Development and analysis of gene resources generated from drought stressed seedlings of *indica* rice *(Oryza sativa* .L*)*

A thesis submitted for the degree of
DOCTOR OF PHILOSOPHY

By
P. RAVINDRA BABU

Dept. of Plant Sciences
School of Life Sciences
University of Hyderabad
Hyderabad 500 046
Andhra Pradesh, India

Enrollment No. 98LPPH02
March 2004

Department of Plant Sciences
School of Life Sciences
University of Hyderabad
Hyderabad-500 046
INDIA

CERTIFICATE

This is to certify that **Ravindra Babu Peram** has carried out the research work embodied in the present thesis entitled "**Development and analysis of gene resources generated from drought stressed seedlings of *indica* rice *(Oryza sativa .L)*"** for the degree of **Doctor of Philosophy** under my supervision in the Department of Plant Sciences, School of Life Sciences, University of Hyderabad.

This work has not been submitted for the award of any degree or diploma of any other University or Institute.

Prof. A. Ramachandra Reddy
Supervisor
Department of Plant Sciences

Head
Dept. of Plant Sciences

Dean
School of Life Sciences

25/2/04

DECLARATION

I hereby declare that the work presented in this thesis entitled **"Development and analysis of gene resources generated from drought stressed seedlings of *indica* rice *(Oryza sativa* .L)"** has been carried out by me under the supervision of **Prof. Arjula Ramachandra Reddy** in the Dept of Plant Sciences, School of Life Sciences, University of Hyderabad and that this work has not been submitted for any degree or diploma of any other University or Institute.

P. Ravindra Babu
Ravindra Babu Peram

Prof. A.R. REDDY
Supervisor
Dept. of Plant Science
University of Hyderabad

Prof. A. Ramachandra Reddy
Supervisor
Department of Plant Sciences

Acknowledgements

It gives me an immense pleasure to express my deep sense of gratitude to **Prof. Arjula R. Reddy** for his supervision to carry out my doctoral work. I am thankful to Prof. A.R.R for introducing me to the genomics, bioinformatics, and take strides into interdisciplinary fields of science. I am indebted for his constant encouragement, support, and inspiring discussions. I thank Mrs A.R.Reddy and their family for the support.

I thank former Deans of Scholl of Life Sciences, **Prof. R. P. Sharma, Prof. T. Suryanarayana** and the present Dean **Prof. A. S. Raghavendra** for providing the facilities in the school. I would like to extend my thanks to former Heads, Department of Plant Sciences, **Prof. A. S. Raghavendra, Prof. M. N. V. Prasad** and the present Head **Prof. P.B. Kirthi**, for their support in the Department of Plant Sciences. I thank all the faculty members of the school for sharing their thoughts and views in the scientific endeavor. I thank the software developers, which were used, in my work. I thank all the rice research groups, specially International Rice Genome Sequencing Program for open access of their data to the scientific community.

I would like to thank my seniors Dr. B. Karuna Sree, Dr. M. Padmavathi, Dr. G. Madhuri, Dr. Nagabhushana Ithal, Dr. Rajendra Kumar, and B. Chadra Sekhar for their collegial nature and guidance. I wish to thank the present lab members L. Sivram Prasad, G. Markandeya, A. Chandra Shekar, A.Madan Mohan Reddy, V. Kiran Kumar, L. Vijay Bhaskar Reddy, Linga Reddy and B. Venkat Reddy for their support and help.

I thank my friends Dr. Lakxmi Narayana, Dr. Bhaskar, E. V. Subbaiah, Y. Sailu, and B. Venkataramana for their moral support and encouragement. I thank Sravan, Sudar, Mahipal Sr, Mallikarjun Reddy, Mahipal Jr, Raju, Vijay Bhaskar Reddy, and Ashok for the moments shared in the University.

I would like to thank Non-teaching staff Misra, Lalan, Srinivas, Raju and Sivram Mohan for their help in the Department. I wish to thank the assistance of Mr. Ravinder and Mr. Poorna in the lab.

I wish to acknowledge the support of Rockefeller foundation project grant to Prof. A.R.R and the financial support received by me as JRF and SRF from this grant is highly appreciated.

I thank all my family members for their support and encouragement.

Ravindra Babu

CONTENTS

List of Tables:

List of illustrations

Abbreviations

ABA: Abscisic acid

Mb: million base pairs

ESTs: Expressed Sequence Tags

NCBI: National center for biotechnogy information

DREB: Dehydration Responsive Element Binding protein

LEA proteins: late embryogenesis abundant proteins

ROS: reactive oxygen species

RWC: relative water content

UTR: untranslated region

PHRED: Phill read length

CAP: contig assembly program

Phrap: phragment assembly program

BLAST: basic local alignment search tool

BAC/PAC: bacterial artificial chromosome /P1-derived artificial chromosome
HTGS: high throughput genomic sequences

PLACE: plant *cis*-acting elements

PEG: Poly Ethylene Glycol

QTL: Quantitative Trait Loci

SNP: Single Nucleotide Polymorphism

SSR: Simple Sequence Repeat

MT: metallothionein

GAAS: Rice Genome Automated Annotation System

MAPK: Mitogen activated protein kinase

Na+/H+ antiporters: sodium/proton antiporters

SAGE: Serial analysis of gene expression

CBF: C-repeat binding factor

RGP: Rice Genome Program

IRGSP: International Rice Genome Sequencing Project

DHLs: Doubled Haploid lines
RILs: Recombinant Inbred lines

1. Introduction

Rice is the most important cereal crop providing staple food for about half of the world's population (Khush 1997). Rice has emerged as the model crop system to understand basis of complex traits such as yield, hybrid vigor, disease resistance, and aboitic stress tolerance. It has now become central to cereal genomics in view of its small genome size, high-resolution genetic maps, and syntenic relationship to other agronomically important cereal species (Devos and Gale 1997), which contribute the major caloric energy source to human consumption. Rice crop occupies about one-third of the world's total area planted to cereals and provides 35–60% of the calories consumed by 2.7 billion people. More than 90% of the world's rice is produced and consumed in Asia (Barker and Herdt 1985, IRRI 1989). Rice is the most widely grown crop under irrigation. More than 80% of the available freshwater resources in Asia are used for irrigation purposes and more than 90% of this water is used for rice production (Bhuiyan 1992). Rice is intimately linked to the food security in most of the developing countries of Asia.

Rice is grown under almost all major agro-climatic conditions across tropical and temperate countries. About 80% of the world's rice is grown under irrigated (55%) and rain fed lowland (25%) ecosystems both of which depend on fresh water resources. The limitation of water availability due to inadequate and erratic rainfalls has contributed to as much as 15% loss (Dey and Upadhaya, 1996) in rice production, twice the damage caused by biotic stresses. Drought stress is the major cause of rice yield instabilities across diverse crop growing regions of Asia. Breeding for drought tolerance has been, and continues to be, a difficult task as drought tolerance is a complex trait controlled by many genes exhibiting incomplete penetrance, quantitative inheritance, and epistatic interactions. Though conventional breeding and genetic studies of drought tolerance yielded rich dividends and valuable information respectively, progress has been rather slow and inadequate due to the complexity of the trait. Furthermore, drought stress response in plants involves an array of different pathways associated with stress perception, signal transduction, gene expression and regulation, leading to synthesis and accumulation of a variety of bio-molecules. Recent developments in functional genomics facilitated the assessment of gene function and expand the scope of biological investigation from studying single genes or proteins to studying all genes or proteins at once in a systematic fashion. This is naturally a preferred approach to unravel the genetic determinants associated with abiotic stress tolerance in plants (Cushman and Bohnert, 2000) and apply to breeding procedures.

The rice genome is estimated to be 430 Mb, roughly four times that of *Arabidopsis* genome. Gene predictions based on the assembled draft genome sequence of rice (Goff *et al.*, 2002; Yu *et al.*, 2002) estimated to have about 32,000 to 50,000 genes. To verify these *in silico* gene predictions based on different models, Expressed Sequence Tags (ESTs) deposited in public domains such as NCBI, were effectively used. Expressed sequence tags generated from high throughput transcriptome sampling represent a snapshot of gene expression during defined developmental stages of rice growth cycle and following specific biotic and abiotic challenges. The potential of ESTs as a valuable gene resource and their applicability to elucidate gene function was rather underestimated previously. With the availability of the genome resources and other allied technologies, ESTs have become essential component for genome analysis, novel gene discovery and accurate genome annotation in many plant systems.

In rice, large-scale ESTs have been generated from cDNA libraries of different callus tissues or organs such as panicle, root and leaf at different developmental stages, environmental stress conditions and hormone treatments. In general, these libraries represent abundant transcripts and usually missed transcripts of low abundance. Owing to the complex gene expression patterns in plants, such libraries may totally miss rare transcripts. This becomes a critical factor in stress responsive gene expression studies. On the other hand normalized cDNA libraries (Bonaldo *et al.*, 1996) greatly reduced clone redundancy and thereby increased the probability of detection of less abundant transcripts. By utilizing this approach, normalized cDNA libraries were constructed (Reddy *et al.*, 2002) from drought-stressed rice seedlings in order to discover genes that control and mediate gene expression under abiotic stress. Differential libraries were also proved to be useful in many instances in uncovering genes controlling this trait in rice and other plants. Molecular and functional genomic studies in model plants, *Arabidopsis* and rice have identified a host of genes induced by abiotic stresses (Bray *et al.*, 2000; Bohnert *et al.*, 2001; Zhu *et al.*, 2001; Seki *et al.*, 2002; Kawasaki *et al.*, 2001; Kreps *et* *al*., 2002). However, the mechanisms of stress responsive gene expression and stress tolerance are yet to be unequivocally elucidated.

To uncover the stress transcriptome associated with drought-stress response in rice, a true representative of monocot lineage diverged from dicots150 millions years ago (Brendel *et al*., 2002), we have analyzed the large-scale gene resources generated from normalized cDNA libraries of drought-stressed seedlings of rice and took a comparative genomic approach to identify candidate genes of drought stress response.

The present study mainly deals with the identification of putative candidate genes associated with drought-stress response in rice through large-scale cDNA analysis and comparative genomic approaches. The objective is to identify genes whose expression is related to the plant survival and adaptation to drought stress.

Objectives

- High-quality rice EST data analysis and submission to the public database
- Analysis of EST sequence profiles
- Development of EST database
- Annotation and mapping the EST s to rice genome sequence.
- Identification of stress responsive genes of rice through homology searches of documented genes of other plants presumed to be associated with abiotic stress.
- Defining the possible candidate gene sequences deduced from expression profiling analysis using the EST dataset.
- Analysis of the gene organization and identification of putative candidate genes at target quantitative trait loci (QTL) associated with drought stress response in rice.
- Cloning of a full-length DREB transcription factor associated with drought-stress response and analysis of organization of the DREB gene family in rice.

2. Literature Review

The variation in the genome size of different organisms attribute limitation to sequence the large genomes and have taken alternative to complement the whole genome through expressed sequence tags (ESTs). Expressed sequence tags (ESTs) are currently the most widely sequenced nucleotide commodity from the plant genomes in terms of the number of sequences and the total nucleotide count (Table 2.1). ESTs provide a robust sequence resource that can be exploited for gene discovery, genome annotation and comparative genomics.

The genomic tools developed initially in the Human Genome program have been applied to plant systems recently through sequencing the genome of model plants and other economically important plants. The results of plant breeding over millennia provide all food, feed and fiber for human use. The population growth, water resources and global environmental changes lead to projections to identify additional strategies to sustain the crop productivity. Much of the yield increase of wheat, rice, and maize from 1967 to 1997 is attributed to plant breeding and selection for high yielding genotypes; however, there has not been an increase in yield potential (Cassman, 1999; Duvick and Cassman, 1999; Peng *et al.*, 2000; Tollenaar and Wu, 2000). In other words, when grown in ideal environments where pests are controlled, resources are not limited, and changes in the environment are minimal, old and new cultivars have the same yield. Therefore, it is clear that understanding plant response to abiotic stress is important at a very applied level of plant improvement. There is also great potential to improve stress tolerance through genetic engineering. Understanding the abiotic stress response and tolerance as a biological phenomenon proved difficult because of trait multigenicity. The genetic studies identified the phenotypic determinants of the organisms as single genes with major affect or polygenes with minor affects along with the environmental component.

To unravel the genetic determinants of the complex traits such as abiotic stress tolerance in plants, which is associated with genes and gene networks to exhibit adaptive response, generation and analysis of ESTs has taken as fruitful approach (Cushman and Bohnert, 2000). Expressed sequence tags generated from high throughput transcriptome sampling represent a snapshot of gene expression during defined developmental stages and following specific biotic and abiotic challenges. The molecular analysis of these transcripts collected as ESTs and the application of these abstractive resources were underestimated previously. With the availability of the genome resources, ESTs have become essential component for genome analysis, novel gene discovery and accurate annotation.

Table 2.1: **Genome size variation in cereals and model plants**

Species	2n	Ploidy	1C value (pg)	% Repetitive DNA	Size Mb
Oryza sativa	24	2×	0.4–0.5	58–66	430
Hordeum vulgare	14	2×	5.1–5.9	>70	5000
Triticum aestivum	42	6×	16.5–19.5	>75	17000
Triticum durum	28	4×	13.2–13.8	>75	
Triticum monoccocum	14	2×	4.0–6.0	>75	5700
Sorghum bicolor	20	2×	0.7	na	1000
Secale cereale	14	2×	8.3–10.5	>75	2500
Zea mays	20	2×	2.4–3.0	60	2670
Tomato					950
Arabidopsis					125
Avena sativa (oat)					26000

The present literature focus on the recent utilization technological advances towards analyzing the global gene expression profiles in understanding complex biological phenomenon such as abiotic stress tolerance in plants and the impact of unraveling the gene functions.

2.1 Transcriptional profiling

2.1.1 Expressed Sequence Tags:

ESTs are partial sequences of cDNA, reversely transcribed from mRNA, and represent a direct supply of coding intron-free sequences of genes. 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 expressed sequence tag (EST) collections (Table 2.2). The relative cheapness of EST sequencing and its associated automation often make EST sequencing the most attractive route for broad sampling of the transcriptome compared to other existing technologies. Large-scale Expressed Sequence Tags have been analyzed from many plant species (Cooke *et al.*, 1996; Shen *et al.*, 1994; Yamamoto and Sasaki, 1997) in view of their utility as powerful tools to catalogue all the genes (Adams *et al.*, 1991; 1995) and deciphering the roles of transcriptionally regulated genes in different tissues. The analysis of the datasets from *Arabidosis* and rice revealed similar patterns of EST abundance, and supported the validity of the digital-northern approach. Few studies have focused on the analysis of transcriptome profiles of rice seedlings subjected abiotic stresses like salt and cold (Kawasaki *et al.*, 2001; de los Reyes *et al.*, 2003).

A special database, dbEST (Boguski *et al.*, 1993), has been established to handle the large amount of ESTs produced. To date dbEST contains 20,004,224 public entries of which more than 25% of the sequences represent from plant kingdom.

dbEST release 020604 http://www.ncbi.nlm.nih.gov/dbEST/dbEST_summary.html

Table 2.2: Number of EST entries for cereals and model plants in dbEST division

Species	**No of ESTs**
Triticum aestivum (wheat)	549,926
Zea mays (maize)	391,145
Hordeum vulgare + subsp. vulgare (barley)	352,924
Oryza sativa (rice)	267,943
Saccharum officinarum	246,301
Sorghum bicolor (sorghum)	161,766
Secale cereale	9194
Triticum monococcum	9973
Triticum turgidum subsp. durum	7847
Avena sativa (oat)	509
Arabidopsis thaliana (thale cress)	196,988
Glycine max (soybean)	345,723
Medicago truncatula (barrel medic)	187,763
Lycopersicon esculentum (tomato)	152,900
Solanum tuberosum (potato)	132,304

ESTs were extensively used as molecular markers for the construction of high-density genetic linkage maps of rice and maize (Harushima *et al.*, 1998; Davis *et al.*, 1999) and for a physical map of rice (Kurata *et al.*, 1997) and comprehensive YAC-based rice transcript map (Wu *et al.*, 2002). Furthermore the sequencing data can be used to study gene families (Cooke *et al.*, 1997; Epple *et al.*, 1997) and they form a basis for SNP development (Cho *et al.*, 1999). Apart from applications in the field of genetic and physical mapping, ESTs are the central resource for the analysis of gene expression with the help of high-density arrays, as demonstrated for *Arabidopsis* (Schena *et al.*, 1995; Girke *et al.*, 2000; Schenk *et al.*, 2000), barley (Ozturk *et al.*, 2002) maize (Wang *et al.*, 2003), and rice (Kawasaki *et al.*, 2001; Rabbani *et al.*, 2003).

2.1.2 SAGE analysis:

Serial analysis of gene expression (SAGE) is also sequence-based approach allowing the identification of a large number of transcripts present in tissues and the quantitative comparison of transcriptomes (Velculescu *et al.*, 1995). The principle of SAGE is to generate a short specific tag (14 bp) from each mRNA present in a sample, resulting in the production of a SAGE tags library representative of this sample. The sequencing of these tags allows a high-throughput determination of their frequencies in the library, which are correlated with the relative amounts of the corresponding mRNAs. SAGE has proven to be a very powerful and robust method for investigating gene expression at the whole-genome scale (Velculescu *et al.*, 1997; Boon *et al.*, 2002; Liang, 2002) and to reflect the actual relative contents of mRNAs in a sample (Chrast *et al.*, 2000; Piquemal *et al.*, 2002; Jung *et al.*, 2003; Matsumura *et al.*, 2003).

The major limitation of SAGE is that in most species, tag to gene assignment is based on EST clusters or on available cDNA sequences. This results in very incomplete identification of the transcripts revealed by SAGE tags, leaving many of them without any match in the databases (Lash *et al.*, 2000; Boheler and Stern, 2003; Pleasance *et al.*, 2003). SAGE in plants faced the same problem, and gene identification using cDNA or EST databases was not possible for up to 70% to 75% of the SAGE tags obtained from rice (*Oryza sativa*) plants and *Arabidopsis* leaves (Matsumura *et al.*, 1999, 2003; Jung *et al.*, 2003).

The described EST- and SAGE tag-sequencing methods result in complete gene expression profiles. Not only genes that are present or absent in different tissues can be studied, but also up- and down-regulations, which is a great advantage as compared to techniques for differential expression which enables isolation of few candidate genes through differential display (Liang and Pardee, 1992), RNA fingerprinting by arbitrarily primed PCR (Welsh *et al.*, 1992), representational difference analysis (RDA) (Hubank and Schatz, 1994), subtractive hybridization (Wang and Brown, 1991) and differential screening of arrayed cDNA clones (Byrne *et al.*, 1995).

2.2 DNA sequence analysis and annotation

Biological research is now generating sequence data at an explosive rate. Management and analysis of the enormous amounts of data will require powerful computational resources. New software and hardware is needed for efficient data processing, assembly and annotation, as well as for gene sequence predictions and functional and structural classifications. Database integration is important to access all kinds of data related to the sequence and the database entries must be automatically updated regularly. Smaller highly annotated data sets for first-pass analyses will be essential to reduce search times.

2.3 Abiotic stress responsive genes and stress tolerance

Different approaches were utilized to manipulate the mechanistic end-point of stress tolerance such as over expression of superoxide dismutase in order to detoxify oxygen radicals produced under stress (McKersie *et al.*, 1996; Roxas *et al.*, 1997). Transgenic plants designed to synthesize high levels of osmoprotectants show elevated levels of stress tolerance, but often suffer from deleterious pleiotropic effects such as dwarfing (Tarczynski *et al.*, 1993; Romero *et al.*, 1997). Several other pathways important to plant abiotic stress response have been relatively well described such as ion homeostasis and salt tolerance (Hasegawa et al., 2000; Zhu, 2002); Na+/H+ antiporters (Apse *et al.*, 1999), ABA response pathway, (Leung and Giraudat 1998); lipid and MAPK signaling pathways (Munnik and Meijer 2001; Jonak *et al.*, 2002). Transcript profiling will improve our understanding of each of these pathways and provide potential avenues for plant improvement. Precise regulatory mechanisms of such pathways need to be understood to effectively deploy the superior alleles in crop improvement for drought tolerance. Transcript profiling is potential tool for identifying candidate genes.

2.4 Analysis of gene expression profiles

Model experimental systems in plants such as *Arabidopsis* and rice are highly amenable to gene expression profiling particularly dealing with abiotic stress. There are now several examples of plant abiotic-stress-related transcriptome profiling that have revealed many new components in stress response pathways (Kawasaki *et al.*, 2001; Desikan *et al.*, 2001; Ozturk *et al.*, 2002; Chen *et al.*, 2002; Kreps *et al.*, 2002; Fowler and Thomashow 2002; Seki *et al.*, 2001, 2002; Rizskhy *et al.*, 2002; Klok *et al.*, 2002; Kim *et al.*, 2003; Yu and Setter 2003). The results reported thus far were derived very differently by using various stress treatments, array formats, species, tissue types, and time courses making it difficult to make direct comparison among studies.

It is now clear that changes in gene expression in response to abiotic stress include both general and specific to a particular stress. The extent of overlap is becoming increasingly clear. It is likely that the experience of multiple stresses by plants more closely mimic field conditions than single stress treatments. These results suggest that transcript profiling using a combination of abiotic stresses may better mimic field conditions and further reveal novel patterns of gene expression. Plants make use of common pathways and components in the stress response relationship. This phenomenon, which is known as cross-tolerance, allows plants to adapt/acclimate to a range of different stresses after exposure to one specific stress. Some stresses themselves are tissue-specific. For example, the roots will perceive soil source abiotic stress, whereas extremes in air temperature will be perceived by the aboveground biomass. It is also true that physiologically, roots and shoots respond to drought stress differently. Roots can continue elongation while at the same time shoot elongation is completely inhibited. The consequence of this differential response is that roots can reach water in deeper soil while inhibition of shoot elongation conserves water consumption (Sharp *et al.*, 1988). Kawasaki *et al.*, (2001) described four stages of change in gene expression across time in salt-stressed rice plants: instantaneous response, early response, early recovery, and stress compensation.

The analysis of plants with a dysfunctional allele (a knock-out) derived by insertional mutagenesis will be of immense value in deciphering molecular mechanisms of stress resistance. Through extensive efforts of both private and public institutes, collections with a knock-out in every gene in the genome (not including lethal mutations) will be available in the near future in rice (http://tos.nias.affrc.go.jp/~miyao/pub/tos17/), maize, (http://www.zmdb.iastate.edu; http://mtm.cshl.org), and *Arabidopsis* (http://www.biotech.wisc.edu; http://signal.salk.edu; http://www.tmri.org.com).

For genes not represented in the mutant collections, post-transcriptional silencing via antisense or dsRNA may provide the necessary genetic material. The use of experimentally induced promoters can also facilitate and increase the precision of this approach. It is not uncommon for knock-out mutants to reveal no detectable mutant phenotype or have a lethal effect. In these circumstances, over expression of candidate genes may imply function. Mutant and over expression transgenic plants are also very useful in revealing gene interactions within complex transcriptional pathways. For example, an increase in expression of ABA-responsive genes was observed from the transcript profile of the *abh1* mutant, which lacks a functional copy of a negative regulator of the ABA response pathway. Moreover, gene expression profiling also identified several new genes in this pathway (Hugouvieux *et al.*, 2001). This approach identified downstream elements that would normally respond to a functional copy of the mutant allele and exclude other candidates. In another example, transcript profiling (~1,300 genes) of *Arabidopsis* plants over-expressing *DREB1a/CBF3* identified 12 genes, 6 of which were previously unknown to be targets of *DREB1a* gene (Seki *et al.*, 2001). A more extensive analysis using an array of ~8,000 genes and ecotype Wassilewskija-2 plants over-expressing one of three regulators of the cold responsive pathway, *CBF1*, *CBF2*, or *CBF3*, identified 306 genes responding to cold treatment (Fowler and Thomashow, 2002). Interestingly, 28% of the cold responsive genes were identified as not belonging to the CBF regulon, suggesting they belong to unidentified cold stress response pathways.

2.5 Transcriptional regulation of stress-responsive genes (SRG)

The study of the promoter regions of stress responsive genes can be considered a key step for the comprehension of the molecular mechanisms leading to gene activation. Sequence analysis has revealed that different motifs are required for stress activation (*cis*-acting elements). Each element responds to a specific hormone or environmental stimulus, therefore the combination of different factors confers specificity to the stress-related gene expression. Functional dissection of ABA-inducible gene promoters revealed specific ABA-responsive elements (ABREs) containing the ACGT core element. In cereals different ABREs have been found. The first examples were the wheat *Em1a* ABRE (Marcotte *et al.*, 1989) and *Motifl* ABRE from rice *rab16* (Mundy *et al.*, 1990). In *Arabidopsis*, a drought responsive element (DRE) corresponding to the sequence TACCGACAT, has been found in the promoter region of several stress-related genes (*lti78/rd29a* and *lti76/rd29b*) (Yamaguchi-Shinozaki and Shinozaki, 1994). The DRE controls the cold and drought induction of *rd29a* in an ABA-independent manner.

Common motifs have been observed upstream of the coding regions of stress-induced genes. Based on the expression profile of *Arabidopsis* root cultures treated with low oxygen across four time points, genes were clustered into six groups using multivariate statistical analysis (Klok *et al.*, 2002). Common motifs specific to each cluster were observed in the 500 bp upstream of the coding regions. Among genes up regulated in *Arabidopsis* treated with saline, drought, or cold, a majority had the DRE-related core motif or the abscisic acid-response element ABRE (Seki *et al.*, 2002). A similar but weaker association of the DRE with the promoters of stress-regulated genes was found in similar studies using a more comprehensive array (Chen *et al.*, 2002; Kreps *et al.*, 2002).

