to answer questions of tolerance to drought by some genotypes. The answers itself provide us the scope for identifying tolerant cultivar and understanding their lying mechanism of tolerance. Moreover, recognition of genes and genetic pathways that involved in response to drought enhances the possibility of promoting crop improvement through direct genetic modification.

Being a woody perennial with a life span of more than 60 years, tea plant experiences several abiotic stresses. Amidst all, drought is an important recurrent constraint to world tea cultivation including India. For the identification, selection and/or production of drought tolerant cultivars, there is an urgent need of comprehensive physiobiochemical and molecular analyses of drought stress responses of this important cash crop. Hence, the present work as embodied in this thesis entitled 'Generation and Characterization of Expressed Sequence Tags of Tea (*Camellia sinensis* (L.) O. Kuntze)' was initiated in the year 2007 with broad objectives of physiobiochemical and transcriptional analyses in tea under drought stress keeping in mind the following specific objectives.

- 1) Physiological and biochemical analyses of drought stress responses.
- 2) Construction of cDNA libraries.
- 3) Sequencing of Expressed Sequence Tags and *in silico* analysis

The status of this work with respect to physiobiochemical and transcriptional analyses of tea and their results and inferences drawn thereof is presented in seven major chapters in the thesis. Besides, there are also supplementary details given as appendix at the end.

## TABLE OF CONTENTS

Acknowledgement	i-iii
List of Abbreviations	iv-x
List of Tables	xi
List of Figures	xii-xiii
Preface	xiv-xv

---

### CHAPTER I INTRODUCTION

1.1	Tea: an overview	1
1.2	Origin and distribution	1-2
1.3	Economic importance	2
1.4	Cytology	3
1.5	Functional genomics	3-4

### CHAPTER II ANATOMICAL AND PHYSIOBIOCHEMICAL ANALYSES

2.1	Review of literature	5-17
2.1.1	Abiotic stress responses in plants	5-6
2.1.2	Physiological responses to drought stress	6-7
2.1.3	Responses of biochemical solutes to drought stress	7-10
2.1.3.1	Carbohydrates	7-8
2.1.3.2	Amino acids, amides and proteins	8-9
2.1.3.3	Quaternary ammonium compounds	9
2.1.3.4	Polyols	10
2.1.4	Antioxidative responses to drought	10-12
2.1.4.1	Non-enzymatic antioxidants	11
2.1.4.2	Enzymatic antioxidants	12
2.1.5	Brief report on works done in tea plants	13-17
2.1.5.1	Growth and physiology	13-15

2.1.5.2	Biochemical changes	15-17
2.2	Materials and methods	18-26
2.2.1	Plant materials, stress induction and sample collection	18
2.2.2	Determination of soil moisture content, photosynthesis rate, transpiration rate and stomatal conductance	19
2.2.3	Determination of relative water content	19
2.2.4	Anatomical studies using light and transmission electron microscopy	19-20
2.2.5	Estimation of TSS, RS, starch, pigments, phenols, proline, ascorbic acids and ABA contents	20-23
2.2.6	Determination of ROS, lipid peroxidation, hydrogen peroxide and membrane stability levels	23-24
2.2.7	Extractions and assays of antioxidative enzymes	25-26
2.2.8	Statistical analyses	26
2.3	Results	27-29
2.3.1	Stress induction and plants growth	27
2.3.2	Anatomical studies	27-28
2.3.3	ROS, lipid peroxidation and membrane stability	28
2.3.4	TSS, RS, starch, protein, proline, phenolics, ascorbic acids and ABA content estimation	28
2.3.5	Activity of antioxidative enzymes	28-29
2.4	Discussion	30-35
2.4.1	Plants growth and anatomical studies	30-32
2.4.2	Determination of ROS, lipid peroxidation and membrane stability levels	32-33
2.4.3	Determination of TSS, RS, starch, protein, proline, phenolics, ascorbic acids and ABA contents	33-34
2.4.4	Activities of antioxidative enzymes	34-35

## CHAPTER III GENERATION AND CHARACTERIZATION OF ESTs

3.1	Review of literature	36-53
3.1.1	Abiotic stress associated genes and genetic mechanisms	36-37
3.1.1.1	Signaling cascades and transcriptional control	36
3.1.1.2	Heat shock proteins and chaperons	37
3.1.1.3	Ion and water transport	37
3.1.2	Drought stress associated genes and their regulation	38-39
3.1.2.1	ABA independent pathways	38
3.1.2.2	ABA dependant pathways	39
3.1.3	Importance for studying tree genomics	39
3.1.4	Techniques for transcriptome analysis	40-41
3.1.4.1	Subtractive hybridization	40-41
3.1.4.2	Advantages of SSH techniques	41
3.1.5	Transcriptome analysis of woody plants: ESTs as a strategy	42-51
3.1.6	Transcriptome analysis of tea	51-53
3.2	Materials and methods	54-82
3.2.1	Sample preparation	54
3.2.2	Protocol optimization for RNA isolation	55
3.2.3	Isolation of total RNA	55-56
3.2.4	Quantity and quality checking of RNA	56
3.2.5	Isolation of mRNA	56-57
3.2.6	Precipitation of mRNA	57
3.2.7	Determination of mRNA quantity and purity	57-58
3.2.8	Construction of drought stressed cDNA libraries	58-68
3.2.8.1	First-strand cDNA synthesis	58
3.2.8.2	Second-strand cDNA synthesis	58-59
3.2.8.3	Restriction digestion of complementary dsDNA	60
3.2.8.4	Evaluation of size ranges of the synthesized dsDNAs	60
3.2.8.5	Adaptor ligation to the restriction digested 'tester' dsDNAs	61-62
3.2.8.6	Analysis of ligation efficiency	62-63

3.2.8.7	PCR select subtraction of differentially expressed transcripts	63-64
3.2.8.8	Amplification of differentially expressed transcripts	64-65
3.2.8.9	Gel analysis of PCR products	65
3.2.8.10	PCR analysis of subtraction efficiency	65-66
3.2.8.11	Cloning of differentially expressed transcripts	66-67
3.2.8.12	Transformation of cloned DNAs	67
3.2.8.13	Blue white screening and calculation of transformation efficiency	67-68
3.2.9	Construction of standard full-length cDNA library	68-74
3.2.9.1	First-strand cDNA synthesis	68
3.2.9.2	Amplification of cDNA by LD PCR	69
3.2.9.3	Proteinase K digestion	69
3.2.9.4	Sfil digestion	70
3.2.9.5	Size fractionation of cDNAs	70-71
3.2.9.6	Ligation of cDNA to $\lambda$ TriplEx2 vector	71
3.2.9.7	Lambda phage packaging of ligated cDNA	71-72
3.2.9.8	Titering of the unamplified library	72
3.2.9.9	Determination of recombinant percentage	72-73
3.2.9.10	Amplification of the library	73
3.2.9.11	Conversion of phage ( $\lambda$ TriplEx2) to plasmid (pTriplEx2)	73-74
3.2.10	Liquid culturing of transformed <i>E. coli</i> cells	74
3.2.11	Extraction of recombinant plasmids	74-76
3.2.12	Sequencing of inserts in plasmid DNAs	76-77
3.2.13	Base calling, vector trimming and homology searching of EST sequences	77-78
3.2.14	Functional annotation and analysis of drought induced unigenes	78-80
3.2.14.1	GC percentage and codon usage analyses	78
3.2.14.2	Mining of ORFs, protein domains and construction of phylogenetic trees	79

3.2.14.3	Mining of signal peptide and peptide anchor in protein domain sequences	79
3.2.14.4	Gene ontology analysis	79-80
3.2.14.5	Identification of drought associated genes and pathways	80
3.2.14.6	Mining of genic microsatellite markers	80
3.2.15	Analyses of standard full-length unigenes	80
3.2.16	Comparative analyses of transcripts under normal growth, drought and winter dormancy stress	81-82
3.2.16.1	Tissue specific expression of genes under drought	81
3.2.16.2	Comparative functional annotation of unigenes	82
3.3	Results	83-95
3.3.1	Protocol optimization for RNA isolation	83-84
3.3.2	Analyses of different steps in SSH library construction	84-85
3.3.2.1	Yield and quality checking of dsDNAs	84
3.3.2.2	Ligation efficiency analysis	84
3.3.2.3	Analysis of differentially expressed transcripts	84-85
3.3.2.4	Screening of transformed cells and isolation of plasmids	85
3.3.3	Generation of ESTs under drought stress and their clustering	85-87
3.3.4	Gene ontology analysis of drought induced unigenes	87
3.3.5	Analyses of drought associated genes and pathways	87-90
3.3.5.1	Ubiquitin-proteasome pathways	88
3.3.5.2	Glutathione biosynthesis and metabolism	89
3.3.5.3	Sugar synthesis, transport and metabolism	89-90
3.3.5.4	Genic microsatellite marker analysis	90
3.3.6	Analyses of different steps of standard full-length library construction	90-93
3.3.6.1	Generation and clustering of full-length unigenes	91
3.3.6.2	Functional annotation of full-length unigenes	92
3.3.6.3	Identification of redundant genes under normal growth and drought stress	92-93

3.3.7	Comparative analysis of unigenes under normal growth, drought and winter dormancy stress	93-95
3.4	Discussion	96-106
3.4.1	Protocol optimization and RNA isolation	96
3.4.2	SSH library construction, analyses and functional annotation of ESTs	96-104
3.4.2.1	Gene ontology analysis	98-99
3.4.2.2	Drought associated genes and genetic pathways	99-103
3.4.2.2.1	Ubiquitin-proteasome degrading pathways	101-102
3.4.2.2.2	Glutathione synthesis and metabolism pathways	102
3.4.2.2.3	Sugar synthesis, transport and metabolism	103
3.4.2.3	Genic microsatellite markers	103-104
3.4.3	Full-length library construction, analyses and functional annotation of ESTs	104-105
3.4.4	Comparative analyses of transcripts under normal growth, drought and winter dormancy stress	105-106

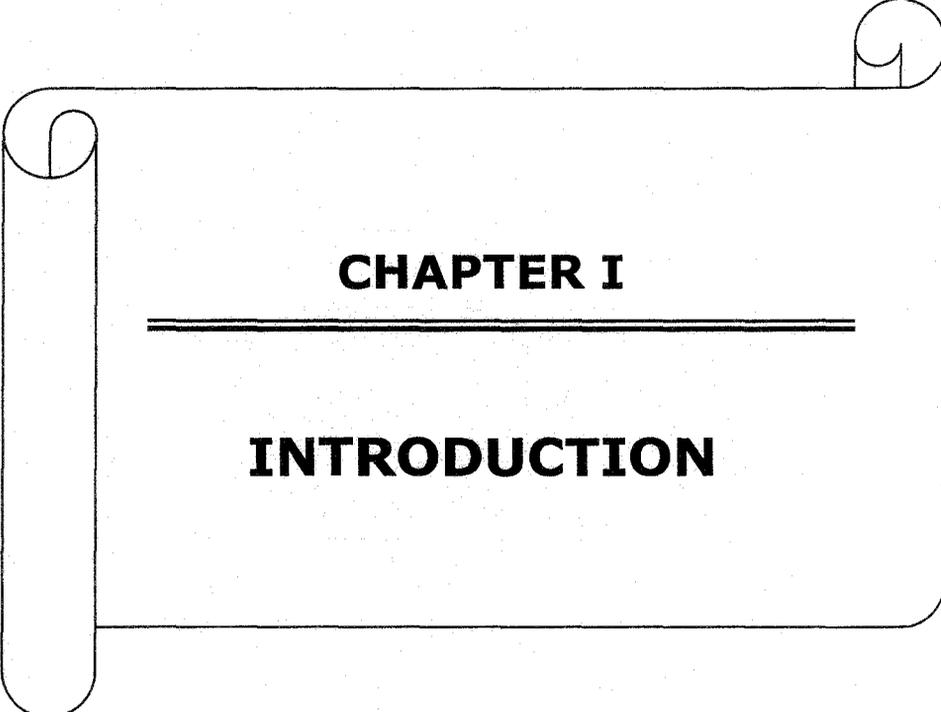
#### **CHAPTER IV COMPUTATIONAL IDENTIFICATION OF miRNAs AND THEIR TARGETS**

4.1	Review of literature	107-111
4.1.1	Biogenesis of miRNAs and their functional roles	107-108
4.1.2	Strategies for miRNA identification and characterization	109-110
4.1.3	Identification of miRNA targets	110-111
4.2	Material and methods	112-114
4.2.1	Collection of reference miRNAs, full-length nucleotides and EST sequences	112
4.2.2	Prediction of miRNAs and their precursors	112-113
4.2.3	Identification of potential miRNA targets	113
4.2.4	Evaluation of conserved nature in miRNAs and their Phylogenetic relationships	113
4.2.5	Nomenclature of miRNAs	114
4.3	Results	115-116

4.3.1	Identification of miRNAs and their precursors	115
4.3.2	Phylogenetic analysis	116
4.3.3	Identification of potential miRNA targets	116
4.4	Discussion	117-119
<b>CHAPTER V</b>	<b>GENERAL DISCUSSION</b>	120-124
<b>CHAPTER VI</b>	<b>SUMMARY</b>	125-127
<b>CHAPTER VII</b>	<b>BIBLIOGRAPHY</b>	128-178

---

Appendix-A	Chemicals, reagents and stock solutions
Appendix-B	Flowchart of SSH & SMART libraries construction techniques
Appendix-C	Standard molecular weight markers
Appendix-D	Bioinformatics analyses of ESTs
Appendix-E	List of published/communicated research papers



**CHAPTER I**

---

**INTRODUCTION**

## INTRODUCTION

### 1.1 Tea: an overview

Tea is the oldest non-alcoholic beverage crop in the world. Tender leaves of tea plants are plucked and processed for using as a drink in different types such as green, black and oolong. It has been socially and habitually consumed by people since 3000 B.C (Lin et al. 2003). Besides its high medicinal value, tea is refreshing, mildly stimulating, and produces a feeling of well-being (Miller-Hamilton 1995). Tea is basically a woody, perennial plantation crop, which is under large-scale cultivation in several countries including India. The cultivated tea plants are of two main varieties under the family Theacea: the small-leafed China tea (*Camellia sinensis*) and the large-leafed Assam tea (*Camellia assamica*). Cultivation of tea and its management has now become a most important agro-based, labor-intensive, employment-generating, export-oriented industry in most of the tea growing countries such as India, Kenya, Sri Lanka, China and Japan. Globally, India is in the first rank of tea production, consumption and exportation. Tea industry plays a pivotal role in Indian national economy with a total annual turnover of Rs.10000 crores. India occupies 1.016 million acres of tea growing land which is 16.4% of the total tea growing areas of the world. Tea industry directly employs over 1.1 million workers and generates income for another 10 million people. In spite of being the largest producer in the world, total net foreign exchange earned per annum is only around Rs. 1847 crores because of its ever increasing domestic consumption (Mondal 2007).

### 1.2 Origin and distribution

Tea is the native of south-east Asia, specifically around the intersection of latitude 29° north and longitude 98° east, the point of confluence of the lands of north-east India, north Burma, south-west China and Tibet. Tea plant was introduced to more than 52 countries, from this 'centre of origin' (Mondal 2007). Since the development of black tea from *C. assamica* by R. Bruce in 1823, tea was introduced as a commercial venture in other parts of the world too. Now, more than 45 countries

are growing tea within the latitudinal range of 45° north to 34° south. Tea was introduced to Japan from China in the early part of the 8<sup>th</sup> century, and later during 17<sup>th</sup> century it was spreaded to Indonesia from Japan. In Sri Lanka, tea was first planted in 1839 when seeds were brought from India. Tea seed was imported to USSR from China for cultivation during the end of last century. Later, from USSR, seeds were exported to Turkey in the year 1939-40. In Europe, tea was introduced in 1740 by the East India Company's captain Goff, but those plants which were planted in the Royal Botanic Garden at Kew in England could not survive (Sealy 1958). First successful introduction was achieved by a British merchant and naturalist John Ellis in 1768 (Aiton 1789, Booth 1830). Then, it was spreaded to the African countries at the end of the 19<sup>th</sup> century.

### 1.3 Economic importance

Tea is an economically important plantation crop with a life span of more than 60 years. It occupies 3.07 million ha land all over the world with an average production of 3.87 million tons annually (Alkan et al. 2009). It is consumed in the form of fermented, semi-fermented or non-fermented which is known as black, oolong or green tea, respectively. In Burma and Thailand, tea leaves are consumed as vegetables which have been known as 'leppet tea' and 'meing tea', respectively. Moreover, tea has been shown to have a wide range of beneficial physiological and pharmacological effects such as strengthening capillaries, anti-inflammatory effect, anti-microbial, anti-oxidant properties and positive role in cardiovascular ailments (Miller-Hamilton 1995). Animal studies have shown that tea and tea constituents inhibit carcinogenesis of the skin, lung, oral cavity, esophagus, stomach, liver, prostate and other organs (Lambert and Yang 2003). Besides, tea seed oil is used in human consumption in certain parts of the world. Apart from being used as drink or food products, few species of this genus are used in different purposes. For example, *C. oleracea* produces oil which is used in the cosmetic industry as well as *C. japonica* and *C. reticulata* which are grown as ornamental plants for their beautiful flowers.

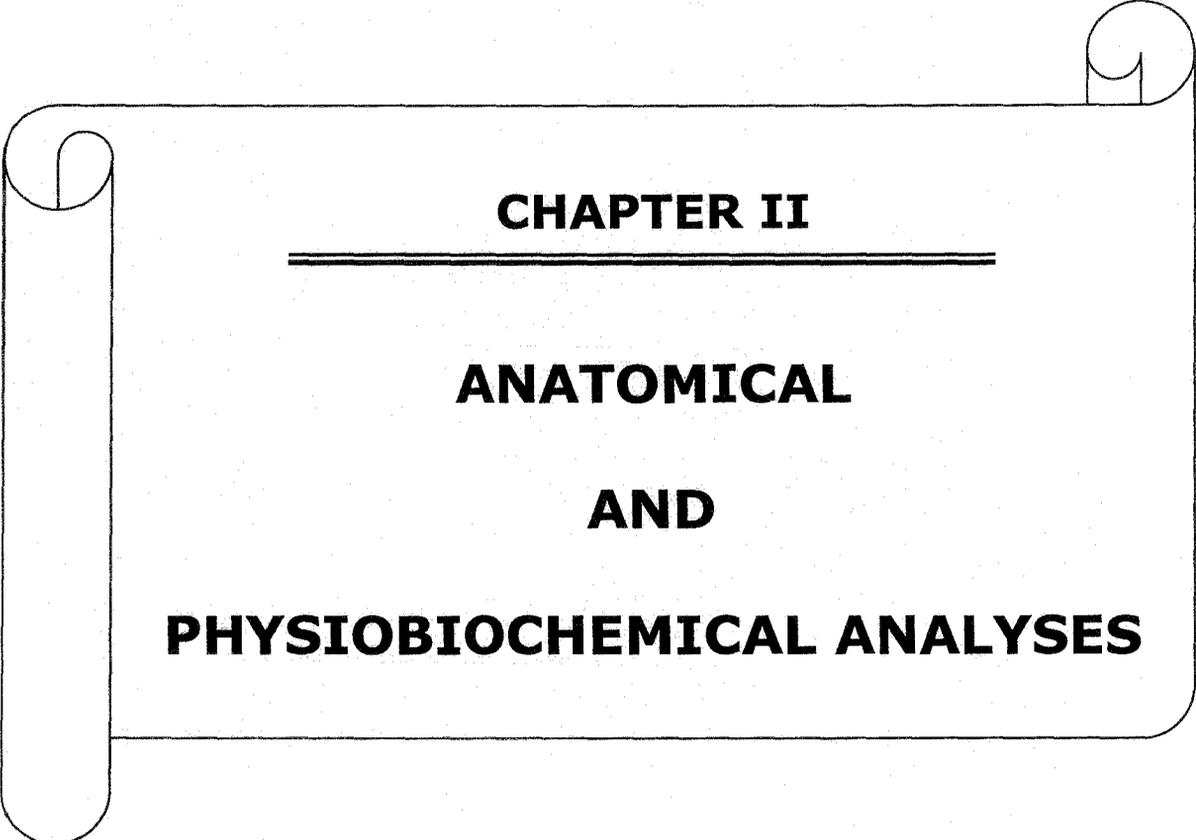
## 1.4 Cytology

Cytogenetic works in *Camellia* have a long history compared to other cultivated plants (Kondo 1977). The first report of the gametic chromosome number in *Camellia* was made from the diploid *Camellia sinensis* ( $n=15$ ) by Morinaga et al. (1929). Polyploidy in *Camellia* was also reported first in *Camellia sinensis* ( $2n=45$ ) itself (Karasawa 1932). Extensive investigations in chromosome number and ploidy level in *Camellia* began in the early 1950's. Later, it has been confirmed that the cultivated tea is mainly diploid ( $2n=30$ ) and their chromosome structures are comparable amongst the wild relatives (Bezbaruah 1971). There are few exceptions such as Indian cultivated clones Sundaram and TV-29 are triploids. Some natural triploids ( $2n=45$ ), tetraploids ( $2n=60$ ), pentaploids ( $2n=75$ ) and aneuploids ( $2n+1$  to 29) have also been sampled in tea populations of Assam, but those are reported to be present in very low numbers (Singh 1980). A natural triploid of tea has also been reported in China (Zhan et al. 1987). Deviation from normal chromosome number is also known in some wild relatives. For example, *Camellia sasanqua* is recorded as a hexaploid species, but some clones show a different ploidy level including tetraploid, pentaploid, heptaploid, octoploid and even a few aneuploids (Kondo 1975). That way, *Camellia reticulata* is also recorded as hexaploid species, however some clones are triploid and heptaploid in chromosome number (Kondo 1977).

## 1.5 Functional genomics

Tea is an important plantation crops that is valued for its rich source of secondary metabolites. Several improved tea varieties were developed through conventional breeding and propagation techniques in the last several decades. However, due to the limitations of conventional breeding coupled with the demand of increasing productivity and quality with lower cost of production, application of biotechnology becomes an alternative approach. Different biotechnological applications have already been adopted successfully in tea such as micropropagation (Dood 1994, Mondal et al. 2002), cell and organ culture (Jain and Newton 1990, Akula et al. 2000, Kuboi et al. 1991), transgenic production (Mondal et al. 2001, Konwar et al. 1998) and DNA fingerprinting (Wachira et al. 1997, Mondal 2002, Matsumoto et al. 1994, Paul et al. 1997). Functional genomics research, which focuses on the

characterization of gene function and demonstration of the spatial, temporal and cell-dependant expression and regulatory mechanisms of genes is relatively new in tea compared to other crop species and progress has also been slow. It was initiated only two decades ago by isolating a cDNA of ribulose-1, 5-biphosphate-carboxylase (Savolainen et al. 1994). Since then, many important genes have been isolated, cloned and characterized such as chalcone synthase (Takeuchi et al. 1994), PAL (Matsumoto et al. 1994), caffeine synthase (Kato et al. 2000) etc. Tea functional genomics study has been accelerated with these successes and as on 20 September, 2010, more than 700 cds are available in NCBI (<http://www.ncbi.nlm.nih.gov>). Of these, many important functional genes related to metabolism, signal transduction and antioxidant responses were cloned, characterized and studied extensively. More recently, progress has been made on EST sequencing and annotation (Park et al. 2004, Chen et al. 2005, Zhao et al. 2008b), elucidation of gene expression profiling (Singh et al. 2008, Singh et al. 2009), establishment and use of cDNA microarrays (Zhao et al. 2006), data mining for marker development such as EST-SSR and STMS (Zhao et al. 2007, Jin et al. 2007, Sharma et al. 2009, Matteo et al. 2010) as well as cloning and expression analysis of genes involved in secondary metabolism and stress defense (Zhao et al. 2006, Wang et al. 2008). In the recent years, several groups have been working on tea functional genomics all over the world, and much information would be expected in the years ahead.



**CHAPTER II**

---

**ANATOMICAL**

**AND**

**PHYSIOBIOCHEMICAL ANALYSES**

## ANATOMICAL AND PHYSIOBIOCHEMICAL ANALYSES

### 2.1 Review of literature

Plant's growth and productivity are adversely affected by nature's wrath in the form of various biotic and abiotic stress factors. Water deficit or commonly known as drought is one of the major abiotic stresses, which adversely affects growth and yield of crops. Drought can be defined as the absence of adequate moisture necessary for normal plant's growth and development to complete the life cycle. Water is a fundamentally important component of the metabolism of all living organisms, facilitating many vital biological reactions by being a solvent, a transport medium and evaporative coolant. In plants, water plays the additional role of providing the energy necessary to drive photosynthesis through autolysis by yielding electrons. Hence, drought reduces plant's growth by affecting various physiological and biochemical processes, such as photosynthesis, respiration, translocation, ion uptake, carbohydrates, nutrient metabolism and growth promoters (Jaleel et al. 2009). Plants have evolved a wide range of defense mechanisms to contend with this problem. Drought is a common occurrence in rain-fed areas, brought about by inadequate rains or poor irrigations. Tea plant is a woody perennial, and as such, encounters a large number of environmental stresses throughout its long life span. Out of all, drought is the most important abiotic stress which causes around 40% crop loss (Barua 1989, Jain 1999).

#### 2.1.1 Abiotic stress responses in plants

Plants must respond and adapt to abiotic stresses to survive in various environmental conditions. Plants have acquired various stress tolerance mechanisms, which are different processes involving physiological and biochemical changes that result in adaptive or morphological changes (Urano et al. 2010) (Fig. 2.1). Different environmental stresses to a plant may result in similar responses at the cellular and molecular level. This is due to the fact that the impacts of the stressors trigger similar strains and downstream signal transduction chains. Drought, cold and salinity are three major forms of abiotic stress that affect the water relations of a plant at

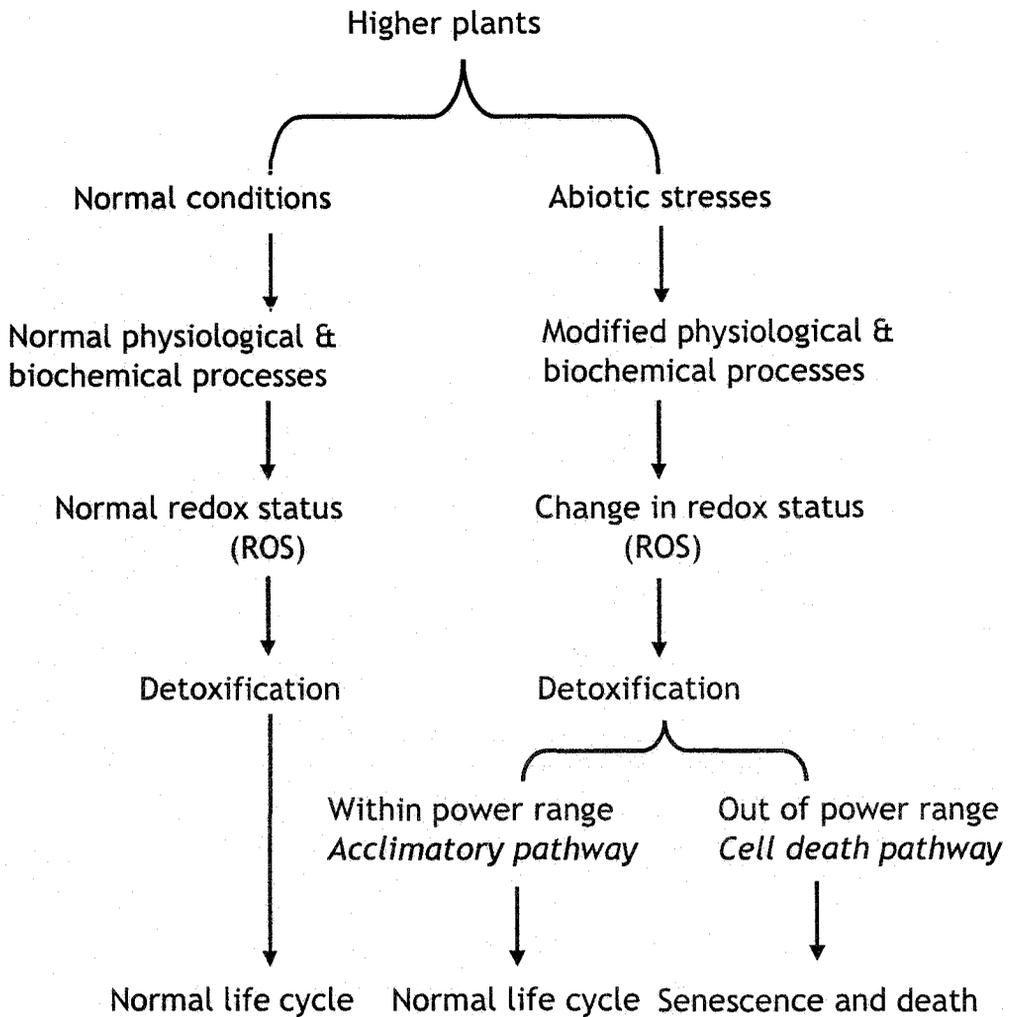


Fig. 2.1: Flowchart of main antioxidants and redox signaling processes in higher plants (modified from Hong-bo et al. 2008).

cellular and whole plant level causing specific as well as unspecific reactions, damages and adaptation reactions (Beck et al. 2007). Drought or cold stress induces various primary effects at cellular level, termed as strains (Levitt 1980), which in turn lead to uncontrolled (damage) or controlled effects (adaptation) on that level. Adaptation to stress occurs when a plant withstands the imposed stress that may arise from either tolerance or a mechanism that permits escape from the situation (Beck et al. 2007). This will have direct and indirect effects on the reduction of net photosynthesis and the overall production of ROS (Cruz de Carvalho 2008). During that period, plants maintain the internal water potential, turgor and water uptake by increasing the level of osmolytes, either by uptake of soil solutes or by synthesis of metabolic (compatible) solutes (Zhu 2002) as well as maintain the toxic oxygen species and compounds through anti-oxidative mechanisms.

### **2.1.2 Physiological responses to drought stress**

Drought is one of the most important manifestations of abiotic stress in plants. The initial reaction to stress-induced drought is to modify stomatal conductance by reducing stomatal aperture, this minimizes the loss of water by transpiration. However, stomatal closure has complex consequences, requiring adjustment at the levels of photosynthesis and respiration, altering ion, nutrient and water fluxes and change in the allocation of carbon and nitrogen (Hsiao 1973). This response operates more or less precociously at the onset of drought period, depending on the plants' inherent strategy (and acclimation) towards drought stress (Cruz de Carvalho 2008). The phytohormones play a crucial role in triggering several molecules for performing cascades of biochemical reactions in response mechanism. For instance, ABA activates the synthesis of  $H_2O_2$  in guard cells by a membrane bound NADPH oxidase and that  $H_2O_2$  mediates stomatal closure by activating plasma membrane  $Ca^{2+}$  channels (Kwak et al. 2003, Pei et al. 2000). As a consequence, drought stress reduces the photosynthesis rate, transpiration rate, stomatal conductance, water use efficiency and RWC in leaf as well as enhances the ABA content and leaf water potential (Jiang and Hartung 2007). It also affects chl content, inhibits roots growth, dry matter production and severely reduces the yield and yield components (Loggini et al. 1999). Physiological responses to drought stress have been reported in numbers of crop plants for the years. For example: Pirdashti et al. (2009) reported a

significant and positive correlation between chl content and RWC and also between these traits and grain yield in four rice cultivars. Efeoğlu et al. (2009) recorded the decrease of RWC, chl a, chl b, total chl (a + b) and carotenoids content in three maize cultivars during drought stress as well as control values after recovery from stress. Bano and Yasmeen (2010) reported that drought stress resulted in maximum decrease in IAA and GA content but maximum increase of ABA and proline in wheat cultivar.

### **2.1.3 Responses of biochemical solutes to drought stress**

Drought stress induces not only physiological but also biochemical changes that finally alter the metabolic status of the stressed plant. To counter with drought stress, many plants increase the osmotic potential of their cells by synthesizing and accumulating osmolytes (Du et al. 2004). Two hypotheses have been suggested to explain the functions of accumulating solutes (Truchet et al. 1991). In the first, accumulation leads to 'osmotic adjustment' through mass action, which results in increased water retention and/or sodium exclusion. The second hypothesis considers accumulating compounds as 'compatible solutes'. In this function they could replace water as a solute in biochemical reactions. Compatible solutes could also, or as an alternative, associate with lipids or proteins and thus prevent membrane disintegration, dissociation of protein complexes, or inactivation of enzymes. The solutes that accumulate vary with the organism and even between plant species (Bray 1997), which include elemental ions, such as  $K^+$  and majority of organic solutes (Cruz de Carvalho 2008). The major category of organic solutes consists of simple sugars (mainly fructose and glucose), sugar alcohols (glycerol and methylated inositols) and complex sugars (trehalose, raffinose and fructans) (Bohnert and Jensen 1996). Others include quaternary amino acid derivatives (proline, glycine betaine,  $\beta$ -alanine betaine, proline betaine, tertiary amines), and sulfonium compounds (choline osulfate, dimethyl sulfonium propionate) (Bray 1997).

#### **2.1.3.1 Carbohydrates**

Carbohydrates such as sugars accumulate under drought stress and play a leading role in osmoprotection, osmotic adjustment, carbon storage and radical scavenging

(Hsiao 1973). Higher accumulation of soluble carbohydrates in response to drought stress has been widely reported in various crop plants such as rice (Vajrabhaya et al. 2001), wheat (Hakimi et al. 1995, Kameli and Losel 1993), maize (Mohammadkhani and Heidari 2008a) and pea (Sánchez et al. 1998). An increase in both reducing and non-reducing sugars and a decrease in starch under drought stress have also been reported in maize (Mohammadkhani and Heidari 2008a), wheat (Nicolas et al. 1985) and coffee (Matta et al. 1997) cultivars. Trehalose, a disaccharide, also accumulates under drought conditions and protects membranes and proteins in cells exposed to stress and reduce aggregation of denatured proteins (Singer and Lindquist 1998). It has a suppressive effect on apoptotic cell death under drought (Yamada et al. 2003). Higher accumulation of trehalose in response to drought has been reported in various plant species such as rice (Garg et al. 2002) and wheat (El-Bashiti et al. 2005).

#### **2.1.3.2 Amino acids, amides and proteins**

Amino acids such as proline, alanine, leucine, valine, aspartate and amides such as glutamine and asparagine have also been reported to accumulate in response to drought in various crop plants. For example, drought treatment in maize seedlings caused an increase in total free amino acid content and a consistent rearrangement of the amino acid pool, with an accumulation of alanine, arginine, valine, aspartate, serine, threonine and tyrosine, as well a decrease in glutamic acid, glycine and methionine (Ranieri et al. 1989). Among the amino acids, proline is osmotically very active and contributes to membranes stability and mitigates the effect of stress on cell membrane disruption (Szekely 2004). Proline accumulation normally occurs in the cytosol where it contributes substantially to the cytoplasmic osmotic adjustment, and even at supra-optimal levels, it does not suppress enzyme activity (Ketchum et al. 1991). In higher plants, proline is synthesized from glutamate or ornithine. It may act as a signaling or regulatory molecule that may be able to activate multiple responses that are component of the adaptation process (Maggio et al. 2002). Proline accumulation under drought and its role in conferring tolerance have been studied extensively from model plants like *Arabidopsis* (Szekely 2004) to crop plants such as rice (Mostajeran and Rahimi-Eichi 2009), chickpea (Najaphy et al. 2010), drought tolerant genotypes of wheat (Nayyar and Walia 2003) and potato

(Schafleitner et al. 2007). Although amides generally accumulate in environmentally stressed plants to a lesser extent than do other nitrogen containing compounds (Mansour et al. 2002), concentration of asparagine and glutamine frequently increase in response to drought. For example, a substantial accumulation of asparagine and glutamine in different drought resistant cultivars of maize were registered under drought (Thakur and Rai 1982). Proteins that accumulate under drought may provide a storage form of nitrogen for re-utilizing later and may play a role in osmotic adjustment. A higher content of soluble proteins in response to drought has been reported in various crop plants such as broad bean (El-Tayeb 2006), mustard (Ashraf and Mehmood 1990) and cotton (Li et al. 2010), however there are also reports of decrease of the same in diverse crop species under drought such as maize (Mohammadkhani and Heidari 2008b) and grapevine (Maroco et al. 2002). In higher plants, numbers of proteins accumulate in response to drought as part of the tolerance mechanism such as molecular chaperons (heat shock proteins and dehydrins), aquaporins and antioxidative enzymes (SOD, APX etc.) (Shinozaki and Yamaguchi-Shinozaki 2007).

### **2.1.3.3 Quaternary ammonium compounds**

Quaternary ammonium compounds such as glycine-betaine and proline-betaine are now well-known to have osmoprotective effects under drought (Cushman 2001). QACs can accumulate to high concentration which may not only increase the osmotic pressure of the cytoplasm without perturbing metabolism but also stabilize enzymes and membranes (Yancey 1994, Kavikishore et al. 2005). The synthetic pathway to glycine-betaine, the most common QAC, therefore has been the target of recent metabolic engineering efforts to improve plant's stress tolerance (McNeil et al. 1999, Rathinasabapathi et al. 2000, Sakamoto and Murata 2000). This organic compound is mainly localized in chloroplast thylakoid membranes, thereby maintaining photosynthetic efficiency and plasma membrane integrity (Huang et al. 2000). Glycine-betaine accumulates in response to drought in many crops including rice (Sawahel 2003), wheat (Wang et al. 2010a), maize (Quan et al. 2004) and sorghum (Yang et al. 2003).

#### 2.1.3.4 Polyols

Polyols, the polyhydric alcohols, are among the compatible solutes involved in osmoregulation which are thought to play a role in plant's drought tolerance (Bohnert et al. 1995). The most common polyols in plants include acyclic forms such as mannitol, glycerol and sorbitol as well as cyclic forms such as ononitol and pinitol. Higher accumulation of different polyols in conferring drought tolerance has been reported in diverse crop plants such as mannitol in wheat (Abebe et al. 2003) and sorbitol in cherry (Ranney et al. 1991) and apple (Wang and Stutte 1992).

#### 2.1.4 Antioxidative responses to drought

Metabolism of higher plants is highly regulated in order to allow effective integration of diverse spectrum of biosynthetic pathways. Higher plants, as other aerobic organisms, require oxygen for the efficient production of energy. During the reduction of  $O_2$  to  $H_2O$ , ROS, namely superoxide radical ( $O_2^{\cdot-}$ ), hydrogen peroxide ( $H_2O_2$ ), and hydroxyl radical ( $OH^{\cdot}$ ) can be formed (Hong-bo et al. 2008). Most of the cellular compartments in higher plants have the potential to become a source of ROS. In order to cope with continuous ROS production, plants have a battery of enzymatic and non-enzymatic antioxidants, which function as an extremely efficient co-operative system (Mittler 2002). Under drought stress,  $CO_2$  fixation is reduced which consequently enhance ROS production through multiple ways and create chances of oxidative damage to cellular membranes and organelles (Cruz de Carvalho 2008). For example, the limitation on  $CO_2$  fixation reduces  $NADP^+$  regeneration through the calvin cycle, hence provoke an over reduction of the photosynthetic electron transport chain. In fact, during photosynthesis under drought stress, there is a higher leakage of electrons to  $O_2$  by the Mehler reaction (Smirnoff 1993). Plant's defense system fights against the higher production of ROS under drought by enhancing the activities of enzymatic antioxidants such as SOD, CAT, APX, POX, GR etc. and by accumulating non-enzymatic antioxidants such as glutathione, ascorbate, tocopherol, proline and betaine etc. (Mittler 2002). In addition to crucial roles in defense system and as enzyme co-factors, antioxidants influence growth and development of higher plants by modifying processes from mitosis and cell elongation to senescence and death (Hong-bo et al. 2008).

### 2.1.4.1 Non-enzymatic antioxidants

A continuous oxidative assault on plants during drought stress has led to the presence of an arsenal of non-enzymatic antioxidant defenses to counter the phenomenon of oxidative stress (Du et al. 2004). AA is an important antioxidant, which reacts not only with  $H_2O_2$ , but also with  $O_2^-$ , OH and lipid hydroperoxidases (Hong-bo et al. 2008). AA is water soluble and has an additional role in protecting or regenerating oxidized carotenoids or tocopherols. It also involved in the neutralization of secondary products of ROS (Mittler 2002). AA is one of the most extensively studied antioxidant and has been detected in the majority of plant cell types, organelles and apoplast (Loewus 1988). Glutathione is also takes part in the control of  $H_2O_2$  levels. Change in the ratio of its reduced (GSH) to oxidized (GSSG) form during the degradation of  $H_2O_2$  is important in redox signaling pathways (Mittler 2002). It has been suggested that the GSH/GSSG ratio is an indicative of the cellular redox balance which may be involved in ROS perception (Cruz de Carvalho 2008). Reduced glutathione (GSH) acts as an antioxidant and is involved directly in the reduction of most active oxygen radicals generated due to drought (Cruz de Carvalho 2008). Tocopherols, found in green parts of plants, also scavenge lipid peroxy radicals through the concerted action of other antioxidants (Hong-bo et al. 2008). Further, they are known to protect lipids and other membrane components by physically quenching and chemically reacting with  $O_2$  in chloroplasts, thus protecting the structure and function of photosystem II (Asada 2006).  $\alpha$ -tocopherol interact with the polyunsaturated acyl groups of lipids, stabilize membranes and scavenge as well as quench various ROS and lipid soluble by-products of oxidative stress (Asada 2006). Singlet oxygen quenching by tocopherols is highly efficient, and it is estimated that a single  $\alpha$ -tocopherol molecule can neutralize upto 120 singlet oxygen molecules *in vitro* before being degraded (Wu and Tang 2004). Other antioxidants such as citrulline and metallothionein are also efficient in scavenging hydroxyl radical and protecting proteins and DNA effectively from oxidative damage (Cruz de Carvalho 2008). Responses of antioxidants under drought stress have been studied in diverse crop plant species through the years, and their observations have revealed differential responses that varied among the crops (Table 2.1).

Table 2.1: Drought stress responses of major antioxidants in various crop plants

Antioxidants	Crops	Responses to drought stress	References
Ascorbic acid	Rice	decreased significantly	Sharma and Dubey 2005
	Soybean	increased highly in tolerant genotypes	Angra et al. 2010
	Wheat	decreased in high percentage	Sairam et al. 1998, Bartoli et al. 1999
Glutathione	Apple	increased under moderate drought stress	Sircelj et al. 2007
	Rice	decreased under severe drought	Sharma and Dubey 2005
Tocopherol	Apple	increased significantly under moderate drought stress	Sircelj et al. 2007
	Wheat	increased significantly	Bartoli et al. 1999, Keles and Oncel 2002
Proline	Apple	decreased under severe drought stress	Sircelj et al. 2007
	Maize	increased at different drought intensities	Mohammadkhani and Heidari 2008a
	Sorghum	increased significantly only at severe drought stress	Waldren and Teare 1974
	Soybean	increased significantly only at severe drought stress	Waldren and Teare 1974
SOD	Olive	increased in different cultivars	Ahmed et al. 2009
	Rice	increased consistently with increasing drought stress	Sharma and Dubey 2005
	Maize	increased in different cultivars	Moussa and Abdel-Aziz 2008
	Wheat	significantly increased at two different stages after anthesis	Sairam et al. 1998
	Wheat	increased or maintained at same level	Navari-Izzo et al. 1993
	Soybean	increased equally in tolerant and susceptible genotypes	Angra et al. 2010
	Oilseed Rape	increased activity in different genotypes	Abedi and Pakniyat 2010
	Olive	increased in different cultivars	Ahmed et al. 2009
	Coffee	increased in both tolerant and susceptible genotypes	Pinheiro et al. 2004

(Table 2.1 continued..)

APX	Rice	increased consistently with increasing drought stress	Sharma and Dubey 2005
	Wheat	significantly increased at two different stages after anthesis	Sairam et al. 1998
	Soybean	increased highly in tolerant genotypes	Angra et al. 2010
	Olive	increased in different cultivars	Ahmed et al. 2009
	Coffee	increased in both tolerant and susceptible genotypes	Pinheiro et al. 2004
GR	Rice	increased under different drought stress level	Sharma and Dubey 2005
	Wheat	decreased under severe drought	Selote and Khanna-Chopra 2010
	Coffee	increased in both tolerant and susceptible genotypes	Pinheiro et al. 2004
CAT	Rice	decreased consistently with increasing drought stress	Sharma and Dubey 2005
	Wheat	significantly increased at two different stages after anthesis	Sairam et al. 1998
	Wheat	increased or maintained at same level	Navari-Izzo et al. 1993
	Maize	increased in different cultivars	Moussa and Abdel-Aziz 2008
	Soybean	increased highly in tolerant genotypes	Angra et al. 2010
	Oilseed Rape	increased activity in different genotypes	Abedi and Pakniyat 2010
	Coffee	increased in both tolerant and susceptible genotypes	Pinheiro et al. 2004
	Olive	increased in different cultivars	Ahmed et al. 2009

---

#### 2.1.4.2 Enzymatic antioxidants

Plants defense system for controlling ROS under drought stress includes several antioxidative enzymes that actively and efficiently participate in the process. The major enzymatic scavenging mechanisms include SOD, enzymes and metabolites from the ascorbate-glutathione cycle and CAT (Bowler et al. 1992, Noctor and Foyer 1998, Willekens et al. 1997). They are located throughout the different compartments of plant cells, with the exception of CAT that is exclusively located in peroxisomes. SOD is the front-line enzyme in ROS attack since it rapidly scavenges superoxide, one of the first ROS to be produced, dismutating it to oxygen and  $H_2O_2$  (Bowler et al. 1992). However, this reaction only converts one ROS to another, and  $H_2O_2$  also needs to be destroyed since it promptly attacks thiol proteins. The major enzymatic cellular scavengers of  $H_2O_2$  are CAT and APX (Noctor and Foyer 1998, Willekens et al. 1997). They have, however, different affinities for this ROS and seem to have different cellular roles in  $H_2O_2$  scavenging. In fact, CAT does not need a reductant to scavenge  $H_2O_2$  making it reducing power-free, whereas APX needs a reductant, ascorbate. The GR is the last enzyme of the ascorbate/glutathione cycle which has a major role in maintaining the intracellular glutathione pool in the reduced state (GSH). GSH can function as an antioxidant either directly or indirectly as a reducing agent that recycles ascorbic acid form to its reduced form by the enzyme dehydroascorbate reductase (Loewus 1988).

Measuring of specific antioxidant enzyme activities and/or expression analysis during water stress treatments has been generally accepted as an approach to assess the involvement of the scavenging system during drought stress (Cruz de Carvalho 2008). However, contradictory results have been gathered through the years (Table 2.1). These differences might be related not only to the plants' age and their tolerance level or strategy towards water stress treatment, but also to the duration and the intensity of the stress treatment. Nevertheless, there are reports of a direct correlation between the level of the induction of antioxidants and the degree of drought tolerance in numbers of crop plants of same genus or species such as maize (Pastori and Trippi 1992), wheat (Loggini et al. 1999), rice (Guo et al. 2006) and coffee (Lima et al. 2002).

## **2.1.5 Brief report on works done in tea plants**

Tea plant being perennial in nature can grow under diverse climatic conditions and is always subjected to various environmental stresses including drought. The effects of drought in tea plantations are well-known which severely impairs growth, crop yield and various morphological, anatomical, physiological and biochemical processes. The changes basically take part as a mechanism of drought adaptation and/or tolerance.

### **2.1.5.1 Growth and physiology**

High temperature and dry weather are unfavorable for the growth of tea plants. Increases in temperature, soil moisture deficit, and vapor pressure deficit create a plant water deficit, which lead to growth retardation (Wijeratne 1992). Young tea plants are more vulnerable to drought stress than the mature ones due to their inefficiency of stress tolerance (Nixon et al. 2001) which show the early symptom of shoot growth retardation (Cheruiyot et al. 2007). It was reported that the clonal tea growth and yields could adversely be affected at temperatures  $>26^{\circ}\text{C}$ , soil moisture deficits  $>30\text{-}50$  mm, and saturation vapor pressure deficits  $>1.2$  kPa (Wijeratne 1994). Burgess and Carr (1996) studied the yield response to drought and temperature of six contrasting tea clones, and concluded that annual yields decreased curvi-linearly as the maximum soil water deficit increased. A reduction of yield about 2.9 kg per ha for each mm increase in the potential soil water deficit was reported by Stephens and Carr (1991).

Dynamic states of water potential of the shoot, synthetic function and metabolic function of tea plant under the condition of different soil water content including drought were studied by Yang et al. (1987). The results showed that leaf water potential dropped under drought stress and slowed down the growth of the leaf. The photosynthesis and synthetic metabolism were also decreased; however, the electrical conductivity of the shoot was increased. Qian (1999) also reported that the net photosynthetic rate, transpiration rate and stomatal conductance under high, moderate and low soil water content were gradually decreased but intercellular  $\text{CO}_2$  concentration was increased with the time of drought stress.

Declining of transpiration rate and stomatal conductance alleviated cellular injury in tea plants under drought which enhances the significant accumulation of phytohormones such as ABA. Dropping of net photosynthesis rate ultimately leads to retardation of the growth of the plants (Xuelan 1987). Affects of drought stress on stomatal conductance and transpiration of tea leaves are well-documented in literature (Fordham 1971, Gee et al. 1982, Saikia and Dey 1984). Clonal differences in the stomatal conductance and diffusion resistance under drought stress have also been reported (Squire 1978, Sandanam et al. 1980). Therefore, drought stress tolerance level is varied on different clones. Ngetich and Bore (1998) found that Sc 31/37 and 303/577 clones tolerated drought stress better than the other clones planted at the same time under a soil water deficit of 391 mm in rehabilitated and replanted fields.

Plants maintain high RWC in leaves as a mechanism of water stress adaptation (Farooqui et al. 2000). Wijeratne et al. (1998) found that drought stress reduced RWC and water potential, and increased diffusive resistance in leaves. In their study, the critical leaf water potential increasing diffusive resistance and reducing transpiration of drought tolerant clone (TRI 2025) was comparatively higher than that of drought susceptible clone (TRI 2023). It was found that the clones having efficient stomatal control for reducing water loss and osmotic adjustments for absorbing water from drier soils can withstand drought.

Maintenance of RWC under drought stress was also reported in TV-1, TV-20, TV-29 and TV-30 cultivars (Upadhyaya et al. 2008). Maximum decrease in RWC was observed in TV-30 (53.07%) cultivar after 20 d of stress imposition, whereas TV-1 (42.03%) showed less decrease. After rehydration, plants were recovered RWC and maintained highest content in TV-1 (91.22%). Decrease in fresh mass was also found highest in TV-1 (51.85%) whereas TV-20 (20.01%) showed least decrease over well-watered plants after 20 d of stress imposition. Declining of RWC in TV-18, TV-26, UPASI-3, UPASI-26, T-78 and HV-39 cultivars under drought were reported by Chakraborty et al. (2002), however, drought tolerant cultivars UPASI-3 and UPASI-26 maintained higher RWC in comparison to the other clones.

It is well-known that differences in rooting depth are one of the factors responsible for the variability of drought resistance in a number of plants (Begg and Turner 1976). In tea, out of three parameters such as rooting depth, roots weight and vertical distribution of roots in the soil, a positive co-relation of rooting depth with drought stress tolerance was reported (Nagarajah and Ratnasuriya 1981). In shallow rooted clones such as NL 4/2, 2026, 2024 and 2023, drought resistance was increased with the increase of rooting depth; however, in deep rooted clones such as 2025, 2027, CY 9 and DN, drought resistance was not associated to rooting depth.

### **2.1.5.2 Biochemical changes**

There is a co-relationship between the drought injury of tea plants and the damage that is caused by ROS. A positive correlation between MDA content and leaked electrolyte level ( $r=0.9938$ ,  $p=0.01$ ) was reported by Lu (1992). Increased activities of SOD and CAT were also recorded at the initial stage of drought which was decreased sharply with increasing drought intensity, but POD activity was increased continuously during drought stress (Lu 1992). Yang et al. (1987) reported that drought stress decrease the activities of shoot POD and POX. Chakraborty et al. (2001) studied the biochemical responses of drought stress in TV-22, TV-23, TV-25, TV-26, TV-27 and TV-17 cultivars by withholding water for 7 and 14 d. Both total phenol and PAL activity were increased moderately in those 4 varieties following 7 d of stress imposition, while longer period of drought resulted in a sharp decline of both phenolics and PAL activity. The POX activity was increased following drought while POD activity registered a marked decline. Protein content was not significantly affected by initial stage of drought; however it was decreased significantly in extended drought period. Higher activities of PAL, POD and POX and phenol content were also reported in TV-18, TV-26, UPASI-3, UPASI-26, T-78 and HV-39 cultivars at the initial stage of drought which were decreased under severe drought condition (Chakraborty et al. 2002). Jeyaramraja et al. (2003) reported a reduction of PAL activity in drought tolerant 'Assam' cultivar UPASI-2, followed by UPASI-8 and UPASI-9. Lower PAL activity correlated well with lower synthesis of flavanols such as EGCG and ECG. Hernández et al. (2006) reported that ECQ and EGCGQ, the oxidation products of EC and EGCG, increased up to 100 and 30 fold, respectively, in tea plants exposed to 19 d of water deficit. The formation of ECQ and EGCGQ were

found negatively correlated with the extent of lipid peroxidation in leaves supporting the protective roles of these compounds under drought stress. Accumulation of proline as a protective osmolyte under drought stress in tea plants is well-documented in literature (Qian 1999, Chakraborty et al. 2001, Chakraborty et al. 2002, Upadhyaya and Panda 2004). Upadhyaya and Panda (2004) reported the decreased activities of guaiacol peroxidase and GR as well as contents of chl, carotenoids, ascorbate and glutathione after 5 d of drought imposition. Simultaneously, increased  $H_2O_2$  and superoxide anion content and lipid peroxidation level as well as CAT and SOD activities were achieved. In another attempt, Upadhyaya et al. (2008) revealed the effect of drought stress by studying the ROS metabolism in TV-1, TV-20, TV-29 and TV-30 cultivars after 20 d of drought stress imposition followed by rehydration. The results showed non-enzymatic antioxidants like ascorbate and glutathione content were decreased with differential responses of antioxidative enzymes. Increased level of lipid peroxidation,  $O_2^-$  and  $H_2O_2$  content were also recorded. The antioxidative efficiency and biochemical tolerance of the selected clones in response to drought stress were observed as TV-30>TV-1>TV-29>TV-20.

Cheruiyot et al. (2007) conducted an experiment to determine the association of tea polyphenols with water stress and their suitability as indicators for drought tolerance. The experiment was conducted in a 'rain-out' shelter, and consisted of six tea clones (BBK 35, TRFK 6/8, TRFK 76/1, TRFK 395/2, TRFK 31/30, and TRFK 311/287) at four different soil water levels (38, 30, 22 and 14% v/v), which were maintained for a period of 12 weeks. The results indicated that declining of soil water content reduced the content of polyphenols. Drought tolerant clones maintained a higher level of polyphenol content under stress, and also showed less fluctuation in total phenolics content. There was a significant ( $P<0.001$ ) correlation of total polyphenol content with shoot growth and water stress of tea, and a linear relationship between soil water content and both water stress index and shoot polyphenol content. This study showed a potential use of polyphenol as an indicator for selection of drought tolerant cultivars. In another study, Cheruiyot et al. (2008) suggested specifically the potential use of tea polyphenol, EC and EGC as indicators in predicting drought tolerance. The study consisted of six tea clones (BBK 35, TRFK 6/8, TRFK 76/1, TRFK 395/2, TRFK 31/30, and TRFK 311/287) and four levels of soil

water content (38, 30, 22, and 14% v/v). Shoot contents of EC and EGC in the clones showed varied responses, with a distinct pattern in the water-stress tolerant clones (TRFK 6/8 and TRFK 31/30). Total catechins showed significant correlation with shoot growth ( $r=0.65$ ,  $P=0.006$ ), soil water content ( $r=0.54$ ,  $P=0.0066$ ), and water stress index ( $r=0.67$ ,  $P=0.0004$ ).

Even though some works on physiological and biochemical responses to drought stress of tea plants have been done so far, these are primarily descriptive in nature. More information is yet to be gathered at physiological, biochemical and molecular level which is essential for better understanding of the mechanism of drought stress responses and which would provide a basis for comprehending the underlying process of their tolerance.

31 JAN 2013

245649



## 2.2 Materials and methods

Plants response to stress is often manifested by its physiological and biochemical reactions, which can provide a basis for better understanding the stress tolerance mechanism as well as screening and selection of individual varieties and germplasms resistant to stress. In this present investigation, drought-tolerant and -susceptible tea cultivars were selected for studying their variability in drought stress responses. The study was undertaken in both leaves and roots for understanding the responses and effects of drought stress at anatomical, physiological and biochemical level.

### 2.2.1 Plant materials, stress induction and sample collection

Two year-old vegetatively propagated well-rooted tea seedlings (~36 inch height) of S.3/A3 (drought-susceptible) and TV-23 cultivars (drought-tolerant) (Konwar 2004) were planted in earthen pots (12 inch dia) under controlled greenhouse conditions at a light intensity of  $300 \mu\text{mol m}^{-2} \text{s}^{-1}$  and  $25 \pm 2^\circ\text{C}$  with relative humidity of 65-70%. A total of 24 seedlings of each cultivar were collected from Nagrakata Experimental Station, Tea Research Association, Jalpaiguri, West Bengal, India. Drought sensitivities of the cultivars were confirmed by breeders and commercially established on the basis of yield stability in drought years. Initially, all the seedlings were watered regularly for 2 months to establish new growth, and subsequently withheld water in half numbers of seedlings in each cultivar. The period of severe drought stress was determined on the basis of soil moisture content and physiological status of the plants. Soil moisture content and physiological status of the plants were evaluated in every 2 d interval from the d of withholding water for confirmation of drought stress induction. On 21<sup>st</sup> d of stress induction, at average 7% of soil moisture content ( $-1.2 \text{ Mpa} \pm 0.20$ ),  $8.73 \mu\text{mol m}^{-2} \text{s}^{-1}$  of photosynthesis rate and  $0.42 \text{ mmol m}^{-2} \text{s}^{-1}$  of stomatal conductance, roots and leaves were collected for various experiments. The 3<sup>rd</sup> mature leaves from top of the shoots and young white fibrous roots (100 mg) of control and drought induced plants were taken and subsequently performed various experiments.

### **2.2.2 Determination of soil moisture content, photosynthesis rate, transpiration rate and stomatal conductance**

Soil moisture content was measured following a gravimetric method according to Black (1965). Photosynthesis (Pn) and transpiration rate (E) as well as stomatal conductance (SC) were recorded for all the treatments independently using an intelligent portable photosynthesis system (LCpro+, ADC, UK) between 9:00 to 10:00 h.

### **2.2.3 Determination of relative water content**

Five mature leaves were considered for estimating RWC. Fresh leaves were weighed (FW) and left saturated in water for three hours followed by measuring their turgid weights (TW). The samples were then dried in an oven at 80°C for 24 h and weighed again (DW). RWC (%) was determined according to Barr and Weatherley (1992) as follows:  $RWC (\%) = (FW - DW) / (TW - DW) \times 100$ .

### **2.2.4 Anatomical studies using light and transmission electron microscopy**

In a potato block, semi-thin sections of tissues such as leaves, petioles and roots from drought stressed and well-watered (control) plants were made by sharp blades for light microscopic observations. Sections were stained with 0.1% of safranin solution (v/v) for 5 min and visualized at 20X resolution under a compound microscope (Leica, DM1000). An impression of stomata was also taken on a piece of transparent cello tap from the abaxial side of leaves with the help of nitrocellulose solution, and subsequently visualized under light microscope. For the study of ultra-structures, 1-2 mm sections of fresh leaves and roots were made after washing with 50 mM potassium phosphate buffer (pH 7). Samples were fixed with 2.5% of glutaraldehyde solution prepared in 50 mM potassium phosphate buffer (pH 6.9) for overnight followed by 3 times of washing for 15 min each with 100 mM sodium cacodylate buffer (pH 6.9) (Sandalio et al. 2001). Samples were then dehydrated with gradual changes of acetone concentration as 30%, 50%, 70%, 80%, 90%, 95% and 100% each of 15 min duration with 2 changes. It was followed by drying with liquid carbon dioxide at its critical point i.e 31.5°C at 1100 psi. Dried specimens were

mounted on brass and coated with gold at 35 nm thickness. Thereafter, clearing was carried out using propylene oxide at room temperature for 15 min with two consecutive changes. The specimens were then infiltrated with propylene oxide and embedding medium in 3:1 ratio for overnight and afterwards transferred into an embedding moulds or BEEM capsules in proper orientation. Pure embedding medium was then poured into the mould or capsule and transferred to an embedding oven at 50°C for overnight. Oven temperature was raised later on to 60°C and kept for another 24 h for polymerization. Thick sections of specimens were made (0.5-1.0 µm) and stained with toluidine blue for 2-3 min. Sections were observed under microscope for precise location of making ultra-thin sections at 60-90 nm thick (silver-yellow color). The ultrathin sections were collected onto grids and dried for overnight. The grids were stained with uranyl acetate for 15 min and lead citrate for 5 min and observed under TEM (JEM 100C x II, Jeol) at an accelerating voltage of 80 kV.

### **2.2.5 Estimation of TSS, RS, starch, pigments, phenols, proline, ascorbic acids and ABA contents**

Total soluble sugar was estimated using anthrone reagent (Devi 2007). For estimation of TSS, the samples were hydrolyzed by keeping in a boiling water bath for 3 h with 5 ml of 2.5 N HCl and cooled at room temperature. The reaction was neutralized with solid sodium carbonate and made up the final vol to 10 ml. After centrifugation at 10000 rpm for 10 min, the supernatant was collected and diluted to 10 times. For analysis, 200 and 600 µl aliquots of the diluted sample were taken and made the final vol to 1 ml with distilled water. In 1 ml vol of sample, 4 ml of anthrone reagent was added and heated in a boiling water bath for 8 min followed by rapid cooling. It produces a green colour solution and absorbance was read at 630 nm using a UV-VIS spectrophotometer (lamda 25, perkinelmer). The amount of soluble sugar was calculated comparing with a standard curve of glucose by the following formula and expressed as mg in per g FW of sample.

$$\text{Amount of TSS (\% mg)} = \left[ \frac{\text{Sugar value from graph (mg)}}{\text{sample used (ml)}} \right] \times \left[ \frac{\text{total vol of extract (ml)}}{\text{weight of sample (mg)}} \right] \times 100$$

Estimation of reducing sugar (RS) was done by the method of Miller (1972). Five hundred mg of sample was treated with 10.0 ml of 80% ethyl alcohol. In 3.0 ml of alcoholic extract, 3.0 ml of DNS reagent was added. The mixture was heated for 5 min in a boiling water bath. After the colour had developed 1.0 ml of 40% rochelle salt was added when the contents of the tubes were still warm. The tubes were cooled under running tap water. Absorbance was recorded using spectrophotometer at 515 nm. Amount of reducing sugar was calculated using standard curve prepared from glucose. The quantity of reducing sugar was expressed as mg in per g FW of sample.

Starch content was measured as liberated glucose using anthrone reagent following hydrolysis of the extracted powders with perchloric acid (MacRae et al. 1974). Samples were ground with 1 ml of hot 80% ethanol and centrifuged at 14000 rpm for 15 min at room temperature. Pellet was taken and washed several times with hot 80% ethanol till become colourless. It was centrifuged again, collected the pellet and dried with speed vacuum. Pellet was dissolved in 1 ml of distilled water and 1.3 ml of 52% perchloric acid (w/v) in a cool condition by incubating for 20 min. Then, it was centrifuged at 14000 rpm for 15 min at room temperature and collected the supernatant finally to adjust 25 ml of vol with 52% perchloric acid. Further, 100  $\mu$ l of diluted sample was taken and adjusted to 1 ml with distilled water followed by addition of 4 ml of anthrone reagent. The solution was heated for 8 min and cooled rapidly in ice. The absorbance was read at 630 nm, and the amount of starch was calculated from a standard curve of glucose and expressed as mg in per g FW of sample.

Pigments such as chl a, chl b and carotenoids were assayed according to the protocol of Lichtenthaler (1987) with minor modifications. Samples were grounded with 2 ml of 80% acetone and centrifuged at 5000 rpm for 5 min at room temperature. Supernatant was collected and the residual amount was grounded again with 2 ml of 80% acetone. Volume of the collected supernatant was finally adjusted to 10 ml with 80% acetone. Absorbance was read at 470, 645 and 663 nm. The amount of chl a, chl b and carotenoids were estimated by the following formula and expressed as  $\mu$ g in per g FW of sample.

$$\text{Chl a } (\mu\text{g/ml}) = (12.21 \times A_{663}) - (2.81 \times A_{646})$$

$$\text{Chl b } (\mu\text{g/ml}) = (20.13 \times A_{646}) - (5.03 \times A_{663})$$

$$\text{Total carotenoids } (\mu\text{g/ml}) = (1000 \times A_{470}) - (3.27 \times \text{chl a}) - (104 \times \text{chl b})/198$$

For converting  $\mu\text{g/ml}$  into  $\mu\text{g}$  per g of FW =  $\mu\text{g/ml} \times \text{final vol made}/\text{weight of tissue taken in g}$

Total phenol was estimated using FCR reagent (Malick and Singh 1980). Samples were grounded in liquid nitrogen and transferred the powder to 1 ml of 80% ethanol. Homogenates were centrifuged at 10000 rpm for 20 min at 4°C. Supernatant was discarded and re-extracted the residue with 80% ethanol in five times of the initial vol, followed by centrifugation and collected the pellet. The residual supernatant was evaporated using a rotary evaporator at 25°C and finally dissolved the pellet in 1 ml of distilled water. The dissolved sample was taken in 0.3 ml and 0.7 ml of aliquots and adjusted the vol to 3 ml with distilled water followed by addition of 0.5 ml of FCR reagent (diluted with distilled water in 1:1 ratio). After incubation for 3 min, 2 ml of 20% sodium carbonate (w/v) solution was added and mixed thoroughly. The reaction was performed exactly for one min by incubating the samples in a boiling water bath. The solution was cooled down and read the absorbance at 650 nm. Total phenol was estimated comparing the absorbance value with known catechol amount in a standard curve and expressed as mg phenol in per g FW of sample.

Proline content was estimated using acid ninhydrin (Bates et al. 1973). Sample homogenates were prepared in 2 ml of 3% sulphosalicylic acid (w/v) and centrifuged at 14000 rpm for 15 min at 4°C. The supernatant was collected and used as sample. In 2 ml of sample, glacial acetic acid and acid ninhydrin were added in 2 ml each. The solution was heated in a boiling water bath for 1 h followed by incubation in ice for few min. Finally, 4 ml of toluene was added and mixed by vortexing. Upper pink layer was collected and absorbance was taken at 520 nm. Proline content was estimated from a standard curve of known proline and expressed as  $\mu\text{mole}$  in per g FW of sample.

Amount of AA was measured following the modified protocol of Oser (1979). Samples were ground in liquid nitrogen and extracted with 5 ml of 6% TCA. The suspension was centrifuged at 7000 rpm at 4°C for 15 min and collected the supernatant. In 4 ml of supernatant, 2 ml of 2% DNPH was added followed by mixing of 1 drop of thiourea. The solution was heated for 15 min in a boiling water bath. After cooling at room temperature, 5 ml 80% sulphuric acid was added keeping the tubes in an ice bath. Absorbance was read at 530 nm and estimated the content of AA using a standard curve of pure AA.

For ABA determination, tissues were homogenized separately in 2 ml of chilled 80% methanol containing BHT (100 mg per litre). Homogenates were stored for 24 h in dark at 4°C followed by centrifugation at 10000 rpm for 10 min at 4°C. The supernatant was filter sterilized and dried at 35°C *in vacuo* in a rotary shaker. Pellet was dissolved in 4 ml of 0.1 M potassium phosphate buffer (pH 8.0), the pH of which was adjusted at 2.5 and extracted for 3 times in an equal vol of diethyl ether. Subsequently, the ether phase was collected, evaporated *in vacuo* and dissolved in 1 ml of methanol. It was filtered through a 0.45 µm filter (Millipore) and finally 20 µl of filtered sample was injected into a RP 18 (10 µm) column protected by a guard column. Elution was carried out with methanol/water (8/2, v/v) with 1% acetic acid at a flow rate of 1 ml per min. Solvents were also filtered through a 0.45 µm filter (Millipore). The column was monitored by a UV detector at 254 nm in a HPLC system (Waters, US) and ABA was measured by referring to an authentic (±) standard ABA (Sigma) (Nagar and Kumar 1996).

### **2.2.6 Determination of ROS, lipid peroxidation, hydrogen peroxide and membrane stability levels**

Estimation of superoxide anion was done by monitoring the nitrate formation from hydroxylamine (Jordan and DeVay 1990). Samples were chopped and transferred to 3 ml of reaction mixture containing 50 mM potassium phosphate buffer (pH 7.8) and 2.5 mM NBT (dissolved in 70% N, N-dimethyl formamide). Samples were vacuum infiltrated and boiled, each step of 30 min duration and followed by centrifugation at 15000 rpm for 15 min at 25°C. The supernatant was taken and adjusted the final vol to 3 ml by N, N dimethyl formamide. Absorbance was read at 580 nm.

Level of lipid peroxidation was measured in terms of malondialdehyde (MDA) according to the method of Kazuhiro et al. (1996) with some modifications. Samples were homogenized by the addition of 2 ml of 0.1% TCA solution in ice cold condition. Homogenates were centrifuged at 10000 rpm for 20 min at 4 °C. In 1.5 ml of supernatant, 1.5 ml of 0.5% TBA (w/v) were added and incubated at 95°C for 30 min followed by rapid cooling in an ice bath. Solution was centrifuged again at 3000 rpm for 5 min and collected the supernatant. Absorbance of the supernatant was read at 532 and 600 nm. The MDA concentration was determined from the difference of absorbance ( $A_{532} - A_{600}$ ) by dividing with molar extinction coefficient (155/mM/cm) and expressed as  $\mu\text{mol}$  in per g FW of sample.

Hydrogen peroxide was estimated following the protocol of Bernt (1974) with some minor modifications. Samples were made powder with liquid nitrogen and homogenized with 2 ml of cold 5% TCA (w/v). Homogenates were centrifuged at 15000 rpm at 4°C for 10 min. In 1.6 ml of supernatant, 0.4 ml of 50% TCA (w/v) and 0.4 ml of 10 mM ferrous ammonium sulphate were added. Solution was mixed well and added 0.2 ml of 2.5 M potassium thiocyanate for colour development. The absorbance was read at 436 nm and estimated the hydrogen peroxide content by the following formula. Molecular extinction co-efficient of  $\text{H}_2\text{O}_2$  at 436 nm is used as 39.4/mM/cm for calculation and subsequently expressed as nmol in per g FW of sample.

$$\text{H}_2\text{O}_2 \text{ (nmol/g FW)} = (A_{436} \times \text{vol of assay solution} \times 100) / (39.4 \times \text{incubation time})$$

Degree of membrane integrity was assessed by the percent of electrolyte leakage. Samples were immersed in 30 ml of distilled water and incubated at 25°C for 2 h in a water bath. Suspension medium was measured for the initial electrical conductivity (EC1). Samples were then autoclaved for 10 min to release all the electrolytes, cooled and the final EC2 was measured at 15 mV using an electrometer (keithley). Percent leakage of electrolytes was calculated using the formula as:  $(\text{EC1}/\text{EC2}) \times 100\%$  (Bernt 1974).

## 2.2.7 Extractions and assays of antioxidative enzymes

For measuring the activity of different antioxidant enzymes, the tissue was ground with liquid nitrogen and suspended in 1.0 ml buffer, containing 50 mM Tris-HCl buffer (pH 7.8), fortified with 1% PVP (w/v). Homogenates were centrifuged at 12000 rpm for 20 min at 4°C and collected the supernatant. Supernatant was used for assaying SOD, POX and CAT enzymes. However, for APX, the extraction buffer was consisting of 50 mM phosphate buffer (pH 7.0), 0.5 mM EDTA, 1 mM AA and 1% of PVP. The SOD was assayed by measuring its ability to inhibit the photochemical reduction of nitroblue tetrazolium (NBT) (Upadhyaya et al. 2008). Assay buffer in a 3 ml final vol containing 50 mM potassium phosphate buffer (pH 7.8), 10 mM methionine, 2 µM riboflavin, 0.1 mM EDTA and 75 µM NBT was used for measuring the activity by 10 µl of crude enzyme extract. Reaction was started by exposing the solution under the light of 400 W bulbs for 15 min. Assay buffer without enzyme and without enzyme as well as NBT was used as reference control and blank, respectively. Absorbance was read immediately at 560 nm. Fifty percent inhibition of the reaction between riboflavin and NBT in the presence of methionine was considered as 1 unit of SOD activity.

Ascorbate peroxidase assay was followed according to Nakano and Asada (1981) with some modifications using hydrogen peroxide as substrate. By 3 ml assay solution consisting of 50 mM potassium phosphate buffer (pH 7.0), 0.5 mM AA, 0.1 mM EDTA and 1 mM freshly diluted hydrogen peroxide, reaction was started adding 20 µl of crude enzyme extract. Decrease in absorbency due to AA was recorded for 3 min at 290 nm. The APX activity was calculated as concentration of AA oxidized (initial reading-final reading = quantity of AA oxidised) by the following formula and expressed as µmole in per min of per mg FW. Extinction co-efficient of ascorbate was used as 2.8 per mM of per cm at 290 nm.

$$(A_{290} \times \text{vol of assay buffer} \times 1000) / (2.8 \times \text{incubation time} \times \text{sample used in mg})$$

Catalase activity was also assayed using hydrogen peroxide as substrate following the modified protocol of Chance and Maehly (1954). Reaction was started by adding 50 µl of enzyme extract in an assay buffer of 3 ml vol containing 50 mM potassium

phosphate buffer (pH 7.0) and 200 mM H<sub>2</sub>O<sub>2</sub>. Decrease in absorbency was recorded at 240 nm for 5 min duration. Decrease of hydrogen peroxide concentration (initial reading-final reading) was finally used for calculating the CAT activity by the following formula and expressed as  $\mu$ mole per min per mg of FW. Extinction coefficient of H<sub>2</sub>O<sub>2</sub> was used as 40 per mM of per cm at 240 nm.

$$(A_{240} \times \text{vol of assay buffer} \times 1000) / (40 \times \text{incubation time} \times \text{sample used in mg})$$

Activity of POX was estimated using pyrogallol as a substrate following the modified protocol of Chance and Maehly (1955). Assay was performed in 3 ml buffer containing 50 mM potassium phosphate buffer (pH 7), 1 mM pyrogallol and 10 mM H<sub>2</sub>O<sub>2</sub> with 50  $\mu$ l of crude enzyme extract by incubating at 25<sup>0</sup>C for 5 min. Absorbance of purpurogallin formed in the reaction was read immediately at 430 nm. Extinction co-efficient of purpurogallin is used as 2.47 per mM of per cm at 430 nm for the calculation POX activity and expressed as  $\mu$ mole in per min of mg FW as follows:

$$(A_{430} \times \text{vol of assay buffer} \times 1000) / (2.47 \times \text{incubation time} \times \text{sample used in mg})$$

### 2.2.8 Statistical analyses

Experimental design was a completely randomized block with 12 replicates of each treatment. All quantitative data were subjected to one way analysis of variance for each parameter. The mean differences were evaluated for Least Significance Difference at probability level 0.05 through INDOSTAT statistical package (Indostat Services, Hyderabad). Data were expressed as the means  $\pm$  standard error for three independent experiments in each analysis.

## 2.3 Results

Tea plants exhibited clear and obvious changes in anatomical and physiobiochemical states under drought stress. The changes were occurred in responses to plant's inherent capability of stress adaptation and/or tolerance.

### 2.3.1 Stress induction and plants growth

Plants subjected to drought stress were observed clear wilting symptoms on 21<sup>st</sup> d onwards of withholding water at which soil moisture content was recorded as 7% (Fig. 2.2, Table 2.2). Rate of photosynthesis, transpiration and stomatal conductance were reduced under stress, however, the reduction was rapid in susceptible cultivar [Fig. 2.3 (A), Table 2.2]. Similarly, chl a, b and carotenoids content were also affected by drought (Table 2.2). These pigments showed a proportionate reduction in each cultivar in comparison to control plants which was comparatively more rapid in S.3/A3. The chl/carotenoids ratio was also enhanced in both the cultivars and it was higher in TV-23. It was 0.64 and 0.52 in control plants whereas 0.72 and 0.57 in drought stressed plants of TV-23 and S.3/A3, respectively. Further, a uniform decrease in RWC was observed as compared to control for both the cultivars, though the rate of reduction was higher in case of susceptible one (Table 2.2). A higher roots length of drought stressed plants of tolerant cultivar was also noted in comparison to control plant [Fig. 2.3 (B), (C) and (D)]. The 25<sup>th</sup> d of water withholding was found as critical point for reviving the plants by rehydration; and after 30<sup>th</sup> d, even the tolerant cultivar was wilted permanently (Fig. 2.2).

### 2.3.2 Anatomical studies

Most of the stomata were partially closed in TV-23 whereas it was vice-versa in S.3/A3 under drought [Fig. 2.4 (image nos 1 & 2)]. There were no remarkable differences of petiole anatomy in between the cultivars under drought [Fig. 2.4 (image nos 3 & 4)]. Structural deformation of vacuoles, chloroplasts and mitochondria were also observed under drought whereas intact organelles were found in normal growth (Fig. 2.5). In each organelle, less membrane integrity and higher structural damage was observed in S.3/A3. It was observed that pore



Fig. 2.2: Drought induction experiment of tea seedlings: (1) potted seedlings of A. S.3/A3 and G. TV-23 (2) control plants of B. S.3/A3 and H. TV-23 (3) shoots of control plants of C. S.3/A3 and I. TV-23 (4) drought induced plants of D. S.3/A3 and J. TV-23 (before wilting, 21<sup>st</sup> d) (5) shoots of drought induced plants of E. S.3/A3 and K. TV-23 (on 25<sup>th</sup> day) (6) shoots of drought induced plants of F. S.3/A3 and L. TV-23 (wilted permanently).