2.6 Construction of Metabolic and regulatory networks associated with stress response through global gene expression analysis.

In general, tolerance to abiotic stresses is associated with a host of morphological and physiological traits; these include root morphology and depth, plant architecture, variation in leaf cuticle thickness, stomatal regulation, osmotic adjustment, antioxidant capacity, hormonal regulation, desiccation tolerance (membrane and protein stability), maintenance of photosynthesis, and the timing of events during reproduction (Bohnert *et al.*, 1995; Shinozaki and Yamaguchi-Shinozaki 1996; Bray 1997;Nguyen *et al.*, 1997; Edmeades *et al.*, 2001). Several genes and pathways involved in the expression of above traits in stress tolerance have been identified.

The metabolic pathways are complex in several ways. Within a cell, a thousand or more metabolites are distributed among several compartments. Expression patterns of a host of genes belonging to different signaling events associated with abiotic stresses such as ABA mediated genes, oxidative stress responsive genes and transcription factors under cold and pathogen attack were analyzed using microarrays to identify candidate genes and their role in stress resistance.

2.7 QTLs and candidate genes associated with abiotic stress resistance:

The recent advances in the genetic analysis and molecular mapping have led to the identification of a great number of single loci, quantitative trait loci (QTLs) associated with abiotic stresses. Many stress-related genes have been mapped and some of them have been shown to co-segregate with stress tolerance QTLs. However, only a few studies provide genetic evidence that the stress tolerance effect explained by a given QTL can be attributed to a co-mapping stress-related gene. In cowpea, the accumulation of the 35-kDa dehydrin was found to be involved in chilling tolerance during seedling emergence. Allelic differences in the coding region of the dehydrin structural gene map to the same position as the dehydrin protein presence/absence trait, which in turn is associated with chilling tolerance/susceptibility (Ismail *et al.*, 1999). These results also demonstrate that allelic variations in a stress-related gene can significantly alter plant stress tolerance ability.

In rice, 14 permanent populations have been reported so far, including two DHLs, nine RILs, and one BIL for QTL mapping with population sizes of 65 to 315 (Xu, 2002). Genetic maps used in QTL studies consists of 113 to 399 molecular markers. The populations of IR64/Azucena and CT9993 X IR62266 have been shared internationally for the mapping of many agronomic traits disease resistance, cold tolerance, and water-stress tolerance. QTL have been studies in multiple environments towards identification of shared or common QTLs. These studies revealed that the QTL with large phenotypic effect are shared more frequently than QTLs with minor effects. For example three of the 11 QTL identified for leaf rolling were shared in three trials with different drought intensities (Courtois *et al.*, 2000). The information provided by genetic mapping about the approximate location of the target loci would identify the possible candidate genes in the target region with the available wealth of genome sequence information. Fine mapping combined with sequence analysis could narrow the chromosomal region associated with quantitative variation down to a specific nucleotide change (QTN).

Comparative genomics has shown that not only the sequences of stress-related genes are conserved, but the genetic basis of stress tolerance also have a common origin, particularly in closely related genomes such as those of the *Triticeae* and rice. For instance, the map interval on chromosome group 5 containing the *Vrn-A1* locus was shown to be homoeologous to the region of rice chromosome 3 where the heading date QTL *Hd-6* has been mapped (Kato *et al.*, 1999). Van Deynze *et al.*, (1995) also identified a possible relationship between the heading date QTL *Hd3a* of rice and the *Vrn-H2* locus of barley on chromosome 4H. A QTL analysis of chilling tolerance in rice seedlings led to the identification of thirteen QTLs located on rice chromosomes 1, 3, 9 and 11 (Misawa *et al.*, 2000). It is known that rice chromosomes 3, 9 and 11 show large regions of synteny with the homoeologous group 5 chromosome of the *Triticeae*, and three cold tolerance QTLs on rice chromosome 3 are located in the same syntenic interval where *Vrn-1A*, *Fr1* and *Fr2* in wheat and QTLs for cold tolerance in barley have been mapped. Interestingly, the genomic region controlling the vernalization response and frost tolerance in winter cereals is orthologous to a region controlling the chilling tolerance in a tropical plant such as rice (Misawa *et al.*, 2000).

The comprehensive physical map of rice genome based on large-insert, low-copy number bacterial clones, namely bacterial artificial chromosomes (BACs) and P1-derived artificial chromosomes (PACs) fingerprints, have been constructed (Chen *et al.*, 2002). RGP high-density genetic map (Harushima *et al.*, 1998) containing more than 2000 well-mapped genetic markers was utilized to integrate the physical map. Many of these markers are conserved among the grass genomes because they represent expressed genes (cDNAs or ESTs). Utilizing the physical-genetic map of rice, high-resolution comparative physical maps of cereal species can be developed. This will facilitate map-based cloning of agronomically important genes in species with large genome sizes, such as maize, wheat, and barley, using rice as a surrogate.

Several candidate genes putatively involved in drought tolerance can be identified in the literature and gene databases (Skriver and Mundy 1990; Bray 1993; Ingram and

Bartels 1996; Bohnert *et al.*, 2000; Seki *et al.*, 2001). The candidate genes potentially involved in the plant's response under water-limited conditions can be divided in three classes: (1) genes for which very strong evidence has already has been published, which demonstrates their significant role under abiotic stresses conditions (Finkelstein and Lynch 2000); (2) genes that have been identified as being of interest, but which require further evaluation (Liu *et al.*, 1998); and (3) those genes that have not been evaluated in a plant, and therefore still need to be confirmed under experimental conditions (Guiltinan *et al.*, 1990).

QTL mapping, gene cloning, ESTs and genome sequencing projects have led to a vast body of genetic information in public databases supplying the scientific community with powerful tools for comparative genomics (Gai *et al.*, 2000; Mekhedov *et al.*, 2000). The integration of genetic information from related species lead to the identification of highly conserved sequences and/or regulatory mechanisms by which it is possible to predict function and location of genes in different organisms that have been traditionally studied separately. The analysis of the stress response of different plant species by sequence comparison of stress-related genes and of *cis-* or *trans*-acting elements or by looking for conserved positions of stress tolerance loci among related genomes are leading to identify the genetic determinants of stress tolerance.

3. Materials and Methods

3.1 Materials

3.1.1 Sequence repositories and software resource used in EST analysis.

EST sequences are from *O. sativa* subsp. *indica* cultivar N22 (Fig.3.1) generated from drought stressed seedlings. Rice genome sequence of the *O. sativa* subsp *japonica* cv Nipponbare generated by the International Rice Genome Sequencing Project (IRGSP) and draft sequence of the *O. sativa* subsp. *indica* cultivar (93–11) generated by the Beijing Genomics Institute (BGI) available in the GenBank. The rice full-length cDNA consortium sequences of Nipponbare cultivar. Full-length cDNA sequences of possible candidate genes derived from *Arabidopsis* expression profiling studies from The *Arabidopsis* Information Resource (TAIR). The nucleotide, protein and EST databases at NCBI were utilized for homology search using BLAST program.

Standard sequence processing tools PHRED, Pharp, and Crossmatch were used with Codoncode InterPhace. Homology search in the NCBI database was carried out using network client software with the DNA Tools interface http://www.crc.dk/dnatools.

Genscan, GeneMark.HMM, RiceHMM, Glimmer R, FGENESH, Rice Genome Automated Annotation System (Rice GAAS) were used for accurate gene prediction. Sim4, LALNVIEW V2.0 were used for identifying gene structure. Plant CARE and PLACE databases were utilized for identifying cis-acting elements in the promoter regions.

3.1.2 Chemicals and reagents used in Molecular Biology

The chemicals and reagents used in molecular biology experiments were obtained from either U S Biochemicals (Amersham Pharmacia) or Sigma. Restriction enzymes and T_4 DNA Ligase were obtained from New England BioLabs. TA cloning kit, *Taq* polymerase and RNaseA were obtained from MBI Fermentas. DNA labeling kit and radio labeled $[\alpha\text{-}^{32}P]$dCTP were obtained from BARC India. Nylon membrane Hybond N+ was purchased from Amersham Pharmacia. Oligo nucleotide primers of HPLC grade purity were synthesized from MWG, Germany. Plasmid isolation and purification kits from Qiagen and Dye terminator ET sequencing reaction kit from Amersham Pharmacia were used in sequencing.

Plant material: Seeds of Nagina 22 rice cultivar were obtained from Directorate of Rice Research (DRR).

3.1.3 EST source: High-quality ESTs generated from normalized cDNA libraries constructed from drought-stressed seedlings of rice cultivar *N22* (Reddy *et al.* 2002). An outline of the EST generation has been shown here.

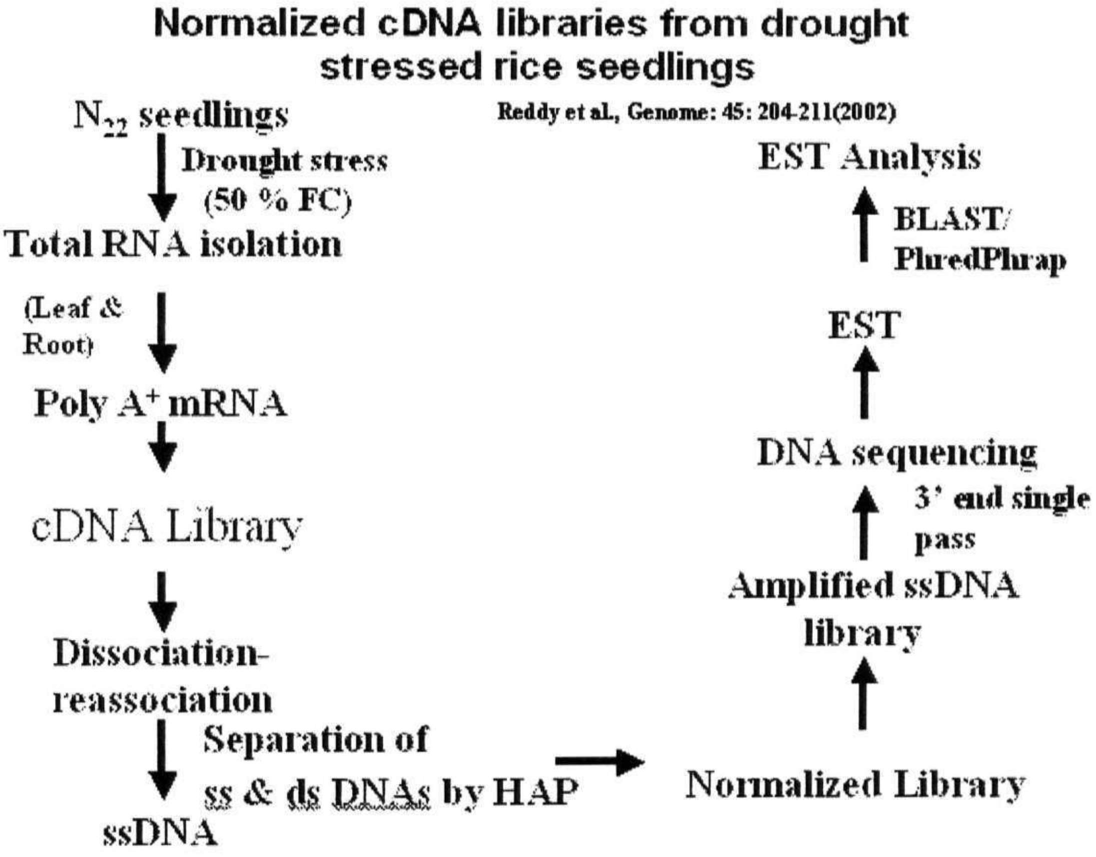


Fig. 3.1: Outline of EST resource used in the analysis

3.2 Methods

3.2.1 Sequence processing and analysis

High quality EST sequences were generated from sequence chromatograms obtained on MegaBACE 500 sequencer using PHRED (Ewing and Green, 1998). The Phred base-calling program provides a quality value for each base called. The quality value is based on the estimated probability that the base call is an error. Quality (Q) and estimated call error are related by the following formula.

$Q= -10 \times \log^{10}$ (probability error)

The estimates of probability of error are based on a number of factors, including peak shape, spacing between peaks, signal strength and background noise (Ewing et al., 1998; Ewing and Green, 1998). The major steps in the quality trimming process involve calculating some average probabilities of error within a certain window along the length of the sequence. The low quality regions preset at the beginning and end of each sequence were trimmed using Phred 20 cutoff value.

Vector screening was performed using CROSS_MATCH program with Codoncode InterPhace software. CROSS_MATCH is used to identify and "mask" vector sequences before sequence assembly. Sequence reads often contain small fragments of vector sequence before the cloning side, which can lead to assembly problems like too many contigs or wrong contig ends (Fig. 3.2). CROSS_MATCH can reduce or eliminate such problems by identifying vector sequences in all individual reads, and masking any vector sequence by replacing each matching base with an 'X'.

CROSS_MATCH is based on a fast implementation of the Smith-Waterman-Gotoh algorithm for local sequence alignments (Smith & Waterman 1981; Gotoh, 1982). In addition, CROSS_MATCH can constrain the comparison to "bands" around word matches between two sequences, which leads to large speed gains without a significant loss of sensitivity. CROSS_MATCH produces local alignments, in contrast to end-to-end alignments produced by variants of the Needleman-Wunsch algorithm. All matches of reads to vector sequences that have an alignment score of at least 20 and are one match of at least 12 bases were replaced by an "X" for each base in the matching region. CROSS_MATCH missed short vector sequences (shorter than 20 bases in the example above), as well as relatively short sequences with high error rates (for example, a 30-base match with one error every 5 bases would be missed).

Sequences were edited for the removal of oligo dT track and other contaminants. A batch file of ESTs having greater than 100 bp length of sequence reads were submitted to the NCBI dbEST division of GenBank.

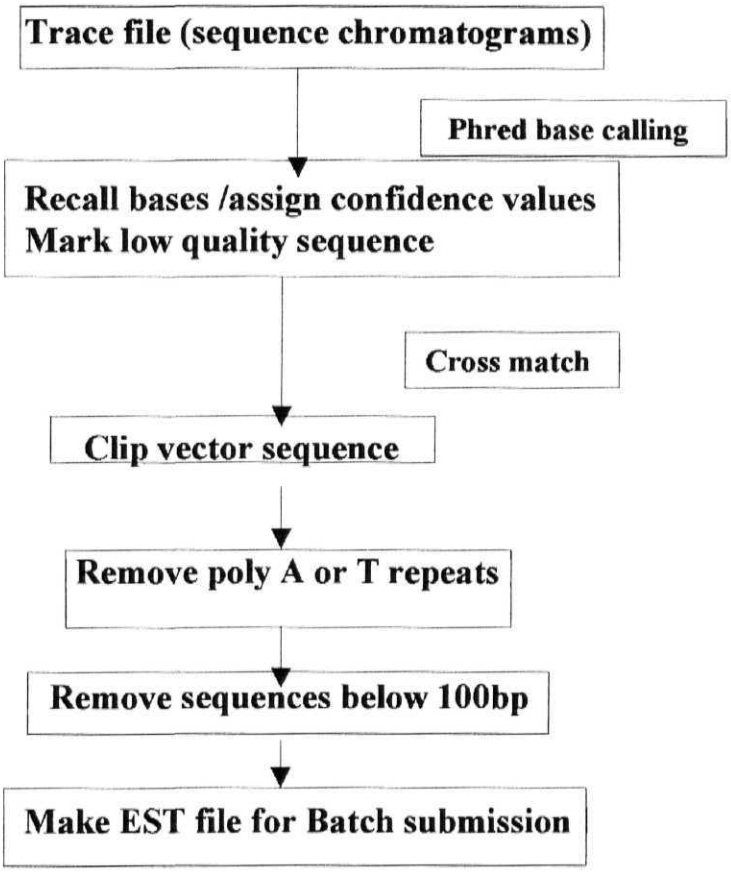


Fig. 3.2: Sequence processing and submission

3.2.2 Clustering

Comparing pairs of ESTs and looking for similarity is the basic element of clustering. This comparison is complex because the underlying sequencing technology is error prone – bases can be inserted, deleted, or misread. Clustering partitions the set into subsets, or clusters, based on similarity. Clustering process reads one sequence from the input file, compares the sequence against every existing cluster, and based on sequence similarity, either add it to an existing cluster or make it the first member of a new cluster. Each EST is a member of at most one cluster. Novelty is computed as the number of clusters identified divided by the number of sequences clustered. Ideally each cluster will uniquely represent a gene. Thus, the goal in constructing a UniGene set is to bring together all of the ESTs sequenced for a given gene into a single cluster. This information is useful for reducing redundant processing and for the annotation of EST sequences.

Phrap and CAP3 (Huang and Madan, 1999) assembly algorithms were used to assemble the individual ESTs into clusters of sequences deriving from the same transcript as tentative consensus sequences (TCs) and singletons representing unique transcripts. Sequence assembly is a process that involves comparison of sequences, finding overlapping fragment pairs, merging as many fragments as possible and creating a consensus sequence from the merged fragments. Accurate assembly algorithms are essential for reconstruction of the original DNA sequence grouping of ESTs in expression profiling. PHRAP identifies all potentially overlapping pairs of sequences. That have at least one "word" of length minmatch (typically 14 bases) in common, and the alignment between the sequences must have an alignment score of at least minscore (default: 30). PHRAP calculate the contig ("consensus") sequences, these quality scores of the contig sequence are conservative estimates of the error probabilities in the contig sequences. The higher sequence diversity in the sequenced 3-prime UTRs of the transcripts is best used to distinguish between gene family members during assembly.

3.3 Gene annotation

Homology search was done against nonredundant (nr) nucleotide and protein sequence databases using BLASTN 2.2.2 and BLASTX 2.2.2 versions of the BLAST programs (Altschul *et al.*, 1997) through BLAST 2.0 network client software with the DNAtools interface (http://www.crc.dk/dnatools). The BLASTN program was used to identify rice EST hits and rice BAC/PAC clones in the non-redundant (nr) nucleotide sequence database, High Throughput Genomic Sequences (HTGS) division of GenBank and the Chinese WGS (whole genome shotgun contigs) draft sequence of *indica* rice genome in the NCBI database were extensively used. The results of the BLAST analysis were manually checked for similarity in the aligned region (Fig. 3.3).

CLUSTALX version 1.8 (Thompson *et al.*, 1997) was used for pair wise and multiple sequence alignments for the analysis of gene families. Gene organization was analyzed using multiple gene prediction programs, Genscan, GeneMark.HMM, Glimmer R, and FGENESH, and aligning the known genes sequences to genomic sequences using SIM4 program (Florea *et al.*, 1998). Gene annotation techniques based on ESTs (Bailey et al., 1998) and gene-predicting algorithms complement each other in the sense that ESTs are often effective in identifying 3-prime ends of genes where the gene finders often fail, while gene finders relatively well determine the 5-prime ends which the oligo(dT)- primed cDNA clones often fail to reach. EST sequences represent spliced genes and are therefore valuable tools for determination of coding sequence in genomic DNA. Comparison between ESTs and genomic sequences immediately revealed the splice sites.

Stress responsive genes were analyzed for the known *cis*-acting elements in the promoter regions by searching in the PLACE database (Higo *et al.*, 1999).

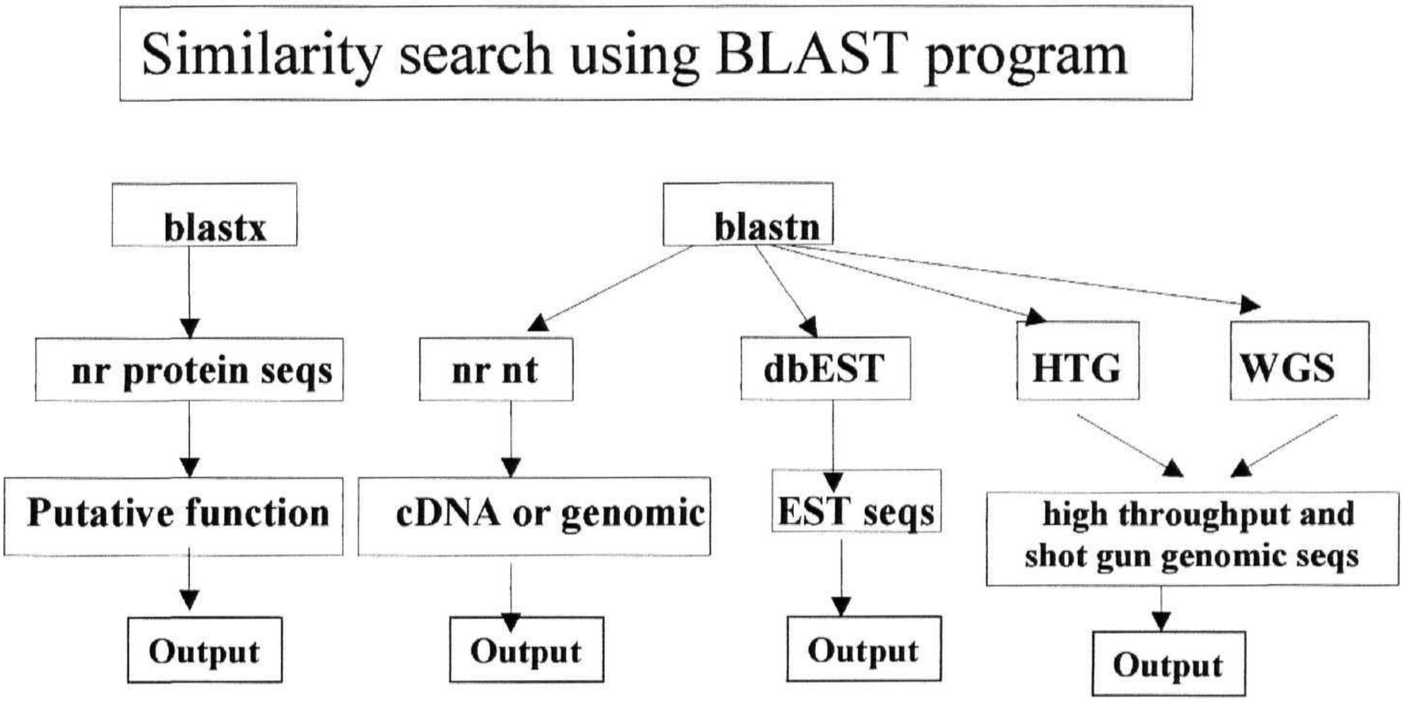


Fig. 3.3: Flow chart of BLAST analysis of high-quality sequences

3.4 Identification of putative stress responsive ESTs associated with abiotic stresses

The ESTs associated with stress response were identified from multiple sources based on the compiled list of stress regulated genes documented or presumed to be relevant to abiotic stress tolerance in more than one plant species (http://stress-genomics.org/stress.fls/expression/expression.html). Further, it is based on the microarray expression profiles of possible candidate gene sequences, 650 from Arabidopsis (Seki *et al.*, 2001, 2002; Kreps *et al.*, 2002), 150 from barley (Ozturk *et al.*, 2002) and 100 from rice (Matsumura *et al.*, 1999; Kawasaki *et al.*, 2001; Rabbani *et al.*, 2003). These were compared to the EST data set using TBLASTX with E-value $>1e^{-20}$. All the stress responsive gene sequences were retrieved from the above studies and a local database was constructed and utilized for BLAST analysis.

The identified putative candidate ESTs were mapped onto genetically anchored BAC/PAC clones to identify the possible candidate genes at the QTL associated with drought tolerance (Babu *et al.*, 2003; Price *et al.*, 2002; Price and Courtois, 2000; Zhang *et al.*, 2001).

3.5 Development of EST Database

The information pertaining to the analyzed ESTs, such as clone information, Accession numbers, and annotation were phrased into MS-Access database.

3.6 Molecular cloning and sequencing of DREB family genes in rice

Primers were designed using the rice genomic sequence information in the NCBI database to amplify DREB (Dehydration Responsive Element Binding) gene from rice using primer 3 program. The forward primer carrying Sal I restriction site (5' acgcgtcgacCC ATC ATC ACC GAG ATC GAC TCG AC – 3') and the reverse primer with Not I restriction site (5'- ataagaatgcggccg CTC ATT GTT CGC TCA CTG GGA G – 3') were used to amplify the target genomic DNA. PCR amplification was carried out in MJ Research thermal cycler with the PCR reaction setup of initial denaturation at 94^0C, 1 min; annealing 54^0C, 1 min; and 72^0C, 1 min for 30 cycles. The amplicons were run on 1% agarose gel, eluted, and dissolved in DNase free water. The amplified product of 1.2 kb was cloned in TA vector pTZ57R using InsT/Aclone PCR product cloning kit

(MBI Fermentas). The ligation reaction was carried out in a total reaction volume of 20μl congaing 100ng of TA vector.

3.7 Transformation of Ligation mixture:

15μl of ligation mixture was added to 50 μl of competent cells and given heat shock at 42^{0}C for 2 min. The cells were recovered by incubating at 37^{0}C for 60 min without antibiotic. 100 μl of transformed cells were plated on LB agar plates containing ampicillin (50μg/ml) for the selection of transformed cells (Sambrook *et al.*, 1989).

3.8 Miniprep of plasmid DNA:

The overnight grown bacterial culture (1.5 ml) was harvested by centrifugation and the bacterial pellet was suspended in 200 μl of TELT buffer (50 mM Tris pH 8.0, 2.5 M LiCl, 62.5 mM EDTA pH 8.0 and 0.4% Triton X--100 added later) and 20 μl of freshly prepared lysozyme (10 mg/ml) were added. After mixing, the suspension was incubated at 95^{0}C for 3 min. The bacterial debris was removed by centrifugation at 15,000 rpm for 20 min at 4^{0}C. The plasmid DNA was recovered by adding 100 μl of ice cold isopropanol and incubated on ice for 15 min. and centrifuged at 12,000 rpm for 20 min at 4^{0}C. The pellet was washed with 70% ethanol, dried and dissolved in appropriate volumes of TE buffer (10 mM Tris, pH 8.0 and 1mM EDTA, pH 8.0).

3.8.1 Sequencing

Sequencing reaction was carried out as prescribed in DYEnamic ET Protocol (Pharmacia.). DNA template, 400 ng was used in each DYEnamic ET sequencing reaction. The PCR product was purified and dissolved in 10 μl of loading solution (70% formamide) and used for injection.

The cloned product was confirmed by sequencing on Mega Bace 500 using M_{13} forward and reverse primers. The sequenced DREB homologue of 1128 nucleotide length was submitted to GenBank.

3.9 Plant material and stress treatment:

N22 seeds were surface sterilized with 5% sodium hypochlorite (NaOCl) for 5 minutes and rinsed with water for several times. Seeds were grown on rough filter papers in see-through germination boxes at room temperature in dark. Seedlings were grown in growth chambers with 16 hr light/ 8 hr dark cycle at 30 ± 2 ^{0}C. Stress treatment of the

10 day old seedlings were treated either with 20% Poly Ethylene Glycol (PEG), 100 μM ABA and 150mM NaCl. Control seedlings received only water. The seedlings were harvested at regular intervals and quick frozen in liquid nitrogen or immediately processed.

3.10 Northern Blotting:

3.10.1 Isolation of Total RNA:

Total RNA was isolated from shoot and leaf samples collected at different time intervals after stress treatment using guanidine isothiocyanate method (Chomezynski and Sacchi, 1987) with minor modifications. The harvested tissues were quick frozen in liquid nitrogen and stored at -70^{0}C. Five grams of freshly harvested or frozen tissues were ground in liquid nitrogen to a fine powder, transferred to tubes with extraction buffer (4M-guanidine isothiocyanate, 20mM EDTA, 20 mM MES, 50mM 2-mercaptoethanol). The extracts were incubated for 10 min at room temperature and centrifuged at 8000 rpm for 10 min at 4^{0}C. The supernatant was extracted twice with phenol-chloroform and once with chloroform. Finally, the aqueous upper phase was precipitated with 1/10 volume of 3 M sodium acetate and 2 volumes of ice-cold ethanol, kept at –80^{0}C for 2 hours and centrifuged at 8500 rpm for 30 min at 4^{0}C. The resultant pellet was suspended in the RNA suspension buffer (2M LiCl, 10 mM Sodium acetate, pH 5.2), incubated at 4^{0}C for 1 hour and centrifuged at 8500 rpm for 10 min at 4^{0}C. The resultant pellet was resuspended in RNase free water. The quality of RNA was checked on 1.2 % denatured agarose formaldehyde gels and quantity was estimated spectrophotometrically at 260 and 280 nm. The concentration of RNA was quantified as follows.

Conc. of RNA (μg/μl) = O.D at 260 x 40 x dilution factor. A solution of RNA absorbance at 260 nm that equals to 1.0 O.D contains approximately 40 μg/μl of RNA. A ratio of 2.0 for O.D 260/ O.D 280 is an indication of RNA purity.

3.10.2 Isolation of poly (A)$^+$ mRNA

Poly (A)$^+$ mRNA was purified from the total RNA using Oligotex (Qiagen) suspension. Nearly 250 μg of total RNA dissolved in 250 μl of Rnase-free water was mixed with 250 μl of buffer OBB (20mM Tris.Cl, pH 7.5, 1 M NaCl, 2m M EDTA and 0.2% SDS) and 15 μl Oligotex suspension. The mixture was incubated at 70^0C for 3 min and then at

room temperature for 10 min. The mixture was spun at 12,000 rpm for 2 min and the pellet was washed twice with 1ml of buffer OW2 (10mM Tris-Cl. pH 7.5, 150mM Nacl, 1mM EDTA) by mixing and centrifugation. After the final wash, the Poly $(A)^+$ mRNA bound to Oligotex particles were eluted by resuspending the pellet in 100 μl of hot (70^0C) buffer OEB(5mM Tris.Cl, pH 7.5) followed by centrifugation at 12000 rpm for 2 min and the step was repeated once for maximum recovery. The supernatants, containing eluted poly $(A)^+$ mRNA, were pooled, precipitated with ethanol, washed in 70% ethanol, dried and dissolved in an appropriate volume of RNase-free water.

3.10.3 Preparation of RNA sample for electrophoresis: The 3 μg of Poly $(A)^+$ mRNA (2 μl) was taken in 18 μl of denaturing buffer (containing 2.5 μl of 10 X MOPS, 4 μl of formamide and 12.5 μl of formaldehyde 37%) was added, kept at 65^0C for 10 min and snap cooled on ice and 3 μl of loading dye was added.

Loading dye: 50% glycerol, 1 mM EDTA, 0.25% bromo-phenol blue, 0.25% xylene cyanol.

3.11 Northern Analysis:

Equal concentrations of poly $(A)^+$ mRNA was separated through 1.2 % agarose formaldehyde gel (each 100 ml solution containing 1.2 g of agarose, 10 ml of 10X MOPS, 73 ml of water and 17 ml of 37% of formaldehyde), transferred to Hybond N+ membrane in the presence of 20X SSC. The filter was UV cross linked for 2 min and/or baked at 80^0C for 60 min. Prehybridization and hybridization were performed at 42^0C for 3 hr and overnight respectively in buffer containing 50% formamide, 5X Denhardt's solution, 5X SSPE, 1.5% SDS and 50 μg/ ml sheared denatured calf thymus DNA. Hybridization solution includes the radio labeled probe. The membranes were washed twice (10 min each) with 2X SSC, 0.5% SDS at RT followed by two washes of 0.5X SSC and 0.2% SDS at 62^0C. The filters were exposed to using an X-ray film with intensifying screens and kept at -70^0C.

Denhardt's solution 50X: 1% BSA, 1% ficoll, 1% PVP

10X MOPS solution: 200 mM MOPS, 100 mM sodium acetate, 10 mM EDTA (pH 8.0)

3.12 Random-primer labelling:

The cDNA fragments were used as template for preparation of radiolabelled probe by random primer labelling reaction. Nearly 100 ng of template DNA in 21 μl of water was denatured at 100^0C for 5 min and snap cooled on ice. To this 20 μl of 2.5 X random primer solution, 10 μM each of dATP, dGTP and dTTP, 50 μCi of [α-^{32}P]dCTP (3000 Ci/mM) and 4 units of Klenow were added and incubated at 37^0C for 60 min. Sephadex G-25 spin column was run to remove unincorporated [α-^{32}P]dCTP. The radioactively labelled cDNA was denatured and used for probing.

4. Results

4.1 Identification of Expressed Sequence Tag features, generation of high quality sequences and submission to GenBank dbEST division

High-quality EST sequences generated by Phred base calling with a cut off of >20 from Standard Chromatogram Format (SCF) files containing the trace data from MegaBACE 500 showed greater than 95% similarity to reported cDNAs and genomic sequences of rice in the database along 90% length of sequences in the aligned region. The integrity of the ESTs, visualized by quality values were used further in down stream applications such as assembly, annotation, and identification of sequence variations. The sequences were further characterized by the3' associated features and restriction site for effective screening of vector and other contaminants (Fig. 4.1). The EST features and the quality values of a representative chromatogram are shown in Fig. 4.2.

```
>NL38_D08 sequence exported from chromatogram file
                         RS      LT      Oligo-dT tract
CTCACTAAAGGGAATAAGCTTGCGGCCGCTCTGTTTTTTTTTTTTTTTTTTTTTTTTTTT
TTTTTTTTAAGAGAAATGGGATGAAAGCAGAGGTAGATTATATTAAGAAGCCTGGCACGC
ATGAGGAGATGGAGCAGGAGTACATGATGAGATCACCATGATGATAGATGATCCATCCAA
GCTAGCTAGGGATCCATCCAACATTGATCGACACACGCACACACTGACAACGACGACGAC
GGTTTTATTTAGGCGTCAGAGCAAAAGGCAAGGCCTGCCTACATATAAGCTGGGTACTAA
TGCAAGCAAATTAAGGCCTCGATCAAAGCTCTGATCGACAGTAGCAGCATCCATACGATA
AATCCTCATGCAGCCGGGCGACGACGACGATAGATTCAGTTGCAGGAGCAGCAGGAGCAG
CCGCTGCAGCTGGTGCCACACTTGCAGGTGTTGCAGCCGCAGCCGCCGTTCTCGCCGCTC
TCCACCGCCATCTCCATTCCTCCAGAGCTCGCCTTGTTGGATGGAGCAGCGAGGACGAAG
GTCTTGGTGGTGGCTGTGGCCTCCACATCAGGGAACATCTTGCATCCGCCGCAGCCGCCG
CCGCACTGGCAGCCGCTGCCGCAGCCGCAGTTGCCGCCGCAGCAAGACATCTTCTTCTTC
TTGGGAGGGCGAAGGACTTAACTAAGCA
```

Fig. 4.1: FASTA sequence showing features associated with cloning and 3'end sequencing. RS: Restriction site used in cloning, LT: Library tag

High-quality insert lengths of the 2069 non-redundant ESTs revealed the presence of 751 reads in the range of 400 to 600bp (Fig.4.4). The read lengths of these were shown in Fig. 4.5.

A total of 5,815 sequences were submitted to dbEST division after screening for vector contamination, RNA contamination, trimming low quality sequence regions (Fig. 4.3) and removing sequences less than 100bp length. (BI305180 to BI306756; BU672765 to BU673915; CB964418 to CB967504). A representative file of batch file submission details has shown below.

Clone Id: NL_15_80 (3')
Insert length: 581
DNA type: cDNA
Sequencing Primers: CGCCAGGGTTTTCCCAGTCACGAC
SEQUENCE

```
TCACTAAAGGGAATAGCTTCGGCCGCTGAAGATGTCGACGTTGCTCCGCTCTAAGTAGTC
GTGGCCTGGCCGCCATGAGGCCGCCCCACGCCTGCAGGTTCCTGATGAGGATGCCGTTGC
CGGCCGCCGTGTAGAGCTCCACGCCGACCGCCGGGTCGGTGCCGGCCTTCCAGATCGAGC
CCGTCTCCACGTAGAGCGTGTACACGCACTGGTTCGACCCCGTCGAGCCCGGGAGCAGCA
GCGCGCCCGTGTCCTCCTCCGACGTTGCTCCGGCGGCCTGCGCATTCAAGGTTATATATG
GTGAAGCTGATCACGAGATGGACTCATAAGCGAGACTCAGAAAATGTAGAACACGTACCA
AAGGCGCCATGGAAATGAGTGCCGTAAAGCAGAGGAGCTTATGCGCCATTGTGGAAGGCT
TGAAGACGCAGACTGTCTCGAGCTTTATCTGCACTTTCGGGATGTGATGTGTACCGACGA
TGATCTACTTACGCTGGCTAATGAGGCTTCTTATAGACGGAGACAAAACGGATAGGGGAA
CTGGAAATGCTACCGTATAAGATTACCTGATTAATGCACTG
```

Entry Created: Oct 7 2002

LIBRARY Lib Name: Drought stress (leaf) Cultivar: Nagina 22 (*indica* sub sp)

Organ: Leaf Tissue

Type: Entire leaf tissue **Develop. stage**: 35 day-old seedlings **Vector**: T7T3Pac

Description: ESTs from normalized leaf cDNA Library from drought stressed seedlings

SUBMITTER Name: Reddy AR., Lab: Department of Plant Sciences, School of Life Sciences, Institution: University of Hyderabad

Address: P.O. Central University, Hyderabad-500 046, A.P, India, Tel: 0091-40-3010265 Fax: 0091-40-3010145 E-mail: arjulsl@uohyd.ernet.in

CITATIONS Title: Novel EST enrichment with normalized cDNA libraries from drought stressed rice (*Oryza sativa* L.cv Nagina 22) (2002)

Authors: Reddy,A.R., Markandeya,G., Ramakrishna,W., Nagabhushana,I., Ravindra Babu,P., Madana Mohan Reddy,A., Chandra Sekhar, A., Bennetzen,J.L. Year: 2002.

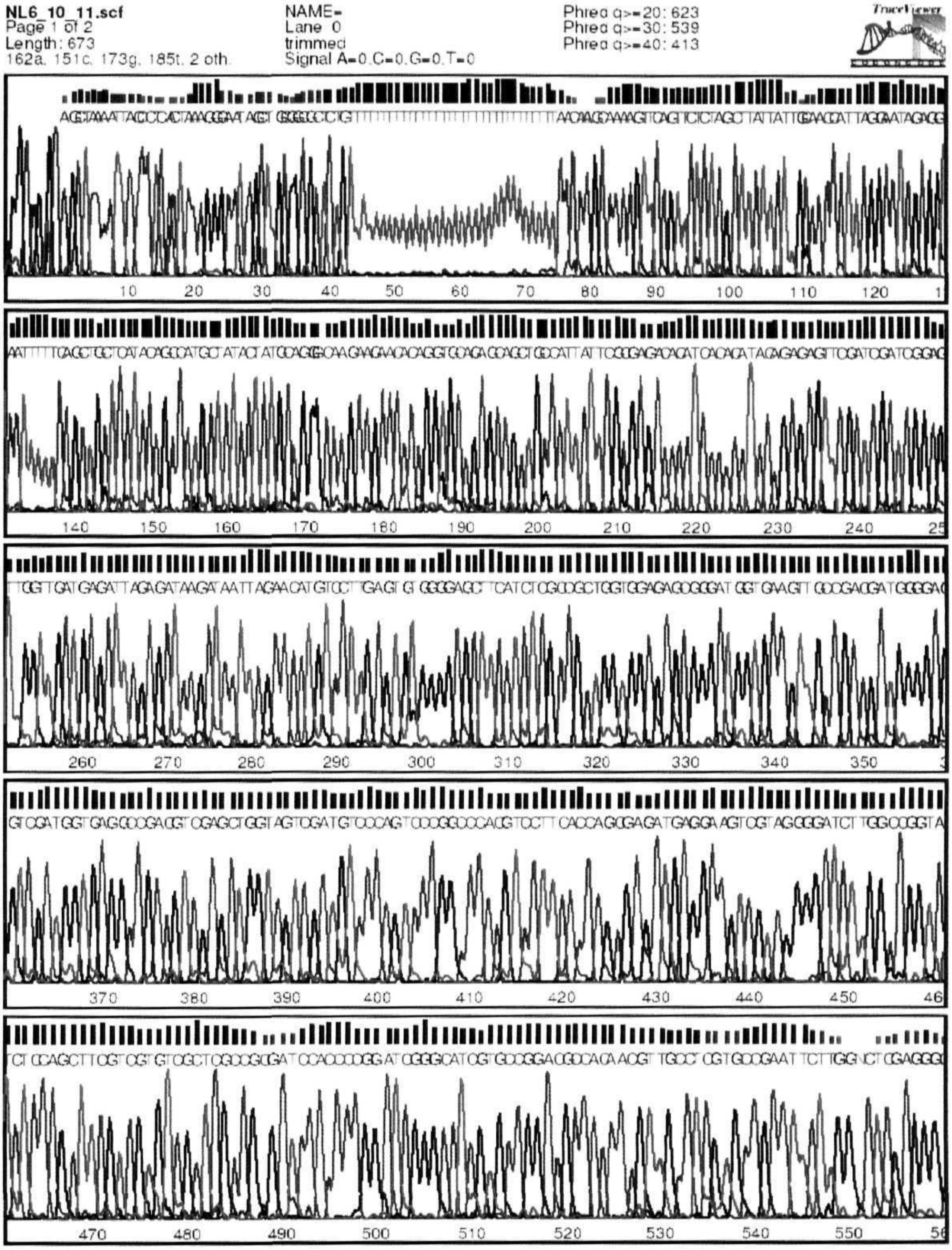


Fig. 4.2: PHRED base called sequence chromatogram with quality values Black bar above sequence represent high quality, red and blue bars show low accuracy of the called base.

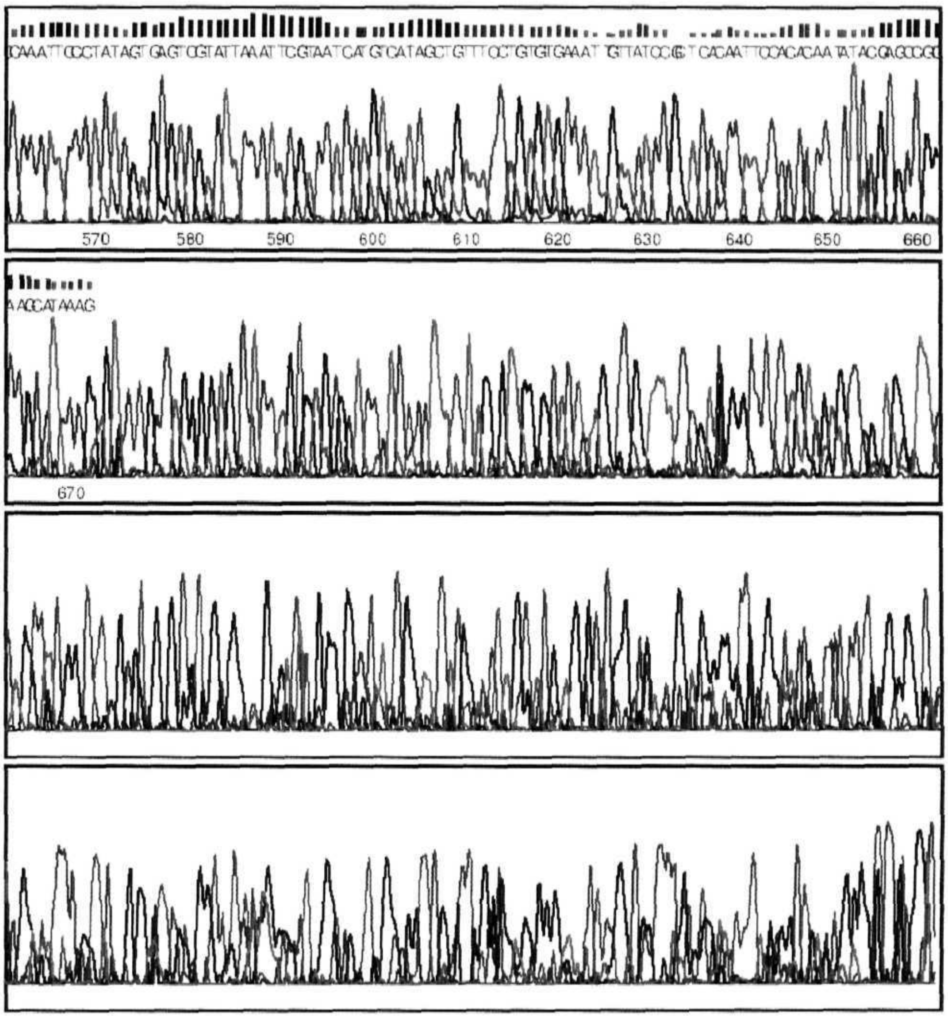


Fig. 4.3: Trimmed low Q20 region in last three panels

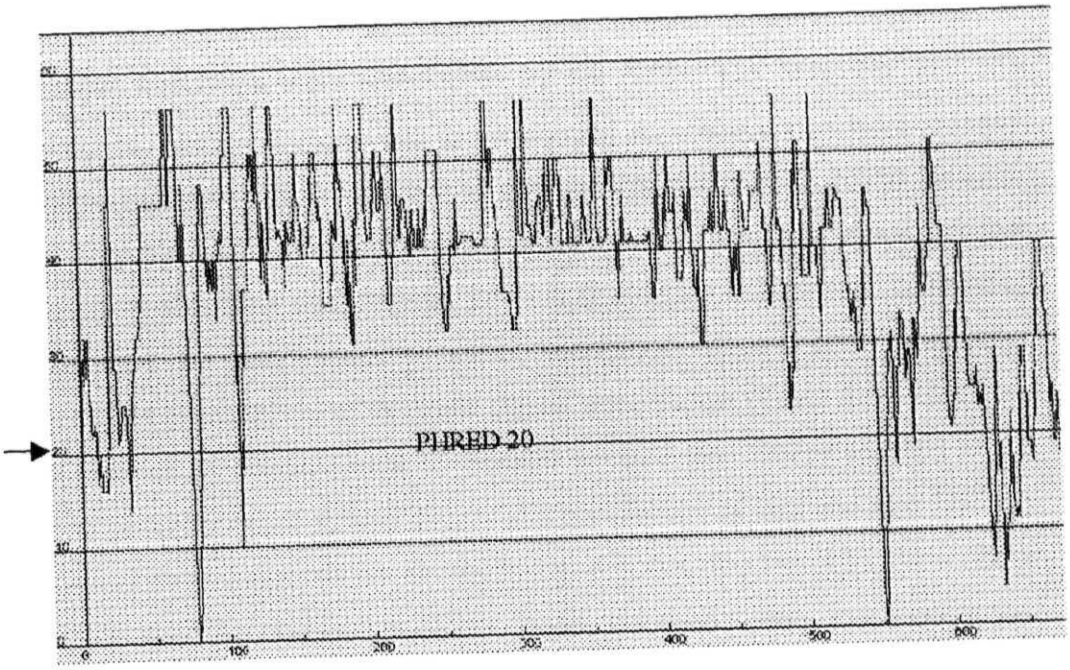


Fig. 4.4: Q20 of average read length

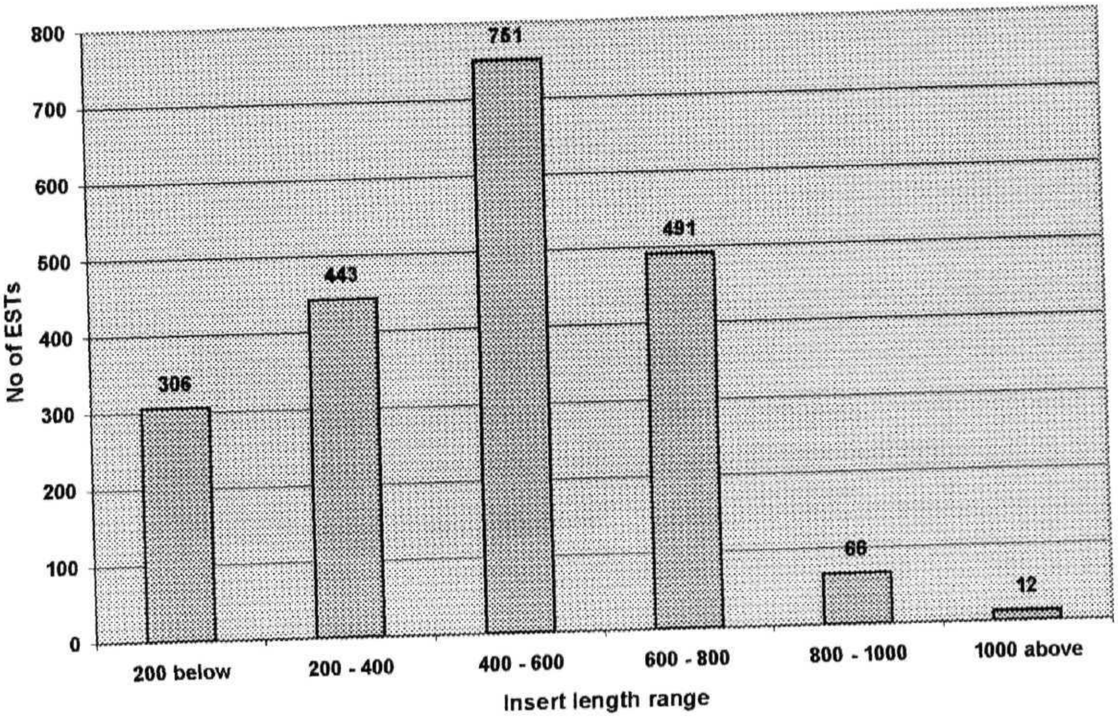


Fig. 4.5: High quality (PHRED 20) insert lengths of 2,069 non-redundant ESTs

4.2 CAP assembly and clustering analysis of ESTs

Assembly and clustering of analysis of ESTs revealed 2,069 (Unigene set of sequences) unique transcripts were represented by our ESTs. The assembly of 5,500 sequences produced 1241 singletons and the remaining 4,260sequences were grouped into 828 contigs (Fig. 4.6). Similar pattern of contigs was observed in NCBI Unigene clustering but singleton sequences were less represented. Identification of transcripts that are highly represented among the analysed ESTs may provide information concerning processes important for acclimation to stress conditions. Above 10X represented transcripts as redundant members of the gene cluster were shown (Table 4.1). This deep coverage and analysis of transcriptome of drought stressed leaf library resulted in the identification of potential stress related genes. The distribution of leaf and root ESTs in CAP assembly and cluster analysis are shown in Fig. 4.7, 4.8, and Fig. 4.9). As we have utilized the ESTs from the Normalized cDNA library the results only reveal genes that are potentially stress-related. The highly represented transcripts were further verified by annotation and comparing with previous studies associated with abiotic stress response in plants.