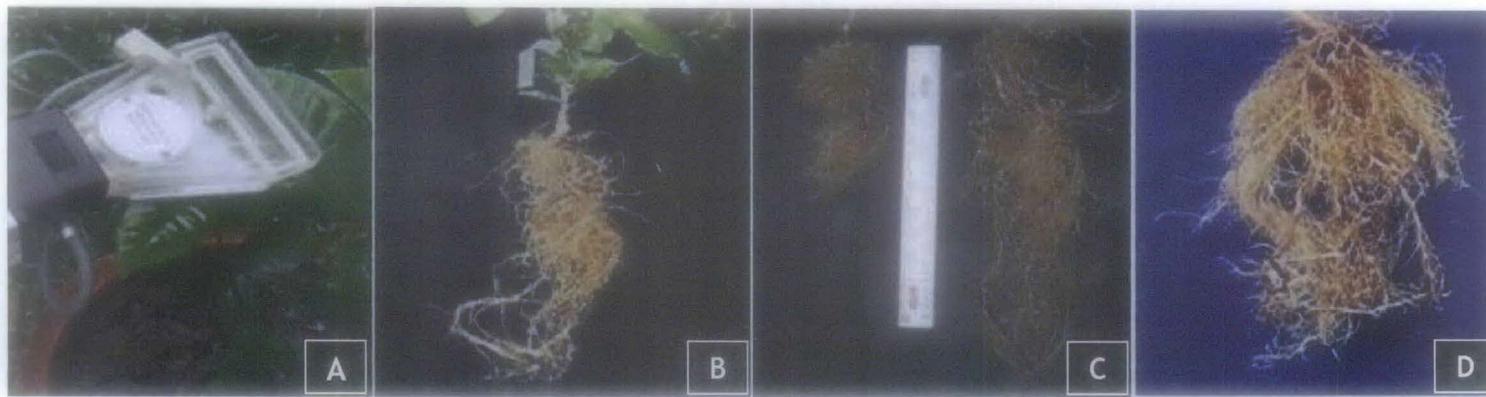


Fig. 2.3: Photographs of (A) photosynthesis rate, transpiration rate and stomatal conductance measurement using a photosynthesis system, (B) an uprooted plant, (C) root length measurement of control (left) and drought induced (right) plants of TV-23 and (D) roots for various experiments.

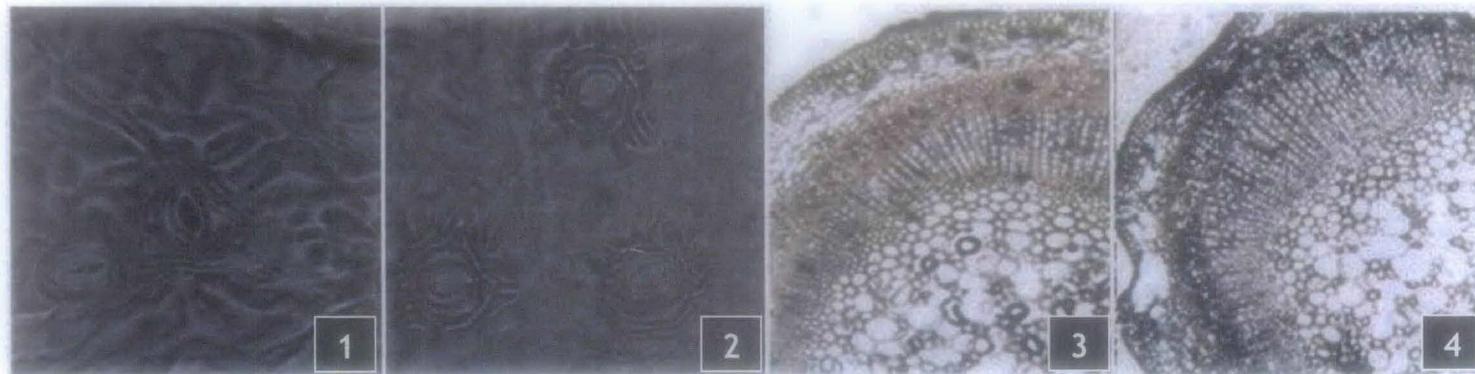


Fig. 2.4: Light microscopic images of stomata: (1) TV-23 and (2) S.3/A3 plants as well as transverse section of petioles: (3) TV-23 and (4) S.3/A3 plants on 21<sup>st</sup> d of drought stress induction.

Table 2.2: Effect of drought stress on RWC, photosynthesis rate (Pn), transpiration rate (E), stomatal conductance (SC) and pigment content of tea leaves\*

Cultivars	Treatment	Soil moisture (%)	RWC (%)	Pn ( $\mu\text{mol m}^{-2} \text{s}^{-1}$ )	E ( $\text{mmol m}^{-2} \text{s}^{-1}$ )	SC ( $\text{mmol m}^{-2} \text{s}^{-1}$ )	Chl a ( $\text{mg g}^{-1} \text{FW}$ )	Chl b ( $\mu\text{g g}^{-1} \text{FW}$ )	Carotenoids ( $\mu\text{g g}^{-1} \text{FW}$ )
TV-23	Control	35.33 $\pm$ 1.33a	92.14 $\pm$ 1.78a	14.38 $\pm$ 0.42a	4.45 $\pm$ 0.04a	0.56 $\pm$ 0.06a	2.38 $\pm$ 0.08a	399 $\pm$ 0.00ab	630 $\pm$ 0.02a
	Exp.	07.03 $\pm$ 1.50b	63.71 $\pm$ 5.81b	10.32 $\pm$ 0.41b	3.47 $\pm$ 0.07ab	0.52 $\pm$ 0.09a	1.85 $\pm$ 0.05b	389 $\pm$ 0.01ab	540 $\pm$ 0.01b
S.3/A3	Control	34.67 $\pm$ 1.33a	91.08 $\pm$ 2.55a	14.61 $\pm$ 0.36a	3.77 $\pm$ 0.06ab	0.56 $\pm$ 0.07a	2.24 $\pm$ 0.02a	430 $\pm$ 0.02a	639 $\pm$ 0.00a
	Exp.	07.21 $\pm$ 0.55b	59.87 $\pm$ 5.40b	07.14 $\pm$ 0.52c	3.04 $\pm$ 0.02c	0.31 $\pm$ 0.02b	1.77 $\pm$ 0.04b	373 $\pm$ 0.01c	534 $\pm$ 0.01b

Table 2.3: Effect of drought stress on ROS, lipid peroxidation (measured as MDA content) and level of membrane stability\*

Cultivars	Treatment	Tissue	MDA ( $\text{nmol g}^{-1} \text{FW}$ )	Electrolyte leakage (%)	Superoxide anion ( $A_{580}$ )	$\text{H}_2\text{O}_2$ ( $\text{nmol g}^{-1} \text{FW}$ )
TV-23	Control	Leaf	29.36 $\pm$ 0.05b	08.55 $\pm$ 0.35c	0.81 $\pm$ 0.03c	06.78 $\pm$ 0.95c
		Root	23.81 $\pm$ 0.05ab	18.31 $\pm$ 0.20b	0.69 $\pm$ 0.48b	42.44 $\pm$ 0.59c
	Exp.	Leaf	36.24 $\pm$ 0.19a	15.41 $\pm$ 0.71b	1.53 $\pm$ 0.09b	12.78 $\pm$ 0.91b
		Root	27.86 $\pm$ 0.03ab	18.75 $\pm$ 0.15b	1.02 $\pm$ 0.72a	78.40 $\pm$ 0.22b
S.3/A3	Control	Leaf	23.19 $\pm$ 0.00c	08.95 $\pm$ 0.25c	0.89 $\pm$ 0.02c	06.50 $\pm$ 1.79c
		Root	17.59 $\pm$ 0.07c	18.20 $\pm$ 0.70b	0.72 $\pm$ 0.51b	42.37 $\pm$ 0.06c
	Exp.	Leaf	31.49 $\pm$ 0.05ab	32.70 $\pm$ 0.00a	1.82 $\pm$ 0.05a	26.12 $\pm$ 2.00a
		Root	27.84 $\pm$ 0.07a	31.60 $\pm$ 2.10a	1.16 $\pm$ 0.82a	85.32 $\pm$ 0.87a

(\*Alphabets indicate least significant differences)

formation or disruption of vacuolar membrane, disorientation of lamellar system or deposition of starch granules in chloroplast, shrinkage and destroyed crista in mitochondria were the consequences of ultra structures under drought (Fig. 2.5).

### **2.3.3 ROS, lipid peroxidation and membrane stability**

Measurement of ROS, lipid peroxidation and membrane stability level were given in Table 2.3. For lipid peroxidation, MDA content in roots and leaves of both the cultivars were determined and found that MDA content was increased in both leaves and roots of each cultivar. However, the rate of increment was more in roots of each cultivar and it was comparatively higher in S.3/A3. Electrolyte leakage was also found high in leaves of each cultivar; however, the rate of increment was more in S.3/A3. Similarly, superoxide anion content was higher in each cultivar and the rate of formation was marginally more in roots of S.3/A3. H<sub>2</sub>O<sub>2</sub> content was also found significantly enhanced in both roots and leaves of each cultivar and the enhancement was remarkably higher in roots and comparatively more in S.3/A3.

### **2.3.4 TSS, RS, starch, protein, proline, phenolics, ascorbic acids and ABA content estimation**

Changes of different biochemical under drought stress were given in Table 2.4. Phenolics, TSS, RS and ABA content were increased in each cultivar and the increment was more in roots of TV-23. Protein, AA and proline content were found more enhanced in leaves of each cultivar, however, the enhancement were more and rapid in TV-23. Starch content was found in decreasing trend both in leaves and roots of each cultivar. The decrement of starch was found significant and more pronounced in roots of TV-23.

### **2.3.5 Activity of antioxidative enzymes**

Activities of different antioxidative enzymes such as APX, SOD, POX and CAT were depicted in Fig. 2.6. There was an overall high activity of APX which was increased 237.02% in leaves of S.3/A3 whereas it was 221.74% in TV-23. Although, APX activity was also showed a similar trend in roots, the enhancement was comparatively more

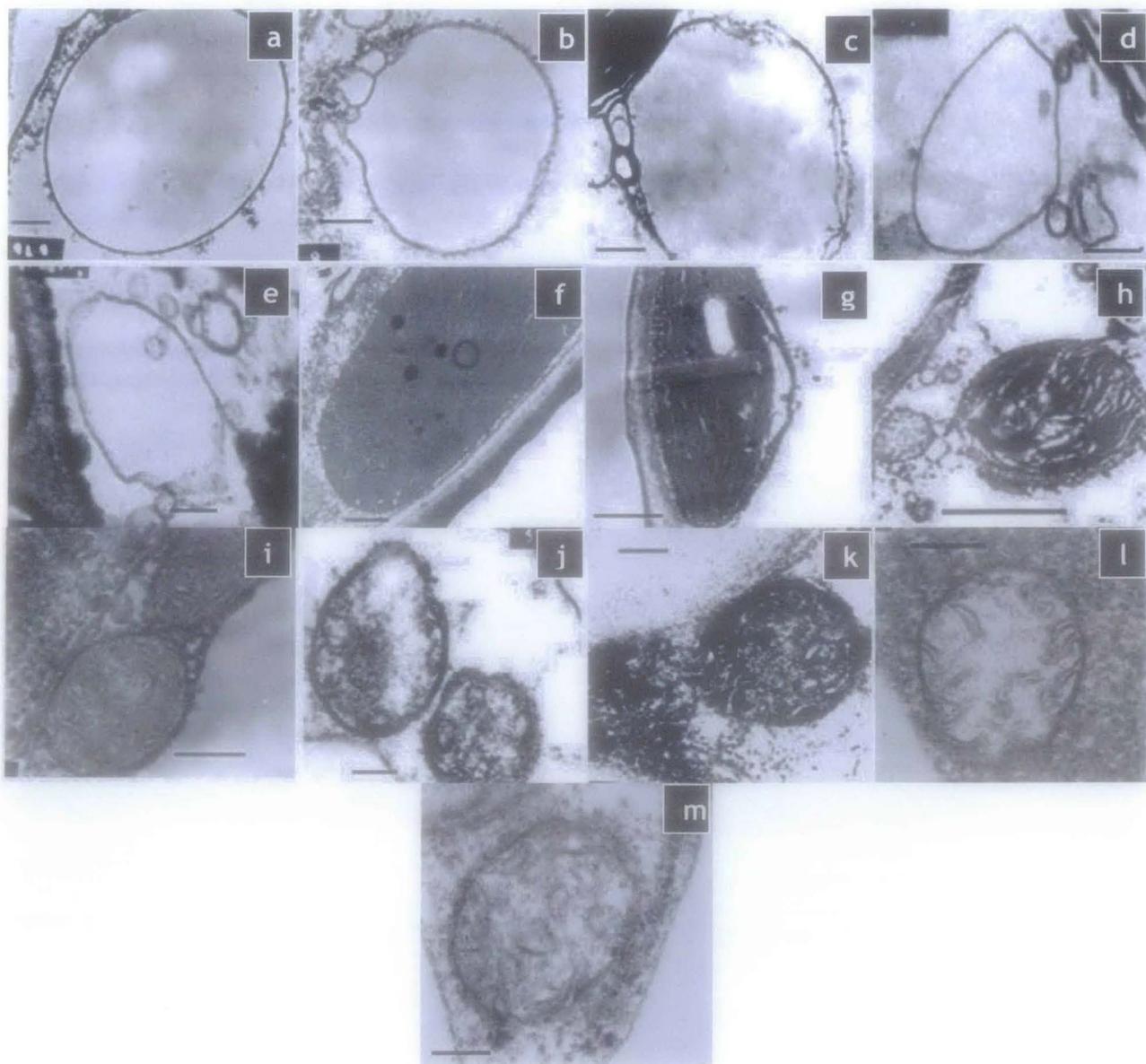


Fig. 2.5: Ultra-structural images of TEM: 1. **vacuoles** of [a] control plants (1.25 cm = 0.37  $\mu$ m); [b-e] drought stressed plants: [b] TV-23 leaf (1.25 cm = 1  $\mu$ m) [c] S.3/A3 leaf (1.25 cm = 0.37  $\mu$ m) [d] TV-23 root (1.25 cm=0.37  $\mu$ m) [e] S.3/A3 root (1.25 cm = 1  $\mu$ m); 2. **chloroplast** of [f] control plants (1 cm=0.71  $\mu$ m); [g, h] drought stressed plants: [g] TV-23 (1.25 cm=1  $\mu$ m) [h] S.3/A3 (1.25 cm = 2.5  $\mu$ m) 3. **mitochondria** of [i] control plants (1.25 cm = 0.25  $\mu$ m) [j-m] drought stressed plants: [j] TV-23 leaf (1 cm = 0.2  $\mu$ m) [k] S.3/A3 leaf (1 cm = 0.2  $\mu$ m) [l] TV-23 root (1 cm = 0.25  $\mu$ m) [m] S.3/A3 root (1 cm = 0.2  $\mu$ m).

Table 2.4: Effect of drought stress on different biochemicals

Cultivars	Treatment	Tissue	Protein (mg g <sup>-1</sup> FW)	Proline (μmol g <sup>-1</sup> FW)	Starch (mg g <sup>-1</sup> FW)	TSS (mg g <sup>-1</sup> FW)	RS (mg g <sup>-1</sup> FW)	Phenolics (mg g <sup>-1</sup> FW)	Ascorbic acid (μmol g <sup>-1</sup> FW)	ABA (mg g <sup>-1</sup> FW)
TV-23	Control	Leaf	15.50 ± 0.10b	5.61 ± 0.32b	27.30 ± 0.08a	31.83 ± 1.02ab	12.33 ± 0.21d	4.48 ± 0.05d	16.31 ± 0.61bc	90.52±0.26b
		Root	06.60 ± 0.40a	2.10 ± 0.04c	29.13 ±0.68a	34.48 ± 0.20c	09.70 ± 0.13d	1.78 ± 0.10c	03.15 ± 0.01c	79.17±0.65b
	Exp.	Leaf	17.20 ± 0.20a	9.34 ± 0.26a	11.61 ± 0.24c	40.57 ± 3.63a	14.76 ± 0.09b	5.76 ± 0.07b	28.78 ± 1.22a	96.31±0.53a
		Root	06.90 ± 0.10a	3.96 ± 0.06a	09.54 ±1.41c	49.66 ± 1.04a	13.11 ± 0.12b	3.43 ± 0.18a	03.64 ± 0.04a	89.35±0.85a
S.3/A3	Control	Leaf	12.00 ± 0.00d	2.38 ± 0.14b	20.49 ± 0.53b	27.80 ± 1.44c	13.28 ± 0.15c	4.10 ± 0.01c	10.49 ± 1.12bc	79.35±0.85c
		Root	06.10 ± 0.10a	2.39 ± 0.13c	25.46 ± 0.82b	34.31 ± 0.37c	11.73 ± 0.00c	1.58 ± 0.14c	02.97 ± 0.02d	75.15±1.20c
	Exp.	Leaf	13.30 ± 0.10c	4.68 ± 0.38c	09.67 ± 0.17d	38.78 ± 2.54a	16.21 ± 0.18a	5.33 ± 0.03a	19.53 ± 2.57b	80.04±0.50c
		Root	06.30 ± 0.10a	2.88 ± 0.07b	07.56 ± 0.54c	42.91 ± 0.16b	15.03 ± 0.06a	2.85 ± 0.13b	03.31 ± 0.00b	79.05±0.89b

(TSS=total soluble sugar, RS=reducing sugar, ABA=abscisic acids, alphabets indicate least significant differences)

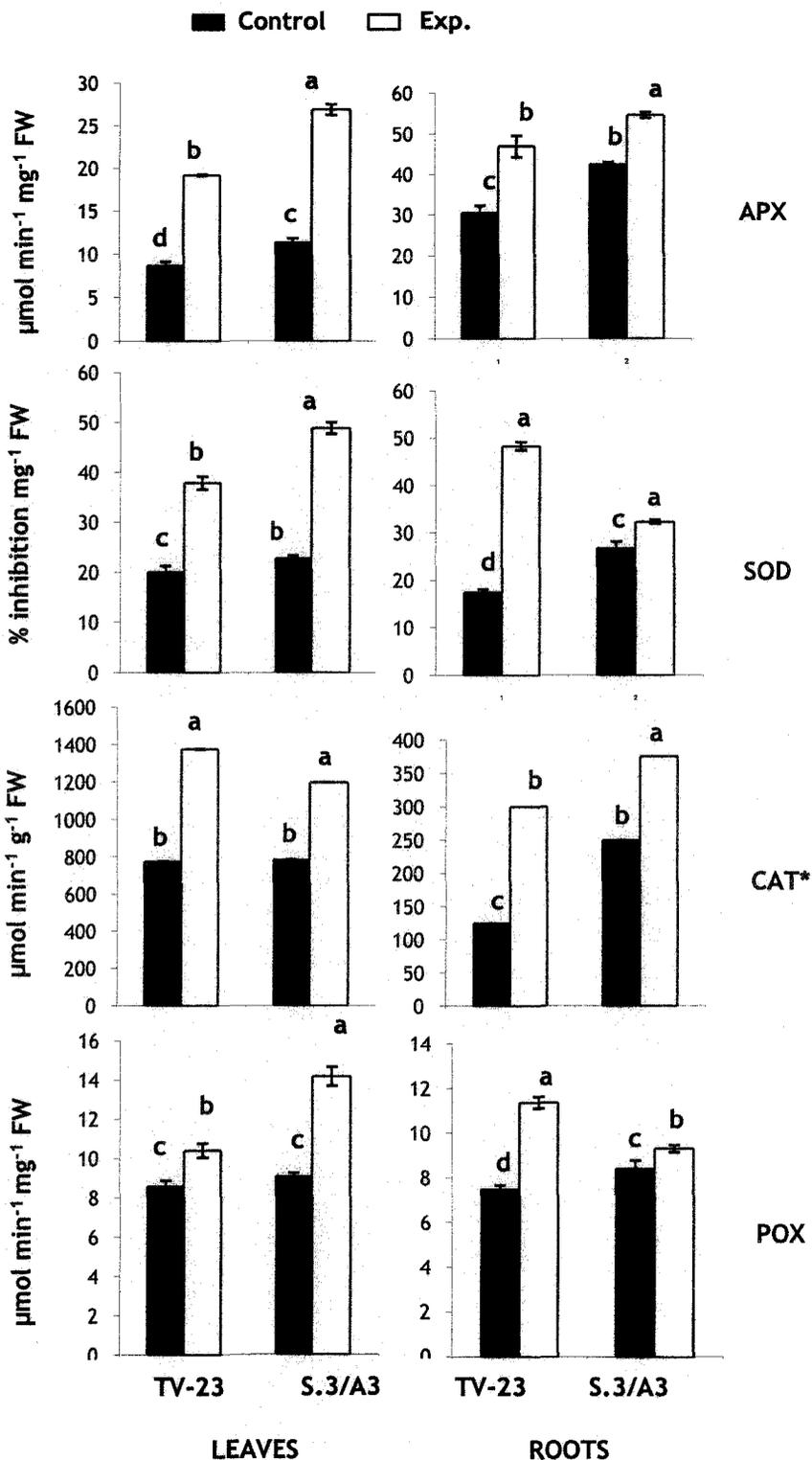


Fig. 2.6: Estimation of different antioxidative enzymes activity of leaves and roots of drought stressed and control tea plants (letters indicate least significant differences; \*standard error was negligible).

in leaves. Similarly, SOD activity was found more in leaves of S.3/A3 (214.21 %) than to TV-23 (188.81%), however, a reverse trend was registered in roots where it was 275.84% in TV-23 and 120.49% in S.3/A3. The CAT activity was also enhanced in each cultivar for both types of tissues, but the increment was higher in TV-23. It was increased 177% and 153% in leaves as well as 240% and 150% in roots of TV-23 and S.3/A3, respectively. Overall, CAT activity was found very high in leaves than to roots. Activity of POX registered a different trend where it was more in leaves of S.3/A3 (158%) in comparison to TV-23 (121%), however, in roots it was higher in TV-23 (151%) than to S.3/A3 (111%).

## 2.4 Discussion

Drought is one of the important environmental factors, which induces significant alterations in plant physiology and biochemistry. Changes occur as part of the tolerance mechanism which is an inherent capacity of the plants that enables to maximize water absorption and minimize transpiration during drought. A clear understanding of tolerance mechanism at cellular level provides us the scope for deciphering the lying mechanism at molecular level. Owing to the fast climatic changes and water limitation, understanding of drought tolerance mechanism for selecting and/or developing drought tolerant cultivar with quality crop is an urgent need in modern agriculture.

### 2.4.1 Plants growth and anatomical studies

A drastic change in plant's morphology and growth were observed under drought. Fresh and dry mass were also decreased more rapidly in susceptible cultivar. Perhaps, it was due to the fact that restriction of water supply from the soil in relation to genotype reduced the plants biomass (Ogbonnaya et al. 2003). Length of roots in tolerant cultivar was increased under drought may be due to facilitating maximum water absorption from water deficit soil. This finding was in accordance with the previous results of various crop plants including tea (Grzesiak et al. 1997, Burgess and Carr 1996). Moreover, Nagarajah and Ratnasuriya (1981) reported that rooting depth confers drought resistance in tea. The rate of CO<sub>2</sub> assimilation in leaf decreases at leaf water deficits or even before leaf water status changes in response to a drop in soil water potential (Gollan et al. 1986). Relative water content of leaves decreased in each cultivar due to drought but it was least in TV-23. Maintenance of high RWC in drought-tolerant cultivar has been reported as an adaptation to water stress in several crop species including tea (Valentovič et al. 2006, Nautiyal et al. 1995, Upadhyaya et al. 2008). Stomatal control of water loss has been identified as an early event in plant responses to water deficit leading to limitation of carbon uptake by the leaves (Chaves 1991). It occurs in response to either a decline in leaf turgor and/or water potential, or to a low-humidity atmosphere (Maroco et al. 1997). As a rule, stomatal responses are more closely linked to soil moisture content than to leaf water status. This suggests that stomata

are responding to chemical signals (e.g. ABA) produced by dehydrating roots (Davies and Zhang 1991). The TV-23 cultivar being a drought-tolerant cultivar restricts stomatal function under stress and improves water use efficiency. Pigment content is an indicator of photosynthetic activity. Pigment loss is associated with environmental stress and the variation in total chl/carotenoids ratio may be a signal of stress in plants (Ünyayar et al. 2005). Present results showed that pigment such as chl-a, chl-b and carotenoids were decreased under drought but the reduction was more and significant in S.3/A3. Chl/carotenoid ratios were also found more increased in S.3/A3 under drought stress. It may be an indication of higher pigment degradation in susceptible cultivar.

Cell division and enlargement as well as intercellular spaces are reduced by water deficit, which contribute to the survival of plants under dry conditions (Levitt 1980). Changes in water regime can strongly modify ultra structural cell characteristics if water stress is severe enough (Ciamporova 1987). Water stress can disrupt the ultra structural features of chloroplasts, mitochondria and vacuoles and alter the nucleus and cell membranes in general (Crevecoeur et al. 1987). Due to higher amount of compatible solutes in vacuole such as sugar and other derivatives for osmotic adjustment, it gets swelled. In increasing stress level, the vacuolar membrane faces higher pressure and gradually, it gets damaged (Blokhina et al. 2003). A high resistance of vacuolar membranes was observed in tolerant cultivar. Mitochondria are the site of ROS production through its electron transport chain (Chen et al. 2003). Under stress condition, higher production of ROS and their accumulation causes destruction of its own inner membranous structures. In susceptible cultivar, the mitochondrial structure was found almost destroyed due to its comparatively higher production and accumulation of ROS and its less capability of detoxification. Due to production, transportation and detoxification of ROS, the inner structures of chloroplast such as thylakoids and lamella were observed as disoriented; however the distortion was severe in susceptible cultivar. The main damage to the chloroplast caused by water stress also reported earlier as excessive swelling, distortion of thylakoids, and the appearance of lipid droplets (Vieira 1976, Berlin et al. 1982). Increase in chloroplast starch accumulation under drought has also been observed in cotton plants (Ackerson and Hebert 1981). The destroyed chloroplast in susceptible cultivar proved its inefficiency in drought tolerance. Changes in water

supply seem to affect the structural characteristics of chloroplasts differentially in various cultivars such as bean, maize, cotton (Stoyanova and Yordanov 1999, Ristic and Cass 1991, Utrillas and Alegre 1997). Hence, observation of the damage to cell membranes, such as plasmalemma and thylakoid membranes, provides valuable information for the ability of a plant to withstand stress (Utrillas and Alegre 1997). Shrinking of cells leads to loss of turgor, osmotic stress and a potential change of membrane potentials. Upon severe water loss from the cells, membrane disintegration and abolition of metabolic processes occur (Mahajan and Tuteja 2006).

#### **2.4.2 Determination of ROS, lipid peroxidation and membrane stability levels**

Reactive oxygen species inactivate enzymes and damage important cellular components. The increased production of toxic oxygen derivatives is considered to be a universal or common feature of stress conditions (Arora et al. 2002). ROS are responsible for stress-dependent peroxidation of membrane lipids (Ratnayaka et al. 2003). Lipid peroxidation is often used as an indicator of increased oxidative damage (Cornic and Briantais 1991). The plant cell membranes are considered as a primary site of stress injury and the membrane destabilization is frequently attributed to lipid peroxidation (Singh et al. 2006). Increased production of MDA by environmental stress exposure has been well-documented in various crop plants such as sunflower, wheat, maize etc. including tea (Gallego et al. 1996, Hong-Bo et al. 2005, Moussa and Abdel-Aziz 2008, Upadhyaya et al. 2008). It has been suggested that plasma membrane stability may be a reliable index of drought and heat resistance in plants (Blokhina et al. 2003). Dehydration in sensitive plants is often accompanied by cytoplasmic membrane injury, resulting in the leakage of solutes. The increase level of electrolyte leakage under drought has been reported in several crops (Panda et al. 2003, Mohammadkhani and Heidari 2007, Pandey et al. 2010). In present study, there was a significant increase in the level of both MDA content and electrolyte leakage in leaves under drought stress. It indicates that leaves are more vulnerable to drought stress than roots. Higher level of lipid peroxidation observed here was probably due to the harmful effect of excessive levels of  $H_2O_2$  or its ROS derivatives in the cellular compartments (Bowler et al. 1992). Excessive levels of ROS may have resulted in damage to cell organelles including the photosynthetic apparatus,

ultimately leading to severe cellular damage and chlorosis of the leaves (Blokhina et al. 2003). Lower level of lipid peroxidation in leaves and roots of TV-23 suggest that, this cultivar is better protected from oxidative damage under drought stress than S.3/A3. This result is in agreement with the results of tea (Upadhyaya et al. 2008) and wheat (Sairam et al. 2005) where the tolerant cultivar showed lower membrane damage under drought stress.

#### **2.4.3 Determination of TSS, RS, starch, protein, proline, phenolics, ascorbic acids and ABA contents**

Sugars are special in that they allow the removal of closely associated water from protein without leading to conformational changes and loss of enzymatic function. According to the water replacement hypothesis, sugars act as a water substitute by satisfying the hydrogen-bonding requirement of polar groups of dried protein surface (Blokhina et al. 2003). Accumulation of soluble and reducing sugars may occur perhaps due to degradation of sub-cellular component affected by drought stress. It was found that starch was decreased under drought in both the cultivars. It indicates that due to less water potential raised soluble sugar fraction was accompanied by a significant decrease in the starch fraction. Similar result was also observed in rice, maize and sugar beet (Vajrabhaya et al. 2001, Mohammadkhani and Heidari 2008a, Hoffmann 2010).

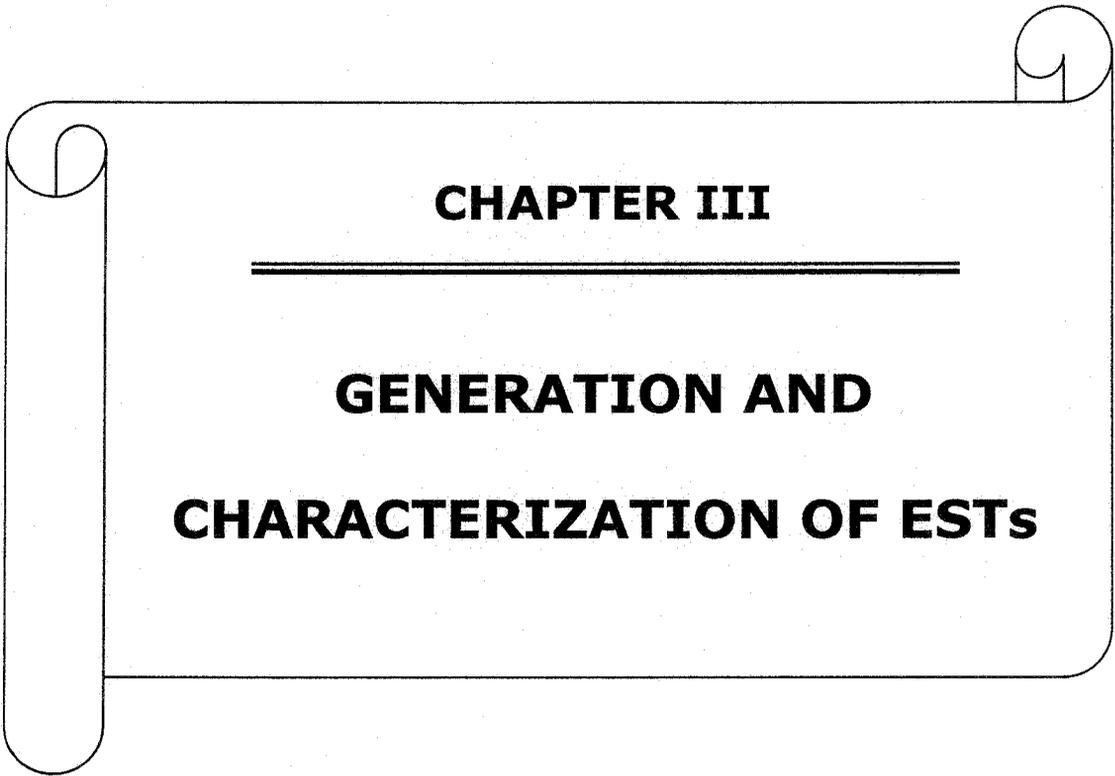
Osmotic adjustment involves an active accumulation of cellular solutes such as proline and soluble proteins within the plant in response to lowering of the soil water potential and reducing the harmful effects of water deficit (Hong-bo et al. 2008). Proline is known to accumulate in response to dehydration treatment which is considered as compatible osmolyte as well as scavenger against ROS (Matysik et al. 2002). In present study, higher proline content in TV-23 confirmed that this cultivar is better protected under drought stress as it helps in water relation; prevent membrane distortion and acts as a hydroxyl radical scavenger. High proline accumulation in response to drought stress was also reported in several crop plants including tea (Upadhyaya et al. 2008). Increased in total phenol and ascorbate contents in TV-23 in response to drought stress suggested the gradual protection of seedling to overcome a drought-induced oxidative damage as reported in numbers of

plant species including tea (Dixon and Steele 1999, Chakraborty et al. 2002, Cheruiyot et al. 2007). Ascorbate is a key substance in the network of antioxidants that includes ascorbate, glutathione,  $\alpha$ -tocopherol, and a series of antioxidative enzymes. Ascorbate has also been shown to play multiple roles in plants growth, such as in cell division, cell wall expansion, and other developmental processes (Smirnoff and Wheeler 2000). Phytohormone ABA plays a key role in plant's adaptation to adverse environmental conditions including drought. Numerous studies have shown that ABA accumulation is a key factor in controlling downstream responses essential for adaptation to stresses (Ramagopal 1987, Shinozaki and Yamaguchi-Shinozaki 1996). In present study also higher accumulation of ABA was found under drought, and it was comparatively high in the tolerant cultivar. It proves that TV-23 cultivar is better in adaptation to the adverse environmental conditions.

#### **2.4.4 Activities of antioxidative enzymes**

To keep the levels of active oxygen species under control, plants have non-enzymatic and enzymatic antioxidant systems to protect cells from oxidative damage (Mittler 2002). Antioxidant enzymes play an important role in defense system of plants against oxidative stress such as SOD, APX, CAT, POX and GR (Hongbo et al. 2008). The balance between ROS production and activities of antioxidative enzymes determines whether oxidative signalling and/or damage will occur. The capability of scavenging ROS and reducing their damaging effects may correlate with the drought tolerance of plants. There are many reports in literature that underline the intimate relationship between enhanced or constitutive antioxidant enzyme activities and increased resistance to drought stress in several plant species, such as rice (Guo et al. 2006), wheat (Khanna-Chopra and Selote 2007), common bean (Turkan et al. 2005) and barley (Acar et al. 2001). In present investigation, activity of POX was enhanced with drought which has also been reported in various plant species including tea (Chakraborty et al. 2002); suggesting POX plays an important protective role against drought stress. Enzyme, SOD is an essential component of plant antioxidation system as it dismutates superoxide radicals to  $H_2O_2$  and  $O_2$  in the cytosol, mitochondria and chloroplast (Salin 1988). Activity of SOD was also found up-regulated by drought stress in numbers of plant species including tea

(Chakraborty et al. 2002, Upadhyaya et al. 2008). Increased level of APX and CAT activities in response to drought are well-established in literature (Srivalli et al. 2003, Upadhyaya and Panda 2004).



**CHAPTER III**

---

**GENERATION AND  
CHARACTERIZATION OF ESTs**

## **GENERATION AND CHARACTERIZATION OF ESTs**

### **3.1 Review of literature**

Studies on the molecular basis of stress responses were started after the observation of altered puffing pattern in *Drosophila* polytene chromosome with the changes of temperature. It revealed that under heat shock conditions protein profile was altered. Thereafter, it was found in plants too where heat shock conditions induce comparable alterations in protein profile (Grover et al. 2001). It has now been an established fact that stress tolerance in plants is a genetically complex event based on activation and regulation of specific stress related genes (Wang et al. 2003). Unlike biotic stress, abiotic stress response is a multigenic trait and the molecular mechanism is more difficult to understand. There are cascades of gene networks that are involved in stress perception, signal transduction, transcriptional control and scavenging of ROS (Vinocur and Altman 2005). A number of genes get turned on during stress by increasing the levels of several metabolites and proteins to fight against the cellular damages. However, various genes get differentially regulated under different abiotic stress conditions (Vandez and Sharma 2008).

#### **3.1.1 Abiotic stress associated genes and genetic mechanisms**

Research into the plant response to abiotic stress is becoming increasingly important due to global climate change. Genes involved in the complex mechanisms of abiotic stress responses can broadly be categorized in three groups (Wang et al. 2003).

##### **3.1.1.1 Signaling cascades and transcriptional control**

Genes involved in signaling cascades are MAP kinases, SOS kinases, Phospholipases, MyC etc. and genes involved in transcriptional control are HSF, C-repeat binding factor/dehydration responsive element binding protein (CBF/DREB), ABA responsive element binding factor/ ABA responsive element (ABF/ ABRE) families (Shinozaki and Yamaguchi-Shinozaki 2007) etc..

### 3.1.1.2 Heat shock proteins and chaperons

These genes are involved in controlling the proper folding and conformation of structural and functional proteins. It includes heat shock proteins and chaperons including the LEA proteins (Vinocur and Altman 2005). Heat shock proteins are mostly encoded by nuclear genes and located in cytoplasm, mitochondria, chloroplast and endoplasmic reticulum (Grover et al. 2001). In broad sense, *hsp* genes can be categorized in two groups on the basis of their expression: i) heat shock cognates (*hsc*) or constitutively expressed and ii) strongly induced under heat stress (*hsp*) (Vierling 1991). Low molecular weight HSPs and HSP70 are the extensively studied and well characterized *hsp* genes. The nucleotide sequence and structural features of the *hsp* genes are found conserved (Singla et al. 1997, Vierling 1991).

### 3.1.1.3 Ion and water transport

These genes are involved mainly in water and ion uptake and their transport. The aquaporin and different transporter genes are of this category. Aquaporin is a large group of protein family including more than 150 membrane channel proteins which are involved in complex and regulated water transport within the plant in order to adapt with different environmental conditions (Maurel 2007). They are present in the plasma membrane and tonoplast to control the water flux into and out of a cytosol or vacuole (Kaldenhoff and Eckert 1999). It contributes to changes in turgor or cell size. It increases the transmembrane water flux when the rapid translocation of large volumes is required and diffusional water flow across the lipid bilayer is insufficient. They belong to a well-conserved and ancient family of proteins called the major intrinsic proteins (MIPs) with molecular weights in the range of 26-34 kDa (Zhao et al. 2008a). For example, *Arabidopsis thaliana*  $\alpha$ TIP,  $\delta$ TIP (tonoplast intrinsic protein), *Nicotiana TobRB7* and *Mesembryanthemum* MIP-A are of this category. There are some examples of aquaporin which expressed differentially during stress condition. For example,  $\gamma$ TIP and  $r$ TIP in *Arabidopsis*, members of the *Craterostigma* PIP family (plasmamembrane intrinsic proteins) and *Oryza*  $r$ TIP have shown up-regulation during drought stress (Kaldenhoff and Eckert 1999).

### 3.1.2 Drought stress associated genes and their regulation

Plant induces a series of physiological and biochemical responses under drought like that of other abiotic stresses. Hence, drought stress associated genes are involved with diverse functions such as stomatal closure, repression of cell growth, photosynthesis, protection against oxidative damage etc. (Fig. 3.1). These genes are either induced or repressed under drought conditions (Shinozaki et al. 2003) through ABA independent or dependent regulatory pathways (Shinozaki and Yamaguchi-Shinozaki 1997).

#### 3.1.2.1 ABA independent pathways

There is a network of genes that are involved in the regulation of drought and other abiotic stress related gene's expression (Fig. 3.2). Two major *cis*-acting elements, ABRE and DRE/CRT present in the promoter region of drought, salinity and cold inducible gene, *RD29A/COR78/LTI78* are involved in the expression of stress inducible gene. However, ABRE and DRE/CRT are involved in both ABA-responsive and ABA independent gene expression under abiotic stress (Shinozaki and Yamaguchi-Shinozaki 2007). *CBF/DREB1* and *DREB2* genes of ERF/AP2 family transcription factors are induced by abiotic stress and may activate other genes involved in cold and drought stress tolerance (Liu et al. 1998). *CBF/DREB1* increased cold, drought and salt tolerance and was found by overexpressing the genes in transgenic *Arabidopsis* plant. This gene is induced rapidly and transiently under abiotic stress (Liu et al. 1998). *DREB2* increased drought tolerance level in transgenic *Arabidopsis* through the trans-activation of stress inducible genes (Sakuma et al. 2006). *DREB2* protein gets activated in the early stage of dehydration stress through its post-translational modification (Shinozaki and Yamaguchi-Shinozaki 2007). *ERD1* gene response to dehydration through ABA-independent pathway and was found to be up-regulated during senescence (Nakashima et al. 1997). Recently, DNA binding proteins interacting with the *cis*-acting element of *ERD1* promoter was identified as NAC domain transcription factor (Tran et al. 2004).

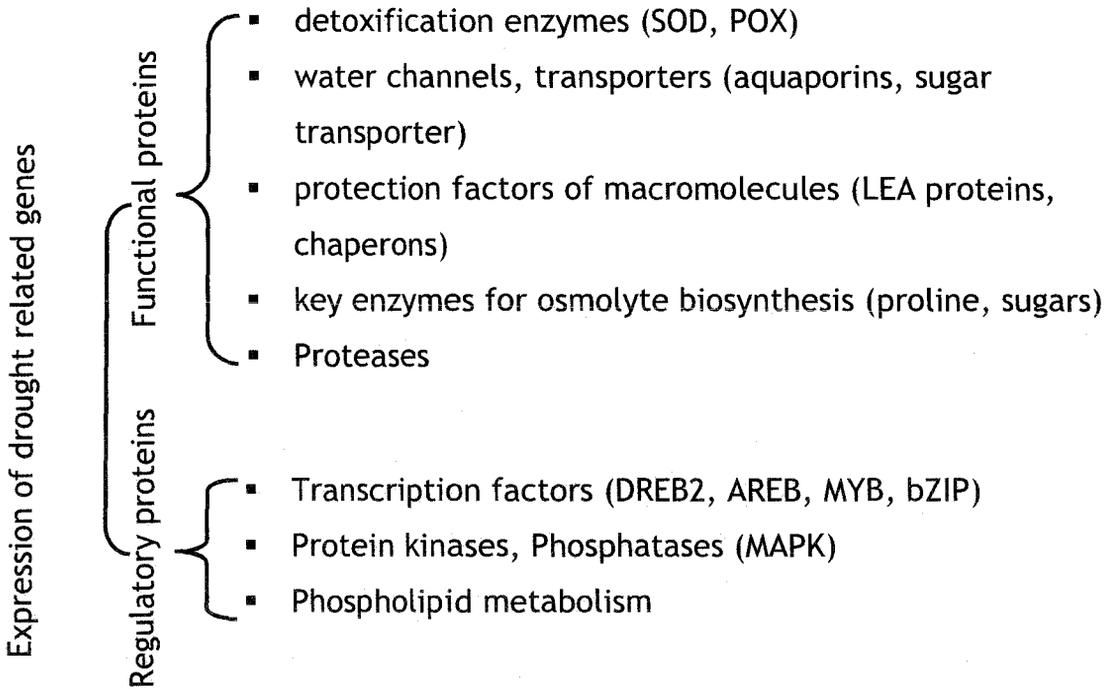


Fig. 3.1: Functions of drought associated genes

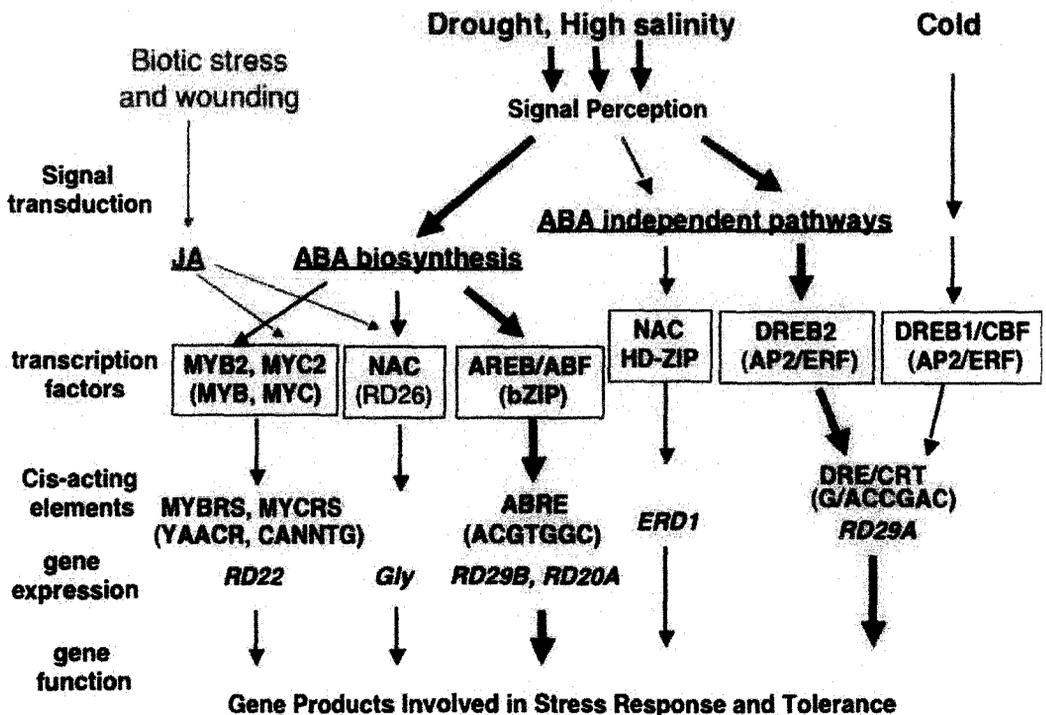


Fig. 3.2: Transcriptional regulatory networks of drought as well as other abiotic stress signals (Shinozaki and Yamaguchi-Shinozaki 2007).

### 3.1.2.2 ABA dependant pathways

An ABA mediated signal is required for the activation of AREB/ABF proteins. ABRE is a major *cis*-acting element in ABA responsive pathway (Uno et al. 2000). This phenomenon is probably due to the ABA dependant phosphorylation of the AREB/ABF proteins (Shinozaki and Yamaguchi-Shinozaki 2007). Overexpression of *ABF3* improved drought tolerance in transgenic *Arabidopsis* (Kang et al. 2002). MYC and MYB transcription factors are synthesized following the accumulation of endogenous ABA that enhanced plant's tolerance to drought stress (Abe et al. 2003). *RD22* gene is induced by ABA under drought stress following the ABA dependant pathway for protein biosynthesis (Shinozaki and Yamaguchi-Shinozaki 2007). *RD26* gene which encodes a NAC transcription factor is also induced by drought, high salinity, ABA and by JA treatments. Recently, it was found that *RD26* mediates between ABA signaling and JA signaling during drought and wounding stress responses (Fujita et al. 2004).

### 3.1.3 Importance for studying tree genomics

Completion of genome projects for the model and other various plant species as well as the development of various genetic, genomic, and biochemical tools offer many possibilities to challenge the unknown biological questions. Small physical size, rapid generation time, straight-forward genetics, high fecundity and small genome size are the reasons for *Arabidopsis* being chosen as a model plant for research. Although this greatly facilitates functional studies, most plant species have different reproductive strategies that make *Arabidopsis* a genetic extreme. Rice, a model species for studying monocotyledonous plant species, is an important crop and has a high degree of synteny with other cereal plant genomes, such as maize, wheat, barley and other grasses. This is because their genomes share a considerable similarity in their sequence and organization (Buell et al. 2005). However, the much accelerated life cycles of *Arabidopsis* and rice also make many traits that are essential in many (or most) plants unimportant such as wood formation and seasonality of growth. Trees represent an opposite extreme to *Arabidopsis* both physiologically and genetically with long life spans, generation times and woody perennial growth habits. Thus, *Populus* offers a new model system to study an expanded repertoire of biological processes that better represent the breadth of

plant biology (Jansson and Douglas 2007). Development of *Populus* as a model system for tree and woody perennial plant biology has been largely driven by the completion of a draft sequence of the *Populus trichocarpa* (black cottonwood) (Tuskan 2006) genome and the rapid development of genomic and molecular biology resources for this genus. However, *Populus* is still in nascent stage to be used as a model plant species for studying tree genomics and *Arabidopsis* remains as the choice for many plant biologists for comparative transcriptome analysis. Tea is an important plantation crop in several countries worldwide. However, there were no reports of comprehensive and comparative genomic analysis of this important cash crop.

### 3.1.4 Techniques for transcriptome analysis

In the last two decades, there is a huge development in the field of genomics. It leads to the understanding of gene expression, organization and its relationship to stress tolerance (Pérez-Torres et al. 2009). Identification and characterization of differential gene expression from tissues subjected to stress has gained much attention in plant science research. Recognition of genes or genetic elements involved in response to a particular stress enhances the possibility of promoting crop improvement through direct genetic modification. Analysis of transcripts expressed in different tissues has been enormously facilitated by the development of PCR. Several different techniques have been established to reveal differential transcripts such as differential display (Liang and Pardee 1992), cDNA-AFLP (Vos et al. 1995), microarray (Zhao et al. 2006), SAGE (Powell 2000) and SSH (Diatchenko et al. 1996).

#### 3.1.4.1 Subtractive hybridization

This technique is based on re-association kinetics of nucleic acids. Here, cDNAs of two different samples are hybridized in order to isolate the differential genes. Transcripts expressed in both the tester (target sample) and driver (sample to be subtracted) would form an mRNA/cDNA hybrid molecule, whereas a cDNA sequence unique to the tester would remain single-stranded. Single and double-stranded molecules were separated using hydroxylapatite chromatography, and the eluted single-stranded cDNA represented differentially expressed genes in tester. The

isolated cDNA can be cloned or directly used as probes for screening libraries. This method was first described in the early 1980s for the purpose of creating cDNA libraries and generating probes of differentially expressed genes. Since then, it has been modified by several workers, using biotin tags (Welcher et al. 1986) or oligo (dT) 30-latex (Hara et al. 1991), PCR amplification to decrease the initial amount of mRNA required and to improve the cloning efficiency of selected transcripts (Hara et al. 1991). Later, this technique was designed to selectively amplify differentially expressed transcripts by suppressing the amplification of abundant transcripts and known as SSH technique (Diatchenko et al. 1996). SSH technique is now widely used for the purpose of identifying differentially expressed genes on a global scale.

#### **3.1.4.2 Advantages of SSH techniques**

Suppression subtractive hybridization is a PCR-based cDNA subtraction method (Diatchenko et al. 1996). It is used to selectively amplify differentially expressed target cDNA fragments and simultaneously suppress the non-target DNA amplification. This subtraction method overcomes the problem of differences in mRNA abundance by incorporating a hybridization step that normalizes (equalizes) sequence abundance during the course of subtraction by standard hybridization kinetics. Hence, there is no requirement of any intermediate step(s) for physical separation of single stranded and double stranded cDNAs and only one subtractive hybridization round is required through which greater than 1000-fold enrichment for differentially expressed cDNAs can be achieved. Furthermore, the subtracted cDNA mixture can be used directly as a hybridization probe for screening recombinant DNA libraries and thereby chromosome-specific and tissue specific expressed sequences can be identified.

Due to the above cited advantages of SSH technique and its recent success of identifying differentially expressed genes in plants (Xiong et al. 2001, Park et al. 2004, Way et al. 2005, Cairney 2006), this technique was applied for the present investigation.

### 3.1.5 Transcriptome analysis of woody plants: ESTs as a strategy

Expressed sequence tags are short and single-pass sequence reads from mRNA (cDNA). They are created by sequencing the 5' and/or 3' ends of randomly isolated gene transcripts that have been converted into cDNA (Adams et al. 1991). They represent a snapshot of genes expressed in a given tissue and/or at a given developmental stage (Chen et al. 2005). A typical EST represents only a portion (~200-900 nt) of a coding sequence and this partial sequence data is of substantial utility. Expressed sequence tags collection is a relatively quick and inexpensive route for discovering new genes, confirming coding regions in genomic sequence, creating opportunities to elucidate phylogenetic relationships, facilitating the construction of genome maps and providing the basis for the development of expression arrays also known as DNA chips (Alba et al. 2004). They help us to understand easily about the genes involved in plant's growth, differentiation, secondary metabolism and biochemical pathways as well as responses to pathogen and environmental stress (Chen et al. 2005).

Generation and characterization of ESTs have been used as a strategy for transcriptome analysis in various crop plant species such as rice, wheat, maize, grape, apple and citrus. It has been studied basically for gene identification and characterization (Way et al. 2005, Clement et al. 2008, Reddy et al. 2008, Fukuoka et al. 2010), expression and comparative studies (Rodrigues et al. 2009, Lata et al. 2010, Gao et al. 2008, Ewing et al. 1999), germplasm characterization (Bhat et al. 2005, Sharma et al. 2009) as well as for genome mapping and analysis (Lazo et al. 2004, Aggarwal et al. 2007). There are more than 10 million ESTs available in dbEST (<http://www.ncbi.nlm.nih.gov/dbEST>) from various crop plant species by 7<sup>th</sup> May, 2010. However, the NCBI database has very little sequence information from woody perennial crop plant species compared to the large-scale single pass sequence entries from annual crop plants. Although similar to annuals in many biochemical pathways, perennials are unique in the fact that they possess long generation times. Without short cycle reproduction as an escape mechanism, perennials have evolved alternative survival mechanisms to pathogen attack and environmental stresses. The study of these alternate strategies by way of functional genomics will greatly increase the understanding of biochemical changes underlining stress responses in

perennials (Bausher et al. 2003). A summary of EST studies in woody plants is provided in Table 3.1.

A brief discussion is given here on transcriptome analysis of woody plants by EST as an approach:-

Shelton et al. (2001) identified 454 clones from a cDNA library of flush growth *Melaleuca alternifolia*, an Australian native tree with monoterpene-rich essential oil. The homology searching of these ESTs revealed similarity with enzymes like cycloartenol synthase, limonene synthase, transcription factors and photosynthesis related proteins. A total of 5% of these ESTs were involved directly with secondary metabolite production. This study illustrates the utility of ESTs as a new method for bio-prospecting of secondary metabolites from simple plant tissues.

Kohler et al. (2003) presented a large-scale production of ESTs from roots of the hybrid cottonwood, *Populus trichocarpa* x *deltoids*. The cDNA libraries were generated from the roots of 2-month-old rooted cuttings and 2.5-month-old 19 d water-stressed cuttings. The partial sequences obtained from 7013 clones were assembled into 1347 clusters and 3527 singletons. This set of ESTs represented 4874 unique transcripts expressed in roots. Putative function was assigned to 3021 (62%) transcripts. There was a significant portion of ESTs which encoded proteins of common metabolic pathways related to energy (5%) and metabolism (8%). Six percent of the total ESTs were involved in signaling pathways and hormone metabolism, while 4% encoded for transporters and channels. They studied the changes in aquaporins and transporter transcripts during adventitious root's development.

Bausher et al. (2003) reported around 6443 numbers of high quality ESTs from a cDNA library of immature sweet orange (*Citrus sinensis*) seedling. Of these, a total of 2272 ESTs (35%) significantly matched with known proteins, 1457 ESTs (23%) matched with unknown proteins and 1619 ESTs (25%) with proteins of putative functions. The remaining 1095 ESTs (17%) did not match with any protein sequence which found in the public databases. The abundant ESTs (6.0% of total) were of photosynthetic proteins chl a/b binding, plastocyanin and ribulose-1,5-bisphosphate

Table 3.1: A snapshot of EST analysis in woody plants

Plants	Tissue	Total ESTs	Unigenes (C+S)*	Remarks	References
Poplar ( <i>Populus trichocarpa</i> X <i>P. deltoides</i> )	Root	8808	4874 (1347+3527)	Two libraries made from the adventitious root of well-watered and 19 d water stressed plants (2-month-old cuttings), respectively for understanding the processes of water & nutrient uptake as well as interactions with rhizosperic micro-organisms; analysed the transcript abundance of genes in water and nutrient absorption using microarray; aquaporin and different transporter found differentially expressed.	Kohler et al. 2003
Cycads ( <i>Cycas rumphii</i> )	Leaf	4200	2458 (1917+ 541)	One library constructed from immature leaves (accession 808/59 A) of this ancient seed plant. Identified related ESTs of angiosperms & gymnosperms; identified and analysed GLR-like gene & their synthesis associated genes.	Brenner et al. 2003
Tea ( <i>Camellia sinensis</i> )	Leaf	588	nr	One SSH library was constructed from leaf sample for trapping the genes involved in secondary metabolism; identified several important metabolism related genes.	Park et al. 2004

Tea ( <i>Camellia sinensis</i> )	Shoot	1684	nr	One library was constructed from tender shoots of Longjing variety for studying the gene expression pattern; identified several important genes.	Chen et al. 2005
Tea ( <i>Camellia sinensis</i> )	Leaf	420	nr	One SSH library was constructed from cold treated leaf sample.	Wang et al. 2008
Tea ( <i>Camellia sinensis</i> )	Root	5115	3482 (901+2581)	One library was constructed for root transcriptome analysis.	Zhao et al. 2008b
Taxus ( <i>T. cuspidata</i> )	Cultured cells	8424	3563 (nr)	One library was constructed from the cultured <i>T. cuspidata</i> cells induced for taxoid biosynthesis using methyl jasmonate; revealed 12 defined genes of Taxol biosynthesis, yielded cDNAs encoding two previously uncharacterized cytochrome P450 taxoid hydroxylases, and provided candidate genes for all but one of the remaining seven steps of taxol biosynthesis.	Jennewein et al. 2004
Conifers ( <i>Cryptomeria Japonica</i> ; <i>Chamaecyparis obtusa</i> )	Leaf, Pollen Bud, Seed/ Pollen cones Cabbium Tissues	18693	8391 (5373+3018)	Eight libraries constructed from different tissues of <i>C. japonica</i> and one from cambium and surrounding tissues of <i>C. obtusa</i> and generated two sets of ESTs (12791 and 5902, respectively) for identifying lineage specific genes of conifers; identified 6% unique ESTs homologs to other conifers and 70% homologs to angiosperms.	Ujino-Ihara et al. 2005
Grape ( <i>Vitis</i> )	Bud, Leaf	146075	- 25746	More than 58 cDNA libraries constructed from	Silva et al. 2005

<i>vinifera</i> )	Stem, Petiole, Flower, Berry, Root etc.			different tissues at different developmental stages and stresses from different species; identified 665 differentially expressed transcripts across 29 libraries; revealed expected associations between plant developmental stages with tissue types, with an exception of abiotic stress treatments.	
Madagascar periwinkle ( <i>Catharanthus roseus</i> )	Leaf, Root	9824	5023 (3327+ 1696)	Two libraries constructed from the base part (composed of the petiole, including 1/3 of the leaf) of young leaves (1.5 cm), and of root tips (3-10 mm in length), respectively; identified 8 MIA biosynthesis related genes for the first time.	Murata et al. 2006
Papaya ( <i>Carica papaya</i> )	Fruit	1171	827 (117+710)	Two libraries constructed from the fruit tissue during developmental stages (immature green, mature green, of four 25% colour break and complete skin colour break) from Taiwanese Tainung Hybrid and Australian Hybrid respectively, identified 14 significant fruit ripening related ESTs.	Devitt et al. 2006
Oak ( <i>Quercus petraea</i> )	Bud	801	233 (137+96)	Three libraries constructed and generated 403, 206 and 192 sequences from the early, late and quiescent stages of bud burst, respectively; identified 233 differentially expressed unique transcripts during bud burst.	Derory et al. 2006

Rubber tree ( <i>Hevea brasiliensis</i> )	Latex tissue	10040	3441 (1380+2061)	One library was constructed from the latex of 15 year-old trees (clone RRIM 600); identified numbers of ESTs related to rubber biosynthesis and stress or defence responses; revealed nine variants of rubber particle membrane protein (RPMP) including rubber elongation factor and small rubber particle protein.	Chow et al. 2007
Cassava ( <i>Manihot esculenta</i> Crantz)	Leaf, stem, petiole, Root meristem	18166	8577 (3194+5383)	Two libraries constructed from pooled samples of dehydration-stressed and control well-watered tissues of cassava genotype, TME117; uncovered a wealth of putative drought-responsive genes.	Lokko et al. 2007
Sweet orange ( <i>C. sinensis</i> , <i>X C. fastidiosa</i> & <i>C. reticulata</i> )	Leaf	19200	12449 (7814+4635)	Two libraries constructed from the leaves of (strain 9a5c) tolerant 'Ponakan' and susceptible 'Pera IAC' sweet orange, respectively; identified 172 differentially expressed genes through <i>in silico</i> differential display.	Souza et al. 2007
Blackberry ( <i>Rubus sp.</i> )	Leaf	3000	2382 (301+2081)	One library constructed from the leaf tissue of Merton Thornless cultivar; detected 673 SSRs and tested 33 primer pairs in two cultivars, of which ten showed average 1.9 polymorphic PCR products.	Lewers et al. 2008
Eucalyptus	Leaf	8737	4941(1062+3879)	One library constructed from leaves of three	Rasmussen et al. 2008

(*E. globules*)

months old plants treated at 4°C for 30 min.

It provides an initial model for genes and regulatory pathways involved in cell-wall biosynthesis at low temperature.

Pine ( <i>Pinus radiate</i> )	Xylem tissue	6389	3304 (952+2352)	Six developing xylem cDNA libraries constructed from earlywood and latewood tissues sampled at juvenile (7 years), transition (11 years) and mature (30 years) ages, respectively; many genes involved in cell-wall biosynthesis and transcription factors in wood formation were identified; relative expression of genes at different stages of wood formation were also identified.	Li et al. 2009
Mulberry ( <i>Morus indica</i> )	Leaf	958	664 (147+517)	One library constructed from the leaf of K2 variety; obtained important ESTs like vacuolar Na <sup>+</sup> /H <sup>+</sup> antiporter, aquaporins, LEA, RD22, dehydrins and HAL3.	Lal et al. 2009
Peach ( <i>Prunus persica</i> )	Fruit	10847	6278 (1413+4865)	Five libraries were constructed from mesocarp and epicarp tissues of peach fruit at four ripening stages from Bolero, OroA and Suncrest genotypes; indicates the volatile metabolism genes in the skin of fruit.	Vecchiatti et al. 2009
Tree peony ( <i>Paeonia</i> )	Flower bud	2241	1300 (363+937)	One library was constructed from flower bud at different developmental stages;	Shu et al. 2009

*Suffruticosa*)

identified 185 putative SNPs and 176 putative SSRs.

Ginkgo  
(*Ginkgo  
biloba*)

Leaf

2039

1437 (249+1188)

One library constructed from mature leaf of female Ginkgo (Hunan Meihe cultivar) for the first time; identified many ESTs similar to photosynthesis, secondary metabolism, and stress-response.

Wang et al. 2010b

---

\*(C+S)=(contigs + singlets), nr = not reported

carboxylase, low molecular weight heat shock proteins, peroxidases, lipid transfer and metallothionein-like proteins which might play a role in the growth and development of *citrus* plants.

Cycads are ancient seed plants endured over 270-280 million years (Mamay 1969). A total of 4200 ESTs were generated from *Cycas rumphii* by Brenner et al. (2003). The ESTs were clustered into 2458 unigenes comprising 1917 contigs 541 singletons, of which 1764 had low-stringency in BLAST similarity to other plant genes. Among those cycad contigs 1718, 1310 and 734 genes were similar to angiosperms, gymnosperms and lower non seed plants, respectively. They successfully purified cycad EST for a GLR-like gene and ESTs potentially involved in the synthesis of the GLR agonist BMAA. This study of cycad EST uncovered conserved and potentially novel genes. Furthermore, the presence of a glutamate receptor agonist, and glutamate receptor-like gene in cycads, supports the hypothesis that such neuroactive plant products are not merely herbivore deterrents but may also serve a role in plant signaling.

Bhat et al. (2005) generated 500 ESTs from the leaves of low moisture-stressed plants of robusta coffee variety 'C x R' (a commercial interspecific hybrid of *Coffea congensis* and *C. canephora*) for surveying SSR motifs (minimum repeat length of 15 nt) ultimately to develop EST-SSR markers. The primer pairs were designed for 14 such identified ESTs. Thus, the validation studies demonstrated the potential utility of the EST-SSRs for genetic analysis of coffee germplasms.

Jennewein et al. (2004) generated ESTs from cultured taxus (*T. cuspidata*) cells after 16 h treatment with methyl jasmonate as an elicitor of taxol production. A total of 10176 clones were randomly sequenced that generated 8424 numbers of quality ESTs which produced 3563 unique transcripts after clustering. This library was used successfully in the isolation of several genes of taxoid metabolism including six cytochrome P450 taxoid hydroxylases and five acyl and aroyl CoA-dependent transferases.

Devitt et al. (2006) generated a total of 1171 ESTs to identify genes involved in papaya (*Carica papaya* L.) fruit ripening. The ESTs were generated from the

randomly selected clones of two independent fruit cDNA libraries derived from yellow and red-fleshed fruit varieties. Homology search of ESTs revealed significant similarity with the fruit softening, aroma and colour biosynthesis related genes. The most abundant sequences were of chitinase, ACC oxidase, CAT and methionine synthase genes.

To identify and characterize lineage-specific genes of conifers, Ujino-Ihara et al. (2005) generated two sets of ESTs from *Cupressaceae* species namely, *Cryptomeria japonica* and *Chamaecyparis obtusa* with 12791 and 5902 ESTs, respectively. After clustering the ESTs of corresponding libraries, they were collapsed into 5373 and 3018 unique transcripts, respectively. The ESTs were compared with non-redundant sets of genes generated from Pinaceae species, other gymnosperms and angiosperms. The analysis revealed that about 6% of the tentative unique genes of *C. japonica* and *C. obtusa* had homology with other conifers but not with angiosperms, and about 70% had apparent homologs in angiosperms. These studies revealed that GC content of orthologous genes are likely to be lower than those of angiosperms. Comparative results of the numbers of homologous genes in each species suggest that copy numbers of genes may be correlated between diverse seed plants. It suggests that the multiplicity of such genes may have arisen before the divergence of gymnosperms and angiosperms.

Silva et al. (2005) reported the analysis and annotation of 146075 ESTs from *Vitis* species mostly from *Vitis vinifera*. This huge number of sequences estimated 25746 unique sequences which were generated from various tissues, developmental stages and during biotic as well as abiotic stresses. Out of these sequences, 17752 unique transcripts were homolog to putative proteins and 1962 transcripts were subdivided into different functional categories. A simple structured vocabulary was developed to describe the relationship between individual EST and cDNA libraries keeping the modules for plant genotype, plant development, and stress that facilitate data mining within the context of a relational database. Searching of enzymes leading from glycolysis through the oxidative/non-oxidative pentose phosphate pathway, and into the general phenylpropanoid pathway, 65 enzymes out of 77 were identified and where 86% of enzymatic steps were represented by paralogous genes. A total of 665 differentially expressed transcripts were identified across 29 cDNA

libraries which represented a range of developmental and stress conditions. An association between plant developmental stages and tissue types with an exception of abiotic stress was also studied. In case of flower and berry development, total 87 differentially expressed transcripts were identified. This combined data set reveals hundreds of genes in *V. vinifera*, the majority of which have not been previously studied within this species.

Apple (*Malus domestica*) becomes the model fruit bearing crop plants for studying different traits such as pest resistance and abiotic stress resistance. Newcomb et al. (2006) collected more than 150000 ESTs from 43 different cDNA libraries representing 34 different tissues and treatments. These sequences were produced a set of 42938 non redundant sequences comprising 17460 tentative contigs and 25478 singletons after clustering.

Cairney et al. (2006) generated 68721 ESTs from four cDNA libraries of zygotic and somatic embryos of loblolly pine (*Pinus taeda* L.). The ESTs clustered into 5274 contigs and 6880 singlets for a total of 12154 non-redundant sequences. A total of 9189 sequences showed a putative homolog against the non-identical amino-acid database. Finally, to assess similarities between angiosperm and gymnosperm embryo development, the ESTs were examined for putative homologs of angiosperm genes implicated in embryogenesis. Out of 108 angiosperm embryogenesis-related genes, 83 homolog genes were reported in pine suggesting that pine contains similar genes for embryogenesis like other angiosperm and gymnosperm.

In order to identify genes that expressed differentially during the quiescent and active stage of bud development, Derory et al. (2006) generated 801 numbers of ESTs from three cDNA libraries of bud tissue during early, late and quiescent stages of bud burst of sessile oak (*Quercus petraea*). The ESTs produced 233 unigenes comprising 137 contigs and 96 singlets. They carried out the gene expression levels for six stages of bud burst using 233 unique genes through colony array hybridization that was evaluated by reverse-northern blotting using cDNA macroarrays. The 10 most regulated genes were quantified through real-time PCR. These differentially expressed genes were reported as relevant candidates for signaling the pathway of bud burst in trees.

Murata et al. (2006) constructed two cDNA libraries from young leaves and root tips of Madagascar periwinkle (*Catharanthus roseus*) which is well-known for the production of anticancer agents, vinblastine and vincristine. For the first time, they reported 9824 random clones which yielded 3327 related sequences and 1696 singletons after sequencing and analysis. They assigned putative functions to 3663 clones from 5023 non-redundant ESTs and identified 8 functionally characterized monoterpenoid indole alkaloids (MIA) biosynthesis genes in leaf base and root to establish a resource for transcriptome analysis and gene discovery of this important medicinal plant.

Miyama et al. (2006) reported 14842 ESTs from burma mangrove (*Bruguiera gymnorrhiza*). They constructed five different cDNA libraries from leaves and roots of 4-10 month-old seedlings either treated with NaCl, hormones, mannitol or dehydration for studying the mechanism of plants adaptation to adverse environmental conditions. After analysis, the ESTs produced 6943 numbers of unique genes. Homology searches showed that 4339 (62.5%) transcripts had a similarity with known proteins. Of these, functions of 2074 genes were assigned. On the basis of EST frequency in each cDNA library, expression profiles were generated that suggested the co-expression and organ specific gene regulation in mangrove. Further, it was believed that a specific mechanism exists in halophytes like mangrove for coping up in the adverse environmental conditions and CL3 and CL10-2 gene families were supposed to be involved in that mechanism.

Natural rubber is commercially produced from *Hevea brasiliensis*. In order to analyse the latex biosynthetic pathway, specifically rubber biosynthesis related genes, 10040 numbers of ESTs were generated (Chow et al. 2007). After assembly, the ESTs collapsed into 3441 unique transcripts including 1380 contigs and 2061 singlets, of which 73.8% showed similarity with the genes of unknown function. A significant proportion of highly abundant ESTs found homolog to proteins with rubber biosynthesis and stress or defence responses. Around 12% of the ESTs were found homolog with rubber particle membrane proteins comprising three families. The characterization of these ESTs revealed nine rubber particle membrane protein (RPMP) variants including the 14 kDa REF (rubber elongation factor) and 22 kDa SRPP (small rubber particle protein). Moreover, comparative sequence analysis revealed

that RPMP are highly similar with the sequences having stress-related functions in the plant kingdom.

Souza et al. (2007) constructed two cDNA libraries from leaf tissues of sweet orange Pêra IAC (*Citrus sinensis* L. Osb.) and mandarin 'Ponakan' (*C. reticulata* Blanco) which are highly susceptible and tolerant to *X. fastidiosa* respectively for studying the genetic responses to the presence of the said bacterium. A total of 19200 EST sequences were generated, of which 17695 met Phred quality standard score of 20 or above. After assembling, the sequences generated 7814 contigs and 4635 singlets. A total of 152 differentially expressed genes due to the infection of *X. fastidiosa* were identified through *in silico* differential display.

Cassava (*Manihot esculenta* Crantz) is a staple food and it has been increasingly used as an industrial crop for starch production. For understanding the drought tolerance in cassava, Lokko et al. (2007) generated 18166 ESTs from a normalized cDNA library prepared from dehydration stressed and well-watered tissues (control). Analysis of the ESTs produced 8577 unique gene comprising 5383 singletons and 3194 contigs. Around 63% of the unigenes were assigned to functional categories while 11% of them found homologous to hypothetical genes with unclear functions. The rest 26% were considered as novel transcripts. The library uncovered numerous ESTs with recognized roles in drought-responses, including those that encode LEA proteins thought to confer osmoprotective functions during water stress, transcription factors, heat-shock proteins as well as proteins involved in signal transduction and oxidative stress. The unigenes were screened for short tandem repeats for further development as microsatellite markers and finally identified a total of 592 clusters contained 646 repeats representing 3.3% of the ESTs queried. This is the first transcriptome studies in cassava for drought tolerance and the unigene set can be utilized for the development of microarrays as well as for gene-derived molecular markers to further dissect the basis of drought tolerance in cassava.

Lewers et al. (2008) reported the first cDNA library of blackberry (*Rubus* L.) constructed from expanding leaf tissue. A total of 18432 clones were produced from Merton Thornless cultivar, a progenitor of many thornless commercial cultivars. Finally, 3000 genes were reported that assembled into 301 contigs and 2081

singletons. The ESTs were mostly involved with energy, metabolism, cells structure and defense metabolism. A total of 673 numbers of putative SSRs were detected in the EST set, of which 33 SSRs tested for amplification in two blackberry cultivars. Ten of them detected an average of 1.9 polymorphic PCR products.

Rasmussen et al. (2008) reported 8737 ESTs from the most important biofuel plant, *Eucalyptus globules* (Jones et al. 2002, Grattapaglia 2004). A cDNA library was constructed from *E. globules* seedlings subjected to low temperature stress. Finally, 9913 randomly selected clones generating 8737 curated ESTs. The sequence assembly produced 1062 contigs and 3879 singleton forming a unigene set of 4941. After BLASTx analysis, 89.3% of the contigs and 88.5% of the singleton revealed similarity to known genes.

Tree peony (*Paeonia suffruticosa* Andrews), a woody deciduous shrub is an important medicinal and ornamental plant. Shu et al. (2009) constructed a cDNA library from flower buds of different developmental stages. A total of 2241 ESTs have been generated through unidirectional sequencing that resulted in a total of 1300 unigenes comprising 363 contigs and 937 singletons. The unigenes were categorized into 13 broad families with biological roles according to the similar functional characteristics. Finally, a total of 185 putative single nucleotide polymorphism and 176 numbers of putative EST-SSR were identified.

Vecchiatti et al. (2009) reported five different cDNA libraries of peach (*Prunus persica* L. Batsch) constructed from two different fruit tissues (skin and mesocarp). A total of 10847 high quality ESTs were generated which produced 6278 numbers of unique transcripts including 1413 contigs. Finally, 21858 numbers of ESTs from the public database were analysed alongwith their in-house ESTs that produced 17858 unique transcripts, of which 70.8% was assigned to putative functions. The analysis of the frequency of genes putatively involved in the metabolism of some volatiles revealed a predominant presence of those in the skin of peach fruit.

*Pinus radiata* D. Don is a commercial plantation tree in Australia and several other countries; however, there are a very less resources in the genomic database. Li et al. (2009) constructed six xylem cDNA libraries from the earlywood and latewood

tissues sampled at juvenile (7 years), transition (11 years) and mature (30 years) ages, respectively. They generated a total of 6389 numbers of high quality ESTs for studying the genes involved in wood formation. After assembly, a total of 3304 unigenes including 952 contigs and 2352 singletons was produced. It was found that about 97.0% ESTs and 96.1% of the unigenes have matches in the UniProt and TIGR databases. More than half of the 5952 ESTs have matches in the Pfam database and represent 772 known protein families. About 18.0% of the 5952 ESTs matched cell wall related genes in the MAIZEWALL database, representing all 18 categories, 91 of all 174 families and possibly 557 genes. The cell wall-related genes and transcription factors were identified.

Mulberry (*Morus indica*) is the host plant of silkworm. Hence, the elite clone of the species is vital for the sericulture industry. However, there is a lack of molecular information for improving the plant through genetic engineering. Lal et al. (2009) reported the construction of a cDNA library from the mature leaf of mulberry based on Universal PCR amplification method. A total of 1500 randomly selected clones were sequenced from the leaf cDNA library of *M. indica* var. K2. Further the ESTs were assigned to functional categories with protein metabolism (15%), transport (8%), stress related proteins (9%), energy (5%), and photosynthesis (4%). Around 22% of the total ESTs did not show any significant homology with known proteins. This study revealed a large number of genes encoding enzymes involved in various secondary metabolites in mulberry.

Senthil et al. (2010) reported 1047 and 1034 numbers of ESTs from leaves and roots of *Withania somnifera* (Ashwgandha), respectively. The libraries were constructed from an *in vitro* cultured 2-months-old plant. The ESTs represented 48.5% and 61.5% unique sequences, respectively comprising 239 and 230 contigs. The unigenes were functionally classified, of which about 70% encoded proteins found similar to characterized or annotated proteins. The identified genes were of mostly photosynthesis (cytochrome p-450), pathogenesis (arginine decarboxylase, chitinase) and withanolide biosynthesis (squalene epoxidase, CDP-ME kinase) related. The protein domain analysis revealed the presence of functional domains in the selected sequences. This comparative expression studies provide a framework for future

research in proteomics and evolutionary genomics for the withanolide biosynthesis in plants.

Ginkgo (*Ginkgo biloba* L.) belongs to the oldest group of living seed plants. Wang et al. (2010b) constructed a cDNA library from the mature leaf of female Ginkgo (Hunan Meihe cultivar) for the first time. A total of 2039 EST sequences were generated which represented 1437 unigenes consisting of 249 contigs and 1188 singletons. Out of 2039 (60.1%), 1235 numbers of ESTs were assigned to putative functions, and the remaining 804 were not similar to any known sequences. Transcripts related to photosynthesis, disease/defense, flavonoid biosynthesis, pathogenesis as well as secondary metabolism and stress responses were recognised. These results provided new information about the mature leaf-specific transcripts of Ginkgo.