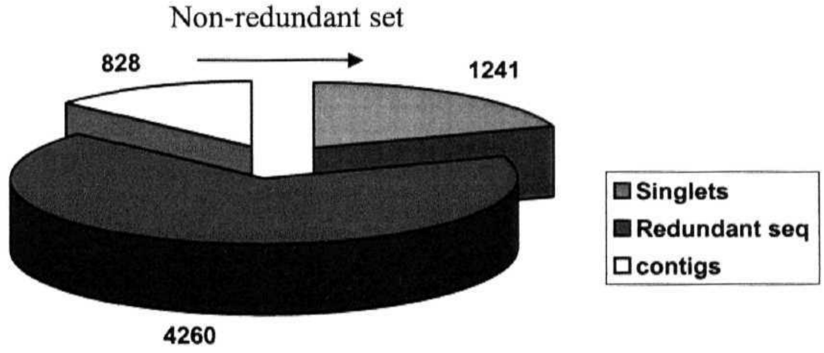


Fig. 4.6: Contig assembly of 5,500 high quality sequences

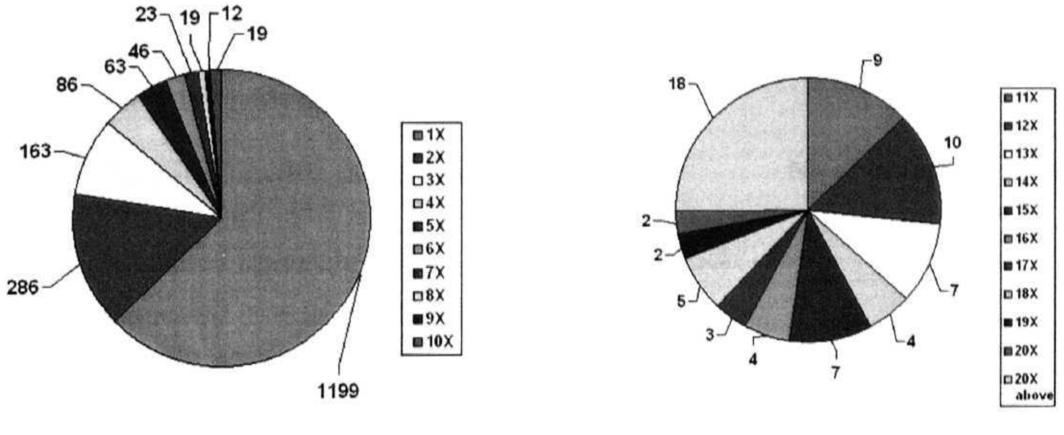


Fig. 4.7: Distribution of 5,300 leaf ESTs in contig assembly

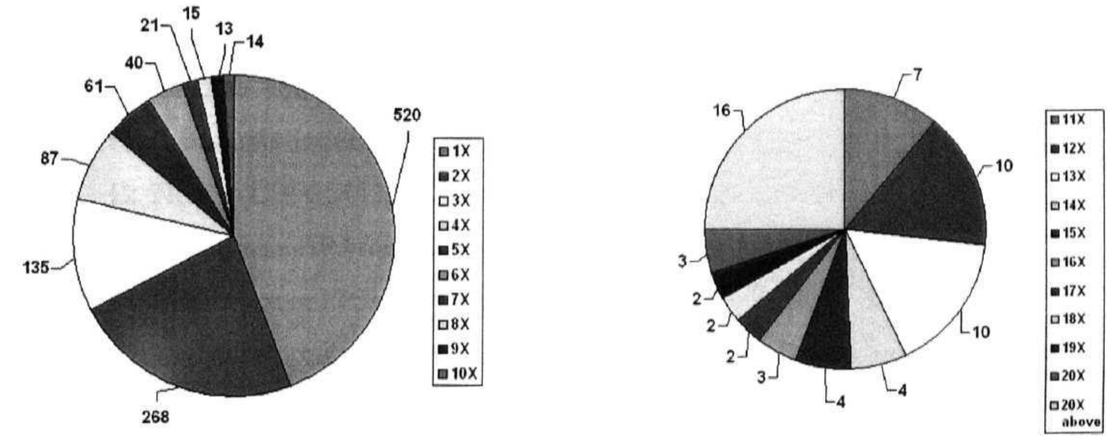
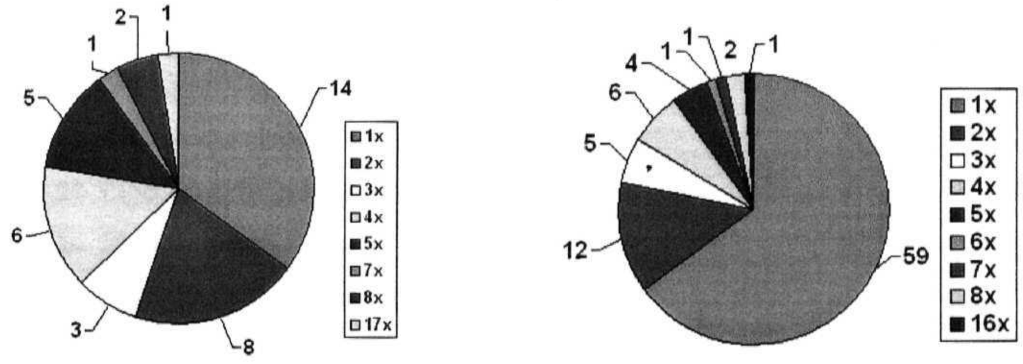


Fig. 4.8: Distribution of 4, 290 leaf ESTs in unigene clusters



Distribution of 128 root ESTs clusters

Contig Assembly (CAP) of 192 root ESTs in 91-unigene set

Fig. 4.9: Distribution of Root ESTs in contig assembly and unigene clusters

Table 4.1: Highly represented transcripts

S.NO	Putative function (Accession number)	No of times represented
1.	RicMT (AB002820)	101
2.	Ribulose bisphosphate carboxylase/oxygenase (L22155)	64
3.	Expressed protein (BI306646)	60
4.	Thioredoxin h (D26547)	55
5.	Putative kinetochore protein (AJ277096)	53
6.	Metallothionein-like protein (AF001396)	49
7.	Metallothionein-like protein (U57638)	47
8.	Rd22 (D10703)	47
9.	Triosephosphate isomerase (L04967)	41
10.	Glycine-rich protein (CAB61838)	31
11.	Jasmonate-induced protein (X98124)	29
12.	No hit (CB965482)	29
13.	Peroxidase BP 1(M73234)	27
14.	Photosystem I PSI-K subunit (L12707)	25
15.	Expressed protein (NM_127785)	23
16.	Lipid transfer protein LPT IV (AF017361)	23
17.	EF-hand Ca2+-binding protein CCD1 (AF181661)	23
18.	Adenylate cyclase (AP003583)	22
19.	Disulfide isomerase A6 precursor (BI305945)	20
20.	Quinone oxidoreductase -like protein (NM_121703)	20
21.	Ubiquinol--cytochrome c reductase (X79275)	19
22.	Ribosomal protein S15 (D10962)	19
23.	Lipid transfer protein precursor (U29176)	18
24.	Zinc-finger-like protein (X60700)	18
25.	Class III chitinase homologue (AB027426)	18
26.	60S ribosomal protein L144 (AF398144)	18
27.	Expressed protein (BI306274)	17
28.	60s ribosomal protein L36 (AL132960)	17
29.	No hit (CB965406)	17

30. Calmodulin (AF042840)	16
31. No hit (CB965220)	16
32. Mitochondrial malate dehydrogenase (X78800)	16
33. No hit (BU673628)	16
34. No hit (BI306445)	15
35. Translation initiation factor (AF094774)	15
36. Expressed protein (NM_117597)	15
37. Cytochrome P450 monooxygenase CYP92A1 (AY072297)	15
38. 60s ribosomal protein L 37(X79074)	15
39. IMP dehydrogenase/GMP reductase (ZP_00107581)	15
40. Chlorophyll a/b-binding protein (U74295)	15
41. No hit (CB965589)	14
42. Pridicted protein (BI306409)	14
43. Beta-oxyacyl-[acyl-carrier protein] reductase (AJ243091)	14
44. Mitochondrial ATP synthase 6 KD subunit (AB055076)	14
45. Dehydration-responsive protein RD22 precursor (Q08298)	13
46. UDP-glucuronic acid decarboxylase (AB079064)	13
47. Expressed protein(NM_129142)	13
48. Hypothetical protein (AJ271079)	13
49. Unknown(AY086234)	13
50. Histone H4(M12277)	13
51. 50S Ribosomal protein L18 (AC007932)	13
52. No hit (BI306584)	12
53. Putative protein (BI306641)	12
54. Ribosomal protein L35A (AF448416)	12
55. Photosystem II 10 kda polypeptide (U86018)	12
56. Putative protein (CB967429)	12
57. No hit (CB967065)	12
58. Alpha 1 subunit of 20S proteasome (AB026558)	12
59. Pridicted protein (CB964534)	12
60. Nonspecific lipid transfer protein (U88090)	12

61. Dof zinc finger protein (AB028132)	12
62. Fructose-1,6-bisphosphatase (AB007193)	11
63. No hit (BI306299)	11
64. Xyloglucan endotransglycosylase(X93175)	11
65. Metallothionein-like protein (U18404)	11
66. Cytosolic glyceraldehyde3phosphate dehydrogenase (AF251217)	11
67. Shoot GS1 for cytosolic glutamine synthetase (X14245)	11
68. No hit (BI306496)	11
69. Chitinase-B(AF402939)	10
70. Glutaredoxin (D86744)	10
71. Beta-D-glucan exohydrolase, isoenzyme exoii (U46003)	10
72. Heat shock protein 82 (Z15018)	10
73. Aldolase C-1 (D50307)	10
74. Catalase (D26484)	10
75. OsMYB1 (D88617)	10
76. Disease resistance response protein (NM_123616)	10
77. Pridicted protein (CB965224)	10
78. 40S ribosomal protein S30-like protein (AY128361)	10
79. Ribosomal protein S31 (D38011)	10
80. HSP90-like protein (AY077617)	10
81. Gigantea-like protein (AJ133787)	10
82. Pathogenesis-related protein (U20347)	10
83. GP28 gene (Z15085)	10
84. Small subunit of ribulose-1,5-bisphosphate carboxylase (D00644)	10
85. Cell division protein ftsh-like protein (NM_111112)	10
86. Vacuolar H+-pyrophosphatase (D45384)	10
87. No hit (CB967146)	10
Root tissue	
1. Cytochrome P450 (X81828)	17
2. Serine/threonine kinase (Y12465)	8
3. Cytochrome P450 (AB023038)	8

4.3 Annotation

The availability of vast gene sequence resources, information pertaining to genetic analysis of these resources, and the powerful Bioinformatics tools supported and shared in the public domain internationally has opened new avenues to understand genetic determinants of complex biological mechanisms.

Annotation of high-quality ESTs through homology search in the NCBI nr nucleotide and protein databases, using BLASTN and BLASTX programs resulted in the identification of putative genes to 78 % of the ESTs. The novel ESTs constitute 22 %, which has no significant homology in nr sequence database and dbEST division of GenBank (Fig. 4.10). Among the putative genes 48% have known functions, 30% are with unknown function (hypothetical and unknown proteins).

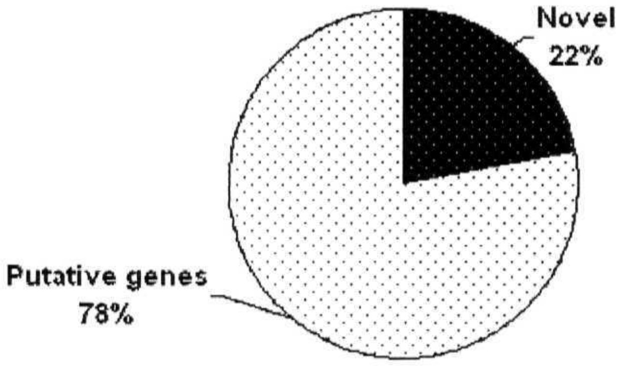


Fig. 4.10: Annotation of 2069 non-redundant ESTs based on similarity search in Nucleotide, Protein, and rice dbEST databases.

Homology search of non-redundant ESTs in rice dbEST divison showed 1157 hits and 912 ESTs do not have significant similarity to rice ESTs. These 912 ESTs constitute novel 3' sequence resource for accurate gene annotation (Fig.4.11).

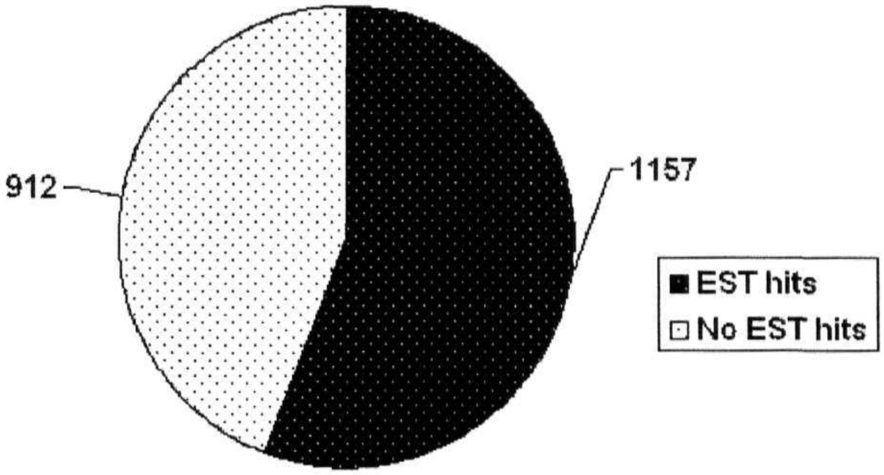
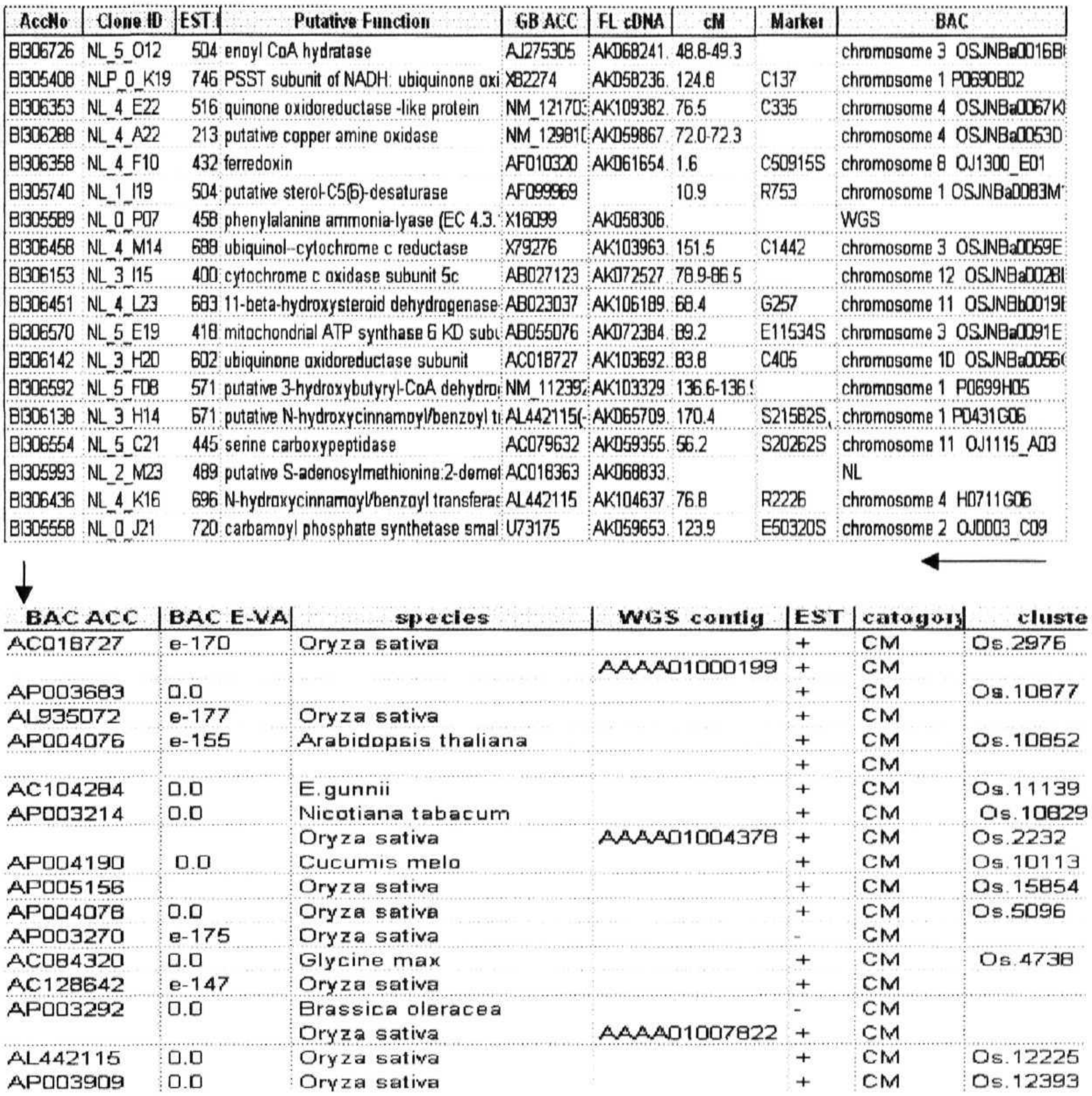


Fig. 4.11: No of ESTs showing hits in the rice db ESTdivision of GenBank

4.4 Development of extensive EST Database for identification and retrieval of information

An extensive EST database has been developed for data storage and analysis using Microsoft Access with several fields such as clone information, putative functions, expectation value, chromosomal location, and nearest marker information (Fig. 4.12).

Fig. 4.12 Database design



AccNo	Clone ID	EST	Putative Function	GB ACC	FL cDNA	cM	Marker	BAC
BI306726	NL_5_O12	504	enoyl CoA hydratase	AJ275305	AK068241.	48.8-49.3		chromosome 3 OSJNBa0016B
BI305408	NLP_0_K19	746	PSST subunit of NADH: ubiquinone oxi	X82274	AK058236.	124.8	C137	chromosome 1 P0690B02
BI306353	NL_4_E22	516	quinone oxidoreductase -like protein	NM_12170	AK109382.	76.5	C335	chromosome 4 OSJNBa0067K
BI306288	NL_4_A22	213	putative copper amine oxidase	NM_12981	AK059867.	72.0-72.3		chromosome 4 OSJNBa0053D
BI306358	NL_4_F10	432	ferredoxin	AF010320	AK061654.	1.6	C50915S	chromosome 8 OJ1300_E01
BI305740	NL_1_I19	504	putative sterol-C5(6)-desaturase	AF099969		10.9	R753	chromosome 1 OSJNBa0083M
BI305589	NL_0_P07	458	phenylalanine ammonia-lyase (EC 4.3.	X16099	AK058306.			WGS
BI306458	NL_4_M14	688	ubiquinol--cytochrome c reductase	X79276	AK103963.	151.5	C1442	chromosome 3 OSJNBa0059E
BI306153	NL_3_I15	400	cytochrome c oxidase subunit 5c	AB027123	AK072527.	78.9-86.5		chromosome 12 OSJNBa0028
BI306451	NL_4_L23	683	11-beta-hydroxysteroid dehydrogenase	AB023037	AK106189.	68.4	G257	chromosome 11 OSJNBb0019
BI306570	NL_5_E19	418	mitochondrial ATP synthase 6 KD sub	AB055076	AK072384.	89.2	E11534S	chromosome 3 OSJNBa0091E
BI306142	NL_3_H20	602	ubiquinone oxidoreductase subunit	AC018727	AK103692.	83.8	C405	chromosome 10 OSJNBa0056
BI306592	NL_5_F08	571	putative 3-hydroxybutyryl-CoA dehydro	NM_11239	AK103329.	136.6-136.		chromosome 1 P0699H05
BI306138	NL_3_H14	671	putative N-hydroxycinnamoyl/benzoyl t	AL442115(-	AK065709.	170.4	S21582S,	chromosome 1 P0431G06
BI306554	NL_5_C21	445	serine carboxypeptidase	AC079632	AK059355.	56.2	S20262S	chromosome 11 OJ1115_A03
BI305993	NL_2_M23	489	putative S-adenosylmethionine:2-demet	AC018363	AK068833.			NL
BI306436	NL_4_K16	696	N-hydroxycinnamoyl/benzoyl transfera	AL442115	AK104637.	76.8	R2226	chromosome 4 H0711G06
BI305558	NL_0_J21	720	carbamoyl phosphate synthetase smal	U73175	AK059653.	123.9	E50320S	chromosome 2 OJ0003_C09

BAC ACC	BAC E-VA	species	WGS contig	EST	catogory	cluste
AC018727	e-170	Oryza sativa		+	CM	Os.2976
			AAAA01000199	+	CM	
AP003683	0.0			+	CM	Os.10877
AL935072	e-177	Oryza sativa		+	CM	
AP004076	e-155	Arabidopsis thaliana		+	CM	Os.10852
				+	CM	
AC104284	0.0	E.gunnii		+	CM	Os.11139
AP003214	0.0	Nicotiana tabacum		+	CM	Os.10829
		Oryza sativa	AAAA01004378	+	CM	Os.2232
AP004190	0.0	Cucumis melo		+	CM	Os.10113
AP005156		Oryza sativa		+	CM	Os.15854
AP004078	0.0	Oryza sativa		+	CM	Os.5096
AP003270	e-175	Oryza sativa		-	CM	
AC084320	0.0	Glycine max		+	CM	Os.4738
AC128642	e-147	Oryza sativa		+	CM	
AP003292	0.0	Brassica oleracea		-	CM	
		Oryza sativa	AAAA01007822	+	CM	
AL442115	0.0	Oryza sativa		+	CM	Os.12225
AP003909	0.0	Oryza sativa		+	CM	Os.12393

4.5 Identification of the rice genomic sequences

To identify candidate genes associated with drought stress response the chromosomal locations were assigned to 1500 ESTs (Babu *et al.*, 2002). This was based on, above 95% identity along 90% of the EST length in the aligned region against rice genomic sequences using BLAST N program. The distribution of the identified chromosomal locations of the ESTs among the 12 rice chromosomes or *indica* whole genome shotgun sequences was shown in Fig. 4.13.

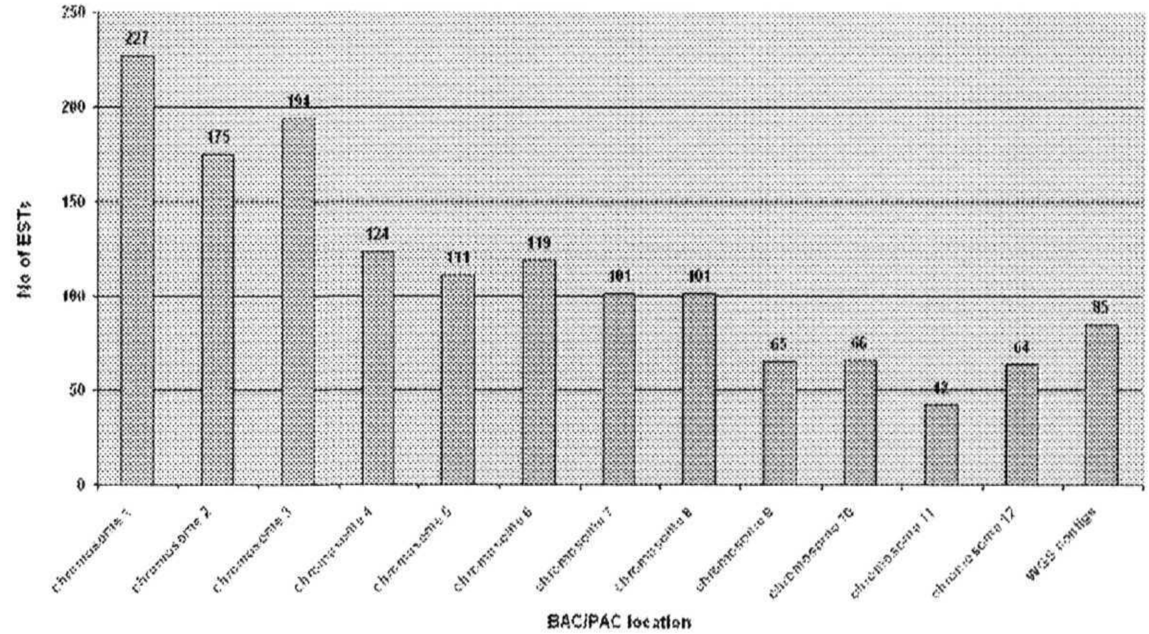


Fig. 4.13: Distrbution of genomic segments of 1475 non-redundant Expressed Sequence Tags

The locations of the ESTs on chromosomes were obtained from positive hit of the genetically mapped rice genomic BAC/PAC clones and the associated marker is considered as the nearest marker for association analysis. All the ESTs which were mapped to rice genomic sequences along with the annotation were shown in (Table 4.2 a and 4.2 b).

Table 4.2(a): BAC/PAC clone localization of non-redundant ESTs
BAC clone: BAC/PAC clone ID of rice genomic fragment
BAC acc: BAC/PAC clone accession in the GenBank
E-value: expectation value

EST Acc	Putative Function	Known Acc	FL cDNA Acc	Location cM	Marker	BAC clone	BAC Acc	E-VALUE
BI306311	Hypothetical protein	AC004146				P0520B06	AP003077	5e-51
BI306587	No hit					P0518C01	AP003277	1e-22
BI305950	50S Ribosomal protein L18	AC007932	AK099043			F11A17	AC007932	
BI306395	q group of receptor for activated C-kinase	D38231	AK062179			P0519D04	AP003455	5e-20
BU673726	No hit		AK070666			OSJNBb0049O23	AP003342	
BI305960	unknown protein	AP002866	AK106098			P0410E01	AP002866	8e-59
BI306425	beta-glucosidase homolog	AP003272	AK067934			P0432B10	AP003570	e-112
BI305588	No hit					P0699D11	AP002817	1e-50
BI305911	No hit					P0415C01	AP003243	2e-28
BI305393	ribosomal protein L29	AB042860	AK105054			P0475H04	AP002871	
BI305601	ribosomal protein S31	D38011	AK098982			P0031D02	AP003230	2e-28
BI305764	zinc finger protein	AF466199	AK068762			P0504D03	AP002970	
BI306561	No hit					P0489G09	AP002745	
BI306677	light-regulated protein	X68807	AK067670	0.0	R687	P0436E04	AP002818	4e-82
BI306702	1-deoxy-D-xylulose 5-phosphate reductoisomerase	AF367205	AK099702	0.3-4.2		P0482C06	AP002845	0.0
CB967361	aldolase		AK099387	5.6	S1543	P0494A10	AP002541	2e-30
CB965039	unknown protein	AAP46641	AK105948	5.6, 7.0	S1543,S10623	P0494A10	AP002541	0.0
BI306548	16.9 kDa heat shock protein	AP003250	AK065690	10.9	C90	P0443D08	AP003250	e-107
BI305921	predicted protein			10.9	R753	OSJNBa0083M16	AP003214	5e-66
BI305740	sterol-C5(6)-desaturase	AF099969		10.9	R753	OSJNBa0083M16	AP003214	0.0
BI305838	ribosomal protein L26	AF093540	AK105075	10.9	E30207S	P0480E02	AP002913	e-158
CB965987	No hit		AK111442	10.9	R753	OSJNBa0083M16	AP003214	2e-20
CB965081	No hit		AK100982	10.9	C90	P0443D08	AP003250	5e-56
BI306073	*wound induced protein homolog*	AB059238	AK059239	*10.9, 9.5*	*R753, S13048*	OSJNBa0083M16 *IAI2*	AP003214	*0.0*
CB964462	putative protein	NP_196193	AK065031	11.4	S5853	P0019D06	AP002483	0.0
CB967284	hypothetical protein			11.4, 11.7	S5853, S20154S	P0019D06	AP002483	0.0
CB966192	No hit		AK102823	11.7	S20154S	P0024G09	AP003311	1e-78
BU673389	glycosyl transferase, group 1 family protein		AK060202	11.7, 11.4	S20154S, S5853	P0024G09	AP003311	e-137
BU672800	Metallothionein-like protein type 2	D89931	AK069318	12.2	C146	P0434B04	AP002540	0.0
BI305763	triosephosphate isomerase	L04967	AK060920	12.2	C146, C30	P0434B04	AP002540	0.0
CB964871	No hit			13.1-16.1		B1189A09	AP003209	0.0
BI306712	protein synthesis inhibitor II	AP002912	AK103707	16.4, 16.1	E222S, R1613	P0701D05	AP003301	1e-62
BI305510	No hit			16.4-19.9		P0013F10	AP002523	2e-15
CB967290	expressed protein			16.4-19.9		OSJNBa0089K24	AP003215	e-142
CB967210	expressed protein	NP_201026	AK067273	16.4-19.9		OSJNBa0089K24	AP003215	0.0
BU673747	GTP-binding protein	AJ307662	AK069903	20.2	R480B	P0406H10	AP002524	e-172
BU673897	gigantea homologue	AJ133787	AK072166	22.6	S1778	P0666G04	AP003047	
BI305947	gigantea-like protein	AJ133787	AK072166	22.6	S1778	P0666G04	AP003047	0.0
BI305390	glutaminyl-tRNA synthetase	P52780	AK105222	24.0	R1944	P0510F03	AP002486	1e-54
CB967198	No hit			24.0-25.4		P0695A04	AP002816	0.0
BI305642	phytochrome-associated protein	AB059238	AK100314	25.4	C101	P0710E05	AP002743	0.0
BI306701	metallothionein-like protein	AF001396	AK058313	28.4	C12072	B1015E06	AP003197	0.0
BU673544	PKF1	X97547	AK069537	28.4	C12072	B1015E06	AP003197	0.0
CB965892	No hit		AK073568	28.4, 28.4-28.9	C12072	B1015E06	AP003197	8e-80
CB964493	putative protein	NP_568659	AK060099	29.7	C319	P0466B10	AP002093	0.0
BI306183	lipase	U38916	AK061118	30.8	S14085	P0515G01	AP001633	e-118
CB965162	lipase - like protein		AK104740	30.8, 30.5	S14085, E31375S	OSJNBa0038J17	AP003104	e-116
BI305241	lipase	AC007508	AK067352	32.4-36.9		P0417G05	AP002835	e-168
CB965323	hypothetical protein		AK062918	37.4	R457	P0431F01	AP001550	e-160
CB965913	pectinesterase family	NP_190491		38.8	R3470S	P0702F03	AP002481	3e-65
BU673240	pectinesterase	AL132956		38.8	R3470S	P0702F03	AP002481	e-175
CB967301	No hit		AK062426	38.8	C955	P0708G02	AP001539	5e -83
CB965373	expressed protein	NP_568985	AK061748	43.2	G317	P0034C11	AP002865	e-154
CB964453	Ramy1	AY072712	AK069098	43.2-45.4		P0711E10	AP000570	9e-58
BU673341	putative protein	AL049481	AK065023	49	C113	P0011D01	AP000969	e-126
BU673250	putative protein	NM_115533	AK073347	49, 49.3	C113, E476S	P0011D01	AP000969	e-168
BI306326	hypothetical protein	AP001081	AK058821	49.6	R2151S	P0693B08	AP001081	e-146
CB965454	expressed protein	NP_190772	AK100277	49.6	R2151S	P0693B08	AP001081	e-140
BU673142	ribosomal protein	AF034949	AK058917	50.8	C1338	P0038F12	AP000836	3e -39
BI306644	hypothetical protein			52.1	C1211	P0025D05	AP001072	0.0
CB965312	leucine zipper-containing protein		AK069333	52.1,52.4	C1211, S10086	P0025D05	AP001072	e-150
BI305581	beta-tubulin 1	U76744		52.7-53.9		P0581F09	AP003631	3e-59
BU673438	beta-tubulin 2	X78142	AK065323	52.7-53.9		P0581F09	AP003631	e-152
BI305836	CP12-like protein			60.6	S2139	P0035H10	AP002881	2e-56
CB965855	putative protein		AK103732	60.6	S2139, E4175S	B1146F03	AP003206	2e-93
CB965366	splicing factor SR1E		AK103534	63.9	S21093S	B1012D10	AP003535	0.0
CB965563	endo-1,3(4)-beta-glucanase		AK106887	63.9	L429	B1011A07	AP003722	0.0
BU673013	AP2 domain protein homolog	U77655		63.9	L429	B1011A07	AP003722	7e-91
CB965111	No hit			64.4	C51420S	P0028G04	AP003921	1e-18
BI305350	unknown protein		AK070437	64.4	C51420S	P0028G04	AP003921	
BI306497	ubiquitin	M60175	AK061988	66.5	C1905	P0537A05	AP002971	0.0
BI306002	hypothetical protein	AP003331	AK066632	67.6,66.5	E3109S, C1905	P0554D10	AP002869	0.0
BI306220	hypothetical protein		AK066107	69.3	S11337	B1088D01	AP003331	0.0
CB964846	Pridicted protein			70.1	E324S	OJ1111_G12	AP003337	0.0
BI305521	ribosomal protein S4	Y15009	AK102423	72.8	C369	P0514H03	AP003275	e-100
BI305381	ribosomal protein S4		AK061826	72.8	C369	P0514H03	AP003275	5e -47
BI305994	lipase	U38916		73.1	E3004S, R2630	P0043B10	AP003236	e-106
CB965243	No hit		AK104340	73.1	E3004S, R2630, S14028	B1329D01	AP006530	3e-65

BU673489	hypothetical protein		AK066633	73.4		OSJNBb0093M23	AP003854	0.0
CB966906	Glycosyl hydrolase family			73.4		OSJNBa0094H06	AP003217	e-172
BU673034	copper chaperone	NM_121751	AK060145	73.4		OSJNBa0094H06	AP003217	0.0
BI305761	sucrose-6F-phosphate phosphohydrolase	AY029159	AK071525	73.4		B1111C09	AP003204	e-155
BI306324	unknown protein	AP003213	AK072161	73.4	S3382S, R1547	P0520B06	AP003077	5e-83
BI305819	26S proteasome subunit	AB070262	AK069825	73.4	S3382S, R1547	P0520B06	AP003077	e-106
BI306054	33kDa oxygen evolving protein of photosystem II	X57408	AK104760	73.4	S10626	B1080D07	AP003203	0.0
BU673468	expressed protein	NM_111122	AK071246	73.4	S3382S,R1547	P0520B06	AP003077	0.0
BI305697	unknown protein	AP002868	AK063717	73.4-73.7		B1108H10	AP003562	2e-84
CB966276	unknown protein		AK072383	73.7	E60808S	B1003B09	AP004222	e-142
BI306591	HMG protein	Y08807	AK066658	78.8-83.7		OSJNBa0047D12	AP003516	
BU673703	dnaJ-like protein	NM_115594		86.0	E60128S	OJ1014_G12	AP003372	e-143
BI306401	hypothetical protein		AK102828	87.1	S13849	B1064G04	AP003924	3e-95
BI305196	unknown protein	AC004747	AK070797	87.1	S13849	B1064G04	AP003924	0.0
CB965333	hypothetical protein			87.1	S13849	B1064G04	AP003924	6e-66
BI306675	ras-related GTP binding protein	S66160		87.1, 86.0-87.1	S13849	B1064G04	AP003924	5e-82
BI305219	Insulinase			87.1, 86.0-87.1	S13849	B1064G04	AP003924	4e-46
BI306595	Col-0 casein kinase I-like protein			87.4-90.0		B1114B07	AP003334	0.0
BI306608	Unknown protein	AC015446		94.5-95.7		P0415C01	AP003243	e-162
BI305756	lysine decarboxylase-like protein	AB006700	AK071695	94.5-95.7		P0415C01	AP003243	0.0
BI306432	unknown protein			96.1	C50509	P0025A05	AP003504	e-128
BI305890	MYB family transcription factor		AK103432	98.0	C61385S, L588	P0681B11	AP003022	e-147
CB965488	nuclear RNA binding protein A	AF110228	AK065107	98.0, 98.5	C61385S, L588	B1085F01	AP003330	0.0
BU673642	hypothetical protein			98.5	S13559	OSJNBa0090K04	AP003216	3e-36
CB965232	hypothetical protein			98.5	S13559	OSJNBa0090K04	AP003216	1e-36
BI305552	small GTP-binding protein	AB015971	AK072731	100.4	R2635	P0410E03	AP002844	e-141
BI306695	subtilisin-chymotrypsin inhibitor 2	Y08625	AK062495	100.4-101.8		P0410E01	AP002866	0.0
BI305625	zinc finger protein	AC069474		100.4-101.8		P0686E09	AP002897	e-108
CB966166	Expressed protein	NP_567805	AK073157	100.4-101.8		P0686E09	AP002897	1e-76
BI306006	Group 4 late embryogenesis-abundant protein	M88321	AK061818	102.3	C1456	P0006C01	AP002744	1e-32
BU672887	Sgt1	AF192467	AK059268	103.1	C409	P0006C01	AP002744	e-174
BI306752	cytochrome P450		AK063764	103.7	R2374, C409	P0688A04	AP002839	0.0
BI305360	cytochrome P450	AB023038	AK105255	103.7, 103.1	R2374, C409	P0688A04	AP002839	0.0
CB966814	hypothetical protein		AK100804	103.7-106.2		B1158C05	AP003681	0.0
BI305815	peptide deformylase	AP003607	AK106980	106.2	R2816	P0004A09	AP003607	3e-73
BU673116	chloroplast carbonic anhydrase	U08404	AK060890	106.2	R2816	P0039G05	AP003855	0.0
BU673302	translation initiation factor 4A	X61206	AK073640	106.2	R2816	P0696E01	AP004367	0.0
CB964567	No hit		AK068661	107.6-110.2		P0038D11	AP003234	7e-45
CB965497	malate dehydrogenase	AF444195	AK071699	110.2	R886	OJ1159_D09	AP003792	1e-78
BU673186	hypothetical protein			110.2-112.7		OSJNBb0032K15	AP003710	2e-54
BI305861	VIP2 protein	AJ251051	AK071200	112.7	C10728S	P0684E06	AP003291	e-138
CB967256	unknown protein	AAO22748	AK105412	113.0, 114.1	R1514, S1457	P0671D01	AP003284	e-178
BI306255	nifU-like protein	NM_118347	AK098887	114.1	S1457	P0671D01	AP003284	e-130
BI305483	No hit		AK101664	114.1-116.5		P0014E08	AP004194	0.0
BI306648	peroxiredoxin	AF203879	AK058509	116.5	S2717	P0485G01	AP003264	0.0
BI306737	unknown protein	AP003223	AK101922	116.5	C50101S, S2717	P0007F06	AP003223	0.0
BI306666	unknown protein	AP001299	AK062194	116.5-118.9		B1144G04	AP003335	0.0
BI305604	No hit		AK069379	118.9	R1012	P0019E03	AP004363	1e-51
CB964630	expressed protein		AK069481	120.5	C922A	P0478H03	AP003452	e-113
BI305395	hypothetical protein			121.0	R178	P0519D04	AP003455	e-152
CB965408	phosphoethanolamine methyltransferase	AAL40895	AK069137	122.1	S11702	P0431H09	AP003248	8e-69
CB966840	No hit		AK102037	122.1	S11702	P0431H09	AP003248	1e-20
BI305825	unknown protein	AP003023	AK063120	122.6	C12174, E30470S	P0684B02	AP003023	e-168
CB964577	casein kinase I			122.6	C1162	P0510F09	AP003273	0.0
CB965809	P-Protein			123.2	F30, R290	B1142C05	AP003410	e-137
BI305978	putative protein	NM_119117		123.2	R290	P0456F08	AP002901	0.0
BI305582	lipase	AP002901	AK106129	123.2	R290	P0456F08	AP002901	0.0
BI305334	GmCK2p	U43839	AK070065	123.5	S1870	P0683B11	AP003288	e-133
CB966367	alliin lyase homolog		AK064938	123.5	S1870, S1676	P0683B11	AP003288	e-136
BI305422	chlorophyll a/b-binding protein	U74295	AK104176	124.8	C137	P0690B02	AP003292	0.0
CB967037	putative protein		AK061237	124.8	C137	P0690B02	AP003292	0.0
BI305408	PSST subunit of NADH: ubiquinone oxidoreductase	X82274	AK058236 ASP	124.8	C137	P0690B02	AP003292	0.0
CB965673	No hit			124.8	C137	P0690B02	AP003292	2e-20
CB965629	GDSL-motif lipase/hydrolase protein	NP_199379	AK104682	126.2	S1776	P0042A10	AP003343	e-169
BI306460	hypothetical protein		AK108262	126.5	C854	OSJNBb0021A09	AP003218	e-145
BU673292	dehydrin-like protein	AF314251	AK065226	126.5	C854	OSJNBb0021A09	AP003218	e-114
BI306285	unknown protein	AP002972		126.5	E21042S	P0638D12	AP002972	0.0
BI306081	hypothetical protein		AK062779	127.3	G393	P0439E07	AP003768	e-169
CB964721	Pridicted protein		AK062377	127.3	R1416	OSJNBa0014K08	AP003376	0.0
BI305760	unknown protein	AP003076	AK066261	129.0	S10712	P0481E12	AP003076	e-157
BU673293	cellulose synthase	AF323039	AK105078	129.3	R2417	P0046E05	AP003237	0.0
BI305473	ribosomal protein L18a	D21301		129.3	R2417, C1282	P0046E05	AP003237	0.0
BI305638	unknown protein	AB006698	AK103612	129.3	C1282	P0435B05	AP003249	e-161
BI306231	hypothetical protein			130.1	L694, C1282	P0435B05	AP003249	0.0
BI305628	epimerase/dehydratase - like protein	NM_122767		132.0	C250	P0460E08	AP003256	0.0
CB966760	glutathione S-transferase II	AF062403	AK099142	132.0	R37	P0403C05	AP003239	5e-73
BI306411	glutathione S-transferase II	AF062403	AK099142	132.0	R37	P0403C05	AP003239	0.0
CB967065	No hit		AK068819	132.8	C30040	OSJNBb0053G03	AP003377	0.0
BI306349	serine proteinase	AP003106	AK070376	133.9,134.2	R3347, R758	P0665A11	AP003106	0.0

CB967461	photosystem II protein psbW	S53025	AK060438	134.7	S10271	P0490D09	AP003265	0.0
BI306474	iron(III)-zinc(II) purple acid phosphatase precursor		AK060314	134.7-135.8		P0413G02	AP003344	0.0
CB966191	unknown protein		AK066370	134.7-135.8		P0413G02	AP003344	0.0
BI305434	acyl-CoA:1-acylglycerol-3-phosphate acyltransferase	AP002039	AK104484	136.1	E30358S	B1100D10	AP003368	0.0
CB967203	conserved hypothetical protein			136.1		B1070A12	AP003406	0.0
BU673575	ribulose 1,5-bisphosphate carboxylase/oxygenase large subunit	AB084766	AK105600	136.6	R2347A	P0557A01	AP003280	0.0
BI306592	3-hydroxybutyryl-CoA dehydrogenase	NM_112392	AK103329	136.6-136.9		P0699H05	AP003299	0.0
BI305530	putative PA domain			136.6-136.9		P0699H05	AP003299	6e-85
BI305405	putative protein	NM_115543		138.0	C2340	P0468B07	AP003260	0.0
BI305964	L24 ribosomal protein	X94296	AK101927	139.4	C1979	OSJNBa0085D07	AP004331	1e-78
BI306229	putative protein	NM_121551		139.4	C1979	OSJNBa0085D07	AP004331	0.0
CB964689	TOG protein			139.4, 139.9	C1979, S13781	B1148D12	AP003411	0.0
CB967415	hypothetical protein	AAO16698		139.9, 140.2	S13781, S14403	B1148D12	AP003411	e-151
BI305715	2,3-bisphosphoglycerate-independent phosphoglycerate mutase	M80912	AK060836	139.9,140.2	S13781,S14403	B1148D12	AP003411	0.0
BU673830	Ubiquitin-conjugating enzyme E2-17 kDa 11		AK058360	140.2-140.5		OJ1529_G03	AP003446	2e-99
BI305484	lipid transfer protein	AF109195	AK071598	140.5	G370, S721	P0485B12	AP003348	0.0
BI306731	No hit		AK073922	140.5-141.6		B1088C09	AP003734	4e-96
CB964444	leucine zipper-containing protein	NP_200651		140.5-141.6		B1088C09	AP003734	3e-93
CB965167	Pridicted protein		AK063898	142.4	R2125	P0506B12	AP003271	0.0
BU672999	expressed protein	NM_129142		142.4	R2125	P0460C04	AP004366	e-174
CB966272	ribosomal protein S25		AK069399	142.4	R2125, C1257	P0506B12	AP003271	e-180
BU672963	protein phosphatase 2C	AP003251	AK071996	143.7	C283	P0446B05	AP003251	0.0
CB966149	Pridicted protein		AK099428	143.7	E60580S	P0529E05	AP003279	e-121
CB965763	hypothetical protein	NP_172208	AK066597	143.7-144.5		B1096A10	AP003442	3e-50
CB967191	hypothetical protein	NP_565685	AK066597	143.7-144.5		B1096A10	AP003442	e-154
BI305545	unknown protein	AL161578	AK067838	145.0	S12805	P0683F02	AP003289	e-164
BI305459	ferredoxin	AB001386	AK069860	145.3	E50125S, C727A	P0489B03	AP003794	0.0
BI305524	fructose-1,6-bisphosphatase	AB007193	AK070516	146.4	S2596	P0505D12	AP003270	e-175
BI306282	oligopeptide transporter	AP003235	AK068840	146.4	R1764, C60583S	P0039A07	AP003235	3e-95
CB966832	fructose-1,6-bisphosphatase	AB007193	AK070516	146.4	S2596	P0505D12	AP003270	1e-75
BI305821	photosystem II subunit (22KDa) precursor	AP003235	AK058284	146.4	E10026S, R1764	P0039A07	AP003235	0.0
BI306654	No hit			146.4	R1764	P0039A07	AP003235	e-101
BU673539	Expressed protein	NM_117597	AK070251	147.5	E30867	P0471B04	AP003261	0.0
BI305955	RING-H2 finger protein RHA1a	AF078683	AK102849	147.5,149.1	C12740, C10419	B1065E10	AP003561	2e-53
BU673272	TGA-type basic leucine zipper protein	AF402608	AK109719	149.1	C10419	B1065E10	AP003561	e-110
BU673848	gibberellin-20 oxidase (Sd-1)	AF465256		149.1	C10419	B1065E10	AP003561	
CB966455	Calreticulin 3 precursor		AK099626	151.0-154.6		B1078G07	AP003407	e-166
CB967129	signal recognition particle 72KD protein		AK103621	151.0-154.6		B1099D03	AP003431	0.0
BI306051	5S ribosomal RNA	X64622	AK104693	154.6	L819	P0674H09	AP003349	2e-53
CB964565	hypothetical protein			155.2	S13312	P0674H09	AP003349	9e-82
BI305583	ATP-dependent RNA helicase-like protein	AC079022	AK064782	157.6	L1082,C346	P0470A12	AP003436	e-123
CB967111	glycine-rich protein	CAB61838		157.6	L1082, C346	P0470A12	AP003436	8e-66
BU673226	amino acid permease	AB022783		157.6	C346, G2200	P0497A05	AP003380	e-148
BU673074	hypothetical protein	AP003436		157.6,158.2	L1082, S10526	P0470A12	AC091088	0.0
BI306510	putative protein	AL021710	AK066408	159.0	C547	P0678F11	AP003437	e-162
BU673066	unknown	AY085105	AK067058	159.0	C547	P0678F11	AP003437	0.0
BU673491	hypothetical protein	NM_113624		159.0	C547	P0678F11	AP003437	0.0
CB966362	MADS-box protein FDRMADS5	AF141964	AK066160	160.4	R2833	P0592G05	AP004672	e-140
BI305562	No hit		AK068358	161.8	G54	OSJNBa0052O12	AP004330	e-134
BI306729	hypothetical protein	T48066	AK072525	163.5	C742	P0423A12	AP003246	6e-86
CB966827	dermal glycoprotein precursor	BAB89709		164.1	C1310S	P0504E02	AP003269	0.0
CB965999	No hit			164.1	C1310S	P0504E02	AP003269	0.0
CB967483	conserved hypothetical protein			164.1	C1310S	P0504E02	AP003269	e-122
BI306511	response regulator 5	AB042267	AK101721	169.5	S10581	P0431G06	AP003683	0.0
BI306138	N-hydroxycinnamoyl/benzoyl transferase	AL442115	AK065709	170.4	S21582S, S10581	P0431G06	AP003683	0.0
CB965782	LRP1		AK102444	170.4, 169.5	S21582S, S10581	B1139B11	AP004368	2e-30
CB964435	fruit-ripening protein	AAA99440	AK063053	170.4-176.3		P0401G10	AP003238	e-154
BI305854	gigantea-like protein	AJ133787	AK058375	170.4-176.3		P0401G10	AP003238	e-157
BU673756	signal recognition particle receptor alpha	AF360125	AK103090	170.4-176.3		OJ1294_F06	AP004326	4e-63
BU673258	bundle sheath cell specific protein 1		AK104613	176.3	R117	P0483G10	AP003263	e-160
BI306263	peroxidase BP 1	M73234	AK104420	176.3	R117	P0483G10	AP003263	e-167
BI306500	dTDP-glucose 4-6-dehydratase	NM_128345	AK070623	178.1	E60152S	OJ1656_A11	AP003448	0.0
CB964503	expressed protein	NP_179842	AK100555	181.8	R2727	P0020E09	AP003228	0.0
BU673120	RicMT	AB002820	AK062796	181.8	C112	P0459B04	AP003627	e-150
BU672810	RNA-binding glycine rich protein	D26182		181.8		P0459B04	AP003627	e-169
BI305247	chloroplast RNA helicase VDL isoform 1	AF261020		181.8	C112	P0518C01	AP003277	9e-98
BI306012	unknown protein		AK061660	181.8	C112	P0518C01	AP003277	0.0
						chromosome 2		
BI306700	fatty acid longase	AJ292770				OSJNBb0056C19	AP005915	7e -87
BU673044	kinetochore protein	AJ277096				OJ1435_F07	AP004187	3e -44
CB964743	vacuolar H+-pyrophosphatase		AK099807			OJ1643_A10	AP004192	0.0
CB966802	28S large subunit rRNA	M19228	AK105438			P0459B01	AP004778	1e-44
BU672776	ferredoxin	AB035645	AK065309	0.0		OJ1435_F07	AP004187	e-127
BI306374	No hit		AK071428	0.0		OJ1435_F07	AP004187	e-140
BI306464	possible apospory-associated protein	U13149	AK098858	0.0		OJ1399_D07	AP004138	5e-47
BI305211	hydroxypyruvate reductase	AC012563	AK067642	0.0		OJ1435_F07	AP004187	
CB966662	hypothetical protein		AK102217	0.0	C1357	OJ1359_D06	AP004851	0.0
BU673499	HUMAN Transcription factor IIIA		AK100473	4.7	C673	P0036E06	AP004867	0.0
CB965833	catalase	X61626	AK061923	4.7	C673	P0036E06	AP004867	e-120
CB964610	endosperm lumenal binding protein	AF006825	AK066006	4.7	C673	P0036E06	AP004867	e-145