### 3.1.6 Transcriptome analysis of tea

Molecular biology works on tea is very little in comparison to other woody crop plants and cereals. There is only around 14000 sequence information (on July, 2010) in NCBI. However, the number is still not enough to understand and decipher the complex mechanism of stress responses and tolerance, growth and development of this highly cross pollinated plant. Nevertheless, a brief review of different transcriptome analysis works of tea is given here.

Park et al. (2004) for the first time applied SSH technique in *Camellia sinensis* to identify the genes involved in secondary metabolism using young and mature leaf cDNA as 'tester' and 'driver', respectively. A total of 588 cDNA clones were sequenced and found that 8.7% of the clones encoded enzymes involved in secondary metabolism, mostly of flavonoid-metabolism pathways (5.1%) which included chalcone synthase (CHS), flavanone 3-hydroxylase (F3H), flavonoid 3' 5'-hydroxylase (F3'5'H), flavonol synthase (FLS), dihydroflavonol 4-reductase (DFR) and leucoanthocyanidin reductase (LCR). It was confirmed that flavonoid biosynthesis genes were differentially regulated in the developmental stages of tea.

Chen et al. (2005) constructed a cDNA library from the tender shoots of Chinese elite green tea cultivar 'Longjing 43'. A total of 4320 clones were randomly picked

and sequenced which finally produced 1684 numbers of quality ESTs after removal of ribosomal, mitochondrial, chloroplast and poor quality sequences. More than 50% of these ESTs found homolog with model plants such as *Arabidopsis*, rice and tobacco. The transcripts which identified with known or putative functions were assigned functional categories as described in *Arabidopsis*.

To identify the drought-modulated ESTs, Sharma and Kumar (2005) conducted a differential display reverse transcriptase PCR experiment. Tea cultivar, T-78 of Darjeeling was used for drought induction in a greenhouse condition and subsequently collected leaf sample for differential display and northern analysis on d 14 and 18. Three drought modulated ESTs from three clones namely *dr1*, *dr2* and *dr3* were identified and validated as: *dr1* (drought responsive), induced only by drought but not by ABA, showed significant scores with PR-5 (pathogenesis related) family of PR-protein gene; *dr2*, repressed by drought but not by ABA, had nucleotide repeats for polyaspartate which was also present in chicken calsequestrin-like mRNA; *dr3*, responded similarly as *dr2* but did not show similarity to known genes.

Wang et al. (2008) constructed a SSH library to identify genes induced in response to low temperature of tea (*C. sinensis* L. cv. Xinyangdaye) leaves using cold treated cDNA as 'tester' and control (without treatment) sample cDNA as 'driver'. For isolating cold induced gene, they applied reverse-northern analysis to screen the subtracted library. The subtracted cDNA was cloned and PCR products of 420 randomly selected clones were arrayed on nylon membrane. A total of ten complementary cDNA clones were induced by low temperature. These transcripts were classified into three functional groups: transcription factor, structural protein and metabolism related. The final interpretation of results gave an indication that cold might comprehensively activate the genes in the flavonoid pathway more than the genes specific to anthocyanin biosynthesis.

A full-length cDNA library was constructed from young roots sample by Zhao et al. (2008b). A total of 5115 clones sequenced which produced 3482 numbers of unigenes including 901 TC and 2581 singletons. There were 1376 (39.5%) ESTs with known functions which subsequently categorized according to their functions assigned in *Arabidopsis*.

Moreover, there are some sequences in dbEST, NCBI (<http://www.ncbi.nlm.nih.gov/dbEST>) without available literature. These include 407 differentially expressed ESTs generated from the leaves of tea plant induced by *Ectropis oblique* feeding using SSH technique, 66 differentially expressed ESTs generated from the root tissue induced by *Ectropis oblique* feeding using cDNA-AFLP technique, 250 ESTs generated for investigating the influence of light on the pathway of phenylpropanoid biosynthesis using cDNA-AFLP, 110 ESTs for transcriptome analysis of Assam tea, 3201 ESTs generated from Assam and Darjeeling tea for analyzing transcripts related to drought, flavor and blister blight infection.

## **3.2 Materials and methods**

In an attempt to identify differentially expressed transcriptome under drought and understanding the associated tolerance mechanism thereof, three drought stressed SSH cDNA libraries comprising one inter-varietal and two intra-varietal as well as one standard full length libraries were constructed. The intra-varietal libraries were each of TV-23 (drought-tolerant) and S.3/A3 (drought-susceptible) cultivars, constructed by taking the individual drought induced plant cDNAs as 'tester' (sample containing the differentially expressed transcripts which to be generated) and control plant cDNA as 'driver' (containing the reference transcripts which to be eliminated). However, the inter-varietal library was constructed taking the drought induced TV-23 plant cDNA sample as 'tester' and S.3/A3 plant cDNA sample as 'driver' for generating those transcripts that expressed only in the tolerant cultivar under drought stress. A full-length standard library was also constructed from the cDNA sample of control plant of TV-23 cultivar for using as reference.

### **3.2.1 Sample preparation**

Fresh tender roots of control and drought induced plants of TV-23 and S.3/A3 cultivars were collected by carefully uprooting the plants on 21<sup>st</sup> d of drought stress induction based on physiological experiments (as described in 2.2). Collected roots were immediately packed individually for each sample in aluminium foil and immersed in liquid nitrogen. Packed samples were stored in -80<sup>o</sup>C until RNA extraction. The whole process was done immediately and carefully to avoid RNA degradation.

### **3.2.2 Protocol optimization for RNA isolation**

Several published protocol and commercial kits were available for RNA isolation from different plant tissues. However, there was no reported protocol for RNA isolation from tea roots. Therefore, RNA isolation was attempted with conventional guanidine-HCl (Longemann et al. 1987), modified guanidine-HCl (Lal et al. 2001), modified CTAB (Gasic et al. 2004) and phenol/SDS (Alemzadeh et al. 2005) based protocol as well as commercial kits such as RNaquous kit (Ambion, USA) and TRIzol

reagent (Invitrogen, USA) for searching the suitable one. Finally, changes were made in some critical steps of SDS based protocols (Alemzadeh et al. 2005, Ikoma et al. 1996) and adopted for the isolation of tea roots RNA. The changes were addition of  $\beta$ -mercaptoethanol as an anti-oxidative agent in the extraction buffer in a right proportion (Lal et al. 2001), precipitation of RNA using isopropanol in addition to acetate (Jaiprakash et al. 2003) instead of lithium chloride (LiCl) which have a tendency for precipitating larger fragments (Basia and Arie 2005), incorporation of DNaseI treatment and purification as well as reduction of isopropanol precipitation period effectively to 30 min.

### 3.2.3 Isolation of total RNA

Total RNA was isolated from 100 mg stored tissue of each control and drought induced samples of both the cultivars. The tissue was grounded in liquid nitrogen and transferred to 800  $\mu$ l of extraction buffer [100 mM Tris-HCl (pH 8.0), 10 mM EDTA (pH 8.0), 100 mM LiCl, 1% SDS and 200 mM  $\beta$ -mercaptoethanol (added just before use)] in an eppendorf tube at room temperature. Tissue was homogenized completely, added 400  $\mu$ l of cold PCI (25:24:1, v/v/v) within 3-5 min of homogenization, mixed by inverting the tube gently for 4-6 times and incubated in ice for 5 min. Tubes were centrifuged at 14000 rpm for 15 min at 4<sup>o</sup>C. The supernatant was taken and added equal vol of cold PCI (25:24:1, v/v/v) as well as mixed by gently inverting the tube for 4-6 times. The tubes were centrifuged at 14000 rpm for 10 min at 4<sup>o</sup>C. The supernatant was collected, added an equal vol of isopropanol and 0.1 vol of sodium acetate (pH 5.2) and subsequently mixed by gentle inversion of the tube. The tubes were incubated at -80<sup>o</sup>C for 30 min and later centrifuged at 14000 rpm for 10 min at 4<sup>o</sup>C. The pellet was taken and added 500  $\mu$ l of 70% ethanol for washing the pellet by inverting the tube for 4-6 times followed by centrifugation at 14000 rpm for 10 min at 4<sup>o</sup>C. The ethanol was decanted and evaporated the residual portion by incubating the tube at 37<sup>o</sup>C for 30 min. The pellet was dissolved in 80  $\mu$ l of DEPC treated autoclaved distilled water and incubated the tube in ice for 10 min for dissolving the pellet completely. For DNase treatment, 10  $\mu$ l of RNase-free DNase I (1 u/ $\mu$ l) and 10X reaction buffer with MgCl<sub>2</sub> were added. The reagents were mixed by gentle inversion and made short spin. The tubes were incubated at 37<sup>o</sup>C for 15 min, and later made up the vol upto 500  $\mu$ l

using DEPC treat autoclaved distilled water. An equal amount of PCI (25:24:1, v/v/v) was added and mixed by inverting the tube gently. The tubes were centrifuged at 14000 rpm for 10 min at 4°C and subsequently repeated the isopropanol precipitation as well as 70% ethanol washing steps for once. The pellet was air dried and dissolved in 30 µl of RNA storage buffer or nuclease free water. It was then stored at -80°C for further use.

#### 3.2.4 Quantity and quality checking of RNA

The isolated RNA was resolved in a denaturing agarose gel (1%) using 1X TBE buffer for 45 min at 70 V (# 164-0300, mini gel apparatus, biorad), and later stained with ethidium bromide (10 mg/ml) for checking its integrity. For running, 5 µl was used from each sample with an equal amount of RNA gel loading dye (ambion). The quality and quantity were verified from 1 µl of the isolated RNA using Biophotometer plus (eppendorf) at an absorbance of 260, 280 and 320 nm. Quality samples were selected for using in further steps on the basis of its integrity and purity.

#### 3.2.5 Isolation of mRNA

Intact and pure poly A<sup>+</sup> RNA is essential for the synthesis of high quality cDNA. In this experiment, mRNA was isolated from each total RNA sample (individual extractions were pooled) using PolyATract mRNA isolation system (# Z5300, promega). There was a biotinylated oligo (dT) primer to hybridize at high efficiency in solution to the 3' poly (A) region present in mature mRNAs. The hybrids were captured and washed at high stringency using streptavidin coupled to paramagnetic particles and a magnetic separation stand. mRNA was eluted from the solid phase by simple addition of ribonuclease-free deionized water. In a sterile, RNase-free 1.5 ml eppendorf tube, total RNA (1.0 mg) was taken and adjusted the final vol to 500 µl with RNase-free water. Tube was placed in a 65°C heating block for 10 min and subsequently added 3 µl of biotinylated-oligo (dT) probe and 13 µl of 20X SSC to the RNA solution. Solution was mixed gently and incubated at room temperature for cooling. A single tube (0.6 ml vol) of the streptavidin-paramagnetic particles (SA-PMPs) was taken for each sample. The particles were resuspended by gently flicking the bottom of the tube and captured them on the wall by placing in a magnetic

stand (~ 30 s). The supernatant was removed carefully. The SA-PMPs were washed for three times with 0.5X SSC (300 µl per wash), in each time they were captured using the magnetic stand and carefully removed the supernatant. The washed SA-PMPs were resuspended in 100 µl of 0.5X SSC. The entire content of the annealing reaction were transferred to the tube containing the washed SA-PMPs and mixed gently. The tube incubated at room temperature for 10 min and mixed gently by inverting in every 2 min. The SA-PMPs were captured using the magnetic stand and removed the supernatant carefully without disturbing the SA-PMP pellet. The particles were washed for four times with 0.1X SSC (300 µl per wash) by gently flicking the bottom of the tube until all of the particles get resuspended. After the final wash, as much of the supernatant as possible was removed without disturbing the SA-PMP particles. Finally, SA-PMP pellet was resuspended in 100 µl of the RNase-free water by gently flicking the tube. The SA-PMPs particles were captured magnetically and transferred the eluted mRNA to a sterile, RNase-free tube. The elution step was repeated by resuspending the particles in 100 µl of RNase-free water.

### **3.2.6 Precipitation of mRNA**

The mRNA fraction isolated with the PolyAtract system was too diluted for the next step of cDNA synthesis in SSH library construction process. Hence, the mRNA pool was precipitated with 0.1 vol of 3 M ammonium acetate (pH 5.2) and 1.0 vol of isopropanol by incubating at -70°C for 3 h. Then, the tube was centrifuged at 12000 rpm for 15 min at 4°C. The pellet was resuspended in 1 ml of 75% ethanol and centrifuged again at 12000 rpm for 10 min at 4°C. Supernatant was removed carefully and air dried the tube. Finally, the pellet was resuspended in 10 µl of RNase-free, deionized water and stored at -70°C for further use.

### **3.2.7 Determination of mRNA quantity and purity**

Quantity and purity of the concentrated mRNA was determined by Biophotometer plus. For analysis, 1 µl of the isolated mRNA was diluted with 49 µl of RNase free water of which absorbance was read at 260, 280 and 320 nm taking RNase free water as blank following the program of RNA considering optical density 1 equal to 40 µg of

RNA and expressed as  $\mu\text{g}/\mu\text{l}$ . Finally, the mRNA was diluted to a final concentration of  $1 \mu\text{g}/\mu\text{l}$  with RNase-free, deionized water.

### **3.2.8 Construction of drought stressed cDNA libraries**

For the construction of three drought stressed cDNA SSH libraries, the isolated mRNAs were transcribed to cDNA using AMV reverse transcriptase. The first strand cDNA was further converted to double-strand using an enzyme cocktail, and subsequently constructed the drought stressed SSH libraries following the user manual of PCR-select cDNA subtraction kit (# 637401, clontech, USA) (Appendix-B 1).

#### **3.2.8.1 First-strand cDNA synthesis**

For each sample, two individual reactions were performed. The following components were combined in a sterile 0.5 ml microcentrifuge tube for each reaction:  $4 \mu\text{l}$  of poly A<sup>+</sup> RNA ( $2 \mu\text{g}$ ) and  $1 \mu\text{l}$  of cDNA synthesis primer ( $10 \mu\text{M}$ ). The contents were mixed gently and made a brief centrifugation. Tubes were incubated at  $70^\circ\text{C}$  for 2 min in a thermal cycler followed by 2 min in ice and brief centrifugation. Afterwards, in each reaction  $2 \mu\text{l}$  of 5X first-strand buffer (250 mM Tris-HCl pH 8.5, 40 mM  $\text{MgCl}_2$ , 150 mM KCl and 5 mM DTT) and  $1 \mu\text{l}$  each of dNTP mix (10 mM each), sterile water, AMV reverse transcriptase ( $20 \text{ u}/\mu\text{l}$ ) were added. The reagents were mixed gently, centrifuged briefly and subsequently incubated at  $42^\circ\text{C}$  for 1.5 h in an air incubator for the synthesis of cDNA. The reaction was terminated by incubating the tube in ice and immediately proceeds with the second-strand synthesis reaction.

#### **3.2.8.2 Second-strand cDNA synthesis**

For the synthesis of second-strand cDNA, the following components were added to the tube containing  $10 \mu\text{l}$  of first-strand cDNA according to the order below:

	per rxn
Sterile H <sub>2</sub> O	48.4 $\mu$ l
5X second-strand buffer 1	6.0 $\mu$ l
dNTP mix (10 mM)	1.6 $\mu$ l
20X second-strand enzyme cocktail	4.0 $\mu$ l
-----	
Final vol	80 $\mu$ l

Second-strand 5X buffer 1 contains KCl (500 mM), ammonium sulfate (50 mM, MgCl<sub>2</sub> (25 mM),  $\beta$ -NAD (0.75 mM), Tris-HCl (100 mM pH 7.5) and BSA (0.25 mg/ml) as well as enzyme cocktail containing DNA polymerase I (6 u/ $\mu$ l), RNase H (0.25 u/ $\mu$ l) and *E. coli* DNA ligase (1.2 u/ $\mu$ l). Reagents were mixed gently followed by a brief centrifugation. The tubes were incubated at 16°C for 2 h in a water bath, and later added 2  $\mu$ l of T4 DNA polymerase (3 u/ $\mu$ l). Reactions were incubated at 16°C for 30 min in a water bath or thermal cycler, and finally terminated by 4  $\mu$ l of 20X EDTA/glycogen mix (0.2 M EDTA, 1 mg/ml glycogen). Afterwards, 100  $\mu$ l of PCI (25:24:1, v/v/v) was added and mixed by gentle inversion of the tubes followed by centrifugation at 14000 rpm for 10 min at room temperature. Top aqueous layer was collected carefully and placed in a fresh 0.5 ml microcentrifuge tube. Then, 100  $\mu$ l of CI (24:1, v/v) was added in each tube, and repeated the centrifugation and subsequent steps. To the collected solution, 40  $\mu$ l of 4 M ammonium acetate and 300  $\mu$ l of 95% ethanol were added. The components were mixed by gentle inversion of the tubes for 3-5 times. Mixed content was centrifuged at 14000 rpm for 20 min at room temperature. Supernatant was collected carefully without disturbing the pellets. The pellet was overlaid with 500  $\mu$ l of 80% ethanol and later centrifuged at 14000 rpm for 10 min. Supernatant was removed and air dried the pellet for about 10 min to evaporate the residual ethanol. Finally, it was dissolved in 50  $\mu$ l of sterile distilled water, and subsequently 6  $\mu$ l kept aside in a fresh tube for checking the yield and size ranges of dsDNA and rest amount stored at -20°C for further use.

### 3.2.8.3 Restriction digestion of complementary dsDNA

Synthesized dsDNAs were digested with RsaI restriction enzyme for generating shorter, blunt-ended fragments. This blunt ended cDNA fragments were required for adaptor ligation which is optimal for further steps of subtraction process. For each reaction of restriction digestion, 43.5  $\mu\text{l}$  of dsDNA, 5.0  $\mu\text{l}$  of 10X RsaI restriction buffer and 1.5  $\mu\text{l}$  of RsaI (10  $\text{u}/\mu\text{l}$ ) components were added. The reagents were mixed well, centrifuged briefly and incubated at 37°C for 1.5 h. Then, 5  $\mu\text{l}$  of the digested mixture was set aside from each reaction tube for analyzing the digestion efficiency. In the rest amount, 2.5  $\mu\text{l}$  of 20X EDTA/glycogen mix was added to terminate the reaction, and later added 50  $\mu\text{l}$  of PCI (25:24:1, v/v/v). The components were mixed gently followed by centrifugation at 14000 rpm for 10 min at room temperature. Top aqueous layer was collected carefully in a fresh 0.5 ml tube, added 50  $\mu\text{l}$  of PCI (24:1, v/v), and subsequently repeated the previous steps of centrifugation as well as collection. In the collected solution, 25  $\mu\text{l}$  of 4 M ammonium acetate and 187.5  $\mu\text{l}$  of 95% ethanol were added as well as mixed by gentle inversion for 3-5 times. Tubes were centrifuged at 14000 rpm for 10 min at room temperature. Supernatant was removed carefully and overlaid with 200  $\mu\text{l}$  of 80% ethanol for washing. Tubes were centrifuged at 14000 rpm for 5 min and carefully removed the supernatant. Pellets were air dried for 5-10 min and later dissolved in 5.5  $\mu\text{l}$  of deionised sterile water. Final solution of two individual isolation was pooled up and finally stored at -20°C for further use.

### 3.2.8.4 Evaluation of size ranges of the synthesized dsDNAs

For checking the size ranges, 2.5  $\mu\text{l}$  of the undigested and 5  $\mu\text{l}$  of RsaI restriction digested dsDNAs were resolved with 1 kb and 100 bp markers (each of 300 ng) side-by-side on an agarose gel (1%) using 1X TAE buffer, run for 40 min at 70 V. The gel was stained with ethidium bromide (10 mg/ml) for 15 min and visualized under UV light in a gel documentation system (Gel Doc XR+, biorad).

### 3.2.8.5 Adaptor ligation to the restriction digested 'tester' dsDNAs

Forward subtraction experiment was designed to enrich the differentially expressed sequences in the cDNA sample of 'tester' but not in the 'driver'. For achieving this, two separate adaptor ligation reactions were performed for each experimental 'tester' cDNAs: one with adaptor 1 (tester 1-1) and the other with adaptor 2R (tester 1-2). After setting the ligation reactions, portion of each reaction were combined (unsubtracted tester control 1-C) so that both adaptors would be ligated to cDNAs. Each unsubtracted tester control cDNA was served as a positive control for ligation and later served as a negative control for subtraction. The adaptors i.e adaptor 1 (5'-CTAATACGACTCACTATAGGGCTCGAGCGGCCGCCCGGGCAGGT-3') and adaptor 2R (5'-CTAATACGACTCACTATAGGGCAGCGTGGTTCGCGGCCGAGGT-3') have the binding site for T7 primer (5'-CTAATACGACTCACTATAGGGC-3') as well its complementary sequences give rise to the binding sites for PCR primer 1 (5'-CTAATACGACTCACTATAGGGC-3'), nested PCR primer 1 (5'-TCGAGCGGCCGCCCGGGCAGGT-3') and nested PCR primer 2R (5'-AGCGTGGTTCGCGGCCGAGGT-3'). Adaptor ligation was not required for 'driver' cDNA.

A ligation master mix was prepared in a 0.5 ml microcentrifuge tube combining 3  $\mu$ l of sterile water, 2  $\mu$ l of 5X ligation buffer and 1  $\mu$ l of T4 DNA ligase (400 u/ $\mu$ l) for each reaction. RsaI digested experimental cDNAs (1  $\mu$ l of each sample) were diluted with 5  $\mu$ l of sterile water. Following reagents were combined according to the given order in a 0.5 ml tube and mixed thoroughly by pipetting up and down for each 'tester' dsDNA sample.

Components	tester 1-1	tester 1-2
Diluted tester cDNA	2 $\mu$ l	2 $\mu$ l
Adaptor 1 (10 $\mu$ M)	2 $\mu$ l	-
Adaptor 2R (10 $\mu$ M)	-	2 $\mu$ l
Master mix	6 $\mu$ l	6 $\mu$ l
-----		
Final vol	10 $\mu$ l	10 $\mu$ l

Reagents were mixed well, and subsequently combined 2  $\mu$ l from each of tester 1-1 and tester 1-2 for using as unsubstracted tester control (1-C). Later, 1  $\mu$ l of unsubstracted tester control sample was diluted to 1 ml with sterile water and stored at  $-20^{\circ}\text{C}$  for further use. Rest amount of the reactions was incubated at  $16^{\circ}\text{C}$  for overnight ligation. Reaction was stopped by adding 1  $\mu$ l of 20X EDTA/glycogen mix. For inactivating the ligase, the samples were heated at  $72^{\circ}\text{C}$  for 5 min followed by a brief centrifugation, and later stored at  $-20^{\circ}\text{C}$ .

### 3.2.8.6 Analysis of ligation efficiency

Ligation efficiency was analyzed through a PCR experiment. The experiment was designed to amplify the adaptor/cDNA junction fragments of tester 1-1 and tester 1-2 where 26S rRNA, a housekeeping gene was taken as experimental control. The 26S rRNA gene is known to be consistently expressed regardless to tissue types or tissue condition in plants. The specific primers of 26S rRNA gene were synthesized according to an earlier report in tea (Cho et al. 2007) as: forward primer 5'-ATGAGTAGGAGGGCGCGGCGGT-3' and reverse primer 5'-GGAGGCACTCGGTCCTCCGGAT-3'. The adaptor ligated dsDNA (1  $\mu$ l of each) was diluted into 200  $\mu$ l of sterile water. The following reagents were combined in four separate tubes as follows:

Components	Tubes ( $\mu$ l)			
	1	2	3	4
Tester 1-1	1	1	-	-
Tester 1-2	-	-	1	1
26S rRNA forward primer (10 $\mu$ M)	1	1	1	1
26S rRNA reverse primer (10 $\mu$ M)	-	1	-	1
PCR primer 1 (10 $\mu$ M)	1	-	1	-
Total vol	3	3	3	3

A master mix was prepared by mixing 18.5  $\mu$ l of sterile water, 2.5  $\mu$ l of 10X PCR reaction buffer, 0.5  $\mu$ l each of dNTP mix (10 mM) and 50X advantage cDNA polymerase mix in a total vol of 22.0  $\mu$ l for each reaction. In the above reaction

tubes, 22  $\mu$ l of master mix was added and mixed by brief centrifugation. The solution was overlaid by mineral oil and incubated at 37°C for 5 min in a thermal cycler to extend the adaptors, and immediately followed by thermal cycling in the following program: 94°C for 4 min followed by 30 cycles at 94°C for 30 s, 55°C for 30 s and 72°C for 1 min with a final extension at 72°C for 5 min. Final product (5  $\mu$ l from each reaction) was analysed on a 2.0% agarose gel run in 1X TAE buffer at 70 V for 50 min and later stained with ethidium bromide. For size determination, 1 kb and 100 bp DNA markers (300 ng each) was loaded side-by-side to the samples.

### 3.2.8.7 PCR select subtraction of differentially expressed transcripts

For the subtraction of differentially expressed transcripts, two hybridization reactions were performed. In the first, an excess of 'driver' cDNA was added to each sample of 'tester' cDNA and subsequently heat denatured and allowed to anneal. The remaining ss cDNAs (available for the second hybridization) were dramatically enriched for differentially expressed sequences because non-target cDNAs present in the 'tester' and 'driver' cDNA form hybrids.

For conducting the first hybridization, 4X hybridization buffer was allowed to warm up at room temperature for 20 min. The following reaction was performed for each forward subtraction in a 0.5 ml tube:

Components	Hybridization samples	
	H1	H2
Rsal digested driver cDNA	1.5 $\mu$ l	1.5 $\mu$ l
Adaptor 1 ligated tester 1-1	1.5 $\mu$ l	-
Adaptor 2R ligated tester 1-2	-	1.5 $\mu$ l
4X hybridisation buffer	1.0 $\mu$ l	1.0 $\mu$ l
-----		
Final vol	4.0 $\mu$ l	4.0 $\mu$ l

Samples were overlaid with one drop of mineral oil, centrifuged briefly and incubated at 98°C for 1.5 min followed by 68°C for 10 h.

In the second hybridization, the two samples from the first hybridization (H1 and H2) were mixed together and fresh driver DNA was added to further enrich the differentially expressed sequences. The following reagents were added into a sterile tube separately for each reaction:

Components	amount
Driver cDNA	1 $\mu$ l
4X Hybridisation Buffer	1 $\mu$ l
Sterile H <sub>2</sub> O	2 $\mu$ l
-----	
Final vol	4 $\mu$ l

From the above mixer, 1  $\mu$ l was placed in a 0.5 ml microcentrifuge tube and overlaid with 1 drop of mineral oil. Tube was incubated at 98<sup>o</sup>C for 1.5 min in a thermal cycler for denaturing the driver dsDNA. Now, the micropipette was set at 15  $\mu$ l and gently touched the tip to the mineral oil/sample interface of the tube containing hybridization sample H2. Entire sample was drawn up and a small amount of air into the pipette tip. Above procedure was repeated for the tube containing the freshly denatured driver. Entire sample was transferred to the tube containing the hybridization sample H1 and mixed by pipetting up and down. Tube was centrifuged briefly and incubated at 68<sup>o</sup>C for overnight. Mixture was diluted by adding 200  $\mu$ l of dilution buffer. Tube was heated at 68<sup>o</sup>C for 7 min in a thermal cycler and later stored the sample at -20<sup>o</sup>C.

### 3.2.8.8 Amplification of differentially expressed transcripts

Differentially expressed cDNAs were selectively amplified through two rounds of PCR. Prior to thermal cycling, the missing strands of the adaptors were filled in by a brief incubation at 75 <sup>o</sup>C. In the first amplification (primary PCR), only dsDNAs with different adaptor sequences on each end were exponentially amplified. In the second amplification (secondary PCR), nested PCR was used to further reduce background and enrich for differentially expressed sequences.

Primary PCR was performed from 1  $\mu$ l of each diluted cDNA of subtracted and the corresponding unsubtracting tester control (1-C). A master mix was prepared by mixing 19.5  $\mu$ l of sterile water, 2.5  $\mu$ l of 10X PCR reaction buffer, 0.5  $\mu$ l of dNTP mix (10 mM), 1.0  $\mu$ l of PCR primer 1 (10  $\mu$ M) and 0.5  $\mu$ l of 50X advantage cDNA polymerase mix in a final vol of 24.0  $\mu$ l. The reagents were mixed by gentle pipetting followed by brief centrifugation. In each of the reaction tube, 24  $\mu$ l of master mix was added and overlaid with 50  $\mu$ l of mineral oil. The reaction mixer was incubated at 75<sup>0</sup>C for 5 min in a thermal cycler and immediately followed the program as: 94<sup>0</sup>C for 25 s for denaturation, either 25, 27 or 30 cycles as 94<sup>0</sup>C for 10 s, 66<sup>0</sup>C for 30 s and 72<sup>0</sup>C for 1.5 min. Final products were stored at -20<sup>0</sup>C.

Primary PCR products (3  $\mu$ l) were diluted into 27  $\mu$ l of sterile water. A secondary PCR was performed for each of the diluted primary PCR products taking 1  $\mu$ l as template. A master mix was prepared by mixing 18.5  $\mu$ l of sterile water, 2.5  $\mu$ l of 10X PCR reaction buffer, 1.0  $\mu$ l of nested PCR primer 1 (10  $\mu$ M), 1.0  $\mu$ l of nested PCR primer 2R (10  $\mu$ M), 0.5  $\mu$ l of dNTP mix (10 mM) and 0.5  $\mu$ l of 50X advantage cDNA polymerase mix and subsequently added 24  $\mu$ l mix to each reaction. One drop of mineral oil was overlaid to each tube. The reaction mix was incubated at 75<sup>0</sup>C for 5 min in the thermal cycler to extend the adaptors. Then, PCR was run immediately for either 9, 11 or 12 cycles using the cycling program as: 94<sup>0</sup>C for 10 s, 66<sup>0</sup>C for 30 s and 72<sup>0</sup>C for 1.5 min. Final products were stored at -20<sup>0</sup>C for further use.

### **3.2.8.9 Gel analysis of PCR products**

The primary (1<sup>0</sup>) and secondary PCR (2<sup>0</sup>) products (8  $\mu$ l from each) were resolved in an agarose gel (2.0%) using 1X TAE buffer for 50 min at 70 V, and later stained with ethidium bromide (10 mg/ml) for visualizing under UV light.

### **3.2.8.10 PCR analysis of subtraction efficiency**

Subtraction efficiency was estimated by PCR analysis comparing the abundance of known cDNAs before and after subtraction. This was done by comparing with the 26S rRNA gene. The following reagents were combined in a 0.5 ml tubes.

	rxn 1	rxn 2
Diluted subtracted cDNA (2 <sup>0</sup> PCR product)	1.0 µl	-
Diluted unsubtractd tester control 1-C (2 <sup>0</sup> PCR product)	-	1.0 µl
26S rRNA forward primer (10 µM)	1.2 µl	1.2 µl
26S rRNA reverse primer (10 µM)	1.2 µl	1.2 µl
Sterile H <sub>2</sub> O	22.4 µl	22.4 µl
10X PCR reaction buffer	3.0 µl	3.0 µl
dNTP mix (10 mM)	0.6 µl	0.6 µl
50X advantage cDNA polymerase mix	0.6 µl	0.6 µl
-----		
Total Vol	30.0 µl	30.0 µl

Reagents were mixed well followed by a brief centrifugation. The solution was overlaid with one drop of mineral oil. The PCR was run for 18 cycles following the cycling program as: 94<sup>0</sup>C for 30 s, 60<sup>0</sup>C for 30 s and 68<sup>0</sup>C for 2 min. From each reaction tube, 5 µl was removed and put the rest back to PCR for running 5 additional cycles. This step was repeated upto 33 cycles in an interval of 5 cycles. Later, each aliquot of 5 µl PCR product was resolved on an agarose gel (2.0%) using 1X TAE buffer for 50 min at 70 V and finally stained with ethidium bromide for visualizing under UV light.

### 3.2.8.11 Cloning of differentially expressed transcripts

Cloning vector, pGEM-T easy (# A3610, promega) was selected for the cloning of differentially expressed transcripts. For vector ligation of subtracted cDNA, the tubes containing pGEM-T easy vector, control inserts DNA of pGEM®-*luc* vector (4 ng/µl) and subtracted cDNA were centrifuged briefly. The 2X rapid ligation buffer was then mixed vigorously by vortexing. Reaction set up was performed as follows:

Components	Reactions		
	sample	positive	negative
2X rapid ligation buffer	5 $\mu$ l	5 $\mu$ l	5 $\mu$ l
pGEM-T easy vector (50 ng/ $\mu$ l)	1 $\mu$ l	1 $\mu$ l	1 $\mu$ l
PCR product	2 $\mu$ l	-	-
Control inserts DNA (4 ng/ $\mu$ l)	-	2 $\mu$ l	-
T4 DNA ligase (3 u/ $\mu$ l)	1 $\mu$ l	1 $\mu$ l	1 $\mu$ l
Deionized water to a final vol of	10 $\mu$ l	10 $\mu$ l	10 $\mu$ l

The components were mixed by pipetting and incubated in a water bath at 16<sup>o</sup>C for overnight and later stored the vector ligated cDNA at -20<sup>o</sup>C.

### 3.2.8.12 Transformation of cloned DNAs

Electrocompetent *E.coli* DH10B cells (invitrogen) were chosen for transformation of insert ligated vector DNA. The DH10B electrocompetent cells were taken out from -80<sup>o</sup>C, thawed and mixed by gentle tapping. Then, 20  $\mu$ l of competent cells was added to each chill microcentrifuge tube containing 5  $\mu$ l either vector ligated control or sample cDNA. The cell/DNA mixture was transferred into a chilled 1 mm electroporation cuvette and subsequently electroporation was done with the help of Electroporator 2510 (eppendorf) at a pulse voltage of 2500 V. It was followed by addition of 1 ml sterile SOC medium to the cells immediately and transferred the solution to a 15 ml snap-cap tube. Tube was incubated at 37<sup>o</sup>C and 225 rpm for 1 h.

### 3.2.8.13 Blue white screening and calculation of transformation efficiency

Transformed cells with control DNA was diluted (1:100) with SOC media. Diluted cells (50  $\mu$ l) were spreaded on a prewarmed (at 37<sup>o</sup>C) LB plates containing ampicillin (100  $\mu$ g/ml), IPTG (0.5 mM) and X-Gal (80  $\mu$ g/ml). The transformed cells with sample DNAs were diluted as 1: 5 ratio with SOC media and spreaded 200  $\mu$ l on a selective plate similar to the one of control sample with the help of plate master. All plates were kept inside the laminar air flow for 10 min for absorbing the media, sealed

with parafilm and incubated in an inverted position at 37°C for overnight. After completion of the incubation periods, the plates were kept at 4°C for some time for developing a distinct blue color of the untransformed cells colony. The numbers of blue and white cells were counted using a colony counter (himedia) for the calculation of transformation efficiency using the following formula:

$$\frac{(\text{cfu on control plate/pg pUC19 DNA}) \times (1 \times 10^6 \text{ pg/ } \mu\text{g} \times \text{vol of transformants})}{(\text{plating vol} \times \text{dilution factor})}$$

Culture plates were stored at 4°C until picking up the individual white colonies for liquid culture.

### 3.2.9 Construction of standard full-length cDNA library

The standard full-length cDNA library was constructed following a PCR based method of SMART protocol (# 634901, clontech, USA) (Appendix-B 2). The SMART protocol was used to preferentially enrich the full-length cDNAs by eliminating T4 DNA polymerase and adaptor ligation steps.

#### 3.2.9.1 First-strand cDNA synthesis

First-strand cDNA was synthesized from 3  $\mu\text{l}$  (250 ng) of mRNA sample by combining 1  $\mu\text{l}$  of SMART IV oligonucleotide (12  $\mu\text{M}$ , 5'-AAGCAGTGGTATCAACGCAGAGTGGCCATTA CGGCCGGG-3') and 1  $\mu\text{l}$  of CDS III/3' PCR primer (12  $\mu\text{M}$ , 5'-ATTCTAGAGGCCGAGGCG GCCGACATG-d(T)<sub>30</sub>N-1N-3' where N = A, G, C, or T; N-1= A, G, or C) in a 0.5 ml tube. The components were mixed by brief centrifugation and incubated at 72°C for 2 min. After cooling for 2 min in ice, following components were added: 2  $\mu\text{l}$  of 5X first-strand buffer (250 mM Tris pH 8.3, 30 mM MgCl<sub>2</sub> and 375 mM KCl), 1  $\mu\text{l}$  of DTT (20 mM), 1  $\mu\text{l}$  of dNTP mix (10 mM) and 1  $\mu\text{l}$  of MMLV reverse transcriptase (20 u/  $\mu\text{l}$ ). The content was mixed by gentle pipetting and brief spinning, added one drop of mineral oil and subsequently incubated at 42°C for 1 h in a thermal cycler. Later, first-strand synthesis reaction was terminated by incubating the tube in ice. The first-strand cDNA solution was stored at -20°C for further use.

### 3.2.9.2 Amplification of cDNA by LD PCR

In a 0.5 ml tube containing 2  $\mu$ l of first-strand cDNA, following components were added in a total vol of 100  $\mu$ l in order to amplify the first-strand cDNAs: 80  $\mu$ l of deionized water, 10  $\mu$ l of 10X advantage 2 PCR buffer, 2  $\mu$ l of 50X dNTP mix, 2  $\mu$ l of 5' PCR primer (12  $\mu$ M, 5'-AAGCAGTGGTATCAACGCAGAGT-3'), 2  $\mu$ l of CDS III/3' PCR primer and 2  $\mu$ l of 50X advantage 2 polymerase mix. The components were mixed by gentle flicking of the tube and centrifuged briefly to collect the content at the bottom of the tube. The reaction mixture was overlaid with 2 drops of mineral oil, placed in a preheated (95°C) thermal cycler and followed the cycling parameters as: 95°C for 1 min, 25 cycles at 95°C for 15 s and 68°C for 6 min. After completing the cycles, 5  $\mu$ l of PCR product was resolved on a 1.1% agarose gel stained with ethidium bromide alongside 1 kb DNA marker (100 ng) at 65 V for 50 min and visualized under UV light, as well the rest amount stored at -20°C for further use.

### 3.2.9.3 Proteinase K digestion

In a sterile 0.5 ml tube, 2  $\mu$ l of proteinase K (20  $\mu$ g/ $\mu$ l) was added to 50  $\mu$ l of the amplified ds cDNA (2  $\mu$ g) in order to digest the contaminated proteins. Content was mixed properly and incubated at 45°C for 20 min. Then, 50  $\mu$ l of deionized water and 100  $\mu$ l of PCI (25:24:1, v/v/v) were added and subsequently mixed by continuous gentle inversion for 1-2 min. Mixture was centrifuged at 14000 rpm for 5 min to separate the phases. Top aqueous layer was collected to a clean 0.5 ml tube. To the collected solution, 100  $\mu$ l of CI (24:1, v/v) was added and mixed by continuous gentle inversion. Tube was centrifuged at 14000 rpm for 5 min to separate the phases. Top aqueous layer was collected to a clean 0.5 ml tube and added 10  $\mu$ l of 3 M sodium acetate, 1.3  $\mu$ l of glycogen (20  $\mu$ g/ $\mu$ l) as well as 260  $\mu$ l of 95% room temperature ethanol. Mixed solution was centrifuged immediately at 14000 rpm for 20 min at room temperature. Supernatant was removed carefully and later washed the pellet with 100  $\mu$ l of 80% ethanol. Pellet was air dried to evaporate off the residual ethanol and added 79  $\mu$ l of deionized water to resuspend it. Resuspended solution was stored at -20°C for further use.

#### 3.2.9.4 SfiI digestion

Following components were combined in order to digest the dsDNA with restriction enzyme in a 0.5 ml tube: 79  $\mu\text{l}$  of purified dsDNA, 10  $\mu\text{l}$  of 10X Sfi buffer, 10  $\mu\text{l}$  of SfiI enzyme (20 u/ $\mu\text{l}$ ) and 1  $\mu\text{l}$  of 100X BSA in a total vol of 100  $\mu\text{l}$ . The components were mixed well and incubated at 50°C for 2 h. Later, 2  $\mu\text{l}$  of 1% xylene cyanol dye was added to the solution and mixed well and followed with size fractionation step immediately.

#### 3.2.9.5 Size fractionation of cDNAs

A column procedure was used for the size fractionation of cDNAs. The chroma spin-400 column was prepared for its drip procedure by resuspending the gel matrix using a 1000  $\mu\text{l}$  pipettor. The column was attached to a ring stand and allowed the storage buffer to drain out. Then, carefully and gently 700  $\mu\text{l}$  of column buffer was added to the top of the column by allowing it to drain out. When this buffer had stopped dripping (~ 15-20 min), 100  $\mu\text{l}$  mixture of SfiI-digested cDNA and xylene cyanol dye was carefully and evenly applied to the top-center surface of the matrix. Sample was allowed to be fully absorbed into the surface of the matrix and subsequently applied 100  $\mu\text{l}$  of column buffer for washing. Buffer was allowed to drain out of the column. A rack containing fourteen 1.5 ml collection tubes was placed under the column, so that the first tube was directly under the column outlet. Then, 600  $\mu\text{l}$  of column buffer was added and immediately begin collecting single-drop fractions (~ 35  $\mu\text{l}$ ) in tubes from 1 to 14 one by one.

Profile of the fractions was evaluated on an agarose gel (1.1%) by running 3  $\mu\text{l}$  of each fraction in adjacent wells, alongside 300 ng of a 1 kb DNA marker at 150 V for 15 min and visualized under UV light. The first four fractions containing cDNA were pooled in a 1.5 ml tube (~ 140  $\mu\text{l}$ ) and subsequently added 1/10<sup>th</sup> vol of 3 M sodium acetate (pH 4.8), 1.3  $\mu\text{l}$  of glycogen (20 mg/ml) and 2.5 vol of 95% cold ethanol. The components were mixed gently by rocking the tube back and forth and placed the tube at -20°C for overnight. Then, the tube was centrifuged at 14000 rpm for 20 min at room temperature and carefully removed the liquid portion. The pellet was air

dried (~ 10 min) and subsequently resuspended in 7  $\mu$ l of deionized water gently as well stored at  $-20^{\circ}\text{C}$  until use in the ligation step.

### 3.2.9.6 Ligation of cDNA to $\lambda$ TriplEx2 vector

Ratio of cDNA to vector in the ligation reaction is a critical factor in determining transformation efficiency, and ultimately the number of independent clones in the library. Hence, for determining the optimal ratio of cDNA to vector, three independent ligation reactions were set up using three different ratios of cDNA to vector in 0.5 ml tubes as follows:

Components	1st ligation ( $\mu$ l)	2nd ligation ( $\mu$ l)	3rd ligation ( $\mu$ l)
cDNA	0.5	1.0	1.5
$\lambda$ TriplEx2 vector (500 ng/ $\mu$ l)	1.0	1.0	1.0
10X ligation buffer	0.5	0.5	0.5
ATP (10 mM)	0.5	0.5	0.5
T4 DNA ligase	0.5	0.5	0.5
Deionized water	2.0	1.5	1.0
-----			
Total vol	5.0	5.0	5.0

The components were mixed gently and incubated at  $16^{\circ}\text{C}$  for overnight.

### 3.2.9.7 Lambda phage packaging of ligated cDNA

Lambda phage packaging of the ligated cDNA was done with the packaging extract of Gigapack III gold (# 200202, stratagene). The tube of packaging extract was removed quickly from  $-80^{\circ}\text{C}$  and held between fingers until the content just begins to thaw. Then, 4  $\mu$ l of (~ 1.0  $\mu$ g) experimental DNA was added immediately to the packaging extract. The content was stirred with a pipet tip to mix well and subsequently spinned the tube quickly for 3-5 s. Tube was incubated at  $22^{\circ}\text{C}$  for 2 h, and later 500  $\mu$ l of SM buffer was added to the content. Then, 20  $\mu$ l of chloroform was mixed

gently and spun the tube briefly to sediment the debris. Supernatant was collected and stored at 4°C. This collected solution was used as unamplified library.

### 3.2.9.8 Titering of the unamplified library

Titering (pfu/ $\mu$ l) of the unamplified library was done for the estimation of number of independent phage and independent clones in the library. A single, isolated colony of XL1-blue *E. coli* cells was picked from the working culture plate and inoculated to 15 ml of LB/MgSO<sub>4</sub>/maltose broth in a 50 ml Erlenmeyer flask. The flask was incubated at 37°C and 140 rpm for overnight until the absorption of the culture at 600 nm reaches 2.0. The culture was centrifuged at 5000 rpm for 5 min. The supernatant was poured off and resuspended the pellet in 7.5 ml of 10 mM MgSO<sub>4</sub>. In the mean time, 90 mm LB/MgSO<sub>4</sub> plates were dried and warmed at 37°C for 10-15 min. The packaging extract from the previous step was diluted in 1X lambda dilution buffer as 1:5, 1:10, 1:15 and 1:20 ratios. Later, 1  $\mu$ l of the diluted phage was added to 200  $\mu$ l of the XL1-blue overnight culture, and allowed the phage to adsorb at 37°C for 10-15 min. Then, 2 ml of melted LB/MgSO<sub>4</sub> top agar (temperature not more than 45°C) was added, mixed quickly by inverting the tube, and immediately poured onto the prewarmed 90 mm LB/MgSO<sub>4</sub> plates. The plates were swirled quickly to allow even distribution of the top agar and kept at room temperature for 10 min to harden the top agar. The plates were incubated at 37°C for 6-18 h and checked periodically to ensure that plaques were developing. The plaques were counted and subsequently calculated the titer of the phage (pfu/ml) as: number of plaques x dilution factor x 10<sup>3</sup>  $\mu$ l/ml divided by  $\mu$ l of diluted phage plated. Titer of the three individual ligation reactions were compared and selected the best one.

### 3.2.9.9 Determination of recombinant percentage

Blue-white screening of *E. coli* XL1-blue cells was performed for the determination of recombinant percentage. The procedure was followed similar to the titering of the unamplified library on LB/MgSO<sub>4</sub> plates, except IPTG and X-gal was added to the melted top agar before plating the phage plus bacterial mixtures. For every 2 ml of melted top agar, 50  $\mu$ l each of the IPTG and X-gal stock solutions (100 mM each) were added. Plates were incubated at 37°C until plaques and blue colors get

developed (~ 16 h). The numbers of white (recombinant) and blue (non-recombinant) plaques were counted and calculated the recombinant percentage as: total white plaques/total blue plaques x 100.

### 3.2.9.10 Amplification of the library

A single, isolated colony from the primary working plate of XL1-blue was picked and inoculated to 15 ml of LB/MgSO<sub>4</sub> maltose broth. The inoculated media was incubated at 37°C and 140 rpm until the absorption of the culture at 600 nm reaches 2.0. The culture was centrifuged for 5 min at 5000 rpm, poured off the supernatant, and resuspended the pellet in 7.5 ml of 10 mM MgSO<sub>4</sub>. In 500 µl of overnight bacterial culture, 50 µl of diluted lysate (diluted as 1:10, 1:50 and 1:100 ratios) was added and incubated at 37°C in a water bath for 15 min. Later, 4.5 ml of melted LB/MgSO<sub>4</sub> soft top agar (not more than 45°C) was added to each tube, mixed quickly and swirled on prewarmed (at 37°C for 10-15 min) LB/MgSO<sub>4</sub> agar plates (90 mm) and let them cool at room temperature for 10 min to harden the top agar. The plates were kept in inverted position at 37°C until (~ 16 h) the plaques touching each other. Then, 4 ml of 1X lambda dilution buffer was added to each plate and stored at 4°C overnight. The plaques were then pooled in 1X lambda dilution buffer to form an amplified library lysate and subsequently incubated at room temperature and 50 rpm for 1 h on a platform shaker. To clear the phage lysate of cell debris and to lyse any remaining intact cells, the phage lysate was mixed well and poured into a sterile, 50 ml polypropylene screw-cap tube. Later, 10 ml of chloroform was added to the lysate, mixed by vortexing for 2 min and subsequently centrifuged at 7000 rpm for 10 min. The supernatant was collected and stored at 4°C. Later, 1 ml aliquots were prepared, added DMSO solution to a final concentration of 7% and stored at -70°C.

### 3.2.9.11 Conversion of phage (λTriplEx2) to plasmid (pTriplEx2)

Conversion of a λTriplEx2 clone to a pTriplEx2 involves *in vivo* excision and circularization of a complete plasmid from the recombinant phage. Release of the plasmid occurs automatically when the recombinant phage is transduced into a bacterial host in which cre recombinase is being expressed. In the present

investigation, *E. coli* BM25.8 was used for cre recombinase activity that had been grown at 31°C. A single, isolated colony of BM25.8 was picked up from the working plate and inoculated into 10 ml of LB broth in a 50 ml Erlenmeyer flask and subsequently incubated at 31°C, 150 rpm until the absorption of the culture at 600 nm reaches in between 1.1 to 1.4. Later, 100 µl of 1 M MgCl<sub>2</sub> was added to the culture. Then, 200 µl of cell culture was infected with 1 µl of phage solution (amplified library) and incubated at 31°C for 30 min and subsequently added 400 µl of LB broth. The infected cell culture was incubated for an additional 1 h at 31°C and 225 rpm. Finally, 10 µl of the cell suspension was diluted with 200 µl of LB broth which spreaded on LB agar plate containing carbenicillin (100 µg/ml) and later incubated the plates at 31°C for overnight. The plates were stored at 4°C until picking individual colonies for liquid culture.

### 3.2.10 Liquid culturing of transformed *E. coli* cells

Individual white colonies of transformed *E. coli* cells were picked up using sterile tooth pick in an aseptic condition and inoculated into 2 ml of LB broth containing either ampicillin (100 µg/ml) for DH10B cells or carbenicillin (100 µg/ml) for BM25.8 cells. The tubes were incubated either at 37°C for DH10B cells or at 31°C for BM25.8 cells under shaking at 200 rpm until absorption of the culture at 600 nm reaches more than 1.5. Later, the cultures were centrifuged for 2 min at 12000 rpm, discarded the liquid portion and collected the cell pellets to be used for plasmid extraction.

### 3.2.11 Extraction of recombinant plasmids

Good quality plasmid DNA is a pre-requisite for sequencing reactions. The plasmids of successfully transformed DH10B and transduced BM25.8 cells were extracted either with the modified protocol of Sambrook et al. (1989) or with the manufacturer protocol of GenElute plasmid miniprep kit (# PLN350; sigma-aldrich). Protocol of Sambrook et al. (1989) was modified and followed as: 100 µl of chilled homogenization solution (50 mM glucose, 25 mM Tris-Cl pH 8.0, 10 mM EDTA pH8.0) was added to the cell pellets and subsequently resuspended by vortaxing. Then, 200 µl of freshly prepared lysis solution (0.2 M NaOH and 1% SDS) was added and mixed

by inversion of the tubes 4/5 times gently. After 5 min of incubation at room temperature, 150  $\mu$ l of ice cold precipitation solution (5 M potassium acetate pH 5.5) was added followed by gentle inversion of the tubes for 4/5 times gently. At this stage, tubes were incubated in ice for 5 min and subsequently centrifuged for 8 min at 14000 rpm. The supernatant was collected, added 300  $\mu$ l of isopropanol and mixed by 4/5 times by gentle inversion. Tubes were centrifuged immediately for 8 min at 14000 rpm. The supernatant was discarded and kept the tubes on tissue paper for 5-10 min for drying the pellets in an inverted position. Later, 300  $\mu$ l of TE buffer (pH 8.0) was added for dissolving the pellets, subsequently added 2  $\mu$ l of RNase (10 mg/ml), mixed thoroughly by brief spin followed by incubation at 37<sup>o</sup>C for 15 min. It was added with 200  $\mu$ l of PCI (pH 8.0), mixed by gentle inversion of the tubes and centrifuged for 8 min at 14000 rpm. Upper aqueous phase was collected carefully, added 0.1 vol of 3 M sodium acetate (pH 5.2) and 2 vol of 100% ethanol, which subsequently incubated for 10 min in ice. Afterwards, the tubes were centrifuged for 8 min at 14000 rpm, discarded the supernatant completely and added 500  $\mu$ l of 70% ethanol. It was centrifuged again for 5 min at 14000 rpm, discarded the liquid portion and air dried the pellets. The dried pellets were dissolved in 25  $\mu$ l of distilled water and stored at -20<sup>o</sup>C until use for sequencing.

The protocol of GenElute plasmid miniprep kit was followed as: the cell pellets were homogenized in 200  $\mu$ l of resuspension solution by vortexing. The cells were lysed in 200  $\mu$ l of lysis solution by gentle inversion of the tubes for 6-8 times followed by incubation for 3-5 min until the mixture becomes clear and viscous. The cell debris was precipitated by adding 350  $\mu$ l of the neutralization or binding solution followed by gentle inversion for 4-6 times. The tubes were centrifuged at 12000 rpm for 10 min. The clear lysate was transferred to a prepared miniprep column. The column was prepared by washing with 500  $\mu$ l of column preparation solution followed by centrifugation at 12000 rpm for 1 min. The column containing the lysate was centrifuged at 12000 rpm for 1 min, and subsequently decanted the flow through liquid. The column was added with 750  $\mu$ l of diluted wash solution and centrifuged at 12000 rpm for 1 min. The flow through liquid was discarded and centrifuged again at 12000 rpm for 1 min to remove the excess liquid. The column was transferred to a new collection tube and added 50  $\mu$ l of sterile water followed by incubation for 3 min. Finally, the tubes containing the column were centrifuged at 12000 rpm for 1

min, collected the plasmid DNA in solution and stored at  $-20^{\circ}\text{C}$  until use for sequencing.

Quality of the extracted plasmid DNAs was evaluated by resolving 5  $\mu\text{l}$  on an agarose gel (1%) stained with ethidium bromide and run for 40 min at 70 V and visualized under UV light.

### 3.2.12 Sequencing of inserts in plasmid DNAs

Sequencing procedure for the extracted plasmid DNAs were followed by three main steps: cycle sequencing PCR, purification of PCR products and electrophoresis of the purified products. The cycle sequencing PCR was performed for each sample of plasmid DNA taking 1  $\mu\text{l}$  (~ 150-300 ng) as template with the reagents of BigDye terminator kit 3.1 (Applied biosystem) as well as either with M13 forward primer (3.2 pmol, 5'-GTAAAACGACGGCCAGT-3') for pGEM-T easy vector or 5' sequencing primer (3.2 pmol, 5'-CGCCTGGAGACGCCATCC-3') for TriplEx2 plasmid. The pGEM -3Z (+) DNA sample was used as template for control reaction (0.2  $\mu\text{g}/\mu\text{l}$ ) with the specific -21 M13 forward primer (3.2 pmol). A PCR master mix was prepared in a 0.5 ml tube by  $1/8^{\text{th}}$  dilution of a full strength reaction and subsequently added 9  $\mu\text{l}$  in each 1  $\mu\text{l}$  template DNA as follows:

Components	full strength	eighth strength
2.5X ready reaction mix	8.0 $\mu\text{l}$	1.0 $\mu\text{l}$
5X dilution buffer	-	1.5 $\mu\text{l}$
Primer (0.8 pmol/ $\mu\text{l}$ )	4.0 $\mu\text{l}$	2.0 $\mu\text{l}$
Sterile water	7.0 $\mu\text{l}$	4.5 $\mu\text{l}$
-----		
Total vol	19.0 $\mu\text{l}$	9.0 $\mu\text{l}$

PCR program was followed as:  $96^{\circ}\text{C}$  for 1 min, 25 cycles at  $96^{\circ}\text{C}$  for 10 s,  $50^{\circ}\text{C}$  for 5 s and  $60^{\circ}\text{C}$  for 4 min for M13 forward primer, as well  $96^{\circ}\text{C}$  for 2 min, 25 cycles at  $96^{\circ}\text{C}$  for 12 s,  $50^{\circ}\text{C}$  for 7 s and  $60^{\circ}\text{C}$  for 5 min for 5' sequencing primer.

Cycle sequencing PCR products were purified following the protocol of BigDye Terminator v 3.1 clean up tube method. The final 10  $\mu$ l product was transferred to a 1.5 ml tube, added 12  $\mu$ l of master mix I containing 10  $\mu$ l of sterile water with 2  $\mu$ l of 125 mM EDTA (freshly prepared) and mixed well. Then, 52  $\mu$ l of master mix II containing 2  $\mu$ l 3 M sodium acetate (pH 4.6) with 50  $\mu$ l of ethanol (room temperature) was added, mixed well and incubated at room temperature for 15 min. The tubes were centrifuged immediately at 12000 rpm for 20 min, decanted the liquid portion, and subsequently added 250  $\mu$ l of 70% ethanol for washing. The tubes were centrifuged again at 12000 rpm for 10 min at room temperature, decanted the supernatant carefully and dried the tubes at 37°C for 30 min. Finally, 12-15  $\mu$ l of Hi-Di formamide was added in each tube and dissolved the pellets. The sample solutions were transferred from 1.5 ml tubes to a 96 well sample plate. The samples were covered with septa, centrifuged at 2000 rpm for 1 min at room temperature, and subsequently incubated at 95°C for 5 min in a hot start PCR for denaturation of dsDNA followed by instant chilling in ice. The plate was fitted with a plate base, retainer and subsequently placed in Genetic Analyzer 3130xl (applied biosystem) for electrophoresis. The running programme was scheduled in Data collection v 3.0 software as: 36 cm or 50 cm capillary, rapid run type, POP7 polymer, oven temperature 65°C.

### **3.2.13 Base calling, vector trimming and homology searching of EST sequences**

After completion of electrophoresis and data acquisition, the raw sequences were passed through several quality parameters: base calling, vector screening and trimming as well as homology searching above certain expectation value.

Raw sequences were analyzed using Sequence analysis software v 5.2 (applied biosystem). Base calling was done using the algorithm of KB basecaller with a minimum quality value (QV) 20, where the chances of incorrect base call is 1%. After quality calling sequences were checked again for its quality statistics such as QVs, length of read (LOR) and signal to noise. Sequences with short read length and poor quality electropherogram were eliminated.

Vector and adaptor sequences present in the 5' and 3' sites of the raw sequences were trimmed using Sequence scanner v 3.1 software (applied biosystem). The trimmed sequences with more than 100 bp in length were selected for further analysis. Vector trimmed sequences were searched for their homolog sequences using the BLAST program in NCBI (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>). Both BLASTn and BLASTx program were used for searching homolog in a non-redundant nt and protein database of NCBI, respectively with a minimum cut off e-value less than 1E-05. The sequences with no significant hits were considered as novel sequences and included for subsequent analysis.

Sequences that passed through quality parameters were considered for clustering by Sequencher 4.1 software (Gene Codes Corporation). Clustering was made using the dirty data algorithm considering ambiguous base calls for poor matches to exact base calls following the criteria as: a) gaps optimization for small inserts and double-called bases through ReAligner algorithm b) prefer 3' gap replacement c) minimum overlap 20 bp d) minimum match percentage 85. The consensus sequences of clustered ESTs (contigs) alongwith the singlets were imported in FASTA format which were together defined as unigene set. The contig sequences were BLASTed against the non-redundant protein database in NCBI for searching their homolog which code for the same protein.

### **3.2.14 Functional annotation and analysis of drought induced unigenes**

Bioinformatics tools and databases were provided the scope of generating useful information from ESTs. In order to analyse and annotate the functions of drought induced unigenes, number of bioinformatics databases, tools and programs were used.

#### **3.2.14.1 GC percentage and codon usage analyses**

Percentage of GC present in the unigene set was calculated using GEECEE program and codon usage as well functional GC percentage were analysed through CUSP program of Pasteur lab (<http://mobylye.pasteur.fr/cgi-bin/portal.py>). Codon usage analysis was done on the basis of a standard codon usage table.

### 3.2.14.2 Mining of ORFs, protein domains and construction of phylogenetic trees

Open reading frames were predicted by translating each sequence into six reading frame using NCBI ORF Finder (<http://www.ncbi.nlm.nih.gov/projects/gorf/>). The ORFs located in the direct strand were translated into amino acid sequences and searched for their homolog domain in pfam database (<http://pfam.sanger.ac.uk/>) for assigning functions. Protein sequences showing e-value less than  $1E-05$  were selected for searching their conserved and multi-domain nature against CDD database of NCBI, (<http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml>). After eliminating the same hit domains, conserved domain sequences (e-value less than  $1E-05$ ) were selected for the construction of maximum likelihood trees depending on JTT matrix-based model (Jones et al. 1992) at bootstrap value 500 using MEGA 4.0 (Tamura et al. 2007) for illustrating their evolutionary relationship.

### 3.2.14.3 Mining of signal peptide and peptide anchor in protein domain sequences

Protein domain sequences found in the unigene set were mined for signal peptide and peptide anchor through SignalP (<http://www.cbs.dtu.dk/services/SignalP/>) program, using both methods, i.e. neural networks and HMM by selecting the organism group as eukaryotes. For signal peptide, most likely to be cleaving site and for peptide anchor high probable anchor were selected and recorded their cleaving site and anchor site amino acid positions, respectively.

### 3.2.14.4 Gene ontology analysis

Gene ontology provides a set of structured, controlled vocabularies for community use in annotating genes, gene products and sequences. There are numbers of computational tools that have been developed and successfully used for GO analysis (<http://www.geneontology.org/GO.tools.browsers.shtml>). Since tea genome is not available, as a model plant, *Arabidopsis* was chosen for GO analysis. WU-BLAST was performed for each of the unigene for finding their homolog against *Arabidopsis* by using TAIR9 transcripts (-introns, +UTRs) (DNA) (<http://www.arabidopsis.org/index.jsp>). The corresponding locus hit of each unigene having e-value less than  $1E-$

10 were selected for GO search against *Arabidopsis* proteome using GOslim program (<http://www.arabidopsis.org/index.jsp>) and presented each sub-categories in percentage according to the following formula:

$$\frac{\text{Annotations to terms in GOslim category}}{\text{Total annotations to terms in this ontology}} \times 100 = \%$$

### 3.2.14.5 Identification of drought associated genes and pathways

Identification of drought associated gene was done based on the protein hits in NCBI database. Those genes of which protein products somehow associated to drought were recorded as drought associated genes based on knowledge of biochemistry and molecular biology. If there were three or more genes found in a single pathway, the corresponding pathways were searched through KEGG analyses (<http://www.genome.jp/kegg/>).

### 3.2.14.6 Mining of genic microsatellite markers

For mining SSR, WebSat program (<http://wsmartins.net/websat/>) was used. The criteria for SSR search were set as: minimum of 9, 5, 4, 4 and 3 times of repeat for di, tri, tetra, penta and hexa repeat SSR, respectively. Primer picking parameters were given as: optimum length 22 nt (minimum 18 and max. 27), T<sub>m</sub> 60 °C (minimum 57 °C and max. 68 °C), minimum GC % 40.0 and rest with default parameters. The primer designed was done either with WebSat program or with VectorNTI (Invitrogen) software.

### 3.2.15 Analyses of standard full-length unigenes

Full-length unigenes were subjected to various bioinformatics analyses for functional annotation. The codon usage, GC percentage and GO analyses were done as described in 3.2.14.

A homology search was performed between the unigenes of drought induced and standard full length unigenes for identifying the redundant sequences using BLAST+

2.2.22 program (Altschul et al. 1997). In order to perform the search, a local nt database of drought induced unigenes was made and subsequently BLASTed the standard full length unigenes against this database at an e-value threshold of 0.0001. The homolog genes were selected on the basis of cut off value as: minimum identity percentage 80% and bit score 100. Finally, the coding proteins of homolog unigenes were identified through BLASTx analysis of NCBI (<http://blast.ncbi.nlm.nih.gov/Blast>) and represented using Venn diagram.

### **3.2.16 Comparative analyses of transcripts under normal growth, drought and winter dormancy stress**

Genomic information available in the open access biological databases gives an opportunity of comparative analyses. An attempt of comparative genomic analysis was made among the transcripts of normal growth, drought and winter dormancy stress. All the available EST sequences of tea leaves in dbEST, NCBI, generated under normal growth, drought and winter dormancy stresses were downloaded using the bulk sequence download system of NCBI and subsequently vector trimmed and checked for quality parameters. The quality sequences of each category were clustered and subsequently used the clustered sequences (or unigenes) for further analyses. The quality parameters of sequences and its clustering methodology were followed as described in 3.2.13.

#### **3.2.16.1 Tissue specific expression of genes under drought**

In order to identify the tissue specific i.e in roots and leaves expression of unigenes under drought, a homology search was performed between the drought induced unigene sets of roots and leaves using BLAST+ 2.2.22 program (Altschul et al. 1997). A local nt database of drought induced unigenes of root tissue was made, and subsequently BLASTed the drought induced unigenes of leaf tissue against the local nt database at an e-value threshold of 0.0001. The parameters for the identification of homolog genes were followed as described in 3.2.15. Besides the homolog unigenes, the rest unigenes were considered as specific either in roots or leaves.

### 3.2.16.2 Comparative functional annotation of unigenes

Number of cds and GC percentage were calculated using the CUSP program (<http://mobyli.pasteur.fr/cgi-bin/portal.py>). On the basis of NCBI, BLASTx analysis, each set of unigenes were classified as: no significant hit sequences, unknown proteins, predicted proteins, chaperones, heat shock proteins, transcription factors, enzymes and other functional proteins at an e-value threshold of 1E-05. The common unigenes between or among the unigene sets of normal growth, drought or winter dormancy stress were identified through homology search between two sets at a time at an e-value threshold of 1E-04, identity percentage 80 and bit score 100, using BLAST+ 2.2.24 program (Altschul et al. 1997) and represented in Venn diagram.

Functional assignment of unigenes from each set of sequences were done on the basis of *Arabidopsis* proteome hits at an e-value threshold of 1E-10 using WU-BLAST, TAIR (<http://www.arabidopsis.org/index.jsp>). The assigned functional categories of unigenes were calculated in percentage as described in 3.2.14.

### 3.3 Results

The findings during the construction of three forward subtracted SSH and one full-length standard cDNA libraries and their subsequent sequence analyses have been presented here in different headings and sub-headings. The SSH libraries of TV-23, S.3/A3 and in between TV-23 and S.3/A3 cultivars were named as FSL1, FSL2 and FSL3, respectively.

#### 3.3.1 Protocol optimization for RNA isolation

A single method of RNA isolation cannot be applied in all plants due to its differences in tissue composition and function. The conventional guanidine-HCl, CTAB/NaCl, SDS/Phenol based protocols were failed to isolate RNA from tea roots. The modified guanidine-HCL protocol yielded not only low RNA but unusable for further analysis (Table 3.2). Commercially available TRIzol reagent from three different sources such as 'Invitrogene', 'MRC' and 'Sigma' were found to be completely unsuccessful (Fig. 3.3, Table 3.2). The RNAaques kit (Ambion, USA) was found to be successful (Fig. 3.3), but the yield was not enough for further analysis (Table 3.2). Hence, there was a need of protocol optimization for RNA isolation. The SDS based protocol was optimized perfectly for isolating RNA from tea roots. Prior to RNA isolation, treatment of roots with detergent followed by washing with DEPC treat water was found fruitful. Moreover, precipitation of RNA using sodium acetate along with isopropanol by incubating only for 30 min was reproducible.

The optimized protocol was followed routinely and found an average yield of 127  $\mu\text{g}$  in per g of FW. Quality was found good as judged by  $A_{260}/A_{280}$  and  $A_{280}/A_{230}$  ratios (Table 3.2) and the appearance of major 26S and 18S bands after gel electrophoresis (Fig. 3.3). The  $A_{260}/A_{230}$  ratio was found higher than 2.0 (Table 3.2) indicating the absence of polyphenol and polysaccharides contamination in the isolated RNA. The  $A_{320}$  was found 0.0 indicating the lack of protein contamination. A comparatively low yield of RNA was found in stress conditioned tissues. Around 50  $\mu\text{g}$  of PolyA<sup>+</sup> RNA was harvested from 1000  $\mu\text{g}$  of isolated total RNA. The  $A_{260}/A_{230}$  and  $A_{260}/A_{280}$  ratio of the

Table 3.2: RNA yield and quality by different methods

Protocols	Ratios*		Yield*
	A <sub>260</sub> /A <sub>230</sub>	A <sub>260</sub> /A <sub>280</sub>	µg g <sup>-1</sup> FW
RNAqueous kit (Ambion)	2.29 ± 0.05	1.66 ± 0.03	28 ± 1.47
Optimised protocol	2.05 ± 0.03	1.80 ± 0.02	127 ± 0.70
Modified guanidine-HCl	1.73 ± 0.08	1.95 ± 0.03	7 ± 1.62
TRIzol reagent (Invitrogene/ MRC/Sigma)	—	—	no yield
Modified CTAB	—	—	no yield

\*mean of 10 different extractions ± standard error



Fig. 3.3: Agarose gel electrophoresis of total RNA isolated from tea roots by different reagents and methods. A, B: RNAqueous kit; C, D: TRIzol reagent; E, F: modified guanidine-HCl; G, H: CTAB method; I-L: using optimized protocol (I, 15 and J, 21 days of drought induced plants, K, L: control plants); M, ssRNA ladder (new england biolab).

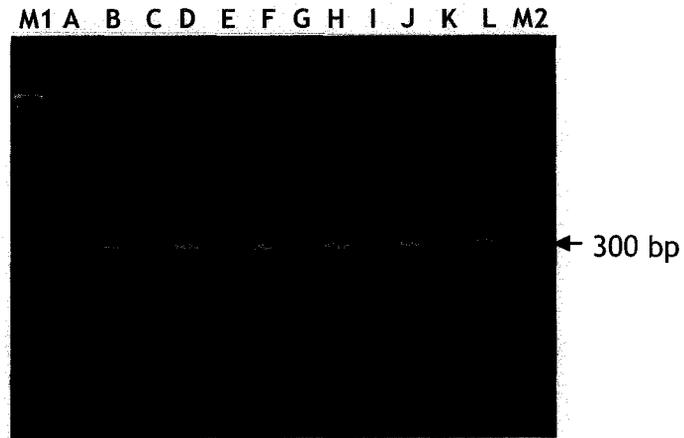
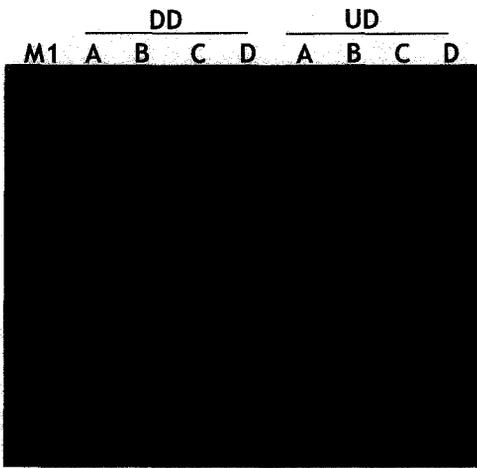


Fig. 3.4 (left) Evaluation of *RsaI* restriction digestion of dsDNAs (DD = digested, UD = undigested): A. TV-23 control, B. TV-23 drought induced, C. S.3/A3 control and D. S.3/A3 drought induced plants.

Fig. 3.5 (right) Adaptor ligation efficiency analyses of 'tester' dsDNAs: (A-D for FSL1 library of TV-23 plant; E-H for FSL2 library of S.3/A3 plant, I-L for FSL3 intervarietal library of TV-23 and S.3/A3 plants) A, E, I, Adaptor 1 ligated tester dsDNA; C, G, K, Adaptor 2 ligated tester dsDNA; B, F, J, Control reaction of 26S rRNA gene for adaptor 1 ligated tester dsDNA; D, H, L, Control reaction of 26S rRNA gene for adaptor 2 ligated tester dsDNA. (M1, 1kb marker and M2, 100 bp marker).

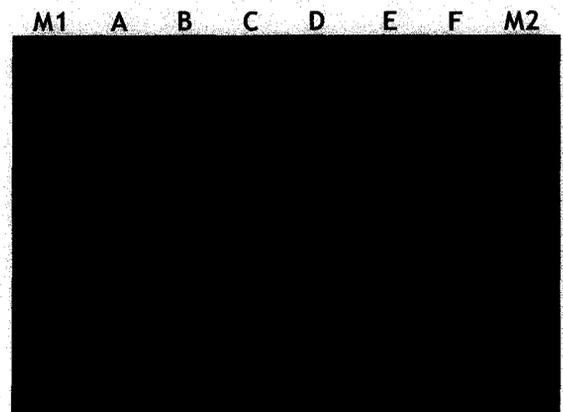


Fig. 3.6 Analyses of primary and secondary PCR products: (left) primary PCR (right) secondary PCR; A, C, E unsubtracted tester control dsDNAs for FSL1, FSL2 and FSL3 libraries, respectively, and B, D, F forward subtracted experimental dsDNA for FSL1, FSL2 and FSL3 libraries, respectively (M1, 1 kb marker and M2, 100 bp marker).

isolated mRNA pool was found on an average 2.2 and 1.8, respectively and  $A_{320}$  was recorded as 0.0 for all samples.

### **3.3.2 Analyses of different steps in SSH library construction**

Results that achieved throughout the SSH library construction process were described below.

#### **3.3.2.1 Yield and quality checking of dsDNAs**

Synthesized dsDNA was visualized as smear in gel. The size distribution of which was observed approximately from 500 bp to more than 12 kb whereas it was found from 300 bp to 10 kb for the *Rsa*I digested products (Fig. 3.4). It was an indication of successful restriction digestion.

#### **3.3.2.2 Ligation efficiency analysis**

Result of each ligation efficiency test was analysed based on the amplicon and their intensities in gel (Fig. 3.5). The housekeeping gene, 26S rRNA was amplified perfectly and formed a band exactly at 300 bp in each sample. The band intensity of 26S rRNA gene was equal in each tested sample. The band intensity of the control gene was compared with the intensity of the corresponding band of each adaptor ligated tester cDNA sample. Band intensity of the experimental sample was found approximately 25% compared to the control sample. Moreover, there were one or more prominent bands in between 1 kb to 400 bp in the tested samples. It had indicated a successful ligation where at least 25% of the cDNAs were with adaptors on both the ends.

#### **3.3.2.3 Analysis of differentially expressed transcripts**

Number of cycles for primary PCR was standardised at 30. Primary PCR products were found as smear ranges from 0.2 to 2 kb with some faint bands in between for the abundant transcripts [Fig. 3.6 (left)]. Smear intensity of control sample was observed higher than that of subtracted sample cDNAs. Secondary PCR was

standardised at 12 cycles. Secondary PCR products were also formed the same smear ranges like that of the primary PCR products [Fig. 3.6 (right)]. However, smear intensity was higher than that of the primary PCR products with a clear difference of the same between control and subtracted cDNAs either with or without distinct bands.

Relative abundance of 26S rRNA gene in subtracted and unsubtracted control cDNAs proved well the efficiency of subtraction. There were only differences at 18<sup>th</sup> PCR cycle among the band intensities of 26S rRNA gene in the unsubtracted control and subtracted cDNA samples of 'FSL1' and 'FSL2' libraries [Fig. 3.7 (image nos 5 & 6)]. These results indicated the poor subtraction of the intra-varietal cDNAs. However, the subjected housekeeping gene was amplified well from 18 PCR cycle onwards in the unsubtracted control cDNAs of inter-varietal 'FSL3' library whereas it was gradually amplified from 28 PCR cycles in the subtracted cDNAs of the same. This indicated an efficient subtraction of the cDNAs of 'FSL3' library [Fig. 3.7 (image no 7)].

#### 3.3.2.4 Screening of transformed cells and isolation of plasmids

A successful identification of transformed cells was done through blue white screening (Fig. 3.8). Transformation efficiency for the control DNA sample was found  $1.0 \times 10^9$  cfu/ $\mu$ g. This value was near to the standard level of electropotent DH10B *E. coli* cells. The transformation efficiencies for the subtracted cDNAs of FSL1, FSL2 and FSL3 libraries were recorded as  $1 \times 10^7$ ,  $3 \times 10^5$  and  $2 \times 10^4$ , respectively. Plasmid isolation was done successfully from the transformed cells following the in-house optimized protocol. There were observed three distinct bands of plasmids in gel without any smearing (Fig. 3.9). Only 1  $\mu$ l of the isolated plasmid DNA solution was found enough for the required concentration (150-300 ng) of sequencing.

#### 3.3.3 Generation of ESTs under drought stress and their clustering

A total of 2250 colonies was randomly picked and sequenced which consisted of 400 colonies from each of FSL1 and FSL2 library and 1450 colonies from FSL3. It has been produced a total of 649 ESTs from both FSL1 and FSL2 libraries as well as 1052 ESTs

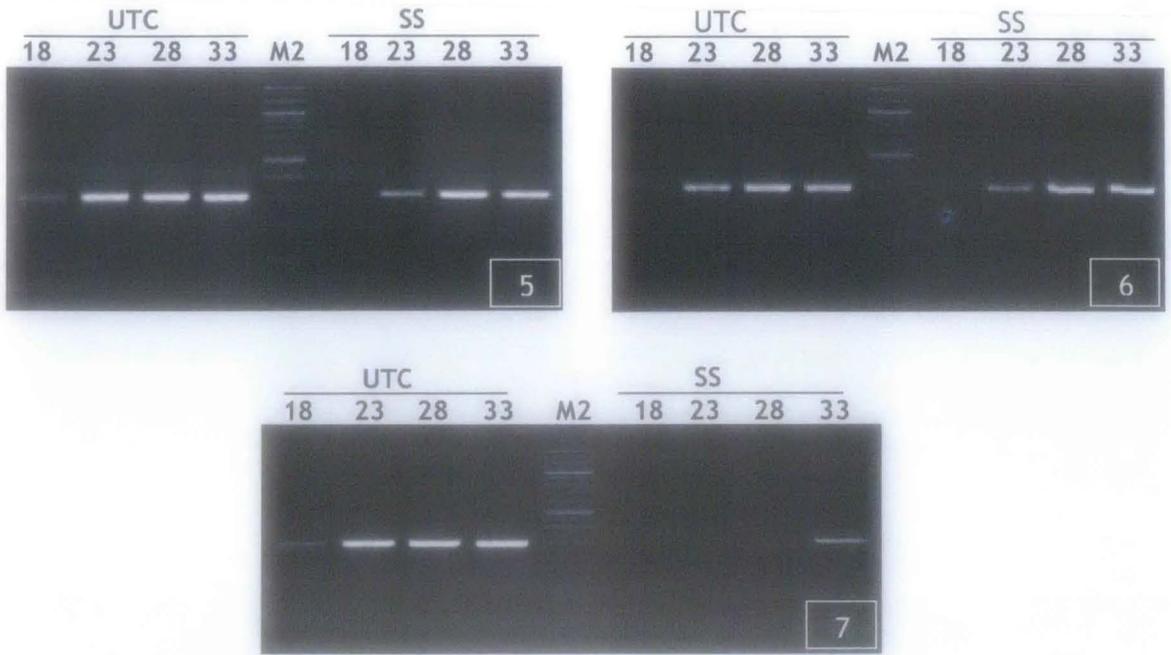


Fig. 3.7 Subtraction efficiency analyses of SSH libraries: (5) FSL1 (6) FSL2 and (7) FSL3 (UTC = unsubstracted tester control, SS = subtracted sample at 18, 23, 28 and 33 PCR cycles, M2 = 100 bp marker).