BI305265	hypothetical protein	AC009978	AK107964	4.7-6.9		OJ1217_F02	AP004084	e-130
CB966349	glycine-rich, zinc-finger DNA-binding protein		AK070711	4.7-6.9		OJ1020_C02	AP004078	e-132
BU672833	Arginine/serine-rich splicing factor RSP41		AK062255	4.7-6.9		P0575F10	AP004885	e-153
BI305755	succinate dehydrogenase subunit 3	AF362741	AK103260	4.7-6.9		OJ1020_C02	AP004078	0.0
BI305468	cyclophilin 2	L29469	AK098919	4.7-6.9		OJ1020_C02	AP004078	0.0
BU673547	sin3 associated polypeptide p18	AY088934		4.7-6.9		OJ1020_C02	AP004078	0.0
BI306016	small nuclear ribonucleoprotein	AL161576	AK063621	6.9	E3626S	P0482F12	AP005311	4e-65
BI305498	expressed protein	NM_102910	AK106464	7.2, 7.9	E30145S, E60459	OJ1007_D04	AP004150	e-147
CB965835	S-ribonuclease binding protein SBP1	AF223395	AK106339	7.2, 7.9	E30145S, E60459	OJ1007_D04	AP004150	0.0
CB965498	No hit		AK106464	7.2, 7.9	E30145S, E60459	P0576F08	AP004886	8e-79
CB964505	ATP synthase		AK066913	7.9	E60459	OJ1007_D04	AP004150	0.0
BI306575	No hit		AK070887	8.9	S21847S	OJ1011_C09	AP004077	0.0
CB964810	hypothetical protein		AK108331	8.9	S21847S	P0030G11	AP004997	0.0
CB964840	glycosyl hydrolase family 17	NP_200617	AK101322	10.8	C11227S	OSJNBa0026E05	AP005647	e-166
CB965729	blast-resistant leaf mRNA			12.8	E290S	P0479D12	AP005191	3e -80
BI306623	putative protein	NM_114287		13.4	E51222S	P0409F01	AP004748	0.0
CB965898	No hit			15.0	R810	OJ9003_G05	AP004126	e-124
BU672770	phosphoglycerate kinase		AK100371	17.6	R2403S	P0669G09	AP005071	3e -59
BI305733	hypothetical protein	NM_129502	AK102098	17.6	R2403S	P0669G09	AP005071	e-167
BU672850	succinic semialdehyde dehydrogenase	AF117335	AK060831	18.4	C1796	OSJNBa0073A21	AP005772	0.0
CB964999	PROTEIN TRANSPORT PROTEIN SEC61 GAMMA SUBUNIT		AK063815	20.9	S1511	P0544B02	AP004840	e-115
BI305841	No hit		AK073452	22.0	G1327	OJ1297_C09	AP004087	e-124
BI306000	hypothetical protein		AK101237	25.5	S21672S	P0437H03	AP000367	0.0
CB966411	KH domain/zinc finger protein	NP_566412		25.5	S21672S	P0437H03	AP000367	0.0
BI305818	fumarylacetoacetate hydrolase-like protein	NM_101077	AK099565	26.9	E30284S, R2344	OJ1524_D08	AP004191	e-151
BI305325	lipoxygenase	L23968	AK061610	26.9	E30284S	OJ1225_F07	AP004184	0.0
CB964922	putative protein	NP_567894	AK106819	30.7	C149	OSJNBb0056C19	AP005915	e-121
BU673399	ribonucleoprotein		AK100904	32.8, 33.1	C1654, C10425S	P0758B01	AP005826	0.0
CB966128	Hypothetical protein	AC034257_6		33.6	C630	P0495C02	AP005003	0.0
BU672768	major intrinsic protein	D17443	AK068806	36.3	G227	OJ1705_E12	AP004070	0.0
BI305504	hypothetical protein			36.3	G227	OJ1705_E12	AP004070	e-108
CB966635	Unknown protein	AF370513	AK103384	36.3	G227	OJ1705_E12	AP004070	e-104
BU673892	ribosomal protein L11			36.8	C2168	OJ1077_A12	AP003991	1e -34
BU673550	shaggy-like kinase etha	Y13437	AK073725	36.8	C2168	OJ1077_A12	AP003991	e-154
BI306222	Bromodomain protein			37.4-39.6		OJ1113_G05	AP004018	e-123
CB965576	alpha-hemolysin -related	NP_187542	AK102256	42.1,42.4	R712, C980	B1178F07	AP006160	2e-72
BI306491	integral membrane protein	NM_105398	AK073967	47.8, 48.1	S2525, R1589	P0413A11	AP004771	0.0
BU673185	40S RIBOSOMAL PROTEIN S3A	D26060	AK099131	50.3	R3191, C626	OJ1756_H07	AP004168	e-174
BI306720	arabinogalactan-like protein	AP004194	AK104533	52.2-53.5		OSJNBb0026D20	AP005476	0.0
BU672791	expressed protein	NM_128956	AK102909	53.5-54.6		P0705A04	AP004891	0.0
BU672821	No hit			54.6		OJ1113_F08	AP004238	e-136
CB965606	hypothetical protein		AK070014	54.6		OJ1521_G01	AP004854	0.0
CB964432	CLE13	NP_683493	AK061156	54.6	C2184B	P0688H12	AP004890	0.0
BI306478	phosphoshikimae1carboxyvinyltransferase	AB052962	AK105036	54.6		OSJNBa0086N11	AP005651	3e-64
BU673546	putative protein	NM_123133	AK104165	54.6	E61832S	OSJNBb0042G06	AP005778	e-104
CB964821	unknown protein		AK068494	54.6	C2184B	P0688H12	AP004890	0.0
BI306403	sterol delta7 reductase		AK100199	54.9	C12318C	OSJNBa0035A24	AP005514	1e -93
CB966371	No hit			57.9	S20944S	OSJNBb0021C10	AP005630	0.0
BU673050	putative protein	NM_123764	AK059542	57.9	S20944S	OJ1288_D09	AP004086	e-156
CB965332	expressed protein		AK102956	58.4	S10847A	OJ1003_F04	AP005285	e-145
BU673525	hypothetical protein		AK073150	58.4	S10847A	OJ1003_F04	AP005285	0.0
CB965750	U6 snRNA-associated Sm-like protein		AK058487	59.5	E30164S	OSJNBa0052M16	AP005841	e-147
BI305632	putative protein	AL133315		59.5	E30164S	OSJNBa0052M16	AP005841	0.0
CB964976	Mei2-like protein		AK062177	62.2-64.7		OJ1001_A04	AP004226	0.0
CB964423	No hit		AK100313	64.7	a0019H14	P0461D06	AP005845	4e-47
BI305489	unknown protein	AB010069		64.7-71.3		OSJNBb0003H22	AP005842	0.0
BI305213	heat stress transcription factor Spl7	AB050095	AK101934	71.3	S21157	P0475F05	AP004879	0.0
BI306586	ribosomal protein S14		AK101459	71.3-77.8		OJ1112_G07	AP004156	4e-89
BU673267	SR3 sucrose-regulated mRNA	U16257		71.3-77.8		P0476C12	AP004789	6e-58
BI305463	hypothetical protein	NM_103154	AK106428	81.4	G1456	OJ1004_H01	AP004038	0.0
CB966732	No hit		AK067395	81.4	G1456	OJ1004_H01	AP004038	3e-60
CB966060	hypothetical protein			81.4	G1456	OJ1004_H01	AP004038	7e-78
CB965340	thylakoid-bound ascorbate peroxidase	BAC79363		81.7	C499	P0470G10	AP004876	2e-61
BI305767	expressed protein	AC005489	AK065290	81.7-83.6		OSJNBb0038F20	AP005808	e-162
CB967000	No hit		AK069453	81.7-83.6		OSJNBb0038F20	AP005808	0.0
BU673520	gamma hydroxybutyrate dehydrogenase	AY044183	AK104168	83.6	E3634S	OJ1712_E04	AP004144	e-162
CB966856	hypothetical protein			86.8-88.2		P0703B01	AP005116	0.0
BI306265	No hit			88.2-89.9		OSJNBa0016D04	AP005800	e-120
BI305561	small GTP-binding protein	L35845	AK061099	89.9	E60561S	OSJNBb0056I22	AP005644	0.0
BI305754	hypothetical protein		AK100801	89.9-91.5		OSJNBb0046O12	AP005643	e-163
CB964679	hypothetical protein		AK100542	89.9-91.5		OSJNBb0046O12	AP005643	0.0
BI306611	serine/threonine protein kinase		AK100504	91.8	S13181	OSJNBa0016G10	AP005874	0.0
BI306646	expressed protein			91.8-93.2		OJ1124_H01	AP004003	e-158
CB965127	N-acetylglucosaminyltransferase	NP_172759	AK060490	91.8-93.2		OJ1126_D09	AP004023	0.0
CB966524	kinase-like protein	AB060276	AK106801	93.2	C1061	OJ1791_B03	AP005385	2e-43
BI305352	S-adenosylmethionine decarboxylase 2	AJ251899	AK059503	95.5	S1785	OJ1004_A05	AP005286	0.0
BI305618	GrpE protein	AJ010819		95.5	S1785	OJ1004_A05	AP005286	0.0
BI306003	expressed protein		AK059024	97.4, 98.2	C41, S10844	OJ1212_C01	AP004083	3e-74

CB966524	kinase-like protein	AB060276	AK106801	93.2	C1061	OJ1791_B03	AP005385	2e-43
BI305352	S-adenosylmethionine decarboxylase 2	AJ251899	AK059503	95.5	S1785	OJ1004_A05	AP005286	0.0
BI305618	GrpE protein	AJ010819		95.5	S1785	OJ1004_A05	AP005286	0.0
BI306003	expressed protein		AK059024	97.4, 98.2	C41, S10844	OJ1212_C01	AP004083	3e-74
BU673856	No hit			97.4, 98.2	C41, S10844	OJ1212_C01	AP004083	0.0
BI305734	phosphoenolpyruvate carboxylase kinase	AF399915		101.2	S2287	B1215B07	AP006523	0.0
CB965589	No hit			101.2	S2287	P0656G12	AP005070	4e -59
BU673745	adenosine kinase	AB050624	AK101791	101.2	S2287	P0656G12	AP005070	e-124
BI305709	GP28 gene	Z15085	AK067801	101.2	S2287	P0656G12	AP005070	0.0
CB967168	GP28	Z15085	AK067801	101.2	S2287	P0656G12	AP005070	0.0
BI306750	aquaporin		AK061782	101.5	C920	P0661A05	AP004889	e-114
BU673411	defensin	AB089942		101.5	C920	P0661A05	AP004889	0.0
BI306754	ATP-dependent Clp protease proteolytic subunit		AK065162	102.9	R2792	OJ1643_A10	AP004192	0.0
BI306533	alpha 2 subunit of 20S proteasome	AB026559	AK072855	103.4	C424	OJ1626_B09	AP004069	e-112
CB964552	hypothetical protein			103.9-105.8		OJ1282_H11	AP005291	0.0
CB966700	auxin-regulated protein			103.9-105.8		OJ1282_H11	AP005291	1e-66
CB967219	cytochrome b5	NP_200168	AK066861	105.8	R857S	P0048B08	AP004868	e-112
BU673099	putative protein		AK068164	107.4	R480A	P0491E01	AP002485	8e-51
CB964906	unknown protein		AK073123	109.3	C1769	chromosome 2	AJ307662	e-175
CB967056	expressed protein	NP_567716	AK108493	109.3	C1769, R63	chromosome 2	AJ307662	e-106
CB967180	Pridicted protein		AK070096	109.3	C1769	P0519E06	AP005006	0.0
BU672860	unknown protein	AY096651	AK103515	110.6, 110.9	S908, C747	P0461B08	AP005108	6e -79
CB964531	hypothetical protein			110.6, 110.9	S908, C747	P0461B08	AP005108	0.0
CB966475	expressed protein		AK101444	110.6, 110.9	S908, C747	P0461B08	AP005108	e-144
BI305883	No hit			111.2	L737	OJ1486_E07	AP004139	e-144
CB964540	Argonaute protein	NP_849784		112.6	C1236	OJ1493_H11	AP004188	0.0
BI305889	hypothetical protein		AK100584	114.0-118.1		P0663F07	AP005823	8e -67
BU673895	No hit			114.0-118.1		P0663F07	AP005823	1e -31
BI305912	conserved hypothetical protein		AK070774	118.1	R2216	OJ1148_D05	AP004118	0.0
BI306714	phosphoribulokinase	X51608	AK099461	118.1-122.8		P0459B01	AP004778	0.0
BU673608	hypothetical protein		AK070100	118.1-122.8		P0459B01	AP004778	0.0
BI305578	gamma-tocopherol methyltransferase	AF213481	AK071763	118.1-122.8		OJ1111_E07	AP003994	e-100
BI305486	25S ribosomal RNA gene	M11585		118.1-122.8		P0459B01	AP004778	0.0
CB966623	xyloglucan endo-transglycosylase	S71223	AK101915	118.1-122.8		B1053A04	AP005859	e-127
CB966485	xyloglucan endo-transglycosylase		AK111242	118.1-122.8		B1053A04	AP005859	0.0
BI305743	18S small subunit ribosomal RNA	AF069218	AK059783	118.1-122.8		P0459B01	AP004778	0.0
CB965981	hypothetical protein		AK070100	118.1-122.8		P0459B01	AP004778	6e-65
BU672778	retinitis pigmentosa GTPase regulator-like protein			118.1-122.8		P0459B01	AP004778	0.0
CB966146	cytochrome P450 family	NP_567581	AK067007	122.8	R418	P0724B10	AP005825	e-132
BI305558	carbamoyl phosphate synthetase small subunit	U73175	AK059653	123.9	E50320S	OJ0003_C09	AP004076	e-155
CB964508	unknown protein		AK065045	125.6	C520	OJ1038_A06	AP003983	e-179
CB964838	unknown protein		AK067559	125.6	C520	OJ1311_H06	AP004161	0.0
CB965104	Omega-6 fatty acid desaturase, endoplasmic reticulum		AK061931	125.9, 126.4	R2643, C92	OJ2056_H01	AP004098	0.0
BI305829	pollen-specific protein BAN102-like protein		AK109216	126.4	C92	OJ2056_H01	AP004098	e-140
BU673273	Dof zinc finger protein	AB028132	AK061000	128.3	C601	B1121A12	AP005284	0.0
BI305346	unknown protein	AB024034	AK061221	128.3	S11127	P0685G12	AP005113	0.0
BI305999	expressed protein	NM_116345	AK058271	129.4	C10187S	P0617A09	AP004888	5e-92
BU673837	hypothetical protein			130.2	C11895S	P0643A10	AP005319	2e-85
BI305577	hypothetical protein		AK104571	130.2	C11895S	OJ1212_E12	AP004050	0.0
BI306042	shoot GS1 for cytosolic glutamine synthetase	X14245	AK104987	131.0		P0487D09	AP004880	e-143
BI305837	small GTP-binding protein OsRac3	AB029510	AK059970	131.0-134.5		P0585G03	AP004800	0.0
CB966304	Pyruvate dehydrogenase E1 component alpha subunit, mitochondrial precursor	BAC57468	AK098950	131.0-134.5		P0684F11	AP005112	e-136
CB964461	NOD26-like membrane integral protein ZmNIP2-1		AK069842	134.5	R685	OJ1734_E02	AP005297	1e-45
BI306402	annexin p35	X98245	AK101787	135.5-138.0	C679	OJ1288_G09	AP004119	0.0
CB965001	glyceraldehyde-3-phosphate dehydrogenase		AK105877	135.5-138.0		OJ1288_G09	AP004119	e-122
CB967127	expressed protein		AK065736	135.5-138.0		OJ1288_G09	AP004119	0.0
BI305249	C3HC4-type RING zinc finger protein	AC004238	AK067456	138.0	C679, S13245	OJ1175_B01	AP004159	
BI305302	3'-5' exoribonuclease	NM_106417	AK072432	138.0	C679	P0627E03	AP005012	0.0
CB965220	immunophilin		AK063850	138.0	S13245	OJ1175_B01	AP004159	e-109
BI305726	low molecular weight heat shock protein precursor (hsp22)	AF035460	AK074003	138.0	C679	P0471A11	AP004814	e-163
CB965213	hypothetical protein		AK064111	138-140.9		OJ1611_C08	AP004068	0.0
CB966637	No hit			138-140.9		OJ1004_A11	AP005287	e-109
BI306075	hypothetical protein		AK103084	138-140.9		P0486G03	AP005002	e-122
BI306260	mitochondrial phosphate transporter	AB016065	AK069611	138-140.9		OJ1004_A11	AP005287	0.0
BU673654	S-adenosylmethionine:2-demethylmenaquinone methyltransferase		AK059956	138-140.9		P0486G03	AP005002	e-134
CB966771	protein kinase		AK066978	140.9	S5450S	OJ1534_E09	AP004140	0.0
BI306641	putative protein		AK101869	142.5	S5424S	OJ1311_D08	AP004849	e-106
CB965895	26S proteasome regulatory particle triple-A ATPase	AB070252		142.5,143.0	S5424S, C253	OJ1311_D08	AP004849	0.0
BU672950	translation initiation factor 3, subunit g	AJ293728	AK103497	143.6	C1119	OSJNBa0054K20	AP005535	0.0
CB966667	progesterone-binding protein homolog Atmp2		AK062552	144.7	G275	P0700F06	AP005115	0.0
BU673403	RING finger-like protein	AL132971	AK104280	144.7-146.6		OJ1695_H09	AP004094	e-110

BI306341	endosomal protein-like	AP000421	AK069053	146.9	R2734	OJ1695_D07	AP005295	0.0
BU672777	NADH-ubiquinone oxireductase		AK072342	146.9	R2734	P0462E07	AP005616	e-143
BI306271	monogalactosyldiacylglycerol synthase		AK100941	147.2	S1730	P0689B12	AP005056	0.0
BI305379	molybdopterin synthase sulphurylase	AF124159		147.2-150.5		OJ1548_F12	AP004240	0.0
BI305232	ABC transporter	AC069158	AK063891	150.5	R3014A	OSJNBa0049O12	AC069158	
CB965189	casein kinase I			150.5	S13187	OJ1116_E04	AP004081	0.0
BU673203	ABC transporter family protein	NM_100271	AK099203	150.5	S13187	OJ1116_E04	AP004081	0.0
BI306534	hypothetical protein		AK060823	151.6	R1521	OJ1293_E04	AP004120	0.0
BI306691	cinnamoyl CoA reductase	AJ428493	AK107665	151.6	R1521	OSJNBa0053L11	AP005691	e-147
BI306584	putative protein		AK061270	152.7	C379, G1234	OJ1293_E04	AP004120	e-120
BU673882	brown planthopper susceptibility protein Hd002A	AAQ54304	AK104525	153.2, 154.1	R2821, R208	OJ1202_E07	AP004048	e-160
BI306519	pumilio/Mpt5 family RNA-binding protein	NM_128471		154.1	R208	OJ1202_E07	AP004048	0.0
BI305824	fibrillarin 2	NM_118695	AK071291	155.5, 156.3	C60012S, C1901	OJ1119_A01	AP004020	2e-75
BI305758	SR3 sucrose-regulated	U16257	AK070498	156.3	R1496	OJ1136_C04	AP004026	0.0
BU673158	auxin-regulated protein	NM_148656	AK108268	157.9		OJ1149_C12	AP004082	e-139
BI305572	small GTP binding protein RACDP	AF218381	AK067504	157.9		P0264G11	AP006187	1e-86
BI305485	No hit		AK101090	157.9		P0452F04	AP004776	9e-47

AccNo	Putative	ACC NO		cM	Marker	chromosome 3	BAC ACC	E-VAL
BI305790	lipoate-protein ligase B					OSJNBb0093J20	AC118674	5e-64
BI305833	expressed protein	NM_106146				OSJNBb0033N16	AC082645	2e-52
CB964817	zinc-finger-like protein					OJ1754_E06	AC104433	e-170
BI305560	ribosomal S29 -like protein	AF457936	AK059123			OSJNBb0094O03	AC092781	
CB967382	transcription coactivator-like protein		AK061882			OSJNBa0072G19	AC137599	0.0
CB966470	Auxin-responsive protein IAA3	AF123504	AK073044			OSJNBa0010N03	AC145379	e-106
BU673143	cytochrome P-450	U32624				OSJNBa0004B24	AC084319	2e-55
BI305985	beta-expansin	AF261275	AK061423			OSJNBb0080O10	AC118673	0.0
BU673878	HSP90-like protein	AY077617		0.0-1.1		OSJNBb0080O10	AC118673	0.0
BI305520	hypothetical protein			1.1	S1792	OSJNBb0043C10	AC105733	e-110
BI306204	unknown protein	AC087182		1.1	S1792	OSJNBb0043C10	AC105733	0.0
BU673543	membrane protein	L13655	AK061443	2.5	R1468B	OSJNBa0056G13	AC134236	422
BU673443	predicted protein		AK073382	2.5	R1468B	OJ1263H11	AC118980	5e -94
CB967383	phosphoribosylanthranilate transferase	AY224452		6.3	C1153	OJ1293G11	AC119747	e-176
CB964972	expressed protein		AK059624	7.6	C51476S	OJ1528D07	AC099739	e-107
CB966828	adenylate kinase		AK070255	7.6	C51476S	P0043E01	AP000615	0.0
CB966336	expansin-like protein 1	AY039022	AK099870	7.9-9.3		OJA1004C08	AC140005	1e-78
BU673645	glutathione S-transferase GST 13	AF244678	AK059226	9.3	R1713	OJ1006F06	AC099399	e-159
BI306740	60S ribosomal protein L17	AY054508	AK110621	10.9	C1279	OSJNBa0003A09	AC118132	e-147
CB964933	protein phosphatase 2C-like protein	NP_195564	AK061618	10.9	C1279	OSJNBa0003A09	AC118132	0.0
BU672804	gamma-Tip	D25534	AK104123	13.3	C725	OSJNBa0067N01	AC090485	0.0
BU673628	No hit		AK063904	13.3	C725	OSJNBa0067N01	AC090485	
CB966788	No hit			14.4	R3131	OSJNBb0050N02	AC105734	e-143
CB965867	hypothetical protein		AK063559	14.8	R1811	OJ1134F05	AC099401	e-174
BI305977	rab28 protein	X59138	AK106743	14.8	R1811	OJ1134F05	AC099401	e-180
BU673448	submergence induced protein 2A	AF068332	AK099497	16.8	E31168	OJ1607A12	AC105729	e-180
BI305516	ribulose-5-phosphate-3-epimerase	AF047444	AK099215	16.8	S13095	OSJNBa0091P11	AC073556	e-152
CB965292	No hit		AK065074	17.9	C3068S	OSJNBa0015N19	AC126221	0.0
BI306102	EF-1 alpha	D63580	AK072648	18.4	R518	OSJNBa0061L19	AC090484	0.0
BU673172	elongation factor 1 alpha	AF136826	AK073196	18.4	R518	OSJNBa0061L19	AC090484	0.0
CB965132	alpha 1 subunit of 20S proteasome	AB026558	AK101031	18.4-20.3		OSJNBa0050H14	AC125472	e-106
BI306645	hypothetical protein		AK106221	18.4-20.3		OSJNBa0050H14	AC125472	4e -87
BI306120	EREBP-like protein	AC079633	AK106163	20.3	S2769	OSJNBa0032G08	AC079633	2e-83
CB965374	AP2 domain containing protein		AK101949	20.3	S2769	OSJNBa0032G08	AC079633	0.0
BI305497	ethylene responsive protein	AJ304840		20.3		OSJNBa0032G08	AC079633	1e-22
BU673457	P40-like protein	AB012702	AK104842	20.3	S2769	OSJNBa0032G08	AC079633	0.0
CB966020	transcription factor like protein		AK066651	24.4	S879	OSJNBb0019D17	AC116604	e-139
BI306685	sugar transporter		AK099273	25.0-25.9		OSJNBb0058P18	AC116426	0.0
BI306420	cyc07	D26060	AK104699	25.0-25.9		OSJNBa0091J11	AC115687	e-117
BI306258	porin-like protein			25.9	C10916SA	OSJNBb0006P09	AC104429	e-103
BI305429	putative protein	AL391143		28.2	S12158	OSJNBa0021B21	AC104179	0.0
BI305682	Sec61 alpha subunit	AY044237	AK062070	28.2	C316	OSJNBb0085F02	AC134241	0.0
CB966912	expressed protein	NP_565300	AK058684	28.2	S12158	OSJNBa0021B21	AC104179	0.0
BI305615	MATE efflux family protein		AK101556	29.1	C30192	OSJNBa0090D11	AC105732	0.0
CB967251	Pridicted protein		AK103846	31.3	C563	OJ1626B05	AC104473	e-132
CB967444	hypothetical protein			31.3	C563	OJ1626B05	AC104473	0.0
CB964905	hypothetical protein		AK061960	31.3	C563	OJ1626B05	AC104473	0.0
BI306492	No hit			33.5, 34.8	S10968, G55	OJ1781E12	AC105927	e-137
CB965109	hypothetical protein		AK069495	36.1	C326	OJ1175C11	AC103891	0.0
BI305677	ubiquitin protein fused to a ribosomal protein	D12629	AK059011	36.1	C326, R2778	OJ1175C11	AC103891	e-105
CB967193	beta-ketoacyl-CoA synthase			36.9-37.7		OSJNBb0096L14	AC129008	0.0
CB967432	senescence-associated protein 15	AAC34858		36.9-37.7		OSJNBb0096L14	AC129008	0.0
BI306613	thaumatin-like protein			36.9-37.7		OSJNBb0042A11	AC134239	0.0
BI305797	enolase	U09450	AK099342	37.7	C706	OSJNBa0030J04	AC134235	5e-70
BU672783	transcription factor Hap5a-like protein	AY072402		39.0	R2293	OSJNBb0081I10	AC134240	0.0
BI306661	chorismate synthase 2	Z21791	AK099850	39.0-40.3		OSJNBa0019J12	AC134234	e-115
BI306322	Pridicted protein			40.3	R1868	OJ1041F02	AC135206	e-124
BI305269	MAP3K beta 1 protein kinase	AJ010093		40.3	R1868	OJ1041F02	AC135206	e-119
BI306072	Pridicted protein		AK067490	40.3	R1868	OJ1041F02	AC135206	0.0
BI306445	conserved hypothetical protein		AK104949	41.9	C60980S	OJ1364E02	AC135208	1e-52
CB965933	putative protein	AAL33781	AK099963	41.9	C60980S	OJ1364E02	AC135208	e-128
BI306130	protein kinase, putative	AC027135	AK100780	41.9	C60980S	OJ1364E02	AC135208	0.0
BU673038	No hit		AK068453	41.9	C60980S	OJ1364E02	AC135208	e-128
BI306637	hypothetical protein		AK101153	41.9	R867	OSJNBa0071M09	AC135209	0.0

BI306387	UDP-glucuronic acid decarboxylase	AB079064	AK100908	44.1	R134	OSJNBa0002D01	AC083942	0.0
BU673208	histone H2B	X82362		44.1	R134	OSJNBa0002D01	AC083942	0.0
BI305599	histone H2A	D38090	AK074018	44.1	R134	OSJNBa0002D01	AC083942	0.0
CB964997	H+-transporting ATP synthase chain 9	AY224460	AK066019	44.1	R134	OSJNBa0002D01	AC083942	0.0
BI306228	No hit		AK072759	44.1	R134	OSJNBa0002D01	AC083942	1e-33
CB964458	hypothetical protein		AK074018	44.1	R134	OSJNBa0002D01	AC083942	
CB965755	hypothetical protein		AK065887	44.1-44.4		OSJNBa0044H10	AC084405	1e-45
CB964533	wound-induced protein WI12		AK068151	47.7	R1158	OSJNBa0036G02	AC118668	0.0
CB966933	Ribosomal protein S7	AJ315794	AK059192	47.7	S941	OSJNBb0036D03	AC118670	e-144
BI306152	26S proteasome regulatory particle triple- A ATPase subunit2b	AB037154	AK058779	47.7	S941	OSJNBa0039F10	AC137931	e-162
CB964692	putative protein	NP_191814	AK071397	48.5-48.8		OJ1275B08	AC135908	2e-75
CB966707	No hit			48.8	R143	OJ1145F05	AC134769	4e-33
BI305554	seed maturation protein PM36	AF169021	AK073626	48.8	R143	OJ1145F05	AC134769	0.0
CB966346	nicotianamine synthase 2	AB023818		48.8	R143	OJ1145F05	AC134769	4e-87
CB964426	nicotianamine synthase	AB046401	AK102205	48.8	R143	OJ1145F05	AC134769	0.0
BI306726	enoyl CoA hydratase	AJ275305	AK068241	48.8-49.3		OSJNBa0016B07	AF461424	e-150
BI305344	serine/threonine kinase	Y12465		50.8	R3156	OSJNBa0042L15	AC119748	
BI306736	photosystem II D1 protein	D21291	AK103503	50.8-54.7		OSJNBb0093J20	AC118674	e-116
BI305659	novel protein, osr40c1	X95402	AK069815	54.7	C51380S	OSJNBb0014I10	AC126222	8e-72
BU673283	No hit			54.7	C51380S	OSJNBb0014I10	AC126222	5e-43
BU673806	expressed protein	NM_129622	AK070095	55.8	R3126	OSJNBb0098H14	AC125496	6e -51
CB967038	expressed protein		AK104744	59.0	E30305SB	OSJNBb0050D18	AC121491	7e-55
BI306686	putative protein	NM_116138	AK108277	59.5	C1135	OSJNBb0094H06	AC118675	e-125
BU673182	40S RIBOSOMAL PROTEIN S21	D12633	AK073724	59.5	C1135	OSJNBa0039O18	AC133930	3e -59
CB966179	phospholipid hydroperoxide glutathione peroxidase	AJ270955	AK073938	61.9-63.3		OSJNBb0033D20	AC137076	e-153
CB965405	phospholipid hydroperoxide glutathione peroxidase	AJ270955	AK073938	61.9-63.3		OSJNBb0033D20	AC137076	e-152
BI305605	vesicle soluble NSF attachment protein receptor	AC082644	AK064875	65.4	C12845S	OSJNBa0013M12	AC082644	0.0
BI305453	putative protein	NM_120129	AK062998	69.7	R2982	OSJNBa0048F08	AC091733	5e-78
CB964447	Pridicted protein		AK070724	72.4	C11260S	OSJNBb0058G04	AC103551	0.0
BI305445	3-deoxy-D-arabino-heptulosonate 7-phosphate synthase	Y14797	AK059247	73.5	R663	OSJNBa0017N12	AC092075	0.0
BI305828	salt-induced protein,lectin	AC084295	AK099576	76.6-83.0		OSJNBa0015K03	AC084295	0.0
CB965557	putative protein		AK099537	76.6-83.0		OSJNBb0074M06	AC133932	e-128
CB965565	unknown protein			84.4-85.2		OSJNBb0056B16	AC137992	3e-72
BI306484	expressed protein	NM_100442	AK101293	85.2	E30331S	OSJNBa0032H19	AC114983	1e-99
BI305877	ribosomal protein S19	D29730	AK099047	85.2	E30331S	OSJNBa0032H19	AC114983	e-172
CB966612	30S ribosomal protein S5	NP_180936	AK065820	86.0		OSJNBb0047D08	AC137925	0.0
BI305784	unknown protein	AC084404	AK100305	86.0	R3235	OSJNBa0026A15	AC084404	0.0
BI305799	expressed under carbonate stress	AB053296		86.0	S2274	OSJNBa0036E17	AC099041	3e-99
BU673058	LOB DOMAIN 37	AF447894	AK071624	86.0		OSJNBa0035N15	AC105743	0.0
CB965429	putative protein		AK065720	86.0	S3287	OSJNBa0054H04	AC106887	e-170
BI305936	unknown protein			86.0	S2274	OSJNBa0036E17	AC099041	0.0
CB966628	ATP synthase subunit		AK101433	86.0		OSJNBa0023A13	AC133337	e-150
BI305261	amino acid permease, putative	AC079041	AK071314	86.5	S2470	OSJNBa0008D12	AL607101	0.0
BI305527	hypothetical protein		AK104941	86.5	S2470	OSJNBa0008D12	AL607101	0.0
CB967259	expressed protein		AK059201	87.1	C816	OJ1785_A05	AC133333	e-168
BU673007	Chain A, Inositol Monophosphatase complexed With Manganese (Ii) And Phosphate		AK071149	87.4	G332	OSJNBa0010D22	AC133003	e-162
CB967023	expressed protein		AK070664	87.4-87.9		OSJNBa0027H16	AC137921	0.0
CB965595	chlorophyll a/b binding protein	AF061577	AK066762	87.9	S10558	OSJNBb0056O10	AC135564	e-127
CB965588	ribosomal protein L15.		AK104774	87.9, 88.7	G1316, R440	OSJNBa0087M10	AC109602	0.0
BI306662	reversibly glycosylated polypeptide	Y18624	AK098933	88.7	R440	OJ1523_A02	AC090874	e-126
BI306570	mitochondrial ATP synthase 6 KD subunit	AB055076	AK072384	89.2	E11534S	OSJNBa0091E13	AC133860	e-109
BI305547	unknown protein	AB010700	AK071403	91.1	C52104S	OSJNBb0007E22	AC136972	0.0
CB965240	No hit			91.1-94.9		OSJNBa0063J18	AC107206	e-103
BU673194	No hit			94.9	E1419S	OSJNBa0002I03	AC091246	e-119
BU672989	unknown protein			96.6-99.0		OSJNBb0029I19	AC128646	7e -33
CB964445	far-red impaired response protein	AAP50996	AK073458	101.6	E729S	OSJNBb0023J24	AC138004	e-168
CB967132	Pridicted protein		AK062338	101.6-101.9		OSJNBa0075A22	AC133859	e-110
CB967303	hypothetical protein		AK066416	101.6-101.9		OSJNBa0075A22	AC133859	0.0
BU672785	No hit		AK103539	122.3	R654	OSJNBa0022C08	AC097277	1e -42
BI305428	putative protein	AL031394	AK062756	122.3	R654	OSJNBa0022C08	AC097277	0.0
CB965390	unknown protein			122.8		OSJNBb0021P10	AC123974	3e-81
BU673604	beta-glucosidase	U28047	AK103027	122.8	R2847	OSJNBa0004L11	AC133334	0.0
CB965233	pescadillo-like protein		AK066652	122.8	S1803	OSJNBb0017F17	AC097368	0.0
CB966826	ARP protein		AK073541	126.8	R1538	OSJNBa0018H01	AC087181	0.0
BU673845	Pridicted protein		AK104418	126.8	R1538	OSJNBa0018H01	AC087181	e-180
BI305502	brain specific protein	D16140	AK103065	126.8	R1538	OSJNBa0018H01	AC087181	0.0
CB966296	glutathione-s-transferase	NP_176758	AK062406	126.8	R1538	OSJNBa0018H01	AC087181	6e-85
CB964809	phragmoplastin			127.7	S10087	OSJNBa0014G15	AC090882	0.0
CB966792	leucine rich repeat receptor-like kinase	Y07748	AK066295	128.3	R2690	OSJNBa0078A17	AC091532	0.0
BU673900	actin	AJ012685	AK100267	128.3	R2690	OSJNBa0078A17	AC091532	8e -87
CB966738	hypothetical protein			130.7-135.1		OSJNBa0005H20	AC135956	e-160
BU672932	No hit			135.1	R689	OSJNBa0079G12	AC103550	6e -33
BU672790	unknown protein	AC069273	AK105387	135.7	R2462	OSJNBa0027J18	AC096689	0.0
BI306414	hypothetical protein		AK071715	135.7	R2462	OSJNBa0027J18	AC096689	0.0
CB967274	unknown protein	AF360123	AK062628	135.7	R2462	OSJNBa0027J18	AC096689	0.0
BI306378	unknown protein		AK060542	136.0	E31057S	OSJNBb0016H12	AC118133	6e-99
CB965908	hypothetical protein		AK068510	136.0	E31057S	OSJNBb0016H12	AC118133	4e-42

BI305778	photoreceptor-interacting protein-like	AB013389		136.5	G249	OSJNBa0057G07	AC117988	0.0
BU673190	IAA1 protein	AJ251791	AK059838	137.6	S770	OJ1365_D05	AC096855	e-116
CB966583	hypothetical protein		AK071734	137.6	S770	OJ1365_D05	AC096855	e-148
BU673557	histone H2A.F/Z			137.9	C1401	OSJNBb0036F07	AC092558	0.0
BI305650	beta-D-glucan exohydrolase, isoenzyme ExoII	U46003	AK065044	138.7	C50518S	OSJNBa0069E14	AC091811	0.0
BU673110	cysteine synthase	AF073697	AK071279	138.7	R1618, C50518S	OJ1124_H03	AC087852	e-134
BI306485	exoglucanase precursor	U46003	AK065044	138.7	C50518S	OSJNBa0069E14	AC091811	e-134
BU673682	Unknown protein	AY054590	AK102404	139.8	E31254	OSJNBa0047E24	AC092556	e-142
BI306467	strictosidine synthase-like	AL589883	AK099001	139.8	E31254	OSJNBa0047E24	AC092556	e-119
BI305623	hypothetical protein		AK109888	139.8	E31254	OSJNBa0047E24	AC092556	9e-66
CB965289	hypothetical protein		AK099001	139.8	E31254	OSJNBa0047E24	AC092556	0.0
CB964598	UDP-glucose dehydrogenase	AAO62313	AK098880	140.1	C2540	OSJNBa0040E01	AC079887	0.0
BU673304	translation initiation factor, eIF-5A	AJ312906	AK060387	140.1	C2540, E3180S	OSJNBb0048A17	AC084282	0.0
BU673647	hypothetical protein	AJ271079	AK106979	143.4	C217	OSJNBb0106M04	AC107207	0.0
BI305607	transposase		AK105528	144.5	C1329	OSJNBa0010I09	AC084748	0.0
CB966384	photosystem-1 F subunit precursor	AF093634	AK069791	145.6	S13014	OSJNBa0010I09	AC084748	6e-66
BU673059	actin depolymerizing factor	AF112887	AK073162	146.1	R3226	OSJNBa0091J19	AC084320	e-154
BI306077	inosine monophosphate dehydrogenase	AJ010201	AK065289	146.1	R3226	OSJNBa0091J19	AC084320	0.0
BI305579	unknown protein	AC084320	AK103025	146.1	R3226	OSJNBa0091J19	AC084320	e-165
CB965825	glycolate oxidase	AF022740	AK098878	146.1-146.4		OSJNBb0024J17	AC133340	1e-70
BI305388	N2,N2-dimethylguanosine tRNA methyltransferase	AC009755	AK067936	146.4	S10656	OSJNBb0024J04	AC084296	0.0
BI306110	No hit		AK068731	149.1	R2224	OSJNBb0021G19	AC092076	0.0
BU672909	S18.A ribosomal protein	AF370463	AK058881	149.1	R2224	OSJNBb0060J21	AC090871	e-142
BU672868	histone-like protein	AF394115	AK099211	151.5	C50771S	OSJNBa0052F07	AC104321	e-176
BI306458	ubiquinol–cytochrome c reductase	X79276	AK103963	151.5	C1442	OSJNBa0059E14	AC135958	e-163
BI305493	histone-like protein		AK067840	151.5	C50771S	OSJNBa0052F07	AC104321	0.0
BI306629	ribosomal protein S15 gene	D10962	AK059639	151.5	C50771S	OSJNBa0052F07	AC104321	0.0
CB967165	BCS1 protein		AK104696	151.5	C393A	OSJNBa0087C10	AC108906	0.0
BI305556	ribosomal protein L36		AK063320	152.3	R3020	OSJNBb0033J23	AC137507	e-135
BI306525	cyclophilin	AJ132763	AK071076	152.3	R3020	OSJNBb0033J23	AC137507	0.0
CB964476	putative protein		AK104712	152.3-153.7		OSJNBa0024F18	AC135594	e-179
BI306455	hypothetical protein		AK066670	153.7	E10579S	OSJNBa0094J08	AC133007	0.0
CB965041	chalcone isomerase	AF474922	AK061390	153.7	E10579S	OJ1754_E06	AC104433	e-177
BU672769	unknown protein		AK066983	153.7	E10579S	OJ1754_E06	AC104433	4e-40
CB965469	hypothetical protein		AK102675	153.7	E10579S	OJ1754_E06	AC104433	4e-55
CB965635	Pridicted protein		AK099739	153.7	E10579S	OSJNBa0094J08	AC133007	e-135
BI305593	Ras-like GTP-binding protein	NM_111825	AK068609	153.7	E10579S	OJ1754_E06	AC104433	e-133
CB966992	hypothetical protein		AK072450	153.7	E10579S	OJ1754_E06	AC104433	0.0
BI306482	Bowman-Birk serine protease inhibitor			154.8	R273	OSJNBb0081B07	AC093018	e-111
CB967474	unknown protein		AK067741	155.8, 156.3	C1219, C1164	OSJNBa0010E04	AC096687	0.0
BU673687	unknown protein		AK064860	157.7	S1509	OSJNBa0042I09	AC104487	1e -97
BI306725	unknown protein		AK060311	158.2	C1354	OSJNBb0043P23	AC099324	e-136
CB965623	sarcoplasmic reticulum protein		AK101661	158.2	R1925	OSJNBa0032G11	AC092852	e-147
BI306268	unknown protein	AC010657	AK069826	159.0	G1318	OSJNBb0096M04	AC092559	0.0
BI306269	possible apospory-associated protein	U13149	AK068082	160.1	C63655SB.	OSJNBa0015N08	AC096688	0.0
BI305361	proteasome regulatory non-ATPase subunit		AK071162	160.1	C63655SB	OSJNBa0015N08	AC096688	4e-58
BU673366	CEO protein	AC027037	AK072499	161.7	S15179S	OSJNBb0006O08	AC120506	e-152
BI306386	vacuolar H+-ATPase	U27098	AK105044	161.7-164.4		OSJNBb0062G19	AC128647	0.0
CB967400	unknown protein	NP_172989	AK058786	161.7-164.4		OSJNBb0062G19	AC128647	7e-62
BI305800	ribosomal protein 30S subunit	M55322	AK071750	161.7-164.4		OSJNBb0062G19	AC128647	0.0
BU673315	No hit			161.7-164.4		OSJNBb0062G19	AC128647	1e -48
BI306530	hypothetical protein		AK098951	164.4	S13122	OSJNBa0059G06	AC096690	e-110

AccNo	Putative	ACC NO		cM	Marker	chromosome 4	BAC ACC	E-VAL
BU673348	nonphototrophic hypocotyl 1b	AB018443				OSJNBa0079M09	AL731609	
CB967429	Pridicted protein					OSJNBa0006M15	OSJN00224	0.0
BI306409	Pridicted protein					OSJNBa0012L21	AL607102	0.0
BI306609	No hit					H0806H05	AL442113	6e-24
BI305801	gag-pol protein	AC084767				B0311F12	AL512548	3e-90
CB966516	ubiquitin-conjugating enzyme E2-17 kDa		AK066232			OSJNBa0043A12	AL606619	1e-82
CB966403	No hit					OSJNBb0088C09	AL731628	e-122
BI305404	hypothetical protein	AP002871				t17804	AL117265	2e-28
BI305826	expressed protein		AK061755	3.1	C60436S	OSJNBa0094O15	AL662935	0.0
BI305872	unknown protein	AC008261	AK060395	6.5	S5925S	OSJNBb0050O03	AL606631	0.0
BI306117	elongation factor EF-2	AY054461	AK103035	6.8	C56	OSJNBa0020P07	AL606450	0.0
BI306434	22 kDakafirincluster	AC080019	AK065870	10.1-11.0		OSJNBb0028M18	AL606654	6e-75
CB964566	Pridicted protein			14.3-16.7		OSJNBa0045O17	AL663018	
CB967485	Pridicted protein			14.3-16.7		OSJNBa0045O17	AL663018	
BI306487	hypothetical protein		AK060268	14.3-16.7		OSJNBa0059D20	AL606997	0.0
BI306711	hypothetical protein			16.7-18.3		OSJNBB0026E15	AL607008	6e-78
BI305242	hypothetical protein		AK067629	18.3-19.6		OSJNBa0052O21	AL606590	
CB965095	glycerol-3-phosphate dehydrogenase		AK104305	19.6	R288	OSJNBb0072N21	AL606634	e-152
BU673822	plasma membrane major intrinsic protein 2	U60148	AK102155	19.6		OSJNBb0093G06	AL731636	0.0
BI306027	No hit		AK058584	19.6	C51175SB	OSJNBb0026i12	AL663002	e-142
CB966911	RING-H2 finger protein RHA2a	NP_172962	AK059146	19.6	C53648S	OSJNBb0043H09	AL606611	0.0
CB965741	No hit			19.6		OSJNBb0093G06	OSJN00281	6e-27
BU673677	Pridicted protein		AK065174	19.6	G124B	OSJNBa0060B20	AL662948	e-150
BI306580	protein phosphatase	AY065090	AK060035	19.9		OSJNBb0067G11	AL663014	0.0
CB966899	eukaryotic release factor 3 homolog	T10058	AK105823	19.9		OSJNBb0067G11	AL663014	e-156
CB964866	expressed protein	NP_191557	AK108051	23.3		OSJNBa0079M09	AL731609	e-105

BU673418	jasmonate-induced protein	X98124		23.3, 24.1	R10945S, E61662S	OSJNBa0033H08	AL662942	0.0
BU673153	Pridicted protein			23.3, 24.1	R10945S, E61662S	OSJNBa0024J22	AL731596	4e -75
CB965518	auxin-induced protein	AAB71969	AK104524	24.9	S5217S	OSJNBa0008A08	AL606589	0.0
BI305467	ribosomal protein S8	U64436	AK068316	30.8	E61384S	OSJNBa0038P21	AL731588	0.0
BU673187	beta-oxyacyl-[acyl-carrier protein] reductase	AJ243091		30.8-41.5		OSJNBb0006N15	AL607003	
BU673909	Pridicted protein			30.8-41.5		OSJNBb0006N15	AL607003	
BI306070	putative protein	NM_125745	AK063862	44.0-48.3		OSJNBa0044M19	AL731601	0.0
BI306094	thioredoxin M		AK061185	52.6-56.1		OSJNBb0012E08	AL606610	e-115
BI305224	protein phosphatase-2c~gene	AB022217		56.1	E11538S	OSJNBB0118P14	AL607005	
CB967028	hypothetical protein		AK069797	58.6	G271	OSJNBa0073L04	AL662947	0.0
CB965823	expressed protein	NP_190845	AK100599	58.6	G271	OSJNBa0073L04	AL662947	e-112
BI306033	hydroxyproline-rich glycoprotein	X61280	AK102919	60.2	C2043	OSJNBb0108J11	AL606618	2e-25
CB966566	Pridicted protein		AK070730	62.1	R93	OSJNBa0042L16	AL606632	0.0
BI305725	expressed protein	NM_105860	AK058925	62.6	R2406	OSJNBa0084A10	AL606458	0.0
CB965224	Pridicted protein		AK067009	66.4	C558	OSJNBa0027P08	AL731593	
BI305935	DNA binding protein	NM_103643	AK069926	67.2	R1721	OSJNBa0064H22	AL606448	e-161
CB964452	Pridicted protein			67.5	C2807	OSJNBa0086B14	AL606615	0.0
BU673196	putative protein	NM_118571	AK100647	67.5	C2807	OSJNBa0086B14	AL606615	e-140
BI306447	expressed protein	NM_127785	AK065178	67.8, 68.3	C10736S, E3080S	OSJNBa0036B21	AL606636	0.0
BI306218	No hit			67.8,68.3	C10736S, E3080S	B0518A01	OSIG00038	2e-69
CB967012	pollen allergen-like protein		AK059231	70.1-70.6		OSJNBb0048E02	AL606653	0.0
CB965946	hypothetical protein		AK069067	70.1-70.6		OSJNBb0048E02	AL606653	e-147
BI306134	cell division protein FtsH-like protein	NM_111112	AK068686	70.6	S1408	OSJNBa0016O02	AL606588	e-158
BU673400	indole-3-glycerol phosphate synthase	AB023039	AK101835	70.6	S1408	OSJNBa0016O02	AL606588	1e -96
BI306165	hypothetical protein		AK102076	72.0-72.3		OSJNBa0053D18	AL606591	0.0
BI306288	copper amine oxidase	NM_129810	AK059867	72.0-72.3		OSJNBa0053D18	AL606591	8e-66
BI306457	cytosolic glyceraldehyde3phosphate dehydrogenase GAPDH	AF251217	AK064960	74.5	R3494	H0302E05	AL627350	e-152
CB967360	hypothetical protein	NP_199119	AK104248	74.5	R896	oj991113_30	AL662946	0.0
BU672803	chitinase-B	AF402939		75.9	E1135S	OSJNBb0091E11	AL606629	0.0
BU673061	zinc finger protein	AB001883	AK061664	75.9	E10707S, E1135S	oj990528_30	AL662938	
BI305534	transporter-like protein	AL356014	AK063469	75.9	R1849	OSJNBa0084K20	AL606613	e-158
BI306544	unknown protein	NM_111635		76.5	C335	OSJNBa0067K08	AL606627	0.0
BI306353	quinone oxidoreductase -like protein	NM_121703	AK109382	76.5	C335	OSJNBa0067K08	AL606627	0.0
CB964624	expressed protein	NP_188108	AK100471	76.5	C335	OSJNBa0067K08	AL606627	e-131
BI306436	N-hydroxycinnamoyl/benzoyl transferase	AL442115	AK104637	76.8	R2226	H0711G06	AL442115	0.0
CB967411	Pridicted protein		AK103812	76.8	R2226	OSJNBa0029H02	OSJN00025	2e-85
CB966342	ribosomal protein L25	AF061508		76.8	R2226	OSJNBa0029H02	AL606594	e-180
BI306384	Expressed protein	NM_119249		76.8	R2226	OSJNBa0029H02	AL606594	0.0
BI306037	unknown protein	NM_127578	AK101840	77.9-78.2		OSJNBa0043L24	AL662969	1e-29
BU673385	CAF-like protein	NM_122039	AK066586	78.2	C1238	OSJNBb0065L13	AL606607	
BI306496	No hit			78.2	R10501S	OSJNBa0043L24	AL662969	e-128
BU673699	glutaredoxin	D86744	AK065962	78.2	R10501S	OSJNBa0043L24	AL662969	
CB964826	Pridicted protein		AK065832	78.2-81.7		OSJNBa0019D11	OSJN00156	0.0
CB966881	expressed protein	NP_565474	AK101663	78.2-81.7		OSJNBa0073E02	AL731616	e-149
CB965143	putative protein	NP_200876	AK065832	78.2-81.7		OSJNBa0019D11	AL662958	0.0
CB966590	bZIP DNA-binding factor		AK105586	81.7	R3351	OSJNBa0081C01	AL662984	
BI305699	remorin 1	AF123265	AK106182	81.7-82.5		OSJNBb0039L24	AL663006	0.0
CB966568	No hit		AK071369	81.7-82.5		OSJNBb0020O11	AL662998	0.0
BU673901	gt-2 gene	X68261	AK107139	83.0	G282	OSJNBa0011L07	AL606587	
BU673913	AP2-related transcription factor			84.1	C11378	OSJNBa0079A21	AL607006	0.0
BI305629	high mobility group protein HMG-beta2			87.1-94.4		OSJNBa0017B10	AL606628	
BI305753	hypothetical protein		AK100392	87.1-94.4		OSJNBA0088I22	AL607001	0.0
CB967097	No hit			87.1-94.4		OSJNBb0032E06	AL663003	7e-55
BI306240	hmgc1 gene	AJ131374	AK105145	87.1-94.4		OSJNBa0017B10	AL606628	1e-37
CB966601	beta-carotene hydroxylase	CAB55625	AK060559	94.4-96.0		OSJNBa0011J08	AL606624	e-178
BI305304	OsCDPK7	AB042550	AK066495	97.4, 97.7	C79, S13536	OSJNBa0088A01	OSJN00188	0.0
BI306067	OsCDPK7	AB042550	AK059786	97.4, 97.7	C79, S13536	OSJNBa0088A01	AL662987	0.0
BU673507	vacuolar membrane ATPase subunit G	AF181688	AK059032	100.7	S2486, R2785	OSJNBa0083N12	AL606683	
BI305474	No hit		AK100925	100.7	R2785	OSJNBa0041A02	AL606638	e-151
CB965119	transcription factor GT-3b	AAP13348	AK105720	100.7	R2785	OSJNBa0041A02	AL606638	0.0
BU673288	ascorbate peroxidase (TL29)	AJ251882	AK069838	100.7	R2785	OSJNBa0041A02	AL606638	
CB967252	WRKY family transcription factor	NP_849559	AK072938	101.8	G102	OSJNBb0015N08	AL662996	8e-57
BI305565	ribosomal protein L7 gene	D29720	AK073472	101.8	G102	OSJNBb0015N08	AL662996	0.0
CB965789	Pridicted protein		AK102588	102.1	R514	OSJNBa0035M09	AL662968	0.0
BI305727	unknown protein		AK067094	102.7	C810	OSJNBa0060D06	AL606691	2e-74
BI305769	hypothetical protein		AK063022	102.7-107.4		OSJNBa0085I10	AL606684	e-164
BU673054	fatty acid elongase-like protein	X88779		102.7-107.4		OSJNBa0085I10	AL606684	
BI305332	I1332.4putative protein	AJ243961		107.4	E50452S	I1332	AJ243961	
BI305803	symbiosis-related protein	NM_104884	AK062573	108.2	S1544	OSJNBa0053K19	AL606645	e-126
CB965303	anthocyanidin reductase	CAD91911	AK059518	109.9	R78	H0323C08	AL732334	4e -47
BI305268	helicase-like transcription factor	AP003224	AK102284	109.9	R78	OSJNBa0089N06	AL662988	0.0
CB966574	wound-induced protein	T01970	AK069933	111.3	L116	OSJNBa0043L09	AL606444	0.0
CB965004	cytochrome P450		AK104741	111.3	L116	OSJNBa0043L09	AL606444	
CB967077	hypothetical protein		AK063036	111.3-113.2		OSJNBb0003B01	AL606649	0.0
BU673790	No hit		AK071468	113.2	R740	H0721B11	AL732335	2e -14
CB966894	Floral homeotic protein APETALA2		AK099991	113.2	R740	OSJNBa0010D21	AL606635	0.0
CB965219	AP2 domain transcription factor		AK061095	113.2	R740	H0609A12	AL512544	e-108
CB965369	Pridicted protein		AK061488	114.0	R634	OSJNBb0059K02	OSJN00105	3e-48
BI306722	hypothetical protein			114.3	C1016	OSJNBa0060N03	AL606690	0.0
CB964667	expressed protein	NP_194768	AK104316	114.3-120.3		OSJNBa0015K02	AL606608	7e-66