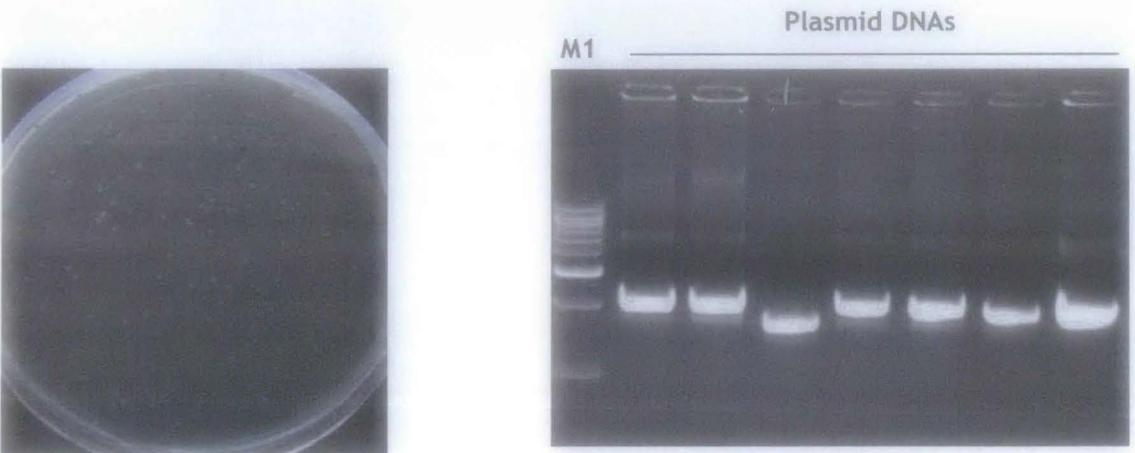


Fig. 3.8 (left) Blue white screening of transformed DH10B *E. coli* cells  
 Fig. 3.9 (right) Extracted plasmids of transformed cells containing the DNAs of interest (M1= 1 kb DNA marker).

from FSL3 library of which read length ranges from 200 to 900 bp. The sequences were deposited in dbEST, NCBI (Genbank accessions, GH734203-GH734851 from FSL1 and FSL2 libraries as well GT968791-GT969386; GW316843-GW317159; GW315010-GW315149 from FSL3 library). However, sequences of FSL1 and FSL2 libraries were not considered for further analyses due to their low quality value and high redundancy. Subsequent clustering and analyses were performed with the sequences of FSL3 library, the inter-varietal one. Out of 1052 sequences, a total of 572 numbers of ESTs passed through the set filtration criteria. Length of the ESTs was found from 178 to 832 bp with an average of 428 bp. There were included 5.94 % (34 out of 572) of ESTs with no significant hits in the database, 6.99% of ESTs (40 out of 572) similar to predicted proteins, 36.19 % (207 out of 572) of ESTs similar to unknown proteins, 3.85% of ESTs (22 out of 572) similar to small heat shock proteins, 3.67% of ESTs (21 out of 572) similar to molecular chaperones and 43.36% ESTs (248 out of 572) similar to different enzymes and functional proteins. After clustering, the quality ESTs collapsed into 54 contigs comprising 2 to 201 numbers of individual ESTs and 192 singlets producing a total of 246 unigenes (Fig. 3.10). The average length of the unigenes was found 513 bp where contigs ranges from 221 to 989 bp and singlets ranges from 178 to 832 bp. Contigs containing more than 4 ESTs and e-value less than  $1E-10$  were depicted in Table 3.3. The unigenes were assigned a name as UEST1, UEST2 and so.

The unigenes were annotated for various functional properties. There was an overall 44.28% of GC content and 1915 numbers of cds containing 44.6% of functional GC in the unigene set. Amino acid, leucine (L) was found as highly coded (9.37%) compared to the other amino acids followed by serine (S) (9.29%) and arginine (R) (6.98%). Interestingly, methionine (M) was found as the least coded amino acid compared to the rest amino acids (1.98%) (Fig. 3.11, Appendix-D 1.1). A total of 832 numbers of ORFs were detected through the translation of six reading frame. It covers 97.23% of the unigenes (out of 246, 239 contains ORF) which have one or more ORFs. Length of the translated sequences of ORFs was ranges from 33 to 236 amino acids. There was identified a total of 85 quality peptide in the direct strand of unigenes through the screening of pfam homologs with an e-value less than  $1E-05$ . Out of 85 peptides, signal peptides and signal anchors were produced by 8 and 37 peptides, respectively (Table 3.4). There were also 74 conserved protein domains as

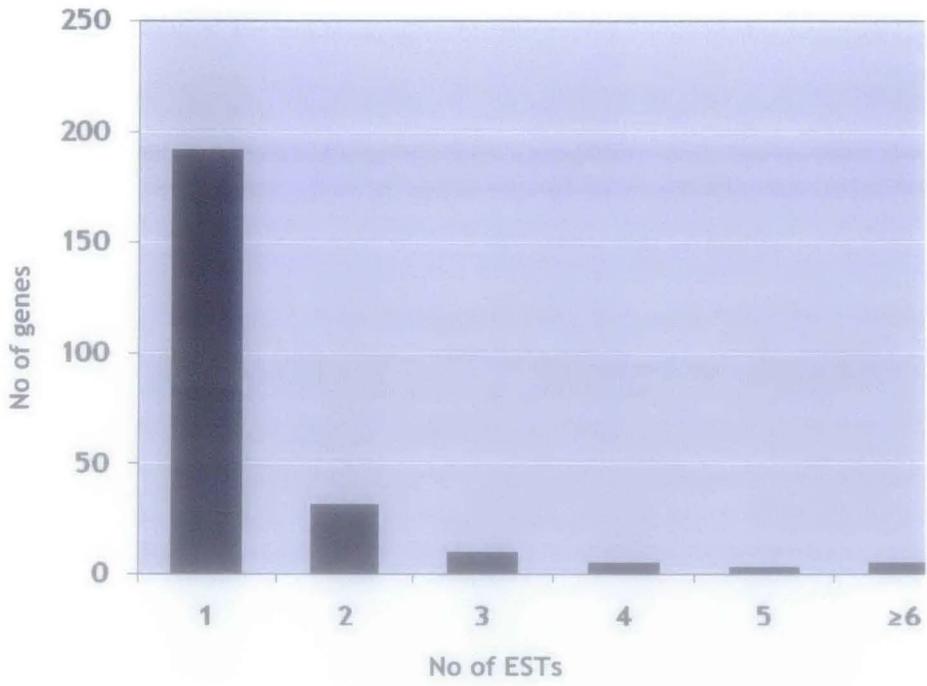


Fig 3.10: Distribution and numbers of assembled drought induced ESTs

Table 3.3: Assembled clusters that contain more than 4 drought induced ESTs and e-value less than 1E-10

Contig IDs	Homology	No of ESTs	Organisms	Gene IDs	E-value	Score
Contig1	18S rRNA gene (partial)	201	<i>Glomus proliferum</i>	FN547501.1	9E-153	549
Contig2	chaperone, mRNA	22	<i>Agave tequilana</i>	DQ515785.1	4E-93	351
Contig4	heat shock protein, mRNA	14	<i>Camellia sinensis</i>	EU727315.1	0.0	1141
Contig5	predicted protein, mRNA	9	<i>Populus trichocarpa</i>	XM002305555.1	2E-80	307
Contig6	ubiquitin carrier protein mRNA, predicted	6	<i>Vitis vinifera</i>	XM002277413.1	1E-17	98.7
Contig7	aspartate aminotransferase, mRNA	5	<i>Daucus carota</i>	M92660.1	3E-101	378
Contig9	heat shock protein, mRNA	5	<i>Hevea brasiliensis</i>	AF521007.1	0.0	699
Contig10	60S acidic ribosomal protein PO, mRNA	4	<i>Euphorbia esula</i>	AF227622.1	4E-115	423
Contig11	evolutionarily conserved C-terminal region 6, mRNA	4	<i>Arabidopsis thaliana</i>	NM_001084699.1	2E-91	345
Contig13	putative remorin protein, mRNA	4	<i>Ricinus communis</i>	XM_002529577.1	2E-101	378

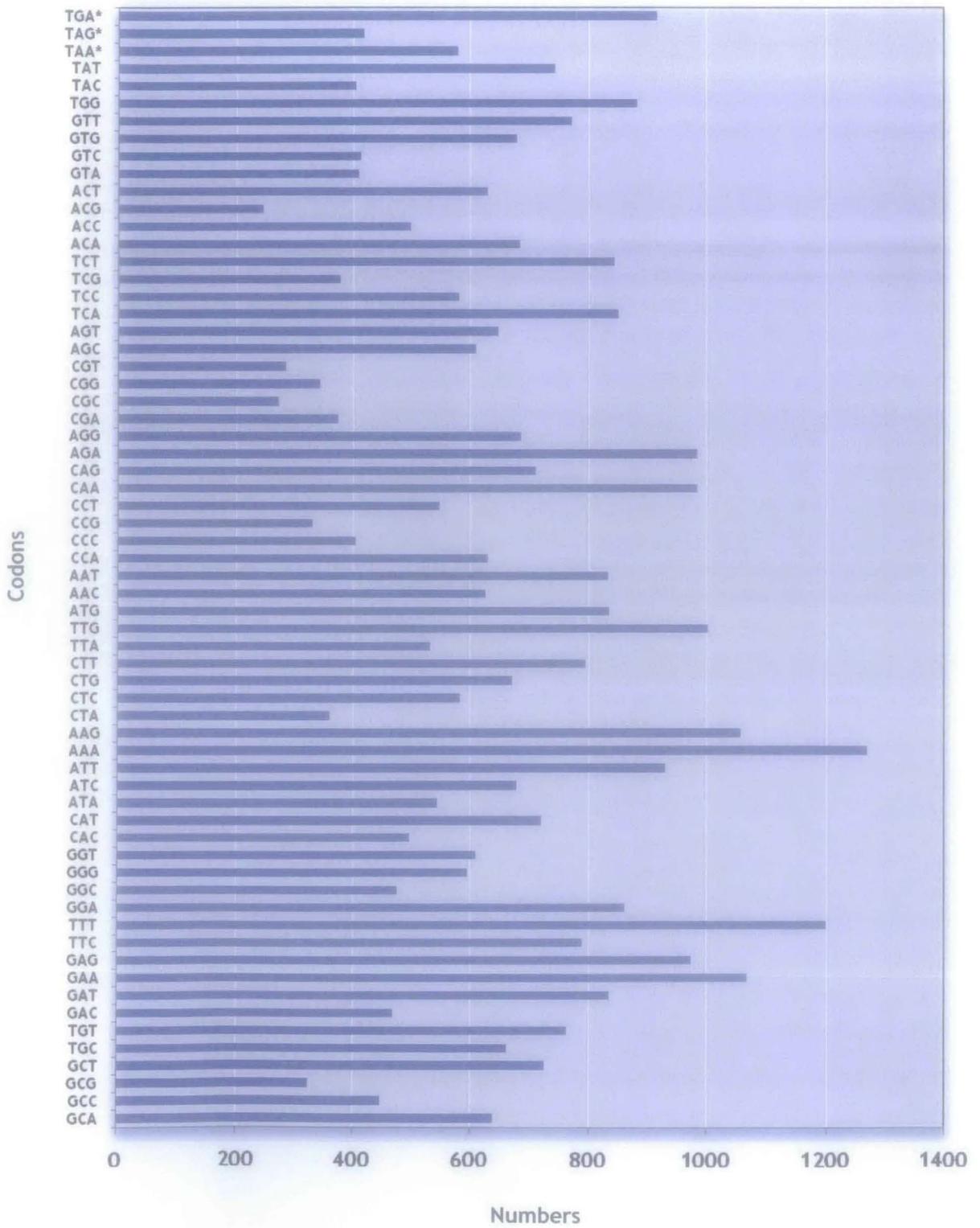


Fig 3.11: Codon usage analyses of drought induced unigenes

on homology search in CDD at defined filtration criteria (Table 3.4). Out of these domains, 16 were of multi-domain in nature which included copper amine oxidase super family, peroxidase like super family, LEA 3 superfamily, NAM super family, cyclase super family etc. Finally, 65 numbers of unique domains were subjected for the construction of maximum likelihood tree. The domains were clustered into 9 different groups and 27 out groups (Fig. 3.12). Group VII incorporated the highest 10 numbers of members having HSP 70, HSP 90, activator of HSP90 ATPase, alpha crystallin, GST etc. and group II, IV and IX incorporated the lowest 2 members having P-loop NTPase and mago nashi protein, Ras like GTPase and MIF as well as MPP and rcd1, respectively.

### 3.3.4 Gene ontology analysis of drought induced unigenes

A total of 187 unigenes (76%) were assigned to functional categories namely cellular components, biological process and molecular functions based on *Arabidopsis* proteome hits (Fig. 3.13, Appendix-D 2.1). In cellular components category, other intracellular components related genes recorded highest in numbers (19.65%) followed by other cytoplasmic components related genes (16.78%). However, endoplasmic reticulum related genes were found least in number (0.88%) followed by cell wall and extracellular related genes with 1.55% each. Highest number of genes in biological process category was recorded for other cellular processes (20.83%) followed by other metabolic processes (18.28%). Besides, signal transduction genes were found in least number (0.95%) followed by electron transport or energy pathway related genes (1.11%). Interestingly, 13.04% of genes responses to stress and 11.45% of genes responses to abiotic or biotic stimulus were also recorded. In molecular function category, the highest percentage was covered by other binding related genes (12.90%) followed by transferase activity related genes (12.20%), in contrast, transcription factor activity related genes were of least in number (2.44%) followed by nucleic acid binding related genes (2.79%).

### 3.3.5 Analyses of drought associated genes and pathways

A total of 123 numbers of drought associated genes were identified in the roots. Details of the identified genes and their protein products as well as corresponding

Table 3.4: Drought induced unigenes containing protein domain, signal peptide and signal anchor

Unigene IDs	Length	ORF loc*	Protein family	E-value	CD*	SP (loc)*	PA (loc)*
UEST2	746	98-565	hsp20/alpha crystallin family	9.30E-33	Y	N	N
UEST4	759	127-579	hsp20/alpha crystallin family	1.20E-33	Y	N	P (15-16)
UEST7	803	73-348	aminotransferase class I and II Domain	6.1E-28	Y	IEL-EI (20-21)	P (14-15)
UEST7	803	364-801	B12D protein family	3.4E-49	N	N	N
UEST9	611	61-609	hsp90 protein family	9.50E-80	Y	N	P (16-17)
UEST10	382	10-381	ribosomal protein L10	1.10E-16	Y	N	N
UEST11	709	75-709	YT521-B-like family	3.40E-17	Y	N	P (18-19)
UEST13	595	242-448	remorin, C-terminal region	1.10E-18	N	N	N
UEST18	989	30-740	tetratricopeptide repeat	2.50E-07	Y	N	P (18-19)
UEST20	695	109-567	hsp20/alpha crystallin family	5.80E-22	Y	TTA-LM (24-25)	N
UEST21	748	128-607	hsp20/alpha crystallin family	1.60E-33	Y	N	P (13-14)
UEST22	604	79-516	ribosomal protein S19e	9.70E-56	Y	N	N
UEST23	693	10-693	ankyrin repeat	6.50E-07	Y	N	P (20-21)
UEST24	720	69-479	core histone H2A/H2B/H3/H4	4.40E-33	Y	N	P (35-36)
UEST25	397	34-198	metallothionein	3.10E-11	N	N	P (21-22)
UEST28	631	85-468	thioredoxin	5.30E-32	Y	N	P (20-21)
UEST30	392	116-313	diacylglycerol binding domain	2.00E-05	N	N	N
UEST31	737	136-411	glutathione S-transferase	1.30E-16	Y	N	N
UEST35	581	49-528	cullin family	2.40E-52	Y	N	N
UEST38	382	24-382	raffinose synthase	2.10E-55	Y	N	N
UEST43	760	11-700	ubiquitin family	4.40E-34	Y	N	P (19-20)
UEST47	518	7-516	calcineurin-like phosphoesterase Domain	9.40E-11	Y	N	N
UEST50	517	27-517	oxygenase superfamily	7.30E-25	Y	N	P (20-21)
UEST53	872	138-650	ABC transporter	1.30E-13	N	N	P (21-22)

(Table 3.4 continued..)

UEST373	464	33-326	rdx family	2.50E-11	Y	N	N
UEST399	263	79-261	cytochrome c oxidase subunit III	1.40E-27	Y	N	N
UEST401	542	7-540	NADPH-dependent FMN reductase Family	4.30E-13	Y	N	P (24-25)
UEST402	501	6-501	nucleotide-sugar transporter Family	5.30E-05	N	N	N
UEST405	723	90-368	dienelactone hydrolase family	3.00E-06	N	N	N
UEST410	368	83-367	RNA recognition motif	4.50E-11	Y	N	P (19-20)
UEST412	382	3-605	RNA recognition motif	5.00E-11	Y	N	N
UEST418	484	13-483	RNA recognition motif	1.50E-13	Y	N	N
UEST429	510	63-332	sugar (and other) transporter family	1.30E-08	N	VMT-IF (34-35)	N
UEST430	682	117-587	heavy-metal-associated domain	1.60E-13	Y	N	P (17-18)
UEST432	644	2-583	D2-small domain of ClpB protein	2.50E-23	Y	N	N
UEST434	687	22-675	dehydrogenase E1 component	1.20E-16	Y	N	P (22-23)
UEST436	395	103-393	A20-like zinc finger family	7.30E-13	Y	N	P (30-31)
UEST445	505	94-504	nudix domain	3.80E-13	Y	N	P (20-21)
UEST447	507	17-406	hsp70 protein family	4.20E-49	Y	N	P (19-20)
UEST449	548	51-548	putative cyclase family	1.70E-12	Y	SDA-YP (28-29)	
UEST452	346	38-346	sin3 associated polypeptide p18 (SAP18)	4.70E-23	Y	N	P (56-57)
UEST453	382	69-335	protease inhibitor/seed storage/LTP family	6.00E-06	N	N	P (30-31)
UEST455	536	7-534	dioxygenase superfamily	1.10E-23	Y	N	P (22-23)
UEST459	545	170-544	integrase core domain	3.40E-09	Y	N	P (15-16)
UEST463	620	61-618	hsp70 protein family	1.40E-73	Y	N	P (15-16)
UEST466	603	49-603	SNW domain family	4.90E-36	Y	N	N
UEST484	395	71-304	ubiquitin-conjugating enzyme Domain	6.30E-24	Y	N	P (39-40)
UEST487	578	79-576	hsp70 protein family	2.40E-64	Y	AGG-VM (29-30)	N
UEST488	49	133-447	no apical meristem (NAM) protein family	2.10E-07	Y	N	N

(Table 3.4 continued..)

UEST490	512	35-397	macrophage migration inhibitory factor	3.60E-22	Y	N	P (19-20)
UEST492	595	106-414	mago nashi protein	2.70E-55	Y	N	N
UEST493	725	26-637	synaptobrevin family	4.60E-15	Y	N	N
UEST496	566	20-322	TFIIB transcription factor	3.60E-09	Y	N	N
UEST501	351	8-351	steroid binding domain	3.60E-17	Y	N	N
UEST502	389	143-505	sugar (and other) transporter family	4.00E-20	Y	N	P (20-21)
UEST505	355	100-306	ribosome associated membrane protein RAMP4	2.50E-27	Y	N	N
UEST509	477	35-379	ribosomal protein L19e family	4.80E-19	Y	N	P (17-18)
UEST513	445	77-433	DNA polymerase delta	3.70E-18	Y	N	P (50-51)
UEST517	678	55-501	nucleoside diphosphate kinase	3.10E-56	Y	N	N
UEST518	721	137-394	cell differentiation family	2.80E-44	Y	LNT-TS (18-19)	N
UEST519	762	1-666	Yip1 domain family	1.40E-08	Y	N	N
UEST522	668	21-668	glycosyltransferase family 17	3.50E-108	Y	N	P (26-27)
UEST524	497	4-246	leuA allosteric (dimerisation) domain	2.00E-11	Y	N	P (20-21)
UEST532	686	34-456	PAN-like domain family	4.10E-12	Y	N	P (27-28)
UEST533	748	247-720	pyridoxal binding domain	4.30E-24	Y	GSQ-IP (15-16)	N
UEST537	571	8-571	pyridine nucleotide-disulphide oxidoreductase	3.40E-20	Y	N	N
UEST540	367	11-367	zinc finger domain	1.40E-09	Y	N	N
UEST541	323	40-321	mitosis protein	6.20E-44	Y	N	P (20-21)
UEST544	666	60-524	phospholipase D	2.70E-33	Y	N	N
UEST545	422	33-422	glutamate-cysteine ligase family	6.20E-28	Y	N	N
UEST549	492	60-492	alpha/beta hydrolase fold domain	1.70E-06	Y	N	N
UEST551	587	60-587	triose-phosphate Transporter family	1.00E-10	Y	N	N
UEST554	697	59-697	ATP synthase subunit C	6.50E-16	Y	N	N

(Table 3.4 continued..)

UEST559	376	31-213	inorganic H <sup>+</sup> pyrophosphatase family	2.70E-22	Y	N	N
UEST570	684	116-684	RNA recognition motif	1.90E-18	Y	N	P (15-16)
UEST573	693	49-693	AP2 domain	8.20E-18	Y	N	N
UEST574	390	70-216	ADP-ribosylation factor family	3.40E-10	Y	N	P (32-33)
UEST575	717	303-557	peroxidase family	1.50E-05	Y	N	N
UEST579	503	21-503	copper amine oxidase	3.70E-44	Y	N	N
UEST582	655	190-480	late embryogenesis abundant protein	9.00E-36	Y	SLA-VT (22-23)	N
UEST583	299	74-229	ribosomal L39 protein	3.80E-22	Y	N	N
UEST587	380	40-264	UBA/TS-N domain	4.20E-09	N	N	N
UEST588	480	26-480	hsp20/alpha crystallin family	9.40E-09	Y	N	N
UEST591	348	114-348	copper/zinc superoxide dismutase	1.30E-22	Y	N	P (15-16)

---

\*ORF loc = ORF location, CD = conserved domain, SP (loc) = signal peptide location, PA (loc) = peptide anchor location, Y = present, N = absent

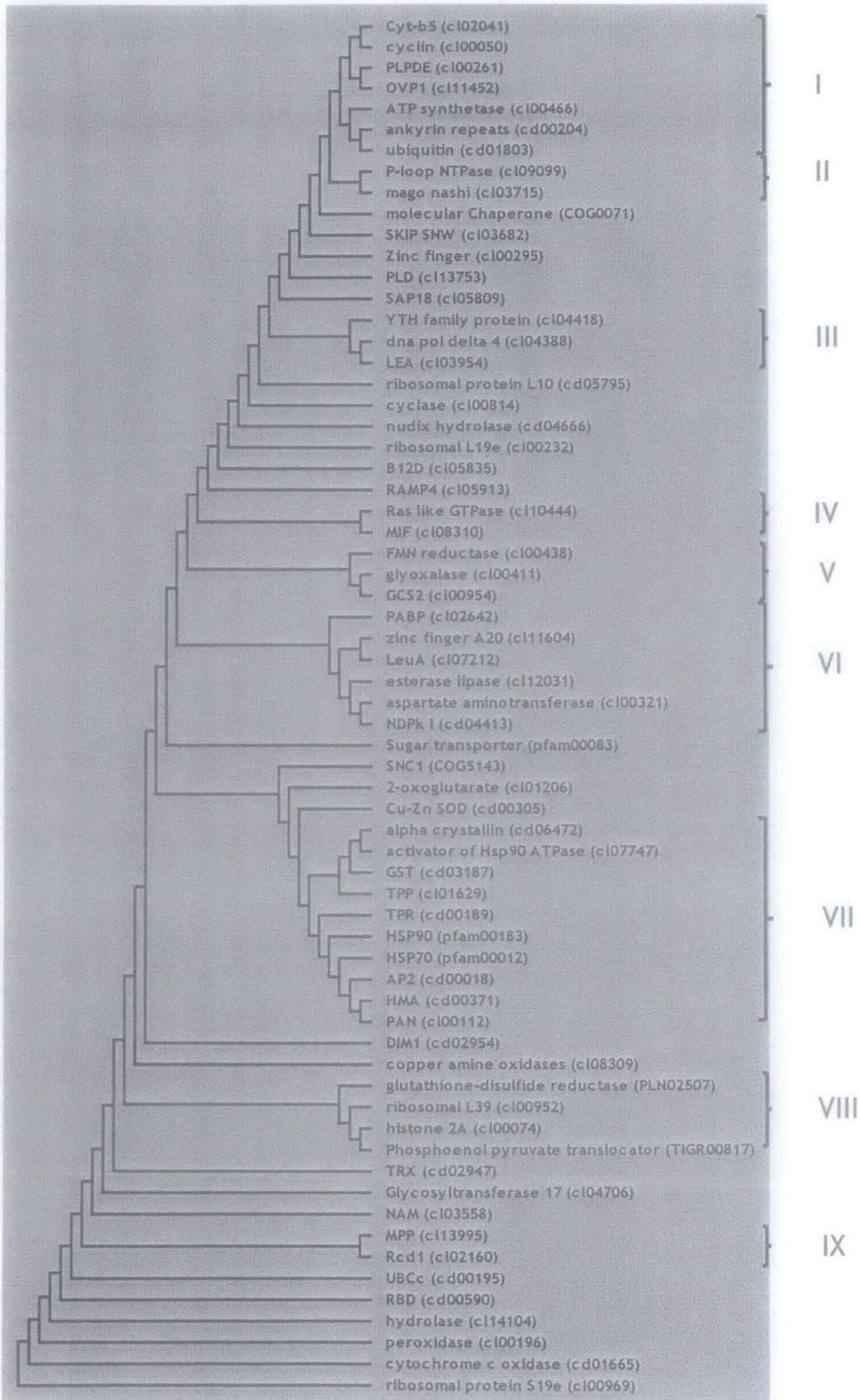
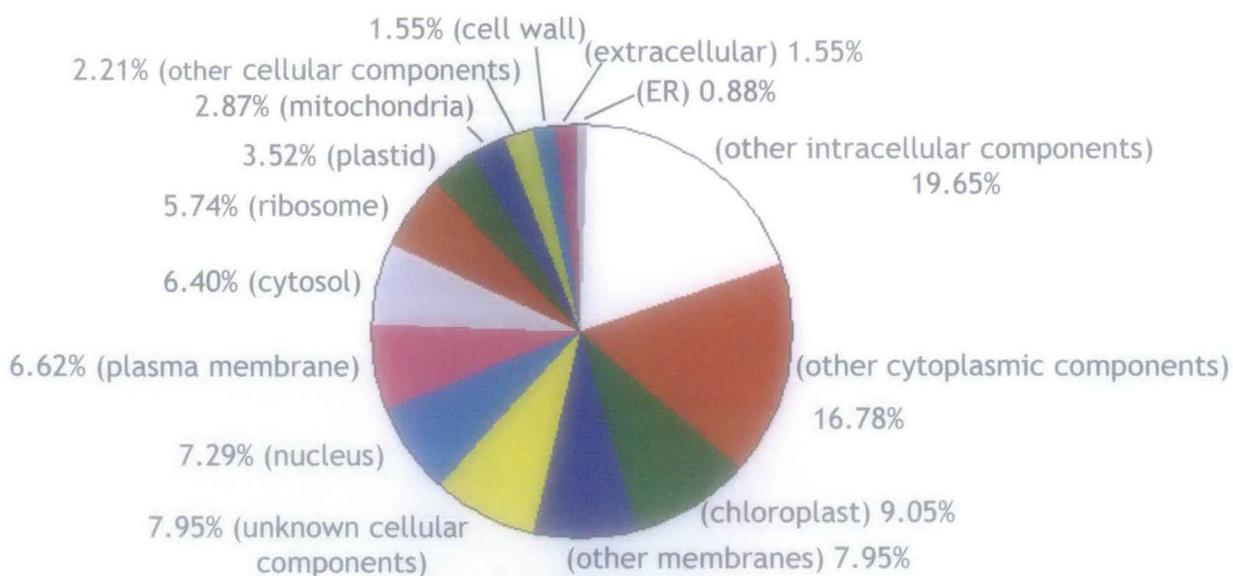
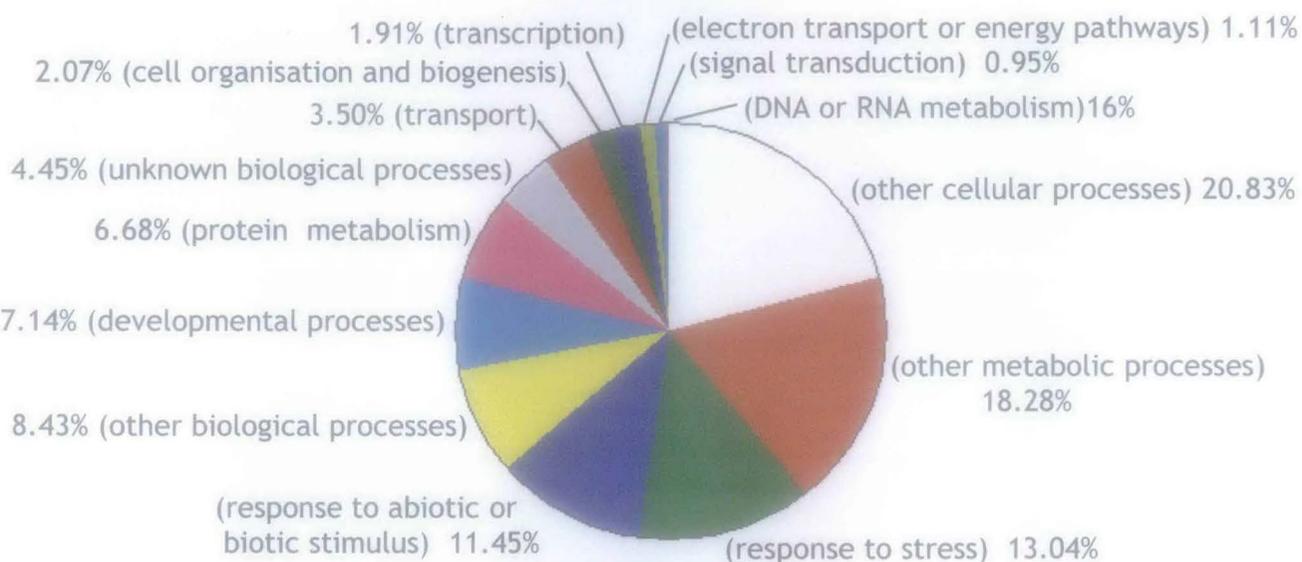


Fig. 3.12: Phylogenetic tree of drought associated conserved protein domain sequences.

### Cellular component



### Biological process



### Molecular function

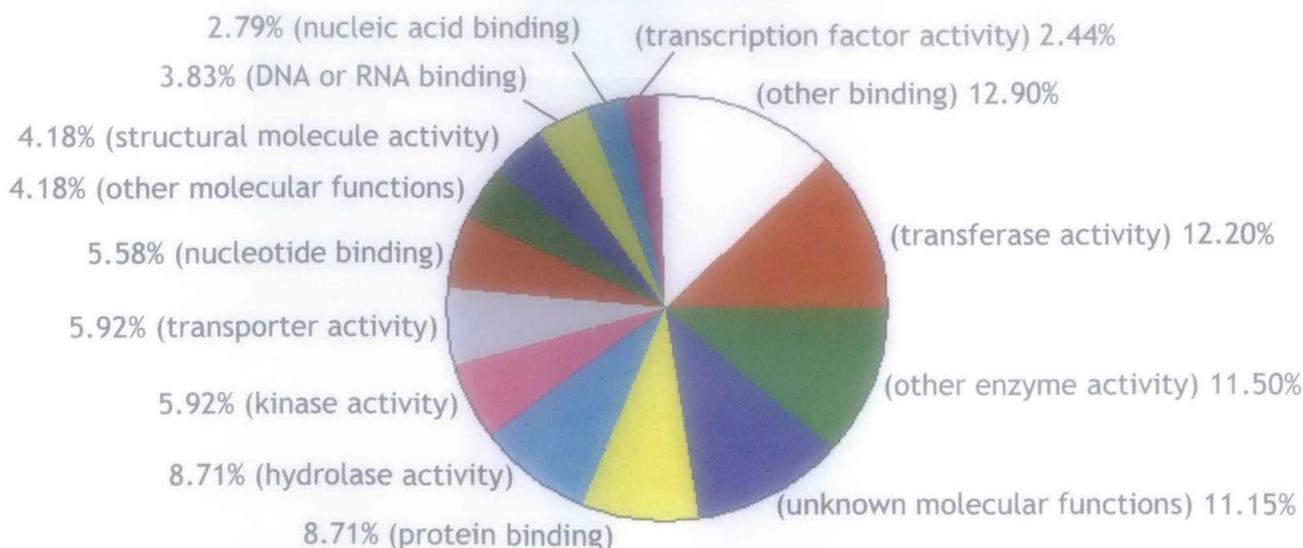


Fig.3.13: Gene ontology classification of drought induced unigenes as defined for *Arabidopsis* proteome

*Arabidopsis* loci were given in Table 3.5. Identified genes were mostly encoded for enzymes (36.59%), transporters (12.19%), transcription factors (6.50%) amino acids (0.81%) and proteins related to various physiological processes (43.91%). It consisted of the well-known drought associated genes such as dehydrin, trehalose phosphate synthase, GR, thioredoxin, ACC oxidase, peroxidase, heat shock proteins, SOD, zinc finger protein, ubiquitin etc. There were candidate genes of some of the important drought associated pathways.

### 3.3.5.1 Ubiquitin-proteasome pathways

Ubiquitination pathway plays an important role in eukaryotic cellular processes. General function of the pathway is to conjugate ubiquitin to Lys residues within substrate proteins, thus targeting them for degradation by the proteasome (Smalle and Vierstra 2004). The ubiquitin protein (76 amino acids) is attached to a substrate through a cascade of reactions consisting of three enzymes: E1 (ubiquitin activating enzyme), E2 (ubiquitin conjugating enzyme) and E3 (ubiquitin ligase). Each E3 has specificity to its substrate, or proteins to be targeted by ubiquitination. The E1 forms a thioester bond with the C terminus of ubiquitin (Ub) in an ATP-dependent manner and transfers the activated Ub to an E2 enzyme. The E2 enzyme either transfers ubiquitin directly or through indirect process. Ideally, this process is then repeated several times to attach multiple ubiquitin molecules to the substrate, and polyubiquitination has been shown to be necessary for degradation of the substrate by 26S proteasome (Moon et al. 2004).

The E3 ubiquitin ligases comprise a large and diverse family of proteins which are classified into four types: HECT type, U-box type, single RING-finger type which includes COP1, ARC1, and multi-subunit RING-finger type which include SCF, CUL3-BTB, and APC complexes (Zheng et al. 2002).

There were E2 (TG969274) and E3 enzymes as well as ubiquitin (GT969157) and polyubiquitin (Contig 43: GT969157, GT969158) found in the unigenes. The E3 enzyme was found as SCF complex comprising of f-box (GT969218), skp1 (GT969255) and cullin (Contig 35: GT969138, GT969139) proteins (Fig. 3.14). Moreover, some other related genes to this pathway were also reported such as zinc finger protein

Table 3.5: Identified ESTs of tea roots which were known to be associated with drought stress.

Genbank IDs Homology	<i>Arabidopsis</i> gene loci	Length	E-value
GT969327 glutathione reductase [ <i>Picrorhiza kurrooa</i> ](ACF93236.1)	AT3G24170.1	571	1.00E-70
GT969133 thioredoxin h1 [ <i>Glycine max</i> ] (ABV71991.1)	AT3G51030.1	501	2.00E-48
GT969341 glucose-6-phosphate [ <i>Pisum sativum</i> ] (AAC08525.1)	AT5G54800.1	587	1.00E-82
GT969343 dehydrin [ <i>Camellia sinensis</i> ] (ACJ65691.1)	AT3G50980.1	239	5.00E-15
GT969344 vacuolar H <sup>+</sup> -ATPase C subunit [ <i>Plantago major</i> ] (CAH58637.1)	AT4G34720.1	697	2.00E-47
GT969349 predicted H <sup>+</sup> -pyrophosphatase [ <i>Vitis vinifera</i> ] (XP_002282358.1)	AT1G15690.1	376	5.00E-30
GT969164 glutathione S-transferase 5 [ <i>Vitis vinifera</i> ] (ABW34390.1)	AT2G30860.1	737	4.00E-54
GT969354 UDP-glucose 4-epimerase, putative [ <i>Ricinus communis</i> ] (XP_002529901.1)	AT2G34850.1	407	4.00E-19
GT969365 peroxidase a [ <i>Eucommia ulmoides</i> ] (AAU04879.1)	AT3G21770.1	717	1.00E-74
GT969366 8-amino-7-oxononanoate synthase-like protein [ <i>Arabidopsis thaliana</i> ] (CAB85568.1)	AT5G04620.2	737	2.00E-11
GT969159 trehalose-6-phosphate synthase [ <i>Solanum lycopersicum</i> ] (BAF98176.1)	AT1G70290.1	352	6.00E-14
GT969176 ACC oxidase [ <i>Camellia sinensis</i> ] (ABI33224.1)	AT1G62380.1	517	9.00E-77
GT969368 GTP-binding protein [ <i>Helianthus annuus</i> ] (AAM12880.1)	AT5G55190.1	488	6.00E-22
GT969369 amine oxidase (copper-containing) precursor, putative [ <i>Ricinus communis</i> ] (XP_002509596.1)	AT4G12280.1	503	1.00E-31
GT969308 cell differentiation protein rcd1, putative [ <i>Ricinus communis</i> ] (XP_002529459.1)	AT3G20800.1	721	1.00E-82
GT969157 ubiquitin ligase-like protein [ <i>Oryza sativa</i> ] (ABS18744.1)	AT4G05320.2	688	1.00E-19
GT969381 Cu/Zn superoxide dismutase [ <i>Camellia sinensis</i> ] (AAU08173.1)	AT1G08830.1	348	4.00E-50

GT969070	heat shock protein 83 [ <i>Ipomoea nil</i> ] (P51819.1)	AT5G52640.1	611	4.00E-82
GT969378	mitochondrial small heat shock protein [ <i>Solanum lycopersicum</i> ] (BAA32547.1)	AT4G25200.1	480	1.00E-37
GT969311	protein transporter [ <i>Zea mays</i> ] (ACG24610.1)	AT4G26670.1	392	4.00E-19
GT969386	ormdl, putative [ <i>Ricinus communis</i> ] (XP_002523307.1)	AT5G42000.1	619	5.00E-77
GT969372	late embryogenesis abundant protein 5 [ <i>Nicotiana tabacum</i> ] (AAC06242.1)	AT1G02820.1	655	1.00E-19
GT969363	putative ethylene responsive element binding protein 2 [ <i>Gossypium hirsutum</i> ] AAX68525.1)	AT3G16770.1	693	3.00E-51
GT969361	NAC domain protein [ <i>Populus trichocarpa</i> ] (XP_002297860.1)	AT2G33480.2	509	7.00E-10
GT969329	chitinase [ <i>Camellia sinensis</i> ] (ACX42261.1)	AT3G12500.1	681	3.00E-112
GT969191	benzoquinone reductase [ <i>Gossypium hirsutum</i> ] (ABN12321.1)	AT4G27270.1	542	2.00E-83
GT969274	ubiquitin-conjugating enzyme [ <i>Hyacinthus orientalis</i> ] (AAT08675.1)	AT1G14400.1	395	1.00E-50
GT969166	profilin-3 [ <i>Hevea brasiliensis</i> ] (Q9M7N0.1)	AT2G19760.1	665	2.00E-60
GT969278	nam-like protein [ <i>Camellia sinensis</i> ] (ACH87170.1)	AT1G77450.1	449	7.00E-39
GT969253	heat shock protein 70 [ <i>Nicotiana tabacum</i> ] (CAA44820.1)	AT1G16030.1	620	3.00E-87
GT969292	sugar transporter [ <i>Citrus unshiu</i> ] (AAN86062.1)	AT1G67300.1	589	4.00E-69
GT969312	glycosyl transferase family 17 protein [ <i>Arabidopsis thaliana</i> ] (NP_189391.1)	AT1G12990.1	668	7.00E-79
GT969323	arginine decarboxylase [ <i>Humulus lupulus</i> ] (ACX56223.1 )	AT4G34710.1	748	4.00E-92
GT969254	adenosine kinase [ <i>Arabidopsis thaliana</i> ] (NP_181262.1)	AT2G37250.1	627	3.00E-49
GT969196	potassium channel beta, putative [ <i>Ricinus communis</i> ] (XP_002518142.1)	AT1G04690.1	705	1.00E-65
GT969079	remorin, C-terminal region [ <i>Medicago truncatula</i> ] (ABN08208.1 )	AT2G41870.1	337	1.00E-16
GT969087	L-aspartate [ <i>Arabidopsis thaliana</i> ] (NP_196713.1)	AT5G11520.1	539	1.00E-33
GT969206	ATP binding protein, putative [ <i>Ricinus communis</i> ] (XP_002531014.1)	AT5G16590.1	391	3.00E-22

GT969062	metallothionein-like protein [ <i>Camellia sinensis</i> ] (ABD97882.1 )	AT1G07600.1	397	2.00E-10
GT969219	hexose transporter 1 [ <i>Solanum lycopersicum</i> ] (ACX47459.1)	AT1G11260.1	510	5.00E-32
GT969220	metal ion binding protein, putative [ <i>Ricinus communis</i> ] (XP_002534419.1)	AT4G38580.1	682	3.00E-67
GT969121	predicted similar to putative ankyrin-repeat protein [ <i>Vitis vinifera</i> ] (XP_002283498.1 )	AT4G35450.2	544	6.00E-60
GT969167	purple acid phosphatase 1 [ <i>Solanum tuberosum</i> ] (AAT37529.1)	AT2G01880.1	518	2.00E-67
GT969225	1-deoxy-D-xylulose 5-phosphate synthase [ <i>Salvia miltiorrhiza</i> ] (ACF21004.1)	AT4G15560.1	424	3.00E-43
GT969021	zinc finger protein [ <i>Camellia sinensis</i> ] (ABI31653.1)	AT1G51200.2	395	2.00E-32
GT969228	calmodulin-related protein [ <i>Arabidopsis thaliana</i> ] (AAM67124.1)	AT1G66400.1	302	6.00E-17
GT969229	mitochondrial malate dehydrogenase [ <i>Solanum lycopersicum</i> ] (AAU29198.1)	AT3G15020.1	625	2.00E-73
GT969120	ozone-responsive stress-related protein, putative [ <i>Arabidopsis thaliana</i> ] (NP_171625.1)	AT1G01170.1	362	2.00E-19
GT969241	mitogen-activated protein kinase [ <i>Citrus sinensis</i> ] (ABM67698.1)	AT3G45640.1	717	7.00E-75
GT969245	lactoylglutathione lyase, putative [ <i>Ricinus communis</i> ] (XP_002518470.1)	AT1G11840.2	536	1.00E-91
GT969137	phosphate high response [ <i>Phaseolus vulgaris</i> ] (ACD13206.1 )	AT2G01060.2	696	3.00E-57
GT969248	glyceraldehyde 3-phosphate dehydrogenase [ <i>Daucus carota</i> ] (AAR84410.2)	AT3G04120.1	359	2.00E-09
GT969216	thioredoxin peroxidase [ <i>Capsicum annuum</i> ] (AAL35363.2)	AT1G65970.1	307	600E-20
GT969214	o-methyltransferase, putative [ <i>Ricinus communis</i> ] (XP_002514167.1)	AT4G26220.1	595	6.00E-67
GT969190	CAXIP1 protein [ <i>Arabidopsis thaliana</i> ] (AAO19647.1)	AT3G54900.1	549	2.00E-35
GT969184	AP-2 complex subunit beta-1, putative [ <i>Ricinus communis</i> ] (XP_002523245.1)	AT4G23460.1	583	2.00E-44
GT969202	ATPase activator/ chaperone binding [ <i>Arabidopsis thaliana</i> ] (NP_200619.1)	AT5G58110.1	682	2.00E-53
GT969208	nucleolysin tia-1, putative [ <i>Ricinus communis</i> ] (XP_002529199.1)	AT3G14100.1	484	3.00E-69
GT969135	arginine/serine-rich splicing factor, putative, expressed [ <i>Oryza sativa</i> ] (ABF96333.1)	AT5G64200.1	511	2.00E-14

GT969163	selenium binding [ <i>Arabidopsis thaliana</i> ] (NP_190314.2)	AT3G47300.1	464	5.00E-19
GT969217	NADK2; NAD+ kinase/ calmodulin binding [ <i>Arabidopsis thaliana</i> ] (NP_564145.1)	AT1G21640.1	446	300E-14
GT969048	galactolipase/ phospholipase [ <i>Arabidopsis thaliana</i> ] (NP_176378.4 )	AT1G61850.1	375	5.00E-26
GT969192	CMP-sialic acid transporter, putative [ <i>Ricinus communis</i> ] (XP_002523445.1)	AT4G35335.1	501	200E-48
GT969189	cytochrome c oxidase subunit 3 [ <i>Carica papaya</i> ] (YP_002608210.1)	AT2G07687.1	263	4.00E-20
GT969255	skp1 [ <i>Medicago sativa</i> ] (AAD34458.1)	AT1G20140.1	573	7.00E-13
GT969199	ATVAMP726 [ <i>Arabidopsis thaliana</i> ] (NP_171968.1)	AT1G04760.1	178	2.00E-15
GT969256	nuclear protein skip, putative [ <i>Ricinus communis</i> ] (XP_002530607.1 )	AT1G77180.1	603	4.00E-74
GT969270	AAA-type ATPase family protein [ <i>Arabidopsis thaliana</i> ] (NP_192327.3)	AT4G04180.1	487	2.00E-22
GT969200	U1snRNP-specific protein, U1A [ <i>Solanum tuberosum</i> ] (CAA90282.1)	AT2G47580.1	368	200E-40
GT969161	global transcription factor group [ <i>Populus trichocarpa</i> ] (XP_002313759.1)	AT4G08350.1	832	2.00E-118
GT969275	phosphatidylinositol synthase, putative [ <i>Ricinus communis</i> ] (XP_002533188.1)	AT1G68000.1	629	5.00E-26
GT969221	nucleotide binding protein, putative [ <i>Ricinus communis</i> ] (XP_002526531.1)	AT2G20330.1	371	9.00E-18
GT969282	mago nashi-like protein 1 [ <i>Physalis pubescens</i> ] (ABQ11262.1)	AT1G02140.1	595	2.00E-24
GT969306	translation initiation factor, putative [ <i>Ricinus communis</i> ] (XP_002514839.1)	AT5G36230.1	709	3.00E-82
GT969307	nucleoside diphosphate kinase [ <i>Nicotiana tabacum</i> ] (Q56E62.1)	AT4G09320.1	678	2.00E-72
GT969291	steroid binding protein, putative [ <i>Ricinus communis</i> ] (XP_002510808.1)	AT3G48890.1	351	1.00E-19
GT969350	inosine monophosphate dehydrogenase [ <i>Vigna unguiculata</i> ] (AAO40253.1)	AT1G16350.1	586	2.00E-40
GT969218	f-box family protein [ <i>Populus trichocarpa</i> ] (XP_002308717.1)	AT2G02310.1	507	4.00E-05
GT969296	cytidyltransferase family [ <i>Arabidopsis thaliana</i> ] (NP_175708.2)	AT1G53000.1	707	5.00E-71
GT969127	histone H3.2 [ <i>Arabidopsis thaliana</i> ] (NP_001078516.1)	AT4G40030.2	664	2.00E-70
GT969320	nudix hydrolase, putative [ <i>Ricinus communis</i> ] (XP_002512761.1)	AT1G79690.1	741	6.00E-78

GT969301	GID1-5 [ <i>Gossypium hirsutum</i> ] (ACN86360.1)	AT5G27320.1	424	1.00E-29
GT969305	radical-induced cell death 1 [ <i>Arabidopsis thaliana</i> ] (NP_564391.1)	AT2G35510.1	646	2.00E-17
GT969309	protein YIPF1, putative [ <i>Ricinus communis</i> ] (XP_002512371.1)	AT5G27490.1	762	4.00E-77
GT969223	rna-dependent RNA polymerase [ <i>Populus trichocarpa</i> ] (XP_002308662.1)	AT3G49500.1	554	6.00E-15
GT969314	predicted similar to putative isopropylmalate synthase [ <i>Vitis vinifera</i> ] (XP_002268942.1)	AT1G74040.1	497	5.00E-21
GT969315	glycosyltransferase UGT72B11 [ <i>Hieracium pilosella</i> ] (ACB56923.1)	AT4G01070.1	680	4.00E-14
GT969321	non-canonical ubiquitin conjugating enzyme, putative [ <i>Ricinus communis</i> ] (XP_002519212.1)	AT3G17000.1	562	4.00E-76
GT969175	calcium-binding EF hand family protein [ <i>Arabidopsis thaliana</i> ] (NP_173582.2)	AT1G21630.1	366	2.00E-32
GT968799	interferon-induced guanylate-binding protein, putative [ <i>Ricinus communis</i> ] (XP_002509420.1)	AT5G46070.1	509	6.00E-07
GT969155	transport protein particle (TRAPP) component Bet3 family protein [ <i>Arabidopsis thaliana</i> ] (NP_187151.1)	AT3G05000.1	523	9.00E-30
GT969335	gamma-glutamylcysteine synthetase [ <i>Zinnia elegans</i> ] (BAD27390.1)	AT4G23100.1	422	2.00E-73
GT969339	salicylic acid-binding protein 2 [ <i>Nicotiana tabacum</i> ] (gb AAR87711.1 )	AT2G23620.1	492	4.00E-51
GT969234	protein kinase APK1A, chloroplast precursor, putative [ <i>Ricinus communis</i> ] (XP_002531440.1)	AT2G28930.2	463	9.00E-70
GT969338	FUS-interacting serine-arginine-rich protein 1, putative [ <i>Ricinus communis</i> ] (XP_002520437.1)	AT1G07350.1	608	9.00E-37
GT969303	delta DNA polymerase, putative [ <i>Ricinus communis</i> ] (XP_002524033.1)	AT1G09815.1	445	7.00E-21
GT969284	leucine-rich repeat family protein [ <i>Glycine max</i> ] (ACM89476.1)	AT1G07650.1	699	5.00E-38

GT969280	light-inducible protein ATLS1 [ <i>Elaeis guineensis</i> ] (ACF06473.1)	AT5G01650.1	512	7.00E-19
GT969147	alkaline alpha-galactosidase [ <i>Pisum sativum</i> ] (ABR19752.1)	AT3G57520.3	382	1.00E-65
GT969360	putative oligouridylate-binding protein [ <i>Prunus dulcis</i> ] (ABR13303.1)	AT3G14100.1	684	6.00E-78
GT969091	light-harvesting complex II protein Lhcb5 [ <i>Populus trichocarpa</i> ] (XP_002329192.1)	AT4G10340.1	708	4.00E-65
GT969364	ADP-ribosylation factor [ <i>Elaeis guineensis</i> ] (ACF06579.1)	AT3G62290.1	390	1.00E-32
GT969233	SGT1 [ <i>Solanum tuberosum</i> ] (AAU04979.1)	AT4G11260.1	582	6.00E-35
GT969156	trafficking protein particle complex subunit 6b, putative [ <i>Ricinus communis</i> ] (XP_002512289.1)	AT3G05000.1	522	9.00E-30
GT969374	cdk8, putative [ <i>Ricinus communis</i> ] (XP_002518859.1   )	AT5G63610.1	355	6.00E-12
GT969224	2-oxoglutarate dehydrogenase, E1 subunit-like protein [ <i>Arabidopsis thaliana</i> ] (CAB75899.1)	AT3G55410.1	687	1.00E-48
GT969031	cysteine proteinase [ <i>Elaeis guineensis</i> ] (ABR19827.1)	AT1G47128.1	690	1.00E-46
GT969139	cullin, putative [ <i>Ricinus communis</i> ] (XP_002522416.1)	AT1G69670.1	505	2.00E-18
GT969108	AtPH1-like protein [ <i>Arabidopsis thaliana</i> ] (AAM61053.1)	AT5G05710.1	573	9.00E-46
GT969100	membrane-anchored ubiquitin-fold protein 2 [ <i>Arabidopsis thaliana</i> ] (NP_568315.1)	AT5G15460.2	534	4.00E-36
GT969152	ER Phosphatidate Phosphatase [ <i>Ricinus communis</i> ] (XP_002526289.1)	AT1G15080.1	570	3.00E-52
GT969097	starch phosphorylase L-2 [ <i>Solanum tuberosum</i> ] (P53535.1)	AT3G29320.1	518	8.00E-39
GT969076	ECT7 (evolutionarily conserved C-terminal region 6) [ <i>Arabidopsis thaliana</i> ] (NP_001078168.1)	AT3G17330.1	706	3.00E-73
GT969330	PRT1 [ <i>Arabidopsis thaliana</i> ] (AAM64697.1)	AT3G24800.1	367	8.00E-24
GT969286	transcription initiation factor IIB [ <i>Glycine max</i> ] (P48513.1)	AT3G10330.1	566	3.00E-43
GT969181	ATP-binding cassette sub-family f member 2 [ <i>Oryza sativa</i> Indica group]] (ABR25685.1)	AT5G60790.1	440	1.00E-31

GT969119	40S ribosomal protein S19 (RPS19C) [ <i>Arabidopsis thaliana</i> ] (NP_200925.1)	AT5G61170.1	604	4.00E-27
GT969334	phospholipase D alpha [ <i>Vitis vinifera</i> ] (ABC59316.1)	AT3G15730.1	666	5.00E-82
GT969072	60S acidic ribosomal protein PO [ <i>Euphorbia esula</i> ] (AAF34767.1)	AT2G40010.1	382	7.00E-21
GT969336	ubiquitin-associated (UBA)/TS-N domain-containing protein [ <i>Arabidopsis thaliana</i> ] (NP_180260.1)	AT2G26920.1	685	7.00E-26
GT969242	histone deacetylase complex subunit SAP18, putative [ <i>Ricinus communis</i> ] (XP_002511764.1)	AT2G45640.1	346	4.00E-29
GT969237	ATP binding [ <i>Arabidopsis thaliana</i> ] (NP_172382.4)	AT1G09080.1	407	8.00E-09
GT969073	60S acidic ribosomal protein P0, putative [ <i>Ricinus communis</i> ] (XP_002526873.1)	AT2G40010.1	382	2.00E-21

---

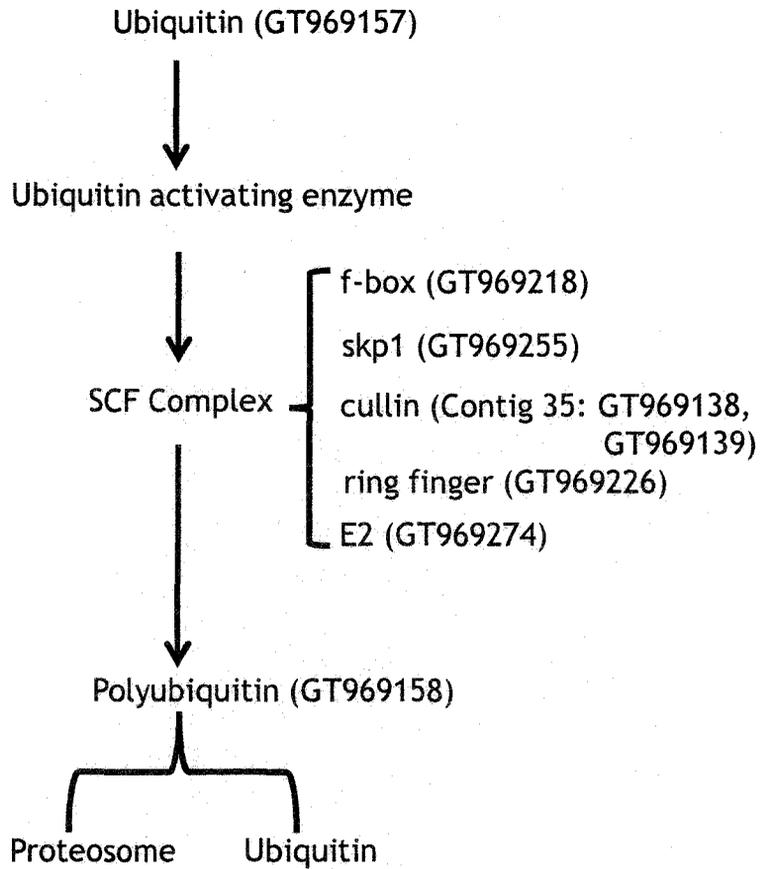


Fig 3.14: SCF complex mediated ubiquitin-proteasome degrading pathway associated ESTs of tea roots under drought stress.

(GT969226), ubiquitin specific protease (GT969211), non-canonical ubiquitin enzyme (GT969321) and 26S proteasome non-ATPase regulatory subunit 11 (GT969263).

### 3.3.5.2 Glutathione biosynthesis and metabolism

Glutathione (GSH) is the most abundant low-molecular weight thiol in the cellular redox system used for both detoxification of ROS and transmission of redox signals (Meyer 2008). The pathway of glutathione biosynthesis is well established: two sequential ATP-dependent reactions allow the synthesis of  $\gamma$ -glutamylcysteine from L-glutamate and L cysteine, followed by the formation of glutathione by addition of glycine to the C-terminal end of  $\gamma$ -glutamylcysteine. These reactions are catalysed by  $\gamma$ -glutamylcysteine synthetase and glutathione synthetase (Noctor et al. 2002).

Glutathione reduces disulfide bond formed within cytoplasmic proteins to cysteine by serving as an electron donor to form glutathione disulfide (GSSG). Under oxidative stress, the enzyme GR, reverts it from its oxidized form. GR catalyses the NADPH-dependent reduction of the disulphide bond of oxidized glutathione. GSSG can then be reduced via reversible thiol-disulphide interchange reactions (Noctor et al. 2002). Moreover, glutathione-s-transferases (GSTs) enzymes catalyze the conjugation of glutathione (GSH) to the electrophilic groups of a large variety of hydrophobic molecules to detoxify cells (Dixon et al. 2008).

There are genes of glutathione synthesis and metabolism i.e thioredoxin (GT969216),  $\gamma$ -glutamylcysteine synthetase (GT969335), GR (GT969327) and glutathione-s-transferase (Contig31: GT969164, GT969353) in drought induced unigene set.

### 3.3.5.3 Sugar synthesis, transport and metabolism

Sugar production in plants is a vital process. Sugar status in plants modulates and coordinates internal regulators and environmental cues that govern growth and development during the entire life cycle (Rolland et al. 2002). In recent years, the study on molecular mechanisms underlying sugar sensing and signaling as well as sugar metabolism in plants under environmental stresses has been gaining interest. There are several genes in the unigene set, protein products of which sense the sugar status in cells such as calmodulins (GT969228), MAP kinase (GT969241). Sugar

transporter acts as a sugar sensor and is involved in phloem loading. There are unigenes for sugar transporter (GT969292) and hexose transporter (GT969219) also. Moreover, there are important genes of glycolysis and gluconeogenesis such as glucose-6-phosphate (GT969341), glyceraldehydes 3-phosphate dehydrogenase (GT969248) and phosphoenolpyruvate carboxykinase (GT969259). Further, several unigenes related to sugar synthesis, metabolism or its related pathways were also found such as trehalose-6-phosphate synthase (GT968834), UDP-glucose 4-epimerase (GT969354), Alpha-1,4 glucan phosphorylase (GT968806), UDP-glucosyltransferase family 1 protein (GT969193), endo-1,3-1,4-beta-d-glucanase (GT969195), 1-deoxy-D-xylulose 5-phosphate synthase (GT969225), UDP-sulfoquinovose synthase (GT969246), glycosyl transferase family 17 protein (GT969312), UDP-glucose glucosyltransferase (GT969315) and inosine-5'-monophosphate dehydrogenase (GT969350).

#### 3.3.5.4 Genic microsatellite marker analysis

A total of 15 microsatellite repeat sequences were identified in drought induced root's unigenes. It has covered 6.09% of the total unigenes. They were designated according to the name and numbers of the corresponding unigene with additional 'M' for marker. Due to less flanking regions either one of the both site i.e 5' or 3'site, primer design could not be done for four identified SSRs. Details of the rest 11 SSR markers and SSR containing unigenes were given in Table 3.6, for which primer design was done successfully with an annealing temperature of 58°C, 59°C or 60°C. There was recorded one compound SSR (UESTM 500). In total, there were 6 dinucleotide, 5 trinucleotide (including the compound one) and 1 hexanucleotide repeats. The microsatellites were repeated from 3 to 17 times of which putative PCR product size ranges from 100-300 bp.

#### 3.3.6 Analyses of different steps of standard full-length library construction

The synthesized dsDNA was visualized as smear in gel ranges from 300 bp to more than 12 kb in size (Fig. 3.15). Absorption ratios of 260/280 and quantity of dsDNA were found as 1.8 and 40 ng/μl. The synthesized dsDNA was size fractionated into 14 fractions, out of which 9 fractions (4-12) yielded visible smear in gels (Fig. 3.16).

Table 3.6: Detection of EST-SSRs and putative functions of the SSR-containing drought induced unigenes

Loci	NCBI protein hits	Repeat units	Primers (5'---3')	Tm (°C)*
UESTM28	thioredoxin h1 [ <i>Glycine max</i> ] (ABV71991.1); 3E-48	(ATT) <sub>5</sub>	F-CACCCCAAGACTGAAGAAGAAAAGG R-CAATCTCTCATAAAAAACAAAGGAAATGC	59
UESTM384	homeodomain 20 transcription factor [ <i>Nicotiana attenuata</i> ] (ADI50265.1); 1E-06	(AG) <sub>12</sub>	F-TGGATAGAGACATGGAATGGTG R-CCGGGAGAATACAAAGTAACAGA	60
UESTM405	beta-d-glucanase, putative [ <i>Ricinus communis</i> ] (XP_002516793.1); 7E-76	(TC) <sub>9</sub>	F-TCCAGATATCTCTGAATCTCTCCTATTCCG R-ATAGAAACCAGCAGCCGCGAC	58
UESTM421	ubiquitin-specific protease 12 [ <i>Nicotiana tabacum</i> ] (ACJ04334.1); 7E-37	(TC) <sub>17</sub>	F-CACATTTGAGAAGCCAGTGAGA R-CCTGGTGTGGGGTTGTTG	60
UESTM457	unknown protein [ <i>Vitis vinifera</i> ] (CAQ58595.1); 3E-51	(TAA) <sub>5</sub>	F-GCTTTCTGCTTACTCTTCTGCC R-CATACGACACAGACTACAAACTGC	60
UESTM472	chlorophyll binding protein [ <i>Lycoris aurea</i> ] (ABJ99590.1); 7E-31	(AG) <sub>11</sub>	F-CGGATAATATGCCAAATAA ACTCTGCAC R-GGTTGGCCATGTTGGGTATGTTG	59
UESTM486	hypothetical protein [ <i>Vitis vinifera</i> ] (XP_002267006.1); 1E-11	(CAG) <sub>6</sub>	F-AGGTAATAGTTTGCCACCTCCA R-CTGTGTTGTTGGCTGATGACTT	60
UESTM498	no significant hit found	(GA) <sub>17</sub>	F-GCCCCAGAACAGAGAGAAACTA R-AGGAGGGAGATGAAGACACAAA	59
UESTM500	transmembrane emp24 domain containing protein 10, putative [ <i>Ricinus communis</i> ] (XP_002513789.1); 3E-09	(GA) <sub>9</sub> (TCT) <sub>5</sub>	F-CTGTCTTCTCACGGCTCAT R-GACGTTGTTCTGGATTTCTTC	60
UESTM512	no significant hit found	(GTTTGT) <sub>3</sub>	F-AGAAGCACCACCCTCTTTTA R-CACCTCTTTTACTCGGTCATCA	60
UESTM578	GTP binding protein [ <i>Helianthus annuus</i> ] (AAM12880.1); 6E-22	(GAT) <sub>5</sub>	F-CTCCAGAAGTGCAGATTGACAT R-ACAATTCCAAACCACCTACAC	60

\*Tm=annealing temperature



Fig. 3.15 (left) Size ranges of ds cDNA synthesized through LD PCR (D, LD PCR product)

Fig. 3.16 (right) Size fractionation of cDNAs through chroma spin 400 column (1-14 serial numbers of fractions collected in corresponding tubes; (M1= 1 kb DNA marker, M2= 100 bp DNA marker).

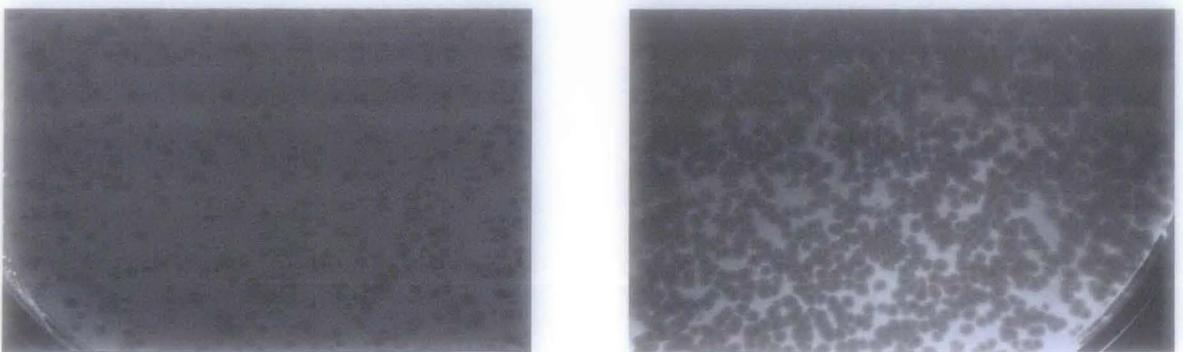


Fig. 3.17 Plaque forming units of unamplified (left) and amplified (right) full-length cDNA libraries.

The size ranges of DNAs in the first three fractions showing smear (4-6) was 400 bp to more than 10 kb.

Out of the three ligation reactions that set up with different vector to cDNA combinations, the 1.5:1 ratio of cDNA to vector ratio was found optimum for ligation. Titer values of the unamplified libraries of each ratio of cDNA to vector i.e 0.5:1, 1:1 and 1.5:1 were found as 0,  $1 \times 10^3$  and  $3.15 \times 10^5$  pfu/ml [Fig. 3.17 (left)], respectively. The unamplified library constructed by the 3<sup>rd</sup> cDNA to vector combination i.e 1.5:1 was selected and amplified subsequently for further use. Titering value of this amplified library was found  $1.823 \times 10^9$  pfu/ml [Fig. 3.17 (right)]. The recombinant percentage of the library was estimated as 83.8%.

### 3.3.6.1 Generation and clustering of full-length unigenes

A total of 1000 cell colonies were randomly picked, cultured and subsequently extracted plasmids. Sequencing of these has been produced a total of 811 ESTs which were deposited in dbEST, NCBI (Genbank accessions: GH623575-GH624058; HS389643-HS389969). Read length of sequences found in between 100 to 700 bp. There were found 586 numbers of vector cleaned and above 100 bp length sequences. Length of vector cleaned sequences ranges from 101 to 691 bp with an average length of 376 bp. Finally, 346 sequences were passed through quality parameters which included sequences of 46% no significant hits in database (160 out of 346), 8.38% unknown proteins (29 out of 346), 22.54% predicted proteins (78 out of 346), 1.73% chaperones and heat shock proteins (6 out of 346) and 21.10% enzymes as well other functional proteins. The quality sequences were clustered into 58 contigs comprising 2 to 28 individual ESTs in a single contig and 149 singlets producing a total of 207 unigenes (Fig. 3.18). Average length of the unigenes was found 401 bp where consensus sequences of contigs ranges from 103 to 685 bp and singlets ranges from 101 to 670 bp. Contigs containing more than 4 ESTs and e-value less than  $1E-10$  were depicted in Table 3.7. The unigenes were assigned a name as USR1, USR2 and so.

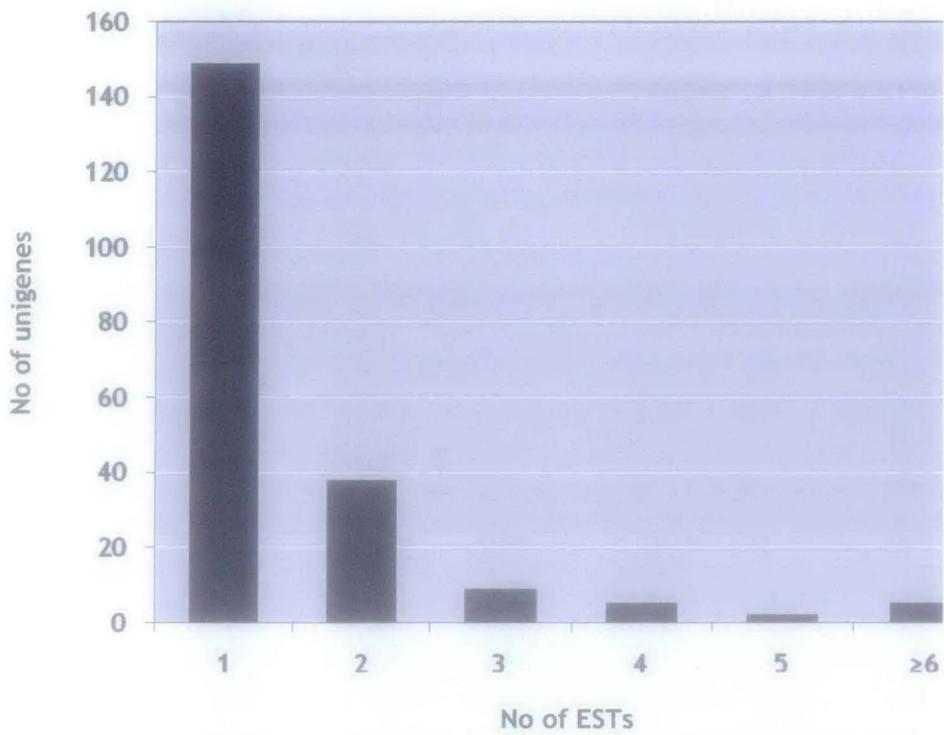


Fig. 3.18 Distribution and numbers of assembled standard full-length ESTs

Table 3.7: Assembled clusters that contain more than 4 full-length ESTs and e-value less than 1E-10

Contig IDs	Homology	No of ESTs	Organisms	Gene IDs	E-value	score
contig45	DnaJ protein	9	<i>Daucus carota</i>	AAG24642.1	253	1.00E-65
contig38	metallothionein-like protein	9	<i>Camellia sinensis</i>	ABD97882	93.6	3.00E-17
contig40	helicase, putative	5	<i>Ricinus communis</i>	XP_002520439.1	101	8.00E-20
contig36	xylem sap protein	4	<i>Solanum lycopersicum</i>	ADQ57297.1	173	1.00E-41
contig39	predicted protein	4	<i>Populus trichocarpa</i>	XP_002316984.1	125	1.00E-27
contig2	proline-rich protein	4	<i>Gossypium hirsutum</i>	ABM05952.1	139	2.00E-31
contig22	Acyl-CoA binding protein (ACBP)	4	<i>Digitalis lanata</i>	CAB56693.1	89	2.00E-16
contig30	conserved hypothetical protein	4	<i>Ricinus communis</i>	XP_002531502.1	97.4	5.00E-19

### 3.3.6.2 Functional annotation of full-length unigenes

Functional annotation of full-length unigene set revealed that there were 42% of overall GC content and 1272 numbers of cds containing 43% coding GC. Codon usage analysis showed that leucine (L) was the highest coded (9.92%) amino acid in comparison to others followed by serine (S) (8.02%) and glycine (G) (6.68%). Besides, the least coded amino acid was found as tryptophane (W) (2.0%) followed by methionine (M) (2.03%) (Fig. 3.19, Appendix-D 1.2).

A total of 35.75% unigenes (74 out of 207) was assigned to functional categories i.e cellular component, biological process and molecular functions as defined in *Arabidopsis* proteome (Fig. 3.20, Appendix-D 2.2). In cellular component category, other intracellular components related genes were found in highest numbers (20.24%) followed by other cytoplasmic components (14.88%) and unknown cellular components (13.31%) related genes. In this category, the least numbers were found golgi apparatus, extracellular and endoplasmic reticulum related genes each with 0.60% followed by mitochondria and cell wall related genes each with 2.38%. In biological process category, other cellular processes related gene was found in highest numbers (26.74%) followed by other metabolic processes (19.79%) and protein metabolism (13.90%) related genes. In this category, the least number was found signal transduction and transcription related genes each with 0.54% followed by DNA or RNA metabolism (1.07%) and transport (2.14%) related genes. In molecular function category, highest number of genes was found in other enzyme activity (23%) followed by other binding, protein binding and unknown molecular functions related genes each with 13%. In this category, least number of genes was found in transporter activity and transcription factor activity related each with 1% followed by kinase activity and nucleic acid binding related each with 2%. Interestingly, 6.95% of genes related to abiotic or biotic stimulus were also found under normal growth of roots.

### 3.3.6.3 Identification of redundant genes under normal growth and drought stress

Homolog analysis of unigenes between drought stressed and standard or normal growth of roots, 246 and 207 numbers, respectively, revealed 10 numbers of

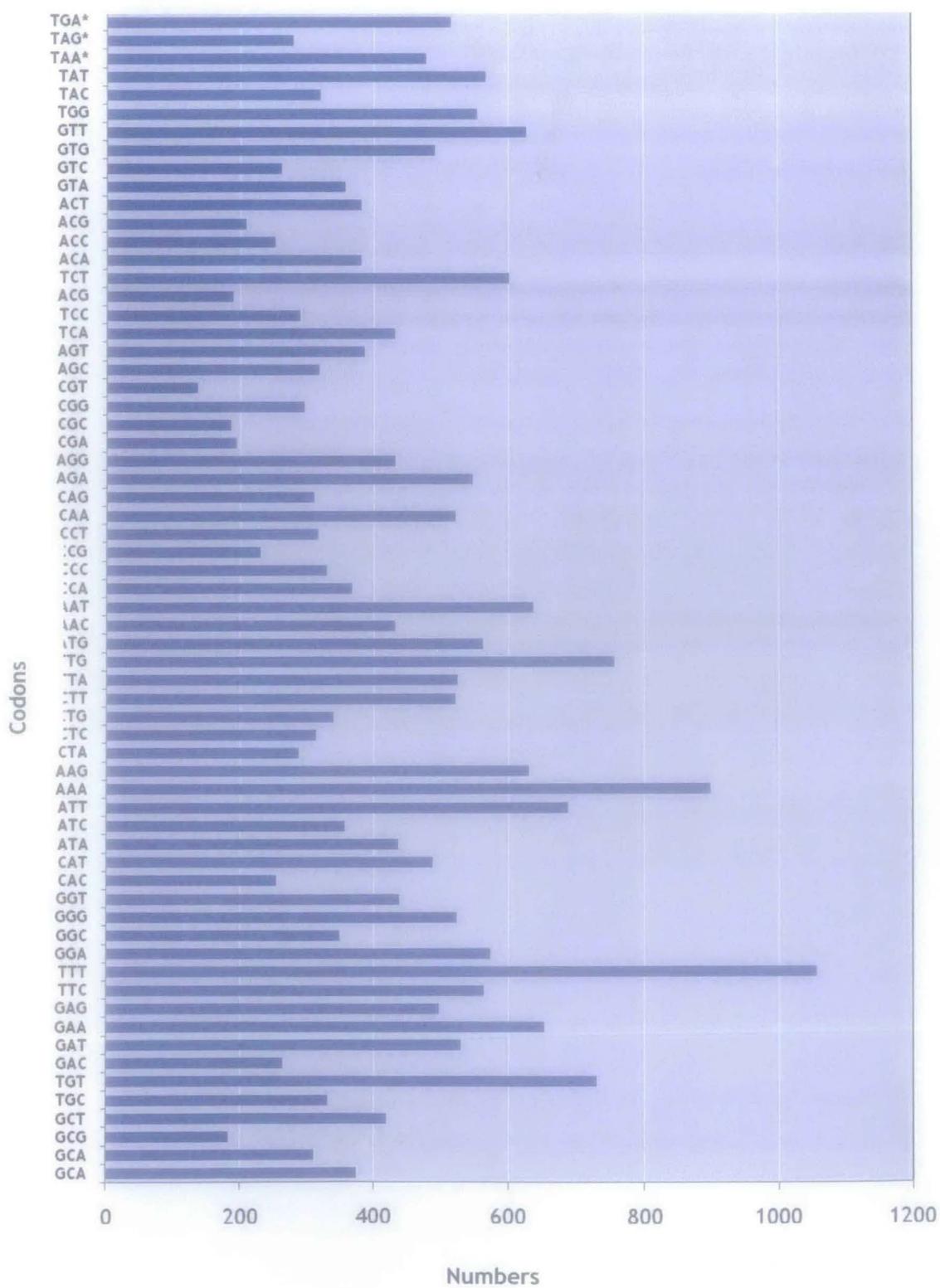
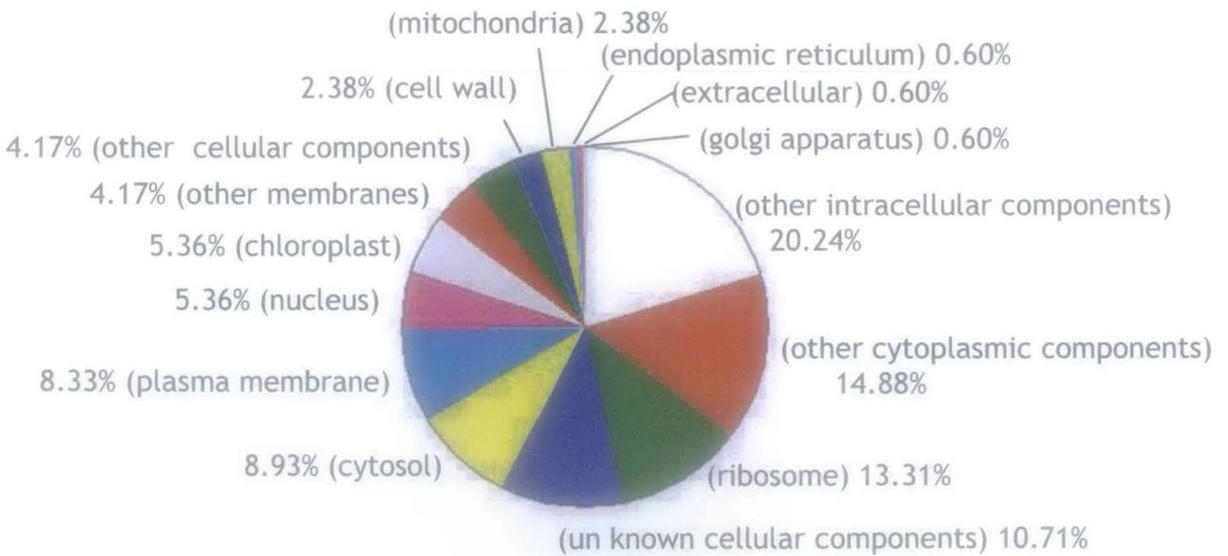
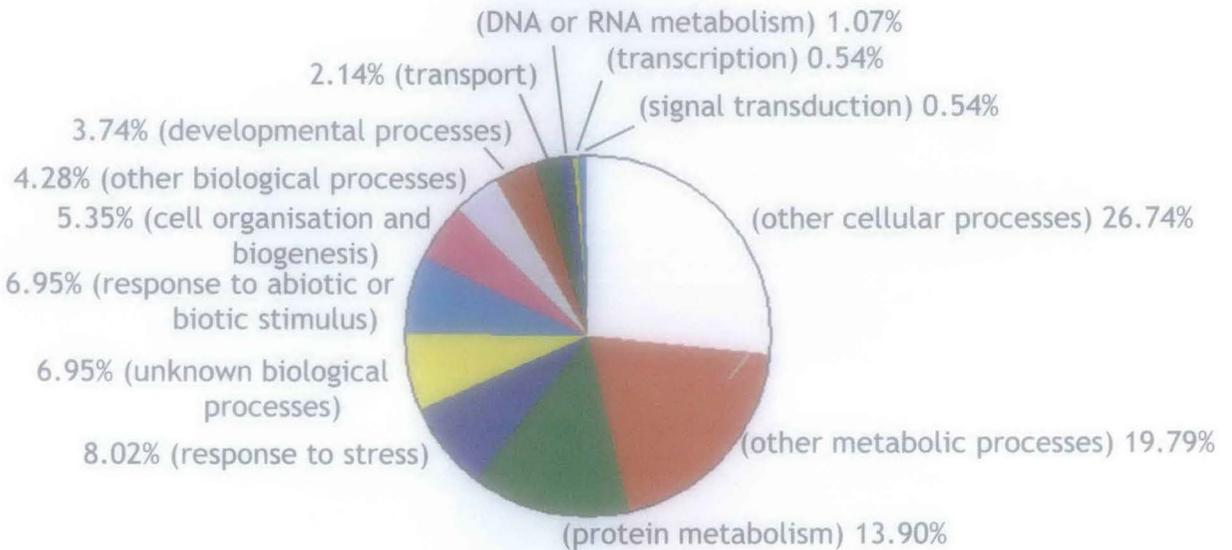


Fig. 3.19 Codon usage analysis of standard full-length unigenes

### Cellular component



### Biological process



### Molecular function

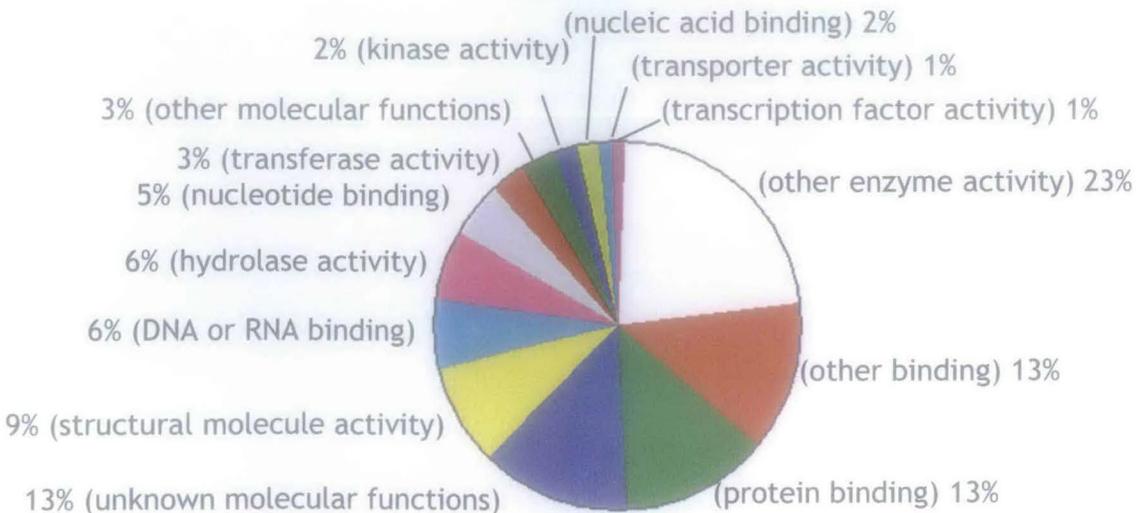


Fig.3.20: Gene ontology classification of full-length unigenes as defined for *Arabidopsis* proteome.

common or redundant unigenes. The identity percentage of the homolog genes was found more than 80 and e-value less than  $1E-05$  (Appendix-D 3.1 & 3.2). This result showed that 4.23% of drought induced genes were also available under normal growth of roots, on contrary, 4.83% of genes required for normal growth of roots were also induced by drought stress. The identified homolog genes coded for heat shock protein, profilin, lipid binding protein, 60S ribosomal protein L31 etc. (Fig. 3.21).

### 3.3.7 Comparative analysis of unigenes under normal growth, drought and winter dormancy stress

In order to comparative analyses of genes under normal growth, drought and winter dormancy stress, a total of 1568 tea leaf ESTs were downloaded from dbEST, NCBI. There were 451 ESTs (GH733750.1-GH734202.1) generated under normal growth and 665 (GH623183-GH623574; GH738509.1-GH738781.1) as well as 490 (FE942774.1-FE943102.1; FF682697.1-FF682833.1; GH454303.1-GH454326.1) ESTs generated under drought and winter dormancy stresses, respectively. Out of these ESTs, a total of 300, 321 and 478 ESTs for normal growth, drought and winter dormancy stress, respectively were passed through quality parameters. There were sequences of no significant hits in database, predicted proteins, unknown proteins, transcription factors, chaperone and heat shock proteins as well as enzymes and other functional proteins (Fig. 3.22). Unigenes were named as UWL for winter dormancy, as UDL for drought stress and USL for normal growth of leaves. Details of each unigene set such as average sequence length, GC content, coding sequences and highest coded amino acid were depicted in Table 3.8 (Appendix-D 1.3-1.5). Leucine was found as highest coded amino acid in normal growth, drought and winter dormancy stress. It was followed by arginine (6.70%), serine (9.12%) and lysine (8.49%) under normal growth, winter dormancy and drought stress, respectively.

A total of 60% unigenes under normal growth, 61% unigenes under winter dormancy stress and 32% of unigenes under drought stress were assigned to various functional categories as established in *Arabidopsis* (Appendix-D 2.3-2.5). Comparative analyses of functional categories among the unigenes of normal growth, winter dormancy and drought stress revealed interesting results (Fig. 3.23). In molecular function

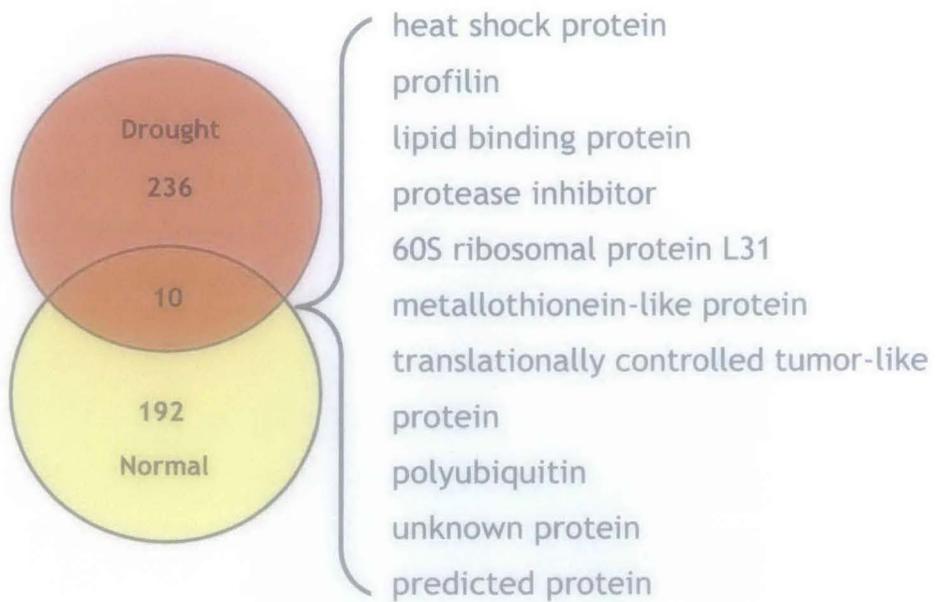


Fig. 3.21: Venn diagram of genes that expressed in roots under drought and normal conditions. The genes were put in three groups: genes of normal conditions, drought and in both drought as well as normal conditions of which the latter category genes were mentioned.

Table 3.8: Details of the ESTs under normal growth, drought and winter dormancy stresses of tea leaves

Conditions	Total ESTs	Quality ESTs	AL*	Unigenes (C+S)*	Coding GC (%)	cds count	Highly coded aa*
Normal growth	453	300	518	152 (24+128)	45	1117	Leucine (10.39%)
Winter dormancy	490	478	425	421 (39+382)	44	2775	Leucine (10.41%)
Drought	665	321	395	116 (45+71)	38	798	Leucine (9.64%)

\*AL= average length, Unigenes (C+S) = unigenes (contigs+singletons), aa=amino acids

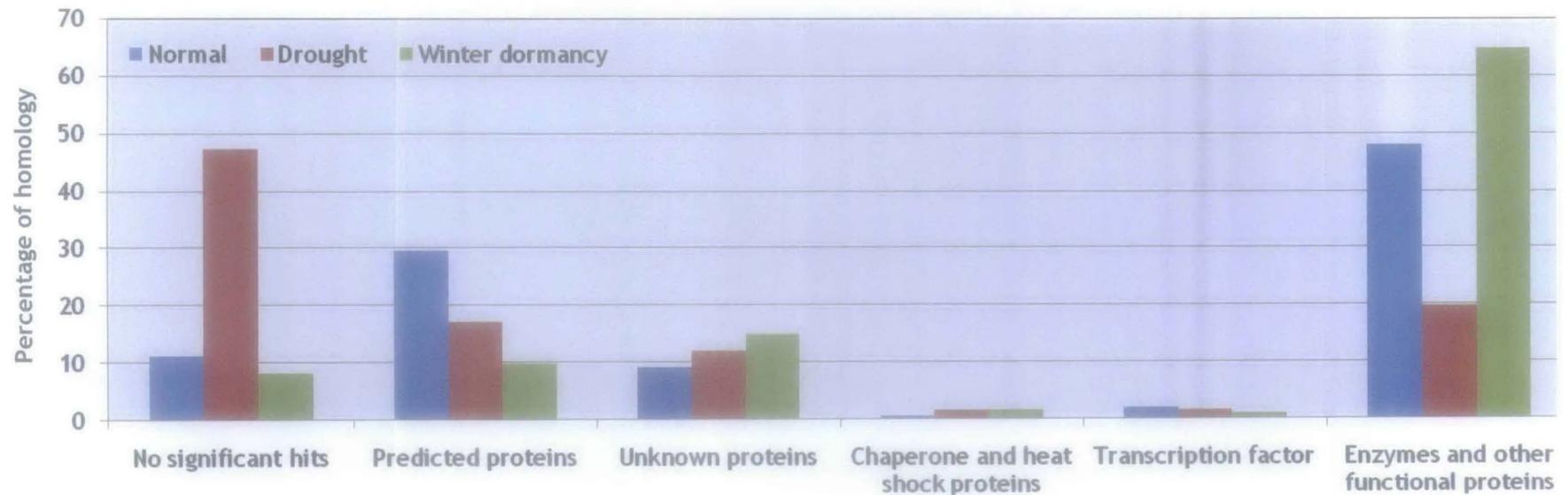


Fig. 3.22: BLASTx analysis of tea leaf unigenes under normal growth, drought and winter dormancy stresses.

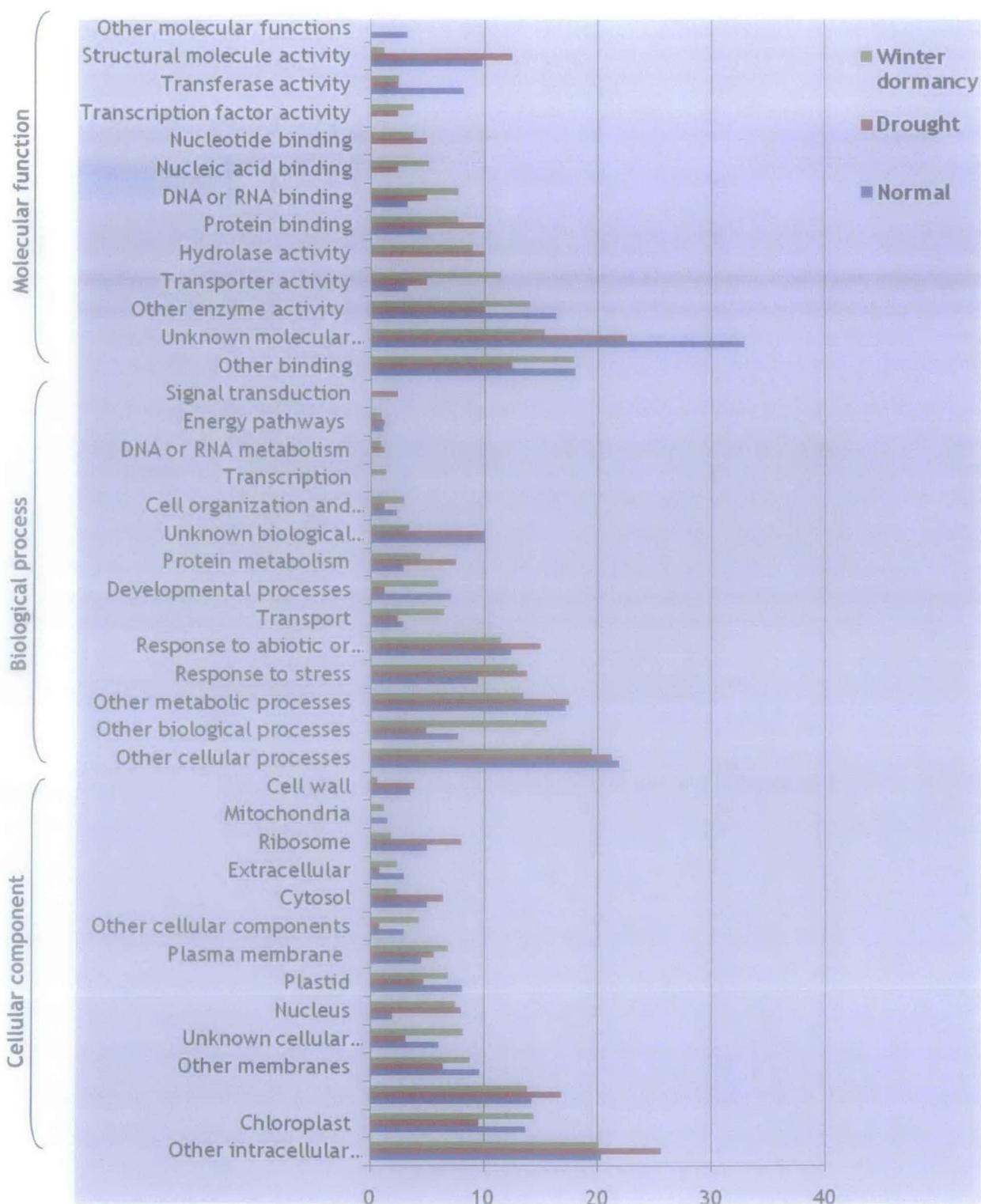


Fig. 3.23 Functional assignment of unigenes under normal growth, drought and winter dormancy stresses of tea leaves as defined for *Arabidopsis* proteome.

category highest numbers of genes were of unknown molecular functions followed by other binding and other enzyme activity related genes. Unknown molecular function related genes were of 32.79% in normal growth, 22.50% under drought and 15.39% under winter dormancy stress. Likewise, other binding related genes were of 18.03%, 12.50% and 17.95% in normal growth, drought and winter dormancy stress, respectively. In this category, least number of genes assigned to function was other molecular functions followed by transcription factor activity and nucleic acid binding. Other molecular functions related genes were found 3.28% only in normal growth. Besides, transcription factor related genes included 2.50% under drought and 3.85% under winter dormancy stress.

In biological process category, highest number of genes was of other cellular processes followed by other biological metabolic processes and response to abiotic or biotic stresses. Other cellular processes related genes were of 21.89% in normal growth, 21.25% under drought and 19.50% under winter dormancy stress. In abiotic or biotic stress response related genes were of 12.43% under normal growth, 15% under drought and 11.50% under winter dormancy stress. In this category, the least number of genes was related to transcription factor followed by DNA or RNA metabolism. Transcription factor related genes were found 1.50% only under winter dormancy stress. Besides, DNA or RNA metabolism related genes were found 1.25% under drought and 1.00% under winter dormancy stress.

In cellular component category, other intracellular components related genes were found highest in number followed by other cytoplasmic components. There were 20.31%, 25.60% and 20.13% of genes related to intracellular components under normal growth, drought and winter dormancy stress, respectively. Other cytoplasmic components related genes were found as 14.21% under normal growth, 16.80% under drought and 13.84% under winter dormancy stress. In this category, the least number of genes were found related to mitochondria followed by extracellular related genes. The mitochondrial related genes were found as 1.52% under normal growth and 1.26% under winter dormancy stress. Besides, extracellular related genes were found 3.05% under normal growth, 0.80% under drought and 2.52% under winter dormancy stress.

Homology search between unigenes of roots and leaves revealed that there were only 12 common unigenes including dehydrin, hexose transporter, histone etc. under drought stress (Fig. 3.24, Appendix-D 3.3 & 3.4). This result showed that 10.35% genes of leaves were also found in roots, on the other hand 5.13% genes of roots were also found in leaves under drought stress condition. That way, there were only 3 genes i.e hexose transporter, ribulose 1, 5-bisphosphate and predicted protein, in common between normal growth and drought stress of leaves (Fig. 3.25). This result indicated that only 1.97% genes of normal growth were found under drought stress of leaves. However, 11.18% genes of normal growth were found under winter dormancy stress (Fig. 3.25). Although, only 1.19% genes of winter dormancy stress were found under drought stress in leaves (Fig. 3.25) (Appendix-D 3.5-3.10).

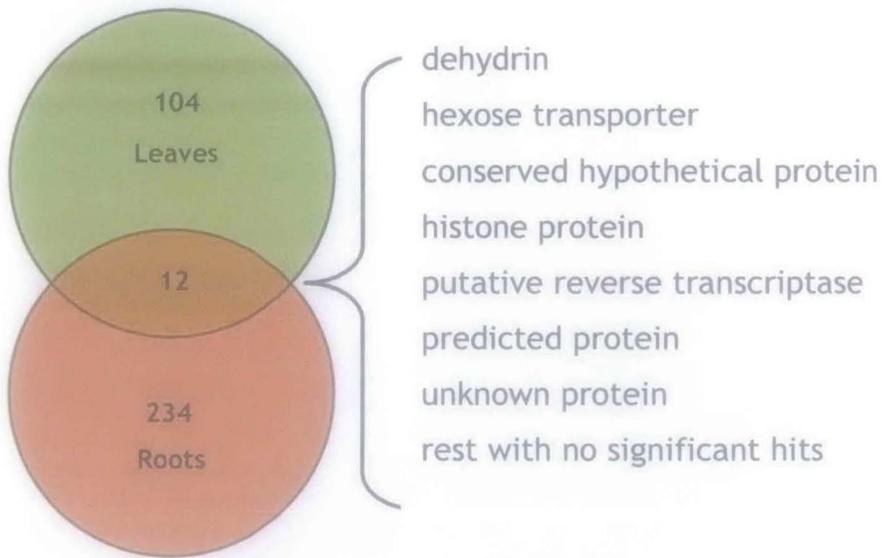


Fig. 3.24: Venn diagram of drought stress induced genes in roots and leaves. The genes were put in three groups: drought induced genes in roots, leaves and in both roots as well as leaves.

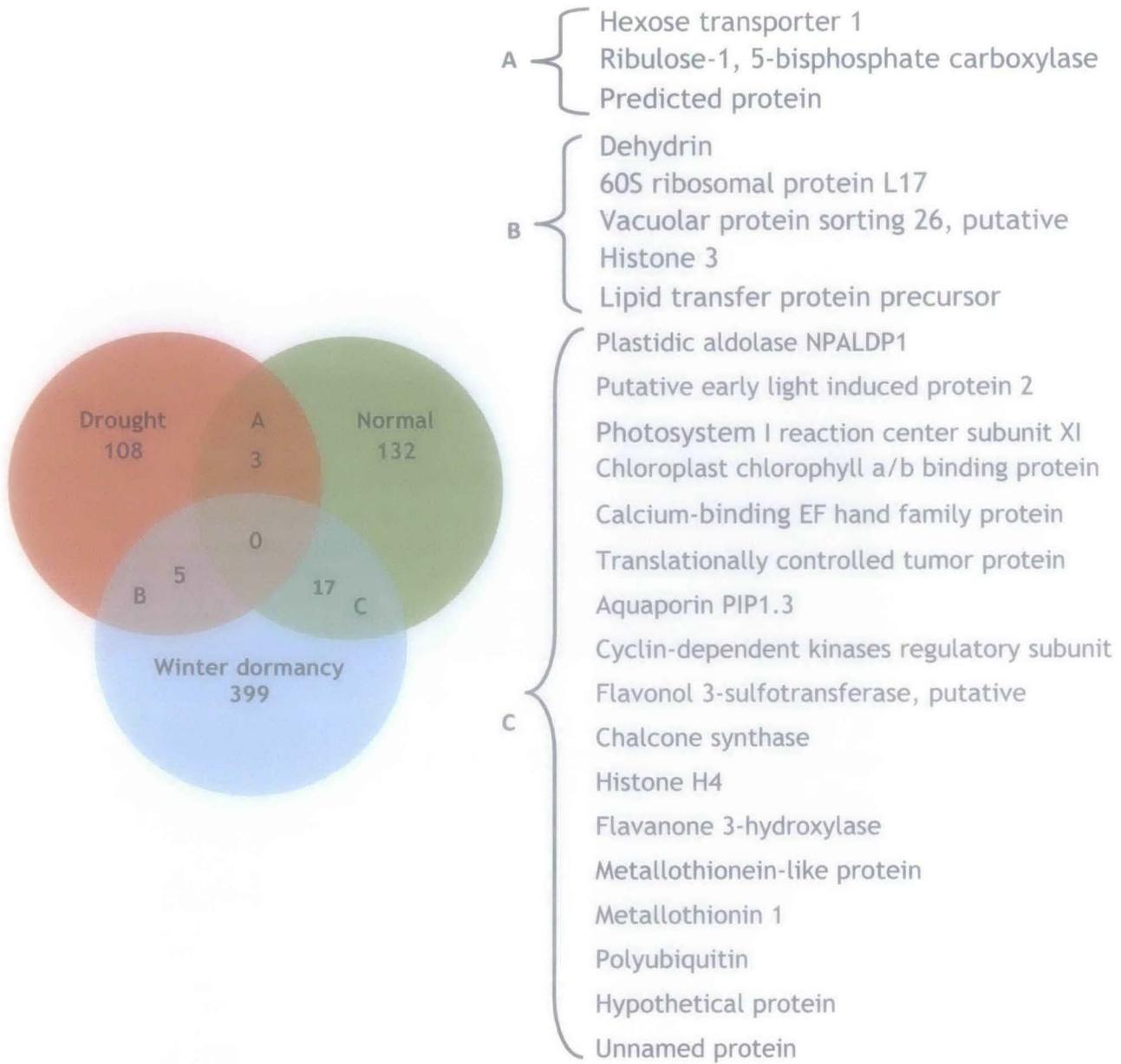


Fig. 3.25 Venn diagram of genes that expressed in leaves under drought, winter dormancy and normal conditions as well as those expressed mutually among these stresses.

### **3.4 Discussion**

Information in cDNA libraries is a powerful tool for identifying gene products of an organism. It lacks information about enhancers, introns and other regulatory elements found in a genomic DNA library. Construction of cDNA library, generation and analysis of ESTs has been proven to be a rapid and efficient way of obtaining the information on genes and its expression patterns (Alba et al. 2004). EST libraries and databases were successfully used in various crop plants for gene discovery, expression analysis, mapping and for comparative transcriptome analyses.

#### **3.4.1 Protocol optimization and RNA isolation**

Root is the plant part where initial stress signal from soil is perceived. Genetic investigation in root tissues of crop plant has been gaining importance to decipher the mechanism of different abiotic stresses. Plant adapted to environmental stresses such as drought by expressing specific stress related genes and metabolites (Alemzadeh et al. 2005). Primary need of gene expression studies is the isolation of quality RNA. Impurities in RNA hinder the process of downstream molecular manipulation such as RT-PCR, cDNA library construction (Azevedo et al. 2003). Woody plant tissue contains high level of polyphenol, polysaccharides and other secondary metabolites which create problem in RNA extraction. These compounds bind and precipitate with nucleic acids (Barron 1997, Basia and Arie 2005). It becomes more problematic in case of stress induced tissues that accumulate higher amount of secondary metabolites for resisting the adverse environmental conditions (Camacho-Villasana et al. 2002). Longer precipitation step increases the chances of impurities that co-precipitate with nucleic acids (Barron 1997). It was observed that there was absolutely no need of 3 to 4 h or overnight incubation period for precipitation as reported in most of the published papers (Scott et al. 1998, Gehrig et al. 2000, Jaiprakash et al. 2003).

#### **3.4.2 SSH library construction, analyses and functional annotation of ESTs**

Suppression subtractive hybridization technique used for construction of cDNA library was found efficient for subtracting differentially expressed cDNAs in tea

roots. Although traditional subtractive hybridization methods have been successful in some cases, they require several rounds of hybridization and are not well-suited for the identification of rare messages (Duguid and Dinauer 1990, Hara et al. 1991). Through the analyses of 572 ESTs, an overall picture of drought specific tea root's transcriptome was drawn. There were 94.06% of ESTs showed significant similarity to sequences present in public databases. The rest sequences (5.94%), without significant hits in the database might indicate their specific roles in tea roots under drought. There were also ESTs similar with unknown proteins (3.85%), however the number was much lower than that of a previous report (22.2%) in tea leaf subtractive library (Park et al. 2005). It may be due to the reason of large number of sequence deposits in the database in the successive years. Average EST's length (428 bp) was found more or less similar to previous findings in tea (Chen et al. 2005). There were most abundant ESTs for ribosomal proteins followed by molecular chaperones and heat shock proteins. Chen et al. (2005) also reported higher number of ribosomal ESTs in a tender shoots cDNA library of tea. Higher number of chaperone and heat shock protein ESTs was found may be due to the up-regulation of transcripts under drought stress. Surprisingly, there were ESTs for the chl a/b binding protein and embryo specific protein in the library. Similar finding was reported during the transcripts analysis of roots in white lupin (*Lupinus albus* L.) (Tian et al. 2009). It was presumed that either they have shared certain conserved motifs between the sequences or they are homologous sequences encoding chloroplast and leucoplast proteins. In present analysis, the coding GC was found higher (44.46%) than that of *Camellia irrawadiensis* (41.95%), although it was similar to apple (44%) (Newcomb et al. 2006) and other dicots represented in the codon usage database (Nakamura et al. 2000). There were 5.56% of non-redundant sequences without a significant hit in *Arabidopsis* genome. This result was similar to the finding in apple ESTs (5%) (Newcomb et al. 2006). Further, detection of ORFs in non-redundant sequences of 97.23% unigenes was similar to the report of Shu et al. (2009) in tree peony. It strongly suggests that most of the ESTs were truly captured and generated from the expressed part of the genome. Searching of domain structure in the ORF revealed 34.55% of unigenes containing a functional domain, most of which (85%) were conserved in nature. The most abundant domain was of heat shock protein domain. Number of unigenes containing a functional domain was lesser than that of tree peony (45.4%) (Shu et al. 2009). The identified conserved

domains suggested their role in drought tolerance in diverse plant species including tea. It includes AP2 domain, LEA super family, NAM super family, cullin family etc. which clustered into nine groups and several out groups in phylogenetic tree. It suggests that most of them have evolutionary relationship. In peptide sequences, signal peptide and anchor sequence position were mined. Since, signal peptides are responsible for directing the post-translational transport of a protein to their target organelles, identification of them based on the amino acid sequence and their validation is a big issue in proteomics. The identified peptide signal and anchor in the protein sequence such as TTA-LM in Hsp20 domain, SLA-VT in LEA domain and VMT-IF in sugar transporter domain, domain provides for the first time a basis for future research in tea.

#### 3.4.2.1 Gene ontology analysis

Gene ontology terms are commonly used for the functional characterization of genes and gene products and to facilitate queries among genes from different organisms (<http://www.geneontology.org/>). Ontology covers three domains: cellular component, the parts of a cell or its extracellular environment; molecular function, the elemental activities of a gene product at the molecular level, such as binding or catalysis; and biological process, operations or sets of molecular events with a defined beginning and end, pertinent to the functioning of integrated living units: cells, tissues, organs, and organisms (Botton et al. 2008). A total of 76% unigenes were assigned to diverse functional categories based on *Arabidopsis* proteome hits. There were reports of GO analysis in tea unigenes generated from leaf for secondary metabolism (Park et al. 2004) and shoots (Chen et al. 2005). However, there was no such report of GO analysis of transcripts generated under stress conditions for any kind of tissues. In present analysis, other intracellular components (19.64%) and other cytoplasmic components (16.78%) were found in highest numbers in cellular components category. It may be due to the rapid synthesis of osmolytes and other metabolites under drought. In normal conditions, other membrane proteins were found in highest number in white lupin (*Lupinus albus* L.) roots (Tian et al. 2009). In biological process category, results were almost similar to Tian et al. (2009) except a high abundance of transcripts for response to stress (13.03%) and response to abiotic or biotic stimulus (11.45%) which was expected. However, it accounted for

9.2% in tea leaf (Chen et al. 2005). Moreover, metabolism category reported 8.5% transcripts in tea leaf (Park et al. 2004) whereas it was found 18.28% in present analysis of roots under drought. In molecular function category, other binding (12.89%) and transferase activity (12.19%) proteins were abundant whereas, in lupin unknown molecular function proteins were found abundant (Tian et al. 2009).

#### 3.4.2.2 Drought associated genes and genetic pathways

Drought tolerance in plant is a complex trait that involves mechanisms to avoid or tolerate water deficit conditions. Genotypes responding differently to drought stress show differences in gene expression patterns of which a portion is involved in drought tolerance mechanism (Rabello et al. 2008). In present study, genes found exclusively in the tolerant genotype were reported for identifying the genes and genetic pathways associated to drought tolerance. Similar to *Arabidopsis* and Rice (Shinozaki and Yamaguchi-Shinozaki 2007, Rabbani et al. 2003), the identified genes for drought tolerance can also be classified into functional and regulatory groups depending on the proteins they code for. In functional category, ion channels and transporters such as vacuolar H<sup>+</sup>-ATPase, potassium channel beta, sugar transporter; detoxification enzymes such as Cu/Zn SOD, glutathione reductase, thioredoxin peroxidase, benzoquinone reductase, arginine decarboxylase; macromolecular protector such as LEA, chaperone; key enzymes for osmolyte biosynthesis such as glucose-6-phosphate, trehalose-6-phosphate synthase,  $\gamma$ -glutamylcysteine synthetase and proteases such as cysteine proteinase etc. were included. In regulatory proteins category, transcription factors such as ATP-binding cassette family protein, transcription initiation factor, zinc finger protein, NAC domain proteins; protein kinases and phosphatases such as MAP kinase, cdk8, ER Phosphatidate Phosphatase, phospholipase, purple acid phosphatase; ABA biosynthesis related proteins such as ethylene responsive element binding protein were included. First group of proteins is probably involved in abiotic stress tolerance and the second group of proteins is involved in regulation of signal transduction and stress responsive gene expression (Shinozaki and Yamaguchi-Shinozaki 2007). Some of the functional and regulatory proteins related to drought stress response and tolerance are discussed in this section and some are discussed with their related pathways in appropriate sections. ATPases, water channel proteins and ion transporters play a crucial role in

maintaining ion homeostasis under stress condition. Proton electrochemical gradient formed by the vacuolar ATPase provides the primary driving force for the transport of numerous ions and metabolites against their electrochemical gradients which is the fundamental requirement of many cellular processes, such as osmoregulation, signal transduction, and metabolic regulation (Sze et al. 1992). The SOD enzyme constitutes a component of the first line of cellular defense against oxidative stress by early scavenging superoxide radicals and converting them to hydrogen peroxide (Chatzidimitriadou et al. 2009). Later, the peroxidases catalyse the conversion of hydrogen peroxide to water using ascorbate, glutathione or thioredoxin as substrate (Rossel et al. 2006). The LEA proteins are involved in protecting from damage caused by environmental stresses, especially drought, which has six different groups (Goyal et al. 2005). Group two has three super families, one of which is known as dehydrins (Goyal et al. 2005). Heat-shock proteins (Hsps) or chaperones such as Hsp70 and Hsp 90 are responsible for protein folding, assembly, translocation and degradation in many normal cellular processes. These proteins stabilize other proteins and membranes and are actively involved in protein refolding under stress conditions including drought (Wang et al. 2004). Organic osmolytes such as amino acids and derivatives, polyols and sugars are used as compatible solutes under drought stress to maintain cell volumes. The enzyme, trehalose-6-phosphate synthase catalyses the first step of trehalose biosynthesis. Trehaloses play an important role in drought tolerance of plants (Lee et al. 2003). Some genes that participate in alterations of glucose metabolism as a result of the limitation caused by low levels of intracellular CO<sub>2</sub> observed during drought stress were also identified. These include phosphoenolpyruvate carboxykinase which converts pyruvate to phosphoenol- pyruvate and glyceraldehydes-3-phosphate dehydrogenase which converts 1,3-bis-phosphoglycerate to glyceraldehydes-3-phosphate during gluconeogenesis. There are genes of metallothionein, a superfamily of low molecular weight proteins that are involved in metal detoxification and scavenging of oxygen-free radicals, which can decrease injury in oxidative tissue (Talame` et al. 2007). Genes associated to maintenance of cell turgor and stress signaling were also identified such as NAD<sup>+</sup> kinase/calmodulin binding proteins and ATP binding proteins (Ranty et al. 2007). There are genes of ethylene responsive element binding protein. Ethylene is a well-characterized phytohormone that may act alone or in combination with ABA in regulating gene expression under drought (Ogawa et al. 2005). NAC

(NAM, ATAF1/2 and CUC2) domain proteins are plant-specific transcriptional factors known to play diverse roles in various plant developmental processes (Hu et al. 2006). NAC domain protein such as RD26 is induced under dehydration stress (Fujita et al. 2004). Zinc finger proteins are well known signal proteins that response to various stress such as drought, cold and salt (Mukhopadhyay et al. 2004). There are also genes of MAP kinases, which include a family of ser/thr protein kinases functioning in many signal transduction pathways. MAP kinase cascades a series of phosphorylation events to amplify and transmit signals (Qiang et al. 2000). Phospholipase D (PLD) and its product, phosphatidic acid, exert their effects by functioning in signal transduction cascades and by influencing the biophysical state of lipid membranes. It has been implicated in multiple stress responses in plants including drought (Bargmann and Munnik 2006). In general, plants respond to abiotic stresses including drought by regulating pathways that counteract oxidative damage and maintain intracellular redox environment (Hicks et al. 2007).

#### **3.4.2.2.1 Ubiquitin-proteasome degrading pathways**

Role of ubiquitination in plants for conferring drought tolerance is well documented in literature. In the eukaryotic ubiquitylation system, there are only one or two ubiquitin activating enzymes (E1), a few dozens of ubiquitin-conjugating enzymes (E2), and a large number of ubiquitin ligases (E3) (Yan et al. 2003). Out of all the ubiquitin ligases, SCF class of E3 ligases has been thoroughly studied in plants. In present investigation, all the genes of SCF complex of E3 were found which indicated its active role in drought tolerance of tea plants. E3 ligases provide the specificity based on which proteins to be ubiquitylated and degraded through the proteasome pathway (Ciechanover et al. 2000) are identified. Amino acid, lysine forms a conjugate with ubiquitin within the substrate and plays a vital role in selective proteasomal degradation (Moon et al. 2004). Drought tolerance in plants has been well-demonstrated in plants through the over expression of ubiquitin or other genes of ubiquitin-proteasome degrading pathways. Tobacco plants transformed with wheat (*Triticum aestivum*) ubiquitin, *Ta-Ub2* demonstrated well that the overexpression of monoubiquitin might be an effective strategy for enhancing drought tolerance in plants (Guo et al. 2008). Enhanced drought and salt tolerance was also reported in *Arabidopsis* through the over expression of soybean

ubiquitin-conjugating enzyme gene *GmUBC2* (Zhou et al. 2009). Jae-Heung et al. (2006) reported that up-regulation of an *Arabidopsis* RING-H2 gene, XERICO, confers drought tolerance through increased ABA biosynthesis. *OsSKIPa*-overexpressing rice showed significantly increased ROS scavenging ability and transcript levels of many stress-related genes, including *SNAC1* and rice homologs of *CBF2*, *PP2C*, and *RD22*, under drought stress conditions (Hou et al. 2009).

#### 3.4.2.2.2 Glutathione synthesis and metabolism pathways

Glutathione is an important antioxidant that takes part in the control of cellular hydrogen peroxide level. The enzyme,  $\gamma$ -glutamylcysteine synthetase catalyses the first and rate limiting step and is important in GSH homeostasis (Hicks et al. 2007). Overexpression of  $\gamma$ -glutamylcysteine synthetase confers freezing tolerance in *Chorispora bungeana* (Wu et al. 2008), herbicide tolerance in poplar (Gullner et al. 2001) as well as heavy metal tolerance in *Arabidopsis* (Li et al. 2006). However, there were no reports on drought tolerance. Change in the ratio of reduced (GSH) to oxidized (GSSG) form of glutathione during the degradation of peroxide is important in redox signaling pathways. Reduced glutathione (GSH) is involved directly in the reduction of most active oxygen radicals generated due to stress (Shao et al. 2008). GR mediates the reduction of GSSG to GSH by using NADPH as an electron donor, and thus a highly reduced state of GSH/GSSG ratios is maintained (Carvalho and Contour-Ansel 2008). GR activity has been shown to increase in various plant species under drought, high light intensity, heavy metals, salinity etc. (Rao and Reddy 2008). Up-regulation of GR gene under drought was reported in cowpea (Contour-ansel et al. 2006) and common bean (Torres-Franklin et al. 2008) as well as it confers drought tolerance in transgenic tobacco (Foyer and Noctor 2005). GST detoxifies endobiotic and xenobiotic compounds by covalently linking glutathione to a hydrophobic substrate, forming less reactive and more polar glutathione s-conjugate (Neuefeind et al. 1997). Overexpression of GST gene under drought was reported in wheat (*Triticum aestivum*) (Gallé et al. 2005), barley (Guo et al. 2009) and in transgenic tobacco plant transformed with a GsGST from wild soybean (*Glycine soja*) (Ji et al. 2010). These results suggest the role of glutathione in drought tolerance mechanism of tea.

### 3.4.2.2.3 Sugar synthesis, transport and metabolism

Sugar synthesis, transport, its storage and metabolism involved in the control of growth and development during the entire plant life cycle. In general, low sugar status enhances photosynthesis, reserve mobilization and export whereas the abundant presence of sugar promotes growth and carbohydrate storage (Rolland et al. 2002). Higher accumulation of sugar under abiotic stress including drought is well-documented in literature (Vajrabhaya et al. 2001, Rolland et al. 2002, Hoffmann 2010). Over expression of sugar synthesis, transport or metabolism related genes in plants confer drought tolerance such as Trehalose-6-phosphate synthase (*TSP1*) transgenic potato (Yeo et al. 2000) as well as glyceraldehyde 3-phosphate dehydrogenase and hexose transporter transgenic *Arabidopsis*, respectively (Yang et al. 1993, Yamada et al. 2010). There are various reports of differential expression of sugar synthesis and metabolism related genes under drought such as phosphoenolpyruvate carboxykinase (Rabello et al. 2008), inosine-5'-monophosphate dehydrogenase (Kushwaha et al. 2009), Alpha-1, 4 glucan phosphorylase (Abebe et al. 2010) etc. Presence of high numbers of sugar synthesis and metabolism related transcripts including those of glycolysis and gluconeogenesis pathway associated genes in present analysis; suggest the role of sugar and related pathway genes for drought tolerance in tea like that of other plants.

### 3.4.2.3 Genic microsatellite markers

Development of genic markers has a major impact on genetic analysis such as gene mapping and marker assisted breeding (Kumpatla and Mukhopadhyay 2005). Compared to the development of genomic tools and technology in other crops, the availability of such tools such as genic markers has been limited in tea. Computational identification and validation of unigene derived microsatellite markers (UESTM) have advantage of assaying variation in the expressed component of the genome with unique identity and position. It reduces the cost and facilitated the gene mapping and marker assisted selection (Sharma et al. 2009). Since UESTM markers containing unigenes corresponded to known proteins in database, it provides the scope to know the consequences of repeat sequences. It was reported that the frequency of EST-SSR was in the range of 2.65% to 10.62% in dicot plants

(Kumpatla and Mukhopadhyay 2005) and 1.5% to 4.7% in monocot plants (Kantety et al. 2002). In present investigation, 6.09% of unigenes were found containing SSRs in the range of dicot species, however higher than that of monocots. This result was supported by previous results in tea where they found 8.9% (Sharma et al. 2009). However, the frequency is significantly influenced by repeat length in the search genome of different plant species and the criteria used for SSR mining (Sharma et al. 2009). In present analysis, it was found that the frequency of dinucleotide repeat was more followed by trinucleotide repeat. This trend of frequency of repeat sequences is in agreement with the earlier reports in tea (Zhao et al. 2007, Sharma et al. 2009). On the basis of previous reports, 60% of the designed primer pairs can be expected to apply functionally (Cordeiro et al. 2001). The functional application of EST-SSR for the assessment of inter- and intra-specific diversity was successfully done in several tea accessions (Jin et al. 2007, Zhao et al. 2007, Sharma et al. 2009). These previous studies have enriched the marker resource in tea, however, there was no trait specific marker reported in tea so far. This present study of EST-SSR considerably broadens the scope of tea diversity studies and may present microsatellite sequences specific to drought, however, wet lab validation remains to be done.

### **3.4.3 Full-length library construction, analyses and functional annotation of ESTs**

Full-length cDNA libraries are very important resources for isolation of the full-length cDNAs. Full-length cDNAs are essential for the correct annotation of transcriptional units and gene products from genomic sequence data and for functional analysis of the genes. The SMART libraries contain a higher percentage of full-length clones than other libraries constructed either by conventional methods or other full-length cDNA synthesis protocols (Okayama and Berg 1982, Kato et al. 1994). By analyzing 346 ESTs, an overall picture of tea root's transcriptome was drawn under normal growth. This transcriptome was used as standard for analyzing the drought induced transcriptome. Average length of the full-length EST sequences (376 bp) was lesser than that of drought induced transcripts (428 bp); however this was due to variation of capillary length used during electrophoresis. In normal growth of roots, most of the transcripts were found either with no significant hits in database (46%) or with predicted proteins (22.54%). It was an indication that

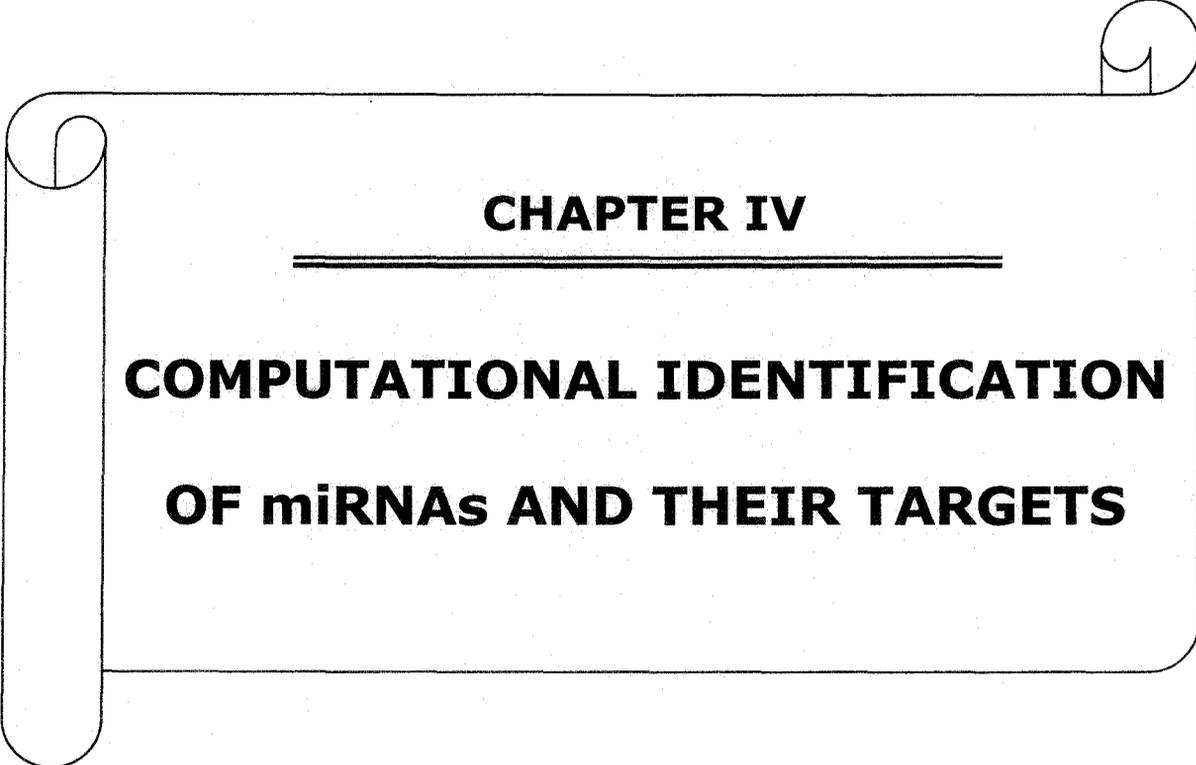
numbers of genes in roots are still not discovered. However, there were less number of transcripts in these categories under drought such as only 5.94% of no significant hits and 6.99% of predicted proteins. It seems that stress associated genes were less in number and most of them were known. There may be different set of genes in plants particularly in roots under normal growth and drought stress. This fact was also supported by the findings that there were only 10 common unigenes in roots under normal growth and drought stress. Probably this was the reason that functional assignment of unigenes generated under normal growth could be done only for 35.75% in contrast to 76% of drought associated unigenes. There were only 1.73% chaperones and heat shock proteins found in normal growth, however, 3.85% of heat shock proteins and 3.67% of molecular chaperones were found under drought stress. It shows that drought stress induces chaperones and heat shock proteins. Higher accumulation of heat shock proteins in roots under drought have been reported in wheat (Sharma and kaur 2009) and mustard (Vartanian et al. 1987). Moreover, synthesis of enzymes and other functional proteins were also higher under drought (43.36%) in comparison to normal growth (21.10%).

#### **3.4.4 Comparative analyses of transcripts under normal growth, drought and winter dormancy stress**

Exposure of a plant to different environmental stresses may result in similar responses at the cellular and molecular level. This is due to the fact that the impacts of the stressors trigger similar strains and downstream signal transduction chains. A good example for an unspecific response is the reaction to stressors which induce water deficiency e.g. drought, salinity and cold (Beck et al. 2007).

Decline in temperature during periods of winter dormancy is one of the most obvious environmental factors experienced by tea plants. It is known that exposure of plants to abiotic factors including low temperature causes oxidative stress in which increased production of ROS is evident. During that period various stress associated genes get up-regulated to adapt or tolerate with the situation (Vyas and Kumar 2005). Comparative transcripts analyses under drought and winter dormancy stresses revealed that there were only a few numbers of common genes in between these stresses. It indicates that there are different gene sets to adapt or tolerate with

each of these stresses. However, winter dormancy stress and normal growth shared a couple of genes which indicates winter dormancy stress is less lethal than to drought stress. Moreover, there was also indication for the presence of tissue specific gene sets to combat with drought stress. However, large number of transcripts analyses may give more conclusive results.



**CHAPTER IV**

---

---

**COMPUTATIONAL IDENTIFICATION  
OF miRNAs AND THEIR TARGETS**

## COMPUTATIONAL IDENTIFICATION OF miRNAs AND THEIR TARGETS

### 4. 1 Review of literature

Genomes of higher eukaryotes encode not only proteins but also diverse non-coding RNAs, particularly small (20 to 30 nt) regulatory RNAs (Zhu et al. 2008). Small RNAs include miRNAs, siRNAs, and piwi-interacting RNAs (piRNAs) (Chapman and Carrington 2007, Matranga and Zamore 2007). The miRNAs are single stranded RNAs that are generated from endogenous hairpin shaped long precursor transcripts (Kim 2005, Zhou et al. 2007). Precursor RNAs (pre-miRNAs) are usually ~60-80 nt in animals and more variable in plants (Murchison and Hannon 2004, Kim 2005). Mature miRNAs always present in one arm of the hairpin precursor without having large internal loops or bulges (Krol et al. 2004). They bind to the complementary sites on target mRNAs and repress post transcriptional gene expression in both animals and plants (Singh and Nagaraju 2008, Zhang et al. 2006a). In higher eukaryotes, miRNAs are one of the largest gene families (>200 members per species) accounting for ~ 1% of the genome which have a key role in diverse regulatory pathways (Kim 2005).

#### 4.1.1 Biogenesis of miRNAs and their functional roles

Since the discovery of first miRNA in *lin-14* gene in *C. elegans* (Lee et al. 1993), it has been extensively studied in both animals and plants for discovering new miRNAs and their biogenesis pathways with functional roles. However, first miRNA in plant was reported only in the year 2002 in *Arabidopsis*, much later than its animal counterpart (Reinhart et al. 2002). It is now known that miRNA genes reside in regions of the genome as distinct transcriptional units as well as in clusters of polycistronic units carrying the information of several miRNAs (Bartel 2004, Chapman and Carrington 2007). Studies suggest that approximately half of known miRNAs reside in non-protein coding RNAs (intron and exon) or within the intron of protein coding genes (Chapman and Carrington 2007). In animals, pre-miRNAs are processed in the nucleus into shorter hairpin RNAs of 65 nt by the microprocessor complex containing RNaseIII enzyme Drosha and its cofactor DGCR8/Pasha, a dsRNA binding protein (Zhu 2008). Precursor miRNA is then exported to the cytoplasm,

where it is further processed by another RNase III enzyme, Dicer, to release a 22 nt miRNA/miRNA duplex (Matranga and Zamore 2007, Han et al. 2006). Dicer function also requires a dsRNA-binding protein, TRBP, as a cofactor. The miRNA is loaded into the effector complex, known as RISC, to direct complementary or partially complementary mRNAs for cleavage or translational repression (Zhu et al. 2008). In plants, the two-step processing of pre-miRNAs into mature miRNAs occurs entirely in the nucleus and is carried out by a single RNase III enzyme, DCL1 (Dicer-like 1) (Jones-Rhoades et al. 2006). In addition to DCL1, genetic analysis revealed that HYL1, a dsRNA binding protein, and SE, a C2H2-type zinc finger, are also required for processing pre-miRNAs and for accumulation of mature miRNAs (Zhu 2008). However, whether DCL1 alone is active in processing pre-miRNAs into miRNAs and how HYL1 and SE may function in the processing steps are not known. Recent studies suggested that DCL1 is inaccurate in catalyzing the release of miRNAs from pre-miRNAs and HYL1 or SE could improve the accuracy resulting in 80% of the *in vitro* processing products being miRNAs (Dong et al. 2008). The importance of DCL1, HYL1, and SE in plant growth and development was evident early on the severe and pleiotropic plant phenotypes exhibited by their loss-of-function mutant alleles (Zhu 2008, Park et al. 2002).

It has been well-demonstrated that miRNAs play critical roles in diverse biological processes such as development, cellular differentiation, cell-cycle control, apoptosis and oncogenesis. As in other organisms, in plants miRNAs play crucial roles in various stages of development and maintenance of organ identity such as leaf morphology and polarity, roots formation, transition from juvenile to adult vegetative phase and vegetative to flowering phase, flowering time, floral organ identity and reproduction (Wang et al. 2007, Mallory and Vaucheret 2006, Sunkar et al. 2007). They are also found to be involved in response to pathogen invasion (Zhang et al. 2006a), hormone signaling (Eckardt 2005, Guo et al. 2005), environmental stress such as cold, salinity, drought stress (Jones-Rhoades and Bartel 2004, Sunkar and Zhu 2004, Lu and Huang 2008) and promotion of anti-viral defence (Lu et al. 2008a).

#### 4.1.2 Strategies for miRNA identification and characterization

Identification of miRNAs and their functional classification has become one of the most active research fields in biology in the recent years. The miRNA identification largely relies on two main reverse genetics strategies: 1) computational and (2) experimental approaches. A third identification approach, forward genetics, is rarely used in miRNA discovery (Unver et al. 2009).

Identification of miRNAs using computational tools is one of the most widely used methods, contributing considerably to the prediction of new miRNAs in both animal and plant systems (Zhang et al. 2006b). This is largely due to low cost, high efficiency, fast and comprehensive methodology of bioinformatics (Griffiths-Jones et al. 2006). The main theory behind this approach is finding homologous sequences of known miRNAs both within a single genome and across genomes of related organisms (Jones-Rhoades and Bartel 2004). Sequence and structure homologies are used for computer-based predictions of miRNAs (Jones-Rhoades and Bartel 2004). Therefore, prior DNA or RNA sequence information of the organism selected for miRNA identification is required in order to run the softwares. On the other hand, the computationally predicted miRNAs should also be confirmed via experimental methods. In recent years, a number of programs and bioinformatics tools have been developed and used successfully for the identification and analysis of miRNAs and their targets (Rajewsky and Socci 2004, Zhang et al. 2009b). Computational identification of miRNAs were reported by homology search in ESTs (Zhang et al. 2005, Guo et al. 2007, Nasaruddin et al. 2007, Jin et al. 2008), GSS (Zhang et al. 2006c, Qiu et al. 2007, Lu and Yang 2010) as well as in whole genome (Zhang et al. 2009a) from diverse plant species. Importantly, large portion of the identified miRNAs were reported from diverse crop plants species (Table 4.1).

Identification and characterization of miRNAs through cloning and sequencing of small RNA libraries represents an experimental approach. Next generation massive sequencing techniques such as pyrosequencing and Illumina are also applied to identify new miRNAs in plants (Moxon et al. 2008, Szittyá et al. 2008). Such approaches for miRNA identification also have some limitations. First, most of the miRNAs are tissue and time specific, and generally their expression level is low. In

Table 4.1: Identification of miRNAs in important crop plants

Number of miRNAs

Crop plants	Total	Exp.	Computational	References
<i>Medicago truncatula</i>	395	356	39	Zhang et al. 2005, Dezulian et al. 2005, Jagadeeswaran et al. 2009, Zhou et al. 2008, Szittyta et al. 2008, Lelandais-Briere et al. 2009
<i>Vigna unguiculata</i>	50	-	50	Wang et al. 2009, Lu and Yang (2010)
<i>Glycine max</i>	166	80	86	Zhang et al. 2006b, Zhang et al. 2005, Subramanian et al. 2008, Wang et al. 2009, Zhang et al. 2008a
<i>Phaseolus sp.</i>	12	8	4	Zhang et al. 2005, Zhang et al. 2006b, Arenas-Huertero et al. 2009, Sunkar and Jagadeeswaran 2008, Jagadeeswaran et al. 2009
<i>Oryza sativa</i>	2641	305	2336	Reinhart et al. 2002, Rhoades et al. 2002, Jones-Rhoades and Bartel 2004, Sunkar et al. 2005, Guddeti et al. 2005, Jones-Rhoades et al. 2006, Archak et al. 2007, Morin et al. 2008, Lu et al. 2008b, Sunkar et al. 2008, Zhang et al. 2006b, Xue et al. 2009, Sunkar et al. 2005, Lindow et al. 2007, Zhu et al. 2008, Lacombe et al. 2008, Wu et al. 2009, Huang et al. 2009, Johnson et al. 2009, Jian et al. 2010

(Table 4.1 continued...)

<i>Brassica sp.</i>	82	70	12	Xie et al. 2007, Wang et al. 2007, Pant et al. 2009, Kutter et al. 2007, He et al. 2008, Hsieh et al. 2009
<i>Zea mays</i>	269	49	220	Zhang et al. 2005, Zhang et al. 2006b, Dezulian et al. 2005, Juarez et al. 2004, Maher et al. 2004, Zhang et al. 2006c Zhang et al. 2009, Johnson et al. 2009
<i>Allium cepa</i>	7	-	7	Zhang et al. 2005, Zhang et al. 2006b
<i>Arachis hypogaea</i>	36	-	36	Zhao et al. 2010
<i>Glycine max</i>	229	201	28	Subramanian et al. 2008, Zhang et al. 2005, Zhang et al. 2008a, Dezulian et al. 2005
<i>Hordeum vulgare</i>	17	-	17	Zhang et al. 2005, Zhang et al. 2006b, Sunkar and Jagadeeswaran 2008
<i>Saccharum officinarum</i>	32	-	32	Sunkar and Jagadeeswaran 2008, Dezulian et al. 2005
<i>Solanum lycopersicum</i>	37	21	16	Zhang et al. 2005, Zhang et al. 2006b, Sunkar and Jagadeeswaran 2008, Zhang et al. 2008b, Moxon et al. 2008
<i>Solanum tuberosum</i>	62	-	62	Zhang et al. 2005, Zhang et al. 2006b, Sunkar and Jagadeeswaran 2008, Zhang et al. 2009b
<i>Sorghum bicolor</i>	150	-	150	Zhang et al. 2005, Zhang et al. 2006b, Sunkar and Jagadeeswaran 2008, Dezulian et al. 2005, Maher et al. 2004, Bedell et al. 2005, Paterson et al. 2009
<i>Triticum aestivum</i>	85	71	14	Zhang et al. 2005, Zhang et al. 2006b, Yao et al. 2007, Wei et al. 2009

(Table 4.1 continued....)

<i>Ricinus communis</i>	87	63	24	Griffiths-Jones et al. 2003, Dezulian et al. 2005, Zeng et al. 2010
<i>Gossypium sp.</i>	89	19	70	Khan Barozai et al. 2008, Qiu et al. 2007, Zhang et al. 2007, Pang et al. 2009, Kwak et al. 2009
<i>Malus x domestica</i>	5	-	5	Zhang et al. 2005, Zhang et al. 2006b
<i>Vitis vinifera</i>	144	-	144	Zhang et al. 2005, Jaillon et al. 2007, Hsieh et al. 2009
<i>Citrus sp.</i>	126	-	126	Zhang et al. 2005, Zhang et al. 2006b, Sunkar and Jagadeeswaran 2008, Wu et al. 2009, Song et al. 2009, Song et al. 2010

---

addition, they mostly express in response to specific environmental stimuli. They also coexist with their cleaved and degraded target mRNAs, hence cloning small RNAs (miRNA and siRNA) is difficult (Unver et al. 2009). Forward genetics approach is time consuming, expensive, and less efficient; hence it is rarely used in plant miRNA identification. There is only one example using a forward genetics experimental approach to identify miRNA in plants. Baker et al. (2005) identified a miRNA loss of function allele by a transposon insertion upstream of the predicted *MIR164c* stem-loop. The miRNA mutant resulted in a flower phenotype with extra petals.

#### 4.1.3 Identification of miRNA targets

Target identification of miRNAs is important to know their functional roles. Specific miRNA targets in plant genomes and transcriptomes have been identified with computational methods and subsequently validated experimentally. Predicting miRNA targets in plants is much easier due to high and significant complementarities to miRNA-mRNA targets (Moxon et al. 2008). High degree of complementarity between plant miRNAs and their target mRNAs has allowed the prediction of targets using computational algorithms and tools (Jones-Rhoades and Bartel 2004). Ability for plant miRNA to target mRNA with perfect sequence complementary matches was first shown with miR171 (Llave et al. 2002). It was shown that miR171 has perfect antisense complementarity with three Scarecrow-like (SCL) transcription factors in the *Arabidopsis* genome. Predicting conserved miRNA targets in different organisms has revealed that homologous mRNAs are targeted by conserved miRNAs within a miRNA family, yet allowing more gaps and more mis-matches between an individual miRNA and its target (Laufs et al. 2004). Successful prediction of miRNA targets have been well demonstrated in several plant species including *Arabidopsis* (Rhoades et al. 2002) and rice (Li et al. 2010). Most of the computationally identified targets in different studies were reported as transcription factors which play important roles in plants growth and developments (Zhang et al. 2006a, Han et al. 2010, Lu and Yang 2010). As with computational approaches, experimental approaches have been utilized widely to predict plant miRNA-mRNA target sites. Genome-wide expression profiling to search for miRNA targets can be applied on expression arrays. In one example, array data showed that five transcripts encoding *TCP* genes were down-

regulated via overexpression of miR319a (miR-JAW) in *Arabidopsis*. Those five TCP transcription factor mRNAs show up to five mismatches, or four mismatches when G:U wobble counts 0.5 mismatch (Palatnik et al. 2003). Additionally Schwab et al. (2005) overexpressed four different miRNAs in each *Arabidopsis* plant and examined each expression profile to experimentally establish parameters for target cleavage guided by plant miRNAs. However, they found no new target mRNAs other than previously identified by computational approaches. At present, the most powerful method to confirm miRNA-mRNA targets is the 5' RACE procedure. The 5' RACE technique has been used successfully to identify miRNA targets in many plants (Sunkar et al. 2005, Palatnik et al. 2003, Mallory et al. 2005). Cleaved mRNA products in plants have two diagnostic properties. One is that the 5' phosphate of a cleaved mRNA product can be ligated to an RNA adaptor with T4 RNA ligase. Second, in general, the precise target cleavage position is that mRNA target nt pair with the tenth nucleotide of miRNA (Sunkar et al. 2005, Kasschau et al. 2003). Cleaved mRNA products by miRNA guided activity can be amplified with ligation of an oligonucleotide adaptor to the 5' end, followed by reverse transcription and PCR amplification with a gene specific primer (Sunkar et al. 2005).

MiRNA studies in plants have already explained a number of biological events in response to abiotic stresses. Improved understanding of molecular mechanisms of miRNA in plants will lead to the development of novel and more precise techniques that will help in better understanding of some post-transcriptional gene silencing in response to stresses. Accumulating knowledge on the roles of plant miRNA's in molecular biology is leading to the development of more efficient and reliable tools for their characterization.

## 4.2 Material and methods

Discoveries of miRNAs in model and crop plants as well as their growing evidence of involvement in a variety of functional roles have produced a great deal of excitement in agricultural biotechnology. Despite the limited genome resources of tea, published ESTs and full length nucleotide sequences have provided the scope to get more genetic information. In this study, new conserved miRNAs were mined in local tea sequence database for the purpose of understanding their roles in regulating growth and development, metabolism and other physiological processes.

### 4.2.1 Collection of reference miRNAs, full-length nucleotides and EST sequences

A total of 9275 plant miRNAs (on 20 September, 2010) and their fold back sequences were obtained from miRBase (Griffiths-Jones et al. 2008). Homolog miRNAs were eliminated and the rest were defined as reference for searching homolog tea miRNAs. Tea nucleotide and EST sequences (14819 as on 8 August, 2010) including those generated in present investigation were downloaded from NCBI's nucleotide and dbEST database (<http://www.ncbi.nlm.nih.gov/>). All redundant and poor quality sequences were eliminated and created a local nucleotide database.

### 4.2.2 Prediction of miRNAs and their precursors

The procedure for searching conserved miRNA homologues in tea is summarized in Fig. 4.1. Since, only mature miRNAs were conserved in plants, the reference sequences were used as a query in homology search against the reference nucleotide database at e-value threshold less than 0.01 through BLAST+ 2.2.22 program (Altschul et al. 1997). Target sequences with no more than four mismatches were considered and predicted the secondary structure using Mfold v 3.2 (Zuker 2003). Precursor sequences were searched in 50 nt upstream or downstream from the location of mature miRNAs with an increment of 10 nt. While selecting a RNA sequence as a candidate miRNA precursor, following criteria were used according to Zhang et al. (2006b) with minor modifications as: i) a RNA sequence can fold into an appropriate stem-loop hairpin secondary structure, ii) a mature miRNA sequence site in one arm of the hairpin structure, iii) miRNAs had less than seven mismatches

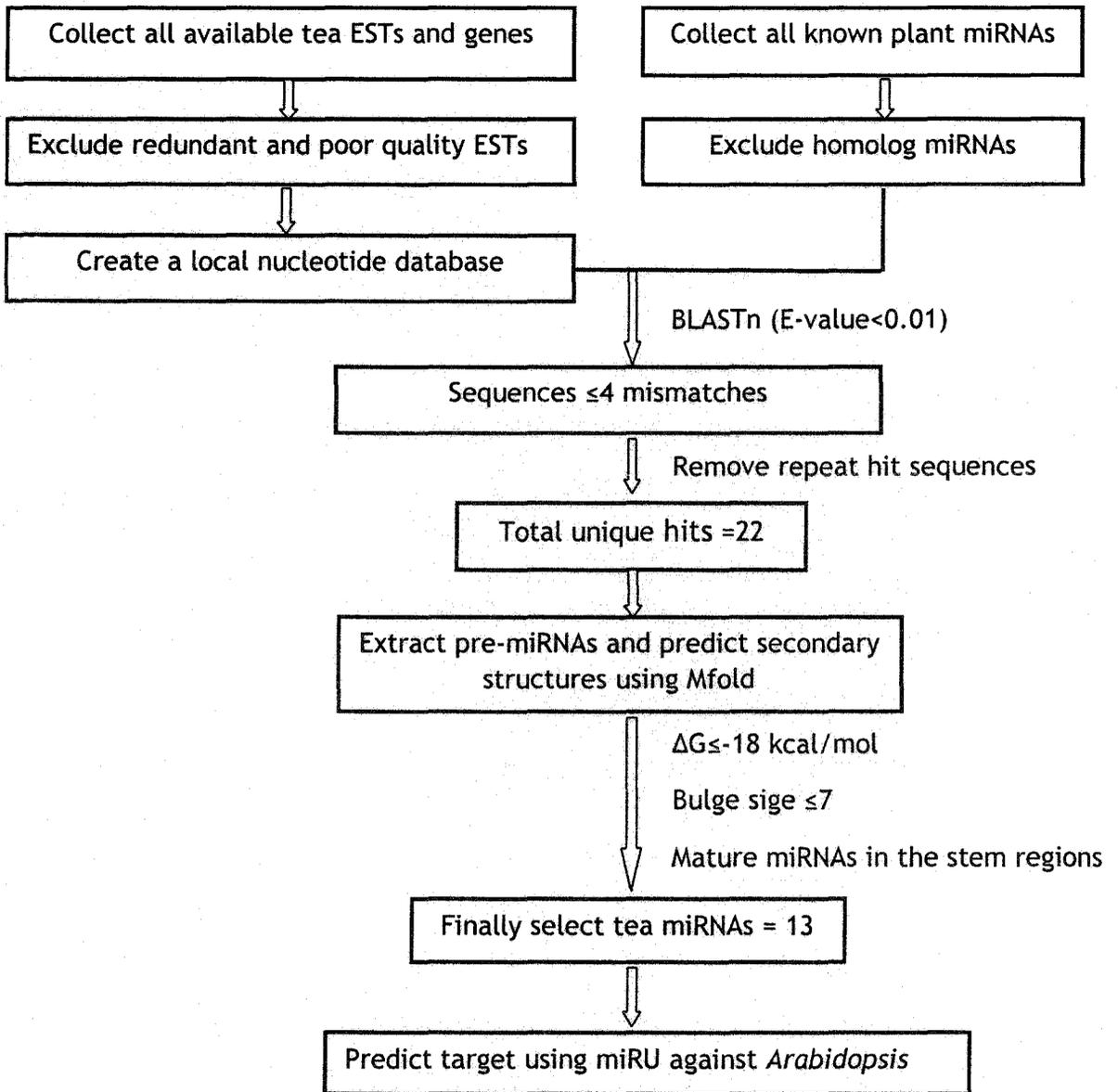


Fig. 4.1: Different steps involved in the identification process of conserved miRNAs and their targets.

with opposite miRNA sequence in the other arm, iv) no loop or break in miRNA sequences, v) predicted secondary structures had higher negative energy MFEs ( $\leq -18$  kcal/mol), and iv) 40-70% A + U contents.

#### 4.2.3 Identification of potential miRNA targets

As for tea, since only few gene sequences are available, *Arabidopsis* was used as a reference system for finding targets of the candidate miRNAs. The predicted tea miRNAs were used as query against the *Arabidopsis thaliana* DFCI gene index (AGI) release 13 using miRU (<http://bioinfo3.noble.org/psRNATarget/>) following the criteria as i) maximum expectation value 3 ii) multiplicity of target sites 2 iii) range of central mismatch for translational inhibition 9 to 11 nt iv) maximum mismatches at the complementary site  $\leq 4$  nt without any gaps.

#### 4.2.4 Evaluation of conserved nature in miRNAs and their phylogenetic relationships

Due to the conserved nature of miRNAs, orthologue discovery can be done through bioinformatics analysis. Here, tea miRNA conservation was analysed with their orthologues. A homology search of candidate tea miRNAs was done against all plant miRNAs using BLAST+ 2.2.22 (Altschul et al. 1997) allowing maximum of 3 mismatches and e-value  $< 0.001$ . The corresponding precursor sequences of homolog small RNA's were identified and collected. The collected sequences of diverse plant species were aligned with the tea miRNA homolog using Clustal W (Thompson et al. 1994).

A query of tea small RNAs against known miRNA families (miRBase, release 15) allowed to identify 3 previously reported large families. The precursor sequences of three known family members were selected along with respective precursor sequences of tea minas. Then, the maximum likelihood trees were constructed for each family based on Tamura-Nei model (Tamura and Nei 1993) with default bootstrap values using MEGA 4.0 (Tamura et al. 2007) to illustrate the evolutionary relationships among the members of the family.

#### 4.2.5 Nomenclature of miRNAs

The predicted tea miRNAs were named in accordance with miRBase (Griffiths-Jones et al. 2006). Mature sequences were designated 'miR', and the precursor hairpins were labeled as 'mir' with prefix '*csi*' for *C. sinensis*, '*cja*' for *C. japonica* and '*cas*' for *Camellia assamica*. In the cases where distinct precursor sequences have identical miRNAs with different mismatch pattern, they were named as *csi-mir-1-a* and *csi-mir-1-b*.

## 4.3 Results

Sequence and structure homologies are the main theory behind the computer-based approach for miRNA prediction. At present, four kinds of databases namely genome, GSS, EST and nucleotide are mainly used for miRNA mining. Considering the unavailability of genome and genomic survey sequences of *Camellia*, both dbEST and nucleotide databases was mined for miRNA identification.

### 4.3.1 Identification of miRNAs and their precursors

A total of 14819 tea sequences containing 2023 full length nt and 12796 ESTs were obtained from available individual database in GenBank. Out of these, 22 sequences had less than five mismatches with previously known plant miRNAs. After carefully evaluating the hairpin structures using the defined criteria, 13 miRNAs were finally identified. Details of tea miRNAs such as source sequences, location in the source sequences, length of precursor sequences and their minimum folding free energies and A + U content were given in Table 4.2. A total of 9 miRNAs were predicted from ESTs whereas 4 were from full length nucleotide sequences. Five of them were located in the direct strand and rests were in the indirect strand. Newly identified pre-miRNAs have minimum folding free energies (mfe) ranging from -186.83 to -18.5 kcal/mol, with an average of about -72.69 kcal/mol and A+U content ranging from 45.24 to 69.83% with an average of 53.79% (Table 4.2, Appendix-4.1). Length of tea miRNA precursors ranges from 65 to 663 nt with an average of 248 nt and mature sequences ranges from 20 to 25 nt. The newly predicted two tea miRNA (cja-miR2910, csi-miRf10132-akr) sequences were perfectly (100%) matched with the corresponding homologues of *populus* and rice, whereas the remaining 11 mature tea miRNA sequences differ by 1 to 4 nt from their homologues. All mature sequences of tea miRNAs were found in the stem portion of the hairpin structures (Fig. 4.2) containing less than 7 mismatches in the other arm without break or loop inside the mature sequences. It was found that tea miRNA (csi-miR408) has been conserved with diverse plant species (Fig. 4.3, Appendix-D 4.2) from monocotyledonous plants such as rice, maize to dicotyledonous plants such as *populus*.

Table 4.2: Predicted miRNAs of *Camellia* sp.

New miRNAs	NS*	Gene ID	Strand	SP*	EP*	NM (nt)	Mature miRNAs	e-value	PL*	A+U (%)	MFE
csi-miR 408	EST	206583693	3'	137	117	3	CUGCACUGCCUCUUCCCUGAG	0.001	336	45.24	-20.10
csi-miR1171	EST	171355265	5'	286	308	1	UGGAGUGGAGUGAAGUGGAGUGG	3E-04	181	56.98	-5.97
csi-miR414a	EST	206583641	3'	757	637	3	UCUUCCUCAUCAUCAUCUUCU	0.001	663	57.32	-3.18
csi-miR414d	EST	284026209	3'	186	166	1	UCAUCGUCAUCGUCAUCAUCU	0.004	193	61.14	-7.72
csi-miR414f	EST	212378632	5'	122	142	1	UCAUCAUCAUCAUCAUCUUCA	6E-05	68	57.35	-18.50
cas-miR1122	FL	214011104	5'	214	237	4	UACUCCUCCGUCCAAAAUAAUG	6E-05	294	69.83	-1.23
csi-miR414g	EST	51453040	3'	474	454	3	CCUUCCUCAUCAUCAUCGUCC	0.001	70	45.71	-25.20
csi-miRf10132-akr	EST	51453383	3'	58	34	0	GCGAGCUUCUCGAAGAUGUCGUUGA	9E-08	200	49.00	-69.50
cja-miR2910	FL	1777723	5'	1262	1282	0	UAGUUGGUGGAGCGAUUUGUC	1E-05	301	49.83	-91.00
csi-miR2914	FL	34787361	5'	345	367	1	UAUGGUGGUGACGGGUGACGGAG	5E-06	65	49.23	-20.90
cas-miRf10185-akr	EST	221071827	3'	232	212	3	GAAAGGGGAAAACAUUGUAGC	0.004	139	48.92	-51.10
cas-miR11590-akr	EST	212379609	3'	113	94	3	UUUUGGUGUGCCUUCAACCU	0.003	75	53.33	-23.80
csi-miR414h	EST	295345415	3'	79	58	4	UCAUCCUCAUCAUCGUCAGAA	0.004	644	55.36	-86.83

\*NS = nucleotide source, FL = full-length, SP = start point, EP = end point, ME = No match, PL = pre-miRNA length

Fig. 4.2: Predicted hairpin secondary structures of candidate pre-miRNAs. The miRNAs were highlighted (red color) in stem portion

**cas-miR1122**

```

A-           A  A           |   GG  U           UUGGU           U
GCU  UACUCCCUCCGUCCCA AAUA UGGUCCCU--UUUGG  AA  CCAACUUUUU  AGUACAA \
UGA  AUGAGGGAGGCAGGGU UUAU AUCAGGGA  AAACC  UU  GGUUGAAAAA  UCAUGUU U
AC           A  A           \  ^   UU  -           UUU--           A

```

**csi-miR414f**

```

|   UGC           GA   UU  UUC-  UU
CUGAG  UGGUGGUG  GGUG  UC   UUC  \
GACUU  ACUACUAC  CUAC  AG   AAG  C
^   CU-           UA   U-  UUCU  UU

```

**csi-miR1171**

```

A  A   AGA GU   G   AGG-|  .-AG           C
CGC CCAUU  U  UCCA UCCA  GCA  UGGAGGAA \
GUG GGUGA  G  AGGU AGGU  UGU  ACCUCCUU C
G  A   A-- UG   G   GAGG^ \  --           U

```

**cja-miR2910**

```

A   U  G   --  C-   A   .-GUUAA  .-AACG   .-CAG  G   CU   --|   U
GUUCUU GUUGG GGA CGAU  UUGU  UGGUU AUUCC  CG   AGACCU  CCU CUA  AGCUA  UGCGGAGG G
UAAGGA CAACU UCU GUUG  AACG  ACUAG UAGGG  GC   UCUGGA  GGA GAU  UCGAU  ACGCCUCC A
G   -  G   UU  UU   A   \  ----- \  ----  \  ---  -  CU   CG^   C

```

**csi-miR2914**

```

C--| UA  A  U  G           AUA
UUUCUG  CC  UCA CU UC AUGGUAGG  \
AGAGGC  GG  AGU GG GG UAUCAUCC  G
AGU^ GC  -  U  -           GGU

```

(Fig. 4.2 continued...)

**csi-miR 408**

```
GAGAGAAGAUGUC      A  -  -  C  U  .-GG  G  .-G|  GG  GU  U
                   UCAGGG AGAG GC AGUG AG GGU   CGGC GC  GCGGU  UGCU  GCC G
                   GGUCCC UCUC CG UUAC UC UCA   GUCG UG  CGUCA  GUGG  UGG A
CC-----          -  U  U  U  U  \  --  G  \  -^  --  --  U
```

**csi-miR414a**

```
U  -  GCAA          UGUCUGCCU      C  C          .-AUGGAGUUUU|  G  GA  A
CUC UCU  UUUUCGUUAUC      GUCUUC UC UCAGUAUCA      GGG UGUU  AGU A
GAG AGA  AGGAGUAGUAG      UAGAAG AG AGUCAUAGU      UCC ACGG  UCG A
G  U  AGA-          -----          A  U          \  -----^  G  G-  A
```

**csi-miR414d**

```
GA  GAUG  U  C  U  G          U  CCC  CU          .-AAUAUAUUA|  AACUUAAC
GGAGAU  UGAC  ACGA GA GA GGC ACUAUGAUUA GC  GG  GCUUGAAUUUG          UGAAC  A
UCUCUG  GCUG  UGUU CU CU UCG UGAUAUUAU CG  UC  CGGAUUUAAAU          AUUUG  A
G-  GG--  -  -  C  A          -  UUU  UC          \  -----^  AUCCCAA
```

**csi-miR414g**

```
A  --          G  AGGAAGGAGAU|  A
GAG AGGAGGA  CGAUGAU AUG          AUC G
CUC UCCUCCU  GCUACUG UGC          UGG G
C          CA  G  CG-----^  U
```

**csi-miRf10132-akr**

```
A  -          A  CGCUC|  G  C  .-A  ----          UAU  AGA  AC  UUACU  C  A
UUCAUCA CG ACAUCUUCGAG AGCU  AG AAGC UCG  GGC  UCGCUAGG  AACA  AGCCU  CA  UCU GGG \
AGGUGGU GC UGUGGAAGUUC UCGA  UC UUUG AGC  UCG  AGUGGUCC  UUGU  UCGGA  GU  AGA CCU G
A  U          -  -----^  A  A  \  -  GUUA          UUC  ---  CU  CGUC-  -  A
```

(Fig. 4.2 continued...)

**cas-miRf10185-akr**

```
      UACAA          GAAG          --|  CCAA          GGGGAA
CCUGC      UGUUUUCCCCUUCUCC      GUUCUCUUUG  CCGC      AUUUUCGG      \
GGGCG      ACGAAGGGGGGAGGGG      CAGGGGAAAC  GGCG      UAAAAGUU      A
      -----          AAAA          UA^  -----          AUAUAA
```

**cas-miR11590-akr**

```
      -----  G  -      AAAA-|  AAG      UU
GUAAGGUUG      AA  GC  ACACC      CA      UUCA  \
CGUCCAAC      UU  CG  UGUGG      GU      AGGU  U
      GGUU  G  U      GGUCG^  AGA      UC
```

**csi-miR414h**

```
      U      A  G      AUG          U  UG      .-G  A      UG  CAU      AA      C  C-      C      .-CUCAA  A      -----|      .-U      G      AUA
GGGA  UCUG  C  AUG      AGGAUGAG  U  GUUUUA      GC  UGG  GU      GGU  CAA  AG  GCU  GGAG          GAG  CAGU      GGAAA      GAUUG  CUA      G
CCCU  AGAC  G  UAC      UCUUACUU  A  CGAGAU      CG  GUC  CG      UCA  GUU  UC  CGA  CCUC          UUC  GUCA      CCUUU      CUAAC  GAU      G
      C      -  -      CG-          U  GU      \  -  A      GU      ---  AC      U  AC      -      \  -----      -      ACUU^      \  -      G      GGU
```

```

vvi-mir408      AAGCUGUUUUUGCUCUACCCAUSSCACUGCCUCUUCGCCUGGCUCUGUCUCUC-----
cpt-mir408      UCUCUGUUUUUGGCUCCTCCCAUUSCACUGCCUCUUCGCCUGGCUCUCUGCCUU-----
ptc-mir408      UACUUGUUUUUGGCUCUACCCADGCAUUSCACUGCCUCUUCGCCUGGCUCUCUGCCUU-----
rcs-mir408      CAUCUGUUUUUGGCUCUACCCCAUUSCACUGCCUCUUCGCCUGGCUCUCUGCCUU-----
ppe-mir408      GAGCUGUUUGGGCUCUACUCACUSCACUGCCUCUUCGCCUGGCUCUCUGCCUU-----
pta-mir408      CGUCUGCUUCUGCCAUUCUUAUSCACUGCCUCUUCGCCUGGCUC-----
smo-mir408      CGGCUCCUUGUGUGGCUUUGAUSCACUGCCUCUUCGCCUGGCUCGGCCAAAGUUA-----
pof-mir408b     UGAGG----GUGUUUGUCCUCASUSCACUGCCUCUUCGCCUGGCUCUC--CCUACAUAGCUCGC
sof-mir408b     GGU-GUUGUUUCUCCUCCCCUUSCACUGCCUCUUCGCCUGGCUCUCCCAACCGUUGCCCUUGC
sof-mir408c     GGU-GUUGUUUCUCCUCCCCUUSCACUGCCUCUUCGCCUGGCUCUCCCAACCGUUGCCCUUGC
sof-mir408a     G----UUGUUUCUCCUCCCCUUSCACUGCCUCUUCGCCUGGCUCUCCCAACCGUUGCCCUUGC
sbi-mir408      GGU-GUUGUUUCUCCUCCCCUUSCACUGCCUCUUCGCCUGGCUC-----
sof-mir408d     GGU-GUUGUUUCUCCUCCCCUUSCACUGCCUCUUCGCCUGGCUCUCCCAACCGUUGCCCUUGC
zma-mir408      --U-GUUGUUUCUCCUCCCCUUSCACUGCCUCUUCGCCUGGCUCUGAUCUCCCAACCGUUGC
tae-mir408      --U-GUUGUUUCUCCUCCCCUUSCACUGCCUCUUCGCCUGGCUCUCCUCC--CAAUCUCUC
osa-mir408      --U-GUUGUUUCUCCUCCCCUUSCACUGCCUCUUCGCCUGGCUCUCCUCCUGCACACUCUCUC
csi-mir408      -----CUGCACUSCCUCUUCGCCUGAG-----
bdi-mir408      ACTAGCTAGCAACAGAAUCCAUSCACUGCCUCUUCGCCUGGCUGGAUCGAUCAAAAGAGG
*****

```

Fig. 4.3: Conservation of tea miRNA (csi-miR408) in diverse plant species. Alignment of pre-miRNAs of different plant species (Appendix-D 4.2) with homolog tea miRNA (csi-miR408) demonstrate the conserve nature of mature miRNAs (highlighted).

### 4.3.2 Phylogenetic analysis

Newly identified tea miRNAs belong to 9 miRNA families including three known independent large miRNA families (mir 408, mir414 and mir1122). There are one tea miRNA namely csi-miR 408 and cas-miR1122 in each family of mir-408 and mir-1122, respectively. However, five members of family mir408 were found in tea (csi-miR414a, csi-miR414d, csi-miR414f, csi-miR414g, and csi-miR414h). Comparison of tea miRNA precursor sequences with other members in the same family showed that most members could be found to have a high degree of sequence similarity with others. Phylogenetic trees among the members of each family illustrated the evolutionary relationships of tea miRNAs (Fig. 4.4, Appendix-D 4.3-4.5).

### 4.3.3 Identification of potential miRNA targets

A total of 37 potential targets were identified for the 7 predicted miRNA families which include 11 miRNAs based on their perfect or nearly perfect complementarity with their target sequences in *Arabidopsis* (Table 4.3, Fig. 4.5). For all the miRNAs, single binding site was found in the targets without any gaps in the complementary region and expectation value ranges from 0 to 3. These potential miRNA targets were belonged to a number of gene families that involved in different biological functions such as regulation of cell cycle, metal ion transportation, starch metabolic processes etc. There were 8% of genes encoding transcription factors, 30% of genes encoding different enzymes and 14% of genes encoding transporters as well as 48% of genes encoding various proteins of physiological and metabolic processes (Table 4.3). The miRNA family 'miR414' showed the highest 30 numbers of independent target genes followed by 'miR408' family with 2 numbers of target genes. Rest miRNA families were with single target genes in *Arabidopsis* (Table 4.3). The 'miR1171' and 'miR1122' miRNA family members did not bind to any target sequences within the defined filtration criteria.

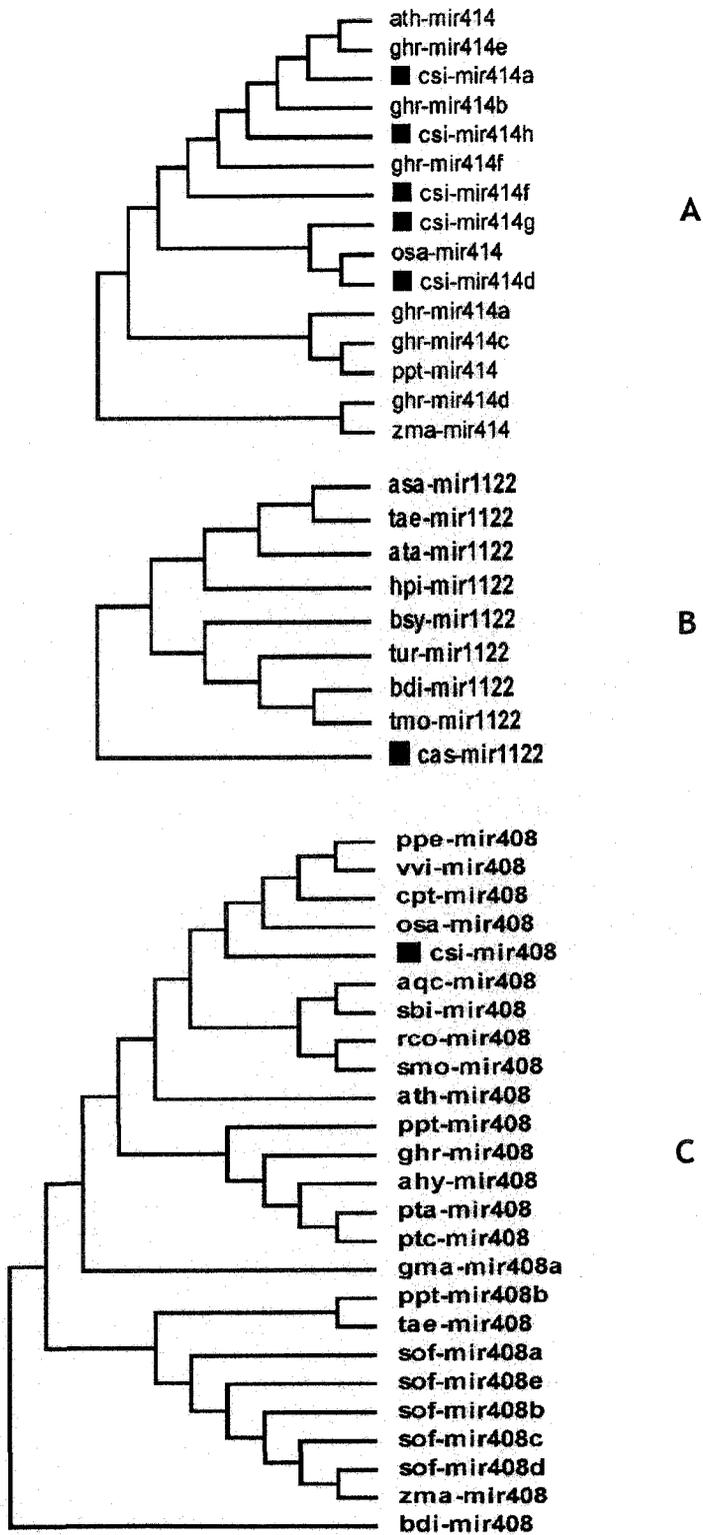


Fig. 4.4. Phylogenetic relationship among the members of each miRNA family, namely: A) miRNA414, B) miRNA1122 and C) miRNA408.

Table 4.3: Potential targets of the candidate miRNA families

Families	TS*	Targeted proteins	Targets involved in	E-value	Gene IDs
miR408	1	Cyclin-dependent protein kinase	regulation of cell cycle	3.0	TC284764
	1	Copper ion binding protein	metal ion transport	3.0	TC300377
miR414	1	RAN GTPase activating protein 2	cytokinesis	1.0	TC283936
	1	50S ribosomal protein L21	translation process	1.0	TC290468
	1	26S proteasome AAA-ATPase subunit	proteasomal protein catabolic process	1.0	TC304323
	1	SEC12p-like transporter	ER to golgi vesicle mediated transport	1.5	NP225634
	1	Nucleotide binding protein	nucleotide binding	2.0	TC280880
	1	MYB transcription factor	regulation of circadian rhythm	2.0	TC283178
	1	Aldose 1-epimerase	carbohydrate metabolic process	0.5	TC285263
	1	Phosphatase 2C-like protein	protein amino acid dephosphorylation	0.5	TC285483
	1	Phosphatidylinositol phosphatase	inositol or phosphatidylinositol activity	0.5	TC312518
	1	Starch branching enzyme class II	starch metabolic process	1.0	AA586097
	1	Reproductive meristem protein 1	regulation of transcription	1.0	TC284340
	1	Calcium ion binding protein	calcium ion binding	1.0	TC299015
	1	Emb protein	RNA processing	1.0	TC294192
	1	Zinc ion binding protein	regulation of transcription	1.0	TC299076
	1	Calmodulin-4	calcium ion binding	1.5	TC294389
1	Translation initiation factor 3 subunit 8	translation initiation	0	TC280751	
1	SMC3 protein	chromosome segregation process	0	TC281006	

(Table 4.3 continued...)

	1	MLO-like protein 3	cell death	0	TC293173
	1	Metalloendopeptidase	proteolysis	0	TC299050
	1	Ubiquitin conjugating enzyme	catabolic process	0.5	CA781750
	1	Zinc finger protein	regulation of transcription	0.5	NP030706
	1	Plastid protein	protein targetting to chloroplast	0.5	TC290688
	1	Ubiquitin thiolesterase	catabolic process	1.5, 1.5	TC280710, TC298096
				1.5	TC305774
	1	Methionyl-tRNA synthetase	methionyl-tRNA aminoacylation	1.5	TC282196
	1	Metal ion binding protein	metal ion binding	1.5	TC305801
	1	Sfc4 protein	xylem or phloem pattern formation	2.0	TC280894
	1	Transcription factor	regulation of transcription	2.0, 2.0	TC293875, TC294793
	1	Synaptosomal-associated protein 25	vesicle mediated transport	2.5	TC293303
	1	ATP binding	protein amino acid phosphorylation	2.5	TC299397
	1	ADP-ribosylation factor-like protein	intracellular protein transport	3.0	TC289959
miRf10132	1	Histone H2B like protein	nucleosome assembly	1.5, 1.5	TC297551, TC313314
				2.5, 1.5	TC313977, TC294144
miR2910	1	Extracellular matrix structural constituent	matrix organisation	0	TC310823,
miR2914	1	Glutamate semialdehyde dehydrogenase	glutamate metabolism	2.0	TC287905
miRf10185	1	Carboxylic ester hydrolase	hydrolase activity	3.0, 3.0	TC298946, TC308821
miR11590	1	FRIGIDA protein	regulation of flower development	2.0	TC309547

---

\*TS=targeted site



#### 4.4 Discussion

With the availability of sequence resources in public databases, computer based miRNA identification methods has been focused more and more in the recent years due to its advantages of low cost and high efficiency. The number and sorts of miRNAs predicted in tea supported the fact that software-based approach is feasible and effective (Khan-Barozai et al. 2008, Jones-Rhoades and Bartel 2004, Singh and Nagraju 2008).

The identified tea miRNAs belong to 9 families where 'miR414' family has 5 members and the rest have single member in each. This familial distribution of miRNAs was also observed in *Arabidopsis*, rice and maize (Bonnet et al. 2004). This may be an indicative of dominant nature of 'miR414' family in miRNA-mediated gene regulation in tea. The identified miRNAs were found diverse in nature such as location of mature miRNA sequences and length of precursor sequences. Average length of precursor sequence was 248 nt; however a majority of them (62%) have 65-200 nt. This finding is similar to other plants where the length of precursors varied in contrast to consistent miRNA length of animal miRNAs (70-80 nt) (Bartel 2004, Ambros 2004). In tea miRNAs, diversity was also observed within the members of same family which was also reported in maize (Zhang et al. 2006c). The identified tea pre-miRNAs fold into hairpin secondary structures using minimum free energies, with an average -72.69 kcal/mol which was lower than the values of *Arabidopsis thaliana* pre-miRNAs and much lower than the folding free energies of tRNA (-27.5 kcal /mol) and rRNA (-33 kcal/ mol) (Khan Barozai et al. 2008).

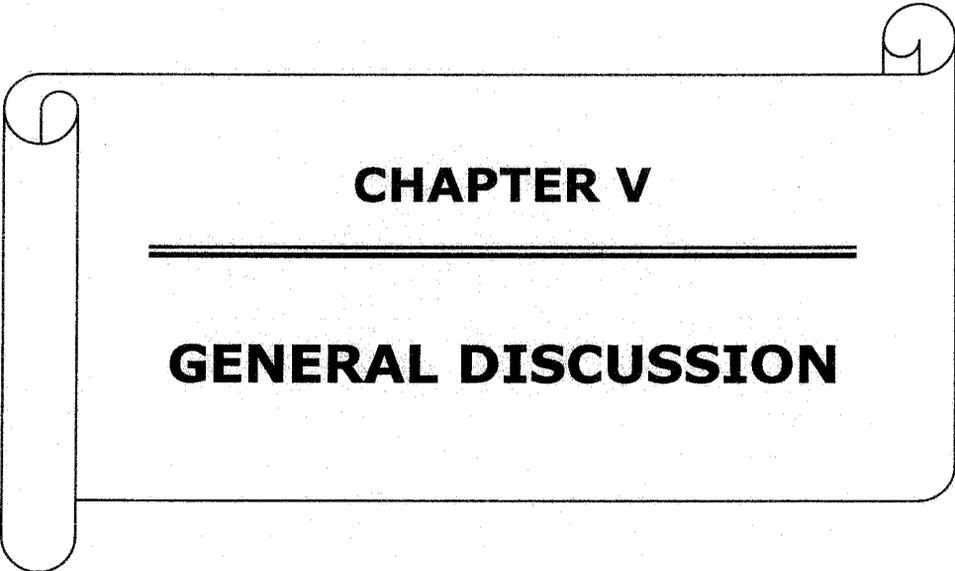
Out of 13 newly identified miRNAs, 10 were from ESTs. There are several reports on miRNA identification from ESTs in various plant species (Zhang et al. 2006b, Qiu et al. 2007). The source sequences of miRNAs show a probable link between expression of miRNAs and their tissues, organs or developmental stages to which it belongs. On that basis, it was recognized that csi-miR414f, cas-miRf10185 and cas-miR11590-akr were expressed in roots and the rest tea miRNAs were expressed in leaf tissues. Moreover, csi-miR1171 and csi-miR414d were found in leaf tissue under the stress of winter dormancy and pest infestation, respectively. Three miRNAs namely cas-miR1122, cja-miR2910 and csi-miR2914 were identified from the full length nt of

RNA polymerase second largest subunit (intron 23) and 18S ribosomal subunit, respectively. Plant miRNAs are highly conserved among distantly related plant species, both in terms of primary and mature miRNAs (Zhang et al. 2006b). This finding is also supported by present results where tea miRNAs were found conserved in diverse plant species from monocotyledonous to dicotyledonous plants. These results suggested that different miRNAs might evolve at different rates not only within the same plant species, but also in different ones.

The miRNA target gene identification is an important step for understanding the role of miRNAs in gene regulatory networks. In the present analysis of target genes for tea miRNAs revealed that more than one gene was regulated by individual miRNA. This result was similar to the recent findings in other plant species (Zhang et al. 2006b, Jones-Rhoades and Bartel 2004) which suggested that miRNA research should be focused on networks rather than individual connections between miRNA and strongly predicted targets. miRNAs may directly target transcription factors which affect plant growth and development, and also specific genes which control metabolism (Zhang et al. 2006a). In this study, a total of 37 potential targets were for the 7 identified miRNA families. The identified 37 potential target genes appeared to be related in diverse biological functions. Transcription factors such as MYB, translation initiation factor TIF3, important proteasome degrading pathway enzyme such as ubiquitin conjugating enzyme, different ion transporters such as copper ion binding protein, carbohydrate metabolism related enzyme such as aldose 1-epimerase, glutamate metabolism related enzymes such as glutamate semialdehyde dehydrogenase, important protein for nucleosome assembly such as histone as well as ribosomal proteins. In an earlier report, 20% target genes of transcription factors and 53% target genes of proteins related to diverse physiological processes were found in contrast to the present findings of 8% target genes of transcription factors and 48% target genes of physiological and metabolic processes, however in the previous investigation, total number of miRNAs was limited to only four (Prabu and Mandal 2010). It becomes clear that tea miRNAs targeted both transcription factors as well as specific genes.

Wet lab based validation of the miRNAs, their expression and functions have been remain as a task to be done in future, through which we can dissect and decipher

the miRNA world in tea. Nevertheless, these findings will considerably broaden the scope of miRNA research in this valuable plantation crop. Further, it shows a way for the prediction and analyses of miRNAs to those species whose genome is not available through bioinformatics tools.



**CHAPTER V**

---

**GENERAL DISCUSSION**

## GENERAL DISCUSSION

Tea is an economically important plantation crop with a life span of 60 years. It grows around three million ha of land world-wide (Alkan et al. 2009). For the last several decades, conventional breeding and propagation techniques were successful for the development of improved tea varieties. However, due to the limitation of these techniques for the development of cultivars with high yield and quality, application of biotechnology becomes an alternative approach (Mondal 2007). Different biotechnological applications have already been adopted in tea for the selection and/or development of quality cultivars such as micropropagation, cell and organ culture, transgenic production, DNA fingerprinting as well as functional genomics (Mondal et al. 2002).

Tea plant is a woody perennial, and hence, encounters a large number of environmental stresses throughout its entire life span. Since most of the world's tea growing areas are prone to drought, tea plant is often subjected to water deficit stress. Drought is the most important recurrent limiting factor of tea cultivation in India and other tea growing countries which incurred around 40% crop loss (Jain 1999). Drought is defined as the absence of adequate water necessary for normal plants growth and development. Water is basically important component of the metabolism of all living organisms, facilitating many vital biological reactions by being a solvent, a transport medium and evaporative coolant. In plants, water provides the necessary energy to drive photosynthesis through autolysis by yielding electrons. Drought stress affects severely the growth, crop yield and various morphological, anatomical, physiological and biochemical processes of tea plantations (Lu 1992, Upadhyaya et al. 2008). The changes basically take part as a mechanism of drought adaptation and/or tolerance. The drought tolerance capability differs from plant to plant depending on the age and genotypes of tea cultivars (Burgess and Carr 1996). Moreover, different environmental stresses to tea plant may result in similar responses at the cellular and molecular level. This is due to the fact that the impacts of the stressors trigger similar strains and downstream signal transduction chains (Urano et al. 2010).

Plants have acquired various stress tolerance mechanisms involving physiological and biochemical changes that result in adaptive or morphological changes (Urano et al. 2010). Drought tolerance is a complex trait, expression of which depends on the action and interaction of different morphological, physiological and biochemical characters (Beck et al. 2007). The level of drought stress correlates with the degree of changes in these characters which provides the scope for identifying drought tolerant cultivars and the study of tolerance mechanism at molecular level thereof. During drought, plants maintain the internal water potential, turgor and water uptake by increasing the level of osmolytes, either by uptake of soil solutes or by synthesis of metabolic solutes (Zhu 2002) as well as maintain the toxic oxygen species level through anti-oxidative mechanisms. For the biosynthesis of these solutes and compounds, numbers of genes and genetic pathways get activated. Hence, a detail knowledge of physiological and biochemical changes under drought provides the scope of correlation with transcriptional analysis, and considerably broaden the prospect of understanding the lying mechanism of tolerance. Understanding the mechanisms of signal perception, transduction and downstream regulation of cellular pathways that are involved in drought stress responses provide valuable information of stress tolerance mechanism for the development of drought tolerant cultivars (Cruz de Carvalho 2008). Understanding the responses of tea plants to their environment in terms of adaptability and performance is of paramount importance for selecting and/or developing cultivars that can withstand the unwanted environmental changes. However, except few physiological and biochemical investigations, no comprehensive analysis of drought stress responses of tea plants have been reported. Hence, present investigation of anatomical, physiological, biochemical and transcriptional changes under drought stress was undertaken in tolerant (TV-23) and susceptible (S.3/A3) cultivars (Konwar 2004) for understanding the drought stress responses of tea plants.

Two year-old vegetatively propagated well-rooted tea seedlings (~ 36 inch height) of S.3/A3 and TV-23 cultivars were planted in earthen pots (12 inch dia) under controlled greenhouse condition (at  $300 \mu\text{mol m}^{-2} \text{s}^{-1}$  light intensity,  $25 \pm 2^\circ\text{C}$  with 65-70% relative humidity) in a completely randomized block design for drought stress imposition. Drought stress was induced by withholding water and subsequently severe stress level was determined on the basis of soil moisture content and

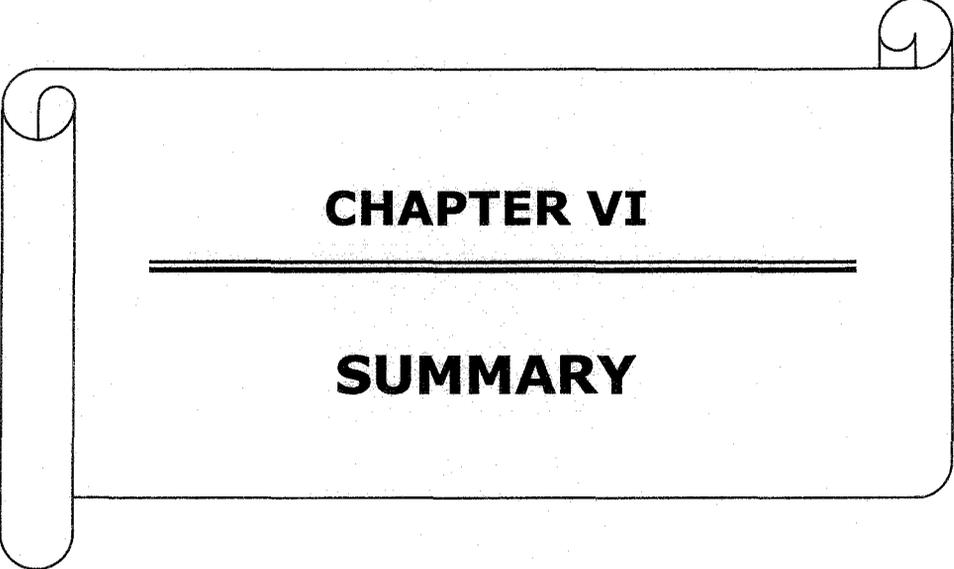
physiological parameters. On average 7% of soil moisture content ( $-1.2 \text{ Mpa} \pm 0.20$ ),  $8.73 \mu\text{mol m}^{-2} \text{ s}^{-1}$  of photosynthesis rate and  $0.42 \text{ mmol m}^{-2} \text{ s}^{-1}$  of stomatal conductance, on 21st d of drought stress induction various experiments were performed. Pigments such as chlorophyll-a (chl-a), chl-b carotenoids and RWC were found decreased in both the cultivar; however the decrease was more rapid in S.3/A3. The quantity of protein, proline, TSS, RS, phenolics, AA and ABA were enhanced more in TV-23 whereas hydrogen peroxide ( $\text{H}_2\text{O}_2$ ), superoxide anion, lipid peroxidation and electrolyte leakage were increased rapidly in S.3/A3. In between tissues, the enhancement of protein, proline, superoxide anion, AA content, lipid peroxidation and electrolyte leakage were found more in leaves as well as TSS, RS, phenolics, ABA and  $\text{H}_2\text{O}_2$  content were found more in roots of each cultivar. The activities of SOD, POX, APX and CAT were more enhanced in S.3/A3 except the activities of POX in leaves as well as CAT and SOD in roots. A decrease in starch content was also recorded in both tissues of each cultivar which was more pronounced in roots of TV-23. There are many reports that underline the intimate relationship between enhanced antioxidative enzyme activities and increased level of drought tolerance in several crop species including rice and wheat (Guo et al. 2006, Khanna-Chopra and Selote 2007). Various cellular organelles membrane damage was observed under drought stress. The damages were more severe in susceptible cultivar which proves its inefficiency in drought tolerance. Hence, observations of the damage to cell membranes, such as crista of mitochondria, thylakoids of plastid and vacuolar membranes, provides valuable information on the ability of a plant to withstand stress (Utrillas and Alegre 1997). Thus, present investigation concluded that oxidative metabolism is the prime defense mechanism under drought stress of tea which brings a cascade of biochemical changes. The closing of stomata, increasing amount of ABA and biochemical as well as rapid synthesis of antioxidative enzymes play an important role in tolerant cultivar to cope up under drought stress. Importantly, leaves were found more vulnerable to oxidative damage in comparison to roots in both tolerant and susceptible cultivars.

For the analysis of differentially expressed transcriptome under drought stress, three SSH libraries were constructed between 21 d drought induced and control plants roots of each TV-23 and S.3/A3 cultivars as well as in between them taking drought induced TV-23 plant's roots as 'tester' and S.3/A3 plants roots as 'driver'. A

total of 572 quality ESTs were generated from the inter-varietal SSH library which gave rise to 246 unigenes, containing 54 contigs and 192 singlets. Average length of the unigenes was of 528 bp with an average GC percentage of 44.28 and coding GC percentage 44.46. There were genes of molecular chaperone, heat shock proteins, transcription factors, transporters and enzymes as well as other functional proteins. The amino acid, leucine was found as highest coded (9.37%) and methionine as least coded (1.98%) in the unigenes. There was 85 quality protein domains detected of which 74 were found to be conserved on the basis of Conserved Domain Database of NCBI as well as 11 simple sequence repeats identified in the unigenes. GO analysis of unigenes as defined in *Arabidopsis* proteome clearly showed that 13.04% of genes were associated to stress. The comparison of tissue specific expression of transcripts under drought suggested the involvement of different genes in leaves and roots. Moreover, a standard cDNA library was also constructed from roots using SMART technology. A total of 346 full-length ESTs were generated which gave rise to 207 unigenes comprising 58 contigs and 149 singlets. There were only 10 numbers of reference genes found to be redundant in drought induced unigenes. This result suggested the involvement of different genes under normal growth and drought stress of roots. Comparative analyses of transcripts showed that drought stress is more lethal to the growth and development of tea plants in comparison to winter dormancy stress. A total of 123 drought associated genes were finally identified including well-known drought associated genes such as dehydrin, trehalose-6-phosphate synthase, Cu/Zn SOD, glutathione reductase etc. Interestingly, there were candidate genes of ubiquitin-proteasome and glutathione metabolism pathways as well as numbers of genes associated to sugar metabolism and transportation. The results strongly suggested the pivotal role of these pathways in drought tolerance mechanism of tea roots. Therefore, this study provides a basis for studying drought tolerance mechanism of this important commercial crop which will also be a valuable resource for functional genomics study of woody plants in future.

Genomes of higher eukaryotes encode diverse non-coding RNAs, particularly 20 to 30 nt regulatory RNAs (Zhu et al. 2008). They bind to the complementary sites on target mRNAs and repress post-transcriptional gene expression (Nagaraju 2008). It has been well-demonstrated that miRNAs play critical roles in diverse biological processes such as development, cellular differentiation, cell-cycle control, apoptosis

and oncogenesis (Sunkar et al. 2007). Due to the advantages of low cost and high efficiency, computer based miRNA identification methods have received more and more attention in recent years. Sequence and structure homologies are the main theory behind the computer-based approach for miRNA prediction. Since, tea genome sequences is not available, both EST and nucleotide databases were mined for miRNA identification. A total of 13 conserved miRNAs were identified belonging to 9 miRNA families where miR414 family has 5 members and the rest have single member in each. Average length of precursor sequence was 248 nt; however a majority of them (62%) have 65-200 nt. The identified tea miRNAs were found conserved in diverse plant species from monocotyledonous to dicotyledonous plants. These results suggested that different miRNAs might evolve at different rates not only within the same plant species, but also in different ones. Target identification of miRNAs is important to know their functional roles. Predicting miRNA targets in plants is much easier due to the high and significant complementarities to miRNA-mRNA targets. A total of 37 target genes were identified for the 9 miRNA families. The target genes were of transcription factors (8%), enzymes (30%), transporters (14%) and others involving various physiological and metabolic processes (48%). Improved understanding of molecular mechanisms of miRNAs in tea plants will help in better understanding of post-transcriptional gene silencing in response to drought and other stresses and in the development of more precise techniques.



**CHAPTER VI**

---

---

**SUMMARY**

## SUMMARY

Tea is an economically important beverage crop in the world. It is a woody perennial with a life span of more than 60 years. Hence, tea plantations often encounter several environmental stresses. Amidst all, drought is an important recurrent limiting factor to world tea cultivation which causes around 40% crop loss. Most of the tea growing areas in India and other countries are prone to drought. Owing to the fast climatic changes and water limitation, selection and/or development of cultivars that can withstand drought condition is an urgent need in modern agriculture. This study was undertaken for investigating the drought stress responses of tea genotypes with contrasting characters of drought tolerance at anatomical, physiobiochemical and molecular level in particular.

A drought experiment was conducted with two year-old tea seedlings of S.3/A3 (drought-susceptible) and TV-23 (drought-tolerant) cultivars under controlled greenhouse conditions (at  $300 \mu\text{mol m}^{-2} \text{s}^{-1}$  light intensity,  $25 \pm 2^\circ\text{C}$  with 65-70% of relative humidity) in a completely randomized block design. On average 7% of soil moisture content ( $-1.2 \text{ Mpa} \pm 0.20$ ),  $8.73 \mu\text{mol m}^{-2} \text{s}^{-1}$  of photosynthesis rate and  $0.42 \text{ mmol m}^{-2} \text{s}^{-1}$  of stomatal conductance), on 21st d of drought stress induction various experiments were performed. Oxidative metabolism was found as the prime defense mechanism of tea plants under drought stress which brings a cascade of physiobiochemical changes. The closing of stomata, elevated level of ABA and biochemical as well as rapid synthesis of antioxidative enzymes play an important role in tolerant cultivar to cope up under drought stress. Importantly, leaves were found more vulnerable to oxidative damage in comparison to roots in both tolerant and susceptible cultivars. A higher level of membrane integrity was also observed in drought-tolerant cultivar under stress.

Construction of a cDNA library and analysis of cDNA clones through ESTs approach have provided several advantages in acquiring data and gathering information on many aspects of plant biology at molecular level. For the analyses of drought responsive transcriptome of tea roots, two intra-varietal and one inter-varietal SSH cDNA libraries were constructed. One standard full-length cDNA library was also

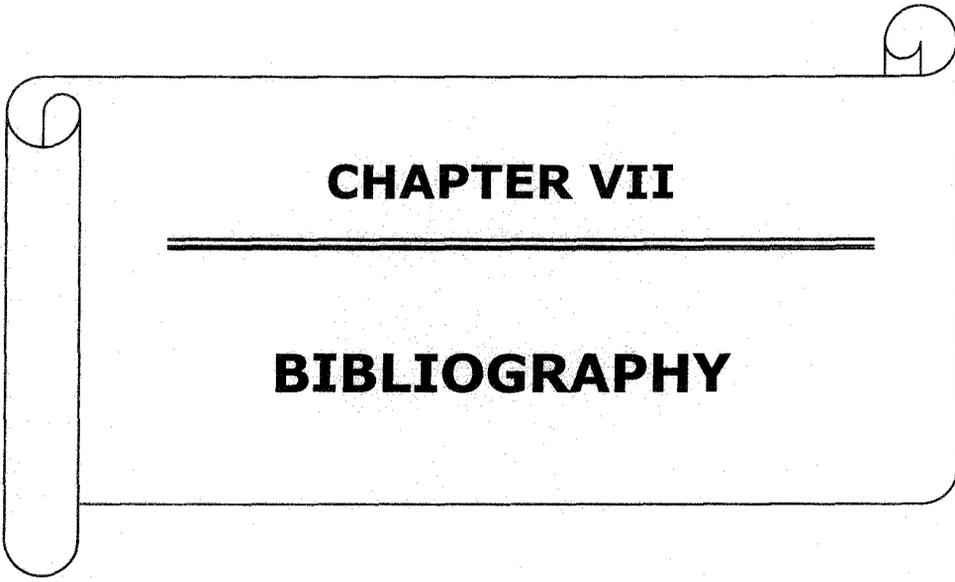
constructed using SMART protocol to be used as reference. A total of 3250 colonies randomly picked and sequenced which finally resulted 1701 drought induced (Genbank accessions: GH734203 to GH734851 of FSL1 and FSL2 libraries as well GT968791 to GT969386, GW316843 to GW317159 and GW315010-GW315149 of FSL3 library) and 811 standard full-length ESTs (Genbank accessions: GH623575 to GH624058 and HS389643 to HS389969). A total of 572 quality ESTs of the inter-varietal SSH library were produced 246 unigenes, containing 54 contigs and 192 singlets. Average length of the unigenes was of 528 bp with an average GC percentage of 44.28 and coding GC percentage 44.46. There were genes of molecular chaperone, heat shock proteins, transcription factors, transporters and enzymes as well as other functional proteins. Amino acid, leucine was found as highest coded (9.37%) and methionine as least coded (1.98%) in the unigenes. There were 85 quality protein domains detected of which 74 found to be conserved on the basis of CDD of NCBI as well as 11 EST-SSRs were identified in the unigenes. A total of 76% drought induced unigenes were assigned to functional categories i.e biological process, cellular component and molecular function, as defined in *Arabidopsis* proteome which clearly revealed that 13.04% of genes were associated to stress. Comparison of tissue specific expression of transcripts i.e leaves and roots under drought suggested the involvement of different genes in each of these tissues.

A total of 346 standard full-length quality ESTs were collapsed into 207 unigenes comprising 58 contigs and 149 singlets. There were only 10 numbers of standard genes were found common with the drought induced unigenes. This result suggested the involvement of different genes under normal growth and drought stress of roots. Moreover, a comparative transcripts analysis under normal growth, winter dormancy stress and drought stress of leaves was also performed with the help of available ESTs in the public domain. It was found that drought stress has been more lethal to the growth and development of tea plants in comparison to winter dormancy stress.

Finally, a total of 123 drought associated genes were identified including well-known drought associated genes such as dehydrin, trehalose-6-phosphate synthase, Cu/Zn SOD, GR etc. Interestingly, there were candidate genes of ubiquitin-proteasome and glutathione metabolism pathways as well as numbers of genes associated to sugar metabolism and transportation. The results strongly suggested the pivotal role of

these pathways in drought tolerance mechanism of tea roots. This study provides a basis for studying drought tolerance mechanism of this important commercial crop for the first time which will also be a valuable resource for functional genomics study of woody plants in future.

In an attempt to identify new conserved miRNAs, previously known plant miRNAs were BLASTed against all the available ESTs and full length nucleotide sequences of tea. The sequences showing homolog no more than four mismatches were predicted for their fold back structures and passed through a series of filtration criteria, which finally led to the identification of 13 conserved miRNAs belonging to 9 miRNA families. A total of 37 potential target genes in *Arabidopsis* were identified subsequently for 7 miRNA families based on their sequence complementarity which encode transcription factors (8%), enzymes (30%) and transporters (14%) as well as other proteins involved in physiological and metabolic processes (48%). These findings will accelerate the way for further researches of miRNAs and their functions in tea.



**CHAPTER VII**

---

**BIBLIOGRAPHY**

## BIBLIOGRAPHY

- Abe H, Urao T, Ito T, Seki M, Shinozaki K, Yamaguchi-Shinozaki K (2003) *Arabidopsis* AtMYC2 (bHLH) and AtMYB2 (MYB) function as transcriptional activators in abscisic acid signaling. *Plant Cell* 15: 63-78.
- Abebe T, Guenzi AC, Martin B, Cushman JC (2003) Tolerance of mannitol-accumulating transgenic wheat to water stress and salinity. *Plant Physiol* 131: 1748-1755.
- Abebe T, Melmaiee K, BergV, Wise RP (2010) Drought response in the spikes of barley: gene expression in the lemma, palea, awn, and seed. *Funct Integr Genomics* 10: 191-205.
- Abedi T, Pakniyat H (2010) Antioxidant enzyme changes in response to drought stress in ten cultivars of oilseed rape (*Brassica napus* L.). *Czech J Genet Plant Breed* 46: 27-34.
- Acar O, Turkan I, Zdemir FO (2001) Superoxide dismutase and peroxidase activities in drought sensitive and resistant barley (*Hordeum vulgare* L.) varieties. *Acta Physiologiae Planta* 3: 351-356.
- Ackerson RC, Hebert RR (1981) Osmoregulation in cotton in response to water stress: I. Alterations in photosynthesis, leaf conductance, translocation and ultrastructure. *Plant Physiol* 67: 484-488.
- Adams MD, Kelley V, Gocayne JD, Dubnick M, Polymeropoulos MH, Xiao H, Merril CR, Wu A et al. (1991) Complementary DNA sequencing: expressed sequence tags and human genome project. *Sci* 252: 1651-1656.
- Aggarwal RK, Hendre PS, Varshney RK, Bhat PR, Krishnakumar V, Singh L (2007) Identification, characterization and utilization of EST-derived genic microsatellite

markers for genome analyses of coffee and related species. *Theor Appl Genet* 114: 359-372.

Ahmed CB, Rouina BB, Sensoy S, Boukhris M, Abdallah FB (2009) Changes in gas exchange, proline accumulation and antioxidative enzyme activities in three olive cultivars under contrasting water availability regimes. *Env Exp Bot* 67: 345-352.

Aiton W (1789) A catalogue of the plants cultivated in the Royal botanic garden, Kew, London.

Akula A, Akula C, Bateson M (2000) Betaine, a novel candidate for rapid induction of somatic embryogenesis in tea (*Camellia sinensis* (L.) O. Kuntze). *Plant Growth Regu* 30: 341-246.

Alba R, Fei Z, Payton P, Liu Y, Moore SL, Debbie P, Cohn JD, Ascenzo M et al. (2004) ESTs, cDNA microarray and gene expression profiling: tools for dissecting plant physiology and development. *Plant J* 39: 697-714.

Alemzadeh A, Fujie M, Usami S, Yamada T (2005) Isolation of high-quality RNA from high-phenolic tissues of eelgrass (*Zostera marina* L.) by keeping temperature low. *Plant Mol Bio Rep* 23: 421-421.

Alkan I, Koprulu O, Alkan B (2009) Latest advances in world tea production and trade, Turkey's aspect. *World J Agric Sci* 5: 345-349.

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

Ambros V (2004) The functions of animal microRNAs. *Nat* 431: 350-355.

Angra S, Kaur S, Singh K, Pathania D, Kaur N, Sharma S, Nayyar H (2010) Water-deficit stress during seed filling in contrasting soybean genotypes: Association of

- stress sensitivity with profiles of osmolytes and antioxidants. *Int J Agric Res* 5: 328-345.
- Archak S, Nagaraju J (2007) Computational prediction of rice (*Oryza sativa*) miRNA targets. *Genomics Proteomics Bioinformatics* 5: 196-206.
- Arenas-Huertero C, Perez B, Rabanal F, Blanco-Melo D, De la Rosa C, Estrada-Navarrete G, Sanchez F, Covarrubias AA, Reyes JL (2009) Conserved and novel miRNAs in the legume *Phaseolus vulgaris* in response to stress. *Plant Mol Biol* 70: 385-401.
- Arora A, Sairam RK, Srivastava GC (2002) Oxidative stress and antioxidative system in plants. *Curr Sci* 82: 1227-1238.
- Asada K (2006) Production and scavenging of reactive oxygen species in chloroplasts and their functions. *Plant Physiol* 141: 391-396.
- Ashraf M, Mehmood S (1990) Response of four Brassica species to drought stress. *Environ Exp Bot* 30: 93-100.
- Azevedo H, Lino-Neto T, Tavares RM (2003) An improved method for high-quality RNA isolation from needle of adult maritime pine trees. *Plant Mol Biol Rep* 21: 333-338.
- Baker CC, Sieber P, Wellmer F, Meyerowitz EM (2005) The early extra petals1 mutant uncovers a role for microRNA miR164c in regulating petal number in *Arabidopsis*. *Curr Biol* 15: 303-315.
- Bano A, Yasmeen S (2010) Role of phytohormones under induced drought stress in wheat. *Pak J Bot* 42: 2579-2587.
- Bargmann OB, Munnik T (2006) The role of phospholipase D in plant stress responses. *Curr Opin Plant Biol* 9: 515-522.

Barr HD, Weatherley PE (1992) A re-examination of the relative turgidity technique for estimating water deficit in leaves. *Aust J Biol Sci* 15: 413-428.

Barron D (1997) RNAagents total RNA isolation system. *Promega Notes Magazine* 61: 33.

Bartel DP (2004) MicroRNAs: Genomics, biogenesis, mechanism, and function. *Cell* 116: 281-297.

Bartoli CG, Simontacchi M, Tambussi E, Beltrano J, Montaldi E, Puntarulo S (1999) Drought and watering-dependent oxidative stress: effect on antioxidant content in *Triticum aestivum* L. leaves. *J Exp Bot* 50:375-383.

Barua DN (1989) Science and practice in tea culture, Tea Research Association, Calcutta, p509.

Basia V, Arie A (2005) Recent advances in engineering plant tolerance to abiotic stress: achievements and limitations. *Curr Opin Biotechnol* 16: 123-132.

Bates LS, Waldren RP, Teare ID (1973) Rapid determination of free proline for water stress studies. *Plant Soil* 39: 205-207.

Bausher M, Shatters R, Chaparro J, Dang P, Hunter W, Niedz R (2003) An expressed sequence tag (EST) set from *Citrus sinensis* L. Osbeck: whole seedlings and the implications of further perennial source investigations. *Plant Sci* 165: 415-422.

Beck EH, Fetting S, Knake C, Hartig K, Bhattarai T (2007) Specific and unspecific responses of plants to cold and drought stress. *J Biosci* 32: 501-510.

Bedell JA, Budiman MA, Nunberg A, Citek RW, Robbins D et al. (2005) Sorghum Genome Sequencing by methylation filtration. *PLoS Biol* 3: e13 doi:10.1371/journal.pbio.0030013.

Begg JE, Turner NC (1976) Crop water deficits. *Advan Agron* 28: 161-217.

Berlin J, Quisenberry JE, Bailey F, Woodworth M, McMichael BL (1982) Effect of water stress on cotton leaves. *Plant Physiol* 70: 238-243.

Bernt EB (1974) Methods of Enzymatic Analysis. In: Bergmeyer HU (ed) Inorganic peroxidases, Academic Press, New York, p2246-2248.

Bezbaruah HP (1971) Cytological investigation in the family Theaceae-I. Chromosome numbers in some *Camellia* species and allied genera. *Caryologia* 24: 421-426.

Bhat PR, Krishnakumar V, Hendre PS, Kumar RP, Varshney RK, Aggarwal RK (2005) Identification and characterization of expressed sequence tags-derived simple sequence repeats markers from robusta coffee variety 'CxR' (an interspecific hybrid of *Coffea canephora* × *Coffea congensis*). *Mol Ecol Notes* 5: 80-83.

Black CA (1965) Methods of soil analysis: part I physical and mineralogical properties. American Society Agronomy, Madison, Wisconsin, USA.

Blokhina O, Vairiolainen E, Fagersted TV (2003) Antioxidants, oxidative damage and oxygen deprivation stress: a review. *Ann Bot* 91: 179-194.

Bohnert HJ, Jensen RG (1996) Strategies for engineering water-stress tolerance in plants. *Trends Biotechnol* 14: 89-97.

Bohnert HJ, Nelson DE, Jensen RG (1995) Adaptations to environmental stresses. *Plant Cell* 7: 1099-1111.

Bonnet E, Wuyts J, Rouze P and Van de Peer Y (2004) Evidence that microRNA precursors, unlike other non-coding RNAs, have lower folding free energies than random sequences. *Bioinformatics* 20: 2911-2917.

Booth WB (1830) History and description of the species of *Camellia* and *Thea*. *Trans Hort Soc, London* 7, p519-562.

Botton A, Galla G, Conesa A, Bachem C, Ramina A, Barcaccia G (2008) Large-scale gene ontology analysis of plant transcriptome-derived sequences retrieved by AFLP technology. *BMC Genomics* 9: 347 doi:10.1186/1471-2164-9-347.

Bowler C, Montagu MV, Inze D (1992) Superoxide dismutase and stress tolerance. *Annu Rev Plant Physiol Plant Mol Biol* 43: 83-116.

Bray EA (1997) Plant responses to water deficit. *Trends Plant Sci* 2: 48-54.

Brenner DE, Stevenson D, McCombie R, Katari SM, Rudd AS, Mayer FK, Palenchar MP, Runko JS, Twigg R, Dai G, Martienssen AR, Benfey NP, Coruzzi MG (2003) Expressed sequence tag analysis in *Cycas*, the most primitive living seed plant. *Genome Biol* 4: R78 doi:10.1186/gb-2003-4-12-r78.

Buell CR, Yuan Q, Ouyang S, Liu J, Zhu W, Wang A, Maiti R et al. (2005) Sequence, annotation, and analysis of synteny between rice chromosome 3 and diverged grass species. *Genome Res* 15: 1284-1291.

Burgess PJ, Carr MKV (1996) Responses of young tea (*Camellia Sinensis*) clones to drought and temperature. I. Yield and yield distribution. *Exp Agric* 32: 357-372.

Cairney J, Cowels AZL, Hsiao J, Zismann V, Liu J, Ouyang S, Nissen TF, Hamilton J, Childs K, Pullman SG, Zhang Y, Oh T, Buell RC (2006) Expressed sequence tags from loblolly pine embryos reveal similarities with angiosperm embryogenesis. *Plant Mol Biol* 62: 485-501.

Camacho-Villasana YM, Ochoa-Alejo N, Walling L, Bray EA (2002) An improved method for isolating RNA from dehydrated and nondehydrated chili pepper (*Capsicum annuum* L.) plant tissues. *Plant Mol Bio Rep* 20: 407-414.

Carvalho MHC, Contour-Ansel D (2008) (h) GR, beans and drought stress. *Plant Sig Behav* 3: 834-835.

- Chakraborty U, Dutta S, Chakraborty B (2001) Drought induced biochemical changes in young tea leaves. *Indian J Plant Physiol* 6: 103-106.
- Chakraborty U, Dutta S, Chakraborty BN (2002) Response of tea plants to water stress. *Biologia planta* 45: 557-562.
- Chance B, Maehly AC (1954) Assay of catalases and peroxidises. *Meth Biochem Anal* 1: 357-424.
- Chapman EJ, Carrington JC (2007) Specialization and evolution of endogenous small RNA pathways. *Nat Rev Genet* 8: 884-896.
- Chatzidimitriadou K, Nianiou-Obeidat I, Madesis P, Perl-Treves R, Tsaftaris A (2009) Expression of SOD transgene in pepper confer stress tolerance and improve shoot regeneration. *Electron J Biotechnol* 12: doi: 10.2225/vol12-issue4-fulltext-10.
- Chaves MM (1991) Effects of water deficits on carbon assimilation. *J Exp Bot* 42: 1-16.
- Chaves MM, Maroco JP, Pereira JS (2003) Understanding plant responses to drought- from genes to the whole plant. *Funct Plant Biol* 30: 239-264.
- Chen L, Zhao LP, Gao QK (2005) Generation and analysis of expressed sequence tags from the tender shoots cDNA library of tea plant (*Camellia sinensis*). *Plant Sci* 168: 359-363.
- Chen Q, Vazquez EJ, Moghaddas S, Hoppel CL, Lesnefsky EJ (2003) Production of reactive oxygen species by mitochondria. Central role of complex III. *J Biol Chem* 278: 36027-36031.
- Cheruiyot EK, Mumera LM, Ng'etich WK, Hassanali A, Wachira F (2007) Polyphenols as potential indicators for drought tolerance in tea (*Camellia sinensis* L.). *Biosci Biotechnol Biochem* 71: 2190-2197.

Cheruiyot EK, Mumera LM, Ng'etich WK, Hassanali A, Wachira F, Wanyoko JK (2008) Shoot epicatechin and epigallocatechin contents respond to water stress in tea [*Camellia sinensis* (L.) O. Kuntze]. *Biosci Biotechnol Biochem* 72: 1-8.

Cho JY, Mizutani M, Shimizu BI, Kinoshita T, Ogura M, Tokoro K, Lin ML, Sakata K (2007) Chemical profiling and gene expression profiling during the manufacturing process of Taiwan oolong tea, "Oriental Beauty". *Biosci Biotechnol Biochem* 71: 1476-1486.

Chow SK, Wan LK, Isa NM, Bahari A, Tan HS, Harikrishna K, Yeang YH (2007) Insights into rubber biosynthesis from transcriptome analysis of *Hevea brasiliensis* latex. *J Exp Bot* 58: 2429-2440.

Ciamporova M (1987) The development of structural changes in epidermal cells of maize roots during water stress. *Biol Planta* 29: 290-294.

Ciechanover A, Orian A, Schwartz AL (2000) Ubiquitin-mediated proteolysis: biological regulation via destruction. *BioEssays* 22: 442-451.

Clement M, Lambert A, Herouart D, Boncompagni E (2008) Identification of new up-regulated genes under drought stress in soybean nodules. *Gene* 426: 15-22.

Contour-ansel D, Torres-franklin ML, Cruz de Carvalho MH, D'arcy-Lameta A, Zuily-Fodil Y (2006) Glutathione Reductase in leaves of cowpea: cloning of two cDNAs, expression and enzymatic activity under progressive drought stress, desiccation and abscisic acid treatment. *Ann Bot* 98: 1279-1287.

Cordeiro GM, Casu R, McIntyre CL, Manners JM, Henry RJ (2001) Microsatellite markers from sugarcane (*Saccharum spp.*) ESTs cross transferable to erianthus and sorghum. *Plant Sci* 160: 1115-1123.

Cornic G, Briantais JM (1991) Partitioning of photosynthetic electron flow between CO<sub>2</sub> and O<sub>2</sub> leaf (*Phaseolus vulgaris* L) at different CO<sub>2</sub> concentration and drought stress. *Planta* 183: 178-184.

- Crevecoeur M, Kevers C, Greppin H, Gaspar T (1987) A comparative biochemical and cytological characterization of normal and habituated sugarbeet calli. *Biol Planta* 29: 1-6.
- Cruz de Carvalho MH (2008) Drought stress and reactive oxygen species. *Plant Sig Behav* 3: 156-165.
- Cushman JC (2001) Osmoregulation in plants: implications for agriculture. *Amer Zool* 41: 758-769.
- Davies WJ, Zhang J (1991) Root signals and the regulation of growth and development of plants in drying soil. *Annu Rev Plant Physiol Plant Mol Biol* 42: 55-76.
- Derory J, Léger P, Garcia V, Schaeffer J, Hauser TM, Salin F, Luschnig C, Plomion C, Glössl J, Kremer A (2006) Transcriptome analysis of bud burst in sessile oak (*Quercus petraea*). *New Phytol* 170: 723-738.
- Devi P (2007) Principles and methods of plant molecular biology biochemistry and genetics, Agrobios, Jodhpur, India.
- Devitt LC, Sawbridge T, Holton TA, Mitchelson K, Dietzgen GR (2006) Discovery of genes associated with fruit ripening in *Carica papaya* using expressed sequence tags. *Plant Sci* 170: 356-363.
- Dezulian T, Palatnik J, Huson D, Weigel D (2005) Conservation and divergence of microRNA families in plants. *Genome Biol* 6: P13 doi:10.1186/gb-2005-6-11-p13.
- Diatchenko L, Lau YF, Siebert PD (1996) Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries. *Proc Nat Acad Sci* 93: 6025-6030.

Dixon DP, Laphorn A, Madesis P, Mudd EA, Day A, Edwards R (2008) Binding and glutathione conjugation of porphyrinogens by plant glutathione transferases. *J Biol Chem* 283: 20268-20276.

Dixon RA, Steele CL (1999) Flavonoids and isoflavonoids-a gold mine for metabolic engineering. *Trends Plant Sci* 4: 394-400.

Dong Z, Han M-H, Fedoroff N (2008) The RNA-binding proteins HYL1 and SE promote accurate in vitro processing of pri-miRNA by DCL1. *Proc Natl Acad Sci USA* 105: 9970-9975.

Dood AW (1994) Tissue culture of tea (*Camellia sinensis* (L.) O. Kuntze)-a review. *Inter J Trop Agric* 12: 212-247.

Du J, Chen X, Li W, Gao Q (2004) Osmoregulation mechanism of drought stress and genetic engineering strategies for improving drought resistance in plants. *Forestry Studies in China* 6: 56-62.

Duguid JR, Dinauer MC (1990) Library subtraction of in vitro cDNA libraries to identify differentially expressed genes in scrapie infection. *Nucl Acids Res.* 18: 2789-2792.

Eckardt NA (2005) MicroRNAs regulate auxin homeostasis and plant development. *Plant Cell* 17: 1335-1338.

Efeoğlu B, Ekmekçi Y, Çiçek N (2009) Physiological responses of three maize cultivars to drought stress and recovery. *South African J Bot* 75: 34-42.

El-Bashiti T, Hamamcı H, Öktem HA, Yücel M (2005) Biochemical analysis of trehalose and its metabolizing enzymes in wheat under abiotic stress conditions. *Plant Sci* 169: 47-54.

El-Tayeb MA (2006) Differential response of two *Vicia faba* cultivars to drought: growth, pigments, lipid peroxidation, organic solutes, catalase and peroxidase activity. *Acta Agronomica Hungarica* 54: 25-37.

Erwin HB, Fettig S, Knake C, Hartig K, Bhattarai T (2007) Specific and unspecific responses of plants to cold and drought stress. *J Biosci* 32: 501-510.

Ewing R, Poirot O, Claverie JM (1999) Comparative analysis of the *Arabidopsis* and rice expressed sequence tag (EST) sets. *In Silico Biol* 1: 197-213.

Farooqui AHA, Kumar R, Fatima S, Sharma S (2000) Responses of different genotypes of lemon grass (*Cymbopogon flexuosus* and *C. pendulus*) to water stress. *J Plant Biol* 27: 277-282.

Fordham R (1971) Stomatal physiology and the water relations of the tea bush. In Carr MKV and Carr S (ed) *Water and the tea plant*, Tea Boards of Kenya, Uganda and Tanzania Tea Authority, p89-100.

Foyer CH, Noctor G (2005) Oxidant and antioxidant signaling in plants: a re-evaluation of the concept of oxidative stress in a physiological context. *Plant Cell Environ* 28: 1056-1071.

Fujita M, Fujita Y, Maruyama K, Seki M, Hiratsu K, Ohme-Takagi M, Tran LSP, Yamaguchi-Shinozaki K, Shinozaki K (2004) A dehydration-induced NAC protein, RD26, is involved in a novel ABA-dependent stress-signaling pathway. *Plant J* 39: 863-876.

Fukuoka H, Yamaguchi H, Nunome T, Negoro S, Miyatake K, Ohyama A (2010) Accumulation, functional annotation, and comparative analysis of expressed sequence tags in eggplant (*Solanum melongena* L.), the third pole of the genus *Solanum* species after tomato and potato. *Gene* 450: 76-84.

Gallé A, Csiszár J, Secenji M, Tari I, Györgyey J, Dudits D, Erdei L (2005) Changes of glutathione S-transferase activities and gene expression in *Triticum aestivum* during polyethylene-glycol induced osmotic stress. *Acta Biol Szeged* 49: 95-96.

Gallego SM, Benavides MP, Tomaro ML (1996) Effect of heavy metal ion excess on sunflower leaves: evidence for involvement of oxidative stress. *Plant Sci* 121: 151-159.

Gao WR, Wanga XS, Liu QY, Peng H, Chen C, Li c JG, Zhang JS, Hub SN, Maa H (2008) Comparative analysis of ESTs in response to drought stress in chickpea (*C. arietinum* L.). *Biochem Biophys Res Commun* 376: 578-583.

Garg AK, Kim JK, Owens TG, Ranwala AP, Choi YD, Kochian LV, Wu RJ (2002) Trehalose accumulation in rice plants confers high tolerance levels to different abiotic stresses. *Proc Nat Acad Sci* 10: 15898-15903.

Gasic K, Hernandez A, Korban SS (2004) RNA Extraction from different apple tissues rich in Polyphenols and polysaccharides for cDNA library construction. *Plant Mol Bio Rep* 22: 437a-g.

Gee GW, Sandanam S, Kulasegaram S, Coomaraswamy A (1982) Effect of shade on leaf water diffusion resistance in clonal tea (*Camellia sinensis* L.). *Tea Q* 51: 12-20.

Gehrig HH, Winter JK, Cushman A, Borland, Taybi T (2000) An improved RNA isolation method for succulent plant species rich in polyphenols and polysaccharides. *Plant Mol Bio Rep* 18: 369-376.

Gollan T, Passioura JB, Munns R (1986) Soil water status affects the stomatal conductance of fully turgid wheat and sunflower leaves. *Aust J Plant Physiol* 13: 459-464.

Goyal K, Walton LJ, Tunnacliffe A (2005) LEA proteins prevent protein aggregation due to water stress. *Biochem J* 388: 151-157.

Grattapaglia D (2004) Integrating genomics into Eucalyptus breeding. *Genet Mol Res* 3: 369-379.

Griffiths-Jones S, Bateman A, Marshall M, Khanna A, Eddy SR (2003) Rfam: an RNA family database. *Nucl Acids Res* 31: 439-441.

Griffiths-Jones S, Grocock RJ, Van DS, Bateman A, Enright AJ (2006) miRBase: microRNA sequences, targets and gene nomenclature. *Nucl Acids Res* 34: D140-D144.

Griffiths-Jones S, Saini HK, Van DS, Enright AJ (2008) miRBase: tools for microRNA genomics. *Nucl Acids Res* 36: D154-D158.

Grover A, Kapoor A, Lakshmi OS, Agarwal S, Sahi C, Agarwal SK, Agarwal M, Dubey H (2001) Understanding molecular alphabets of the plant abiotic stress responses. *Curr Sci* 80: 2-25.

Grzesiak S, Iijima M, Kono Y, Yamauchi A (1997) Differences in drought tolerance between cultivars of field bean and field pea. Morphological characteristics, germination and seedling growth. *Acta Physiologiae Planta* 19: 339-348.

Guddeti S, Zhang DC, Li AL, Leseberg CH, Kang H, Li XG, Zhai WX, Johns MA, Mao L (2005) Molecular evolution of the rice miR395 gene family. *Cell Res* 15: 631-638.

Gullner G, Kömives T, Rennenberg H (2001) Enhanced tolerance of transgenic poplar plants overexpressing  $\gamma$ -glutamylcysteine synthetase towards chloroacetanilide herbicides. *J Exp Bot* 52: 971-979.

Guo HS, Xie Q, Fei JF, Chua NH (2005) MicroRNA directs mRNA cleavage of the transcription factor NAC1 to down regulate auxin signals for *Arabidopsis* lateral root development. *Plant Cell* 17: 1376-1386.

Guo P, Baum M, Grando S, Ceccarelli S, Bai G, Li R, Korff MV, Varshney RK, Graner A, Valkoun J (2009) Differentially expressed genes between drought-tolerant and drought-sensitive barley genotypes in response to drought stress during the reproductive stage. *J Exp Bot* 60: 3531-3544.

Guo Q, Xiang A, Yang Q, Yang YZ (2007) Bioinformatic identification of microRNAs and their target genes from *Solanum tuberosum* expressed sequence tags. Chinese Sci Bullet 52: 2380-2389.

Guo Q, Zhang J, Gao Q, Xing S, Li F, Wang W (2008) Drought tolerance through overexpression of monoubiquitin in transgenic tobacco. J Plant Physiol 165: 1745-1755.

Guo Z, Ou W, Lu S, Zhong Q (2006) Differential responses of antioxidative system to chilling and drought in four rice cultivars differing in sensitivity. Plant Physiol Biochem 44: 828-836.

Hakimi A, Monneveux AP, Galiba G (1995) Soluble sugars, proline and relative water content (RWC) as traits for improving drought tolerance and divergent selection for RWC from *T. polonicum* into *T. durum*. J Genet Breed 49: 237-244.

Han J, Lee Y, Yeom KH, Nam JW, Heo I, Rhee JK, Sohn SY, Cho Y, Zhang BT, Kim VN (2006) Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. Cell 125: 887-901.

Han Y, Luan F, Zhu H, Shao Y, Chen A, Lu C, Luo Y, Zhu B (2009) Computational identification of microRNAs and their targets in wheat (*Triticum aestivum* L.). Sci in China Series C: Life Sciences 52: 1091-1006.

Hara E, Kato T, Nakada S, Sekiya S, Oda K (1991) Subtractive cDNA cloning using oligo (dT) 30-latex and PCR: isolation of cDNA clones specific to undifferentiated human embryonal carcinoma cells. Nucl Acids Res 19: 7097-7104.

He XF, Fang YY, Feng L, Guo HS (2008) Characterization of conserved and novel microRNAs and their targets, including a TuMV-induced TIR-NBS-LRR class R gene-derived novel miRNA in Brassica. FEBS lett 582: 2445-2452.

Hernández I, Alegre L, Munné-Bosch S (2006) Enhanced oxidation of flavan-3-ols and proanthocyanidin accumulation in water-stressed tea plants. *Phytochem* 67: 1120-1126.

Hicks LM, Cahoon RE, Bonner ER, Rivard RS, Sheffield J, Jez JM (2007) Thiol-based regulation of redox-active glutamate-cysteine ligase from *Arabidopsis thaliana*. *Plant Cell* 19: 2653-2661.

Hoffmann CM (2010) Sucrose accumulation in sugar beet under drought stress. *J Agron Crop Sci* 196: 243-252.

Hong-bo S, Li-ye C, Ming-an S, Jaleel CA, Hong-mei M (2008) Higher plant antioxidants and redox signaling under environmental stresses. *C R Biologies* 331: 433-441.

HongBo S, ZongSuo L, MingAn S (2005) Changes of anti-oxidative enzymes and MDA content under soil water deficits among 10 wheat (*Triticum aestivum* L.) genotypes at maturation stage. *Colloids Surf B Biointerfaces* 45: 7-13.

Hou X, Xie K, Yao J, Qi Z, Xiong L (2009) A homolog of human ski-interacting protein in rice positively regulates cell viability and stress tolerance. *Proc Nat Acad Sci* 106: 6410-6415.

Hsiao TC (1973) Plant responses to water stress. *Ann Rev Plant Physiol* 24: 519-570.

Hsieh LC, Lin SI, Shih AC, Chen JW, Lin WY, Tseng CY, Li WH, Chiou TJ (2009) Uncovering small RNA-mediated responses to phosphate deficiency in *Arabidopsis* by deep sequencing. *Plant Physiol* 151: 2120-2132.

Hu H, Dai M, Yao J, Xiao B, Li X, Zhang Q, Xiong L (2006) Overexpressing a NAM, ATAF, and CUC (NAC) transcription factor enhances drought resistance and salt tolerance in rice. *Proc Nat Acad Sci* 103: 12987-12992.

Huang J, Hirji R, Adam L, Rozwadowski KL, Hammerlindl JK, Keller WA, Selvaraj G (2000) Genetic engineering of glycinebetaine production toward enhancing stress tolerance in plants: metabolic limitations. *Plant Physiol* 122: 747-756.

Huang SQ, Peng J, Qiu CX, Yang ZM (2009) Heavy metal-regulated new microRNAs from rice. *J Inorg Biochem* 103: 282-287.

Ikoma Y, Yano M, Ogawa K, Yoshioka T, Xu ZC, Hisada S, Omura M, Moriguchi T (1996) Isolation and evaluation of RNA from polysaccharide-rich tissues in fruit for quality by cDNA library construction and RT-PCR. *J Japan Soc Hort Sci* 64: 809-814.

Jae-Heung KO, Seung Y, Kyung-Hwan H (2006) Upregulation of an *Arabidopsis RING-H2* gene, *XERICO*, confers drought tolerance through increased abscisic acid biosynthesis. *Plant J* 47: 343-355.

Jagadeeswaran G, Zheng Y, Li YF, Shukla LI, Matts J, Hoyt P, Macmil SL, Wiley GB, Roe BA, Zhang W et al. (2009) Cloning and characterization of small RNAs from *Medicago truncatula* reveals four novel legume-specific microRNA families. *New phytol* 184: 85-98.

Jaillon O, Aury JM, Noel B, Policriti A, Clepet C, Casagrande A, Choisne N, Aubourg S, Vitulo N, Jubin C et al. (2007) The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nat* 449: 463-467.

Jain NK (1999) *Global advances in tea science*, Aravali Books International, New Delhi, p882.

Jain SM, Newton RJ (1990) Prospects of biotechnology for tea improvement. *Proc Indian Natl Sci Acad* 6: 441-448.

Jaiprakash RM, Pillai B, Venkatesh P, Subramanian N, Sinkar PV, Sadhale PP (2003) RNA Isolation from high-phenolic freeze-dried tea (*Camellia sinensis*) Leaves. *Plant Mol Bio Rep* 21: 465a-g.

Jaleel CA, Manivannan P, Wahid A, Farooq M, Al-Juburi HJ, Somasundaram R, Panneerselvam R (2009) Drought stress in plants: a review on morphological characteristics and pigments composition. *Int J Agric Biol* 11: 100-105.

Jansson S, Douglas JC (2007) *Populus*: A model system for plant biology. *Annu Rev Plant Biol* 58: 435-58.

Jennewein S, Wildung RM, Chau M, Walker K, Croteau R (2004) Random sequencing of an induced *Taxus* cell cDNA library for identification of clones involved in Taxol biosynthesis. *Proc Nat Acad Sci* 101: 9149-9154.

Jeyaramraja PR, Pius PK, RajKumar R, Jayakumar D (2003) Soil moisture stress-induced alterations in bioconstituents determining tea quality. *J Food Agric* 83: 1187-1191.

Ji W, Zhu Y, Li Y, Yang L, Zhao X, Cai H, Bai X (2010) Over-expression of a glutathione S-transferase gene, *GsGST*, from wild soybean (*Glycine soja*) enhances drought and salt tolerance in transgenic tobacco. *Biotech Lett* 32: 1173-1179.

Jian X, Zhang L, Li G, Zhang L, Wang X, Cao X, Fang X, Chen F. (2010) Identification of novel stress-regulated microRNAs from *Oryza sativa* L.. *Genomics* 95: 47-55.

Jiang F, Hartung W (2007) Long-distance signalling of abscisic acid (ABA): the factors regulating the intensity of the ABA signal. *J Exp Bot* doi: 10.1093/jxb/erm127.

Jin JQ, Cui HR, Gong XC, Chen WY, Xin Y (2007) Studies on tea plants (*Camellia sinensis*) germplasms using EST-SSR marker. *Yi Chuan* 29: 103-108 (In Chinese).

Jin W, Li N, Zhang B, Wu F, Li W, Guo A, Deng Z (2008) Identification and verification of microRNA in wheat (*Triticum aestivum*). *J Plant Res* 121: 351-355.

Johnson C, Kasprzewska A, Tennessen K, Fernandes J, Nan GL, Walbot V, Sundaresan V, Vance V, Bowman LH (2009) Clusters and superclusters of phased small RNAs in the developing inflorescence of rice. *Genome Res* 19: 1429-1440.

Jones DT, Taylor WR, Thornton JM (1992) The rapid generation of mutation data matrices from protein sequences. *Comput Appl Biosci* 8: 275-282.

Jones RC, Steane DA, Potts BM, Vaillancourt RE (2002) Microsatellite and morphological analysis of *Eucalyptus globulus* populations. *Can J Forest Res* 32: 59-66.

Jones-Rhoades MW, Bartel DP (2004) Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. *Mol cell* 14: 787-799.

Jones-Rhoades MW, Bartel DP, Bartel B (2006) MicroRNAs and their regulatory roles in plants. *Annu Rev Plant Biol* 57: 19-53.

Jones-Rhoades WM, Bartel DP (2004) Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. *Mol Cell* 14: 787-799.

Jordan CM, DeVay JE (1990) Lysosome disruption associated with hypersensitive reaction in the potato-*Phytophthora infestans* host-parasite interaction. *Physiol Plant Pathol* 36: 221-236.

Juarez MT, Kui JS, Thomas J, Heller BA, Timmermans MC (2004) MicroRNA-mediated repression of rolled leaf1 specifies maize leaf polarity. *Nat* 428: 84-88.

Kaldenhoff R, Eckert M (1999) Features and function of plant aquaporins. *J Photochem Photobiol* 52: 1-6.

Kameli A, Losel DM (1993) Carbohydrates and water status in wheat plants under water stress. *New Phytol* 125: 609-614.

- Kang JY, Choi HI, Im MY, Kim SY (2002) *Arabidopsis* basic leucine zipper proteins that mediate stress-responsive abscisic acid signaling. *Plant Cell* 14: 343-357.
- Kantety RV, La Rota M, Matthews DE, Sorrells ME (2002) Data mining for simple sequence repeats in expressed sequence tags from barley, maize, rice, sorghum and wheat. *Plant Mol Biol* 48: 501-510.
- Karasawa K (1932) On triploid *Thea*. *Bot Mag Tokyo* 46: 458-460.
- Kasschau KD, Xie Z, Allen E, Llave C, Chapman EJ, Krizan KA, Carrington JC (2003) P1/HC-Pro, a viral suppressor of RNA silencing, interferes with *Arabidopsis* development and miRNA function. *Dev Cell* 4: 205-217.
- Kato M, Mizuno K, Crozier A, Fujimura T, Ashihara H (2000) Caffeine synthase gene from tea leaves. *Nat* 406: 956-957.
- Kato S, Sekine S, Oh SW, Kim NS, Umezawa Y, Abe N, Yokoyama-Kobayashi M, Aoki T (1994) Construction of a human full-length cDNA bank. *Gene* 150: 243-250.
- Kavikishore PB, Sangam S, Amrutha RN, Srilaxmi P, Naidu KR, Rao KRSS, Rao S, Reddy KJ, Theriappan P, Sreenivasulu N (2005) Regulation of proline biosynthesis, degradation, uptake and transport in higher plants: its implications in plant growth and abiotic stress tolerance. *Curr Sci* 88: 424-438.
- Kazuhiro D, Masayasu N, Yamashita I (1996) Changes in lipid peroxidation and antioxidants content in cotyledons of Japanese radish. *J Japan Soc Hort Sci* 65: 603-608.
- Keles Y, Oncel I (2002) Response of the antioxidative defence system to temperature and water stress combinations in wheat seedlings. *Plant Sci* 163: 783-790.

Ketchum REB, Warren RC, Klima LJ, Lopez-Gutierrez F, Nabors MW (1991) The mechanism and regulation of proline accumulation in suspension cultures of the halophyte grass *Distichlis spicata* L. *J Plant Physiol* 137: 368-374.

Khan Barozai MY, Irfan M, Yousaf R, Ali I, Qaisar U, Maqbool A, Zahoor M, Rashid B, Hussnain T, Riazuddin S (2008) Identification of micro-RNAs in cotton. *Plant Physiol Biochem* 46: 739-751.

Khanna-Chopra R, Selote DS (2007) Acclimation to drought stress generates oxidative stress tolerance in drought-resistant than susceptible wheat cultivar under field conditions. *Environ Exp Bot* 60: 276-283.

Kim VN (2005) MicroRNA biogenesis: co-ordinated cropping and dicing. *Nat Rev Mol Cell Biol* 6: 376-385.

Kohler A, Delaruelle C, Martin M, Encelot N, Martin F (2003) The poplar root transcriptome: analysis of 7000 expressed sequence tags. *FEBS Lett* 542: 37-41.

Kondo K (1975) Cytological studies in cultivated species of *Camellia*. PhD Dissertation, Univ N Carol, Chapel Hill, p260.

Kondo K (1977) Cytological studies in cultivated species of *Camellia* II. Hexaploid species and their hybrids. *Japan J Breed* 27: 333-344.

Konwar BK (2004) Evaluation and development of planting material. In Notes on tea management, Tea Research Association, Tocklai Experimental Station, Assam, India, p1-6.

Konwar BK, Das SC, Bordoloi BJ, Dutta RK (1998) Hairy root development in tea through *Agrobacterium rhizogenes*-mediated genetic transformation. *Two and a Bud* 45: 19-20.

Krol J, Sobczak K, Wilczynska U, Drath M, Jasinska A, Kaczynska D, Krzyzosiak WJ (2004) Structural Features of MicroRNA (miRNA) Precursors and their relevance to

miRNA biogenesis and small interfering RNA/short hairpin RNA design. *J Biol Chem* 279: 42230-42239.

Kuboi T, Suda M, Konishi S (1991) Preparation of protoplasts from the leaves. *Proc Int Symp Tea Sci*, Shizuoka, Japan, p427-431.

Kumapala SP, Mukhopadhyay S (2005) Mining and survey of simple sequence repeats in expressed sequence tags of dicotyledonous species. *Genome* 48: 985-998.

Kushwaha HR, Singh AK, Sopory SK, Singla-Pareek SL, Pareek A (2009) Genome wide expression analysis of CBS domain containing proteins in *Arabidopsis thaliana* (L.) Heynh and *Oryza sativa* L. reveals their developmental and stress regulation. *BMC Genomics* 10:200 doi:10.1186/1471-2164-10-200.

Kutter C, Schob H, Stadler M, Meins F, Si-Ammour A (2007) MicroRNA-mediated regulation of stomatal development in *Arabidopsis*. *Plant cell* 19: 2417-2429.

Kwak JM, Mori IC, Pei Z-M, Leonhardt N, Torres MA, Dangl JL, Bloom RE, Bodde S, Jones JDG, Schroeder JI (2003) NADPH oxidase *AtrbohD* and *AtrbohF* genes function in ROS dependent ABA signaling in *Arabidopsis*. *EMBO J* 22: 2623-2633.

Kwak PB, Wang QQ, Chen XS, Qiu CX, Yang ZM (2009) Enrichment of a set of microRNAs during the cotton fiber development. *BMC Genomics* 10: 457 doi:10.1186/1471-2164-10-457.

Lacombe S, Nagasaki H, Santi C, Duval D, Piegu B, Bangratz M, Breitler JC, Guiderdoni E, Brugidou C, Hirsch J et al. (2008) Identification of precursor transcripts for 6 novel miRNAs expands the diversity on the genomic organisation and expression of miRNA genes in rice. *BMC Plant Biol* 8: 123 doi:10.1186/1471-2229-8-123.

Lal L, Sahoo R, Gupta RK, Sharma P, Kumar S (2001) RNA isolation from high-phenolic tea leaves and apical buds. *Plant Mol Bio Rep* 19: 181a-f.

Lal S, Ravi V, Jitendra P, Khurana, Khurana P (2009) Repertoire of leaf expressed sequence tags (ESTs) and partial characterization of stress-related and membrane transporter genes from mulberry (*Morus indica* L.). *Tree Genet Genom* 5: 359-374.

Lambert JD, Yang CS (2003) Mechanisms of cancer prevention by tea constituents. *J Nutr* 133: 3262S-3267S.

Lata C, Sahu PP, Prasad M (2010) Comparative transcriptome analysis of differentially expressed genes 3 in foxtail millet (*Setaria italica* L.) during dehydration stress. *Biochem Biophys Res Commun* 393: 720-727.

Laufs P, Peaucelle A, Morin H, Traas J (2004) MicroRNA regulation of the CUC genes is required for boundary size control in *Arabidopsis* meristems. *Dev* 131: 4311-4322.

Lazo GR, Chao S, Hummel DD, Edwards H, Crossman CC, Lui N, Matthews DE, Carollo VL, Hane DL et al. (2004) Development of an expressed sequence tag (EST) resource for wheat (*Triticum aestivum* L.): EST generation, unigene analysis, probe selection and bioinformatics for a 16,000-locus bin-delineated map. *Genet* 168: 585-593.

Lee R C, Feinbaum RL, Ambros V (1993). The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* 75: 843-854.

Lee SB, Kwon HB, Kwon SJ, Park SC, Jeong MJ, Han SE, Byun MO, Daniell H (2003) Accumulation of trehalose within transgenic chloroplasts confers drought tolerance. *Mol Breed* 11: 1-13.

Lelandais-Briere C, Naya L, Sallet E, Calenge F, Frugier F, Hartmann C, Gouzy J, Crespi M (2009) Genome-wide *Medicago truncatula* small RNA analysis revealed novel microRNAs and isoforms differentially regulated in roots and nodules. *Plant Cell* 21: 2780-2796.

Levitt J (1980) Chilling, freezing, and high temperature stresses. In responses of plants to environmental stress (Ed), Academic Press, New York, p23-64.

Lewers SK, Saski AC, Cuthbertson JB, Henry CD, Staton EM, Main SD, Dhanaraj LA, Rowland JL, Tomkins PJ (2008) A blackberry (*Rubus* L.) expressed sequence tag library for the development of simple sequence repeat markers. *BMC Plant Biol* 8: 69 doi:10.1186/1471-2229-8-69.

Li D, Li C, Sun H, Wang W, Liu L, Zhang Y (2010) Effects of drought on soluble protein content and protective enzyme system in cotton leaves. *Frontiers Agric China* 4: 456-462.

Li X, Wu XH, Dillon KS, Southerton GS (2009) Generation and analysis of expressed sequence tags from six developing xylem libraries in *Pinus radiata* D. Don. *BMC Genomics* 10: 41 doi:10.1186/1471-2164-10-41.

Li Y, Dankher OP, Carreira L, Smith AP, Meagher RB (2006) The shoot-specific expression of  $\gamma$ -Glutamylcysteine synthetase directs the long-distance transport of thiol-peptides to roots conferring tolerance to mercury and arsenic. *Plant Physiol* 141: 288-298.

Li YF, Zheng Y, Addo-Quaye C, Zhang L, Saini A, Jagadeeswaran G, Axtell MJ, Zhang W, Sunkar R (2010) Transcriptome-wide identification of microRNA targets in rice. *The Plant J* 62: 742-759.

Liang P, Pardee A (1992) Differential Display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Sci* 257: 967-971.

Lichtenthaler HK (1987) Chlorophylls and carotenoids, pigments of photosynthetic biomembranes. *Meth Enzym* 148: 350-382.

Lima ALS, DaMatta FM, Pinheiro HA, Totola MR, Loureiro ME (2002) Photochemical responses and oxidative stress in two clones of *Coffea canephora* under water deficit conditions. *Environ Exp Bot* 47: 239-247.

Lin YS, Tasi YJ, Tsay JS, Lin JK (2003) Factors affecting the levels of tea polyphenols and caffeine in tea leaves. *J Agric Food Chem* 51: 1864-1873.

Lindow M, Jacobsen A, Nygaard S, Mang Y, Krogh A (2007) Intragenomic matching reveals a huge potential for miRNA-mediated regulation in plants. *PLoS comput biol* 3, e238 doi:10.1371/journal.pcbi.0030238.

Liu Q, Sakuma Y, Abe H, Kasuga M, Miura S, Yamaguchi-Shinozaki K, Shinozaki K (1998) Two transcription factors, DREB1 and DREB2, with an EREBP/AP2 DNA binding domain, separate two cellular signal transduction pathways in drought and low temperature-responsive gene expression, respectively, in *Arabidopsis*. *Plant Cell* 10: 1491-1406.

Llave C, Xie Z, Kasschau KD, Carrington JC (2002) Cleavage of Scarecrow-like mRNA targets directed by a class of *Arabidopsis* miRNA. *Sci* 297: 2053-2056.

Loewus FA (1988) Ascorbic acid and its metabolic products. In Preiss J (ed) *The biochemistry of plants*, Academic Press, New York 14, p85-107.

Loggini B, Scartazza A, Brugnoli E, Navari-Izzo F (1999) Antioxidative defense system, pigment composition, and photosynthetic efficiency in two wheat cultivars subjected to drought. *Plant Physiol* 119: 1091-1109.

Lokko Y, Anderson JV, Rudd S, Raji A, Horvath D, Mikel MA, Kim R, Liu L, Hernandez A, Dixon AGO, Ingelbrecht IL (2007) Characterization of an 18,166 EST dataset for cassava (*Manihot esculenta* Crantz) enriched for drought-responsive genes. *Plant Cell Rep* 26: 1605-1618.

Longemann, Schell J, Willmitzer L (1987) Improved method for the isolation of RNA from plant tissues. *Anal Biochem* 163: 16-20.

- Lu C, Jeong DH, Kulkarni K, Pillay M, Nobuta K, German R, Thatcher SR, Maher C, Zhang L, Ware D et al. (2008b) Genome wide analysis for discovery of rice microRNAs reveals natural antisense microRNAs (nat-miRNAs). *Proc Nat Acad Sci* 105: 4951-4956.
- Lu D (1992) Studies on the drought injuries with relation to activated oxygen metabolism in tea plant (*Camellia sinensis*). *Acta Agriculturae Universitatis Zhejiangensis* 18: 56-62.
- Lu XY, Huang XL (2008) Plant miRNAs and abiotic stress responses. *Biochem Biophys Res Commun* 368: 458-462.
- Lu Y, Gan Q, Chi X, Qin S (2008a) Roles of microRNA in plant defense and virus offense interaction. *Plant Cell Rep* 27: 1571-1579.
- Lu Y, Yang X (2010) Computational Identification of novel microRNAs and their targets in *Vigna unguiculata*. *Comp Funct Genomics* doi:10.1155/2010/128297.
- MacRae JC, Smith D, McCready RM (1974) Starch estimation in leaf tissue - a comparison of results using six methods. *J Sci Food Agric* 25: 1465-1469.
- Maggio A, Miyazaki S, Veronese P, Fujita T, Ibeas JI, Damasz B, Narasimhan ML, Hasegawa PM, Joly RJ, Bressan RA (2002) Does proline accumulation play an active role in stress-induced growth reduction. *Plant J* 31: 699-712.
- Mahajan S, Tuteja N (2006) Cold, salinity and drought stresses: an overview. *Arch Biochem Biophys* 444: 139-158.
- Maher C, Timmermans M, Stein L, Ware D (2004) Identifying microRNAs in plant genomes. *Proc IEEE CSB* 6: 718-723.
- Malick CP, Singh MB (1980) *Plant enzymology and histoenzymology*, Kalyani publications, New Delhi, p286.

Mallory AC, Bartel DP, Bartel B (2005) MicroRNA-directed regulation of *Arabidopsis* auxin response factor 17 is essential for proper development and modulates expression of early auxin response genes. *Plant Cell* 17: 1360-1375.

Mallory AC, Vaucheret H (2006) Functions of micro-RNAs and related small RNAs in plants. *Nat Genet* 38: S31-S36.

Mamay SH (1969) Cycads: fossil evidence of late paleozoic origin. *Sci* 164: 295-296.

Mansour MMF, Al-Mutawa MM, Salama KHA, Abou Hadid AF (2002) Salt acclimation of wheat salt sensitive cultivar by polyamines. In Ahmad R, Malik KA (ed) prospect for saline agriculture, Kluwer Academic Publishers, Dordrecht.

Maroco JP, Pereira JS, Chaves MM (1997) Stomatal responses to leaf-to-air vapour pressure deficit in sahelian species. *Aust J Plant Physiol* 24: 381-387.

Maroco JP, Rodrigues ML, Lopes C, Chaves MM (2002) Limitations to leaf photosynthesis in field-grown grapevine under drought-metabolic and modelling approaches. *Funct Plant Biol* 29: 451-459.

Matranga C, Zamore PD (2007) Small silencing RNAs. *Curr Biol* 17: R789-R793.

Matsumoto S, Takeuchi A, Hayatsu M, Kondo S (1994) Molecular cloning of phenylalanine ammonia-lyase cDNA and classification of varieties and cultivars of tea plants (*Camellia sinensis*) using the tea PAL cDNA probes. *Theor Appl Genet* 89: 671-675.

Matta Da FM, Maestri M, Barros RS (1997) Photosynthetic performance of two coffee species under drought. *Photosynthetica* 34: 257-264.

Matteo C, Daniela TM, Valentina S (2010) Microsatellite-based genetic relationships in the genus *Camellia*: potential for improving cultivars. *Genome* 53: 384-399.

Matysik J, Alia, Bhalu B, Mohanty P (2002) Molecular mechanisms of quenching of reactive oxygen species by proline under stress in plants. *Curr Sci* 82: 525-532.

Maurel C (2007) Plant aquaporins: novel functions and regulation properties. *FEBS Lett* 581: 2227-2236.

McNeil SD, Nuccio ML, Hanson AD (1999) Betaines and related osmoprotectants: targets for metabolic engineering of stress resistance. *Plant Physiol* 120: 945-949.

Meyer AJ (2008) The integration of glutathione homeostasis and redox signaling. *J Plant Physiol* 165 1390-1403.

Miller GL (1972) Use of dinitrosalicylic acid reagent for determination of reducing sugars. *Anal Chem* 31: 426-428.

Miller-Hamilton JMT (1995) Antimicrobial Properties of Tea (*Camellia sinensis* L.) In: *Antimicrobial Agents and Chemotherapy*, Am Soc Microbiol 39, p2375-2377.

Mittler R (2002) Oxidative stress, antioxidants and stress tolerance. *Trends Plant Sci* 7: 405-410.

Miyama M, Shimizu H, Sugiyama M, Hanagata N (2006) Sequencing and analysis of 14,842 expressed sequence tags of burma mangrove, *Bruguiera gymnorrhiza*. *Plant Sci* 171: 234-241.

Mohammadkhani N, Heidari R (2007) Effects of drought stress on protective enzyme activities and lipid peroxidation in two maize cultivars. *Pak J Biol Sci* 10: 3835-3840.

Mohammadkhani N, Heidari R (2008a) Drought-induced accumulation of soluble sugars and proline in two maize varieties. *World App Sci J* 3: 448-453.

Mohammadkhani N, Heidari R (2008b) Effects of drought stress on soluble proteins in two maize varieties. *Turk J Biol* 32: 23-30.

- Mondal TK (2002) Detection of genetic diversity among the Indian tea (*Camellia sinensis*) germplasm by inter-simple sequence repeats (ISSR). *Euphytica* 128: 307-315.
- Mondal TK (2007) Tea, In Pua EC, Davey MR (ed) *Biotechnology in Agriculture and Forestry, Transgenic Crops V*, Berlin, Springer 60, p519-535.
- Mondal TK, Bhattacharya A, Ahuja PS, Chand PK (2001) Factor effecting *Agrobacterium tumefaciens* mediated transformation of tea (*Camellia sinensis* (L.) O. Kuntze). *Plant Cell Rep* 20: 712-720.
- Mondal TK, Bhattacharya A, Sood A, Ahuja PS (2002) Propagation of tea (*Camellia sinensis* (L.) O. Kuntze) by shoot proliferation of alginate-encapsulated axillary buds stored at 4°C. *Curr Sci* 83: 941-944.
- Moon J, Parry G, Estelle M (2004) The ubiquitin-proteasome pathway and plant development. *Plant Cell* 16: 3181-3195.
- Morin RD, Aksay G, Dolgosheina E, Ebhardt HA, Magrini V, Mardis ER, Sahinalp SC, Unrau PJ (2008) Comparative analysis of the small RNA transcriptomes of *Pinus contorta* and *Oryza sativa*. *Genome Res* 18: 571-584.
- Morinaga T, Fukushima E, Kano T, Maruyama Y, Yamasaki Y (1929) Chromosome numbers of cultivated plants II. *Bot Mag Tokyo* 43: 589-594.
- Mostajeran A, Rahimi-Eichi V (2009) Effects of drought stress on growth and yield of rice (*Oryza sativa* L.) cultivars and accumulation of proline and soluble sugars in sheath and blades of their different ages leaves. *American-Eurasian J Agric Environ Sci* 5: 264-272.
- Moussa HR, Abdel-Aziz SM (2008) Comparative response of drought tolerant and drought sensitive maize genotypes to water stress. *Aust J Crop Sci* 1: 31-361.

Moxon S, Jing R, Szittyá G, Schwach F, Rusholme Pilcher RL, Moulton V, Dalmay T (2008) Deep sequencing of tomato short RNAs identifies microRNAs targeting genes involved in fruit ripening. *Genome Res* 18: 1602-1609.

Mukhopadhyay A, Vij S, Tyagi AK (2004) Overexpression of a zinc-finger protein gene from rice confers tolerance to cold, dehydration, and salt stress in transgenic tobacco. *Proc Nat Acad Sci* 101: 6309-6314.

Murata J, Bienzleb D, Brandlec EJ, Sensend WC, Luca DV (2006) Expressed sequence tags from madagascar periwinkle (*Catharanthus roseus*). *FEBS Lett* 580: 4501-4507.

Murchison EP, Hannon GJ (2004) miRNAs on the move: miRNA biogenesis and the RNAi machinery. *Curr Opin Cell Biol* 16: 223-229.

Nagar PK, Kumar A (1996) Changes in endogenous abscisic acid and phenols during winter dormancy in tea (*Camellia sinensis* (L.) O. Kuntze). *Acta Physiol Plant* 18: 33-38.

Nagarajah S, Ratnasuriya GB (1981) Clonal variability in root growth and drought resistance in tea (*Camellia sinensis*). *Plant Soil* 60: 153-155.

Najaphy A, Khamssi NN, Mostafaie A, Mirzaee H (2010) Effect of progressive water deficit stress on proline accumulation and protein profiles of leaves in chickpea. *African J Biotechnol* 9: 7033-7036.

Nakamura Y, Gojobori T, Ikemura T (2000) Codon usage tabulated from the international DNA sequence databases: status for the year 2000. *Nucl Acids Res* 28: 292.

Nakano Y, Asada K (1981) Hydrogen peroxide is scavenged by ascorbate specific peroxidase in spinach chloroplasts. *Plant Cell Physiol* 22: 867-880.

Nakashima K, Kiyosue T, Yamaguchi-Shinozaki K, Shinozaki K (1997) A nuclear gene, *erd1*, encoding a chloroplast-targeted Clp protease regulatory subunit homolog is not

only induced by water stress but also developmentally up-regulated during senescence in *Arabidopsis thaliana*. Plant J 12: 851-861.

Nasaruddin NM, Harikrishna K, Othman RY, Hoon LS, Harikrishna JA (2007) Computational prediction of microRNAs from Oil Palm (*Elaeis guineensis* Jacq.) expressed sequence tags. AsPac J Mol Biol Biotechnol 15: 107-113.

Nautiyal PC, Ravindra V, Joshi YC (1995) Gas exchange and leaf water relations in two peanut cultivars of different drought tolerance. Biologia Planta 37: 371-374.

Navari-Izzo F, Milone MTA, Quartacci MF, Pinzino C (1993). Metabolic changes in wheat plants subjected to a water-deficit stress programme. Plant Sci 92: 151-157.

Nayyar H, Walia DP (2003) Water stress induced proline accumulation in contrasting wheat genotypes as affected by calcium and abscisic acid. Biologia Planta 46: 275-279.

Neuefeind T, Reinemer P, Bieseler B (1997) Plant glutathione-S-transferases and herbicide detoxification. Biol Chem 378: 199-205.

Newcomb DR, Crowhurst NR, Gleave PA, Rikkerink AHE, Allan CA, Beuning LL, Bowen HJ, Gera E, Jamieson RK, Janssen JB, Laing AW (2006) Analyses of expressed sequence tags from apple. Plant Physiol 141: 147-166.

Ngetich WK, Bore JKA (1998) Performance of clonal tea in rehabilitated and replanted fields. 2. effects of soil water deficits. Tea 19: 77-83.

Nicolas ME, Gleadow RM, Dalling MJ (1985) Effect of post-anthesis drought on cell division and starch accumulation in developing wheat grains. Annal Bot 55: 433-444.

Nixon DJ, Burgess PJ, Sanga BNK, Carr MKV (2001) A comparison of the responses of mature and young tea to drought. Exp Agric 37: 391-402.

Noctor G, Foyer CH (1998) Ascorbate and glutathione: keeping active oxygen under control. *Annu Rev Plant Physiol Plant Mol Biol* 49: 249-279.

Noctor G, Gomez L, Vanacker H, Foyer CH (2002) Interactions between biosynthesis, compartmentation and transport in the control of glutathione homeostasis and signalling. *J Exp Bot* 53: 1283-1304.

Ogawa T, Pan L, Kawai- Yamada M, Yu LH, Yamamura S, Koyama T, Kitajima S, Ohme-Takagi M, Sato F, Uchimiya H (2005) Functional analysis of *Arabidopsis* ethylene-responsive element binding protein conferring resistance to wax and abiotic stress-induced plant cell death. *Plant Physiol* 138: 1436-1445.

Ogbonnaya CI, Sari B, Brou C, Diouf O, Diop NN, Roy-Macauley H (2003) Selection of cowpea genotypes in hydroponics, pots, and field for drought tolerance. *Crop Sci* 43: 1114-1120.

Okayama H, Berg P (1982) High efficiency cloning of full-length cDNA. *Mol Cell Biol* 2:161-170.

Oser BL (1979) *Hawks physiological chemistry*, Mc Graw Hill, New York, USA.

Palatnik JF, Allen E, Wu X, Schommer C, Schwab R, Carrington JC, Weigel D (2003) Control of leaf morphogenesis by microRNAs. *Nat* 425: 257-263.

Panda SK, Khan LB, Khan MH (2003) Does aluminium phytotoxicity induce oxidative stress in greengram (*Vigna radiata*)?. *Bulg J Plant Physiol* 29: 77-86.

Pandey HC, Baig MJ, Chandra A, Bhatt RK (2010) Drought stress induced changes in lipid peroxidation and antioxidant system in genus *Avena*. *J Environ Biol* 31: 435-440.

Pang M, Woodward AW, Agarwal V, Guan X, Ha M, Ramachandran V, Chen X, Triplett BA, Stelly DM, Chen ZJ (2009) Genome-wide analysis reveals rapid and dynamic changes in miRNA and siRNA sequence and expression during ovule and fiber

development in allotetraploid cotton (*Gossypium hirsutum* L.). *Genome Biol* 10: R122 doi:10.1186/gb-2009-10-11-r122.

Pant BD, Musialak-Lange M, Nuc P, May P, Buhtz A, Kehr J, Walther D, Scheible WR (2009) Identification of nutrient-responsive *Arabidopsis* and rapeseed microRNAs by comprehensive real-time polymerase chain reaction profiling and small RNA sequencing. *Plant Physiol* 150: 1541-1555.

Park JS, Kim JB, Haha BS, Kim KH, Ha SH, Kim JB, Kim YH (2004) EST analysis of genes involved in secondary metabolism in *Camellia sinensis* (tea), using suppression subtractive hybridization. *Plant Sci* 166: 953-961.

Park W, Li J, Song R, Messing J, Chen X (2002) Carpel factory, a Dicer homolog, and HEN1, a novel protein, act in microRNA metabolism in *Arabidopsis thaliana*. *Curr Biol* 12:1484-1495.

Pastori GM, Trippi VS (1992) Oxidative stress induces high rate of glutathione reductase synthesis in a drought resistant maize strain. *Plant Cell Physiol* 33: 957-961.

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

Paul S, Wachira FN, Powell W, Waugh R (1997) Diversity and genetic differentiation among population of Indian and Kenyan tea (*Camellia sinensis* (L.) O. Kuntze) revealed by AFLP markers. *Theor Appl Genet* 94: 255-263.

Pei ZM, Murata Y, Benning G, Thomine S, Klusener B, Allen GJ, Grill E, Schroeder JI (2000) Calcium channels activated by hydrogen peroxide mediate abscisic acid signalling in guard cells. *Nat* 406: 731-734.

Pérez-Torres E, Paredes CM, Polanco V, Becerra BV (2009) Gene expression analysis: a way to study tolerance to abiotic stresses in crop species. *Chilean J Agric Res* 69: 260-269.

Pinheiro HA, DaMatta FM, Chaves ARM, Fontes EPB, Loureiro ME (2004) Drought tolerance in relation to protection against oxidative stress in clones of *Coffea canephora* subjected to long-term drought. *Plant Sci* 167: 1307-1314.

Pirdashti H, Sarvestani ZT, Bahmanyar MA (2009) Comparison of physiological responses among four contrast rice cultivars under drought stress conditions. *World Acad Sci Eng Tech* 49: 52-53.

Powell J (2000) SAGE: The serial analysis of gene expression. *Methods Mol Biol* 99: 297-319.

Prabu GR, Mandal AKA (2010) Computational identification of miRNAs and their target genes from expressed sequence tags of tea (*Camellia sinensis*). *Genomics Proteomics Bioinformatics* 8: 113-121.

Qian L PGLY (1999) Physiological responses of tea (*Camellia sinensis*) to different soil moisture regimes. *J Tea Sci* cnki:ISSN:0577-8921.0.1999-04-007.

Qiang L, Yong Z, Shouyi C (2000) Plant protein kinase genes induced by drought, high salt and cold stresses. *Chinese Sci Bullet* 45: 1153-1157.

Qing YS, Elisabeth W, Zheng AL, Hong XR, Xiao YH, Qing H, Fen FG, Su XX, Liang SW (2009) Functional annotation of expressed sequence tags as a tool to understand the molecular mechanism controlling flower bud development in tree peony. *Physiol Plant* 135: 436-449.

Qiu XC, Xie LF, Zhu YY, Guo K, Huang QS, Nie L, Yang MZ (2007) Computational identification of microRNAs and their targets in *Gossypium hirsutum* expressed sequence tags. *Gene* 395: 49-61.

Quan R, Shang M, Zhang H, Zhao Y, Zhang J (2004) Engineering of enhanced glycine betaine synthesis improves drought tolerance in maize. *Plant Biotechnol J* 2: 477-486.

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

Rabello AR, Guimarães CM, Range P HN, Silva FR, Seixas D, Souza E, Brasileiro A CM, Spehar CR, Ferreira ME, Mehta A (2008) Identification of drought-responsive genes in roots of upland rice (*Oryza sativa* L) *BMC Genomics* 9: 485 doi:10.1186/1471-2164-9-485.

Rajewsky N, Succi ND (2004) Computational identification of microRNA targets. *Dev Biol* 267: 529-535.

Ramagopal S (1987) Differential mRNA transcription during salinity in barley. *Proc Nat Acad Sci* 84: 94-98.

Rasmussen PS, Valdés J, Gamboa MC, Valenzuela P DT, Krauskopf E (2008) Generation and analysis of an *Eucalyptus globulus* cDNA library constructed from seedlings subjected to low temperature conditions. *Elect J Biotech* 11: DOI: 10.2225/vol11-issue2-fulltext-14

Ranieri A, Bernardi R, Lanese P, Soldatini GF (1989) Changes in free amino acid content and protein pattern of maize seedlings under water stress. *Environ Exp Botany* 29: 351-357.

Ranney TC, Bassul NL, Whitlow TH (1991) Osmotic adjustment and solute constituents in leaves and roots of water-stressed cherry (*Prunus*) trees. *J Amer Soc Hort Sci* 116: 684-688.

Ranty B, Aldon D, Galaud JP (2006) Plant calmodulins and calmodulin-related proteins. *Plant Sig Behav* 1: 96-104.

Rao ASV, Reddy AR (2008) Sulfur assimilation and abiotic stress in plants. In Khan NA, Singh S, Umar S (ed) Glutathione reductase: a putative redox regulatory system in plant cells, Springer-Verlag-Berlin, Heidelberg, p111-147.

Rathinasabapathi B, Sigua C, Ho J, Gage DA (2000) Osmoprotectant  $\beta$ -alanine betaine synthesis in the Plumbaginaceae: S-adenosyl-l-methionine dependent N-methylation of  $\beta$ -alanine to its betaine is via N-methyl and N,N-dimethyl  $\beta$ -alanines. *Physiol Plant* 109: 225-231.

Ratnayaka HH, Molin WT, Sterling TM (2003) Physiological and antioxidant responses of cotton and spurred anoda under interference and mild drought. *J Exp Bot* 54: 2293-2305.

Reddy PCO, Sairanganayakulu G, Thippeswamy M, Reddy PS, Reddy MK, Sudhakar C (2008) Identification of stress-induced genes from the drought tolerant semi-arid legume crop horsegram (*Macrotyloma uniflorum* (Lam.) Verdc.) through analysis of subtracted expressed sequence tags. *Plant Sci* 175: 372-384.

Reinhart BJ, Weinstein EG, Rhoades MW, Bartel B, Bartel DP (2002) MicroRNAs in plants. *Genes Dev* 16: 1616-1626.

Rhoades MW, Reinhart BJ, Lim LP, Burge CB, Bartel B, Bartel DP (2002) Prediction of plant microRNA targets. *Cell* 110: 513-520.

Ristic Z, Cass D (1991) Chloroplast structure after water shortage and high temperature in two lines of *Zea mays* L. that differ in drought resistance. *Bot Gaz* 152: 186-194.

Rodrigues FA, Laia MLD, Zingaretti SM (2009) Analysis of gene expression profiles under water stress in tolerant and sensitive sugarcane plants. *Plant Sci* 176: 286-302.

Rolland F, Moore B, Sheen J (2002) Sugar sensing and signaling in plants. *Plant Cell* 14: S185-S205.

Rossel JB, Walter PB, Hendrickson L, Chow WS, Poole A, Mullineaux PM, Pogson BJ (2006) A mutation affecting ascorbate peroxidase 2 gene expression reveals a link between responses to high light and drought tolerance. *Plant Cell Environ* 29: 269-281.

Saikia DN, Dey SK (1984) Leaf water potential and stomatal resistance of tea leaf as influenced by soil moisture and potash application. *Indian J Agric Sci* 54: 727-732.

Sairam RK, Deshmukh PS, Saxena DC (1998) Role of antioxidant systems in wheat genotypes tolerance to water stress. *Biol Planta* 41: 387-394.

Sairam RK, Srivastava GC, Agarwal S, Meena RC (2005) Differences in antioxidant activity in response to salinity stress in tolerant and susceptible wheat genotypes. *Biol Plant* 49: 85-91.

Sakamoto A, Murata N (2000) Genetic engineering of glycinebetaine synthesis in plants: current status and implications for enhancement of stress tolerance. *J Exp Bot* 51: 81-88.

Sakuma Y, Maruyama K, Osakabe Y, Qin F, Seki M, Shinozaki K, Yamaguchi-Shinozaki K (2006) Functional analysis of an *Arabidopsis* transcription factor, DREB2A, involved in drought responsive gene expression. *Plant Cell* 18: 1292-1309.

Salin ML (1988) Toxic oxygen species and protective systems of the chloroplast. *Physiol Plant* 72: 681-689.

Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular cloning-a laboratory manual* (2nd Ed), Cold Spring Harbor Laboratory Press, USA.

Sánchez FJ, Manzanares M, Andres de EF, Tenorio JL, Ayerbe L (1998) Turgor maintenance, osmotic adjustment and soluble sugar and proline accumulation in 49 pea cultivars in response to water stress. *Field Crops Res* 59: 225-235.

- Sandalio LM, Dalurzo HC, Gomez M, Romero-Puertas MC, Rio LA (2001) Cadmium-induced changes in the growth and oxidative metabolism of pea plants. *J Exp Bot* 364: 2115-2126.
- Sandanam S, Gee GW, Mapa RB (1980) Leaf water diffusion resistance in clonal tea (*Camellia sinensis* L.): effects of water stress, leaf age and clones. *Ann Bot* 47: 339-349.
- Savolainen V, Manen JF, Douzery E, Spichiger R (1994) Molecular phylogeny of families related to Celastrales based on rbcL 5' flanking sequences. *Phylogenet Evol* 3: 27-37.
- Sawahel W (2003) Improved performance of transgenic glycinebetaine-accumulating rice plants under drought stress. *Biologia planta* 47: 39-44.
- Schafleitner R, Gaudin A, Rosales ROG, Aliaga CAA, Bonierbale M (2007) Proline accumulation and real time PCR expression analysis of genes encoding enzymes of proline metabolism in relation to drought tolerance in Andean potato. *Acta Physiologiae Planta* 29: 19-26.
- Schwab R, Palatnik JF, Riester M, Schommer C, Schmid M, Weigel D (2005) Specific effects of microRNAs on the plant transcriptome. *Dev Cell* 8: 517-527.
- Scott DL, Clark C, Deahl LK, Prakash SC (1998) Isolation of functional RNA from periderm tissue of potato tubers and sweet potato storage roots. *Plant Mol Bio Rep* 16: 3-8.
- Sealy JR (1958) A revision of the genus *Camellia*. R Hortic Soc, London.
- Selote DS, Khanna-Chopra R (2010) Antioxidant response of wheat roots to drought acclimation. *Protoplasma* 245: 153-163.

Senthil K, Wasnik NG, Kim YJ, Yang DC (2010) Generation and analysis of expressed sequence tags from leaf and root of *Withania somnifera* (Ashwgandha). *Mol Biol Rep* 37: 893-902.

Shao HB, Chu LY, Lu ZH, Kang CM (2008) Primary antioxidant free radical scavenging and redox signaling pathways in higher plant cells. *Int J Biol Sci* 4: 8-14.

Sharma AD, Kaur P (2009) Combined effect of drought stress and heat shock on cyclophilin protein expression in *Triticum aestivum*. *General Appl Plant Physiol* 35: 88-92.

Sharma P, Dubey RS (2005) Drought induces oxidative stress and enhances the activities of antioxidant enzymes in growing rice seedlings. *Plant Grow Reg* 46: 209-221.

Sharma RK, Bhardwaj P, Negi R, Mohapatra T (2009) Identification, characterization and utilization of unigene derived microsatellite markers in tea (*Camellia sinensis* L.). *BMC Plant Biol* 9:53 doi:10.1186/1471-2229-9-53.

Shelton DA, Leach DN, Baverstock PR, Henry RJ (2001) Analysis of expressed sequence tags (ESTs) isolated from immature leaves of *Melaleuca alternifolia*. *Plant and Animal Genome IX Conference, San Diego, California, USA.*

Shinozaki K, Yamaguchi-Shinozaki K (1996) Molecular responses to drought and cold stress. *Curr Opin Biotech* 7: 161-167.

Shinozaki K, Yamaguchi-Shinozaki K (1997) Gene expression and signal transduction in water stress response. *Plant Physiol* 115: 327-334.

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

Shinozaki K, Yamaguchi-Shinozaki K, Seki M (2003) Regulatory network of gene expression in the drought and cold stress responses. *Curr Opin Plant Biol* 6: 410-417.

Shu QY, Wischnitzki E, Liu ZA, Ren HX, Han XY, Hao Q, Gao FF, Xu SX, Wang LS (2009) Functional annotation of expressed sequence tags as a tool to understand the molecular mechanism controlling flower bud development in tree peony. *Physiol Plant* 135: 436-449.

Silva GF, Iandolino A, Kayal AF, Bohlmann CM, Cushman AM, Lim H, Ergul A, Figueroa R et al. (2005) Characterizing the grape transcriptome. Analysis of expressed sequence tags from multiple *Vitis species* and development of a compendium of gene expression during berry development. *Plant Physiol* 139: 574-597.

Singer MA, Lindquist S (1998) Multiple effects of trehalose on protein folding in vitro and in vivo. *Mol Cell* 1: 639-648.

Singh ID (1980) Non-conventional approaches to the breeding of tea in north-east India. *Two and a bud* 27: 3-6.

Singh J, Nagaraju J (2008) *In silico* prediction and characterization of microRNAs from red flour beetle (*Tribolium castaneum*). *Insect Mol Biol* 17: 427-436.

Singh K, Paul A, Kumar S, Ahuja PS (2008) Cloning and differential expression of QM like protein homologue from tea [*Camellia sinensis* (L.) O. Kuntze]. *Mol Biol Rep* 36: 921-927.

Singh K, Rani A, Paul A, Dutt S, Joshi R, Gulati A, Ahuja PS, Kumar S (2009) Differential display mediated cloning of anthocyanidin reductase gene from tea (*Camellia sinensis*) and its relationship with the concentration of epicatechins. *Tree Physiol* 29: 837-846.

Singh S, Fapen S, D'Souza SF (2006) Cadmium accumulation and its influence on lipid peroxidation and antioxidative system in an aquatic plant *Bacopa monnieri* L. *Chemosphere* 62: 233-246.

Singla SL, Pareek A, Grover A (1997) High temperature, In Prasad MNV (ed) Plant Ecophysiology, John Wiley and Sons, p101-127.

Sircelj H, Tausz M, Grill D, Batič F (2005) Biochemical responses in leaves of two apple tree cultivars subjected to progressing drought. J Plant Physiol 162: 1308-1318.

Smalle J, Vierstra RD (2004) The ubiquitin 26S proteasome proteolytic pathway. Annu Rev Plant Physiol Plant Mol Biol 55: 555-590.

Smirnoff N (1993) The role of active oxygen in the response of plants to water deficit and desiccation. New Phytol 125: 27-58.

Smirnoff N, Wheeler GL (2000) Ascorbic acid in plants: biosynthesis and function. Crit Rev Biochem Mol Biol 35: 291-314.

Song C, Fang J, Li X, Liu H, Thomas Chao C (2009) Identification and characterization of 27 conserved microRNAs in citrus. Planta 230: 671-685.

Song C, Wang C, Zhang C, Korir NK, Yu H, Ma Z, Fang J (2010) Deep sequencing discovery of novel and conserved microRNAs in trifoliolate orange (*Citrus trifoliata*). BMC Genomics 11: 431 doi:10.1186/1471-2164-11-431.

Souza de AA, Takita MA, Coletta-Filho HD, Campos MA, Teixeira JEC, Targon MLPN, Carlos EF, Ravasi JF, Fischer CN, Machado MA (2007) Comparative analysis of differentially expressed sequence tags of sweet orange and mandarin infected with *Xylella fastidiosa*. Genet Mol Biol 30: 965-971.

Squire GR (1978) Stomatal behavior of tea (*Camellia sinensis*) in relation to environment. J Appl Ecol 15: 287-301.

Srivalli B, Sharma G, Khanna-Chopra R (2003) Antioxidative defense system in an upland rice cultivar subjected to increasing intensity of water stress followed by recovery. Physiol Plant 119: 503-512.

- Stephens W, Carr MKV (1991) Responses of *Camellia sinensis* to irrigation and fertilizer IV. water use. *Exp Agric* 27: 192-210.
- Stoyanova D, Yordanov I (1999) Influence of drought, high temperature, and carbamide cytokinin 4-PU-30 on photosynthetic activity of plants. 2. chloroplast ultrastructure of primary bean leaves. *Photosynthetica* 37: 621-625.
- Subramanian S, Fu Y, Sunkar R, Barbazuk WB, Zhu JK, Yu O (2008) Novel and nodulation-regulated microRNAs in soybean roots. *BMC Genomics* 9: 160 doi:10.1186/1471-2164-9-160.
- Sunkar R, Chinnusamy V, Zhu J, Zhu JK (2007) Small RNAs as big players in plant abiotic stress responses and nutrient deprivation. *Trends Plant Sci* 12: 301-309.
- Sunkar R, Girke T, Jain PK, Zhu JK (2005) Cloning and characterization of microRNAs from rice. *Plant cell* 17: 1397-1411.
- Sunkar R, Jagadeeswaran G (2008) *In silico* identification of conserved microRNAs in large number of diverse plant species. *BMC Plant Biol* 8: 37 10.1186/1471-2229-8-37.
- Sunkar R, Zhou X, Zheng Y, Zhang W, Zhu JK (2008) Identification of novel and candidate miRNAs in rice by high throughput sequencing. *BMC Plant Biol* 8: 25 doi:10.1186/1471-2229-8-25.
- Sunkar R, Zhu JK (2004) Novel and stress-regulated micro-RNAs and other small RNAs from *Arabidopsis*. *Plant Cell* 16: 2001-2019.
- Sze H, Ward JM, Lai S (1992) Vacuolar H(+)-translocating ATPases from plants: structure, function, and isoforms. *J Bioenerg Biomembr* 24: 371-81.
- Szekely G (2004) The role of proline in *Arabidopsis thaliana* osmotic stress response. *Acta Biologica Sze* 48: 81.

Szittyá G, Moxon S, Santos DM, Jing R, Fevereço MP, Moulton V, Dalmay T (2008) High-throughput sequencing of *Medicago truncatula* short RNAs identifies eight new miRNA families. *BMC Genomics* 9: 593 doi:10.1186/1471-2164-9-593.

Takeuchi A, Matsumoto S, Hayatsu M (1994) Chalcone synthase from *Camellia sinensis*: Isolation of the cDNAs and the organ-specific and sugar-responsive expression of the genes. *Plant Cell Physiol* 35: 1011-1018.

Talame V, Ozturk NZ, Bohnert HJ, Tuberosa R (2007) Barley transcript profiles under dehydration shock and drought stress treatments: a comparative analysis. *J Exp Bot* 58: 229-240.

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol Biol Evol* 24: 1596-1599.

Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol Biol Evol* 10: 512-526.

Thakur PS, Rai VK (1982) Dynamics of amino acid accumulation of two differentially drought resistant *Zea mays* cultivars in response to osmotic stress. *Environ Exp Bot* 22: 221-226.

Thompson DJ, Higgins GD, Gibson JT (1994) Clustal W: improving the sensitivity of progressive multiple sequence alignments through sequence weighting, position specific gap penalties and weight matrix choice. *Nucl Acids Res* 22: 4673-4680.

Tian Li, Peel GJ, Lei Z, Aziz N, Dai X, He J, Watson B, Zhao PX, Sumner LW, Dixon RA (2009) Transcript and proteomic analysis of developing white lupin (*Lupinus albus* L.) roots. *BMC Plant Biol* 9:1. doi: 10.1186/1471-2229-9-1.

Torres-Franklin ML, Contour-Ansel D, Zuily-Fodil Y, Pham-Thi AT (2008) Molecular cloning of glutathione reductase cDNAs and analysis of GR gene expression in cowpea

and common bean leaves during recovery from moderate drought stress. *J Plant Physiol* 165: 514-521.

Tran LSP, Nakashima K, Sakuma Y, Simpson SD, Fujita Y, Maruyama K, Fujita M, Seki M, Shinozaki K, Yamaguchi-Shinozaki K (2004) Isolation and functional analysis of *Arabidopsis* stress-inducible NAC transcription factors that bind to a drought-responsive cis-element in the early responsive to dehydration stress 1 promoter. *Plant Cell* 16: 2481-2498.

Truchet G, Roche P, Lerouge P, Vasse JI Camut S, de Billy F, Promé JC, Dénarié J (1991) Sulfated lipo-oligosaccharide signals of *Rhizobium meliloti* elicits root nodule organogenesis in alfalfa. *Nat* 351: 670-673.

Turkan I, Bor M, Ozdemir F, Koca H (2005) Differential responses of lipid peroxidation and antioxidants in the leaves of drought-tolerant *P. acutifolius* gray and drought-sensitive *P. vulgaris* L. subjected to polyethylene glycol mediated water stress. *Plant Sci* 168: 223-231.

Tuskan GA, DiFazio S, Jansson S, Bohlmann J, Grigoriev I (2006) The genome of black cottonwood, *Populus trichocarpa* (Torr & Gray ex. Brayshaw). *Sci* 313:1596-1604.

Ünyayar S, Keleş Y, Çekiç FO (2005) The antioxidative response of two tomato species with different drought tolerances as a result of drought and cadmium stress combinations. *Plant Soil Environ* 51: 57-64.

Ujino-Ihara T, Kanamori H, Yamane H, Taguchi Y, Namiki N, Mukai Y, Yoshimura K, Tsumura Y (2005) Comparative analysis of expressed sequence tags of conifers and angiosperms reveals sequences specifically conserved in conifers. *Plant Mol Biol* 59: 895-907.

Uno Y, Furihata T, Abe H, Yoshida R, Shinozaki K, Yamaguchi-Shinozaki K (2000) *Arabidopsis* basic leucine zipper transcriptional transcription factors involved in an abscisic acid-dependent signal transduction pathway under drought and high-salinity conditions. *Proc Nat Acad Sci* 97: 11632-11637.

- Unver T, Namuth-Covert DM, Budak H (2009) Review of current methodological approaches for characterizing microRNAs in Plants. *Int J Plant Genomics* doi: 10.1155/2009/262463.
- Upadhyaya H, Panda SK (2004) Responses of *Camellia sinensis* to drought and rehydration. *Biol Planta* 48: 597-600.
- Upadhyaya H, Panda SK, Dutta BK (2008) Variation of physiological and antioxidative responses in tea cultivars subjected to elevated water stress followed by rehydration recovery. *Acta Physiol Planta* 30: 457-468.
- Urano K, Kurihara Y, Seki M, Shinozaki K (2010) 'Omics' analyses of regulatory networks in plant abiotic stress responses. *Curr Opinion Plant Biol* 13: 132-138.
- Utrillas MJ, Alegre L (1997) Impact of water stress on leaf anatomy and ultrastructure in *Cynodon dactylon* (L.) pers. under natural conditions. *Int J Plant Sci* 158: 313-324.
- Vajrabhaya M, Kumpun W, Chadchawan S (2001) The solute accumulation: the mechanism for drought tolerance in RD23 Rice (*Oryza sativa* L) lines. *SciAsia* 27: 93-97.
- Valentovič P, Luxová M, Kolarovič L, Gašparíková O (2006) Plant effect of osmotic stress on compatible solutes content, membrane stability and water relations in two maize cultivars. *Soil Environ* 52: 186-191.
- Vandez VBP, Sharma KK (2008) Transgenic approaches for abiotic stress tolerance in plants: retrospect and prospects. *Plant Cell Rep* 27: 411-424.
- Vartanian N, Damerval C, Vienne D (1987) Drought-induced changes in protein patterns of *Brassica napus* var. *oleifera* roots. *Plant Physiol* 84: 989-992.
- Vecchietti A, Lazzari B, Ortugno C, Bianchi R, Malinverni R, Caprera A, Mignani I, Pozzi C (2009) Comparative analysis of expressed sequence tags from tissues in ripening stages of peach (*Prunus persica* L. Batsch). *Tree Genet Genom* 5: 377-391.

Vieira SJ (1976) Water stress, ultrastructure and enzymatic activity. In Lange OL, Kappen L, Schulze E-D (ed) Water and plant life, Springer-Verlag, New York, pp 207-224.

Vierling E (1991) Roles of heat shock proteins in plants. *Annu Rev Plant Physiol Plant Mol Biol* 42: 579-620.

Vinocur B, Altman A (2005) Recent advances in engineering plant tolerance to abiotic stress: achievements and limitations. *Curr Opin Biotechnol* 16: 1-10.

Vos P, Hogers R, Bleeker M, Reijans M, Van de LT, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. *Nucl Acids Res* 23: 4407-4414.

Vyas D, Kumar S (2005) Tea (*Camellia sinensis* (L.) O. Kuntze) clone with lower period of winter dormancy exhibits lesser cellular damage in response to low temperature. *Plant Physiol Biochem* 43: 383-388.

Wachira FN, Powell W, Waugh R (1997) An assessment of genetic diversity among *Camellia sinensis* L. (cultivated tea) and its wild relatives based on randomly amplified polymorphic DNA and organelle-specific STS. *Heredity* 78: 603-611.

Waldren RP, Teare ID (1974) Free proline accumulation in drought-stressed plants under laboratory conditions. *Plant Soil* 40: 689-692.

Wang GP, Li F, Zhang J, Zhao MR, Hui Z, Wang W (2010a) Over accumulation of glycine betaine enhances tolerance of the photosynthetic apparatus to drought and heat stress in wheat. *Photosynthetica* 48: 30-41.

Wang L, Li X, Zhao Q, Jing S, Chen S, Yuan H (2008) Identification of genes induced in response to low-temperature treatment in tea leaves. *Plant Mol Biol Rep* 27: 257-265.

- Wang L, Wang MB, Tu JX, Helliwell CA, Waterhouse PM, Dennis ES, Fu TD, Fana YL (2007) Cloning and characterization of microRNAs from *Brassica napus*. FEBS Lett 581: 3848-3856.
- Wang W, Vinocur B, Altman A (2003) Plant responses to drought, salinity and extreme temperatures: towards genetic engineering for stress tolerance. Planta 218: 1-14.
- Wang W, Vinocur B, Shoseyov O, Altman A (2004) Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. Trends Plant Sci 9: 244-252.
- Wang Y, Li P, Cao X, Wang X, Zhang A, Li X (2009) Identification and expression analysis of miRNAs from nitrogen-fixing soybean nodules. Biochem Biophys Res Commun 378: 799-803.
- Wang YQ, Shen JK, Berglund T, Ohlsson AB, Tang XF, Zhou ZK, Wu RY, Zhou XH, Chen JN (2010b) Analysis of expressed sequence tags from ginkgo mature foliage in China. Tree Genet Genom 6: 357-365.
- Wang Z, Stutte (1992) The role of carbohydrates in active osmotic adjustment in apple under water stresses. J Amer Soc Hort Sci 117: 816-823.
- Way H, Chapman S, McIntyre L, Casu R, Xue GP, Manners J (2005) Shorter R: Identification of differentially expressed genes in wheat undergoing gradual water deficit stress using a subtractive hybridisation approach. Plant Sci 168: 661-670.
- Wei B, Cai T, Zhang R, Li A, Huo N, Li S, Gu YQ, Vogel J, Jia J, Qi Y et al. (2009) Novel microRNAs uncovered by deep sequencing of small RNA transcriptomes in bread wheat (*Triticum aestivum* L.) and *Brachypodium distachyon* (L.) Beauv. Funct Integr Genomics 9: 532-541.
- Welcher AA, Torres AR, Ward DC (1986) Selective enrichment of specific DNA, cDNA and RNA sequences using biotinylated probes, avidin and copper-chelate agarose. Nucl Acids Res 14:10027-10044.

Wijeratne MA (1992) Vulnerability of Sri Lanka tea production to global climate change. *Water Air Soil Pollution* 92: 87-94.

Wijeratne MA (1994) Effect of climatic factors on the growth of tea (*Camellia sinensis* L.) in the low country wet zone of Sri Lanka, PhD Thesis, University of London.

Wijeratne MA, Fordhan R, Anandacumaraswamy A (1998) Water relations of clonal tea (*Camellia sinensis* L.) with reference to drought resistance: II. Effect of water stress. *Tropical Agric Res Exten* 1: 74-80.

Willekens H, Chamnongpol S, Davey M, Schraudner M, Langebartels C, Van Montagu M, Inzé D, Van Camp W (1997) Catalase is a sink for H<sub>2</sub>O<sub>2</sub> and is indispensable for stress defense in C3 plants. *EMBO J* 10: 1723-1732.

Wu J, Qu T, Chen S, Zhao Z, An L (2008) Molecular cloning and characterization of a  $\gamma$ -glutamylcysteine synthetase gene from *Chorisporea bungeana*. *Protoplasma* 235: 27-36.

Wu L, Zhang Q, Zhou H, Ni F, Wu X, Qi Y. (2009) Rice MicroRNA effector complexes and targets. *Plant Cell* 21: 3421-35.

Wu YS, Tang KX (2004) MAP kinase cascades responding to environmental stress in plants. *Acta Bot Sin* 46: 127-136.

Xie FL, Huang SQ, Guo K, Xiang AL, Zhu YY, Nie L, Yang ZM (2007) Computational identification of novel microRNAs and targets in *Brassica napus*. *FEBS Lett* 581, 1464-1474.

Xiong L, Lee MW, Qi M, Yang Y (2001) Identification of defense-related rice genes by suppression subtractive hybridization and differential screening. *Mol Plant Microbe Interact* 14: 685-692.

Xue LJ, Zhang JJ, Xue HW (2009) Characterization and expression profiles of miRNAs in rice seeds. *Nucl Acids Res* 37: 916-930.

Xuelan Hu Haibo YYZ (1987) Effect of the soil water content on the physiological process of tea plant. *J Tea Sci* doi.ISSN:1000-369X.0.1987-01-006.

Yamada K, Osakabe Y, Mizoi J, Nakashima K, Fujita Y, Shinozaki K, Yamaguchi-Shinozaki K (2010) Functional analysis of an *Arabidopsis thaliana* abiotic stress-inducible facilitated diffusion transporter for monosaccharides. *J Biol Chem* 285: 1138-1146.

Yamada T, Takatsu Y, Manabe T, Kasumi M, Marubashi W (2003) Suppressive effect of trehalose on apoptotic cell death leading to petal senescence in ethylene-insensitive flowers of gladiolus. *Plant Sci* 164: 213-221.

Yan J, Wang J, Li Q, Hwang JR, Patterson C, Zhang H (2003) AtCHIP, a U-box-containing E3 ubiquitin ligase, plays a critical role in temperature stress tolerance in *Arabidopsis*. *Plant Physiol* 132: 861-869.

Yancey PH (1994) Compatible and counteracting solutes. In Strange K (ed) *Cellular and molecular physiology of cell, volume Regulation*, CRC Press, Boca Raton, p81-109.

Yang WJ, Rich PJ, Axtell JD, Wood KV, Bonham CC, Ejeta G, Mickelbart MV, Rhodes D (2003) Genotypic variation for glycinebetaine in sorghum. *Crop Sci* 43: 162-169.

Yang Y, Kwon HB, Peng HP, Shih MC (1993) Stress responses and metabolic regulation of glyceraldehyde-3-phosphated ehydrogenase genes in *Arabidopsis*. *Plant Physiol* 101:209-216.

Yang Y, Zhuang X, Hu H (1987) Effect of soil water content on the physiological process of tea plant. *J Tea Sci* 7: 23-28.

Yao Y, Guo G, Ni Z, Sunkar R, Du J, Zhu JK, Sun Q (2007) Cloning and characterization of microRNAs from wheat (*Triticum aestivum* L.). *Genome Biol* 8: R96 doi: 10.1186/gb-2007-8-6-r96.

Yeo ET, Kwon HB, Han SE, Lee JT, Ryu JC, Byu MO (2000) Genetic engineering of drought resistant potato plants by introduction of the trehalose-6-phosphate synthase (TPS1) gene from *Saccharomyces cerevisiae*. *Mol Cells* 10: 263-268.

Zeng C, Wang W, Zheng Y, Chen X, Bo W, Song S, Zhang W, Peng M. (2010) Conservation and divergence of microRNAs and their functions in Euphorbiaceous plants. *Nucl Acids Res* 38: 981-985.

Zhan Z, Ke N, Chen B (1987) The cytology of tea clonal cultivars fujian shuixian and their infertile mechanism. *Proc Int Tea quality-Human Health symp, China*, p46.

Zhang B, Pan X, Anderson TA (2006c) Identification of 188 conserved maize microRNAs and their targets. *FEBS Lett* 580: 3753-3762.

Zhang B, Pan X, Cannon CH, Cobb GP, Anderson TA (2006b) Conservation and divergence of plant microRNA genes. *Plant J* 46: 243-259.

Zhang B, Pan X, Cobb PG, Anderson TA (2006a) Plant microRNA: a small regulatory molecule with big impact. *Dev Biol* 289: 3-16.

Zhang B, Pan X, Stellwag EJ (2008a) Identification of soybean microRNAs and their targets. *Planta* 229: 161-182.

Zhang B, Wang Q, Wang K, Pan X, Liu F, Guo T, Cobb GP, Anderson TA (2007) Identification of cotton microRNAs and their targets. *Gene* 397: 26-37.

Zhang BH, Pan XP, Wang QL, Cobb GP, Anderson TA (2005) Identification and characterization of new plant microRNAs using EST analysis. *Cell Res* 15: 336-360.

Zhang J, Zeng R, Chen J, Liu X, Liao Q (2008b) Identification of conserved microRNAs and their targets from *Solanum lycopersicum* Mill. *Gene* 423: 1-7.

Zhang L, Chia JM, Kumari S, Stein JC, Liu Z, Narechania A, Maher CA, Guill K, McMullen MD, Ware D (2009a) A genome-wide characterization of microRNA genes in maize. *PLoS Genet* 5: e1000716 doi:10.1371/journal.pgen.1000716.

Zhang W, Luo Y, Gong X, Zeng W, Li S (2009b) Computational identification of 48 potato microRNAs and their targets. *Comput Biol Chem* 33: 84-93.

Zhao CX, Shao HB, Chu LY (2008a) Aquaporin structure-function relationships: water flow through plant living cells. *Colloids Surf B Biointerfaces* 62: 163-172.

Zhao CZ, Xia H, Frazier TP, Yao YY, Bi YP, Li AQ, Li MJ, Li CS, Zhang BH, Wang XJ (2010) Deep sequencing identifies novel and conserved microRNAs in peanuts (*Arachis hypogaea* L.). *BMC Plant Biol* 10: 3 doi:10.1186/1471-2229-10-3.

Zhao LP, Gao QK, Chen L, Wang XC, Yao MZ (2006) Development and preliminary application of cDNA microarray of tea plant (*Camellia sinensis*). *J Tea Sci* 26: 166-170 (In Chinese).

Zhao LP, Liu Z, Chen L, Yao MZ, Wang XC (2007) Generation and characterization of 24 novel EST derived microsatellites from tea plant (*Camellia sinensis*) and cross-species amplification in its closely related species and varieties. *Conserv Genet* 9: 1327-1331.

Zhao LP, Ma CL, Chen L (2008b) Construction and expressed sequence tags (ESTs) analysis of young roots cDNA library of tea plant [*Camellia sinensis* (L.) O. Kuntze]. *Mol Plant Breed* 6: 893-898 (In Chinese).

Zheng N, Schulman BA, Song L, Miller JJ, Jeffrey PD, Wang P, Chu C, Koepp DM et al. (2002) Structure of the Cul1-Rbx1-Skp1-F boxSkp2 SCF ubiquitin ligase complex. *Nat* 416: 703-709.

Zhou GA, Chang RZ, Qiu LJ (2010) Overexpression 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: 357-367.

Zhou X, Ruan J, Wang G and Zhang W (2007) Characterization and Identification of microRNA core promoters in four model species. *PLoS comput Biol* 3: e37 doi: 10.1371/journal.pcbi.0030037.

Zhou ZS, Huang SQ, Yang ZM (2008) Bioinformatic identification and expression analysis of new microRNAs from *Medicago truncatula*. *Biochem Biophys Res Commun* 374: 538-542.

Zhu JK (2002) Salt and drought stress signal transduction in plants. *Annu Rev Plant Biol* 53: 247-273.

Zhu JK (2008) Reconstituting plant miRNA biogenesis. *Proc Nat Acad Sci* 105: 9851-9852.

Zhu QH, Spriggs A, Matthew L, Fan L, Kennedy G, Gubler F, Helliwell C (2008) A diverse set of microRNAs and microRNA-like small RNAs in developing rice grains. *Genome Res* 18: 1456-1465.

Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucl Acids Res* 31: 3406-3415.

## APPENDIX-A

### CHEMICALS, REAGENTS AND MEDIA

## CHEMICALS AND PREPARATION OF STOCK SOLUTIONS, REAGENTS & MEDIA

### A1. CHEMICALS AND REAGENTS

Chemicals for buffer preparation and organic solvents were purchased from Indian manufacturers' viz., Qualigens Fine Chemical Company (GlaxoSmithkline) and Ranbaxy Laboratories Ltd., Sisco Research Laboratories, Pvt. Ltd. and E. Merck (India) Ltd. Molecular biology grade chemicals and reagents as well as media were purchased from Sigma Chemical Company (USA), Bangalore Genei Pvt. Ltd. (India), HiMedia Laboratories (India), New England Biolabs (USA), Stratagene Corporation (US), Clontech Laboratories (USA) and Invitrogen (US). All plastic wares were purchased from Imperial Biomedical Pvt. Ltd. and Tarsons Products Pvt. Ltd. Glasswares were purchased from Riviera Glass Pvt. Ltd. and Borosil Glass Pvt. Ltd., India.

### A2. PREPARATION OF STOCK SOLUTIONS AND MEDIA

#### 5 M NaCl

Solid NaCl (292.2 g) was dissolved in 800 ml of water by slightly warming and stirring with a magnetic stirrer to dissolve the salt completely. Total volume was made up to 1000 ml with distilled water and sterilized by autoclaving.

#### 3 M Sodium acetate (pH 5.2)

Sodium acetate (408.1 g) was dissolved in 800 ml of water and adjusted the pH 5.2 with glacial acetic acid. Then, the final volume was made up to 1000 ml with distilled water and sterilized by autoclaving.

#### 0.5 mM EDTA (pH 8.0)

Disodium EDTA-dihydrate (186.1 g) was added to 800 ml of distilled water and mixed vigorously on a magnetic stirrer. The pH was adjusted to 8.0 with NaOH (~ 20 g of NaOH pellets). Final volume was made to 1000 ml with distilled water.

### **0.5 M Tris-HCl**

Tris base 60.62 g was dissolved in 800 ml of water and pH was adjusted the pH 8.0 or 7.5 or 7.0 according to the need with concentrated HCl, made up the final vol to 1000 ml and sterilized by a 0.22  $\mu$ m filter.

### **1 M Glucose**

Solid glucose (18 g) was dissolved in 90 ml of distilled water. After dissolving, vol was adjusted to 100 ml with distilled water and sterilized by passing it through a 0.22  $\mu$ m filter.

### **10% SDS**

SDS powder (100 g) was mixed in 900 ml of distilled water. The solution was heated to 68°C and stir with a magnetic stirrer to assist dissolution. Finally, the vol was made to 1000 ml with distilled water and stored at room temperature.

### **100% TCA**

To a previously unopened bottle containing 500 g of trichloroacetic acid, 227 ml of distilled water was added.

### **0.2 M Sodium cacodylate buffer**

Solid sodium cacodylate 21.4 g was mixed with 3.45 ml of 1 N HCl and finally made up the vol to 500 ml with distilled water (pH 6.9) and sterilized by autoclaving.

### **5X TBE**

Tris base 54 g and boric acid 27.5 g were mixed in 800 ml of distilled water and dissolved by stirring on a magnetic stirrer. After that, 20 ml of 0.5 M EDTA (pH 8.0) was added to it and stirred again to mix well. The final volume was then made up to 1000 ml and autoclaved to sterilize before use.

## **50X TAE**

Tris base (242 g) was dissolved in 100 ml of 0.5 M EDTA (pH 8.0) by stirring on a magnetic stirrer and added 57.1 ml of glacial acetic acid. Final volume was made up to 1000 ml using sterile distilled water (pH 8.3) and autoclaved before use.

## **50X TE**

50 ml of 1 M Tris-HCl (pH 8.0) and 10 ml of 0.5 M EDTA (pH 8.0) were mixed and made the volume up to 100 ml. It was autoclaved to sterilize and stored at 4°C. The pH of Tris-HCl determines the pH of the TE buffer.

## **2.0% Agarose gel**

Agarose powder (2.0 g) was taken in a conical flask and 0.5x TBE/1X TAE buffer was added to make up the final volume to 100 ml. The mixture was then properly boiled in a microwave oven and poured on a gel casting plate when its temperature is about 55-60°C.

## **1.0% Agarose gel**

Agarose powder (1.0 g) was taken in a conical flask and 0.5X TBE/1X TAE buffer was added to make the final volume to 100 ml. The mixture was then properly boiled in a microwave oven and poured on a gel casting plate when its temperature is about 55-60°C.

## **PCI (25:24:1, v/v/v)**

Crystal phenol was dissolved at 68°C and subsequently mixed with chloroform and isoamyl alcohol at a proportion of 25:24:1. In order to prepare 1000 ml of PCI, 500 ml of phenol, 480 ml of chloroform and 20 ml of isoamyl alcohol were mixed and stored below 15°C.

### **CI (24:1, v/v)**

Chloroform and isoamyl alcohol was mixed at a proportion of 24:1. In order to prepare 1000 ml CIA, 960 ml chloroform and 40 ml of iso-amyl alcohol were mixed and stored below 15°C.

### **10X DNA loading dye**

Bromophenol blue and xylene cyanol FF, each of 40 mg and ficoll (type 400) 2.5 g were weighed and added approximately 8 ml of water. The mixture was stirred for several h on a magnetic stirrer till all powders get dissolved. Final volume was made up to 10 ml by adding distilled water and autoclaved to sterilize. (Note: 30% glycerol can be used alternatively to 2.5 g of ficoll or 0.25% each of bromophenol blue and xylene cyanol in 30% glycerol can be used to prepare 10 ml of 6x loading dye)

### **10 mg per ml ethidium bromide**

To prepare 25 ml of 10 mg/ml ethidium bromide, 250 mg of ethidium bromide powder was dissolved in 20 ml water and stirred on a magnetic stirrer for several hours to dissolve the dye completely. The volume was made up to 25 ml and the container was wrapped with aluminium foil and kept in a dark bottle.

### **LB broth**

LB powder (25 g) was mixed in 800 ml of distilled water by stirring on a magnetic stirrer and finally adjusted the vol to 1000 ml (pH 7) and sterilized by autoclaving for 15 min.

### **LB agar (1.5%)**

In 1000 ml of LB broth, 15 g of agar powder was added and mixed by stirring on a magnetic stirrer. The agar was dissolved by boiling the broth and then sterilized by autoclaving.

### **LB-MgSO<sub>4</sub> soft top agar**

In 1000 ml of LB broth, 10 ml of 1 M MgSO<sub>4</sub> (10 mM final concentration) and 7.2 g of agar powder were added, autoclaved and stored at 4°C.

### **SM buffer**

Sodium chloride (5.8 g) and magnesium sulphate (2 g) were mixed in 50 ml of 1 M Tris-Cl buffer (pH 7.5) and 5 ml of 2% gelatin solution. Final volume was made upto 1000 ml with distilled water and sterilized by autoclaving for 20 min at 15 psi. A 2% gelatin solution was made by adding 2 g of gelatin to a total vol of 100 ml distilled water and sterilized by autoclaving for 15 min.

### **SOC Medium**

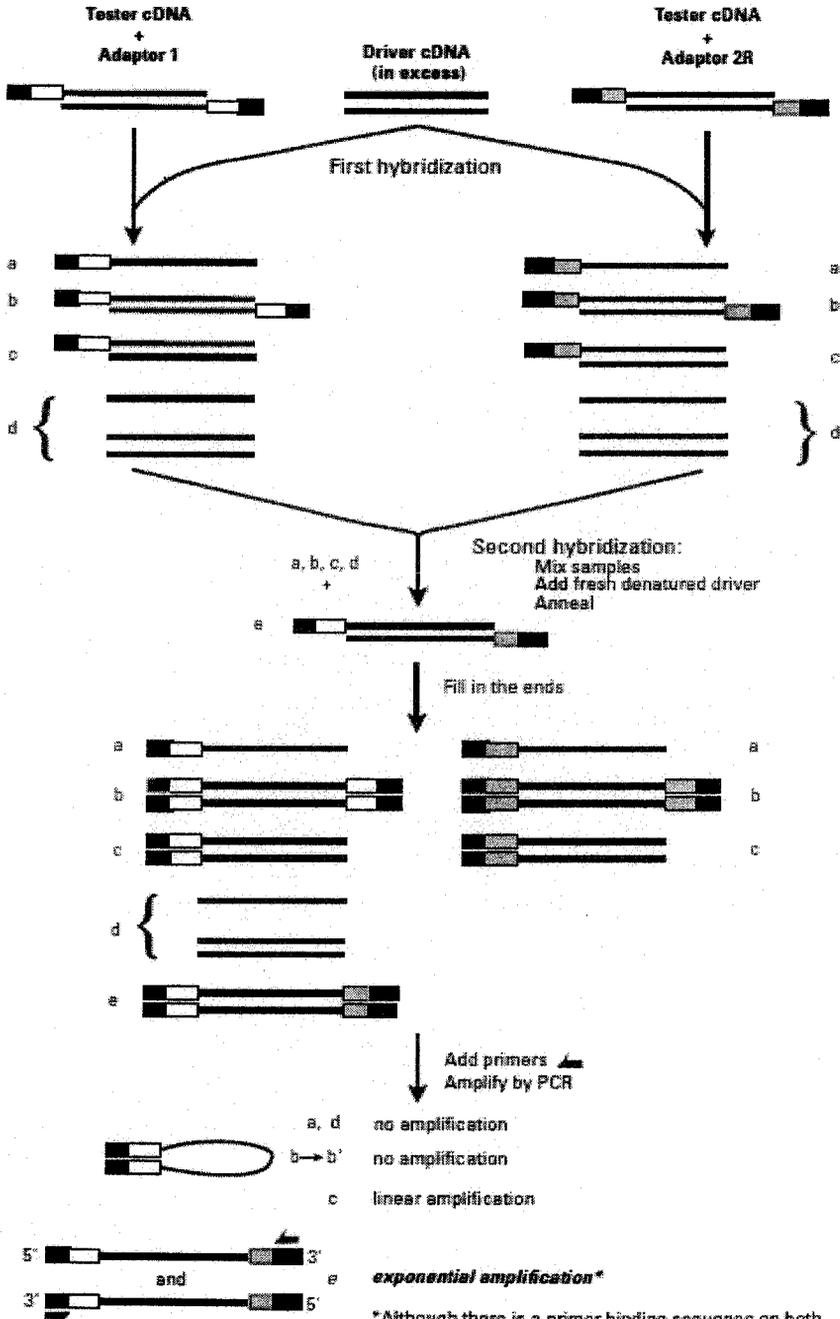
Tryptone (20 g), yeast extracts (5 g) and NaCl (0.5 g) were mixed in 950 ml of distilled water and dissolved by stirring on a magnetic stirrer. Then, 10 ml of a 250 mM solution of KCl was added and adjusted the pH to 7.0 with 5 N NaOH (~ 0.2 ml). Finally, the vol was made up to 1000 ml with distilled water, sterilized by autoclaving for 20 min and allowed to cool to 60°C or less. Then, 20 ml of a sterile 1 M solution of glucose was added. 250 mM KCl solution was made by dissolving 1.86 g of KCl in 100 ml of distilled water.

## APPENDIX-B

# FLOWCHART OF SSH AND SMART TECHNIQUES

# B 1. FLOWCHART OF SSH LIBRARY CONSTRUCTION TECHNIQUE

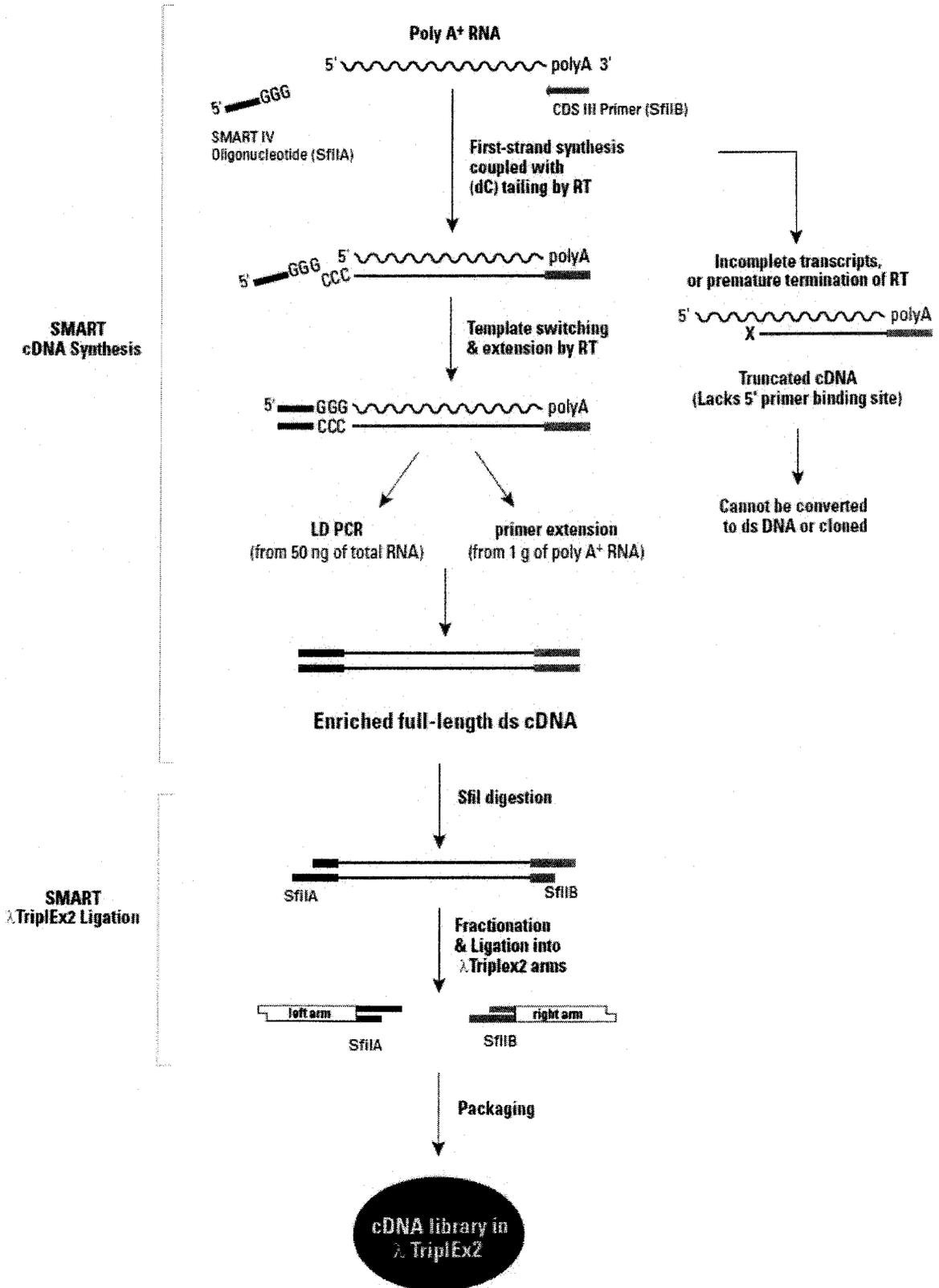
(SMART cDNA library construction kit user manual, Clontech, USA)



\*Although there is a primer binding sequence on both ends of the type e molecules, the shorter overall homology at the two ends effectively negates the suppression PCR effect—except for very short molecules. See Appendix A for more details on suppression PCR.

## B 2. FLOWCHART OF SMART LIBRARY CONSTRUCTION TECHNIQUE

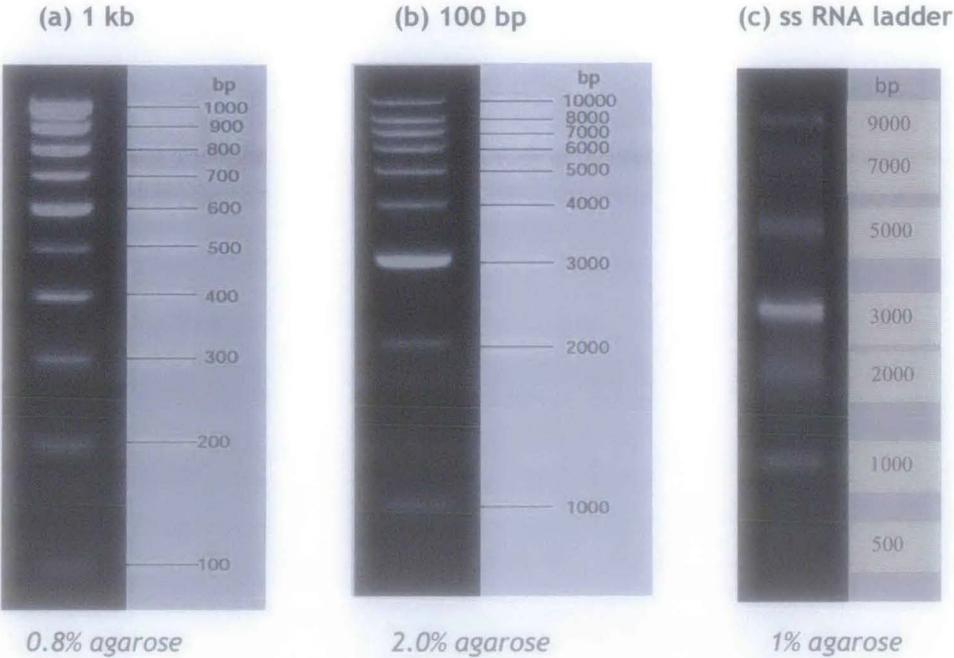
PCR-Select cDNA subtraction kit user manual, Clontech, USA



## APPENDIX-C

# STANDARD MOLECULAR WEIGHT MARKERS

C. STANDARD MOLECULAR WEIGHT MARKERS USED IN THE PRESENT STUDY



## APPENDIX-D

# BIOINFORMATICS ANALYSES OF ESTs

D 1: CODON USAGE ANALYSES OF UNIGENES OF TEA ROOTS AND LEAVES

D 1.1: CODON USAGE ANALYSIS OF UNIGENES OF DROUGHT STRESSED ROOTS

cds count : 1915  
 Coding GC : 44.38%  
 1st letter GC : 44.64%  
 2nd letter GC : 44.47%  
 3rd letter GC : 44.03%

Codon AA Fraction Frequency Number

Codon AA Fraction Frequency Number

GCA	A	0.299	15.153	638
GCC	A	0.210	10.616	447
GCG	A	0.152	7.671	323
GCT	A	0.340	17.195	724
TGC	C	0.465	15.723	662
TGT	C	0.535	18.098	762
GAC	D	0.360	11.139	469
GAT	D	0.640	19.808	834
GAA	E	0.523	25.365	1068
GAG	E	0.477	23.109	973
TTC	F	0.396	18.739	789
TTT	F	0.604	28.524	1201
GGA	G	0.339	20.473	862
GGC	G	0.187	11.305	476
GGG	G	0.234	14.155	596
GGT	G	0.240	14.488	610
CAC	H	0.409	11.828	498
CAT	H	0.591	17.076	719
ATA	I	0.253	12.968	546
ATC	I	0.315	16.103	678
ATT	I	0.432	22.111	931
AAA	K	0.546	30.163	1270
AAG	K	0.454	25.104	1057
CTA	L	0.092	8.574	361
CTC	L	0.148	13.846	583
CTG	L	0.170	15.913	670
CTT	L	0.201	18.858	794
TTA	L	0.135	12.683	534
TTG	L	0.254	23.821	1003
ATG	M	1.000	19.831	835

AAC	N	0.429	14.868	626
AAT	N	0.571	19.760	832
CCA	P	0.329	14.963	630
CCC	P	0.211	9.595	404
CCG	P	0.172	7.838	330
CCT	P	0.287	13.063	550
CAA	Q	0.581	23.394	985
CAG	Q	0.419	16.839	709
AGA	R	0.335	23.370	984
AGG	R	0.233	16.245	684
CGA	R	0.127	8.883	374
CGC	R	0.092	6.436	271
CGG	R	0.116	8.123	342
CGT	R	0.097	6.745	284
AGC	S	0.156	14.488	610
AGT	S	0.165	15.366	647
TCA	S	0.218	20.211	851
TCC	S	0.149	13.823	582
TCG	S	0.097	8.978	378
TCT	S	0.216	20.045	844
ACA	T	0.332	16.221	683
ACC	T	0.243	11.875	500
ACG	T	0.119	5.795	244
ACT	T	0.306	14.963	630
GTA	V	0.181	9.761	411
GTC	V	0.182	9.809	413
GTG	V	0.298	16.103	678
GTT	V	0.339	18.311	771
TGG	W	1.000	20.971	883
TAC	Y	0.353	9.619	405
TAT	Y	0.647	17.623	742
TAA	*	0.302	13.751	579
TAG	*	0.219	9.951	419
TGA	*	0.479	21.779	917

## D 1.2: CODON USAGE ANALYSIS OF STANDARD FULL-LENGTH UNIGENES OF ROOTS

					Codon AA Fraction Frequency Number				
cds count	:		1117		ATG	M	1.000	18.915	499
Coding GC	:		44.60%		AAC	N	0.414	15.162	400
1st letter GC	:		44.61%		AAT	N	0.586	21.493	567
2nd letter GC	:		44.30%		CCA	P	0.359	19.180	506
3rd letter GC	:		44.88%		CCC	P	0.241	12.888	340
					CCG	P	0.193	10.310	272
					CCT	P	0.207	11.069	292
					CAA	Q	0.658	24.184	638
					CAG	Q	0.342	12.547	331
					AGA	R	0.277	18.574	490
					AGG	R	0.207	13.874	366
					CGA	R	0.117	7.809	206
					CGC	R	0.116	7.771	205
					CGG	R	0.219	14.670	387
					CGT	R	0.064	4.321	114
					AGC	S	0.154	13.646	360
					AGT	S	0.147	13.040	344
					TCA	S	0.199	17.664	466
					TCC	S	0.151	13.381	353
					TCG	S	0.117	10.386	274
					TCT	S	0.232	20.621	544
					ACA	T	0.303	14.670	387
					ACC	T	0.283	13.722	362
					ACG	T	0.148	7.164	189
					ACT	T	0.265	12.850	339
					GTA	V	0.194	10.310	272
					GTC	V	0.187	9.893	261
					GTG	V	0.296	15.693	414
					GTT	V	0.323	17.134	452
					TGG	W	1.000	21.227	560
					TAC	Y	0.444	11.637	307
					TAT	Y	0.556	14.594	385
					TAA	*	0.354	14.973	395
					TAG	*	0.226	9.552	252
					TGA	*	0.421	17.816	470
GCA	A	0.295	13.987	369					
GCC	A	0.260	12.319	325					
GCG	A	0.167	7.884	208					
GCT	A	0.278	13.153	347					
TGC	C	0.358	12.547	331					
TGT	C	0.642	22.516	594					
GAC	D	0.379	9.818	259					
GAT	D	0.621	16.110	425					
GAA	E	0.603	22.327	589					
GAG	E	0.397	14.708	388					
TTC	F	0.392	21.986	580					
TTT	F	0.608	34.078	899					
GGA	G	0.277	17.702	467					
GGC	G	0.216	13.836	365					
GGG	G	0.273	17.437	460					
GGT	G	0.234	14.935	394					
CAC	H	0.409	13.002	343					
CAT	H	0.591	18.764	495					
ATA	I	0.280	15.504	409					
ATC	I	0.276	15.314	404					
ATT	I	0.444	24.639	650					
AAA	K	0.607	30.818	813					
AAG	K	0.393	19.976	527					
CTA	L	0.093	9.628	254					
CTC	L	0.165	17.134	452					
CTG	L	0.146	15.125	399					
CTT	L	0.197	20.431	539					
TTA	L	0.157	16.262	429					
TTG	L	0.244	25.321	668					

### D 1.3: CODON USAGE ANALYSIS OF UNIGENES OF WINTER DORMANCY STRESSED LEAVES

**cds count** : 2775  
**Coding GC** : 44.40%  
**1st letter GC** : 43.17%  
**2nd letter GC** : 45.07%  
**3rd letter GC** : 44.96%

#### Codon AA Fraction Frequency Number

CCA	P	0.353	16.058	985
CCC	P	0.232	10.531	646
CCG	P	0.162	7.385	453
CCT	P	0.253	11.477	704
CAA	Q	0.594	22.497	1380
CAG	Q	0.406	15.389	944
AGA	R	0.365	24.975	1532
AGG	R	0.269	18.389	1128
CGA	R	0.119	8.151	500
CGC	R	0.070	4.809	295
CGG	R	0.099	6.798	417
CGT	R	0.078	5.347	328
AGC	S	0.165	15.014	921
AGT	S	0.161	14.656	899
TCA	S	0.227	20.687	1269
TCC	S	0.152	13.840	849
TCG	S	0.089	8.151	500
TCT	S	0.207	18.845	1156
ACA	T	0.327	17.329	1063
ACC	T	0.263	13.938	855
ACG	T	0.127	6.716	412
ACT	T	0.283	14.998	920
GTA	V	0.155	8.298	509
GTC	V	0.178	9.569	587
GTG	V	0.331	17.769	1090
GTT	V	0.336	18.063	1108
TGG	W	1.000	23.899	1466
TAC	Y	0.337	8.135	499
TAT	Y	0.663	16.009	982
TAA	*	0.257	11.623	713
TAG	*	0.217	9.830	603
TGA	*	0.526	23.785	1459

#### Codon AA Fraction Frequency Number

GCA	A	0.317	14.656	899
GCC	A	0.213	9.863	605
GCG	A	0.117	5.412	332
GCT	A	0.353	16.318	1001
TGC	C	0.424	14.802	908
TGT	C	0.576	20.133	1235
GAC	D	0.355	10.254	629
GAT	D	0.645	18.633	1143
GAA	E	0.529	22.562	1384
GAG	E	0.471	20.052	1230
TTC	F	0.400	19.285	1183
TTT	F	0.600	28.871	1771
GGA	G	0.325	20.736	1272
GGC	G	0.191	12.178	747
GGG	G	0.237	15.112	927
GGT	G	0.246	15.699	963
CAC	H	0.432	11.574	710
CAT	H	0.568	15.210	933
ATA	I	0.260	13.025	799
ATC	I	0.305	15.308	939
ATT	I	0.435	21.828	1339
AAA	K	0.502	25.822	1584
AAG	K	0.498	25.643	1573
CTA	L	0.109	11.330	695
CTC	L	0.145	14.900	914
CTG	L	0.163	16.954	1040
CTT	L	0.174	18.144	1113
TTA	L	0.141	14.704	902
TTG	L	0.269	28.023	1719
ATG	M	1.000	24.160	1482
AAC	N	0.444	15.911	976
AAT	N	0.556	19.937	1223

**D 1.4: CODON USAGE ANALYSIS OF UNIGENES OF DROUGHT STRESSED LEAVES**

**cds count** : 798  
**Coding GC** : 37.91%  
**1st letter GC** : 37.44%  
**2nd letter GC** : 37.70%  
**3rd letter GC** : 38.59%

**Codon AA Fraction Frequency Number**

GCA	A	0.315	12.592	197
GCC	A	0.260	10.419	163
GCG	A	0.142	5.689	89
GCT	A	0.283	11.314	177
TGC	C	0.403	13.487	211
TGT	C	0.597	19.942	312
GAC	D	0.330	8.054	126
GAT	D	0.670	16.363	256
GAA	E	0.581	21.668	339
GAG	E	0.419	15.596	244
TTC	F	0.275	22.563	353
TTT	F	0.725	59.380	929
GGA	G	0.314	13.742	215
GGC	G	0.178	7.798	122
GGG	G	0.248	10.866	170
GGT	G	0.260	11.377	178
CAC	H	0.372	11.377	178
CAT	H	0.628	19.175	300
ATA	I	0.316	19.431	304
ATC	I	0.264	16.235	254
ATT	I	0.419	25.759	403
AAA	K	0.759	64.046	1002
AAG	K	0.241	20.326	318
CTA	L	0.115	10.994	172
CTC	L	0.138	13.231	207
CTG	L	0.128	12.208	191
CTT	L	0.186	17.769	278
TTA	L	0.180	17.258	270
TTG	L	0.254	24.289	380
ATG	M	1.000	21.285	333
AAC	N	0.382	16.491	258
AAT	N	0.618	26.654	417
CCA	P	0.361	14.957	234
CCC	P	0.207	8.565	134
CCG	P	0.160	6.647	104

**Codon AA Fraction Frequency Number**

CCT	P	0.272	11.250	176
CAA	Q	0.652	21.349	334
CAG	Q	0.348	11.377	178
AGA	R	0.348	17.194	269
AGG	R	0.226	11.186	175
CGA	R	0.107	5.305	83
CGC	R	0.098	4.858	76
CGG	R	0.132	6.520	102
CGT	R	0.088	4.346	68
AGC	S	0.134	11.122	174
AGT	S	0.148	12.208	191
TCA	S	0.249	20.582	322
TCC	S	0.160	13.231	207
TCG	S	0.072	5.944	93
TCT	S	0.237	19.623	307
ACA	T	0.388	18.217	285
ACC	T	0.216	10.163	159
ACG	T	0.095	4.474	70
ACT	T	0.301	14.126	221
GTA	V	0.210	10.291	161
GTC	V	0.166	8.118	127
GTG	V	0.267	13.103	205
GTT	V	0.357	17.514	274
TGG	W	1.000	16.363	256
TAC	Y	0.390	12.848	201
TAT	Y	0.610	20.134	315
TAA	*	0.326	16.619	260
TAG	*	0.226	11.505	180
TGA	*	0.449	22.883	358

**D 1.5: CODON USAGE ANALYSIS OF STANDARD FULL-LENGTH UNIGENES OF LEAVES**

**cds count** : 1117  
**Coding GC** : 44.60%  
**1st letter GC** : 44.61%  
**2nd letter GC** : 44.30%  
**3rd letter GC** : 44.88%

**Codon AA Fraction Frequency Number**

Codon	AA	Fraction	Frequency	Number
GCA	A	0.295	13.987	369
GCC	A	0.260	12.319	325
GCG	A	0.167	7.884	208
GCT	A	0.278	13.153	347
TGC	C	0.358	12.547	331
TGT	C	0.642	22.516	594
GAC	D	0.379	9.818	259
GAT	D	0.621	16.110	425
GAA	E	0.603	22.327	589
GAG	E	0.397	14.708	388
TTC	F	0.392	21.986	580
TTT	F	0.608	34.078	899
GGA	G	0.277	17.702	467
GGC	G	0.216	13.836	365
GGG	G	0.273	17.437	460
GGT	G	0.234	14.935	394
CAC	H	0.409	13.002	343
CAT	H	0.591	18.764	495
ATA	I	0.280	15.504	409
ATC	I	0.276	15.314	404
ATT	I	0.444	24.639	650
AAA	K	0.607	30.818	813
AAG	K	0.393	19.976	527
CTA	L	0.093	9.628	254
CTC	L	0.165	17.134	452
CTG	L	0.146	15.125	399
CTT	L	0.197	20.431	539
TTA	L	0.157	16.262	429
TTG	L	0.244	25.321	668

**Codon AA Fraction Frequency Number**

ATG	M	1.000	18.915	499
AAC	N	0.414	15.162	400
AAT	N	0.586	21.493	567
CCA	P	0.359	19.180	506
CCC	P	0.241	12.888	340
CCG	P	0.193	10.310	272
CCT	P	0.207	11.069	292
CAA	Q	0.658	24.184	638
CAG	Q	0.342	12.547	331
AGA	R	0.277	18.574	490
AGG	R	0.207	13.874	366
CGA	R	0.117	7.809	206
CGC	R	0.116	7.771	205
CGG	R	0.219	14.670	387
CGT	R	0.064	4.321	114
AGC	S	0.154	13.646	360
AGT	S	0.147	13.040	344
TCA	S	0.199	17.664	466
TCC	S	0.151	13.381	353
TCG	S	0.117	10.386	274
TCT	S	0.232	20.621	544
ACA	T	0.303	14.670	387
ACC	T	0.283	13.722	362
ACG	T	0.148	7.164	189
ACT	T	0.265	12.850	339
GTA	V	0.194	10.310	272
GTC	V	0.187	9.893	261
GTG	V	0.296	15.693	414
GTT	V	0.323	17.134	452
TGG	W	1.000	21.227	560
TAC	Y	0.444	11.637	307
TAT	Y	0.556	14.594	385
TAA	*	0.354	14.973	395
TAG	*	0.226	9.552	252
TGA	*	0.421	17.816	470

## D 2. ARABIDOPSIS LOCUS HITS OF UNIGENES OF TEA ROOTS AND LEAVES

### D 2.1: ARABIDOPSIS LOCUS HITS OF UNIGENES OF DROUGHT STRESSED ROOTS FOR GO ANALYSIS

AT1G07400.1, AT1G47128.1, AT2G29500.1, AT1G61850.1, AT5G11520.1, AT5G52640.1, AT2G40010.1, AT3G17330.1, AT2G41870.1, AT3G29320.1, AT5G15460.2, AT5G48570.1, AT5G05710.1, AT3G46230.1, AT2G29500.1, AT5G61170.1, AT4G35450.2, AT4G40040.2, AT3G51030.1, AT2G30860.1, AT3G43867.1, AT5G64200.1, AT2G01060.1, AT1G69670.1, AT5G02020.1, AT3G57520.3, AT5G10360.1, AT1G15080.1, AT3G05000.1, AT4G05320.2, AT1G23870.1, AT2G19760.1, AT1G21750.1, AT2G01880.1, ATCG00950.1, AT1G62380.1, ATCG00950.1, AT5G60790.1, AT5G05110.1, ATCG01210.1, AT4G08350.1, AT3G47300.1, AT2G46680.2, AT1G21630.1, ATCG00080.1, AT3G20300.1, ATMG00730.1, AT3G54900.1, AT4G27270.1, AT4G35335.1, AT1G22380.1, AT3G23570.1, AT1G04690.1, AT1G04760.1, AT2G47580.1, AT1G34220.2, AT5G58110.1, ATCG01210.1, AT3G14100.1, AT3G11910.1, AT3G62580.1, AT4G26220.1, AT3G11250.1, AT1G65970.1, AT1G21640.1, AT1G11260.1, AT4G38580.1, AT2G20330.1, AT1G74310.1, AT3G49500.1, AT3G55410.1, AT4G15560.1, AT1G51200.2, AT3G53570.3, AT1G66400.1, AT3G15020.1, AT5G19550.1, AT2G45380.1, AT5G13430.1, AT4G11260.1, AT2G28930.1, AT2G01670.1, AT1G09080.1, AT1G44542.1, AT3G45640.1, AT2G45640.1, AT5G05960.1, AT1G11840.2, AT4G33030.1, AT5G16550.1, AT3G04120.1, AT4G04410.1, AT3G56740.1, AT1G16030.1, AT2G37250.1, AT1G20140.1, AT1G77180.1, ATCG01210.1, AT1G15080.1, AT1G61520.2, AT3G05590.1, AT4G01280.2, AT1G14400.1, AT1G68000.1, AT3G09440.1, AT1G77450.1, AT3G62140.1, AT5G01650.1, AT3G53620.1, AT1G02140.1, AT5G58060.1, AT5G58040.1, AT3G10330.1, AT3G48890.1, AT1G67300.1, AT3G16640.1, ATCG01210.1, AT1G27350.1, AT1G53000.1, ATMG00020.1, AT2G19740.1, AT3G16780.1, AT1G79690.1, AT5G27320.1, AT1G09815.1, AT5G10000.1, AT5G36230.1, AT4G09320.1, AT3G20800.1, AT5G27490.1, AT5G48380.1, AT4G26670.1, AT1G12990.1, AT5G64400.1, AT1G74040.1, AT4G01070.1, AT3G03740.1, AT1G56423.1, ATCG00950.1, AT3G17000.1, AT4G34710.1, AT5G02960.1, AT5G03370.1, AT1G09920.1, AT3G24170.1, AT5G64170.1, AT3G12500.1, AT3G24800.1, AT5G08290.1, AT5G33400.1, AT3G15730.1, AT4G23100.1, AT1G07350.1, AT2G23620.1, AT3G50000.1, AT5G54800.1, AT3G50980.1, AT4G34720.1, AT5G62150.1, AT3G53980.1, AT1G15690.1, AT1G16350.1, AT5G08500.1, AT1G34370.3, AT2G34850.1, AT4G39680.2, AT3G14100.1, ATCG00950.1, AT3G16770.1, AT3G62290.1, AT1G05260.1, AT5G04620.2, AT5G55190.1, AT4G12280.1, AT2G15890.1, ATCG01210.1, AT1G02820.1, AT4G31985.1, AT2G17200.1, AT4G25200.1, AT5G64920.1, AT4G30960.1, AT1G08830.1, AT4G34720.1, AT2G17420.1, AT2G41475.1, AT5G42000.1

## D 2.2: ARABIDOPSIS LOCUS HITS OF STANDARD FULL-LENGTH UNIGENES OF ROOTS FOR GO ANALYSIS

AT2G40880.1, AT2G33830.2, AT2G27960.1, AT2G07747.1, AT3G09390.1, AT3G52590.1, AT1G31812.1, AT1G27350.1, AT3G26932.2,  
AT2G30110.1, AT2G36060.2, AT5G05960.1, AT5G02380.1, AT1G50940.1, AT3G46960.1, AT2G16600.1, AT2G01250.1, AT3G44110.1,  
AT3G24503.1, AT1G69410.1, AT3G57520.3, AT3G55350.1, AT5G02380.1, AT1G71950.1, AT3G46030.1, AT4G21320.1, AT5G20500.1,  
AT4G02890.4, AT4G14320.1, AT2G14170.2, AT4G31300.3, AT5G51950.2, AT5G62740.1, AT1G26770.2, AT4G18100.1, AT4G39200.1,  
AT4G26230.1, AT5G56010.1, AT3G12760.1, AT2G40470.1, AT3G22110.1, AT5G06480.1, AT1G02630.1, AT2G29500.1, AT3G11050.1,  
AT1G31050.1, AT1G02820.1, AT5G02560.1, AT5G61030.1, AT3G04230.1, AT4G33865.1, AT5G59970.1, AT3G04830.1, AT1G48430.1,  
AT5G41700.5, AT3G08690.2, AT3G45180.1, AT5G65780.1, AT1G15405.1, AT4G01897.1, AT4G05050.4, AT4G29350.1, AT5G01650.1,  
AT5G54660.1, AT3G16640.1, AT2G30490.1, AT5G50260.1, AT3G24100.1, AT5G06480.1, AT4G28600.1, AT4G03200.1, AT2G47650.1,  
AT4G14615.1, AT2G14170.2

## D 2.3: ARABIDOPSIS LOCUS HITS OF UNIGENES OF WINTER DORMANCY STRESSED LEAVES FOR GO ANALYSIS

AT4G28050.1, AT1G02820.1, AT5G01870.1, AT5G20010.1, AT1G01470.1, AT1G63770.1, AT1G28330.4, AT1G72040.1, AT3G09600.1,  
AT5G10400.1, AT1G70600.1, AT5G42760.1, AT1G09200.1, AT3G04120.1, AT3G56940.1, AT1G61720.1, AT2G33470.1, AT2G16850.1,  
AT5G16250.1, AT3G20670.1, AT1G30690.1, AT5G03240.3, AT3G27310.1, AT2G38750.1, AT3G46030.1, AT1G69770.1, AT2G34430.1,  
AT3G11050.1, AT5G10400.1, AT5G09360.1, AT2G22540.1, AT4G27700.1, AT4G21610.1, AT4G13540.1, AT4G35985.1, AT5G49400.1,  
AT4G39090.1, AT5G02780.2, AT3G21180.1, AT1G64760.1, AT1G64760.1, AT4G02570.3, AT3G20500.1, AT2G28900.1, AT4G13530.1,  
AT5G60360.1, AT4G38520.1, AT5G55700.1, AT3G45600.1, AT1G13950.1, AT1G54410.1, AT4G01870.1, AT3G48570.1, AT4G12280.1,  
AT2G16130.1, AT1G14340.1, AT1G15130.1, AT3G05540.1, AT3G05540.1, AT1G70850.3, AT1G23750.1

#### D 2.4: ARABIDOPSIS LOCUS HITS OF STANDARD FULL-LENGTH UNIGENES OF LEAVES FOR GO ANALYSIS

AT3G09390.1, AT3G22840.1, AT3G09390.1, AT1G29910.1, AT3G59540.1, AT1G28480.1, AT2G30570.1, ATMG00030.1, AT3G47070.1,  
AT5G03850.1, AT5G38410.1, AT5G20630.1, AT5G03240.3, AT3G48140.1, AT5G24105.1, AT5G02380.1, AT3G16640.1, AT1G29920.1,  
AT2G20920.1, ATCG00070.1, AT5G05270.1, AT5G18380.1, AT2G33120.2, AT2G47400.1, AT5G20620.1, AT5G37770.1, AT1G11185.1,  
AT2G27720.3, AT1G08830.1, AT5G53340.1, AT3G51030.1, AT1G70280.2, AT1G12010.1, AT1G67920.1, AT2G30570.1, AT3G45180.1,  
AT5G22430.1, AT1G66240.3, AT2G20562.1, AT2G44310.1, AT2G20940.1, AT4G24010.1, AT1G29930.1, AT3G07880.1, AT5G07250.1,  
AT5G09810.1, AT5G59690.1, AT1G79040.1, AT1G07890.6, AT4G13010.1, AT4G31130.1, AT1G58983.1, ATCG01090.1, AT2G46390.1,  
AT1G72610.1, AT3G59540.1, AT4G32480.1, AT2G07635.1, AT3G46900.1, AT3G13662.1

#### D 2.5: ARABIDOPSIS LOCUS HITS OF UNIGENES OF DROUGHT STRESSED LEAVES FOR GO ANALYSIS

AT1G12760.2, AT5G19430.1, AT3G28730.1, AT3G05590.1, AT2G15220.1, AT3G05590.1, AT5G66400.1, AT4G29260.1, AT1G03220.1,  
AT1G11260.1, AT5G38410.1, ATCG00950.1, AT1G08380.1, AT3G47470.1, AT2G39460.2, AT2G26550.2, AT1G67430.1, AT1G15240.2,  
AT2G36060.2, AT3G25430.1, AT2G07360.1, AT1G71900.1, AT3G07565.4, AT3G10220.1, AT4G37150.1, AT4G05634.1, AT3G21600.2,  
AT4G27450.1, AT5G62700.1, AT3G06240.1, AT2G01540.1, AT2G39550.1

### D 3. ANALYSES OF HOMOLOG UNIGENES

#### D 3.1: ANALYSIS OF HOMOLOGS BETWEEN UNIGENES OF STANDARD AND DROUGHT STRESSED ROOTS

Unigenes		E-value	Score	% Identity
Standard	Drought			
USR2	UEST25	9.00E-07	242	86%
USR15	UEST8	0.0	1136	99%
USR17	UEST453	0.0	698	98%
USR1230	UEST352	7.00E-02	93	93%
USR1234	UEST43	5.00E-05	188	86%
USR1260	UEST22	2.00E-07	246	100%
USR1270	UEST508	2.00E-05	176	89%
USR1279	UEST476	8.00E-94	333	98%
USR1323	UEST2	1.00E-13	452	96%
USR1524	UEST45	1.00E-11	391	87%

#### D 3.6: ASSEMBLY INFORMATION OF HOMOLOG UNIGENES BETWEEN STANDARD AND DROUGHT STRESSED ROOTS

	Unigene IDs	EST nos	Genbank IDs
Normal growth	USR2	4	GH623676, HS389648, HS389644, HS389820
	USR15	2	HS389823, HS389824
	USR17	2	HS389732, HS389936
	USR1230	1	HS389643
	USR1234	1	HS389647
	USR1260	1	HS389673
	USR1270	1	HS389683
	USR1279	1	HS389692
	USR1323	1	HS389736
	USR1524	1	HS389937
Drought stressed	UEST25	2	GT969236, GT969062
	UEST8	5	GT969091, GT969092, GT969109-GT969111
	UEST453	1	GT969243
	UEST352	1	GT969142
	UEST43	2	GT969157, GT969158
	UEST22	3	GT969118-GT969120
	UEST508	1	GT969298
	UEST476	1	GT969266
	UEST2	22	GT968992-GT969012, GT969209
	UEST45	2	GT969165, GT969166

### 3.3: ANALYSIS OF HOMOLOGS BETWEEN UNIGENES OF DROUGHT STRESSED ROOTS AND LEAVES

Unigenes		Score	E-value	% Identity
Roots	Leaves			
UEST1	UDL16	394	1E-112	94%
UEST3	UDL17	432	1E-123	99%
UEST24	UDL118	347	1E-098	85%
UEST52	UDL35	163	1E-042	100%
UEST429	UDL53	995	0.0	100%
UEST512	UDL42	688	0.0	99%
UEST527	UDL40	293	1E-081	96%
UEST553	UDL48	444	1E-127	98%
UEST566	UDL108	412	1E-118	100%
UEST592	UDL28	182	2E-048	85%
UEST12	UDL34	805	0.0	100%
UEST54	UDL1	515	1E-148	98%

D 3.4: ASSEMBLY INFORMATION OF HOMOLOG UNIGENES BETWEEN DROUGHT STRESSED ROOTS AND LEAVES AND THEIR ACCESSION NUMBERS

	Unigene IDs	EST nos	Genbank IDs
Leaves	UEST1	201	GT968791-GT968991
	UEST3	2	GT969013, GT969031
	UEST24	3	GT969127, GT969128, GT969129
	UEST52	2	GT969186, GT969093
	UEST429	1	GT969219
	UEST512	1	GT969302
	UEST527	1	GT969317
	UEST553	1	GT969343
	UEST566	1	GT969356
	UEST592	1	GT969382
	UEST12	4	GT969061, GT969063, GT969064, GT969065
UEST54	2	GT969185, GT969186	
Roots	UDL1	24	GH623574.1, GH623440.1, GH623200.1, GH623476.1, GH623217.1, GH623207.1, GH623330.1, GH623449.1, GH623328.1, H623340.1, GH623443.1, GH623275.1, GH623238.1, GH623553.1, GH623506.1, GH623344.1, GH623498.1, GH623462.1, GH623552.1, GH623523.1, GH623428.1, GH623388.1, GH623318.1, GH623412.1
	UDL17	11	GH623378.1, GH623307.1, GH738651.1, GH623232.1, GH623436.1, GH738733.1, GH738706.1, GH738705.1, GH738618.1, GH738647.1, GH738620.1
	UDL118	1	GH738707.1
	UDL35	2	GH623183.1, GH623503.1
	UDL53	12	GH738571.1, GH738751.1, GH738763.1, GH738754.1, GH738757.1, GH738762.1, GH738748.1, GH738621.1, GH738589.1, GH738619.1, GH738689.1, GH738743.1
	UDL42	3	GH738699.1, GH738657.1, GH738640.1
	UDL40	7	GH738713.1, GH738613.1, GH623560.1, GH623544.1, GH738664.1, GH738752.1, GH738761.1
	UDL48	21	GH738637.1, GH738700.1, GH738660.1, GH738610.1, GH738693.1, GH738722.1, GH738572.1, GH738607.1, GH738602.1, GH738710.1, GH738652.1, GH738634.1, GH738655.1, GH738731.1, GH738623.1, GH738654.1, GH738661.1, GH738644.1, GH738669.1, GH738592.1, GH738735.1
	UDL108	1	GH623230.1
	UDL28	6	GH623300.1, GH738612.1, GH738641.1, GH738681.1, GH738630.1, GH738773.1
	UDL34	11	GH623197.1, GH623549.1, GH623434.1, GH623252.1, GH623446.1, GH623221.1, GH623288.1, GH623198.1, GH623195.1, GH623289.1, GH623213.1
	UDL16	20	GH623399.1, GH623342.1, GH623301.1, GH623427.1, GH623524.1, GH623251.1, GH623282.1, GH623536.1, GH623466.1, GH623193.1, GH623528.1, GH623312.1, GH623201.1, GH623527.1, GH623313.1, GH623320.1, GH623194.1, GH623267.1, GH623227.1, GH623298.1, GH623242.1, H623569.1, GH623547.1, GH623279.1, GH623261.1

**D 3.5: ANALYSIS OF HOMOLOGS BETWEEN UNIGENES OF STANDARD AND WINTER DORMANCY STRESSED LEAVES**

Unigenes		Score	E-value	% Identity
Standard	Winter dormancy			
USL1	UWL40	890	0.00	99%
USL2	UWL41	357	1.00E-101	96%
USL3	UWL42	434	1.00E-124	99%
USL16	UWL31	438	1.00E-125	92%
USL20	UWL43	210	4.00E-06	93%
USL21	UWL36	1124	0.00	96%
USL25	UWL44	98	1.00E-02	80%
USL26	UWL45	280	2.00E-08	99%
USL27	UWL46	1102	0.00	99%
USL28	UWL47	319	3.00E-89	100%
USL29	UWL48	141	5.00E-36	85%
USL30	UWL49	815	0.00	97%
USL31	UWL50	902	0.00	99%
USL32	UWL51	630	0.00	99%
USL33	UWL52	188	6.00E-50	83%
USL34	UWL53	979	0.00	98%
USL35	UWL54	474	1.00E-136	95%

**D 3.6: ASSEMBLY INFORMATION OF HOMOLOG UNIGENES BETWEEN STANDARD AND WINTER DORMANCY STRESSED LEAVES**

Unigene IDs	EST nos	Genbank IDs
USL1	4	GH733954, GH733888, GH733975, GH733855
USL2	3	GH733942, GH733822, GH733775
USL3	44	GH733928, GH733893, GH734019, GH734105, GH733786, GH733835, GH734182, GH733926, GH734196, GH734091, GH734163, GH734168, GH733763, GH734115, GH734154, GH733793, GH734063, GH734011, GH733913, GH733917, GH734042, GH733856, GH733847, GH734002, GH734049, GH734092, GH733752, GH734155, GH734097, GH734200, GH733868, GH733757, GH734180, GH734099, GH734193, GH733751, GH734121, GH734111, GH733780, GH733828, GH734098, GH733839, GH734176, GH733783, GH734030, GH733759, GH734040
USL16	2	GH733753, GH733754
USL20	3	GH734032, GH734160, GH733755
USL21	7	GH734025, GH734143, GH734194, GH733933, GH733935, GH734016, GH733834
USL25	1	GH734197.1
USL26	1	GH734061.1
USL27	1	GH734158.1
USL28	1	GH733750.1
USL29	1	GH733979.1
USL30	1	GH733807.1
USL31	1	GH734140.1
USL32	1	GH733804.1
USL33	1	GH734095.1
USL34	1	GH733827.1
USL35	1	GH733760.1
UWL40	1	FF682716.1
UWL41	1	FF682752.1
UWL42	1	FF682804.1
UWL31	4	FE942971, FE942875, FE942963, FE943066
UWL43	1	FF682790.1
UWL36	6	FE942929, FE943022, FE943060, FE942898, FE943054, FE942801
UWL44	1	FF682772.1
UWL45	1	FF682733.1
UWL46	1	FF682709.1
UWL47	1	FE942799.1
UWL48	1	FE942796.1
UWL49	1	FE943019.1
UWL50	1	FE942954.1
UWL51	1	FE942912.1
UWL52	1	FE942901.1
UWL53	1	FE942873.1
UWL54	1	FE942866.1

**D 3.7: ANALYSIS OF HOMOLOGS BETWEEN UNIGENES OF WINTER DORMANCY AND DROUGHT STRESSED LEAVES**

Unigenes		Score	E-value	% Identity
Winter dormancy	Drought stress			
UWL56	UDL48	627	0	98%
UWL57	UDL67	204	2.00E-54	92%
UWL58	UDL116	607	1.00E-175	98%
UWL14	UDL118	129	1.00E-31	81%
UWL3	UDL121	517	1.00E-149	98%

**D 3.8: ASSEMBLY INFORMATION OF HOMOLOG UNIGENES BETWEEN WINTER DORMANCY AND DROUGHT STRESSED LEAVES**

	Unigene IDs	EST nos	Genbank IDs
Winter dormancy	UWL56	1	FF682741.1
	UWL57	1	FF682754.1
	UWL58	1	FE943091.1
	UWL14	3	FE942823, FE942933, FE942826
	UWL3	2	FF682813, FE943007
Drought stress	UDL48	21	GH738637.1, GH738700.1, GH738660.1, GH738610.1, GH738693.1, GH738722.1, GH738572.1, GH738607.1, GH738602.1, GH738710.1, GH738652.1, GH738634.1, GH738655.1, GH738731.1, GH738623.1, GH738654.1, GH738661.1, GH738644.1, GH738669.1, GH738592.1, GH738735.1
	UDL67	1	GH623477.1
	UDL116	1	GH738709.1
	UDL118	1	GH738707.1
	UDL121	1	GH738680.1

**D 3.9: ANALYSIS OF HOMOLOGS BETWEEN UNIGENES OF STANDARD AND DROUGHT STRESSED LEAVES**

Unigenes		score	E- value	% Identity
Standard	Drought stress			
USL36	UDL35	170	5.00E-45	100%
USL37	UDL53	892	0	99%
USL14	UDL54	684	0	99%

**D 3.10: ASSEMBLY INFORMATION OF HOMOLOG UNIGENES BETWEEN STANDARD AND DROUGHT STRESSED LEAVES**

	Unigene IDs	EST nos	Genbank IDs
Normal growth	USL36	1	GH734052.1
	USL37	1	GH733823.1
	USL14	22	GH734088, GH734078, GH734047, GH734187, GH734043, GH734005, GH734170, GH734083, GH733846, GH734142, GH734138, GH734033, GH733813, GH733931, GH733892, GH733825, GH733927, GH734135, GH733876, GH733936, GH733920, GH733959
Drought stress	UDL35	2	GH623183.1, GH623503.1
	UDL53	12	GH738571.1, GH738751.1, GH738763.1, GH738754.1, GH738757.1, GH738762.1, GH738748.1, GH738621.1, GH738589.1, GH738619.1, GH738689.1, GH738743.1
	UDL54	2	GH738567.1, GH738526.1

#### D 4. NUCLEOTIDE SEQUENCES OF PRE-miRNAs INCLUDED IN THE PRESENT STUDY

##### D 4.1: FOLD BACK NUCLEOTIDE SEQUENCES OF THE IDENTIFIED TEA miRNAs

###### >csi-mir408 (*Camellia sinensis*)

GAGAGAAGAUGUCUCAGGGAAGAGGCAGUGCAGUGGUGGCGGCGGCGGCGGUGGUGCUGUGCCUGAUGGUGGU  
GACUGCAGAGGCAGCUACCUAUACGGUGGGUGGUGCCGGUGGUGGUGGACCUUUAAACAGUGUCAGCUGGCCCAAGG  
GGAAGCGCUUUAGAGCUGGUGACAUACUUGUGUCCAAUUACAGCCCACCGGCACACAAUGUGGUCAGUGUGAACA  
AGGCUGGUUAUGAUAGUUGCAAGGCACAAGCUGGAGCCAGAGUGUUUUUCUUCUGGGAAAGAUCAAAUCAAGCUUG  
UCAAGGGUCAGAACUUCUUCAUUUGCUCUCUCCUGGCC

###### >csi-mir1171 (*Camellia sinensis*)

ACGCACCAUUAGAUGUCCAGUCCAAGGGCAAGUGGAGGAACCUUJCCUCCAGAAUUUAUACAAAUUUUUAAGCA  
AGAACCACCAGGGGUGGAAACA AUUUUUUAGAAU AUUGCUUCUUUUUGGGAUAAGGACGUGCUUGUGAGAGUGUA  
UGUGGAGUGGAGUGGAGUGAAGUGGAGUGG

###### >csi-mir414a (*Camellia sinensis*)

UCUCUCUGCAAUUUUCGUUAUCUGUCUGCCUGUCUCCUCCUCAGUAUCAUUGGAGUUUUGGGGUGUUGAAGUAA  
AAGCUGGGCAGCCUCUUAAGGUCAAUCCUGAAGUAGGCAGCGUCAUACACAUUCCAGGCAGCACUGGGUGAGG  
GAAAGAAGGGAAAAGGAAAUGACAUUGUCCUCUJUGGGUGAAUUAUAAUGGCAAGAAAUUAGUUCUAGGAUCAC  
UUUCUGCAGAGAAGUUCCUCAAGUUUAUJUGAUUUGGUGUUUGAGAAAGAAUUUGAGCUGUCACAUGACUGGA  
AAGAUGGGAGUGUCUACUUUUGGGAUAUUCGGCAGAUAAUCAGUUUGACGAAGAUGUAGAAGAUGAGUUUGAU  
GAGGAUGAUUCUUCUGAGGAUGAAGAUUACGCCGACUCAUCACAGAGAAUGGGAAAGAUGAAUCUAAAGGUGAG  
GAAGCAAAGCCUGCUGCAAGUAAGGGUAAAGCUGCCAAGCUGCCAAGGUGACUGUUGUGGAGCCAAAAAAGAUG  
AUGACAGUGAUGAAGAUGAUGAUACAGAUACGAUGAUGAAGAUGUUGACAAGGUCAUGAUGGAUGCUUCAGAU  
AUAGUGACUCGGAGGACGAAGAUGAUACUGAUGAAGAAGAUGAUGAUGAGGAAGAAGAUGAGG

###### >csi-mir414d (*Camellia sinensis*)

GGAGAUGAUGACGAUGACGAUGACGAUGGCGACUAUGAUUAUGCCCCGGCUGCUUGAAUUUGAAUAUUAUGAA  
CAACUUAACAAAAACCCUAGUUUAAAAAACAUAAAACCCUAAAUUUGUAUGUCCCAAACCCUAAAUUUAGGCCUCU  
UUUGCUAAUUAUAGUAGCUCUCUCUUGUGGGUCGGGUCUCU

###### >csi-mir414f (*Camellia sinensis*)

CUGAGUGCUGGUGGUGGAGGUGUUUCUUCUUCUUCUUGAAUCUUGAUCAUCAUCAUCAUCUUCAG

###### >cas-mir1122 (*Camellia assamica*)

GCUAUACUCCCUCCGUCCAAAAUAAUGGUCCUUUUUGGGAAUCCAACUUUUUUUGGUAGUACAAUUAUUGUACU  
UUUAAAAAGUUGGUUUUCCAAAAUUACCCUUUGAUAGUACACUUUUUUGAGAAUUAUCAAUGUCAUAAUUGUACUUU  
UUUAUUAUUUAUGAAUUUAAAAGAAGGGCAAAAUGGGAAAUUUUAUUAUUUUUGUACUUGAUUUUUGAAGAGG  
GACUAAUUAUUUGGAACAAUGAAAAUGGAAUAAGGGACUAAUUAUUGGGACGGAGGGAGUACAAGU

**>csi-mir414g (*Camellia sinensis*)**

GAGAAGGAGGACGAUGAUGAUGAGGAAGGAGAUAAUCAGGUGGGUGCCGUGGUCAUCGACUCCUCCUCCUC

**>csi-miRf10132-akr (*Camellia sinensis*)**

UUCAUCAACGACAUCUUCGAGAAGCUCGCUCAGGAAGCCUCGAGGCUCGCUAGGUUAACAAGAAGCCUACCAUUA  
CUUCUCGGGAGAUCAGACUGCUGUCAGGCUUGUUCUUCUGGGUGAAUUGGCUAAGCACGCUGUGUCUGAGGGCA  
CUAAGGCGGUCACGAAGUUACUAGCUCUUGAAGGUGUUCGAUGGUGGA

**>cja-mir2910 (*Camellia japonica*)**

GUUCUUAGUUGGUGGAGCGAUUUGUCUGGUUAAUUCGUUAACGAACGAGACCUCAGCCUGCUAACUAGCUAUGC  
GGAGGUGACCCUCCGCAGCUAGCUUCUUGAGAGGGACUAUGGCCCUUCAGGCCACGGAAGUUGAGNCAUAACAG  
GUCUGUGAUGCCCUUAGAUGUUCUGGRCCGCACGCGCGCUACACUGAUGUAUUAACGAGUCUAUAGCCUUGGCC  
GACAGGCCCGGGUAAUCUUUGAAAUUCUUCGUGAUGGGGAUAGAUAUUGCAAUUGUUGGUCUUAACGAGGAA  
U

**>csi-mir2914 (*Camellia sinensis*)**

UUUCUGCCCUAUCAACUUUCGAUGGUAGGAUAGUGGCCUACUAUGGUGGUGACGGGUGACGGAGA

**>cas-mirf10185-akr (*Camellia assamica*)**

CCUGCUACAAUGUUUJCCCCUUCUCCGAAGGUUCUCUUGCCGCCCAAUUUUCGGGGGGAAAAUUAUUGAAA  
AUGCGGAUCAAAAGGGGACAAAAGGGGGGACAAAGAGGCCCAAAGGGGGGAAGCAGCGGG

**>cas-mirf11590-akr (*Camellia sinensis*)**

GUAAGGUUGAAGGCACACCAAACAAGUUCUUCUUGGAAGAUGGCUGGGGUGUUGCGUUUUGGCAACCUUGC

**>csi-mir414h (*Camellia sinensis*)**

GGGAUUCUGACGAUGAUGAGGAUGAGUUUGGUUUUAGGCAUGGUGGUCAUGGUAAACAACAGCGCUCGGAGCUA  
AGAGACAGUGGAAAUGAUUGGCUAAUAGGUGGUAGGCAAUUCUUGGGCGGUCAUCCUUUGGAAAUAGGGACAGA  
AGCUUUGGUGGUGCAUGUUCAAUUGUGGGCGGUCUGGCCAUAGGGCAUCAGAAUGCCCCACCAAGAAGCAAAGC  
UAUUAGUUUCCGUGCCCUGUCUGCUGCACCAGGUUCCUUAACUGCUUGUAAACAGAUAGAGAGUAACUUCUG  
UAACUCCAGCCACUUUUGCAACUGCUGCUGAGCAUUGACGUGAGAUACUGUGUCUCAAGUAAAGAGACCGUUUGA  
GAGGUCUGCAAGGUGGAAAAAAGCUCCUGUGAAGAGUUGCAGACAAGCAGCUUCUGUGUCGUCUCCUGCAAACC  
UCACUUCACCAACCUCAUCUCCCAAUUUUAUGUGAUUUUAUUGACAGGCCAAUUUUUUGAUGUCUAAUUGGCUUCU  
GACCUUAAUUUAAUGUUUGUGUGUUUUUGUUCUUUUGCUUUUCCUGCCAUUGGUUUUAACCCUAUGUAGACCUCAG  
UUUCAUUUCUAUAGAGCUGAUUUCAUUCUGCCAUGCAGACUCCC

**D 4.2: NUCLEOTIDE SEQUENCES OF PRE-miRNAs OF miRNA408 FAMILY MEMBERS USED FOR THE STUDY OF CONSERVED NATURE**

**>zma-mir408 (*Zea mays*)**

GGGGUUGGUUUUGAUUUGGAGACAGGGAUGAGACAGAGCAUGGGAUGGGGCAUCAACAAAGUGGAGGGACUAG  
CUUGCGAGGCAGAAAGAAGGUGCCAGUJCCGGUGCCUCCCCGGUGAAACGAUGAUGGGAGUGUUGUUGCUCCCUC  
CCUGCACUGCCUCUUCCCUGGCUCCGAUCCCCACCGUUGC

**>tae-mir408 (*Triticum aestivum*)**

AUUUUGUGAGUGGAGAGGGGGAGGAGACAGGGAUGGAGCAGAGCAAGGGAUGAGGCAAGCAACAAAUUUACCA  
CCUGAUUAUGAGAAGAGGGAGAGAGUUGCCAGAGCUUCUGUUGCUGUUGUUGCUCCCUCUCCUGCACUGCCUCUUC  
CCUGGCUCCCCUCCAAAUCUCUCCCUCUCCCCUCUCU

**>sof-mir408a (*Saccharum officinarum*)**

AGAAGAUGGGUAUGGUUGGAGACAGGGAUGAGGCAGAGCAUGGGAUGAGGGCAUCAACAAAUUUCCAUUUCUG  
UCCUCCGCUAGGCCGCUACUGCAUUUCUGUUUGCUUUGCUCACAAAACGGAGGGAUUUGUGAGAGUUUACAGGCAG  
AAAGAACAAAAGAAGGUGCCUCCCUGGUGAAGUGGUGAUGGCCUGACCUGAGAGCGGCUGAGAGCUCAGCUGGUGUC  
CUGUUGUUGCUUCCUCCCCUGCACUGCCUCUUCCCUGGCUCCCCACCGUUGCCCUUGC

**>sof-mir408b (*Saccharum officinarum*)**

AGAAGAUGGGUAUGGUUGGAGACAGGGAUGAGGCAGAGCAUGGGAUGGGGCAUCAACAAAUUUCCAUUUCUG  
UCCUCCGCUAGGCCGCUACUGCAUUUAUGUUUGCUUUGCUCACAAAACGGAGGGAUUUGUGAGAGUUUACAGGCAG  
AAAGAACAAAAGAAGGUGCCUCCCUGGUGAAGUGGUGAUGGCCUGACCUGAGAGGGCUGAGAGCUCAGCUGGUGUC  
CUGGUGUUGUUGCUUCCUCCCCUGCACUGCCUCUUCCCUGGCUCCCCACCGUUGCCCUUGC

**>sof-mir408c (*Saccharum officinarum*)**

AGAAGAUGGGUAUGGUUGGAGACAGGGAUGAGGCAGAGCAUGGGAUGGGGCAUCAACAAAUUUCCAUUUCUG  
UCCUCCGCUAGGCCGCUACUGCAUUUCUGUUUGCUUUGCUCACAAAACGGAGGGAUUUGUGAGAGUUUACAGGCAG  
AAAGAACAAAAGAAGGUGCCUCCCUGGUGAAGUGGUGAUGGCCUGACCUGAGAGGGCUGAGAGCUCAGCUGGUGUC  
CUGGUGUUGUUGCUUCCUCCCCUGCACUGCCUCUUCCCUGGCUCCCCACCGUUGCCCUUGC

**>sof-mir408d (*Saccharum officinarum*)**

GGAAGGUAUGUUUGAUUUGGAGACAGGGACGAGGCAGAGCAUGGGAUGGGGCAUCAACAAAUUUCCAUUUCCG  
UUUGCUUUGCUCACAAAACGAAGGUGCCUGCCUCCCUGGUGAUGGCCUGACCUGAGAGGGCUGAGAGCUCAGCUGG  
UGUCCUGGUGUUGUUGCUUCCACCCUUGCACUGCCUCUUCCCUGGCUCCCCACCGUUGCCCUUGC

**>sbi-mir408 (*Sorghum bicolor*)**

ACAGGGACGAGGCAGAGCAUGGGGAUGGGGCCAUCAACAACAAAAUUCCAAUUUCCGUUUGCUUGCCCACAAAUG  
GAGGGACUUGUCAGGAGAGGUAUCAGGCAGAGAGAAGGUGCCUCCUGGUGAAAUGGUGAUGGCCUGAGAGAGC  
UCAGCUGGUGUCCUGGUGUUGUUGCUUCCUCCCCUGCACUGCCUCUUCCUGGCU

**>osa-mir408 (*Oryza sativa*)**

GGGAGUUCUGUGAUUGGAGAGGAGAGGAGACAGGGAUGAGGCAGAGCAUGGGGAUGGGGCUAUCAACAGAUGUAG  
AUUAUCCUUGCACAAGAGAUGAUGAUGAGCUGUGAAUGAGUUCUGAGAGAUUGCCUGGUGUUGUUGUUGCUCCC  
UCCCCUGCACUGCCUCUUCCUGGCUCCCCUGCACACCUCUCUCUCUCUCUCUCUCUCUGUGU

**>vvi-mir408 (*Vitis vinifera*)**

AAGAGGAAGACGGGGACGAGGUAGUGCAUGGAUGGAACUUAUUAACAGAAGAAUGUUAAGCUGUUUUUGCUCUACC  
CAUGCACUGCCUCUUCCUGGCUCUGUCUCUC

**>smo-mir408 (*Selaginella moellendorffii*)**

AUUUAGUAGUAGCUAGGGAGAGACAUUGCAUGAUAGCUACACCAAAGAGCAACCUCGGCUCCUUGUGUGGCUUUG  
AUGCACUGCCUCUUCCUGGCUGCGGCCAAGUUA

**>rco-mir408 (*Ricinus communis*)**

AAAGACUGGGAACAGGCAGUGCAUGGAUGGGGCUACUAACAGAAAACAUUCUGUUUUUGGCUCUACCCCUGCACUGCC  
UCUUCCUGGCUUCCG

**>pta-mir408 (*Pinus taeda*)**

GAGACAGGGACGAGUUAGGGCAUGGGAGUUGCAUAUGCAGAAACGUCUGCUUCUGCCAUUCUUAUGCACUGCCUC  
UUCCUGGCUC

**>ppt-mir408b (*Physcomitrella patens*)**

GUGGAAGAGAGAGUGGUGGAAGGGAGGGAAGCCAGCGUGAGGCAAUGCAUGACAACAGCAUGCCCAGGAGGUC  
CUGAGGGUGUUGUCCUCAUGCACUGCCUCUUCCUGGCUCCCUACAUAGCUCGCCAUUCUUGUGCUCU

**>ppe-mir408 (*Prunus persica*)**

ACUGAUAAAGACAGGGAACAGGUAGAGCAUGGAUGGAGUUCCCAACAGAAAAUAGAGCUGUUGUGGCUCUACUCA  
UGCACUGCCUCUUCCUGGCUGCCGUCCUC

**>bdi-mir408 (*Brachypodium distachyon*)**

GCUCAAGGUGAAACCAACCGCUCUGCUCUCCGGCUCAGUGGAGCCAAACUAGCUAGCAACAGAAUCCAUGCACUG  
CCUCUUCCUGGGGAUCGAUCAUCAAAGAGGGCGAGCUUGACAAGGACGGCGGGG

**>cpt-mir408 (*Citruspon cirus trifoliata*)**

AGACAAAAGACGGGGAACAGGCAGAGCAUGGAUGGAACCAUUAACAGGUUCUCUGUUUUUGGCUCCUCCAUGCACU  
GCCUCUUCCCUGGCUCUCUGCCU

**>ptc-mir408 (*Populus trichocarpa*)**

AGAGACAGAUGAAGACGGGGAACAGGCAGAGCAUGGAUGGAGCUACUAACAGAAGUACUUGUUUUUGGCUCUACCC  
AUGCACUGCCUCUUCCCUGGCUUGUGGCUC

**>csi-mir408 (*Camellia sinensis*)**

GAGAGAAGAUGUCUCAGGGAAGAGGCAGUGCAGUGGUGGGCGGCGGCGGCGGUGGUGCUGUGCCUGAUGGUGGU  
GACUGCAGAGGCAGCUACCUAUACGGUGGGUGGUGCCGGUGGCUGGACCUUUAACAGUGUCAGCUGGCCCAAGG  
GGAAGCGCUUUAGAGCUGGUGACAUACUUGUGUCCA AUUACAGCCCACCGGCACACAAUGUGGUCAGUGUGAACA  
AGGCUGGUUAUGAUAGUUGCAAGGCACAAGCUGGAGCCAGAGUGUUUUUCUUCUGGGAAGAUA AAAUCAAGCUUG  
UCAAGGGUCAGAACUUCUUCAUUUGCUCUCUCCCUGGCC

**D 4.3: NUCLEOTIDE SEQUENCES OF PRE-miRNAs OF miRNA408 FAMILY MEMBERS USED FOR THE  
CONSTRUCTION OF PHYLOGENETIC TREE**

**>csi-mir408 (*Camellia sinensis*)**

GAGAGAAGAUGUCUCAGGGAAGAGGCAGUGCAGUGGUGGGCGGCGGCGGCGGUGGUGCUGUGCCUGAUGGUGGU  
GACUGCAGAGGCAGCUACCUAUACGGUGGGUGGUGCCGGUGGCUGGACCUUUAACAGUGUCAGCUGGCCCAAGG  
GGAAGCGCUUUAGAGCUGGUGACAUACUUGUGUCCA AUUACAGCCCACCGGCACACAAUGUGGUCAGUGUGAACA  
AGGCUGGUUAUGAUAGUUGCAAGGCACAAGCUGGAGCCAGAGUGUUUUUCUUCUGGGAAGAUA AAAUCAAGCUUG  
UCAAGGGUCAGAACUUCUUCAUUUGCUCUCUCCCUGGCC

**>ahy-mir408 (*Arachis hypogaea*)**

GAAGAAGAGAUGACAAAGAACUJGGGAACAGGCAGAGCAUGAAUGGAACUAUCAAUAGACACAUUUUGUUCAUUGAC  
GCUCAUGCACUGCCUCUUCCCUGGCUCUCUCUUCUUCUUUUUCCUUCU

**>aqc-mir408 (*Aquilegia coerulea*)**

AGCCAGGGAAGAGGCAGUGCAUGGAAAGAGCAACAACAUAUUUGUGUUGAUUGGUUCAUCCAUGCUCUGCCUC  
AUCCUUGUCUGU

**>ath-mir408 (*Arabidopsis thaliana*)**

AAGGUUAGAUJGGUAUUUGCAAUGAAAGAAGACAAAGCGUAAUUGAGAGAGAGACAGGGAACAAGCAGAGCAUGGA  
UUGAGUUUACUAAAACAUUAAACGACUCUGUUUUGUCUCUACCCAUGCACUGCCUCUCCCUGGCUCCCUCUUUUU  
UUCUCUAUAUUUCUCUCUCUCCUUUCAUUUCACAGCUUUCAAUGGAAUUUUUAUUGCUACUGCUAACG

**>bdi-mir408 (*Brachypodium distachyon*)**

GCUCAAGGUGAAACCAACCGCUCUGCUCUCCGGCUCCAGUGGAGCCAAACUAGCUAGCAACAGAAUCCAUGCACUG  
CCUCUCCCUGGGGAUCGAUCAAAAAGAGGGCGAGCUUGACAAGGACGGCGGCGG

**>cpt-mir408 (*Citruspon cirus trifoliata*)**

AGACAAAAGACGGGGAACAGGCAGAGCAUGGAUGGAACCAUUAACAGGUUCUCUGUUUUGGCUCCUCCCAUGCACU  
GCCUCUUCCUGGCUCUCUGCCUU

**>ghr-mir408 (*Gossypium hirsutum*)**

AGGCACUGCGUCUUCCUGGAUUCUGGUGGUGUCCAUCAGAAGGGCCGAGGACGGAGGAGCGGCGCCA

**>gma-mir408a (*Glycine max*)**

GCUGGGAACAGGCAGGGCACGAAUGGAGCUAUC AACAGAAAUGGUAAAAGUGAGAAUGAAAGGAGAGAGAGAGAG  
AAGAGAUCUGUUGUGGCUACGCUCAUGCACUGCCUGUUCCUGGC

**>osa-mir408 (*Oryza sativa*)**

GGGAGUUCUGUGAUUGGAGAGGAGAGGACAGGGAUGAGGCAGAGCAUGGGAUUGGGCUAUC AACAGAUGUAG  
AUUAUUCCUUGCACAAGAGAUGAUGAUGAGCUGUGAAUGAGUUCUGAGAGAUGGCUGGUGUUGUUGUUGCUCCC  
UCCCCUGCACUGCCUCUUCCUGGCUCCCCUGCACACCUCUCUCUCUCUCUCUCUCUCUGUGU

**>ppe-mir408 (*Prunus persica*)**

ACUGAUAAGACAGGGAACAGGUAGAGCAUGGAUGGAGUUCCCAACAGAAAAAUAGAGCUGUUGUGGCUCUACUCA  
UGCACUGCCUCUUCCUGGCUGCCGUCUC

**>ppt-mir408 (*Physcomitrella patens*)**

ACAACCAGCACUGCACUGCAUCUUCCUGUGCCAUCUUCGUCGAUUGUCUCGUACCGAAAGUGAUUCAACCACAGC  
GAGAUCAGACUCAACUGUUAAGCUGAUGACACUGAUACGAGAUGAUGAAGACAUAAGGCAGGGCAGCCAGUG

**>ppt-mir408b (*Physcomitrella patens*)**

GUGGAAGAGAGAGAGUGGUGGAAGGGAGGGAAAGCCAGCGUGAGGCAAUGCAUGACAACAGCAUGCCCAGGAGGUC  
CUGAGGGUGUUGUCCUCAUGCACUGCCUCUUCCUGGCUCCCUACAUAGCUCGCCAUUCUUGUGCUCU

**>pta-mir408 (*Pinus taeda*)**

GAGACAGGGACGAGUUAGGGCAUGGGAGUUGCAUAUGCAGAAACGUCUGCUUCUGCCAUUCUUAUGCACUGCCUC  
UCCCCUGGCUC

**>rco-mir408 (*Ricinus communis*)**

AAAGACUGGGAACAGGCAGUGCAUGGAUGGGGCUACUAACAGAAAACAUCUGUUUUGGCUCUACCCCUGCACUGCC  
UCUUCCUGGCUUCGG

**>sbi-mir408 (*Sorghum bicolor*)**

ACAGGGACGAGGCAGAGCAUGGGAUUGGGCCAUCAACAACAAAUUUCCAUUUCCGUUUUGCUUGCCCACAAAUG  
GAGGGACUUGUCAGGAGAGGUAUCAGGCAGAGAGAAGGUGCCUCCUGGUGAAAUGGUGAUGGCCUGAGAGAGC  
UCAGCUGGUGUCCUGGUGUUGUUGCUUCCUCCCCUGCACUGCCUCUUCCUGGCUCU

**>smo-mir408 (*Selaginella moellendorffii*)**

AUUUAGUAGUAGCUAGGGAGAGACAUUGCAUGAUAGCUACACCAAAGAGCAACCUCGGCUCCUUGUGUGGCUUUG  
AUGCACUGCCUCUUCCUGGCUGCGCCAAGUUA

**>sof-mir408a (*Brachypodium distachyon*)**

AGAAGAUGGGUAUGGUUGGAGACAGGGGAUGAGGCAGAGCAUGGGGAUGAGGGCCAUCAACAAAAUUUCCAUUUCUG  
UCCUCCGCUAGGCCGCUACUGCAUUUCUGUUUGCUUGCUCACAAAACGGAGGGAUUUGUGAGAGUUUAUCAGGCAG  
AAAGAACAAGAAGGUGCCUCCUGGUGAAGUGGUGAUGGGCCUGACCUGAGACGGCUGAGAGCUCAGCUGGUGUC  
CUGUUGUUGCUUCCUCCCCUGCACUGCCUCUUCCUGGCUCCCCACCGUUGCCCUUGC

**>sof-mir408b (*Brachypodium distachyon*)**

AGAAGAUGGGUAUGGUUGGAGACAGGGGAUGAGGCAGAGCAUGGGGAUGGGGCCAUCAACAAAAUUUCCAUUUCUG  
UCCUCCGCUAGGCCGCUACUGCAUUUAUGUUUGCUUGCUCACAAAACGGAGGGAUUUGUGAGAGUUUAUCAGGCAG  
AAAGAACAAGAAGGUGCCUCCUGGUGAAGUGGUGAUGGGCCUGACCUGAGAGGGCUGAGAGCUCAGCUGGUGUC  
CUGGUGUUGUUGCUUCCUCCCCUGCACUGCCUCUUCCUGGCUCCCCACCGUUGCCCUUGC

**>sof-mir408c (*Brachypodium distachyon*)**

AGAAGAUGGGUAUGGUUGGAGACAGGGGAUGAGGCAGAGCAUGGGGAUGGGGCCAUCAACAAAAUUUCCAUUUCUG  
UCCUCCGCUAGGCCGCUACUGCAUUUCUGUUUGCUUGCUCACAAAACGGAGGGAUUUGUGAGAGUUUAUCAGGCAG  
AAAGAACAAGAAGGUGCCUCCUGGUGAAGUGGUGAUGGGCCUGACCUGAGAGGGCUGAGAGCUCAGCUGGUGUC  
CUGGUGUUGUUGCUUCCUCCCCUGCACUGCCUCUUCCUGGCUCCCCACCGUUGCCCUUGC

**>sof-mir408d (*Brachypodium distachyon*)**

GGAAGGUAUGUUUGAUUGGAGACAGGGACGAGGCAGAGCAUGGGGAUGGGGCCAUCAACAAAAUUUCCAUUUCCG  
UUUGCUUGCUCACAAAACGAAGGUGCCUGCCUCCUGGUGAUGGCCUGACCUGAGAGGGCUGAGAGCUCAGCUGG  
UGUCCUGGUGUUGUUGCUUCCACCCUGCACUGCCUCUUCCUGGCUCCCCACCGUUGCCCUUGC

**>sof-mir408e (*Brachypodium distachyon*)**

AGAAGAUGGGUAUGGUUGGAGACAGGGGAUGAGGCAGAGCAUGGGGAUGAGGGCCAUCAACAAAAUUUCCAUUUCUG  
UCCUCCGCUAGGCCGCUACUGCAUUUAUGUUUGCUUGCUCACAAAACGGAGGGAUUUGUGAGAGUUUAUCAGGCAG  
AAAGAACAAGAAGGUGCCUCCUGGUGAAGUGGUGAUGGGCCUGACCUGAGACGGAUGAGAGCUCAGCUGGUGUC  
CUGUUGUUGCUUACUUCCUGCACUGACUCUUCCUGGCUCCCCACCGUUGCCCUUGC

**>tae-mir408 (*Triticum aestivum*)**

AUUUUGUGAGUGGAGAGGGGGGAGGAGACAGGGAUGGAGCAGAGCAAGGGAUGAGGCAAGCAACAAAAUUUACCA  
CCUGAUUAUGAGAAGAGGGAGAGAGUUUGCCAGAGCUUCUGUUGCUGUUGUUGCUCCCCUCCUGCACUGCCUCUCU  
CCUGGCUCCCCUCCAAAUCUCUCCUCCCCUCUCU

**>vvi-mir408 (*Vitis vinifera*)**

AAGAGGAAGACGGGGACGAGGUAGUGCAUGGAUGGAACUAUUAAACAGAAGAAUGUUAAGCUGUUUUUGCUCUACC  
CAUGCACUGCCUCUUCCUGGCUCUGUCUCUC

**>zma-mir408 (*Zea mays*)**

GGGGUUGGUUUUGAUUUGGAGACAGGGAUGAGACAGAGCAUGGGAUGGGGCCAUCAACAAAGUGGAGGGACUAG  
CUUGCGAGGCAGAAAGAAGGUGCCAGUGCCGGUGCCUCCCCGGUGAAACGAUGAUGGGAGUGUUGUUGCUCCUC  
CCCUGCACUGCCUCUUCCUUGGCUCCGAUCCCCACCGUUGC

**>ptc-mir408 (*Populus trichocarpa*)**

AGAGACAGAUGAAGACGGGGAACAGGCAGAGCAUGGAUGGAGCUACUAACAGAAGUACUUGUUUUGGCUCUACCC  
AUGCACUGCCUCUUCCUUGGCUUGUGGCUC

**D 4.4: NUCLEOTIDE SEQUENCES OF PRE-MiRNAs OF miRNA414 FAMILY MEMBERS USED FOR THE  
CONSTRUCTION OF PHYLOGENETIC TREE**

**>ath-mir414 (*Arabidopsis thaliana*)**

UCAUCAUUAUCAUCAUCAUUAUUAUCUUAUCAUCAUCGUCAUCAUCAUCAUCAUCGUAUGAGAAGAUAG  
AGAAGAGUGAGAGUAUGAGAUUUGAGUUGAGA

**>ghr-mir414a (*Gossypium hirsutum*)**

AGGAGCUGGUGUAGGAUUAAGUUCAUUAUCAUCCACAUCUCAUCAUCAUCUUCU

**>ghr-mir414b (*Gossypium hirsutum*)**

UUUUCUUCAUCAUCAUCGUCAUCAAUGUUACUCCUCAGCCGAAGGAAGGGCUAACGAGGUUGGUCCGCCGCCG  
UUCUUGACUAAAACUUUUGAUUUGGUGGAAGACCCUUCGACUGAUUCGGUGGUUUC AUGGAGUGAAG

**>ghr-mir414c (*Gossypium hirsutum*)**

GGAUGAUGAGUUUGACGUAGAUCUGGAGGAAAUAUGUCAUGGGAAGCAAUUUGGCAGUCAAUUCAGGAGAACAG  
CAGACACAGAAAGUCUAACAAUGGAGAUGCUGCUUCUUCGGUACAUGUUUCGGUGAUCGCUAUGUCUGACCAGCU  
AGGGCCACAGUGGCCCGUUCAUCAUCAUCGUCAUCAUCAUCAUCU

**>ghr-mir414d (*Gossypium hirsutum*)**

GAUGAUGAGUUUCACGUAGAUCUGGAGGAAAUAUGGUCAUCGAAGCAAUUUGGCAGUCAAUUCAGGAGAACAGC  
AGACACAGAAAGUCUAACAAUGGAGAUGCUGCUUCUUCGGUACAUGUUUCGGUAGAUCGCUAUGUCUCACCAGCU  
AUGGCCACAGUGGCCGUUCAUCAUCAUCGUCAUCAUCAUCAUCU

**>ghr-mir414e (*Gossypium hirsutum*)**

UCAUCUUCAUCAUCAUCCUCCGCUUCAUCGUCGUCGUCGGCUUCUUCUUCGUCCAACGGCAGCGACGGCGGAGAU  
AGCAGCAGUGCCAGUGGAAGCGCUAGCAGCGGAGGAGAAGAGGAAAUGG

**>ghr-mir414f (*Gossypium hirsutum*)**

UCAUCAUCAUCAUCAUCUUCUUCUUCUUCUUCUUCGAAUGAUGACAAAGAUGAUUGUUUUGGUGU

**>osa-mir414 (*Oryza sativa*)**

GUUGCCGUGCCGUAUCCUCAUCAUCAUCGUCCUUGCAUGAAACCGCCAGCAAUCCUUGCGCGGGGGAGGGG  
GAAGGGGAGGGGGAGGGGUCGGAGGGGCGGC

**>ppt-mir414 (*Physcomitrella patens*)**

AGACGUUCAAGAACAUAUJAGAGAAGACCAUUGCUJAGAAUGUGGGGCUCACCAGAUUUUCCAACUGCAUUGGCUAGC  
UCUGCUAGCGCUUUGGCAGUAACCAUAUCUUUGUCAUCUUCAUCAUCGUCAUCUUCGUCAUCAUCUUCGUCAUCCU  
CAUCAUCCUCGUCCUCAGCAUCAG

**>zma-mir414 (*Zea mays*)**

GUCGUCAUCCUCAUCAUCGUUCGUCCUAGAGUCGUCUGGGUCGUCACCUCGGGUGCAGGAGCUUGAACGGCUGG  
UGCAGGCGGUGCCGAGGUCGAUGAC

**>csi-mir414a (*Camellia sinensis*)**

UCUCUCUGCAAUUUUCGUUAUCUGUCUGCCUGUCUCCUCCUCAGUAUCAUUGGAGUUUUGGGGUGUUGAAGUAA  
AAGCUGGGCAGCCUCUUAAGGUCAAUCCUGAAGUAGGCAGCGUCAUACACAUUCCAGGCAGCACUGGGUGAGG  
GAAAGAAGGGAAAAGGAAUUGACAUUGUCCUCUUGGGUGAAUUAUAAUGGCAAGAAUUAGUUCUAGGAUCAC  
UUUCUGCAGAGAAGUUUCCUCAAGUUUCAUUGAUUUUGGUGUUUGAGAAAGAAUUUGAGCUGUCACAUGACUGGA  
AAGAUGGGAGUGUCUACUUUUGUGGAUUAUCGGCAGUAACAGUUUGACGAAGAUJAGAAGAUGAGUUUGAUGA  
GGAUGAUUCUUCUGAGGAUGAAGAUUAACGCCGACUCAUCACAGAGAAUGGGAAAGAUGAAUCUAAAAGGUGAGGA  
AGCAAAGCCUGCUGCAAGUAAGGGUAAAAGCUGCCAAGCUGCCAAGGUGACUGUUGUGGAGCCAAAAAAGAUGAU  
GACAGUGAUGAAGAUGAUGAUACAGAUAGCAGUAUGAUGAAGAUGUUGACAAGGUCAUGAUGGAUGCUUCAGAU  
AGUGACUCGGAGGACGAAGAUGAUACUGAUGAAGAAGAUGAUGAUGAGGAAGAAGAUGAGG

**>csi-mir414d (*Camellia sinensis*)**

GGAGAUGAUGACGAUGACGAUGACGAUGGGCGACUAUGAUUAUGCCCCGGCUGCUUGAAUUUGAAUUAUUAUGAA  
CAACUUAACAAAAACCCUAGUUUAAAAAACAUAAAACCCUAAAUUUGUAUGUCCCCAAACCCUAAAUUUAGGCCUCU  
UUUGCUAAUUAUAGUAGCUCUCUCUUGUGGGUCGGGUCUCU

**>csi-mir414f (*Camellia sinensis*)**

CUGAGUGCUGGUGGUGGAGGUGUUUCUUCUUCUUCUUGAAUCUUGAUCAUCAUCAUCAUCUUCAG

**>csi-mir414g (*Camellia sinensis*)**

GAGAAGGAGGACGAUGAUGAUGAGGAAGGAGAUAAUCAGGUGGGUGCCGUGGUAUCGACUCCUCCUCCUC

**>csi-mir414h (*Camellia sinensis*)**

UCUGACGAUGAUGAGGAUGAGUUUGGUUUUJAGGCAUGGUGGUAUGGUAACAACAGCGCUCGGAGCUCAAGAGAC  
AGUGGAAAUGAUUGGCUAAUAGGUGGUJAGGCAAUCUUCUGGGCGGUAUCCUUGGAAUJAGGGACAGAAGCUUU  
GGUGGUGCAUGUUUCAAUUGUGGGCGGUCUGGCCAUJAGGGCAUCAGAAUGCCCCACCAAGAAGCAAAGCUAUUJAG  
UUUCCGUGCCCUGUGCUGCUGCACCAGGUUCCUJCAACUGCUUGUAACAGAUJAGAGAUAAUUCUGUAACUC  
CAGCCACUUUUGCAACUGCUGCUGAGCAUUGACGUGAGAUACUGUGUCUCAAGUAAAGAGACCGUUUGAGAGGUC  
UGCAAGGUGGAAAAAAGCUCUUGAAGAGUUGCAGACAAGCAGCUUCUGUGUCGUCUCCUGCAAACCUACUUC

ACCAACCUCAUCUCCCAAUUUAUGUGAUUUUAUUGACAGGCCAAUUUUUUGAUGUCUAAUUGGCUUCUGACCUUA  
UUUAAUGUUUGUGUGUUUUUGUUCUUUGCUUUUJCCUGCCAUGGUUUUAACCCUAUGUAGACCUCAGUUUCAU  
UUCUAUAGAGCUGAUUUCAUUCUGCCAUGCAGA

#### D 4.5: NUCLEOTIDE SEQUENCES OF PRE-miRNAs OF miRNA1122 FAMILY MEMBERS USED FOR THE CONSTRUCTION OF PHYLOGENETIC TREE

##### >asa-mir1122 (*Avena sativa*)

AAUGCUAGUACGAAUUUUUGAACUACUCCUCCGUCACAAAAAGGAUGUCUCAAGUUUGUUUGAAUUUGGACGAU  
CUAUUAUUUUAAUAGUGUCUAGAUACAUCCAAGUUUAGACAAAUUUGAUACAUCUUUUCGGGACGGAGGGAGUAC  
AUGUUUGGACAGGUUGUGGUGAAACCG

##### >ata-mir1122 (*Arabidopsis thaliana*)

AAUUGUUCAUGUUGUGCUGUAUUUAUCUCCUCCGUCCGAAUUUUUUGUCGGAGAAAUGGAUGUAUCUAGACG  
UAUUUUAGUUCUAGAUACAUCCAUUUUUAUCCAUUUCUGCAACAAGUAGUUCGGACGGAGGGAGUAUCAUUUA  
CAAUAUAUGCAAUGUUCGA

##### >bdi-mir1122 (*Brachypodium distachyon*)

UCGUCCAUGUAUCCAUAUUUCAUUGCCGAAUAUUACAUGUGUAUCUAGACGUUUUUUACACAUAGAUACAUCCGU  
AUUUGGACAAAUUUGAGACGAGUAUUAGAGAUCCGGAUGGAGUACA AUUUUUUCCCUUCUUAUCUGGUCACCACA  
UACACAAAUUUGCACAGAACUGAACCCAG

##### >bsy-mir1122 (*Bachypodium sylbaticum*)

AUCUUAUGCCAUUUAGCAAAGCUGUACUCCUCCGUCCAACAAAAGAUGUCUCAAGUUUGUCAAAAUUUGGAUGU  
AUCUAGACAUGACUUAGUGUAUAGAUGCAUUCAAUUUAGUCAAAAUUUGAGACAUCUUUGUUGGACGGAGGAAG  
UACCAUGUUUAAUUUACCUCAUUUAAGAUG

##### >hpi-mir1122 (*Heterantherium piliferum*)

ACUUGUUGAAAUCUCUAGAAAGACAAUACUCCUCCGUCCCAAAAUUCUUGUCAUAGAUUUGUCUAAAUACGGAU  
GUAUCUAGUCAUGUUUUAGUGUUAGACACAUUCAUAUCUAAACAAAUCUAAGACNAGAAUUUUGGGACGGAGGA  
GUAUUUAUGAACGGAAGGAGUAGUUGCUA

##### >tae-mir1122 (*Triticum aestivum*)

AUAAGUACUCCUCUGUCCUAAAUACUCCUCCGUCCCAAAAUUUUGGCUUAGAUUUGUCUAAAUACGUAUGUA  
UCAAGUCAUGUUUUAGUAUUAGAUACAUCCGUAUCUAGACUAAUCUAAGACAAGAAUUUUGGGACGGAGGGAGUA  
UAAGUCUUUUUAGAGAUUCCACUACAACUACAUCGGAUGUAUAUAGACAUUUUUUAGAGUGUAGAUUCACUCAU  
UUUACUUCGUAUGUAGUCCAUGUGGAAUCUCUAAGAAGGCAUAUAUUUAGGAACAGAGGGUGUACCA

**>tmo-mir1122 (*Triticum monococcum*)**

UUUUGGCCUACAUAGCCCACUUUUGUACUCCUCCGUCCCAAAAUUGUUGUCUUAAAAUUUGUCUAAAUAUGGAUG  
UAUGUAGUUACAUUUUAGUGUUAGAUACAUCCGUAUCUAAACAAAUGUAAGACAAGAAUUUUGGGAUGGAGGGAA  
UACUACCUCAUUUUAGUAUAAAAUUUGA

**>tur-mir1122 (*Triticum turgidum*)**

CAAAUUUGCAGGUAACAUGAAAUGUACUCCUUUUGUCCUAAAAUUUGUGUCUUAGAUUUUGUCUACAUACGAAUGU  
AUUUUAUUCACAUUUUAGUGUUAGAUACAUCGUAUUUAAAAUUUUUAAAGAUAAAGAAUUUUGGACAGAGGAAGU  
AUUAACAAUCCUUAUAAAACUUUAAACUU

**>cas-mir1122 (*Camellia assamica*)**

GCUAUACUCCUCCGUCCCAAAUAUUGGUCCUUUUGGGGAUCCAACUUUUUUUGGUAGUACAAUUAUUGUACU  
UUUAAAAAGUUGGUUUUCCAAAAUUACCCUUUGAUAGUACACUUUUUUGAGAAUAUCAUAAUUGUACUUU  
UUUAUUUAUUUAUGAAUUUAAAAGAAGGGCAAAUUGGAAAUUUAUUUUUUUGUACUUGAUUUUUUGAAGAGG  
GACUAAUUAUUUGGAACAAUGAAAAUUGGAAUAAGGGACUAAUUAUUUGGGACGGAGGGAGUACAAGU

**APPENDIX-E**

**LIST OF PUBLISHED/COMMUNICATED  
RESEARCH PAPERS**

## E. LIST OF PUBLISHED/COMMUNICATED RESEARCH PAPERS

- Pranay Bantawa, **Akan Das**, Partha Deb Ghosh, Tapan Kumar Mondal. Detection of Natural Genetic Diversity of *Gaultheria fragrantissima* Landraces by RAPDs: An Endangered Woody Oil Bearing Plants of Indo-China Himalayas. Journal of Biotechnology, 2010 (accepted).
- **Akan Das**, Tapan K Mondal (2010) Computational identification of conserved miRNAs and their targets in tea (*Camellia sinensis*). American Journal of Plant Sciences 1: 77-86.
- **Akan Das**, Dipanwita Saha, Mridul Hazarika, Sudripta Das, Tapan K Mondal (2010) Isolation of RNA from the root of tea (*Camellia sinensis* L.) for downstream molecular biology work. Plant Biotechnology Reports (communicated).
- **Akan Das**, Pranay Bantawa, Mainak Mukhopadhaya, Dipanwita Saha, Tapan K Mondal (2010) Differences in antioxidant responses of sensitive and tolerant cultivars of tea (*Camellia assamica* L.) to drought stress. Acta Physiologiae Plantarum (communicated).
- **Akan Das**, Sudripta Das, Tapan K Mondal (2011) Analysis of differentially expressed transcriptome of tea (*Camellia assamica* L.) under drought stress. Tree Genetics and Genomes (communicated).
- Mainak Mukhopadhaya, **Akan Das**, Pranay Bantawa, Pratap Subba, Bipasa Sarkar, Partho D Ghosh, Tapan K Mondal (2011) Evaluation of antioxidant responses and gene expression patterns in *Camellia sinensis* L. under zinc stress. Journal of Plant Nutrition and Soil Science (under preparation).