BU673671	amylogenin	Y18623	AK071012	114.3-120.3		OSJNBa0015K02	AL606608	0.0
BI306327	Protein phosphatase 2C-like protein	NM_122403	AK072534	114.3-120.3		OSJNBa0011F23	AL662953	0.0
BU673356	precursor chloroplastic glutamine synthetase	AF480497	AK099252	114.3-120.3		OSJNBa0015K02	AL606608	e-153
BI305706	No hit		AK069874	120.3-122.9		OSJNBa0087O24	AL606646	5e-51
CB967019	disease-resistent-related protein		AK104901	120.3-122.9		OSJNBa0086B14	AL606615	
CB967087	ubiquitin-conjugating enzyme			122.9	R1427	B0811B10	AL732340	e-141
BI305983	thioredoxin-like protein		AK067891	122.9	R1427	H0818H01	AL732346	e-130
CB967044	amine oxidase family	NP_181830	AK106044	122.9	R1427	OSJNBa0043A12	AL606619	0.0
BI306017	unknown protein		AK106330	122.9	R1427	OSJNBa0043A12	AL606619	3e-50
BU673346	amine oxidase	NM_129863	AK072414	122.9, 123.8	R1427, R2231	H0624F09	AL732356	0.0
BU672839	unknown protein		AK071744	128.5	C107	OSJNBa0032F06	AL606641	e-111
CB967281	elicitor-responsive gene	AAM63058	AK070539	128.5	C107	OSJNBa0032F06	AL606641	e-109
CB966582	No hit		AK066799	129.1	C445	OSJNBa0088H09	AL606651	1e-44
CB966958	hypothetical protein		AK065557	129.1	C445	OSJNBa0088H09	AL606651	0.0
BU673488	diadenosine 5',5'''-P1,P4-tetraphosphate hydrolase		AK071345	129.1, 129.6	C445, E3142S	H0723C07	AL732347	e-129
CB967338	unknown protein			129.6	E3142S	OSJNBa0070M12	AL606686	2e-84
BI306589	glucose 6 phosphate/phosphate translocator, putative	NM_102035	AK103611	129.6		OSJNBb0020J19	AL606656	0.0
CB966539	30S ribosomal protein S17	AF095707	AK099298	129.6		OSJNBb0020J19	AL606656	0.0
AccNo	**Putative**	**ACC NO**		**cM**	**Marker**	**chromosome 5**	**BAC ACC**	**E-VAL**
BU672930	putative protein	AL353994	AK062080			OJ1668B04	AC104715	0.0
CB965495	No hit					OJ1575_E09	AC104714	8e-49
CB966357	chloroplast ORF109		AK110818	0.0-3.0		P0036D10	AC073405	0.0
BU673732	unknown	AY086935	AK066339	3.0	C2526	P0668H12	AC084818	0.0
BU673225	unknown	AY086935	AK062996	3.0	C2526	P0668H12	AC084818	0.0
BI305337	No hit		AK105753	3.0	C2526	P0036D10	AC073405	
BI305762	transcription factor IIA small subunit	AC079022	AK065182	4.6	C568	P0574H01	AC079022	0.0
BI305631	cytochrome B5	X75670	AK064821	4.6	C568	P0574H01	AC079022	e-117
BI305442	tRNA synthase	AC079022	AK065542	4.6	C568	P0574H01	AC079022	0.0
CB965547	No hit		AK099779	4.6 ,6.5	C568 ,S12936	P0016H04	AC079356	6e -76
BI306237	metallothionein-like protein	U57638	AK058529	6.5	S12936	P0016H04	AC079356	0.0
BI305712	amino acid selective channel protein	AJ011921	AK101693	6.5	S12936	P0016H04	AC079356	0.0
BI306659	metallothionein-like protein	U77294	AK058529	6.5	S12936	P0016H04	AC079356	3e-44
BI306418	R2R3MYB-domain protein	AF099432	AK099223	8.1,9.5	C2067, E732S	OJ1654B10	AC108504	0.0
BI306705	legumin-like protein	AP003627	AK058320	9.5	E732S	OJ1654B10	AC108504	e-162
BI306300	legumin-like protein	AP003627	AK058320	9.5	E732S	OJ1654B10	AC108504	0.0
BI306174	SF16 protein		AK105073	14.1	S2136	P0008A07	AC079021	3e-67
BU672916	zinc transporter		AK100735	17.9	E1455S	P0017E12	AC130608	0.0
CB964798	unknown protein		AK058597	19.0	R3166	OSJNBa0077L08	AC118288	e-163
BU673371	peroxidase		AK061809	20.1	R2213S	P0519E07	AC087552	0.0
BI305418	S-adenosylmethionine synthetase	AJ296743	AK104875	20.1	R2213S	P0519E07	AC087552	0.0
CB964882	aspartic proteinase	D12777	AK065206	21.1	R3332	OJ1127_B08	AC093490	3e-89
CB964750	early drought induced protein	AF503583	AK099873	21.7	R708	OSJNBa0069I13	AC137616	e-144
BU673487	glutaredoxin	AP001305		22.5-24.7		OJ1504G04	AC105772	e-180
BI306527	PWWP domain			22.5-24.7		OJ1504_G04	AC105772	1e-57
BU673352	tonneau 2	AF280057	AK070081	22.5-24.7		OJ1504G04	AC105772	0.0
CB967040	No hit			27.7	R1838, C119	P0431G05	AC087551	1e-21
CB966108	60S RIBOSOMAL PROTEIN L18		AK061929	27.7	R1838	P0431G05	AC087551	3e-77
BI305945	disulfide isomerase A6 precursor (P5)		AK098931	27.7	R1838, C119	P0676G05	AC087425	e-115
BI305899	VIP2 protein	AFA251051	AK104006	27.7	R1838	P0431G05	AC087551	0.0
BI306440	putative protein	NM_120678	AK110485	29.1, 30.7	S12447,E4433S	OJ1231F08	AC104276	e-118
BI306209	RING finger protein	Y09539	AK073342	30.7	S14280S	OSJNBb0099P06	AC124144	0.0
BI306005	expressed protein	NM_111934	AK066130	31.5	R2232	P0419C04	AC084817	3e-85
CB964962	beta 2 subunit of 20S proteasome	AB026564	AK103126	37.2	R3572	OJ1097_A12	AC093954	e-125
CB965585	hypothetical protein			38.3	S10613S	OJ1116_A10	AC093489	e-119
BI305398	metallothionein-like protein	AF009959	AK058313	42.2-44.7		P0015G13	AC135920	2e -76
BU673424	No hit		AK069043	45.1	S1873	OSJNBb0016G07	AC130599	0.0
BU672774	RD22		AK106205	46.9, 47.2	C282A, E11511S	OJ1076H08	AC108498	2e -58
CB967484	Dehydration-responsive protein RD22 precursor		AK104940	46.9, 47.2	C282A, E11511S	OJ1076H08	AC108498	e-109
BU673270	proline-rich protein RiP-15	AF221552	AK100217	49.4	R566	OJ1212C10	AC104275	e-178
BI305256	vegetative storage protein	L20233	AK101642	49.4	R566	OJ1212C10	AC104275	e-166
BI305750	class III chitinase homologue	AB027426	AK065866	53.5	S2351	OJ1037G10	AC104270	0.0
BU673290	chitinase	D55708	AK062114	53.5	S2351	OSJNBa0037H06	AC119290	0.0
BI305698	putative protein		AK071399	54.3-54.6		OSJNBa0018H09	AC137610	0.0
CB965535	acetyl-CoA carboxylase	AF359519		54.6	R2558	P0411C02	AC135922	0.0
BI305873	hypothetical protein	AP000367		54.6-55.4		OSJNBb0022L07	AC136225	2e-99
CB966604	hsp 70-like protein	AF074969	AK060410	54.6-55.4		OJ1234_D05	AC098572	0.0
BU673715	manganese superoxide dismutase	L34039	AK104160	58.7	S32S	P0018A03	AC134346	e-133
BI305853	mitochondrial carrier protein	AC002535	AK072064	62.7,64.1	E31132S, E2801S	OJ1045C06	AC104272	0.0
BI306270	No hit			65.5	C62663S	P0692D12	AC135929	2e-98
BI306194	No hit			70.5-73.9		OSJNBa0088M05	AC136222	7e-60
BI305584	U6 snRNA-associated Sm-like protein-like	AB012242	AK059190	70.5-73.9		OJ1562H01	AC105773	1e-61
BU673787	Expressed protein	NM_118306	AK071730	75.0	C2269S	OSJNBb0006J12	AC120991	e-148
CB967178	plasma membrane associated protein	NP_197398	AK062513	75.0	S811, C2269S	OSJNBa0035J16	AC135418	e-147
BU672779	ubiquinol–cytochrome-c reductase-like protein	AL353912	AK062745	75.0	S811, C2269S	P0605G01	AC132492	e-170
CB966116	protein kinase family	NP_175879	AK106842	75.0	S811	P0605G01	AC132492	0.0
BI306096	predicted protein			75.0	S811	OSJNBb0092G21	AC134932	0.0
BI305681	cytochrome P450	AB023038		75.0-77.4		OJ1174H11	AC104708	0.0
BU673716	beta-N-acetylhexosaminidase-like protein	AL132954	AK102449	77.4-78.2		OJ1123C08	AC108875	e-126

CB967107	nucleic acid binding protein	T02745	AK059311	80.4	R3103	P0426G01	AC137623	e-147
CB966830	hypothetical protein	NP_201175		80.4	R3103	P0426G01	AC137623	e-139
CB965612	hypothetical protein		AK061637	80.7	S10569	OSJNBa0044P19	AC135419	0.0
BI306059	OSMYB1	D88617	AK066834	85.7	C308	OSJNBb0048I21	AC130600	0.0
BU673198	unknown	AY087697	AK058810	86.0	E4361S	OJ1378A04	AC108502	e-173
CB966586	ribosomal protein A2		AK058730	88.5-89.6		P0615D12	AC137004	e-127
BI305492	unknown protein		AK108341	89.6	E31389S	OJ1576F01	AC097176	0.0
CB964985	50S ribosomal protein L33	NP_187283		89.6-92.0	C43	OJ1576_F01	AC097176	0.0
CB966820	No hit		AK107796	94.5	C903	OJ1525_A02	AC108876	202
BU673792	60S RIBOSOMAL PROTEIN L36	AL132960	AK058918	94.5	C903	OJ1281H05	AC117265	e-123
BU672964	No hit		AK099763	94.5	C903, C128	OJ1281H05	AC117265	e-158
CB965843	calcium dependent protein kinase		AK101942	95.3	C2161	OJ1387_F08	AC108503	e-134
BI305573	hypothetical protein		AK073713	95.3	C2161	OJ1280A04	AC108500	0.0
BI305244	ornithine/acetylornithine aminotransferase	AF022915	AK060423	96.4	R2117A	P0009H09	AC144743	9e-85
BI306501	ADP-ribosylation factor	D17760	AK067029	101.5	R3149S	OJ1119H02	AC097175	e-145
BI305768	ribosomal protein L30	AF034949	AK062929	101.5	R3149S	OJ1119H02	AC097175	e-131
BI305965	ADP-ribosylation factor	D17760	AK067029	101.5	R3149S	OJ1119H02	AC097175	4e-56
BI306086	hypothetical protein	NM_117949	AK071818	101.5-102.8		OSJNBa0088I06	AC129718	0.0
BU673080	osERF3	AB036883		102.8	G81	OJ1118C04	AC108523	e-147
CB966407	chloroplast phosphoglycerate kinase	X15233	AK062214	102.8	G81	OJ1118_C04	AC108523	e-144
BU673762	thioredoxin h	AB053294	AK106758	104.4	S11241	OSJNBb0108E17	AC130601	e-157
BU672987	trehalose-6-phosphate synthase		AK072066	104.7-107.4		OSJNBb1130G10	AC130603	e-161
CB967004	protein kinase	NP_193628		104.7-107.4		OSJNBb1130G10	AC130603	0.0
CB964631	expressed protein		AK098851	104.7-107.4		P0022D06	AC132485	0.0
CB964504	trehalose-6-phosphate synthase	NM_106505	AK067066	107.4	E4473	P0599F04	AC132491	e-151
CB965594	hypothetical protein		AK102422	107.4	E4473	P0599F04	AC132491	4e-34
CB965436	SnRK1a protein kinase	AB101655	AK067158	108.5	C11368	OJ1131_E09	AC111015	5e-69
BU672928	expressed protein	NM_130219		109.0	C2782A	OJ1362_G11	AC104713	6e-70
BI305918	expressed protein	NM_130219		109.0	C2782A	OJ1362G11	AC104713	e-151
CB964553	FEN-1	AB021666	AK103819	109.0	C2782A	OJ1362_G11	AC104713	5e-84
BI305507	unknown protein		AK071703	109.0	C2782A	OJ1362G11	AC104713	0.0
BI305670	hypothetical protein		AK106948	109.0	C2782A	OJ1362G11	AC104713	0.0
CB967047	hypothetical protein		AK062488	109.0	C2782A	OJ1362_G11	AC104713	0.0
CB966889	Pridicted protein			109.0	C2782A	OJ1362_G11	AC104713	e-174
BI305649	expressed protein	NM_102725	AK058330	110.7	R3085	OJ1345B12	AC104278	e-149
BU673587	nonspecific lipid transfer protein	U88090	AK062463	111.3, 111.0	G1103, S11116	P0494H05	AC130725	0.0
BI306008	ribosomal protein L33	AF334840	AK102775	112.4	S974	OJ1214E03	AC104709	1e-95
BU673655	ribosomal protein L35A	AF448416	AK102775	112.4	S974	OJ1214E03	AC104709	e-124
BU672967	putative protein	NM_116040	AK058980	112.4	S974	OJ1214E03	AC104709	e-157
CB964938	hypothetical protein		AK072756	112.4-115.7		OSJNBa0001A14	AC144735	e-131
BU673017	hypothetical protein		AK072829	115.7	S14121	OSJNBb0053D02	AC124143	0.0
BU673112	IMP dehydrogenase/GMP reductase	ZP_00107581	AK104836	115.7	S14121	OSJNBb0053D02	AC124143	7e -85
BI305531	FtsJ cell division protein			116.2	R3134	OJ1781H11	AC120986	0.0
BI305543	mitochondrial malate dehydrogenase	X78800	AK073698	116.5-117.9		OJ1735C10	AC104284	0.0
BI306331	CTP synthase	AP002897	AK070411	116.5-117.9		OJ1735C10	AC104284	e-110
BI305885	acetohydroxy acid isomeroreductase	AJ251333	AK065295	116.5-117.9	-	OJ1735C10	AC104284	e-100
CB966981	hypothetical protein		AK100204	117.9,119.0	C444, E4370S	OJ1651_G11	AC098573	0.0
CB965176	expressed protein		AK072537	119.0, 120.6	E4370S, S11036	OSJNBb0035N21	AC134929	2e-83
CB966136	No hit		AK070346	120.6	S11036	OJ2097_B11	AC098836	7e-98
CB964762	UDP-glucose 4-epimerase	AB087745		122.3		OSJNBa0030I14	AC136217	1e-85
AccNo	**Putative**	**ACC NO**		**cM**	**Marker**	**chromosome 6**	**BAC ACC**	**E-VAL**
BI306332	No hit	AP000367	AK063737			OSJNBa0041F13	AP003708	2e-49
BI305401	putative protein	NM_12301	AK071277			P0036H01	AP004724	0.0
BI306235	EF-hand Ca2+-binding protein CCD1	AF181661				P0009H10	AP003766	
BU673472	plastocyanin precursor	Y00704	AK070447	0.6	S15909	P0681G01	AP005621	e-112
CB965372	hypothetical protein		AK067892	0.9	C1003B	OSJNBa0075G19	AP002842	e-117
CB965236	unknown protein		AK106556	1.4	S924	P0644B06	AP001129	e-168
BI305433	Scl1 protein	AF149807	AK106868 ASP	1.4	S924	P0644B06	AP001129	0.0
CB964549	polygalacturonase		AK105782	1.4	S924	P0644B06	AP001129	8e-41
BU673221	unknown	AY086234		1.4	S924	P0644B06	AP001129	2e-67
BI306078	RING-H2 finger protein RHA1a	AF078683		1.7, 2.8	G8018, G8017	P0514G12	AP000616	0.0
BU673667	chaperonin	Z68903	AK101334	3.6	S1515	P0541H01	AP001389	e-136
BU672972	pathogenesis-related protein	U20347	AK066771	3.6	S1515	P0541H01	AP001389	e-127
CB965202	Acyl-CoA-binding protein		AK058833	3.6	S1515	P0541H01	AP001389	1e-77
BU673373	unknown protein	AY035178	AK098848	3.6	S1515	P0541H01	AP001389	e-170
BI306208	60 KD Chaperonin Beta subuni	Z68903	AK101334	3.6	S1515	P0541H01	AP001389	1e-94
BI306431	hypothetical protein		AK108643	5.1	S2107	P0542E10	AP003456	0.0
BI305997	DREB1A	AB007787		6.0-6.3	R3139	OSJNBa0038F22	AP002838	7e-50
CB964551	No hit		AK062712	6.9, 8.2	C425A, E10330	P0493C11	AP000559	e-134
CB966914	No hit		AK062712	6.9, 8.2	C425A, E10330	P0493C11	AP000559	5e-86
CB967287	40S subunit ribosomal protein	D12632	AK062184	8.2	E10330, C76A	P0679C08	AP002542	0.0
BI306238	putative protein	NM_124534	AK105690	8.2	E10330	P0679C08	AP002542	0.0
CB965291	No hit		AK105053	8.2	E10330, C76A	P0679C08	AP002542	4e-45
BI305867	RING3-like bromodomain protein	AC004238	AK063880	9.0, 9.8	C52026, C1084	OSJNBa0041F13	AP003708	e-124
BU673824	cytochrome c oxidase subunit Vb		AK071423	10.4	R1952	P0535G04	AP000399	0.0
CB965658	Pridicted protein		AK058823	10.4	R1952	P0535G04	AP000399	0.0
BI305646	profilin	U49505	AK058798	10.7	R2749	P0710H01	AP004806	0.0
CB966430	mitogen-activated protein kinase	AJ535841		11.5	C3010S, E1178SA	B1160E02	AP004844	0.0
CB967209	root-specific GOS9	X51909		12.9-13.5		OSJNBa0033B09	AP002864	0.0
CB967490	wall-associated kinase 2		AK069490	12.9-13.5		P0680A03	AB023482	0.0

BU672920	40S ribosomal protein S30-like protein	AY128361	AK060583	13.5		OSJNBa0014B15	AP002854	e-167
CB965760	membrane-associated salt-inducible protein	AL022198	AK103771	13.5	C226A	P0675A05	AP002071	0.0
BU673535	ubiquinol--cytochrome c reductase	X79275		13.5	R845	P0015E04	AP002069	e-150
BU673227	vacuolar H+-pyrophosphatase	D45384	AK066933	13.8	S1434	OSJNBa0035I03	AP003019	0.0
BI306444	hypothetical protein	AP001072	AK065263	13.8	S1434	OSJNBa0035I03	AP003019	e-166
CB967086	RSZp22 splicing factor		AK063879	15.8	S1520	P0470C02	AP003508	e-127
BU673818	hypothetical protein			18.0	R1966	P0492A09	AP004697	1e -12
BU672978	peroxiredoxin Q	AB037598	AK059845	19.1	R1954	P0492A09	AP004697	e-126
CB966161	DnaJ domain		AK111076	19.1	R1954	P0492A09	AP004697	e-145
BI305831	sucrose synthase	Z15028	AK098923	19.1	R1954	P0492A09	AP004697	3e-77
BI306555	chaperonin 10 gene	D29698	AK104352	19.1	R1954	OSJNBb0026P21	AP004682	0.0
BI306551	Beta-ketoacyl-ACP synthase I (KAS I)		AK060515	19.1	R1954	OSJNBb0026P21	AP004682	0.0
CB967441	unknown protein		AK058843	19.1-31.3		B1172G12	AP006056	0.0
BU673597	unkown protein	NM_100379	AK065860	19.1-31.3		P0664C05	AP004758	1e -59
CB965883	No hit			31.3	E3188	P0021C04	AP004687	2e-39
BU673322	Luminal binding protein 2 precursor (BiP2)		AK104048	31.3, 32.7	E3188, L688	P0701E03	AP003458	0.0
BU673610	phytocyanin		AK104260	33.5	E107S	P0644A02	AP005425	e-156
BI306583	RNA binding protein	NM_121073	AK060429	33.5	E107S	P0436F11	AP003488	e-118
BI306405	hypothetical protein		AK071776	38.3	C1478	P0479H10	AP005522	0.0
BI306573	glutathione dependent dehydroascorbate reductase precursor	AF301597	AK070471	38.3	C1478	P0479H10	AP005522	e-173
CB964418	No hit			40.2	P126	OSJNBa0068B06	AP004995	e-165
BU673730	unknown protein	NM_126499	AK107758	40.2	P126	OSJNBa0068B06	AP004995	e-109
BI305611	glucose-inhibited division protein B		AK071932	40.2	P126	OSJNBa0068B06	AP004995	e-158
BU672984	MutT/nudix family protein		AK064894	51.0	R2171	P0592E11	AP003490	1e -82
BI305515	RAD23	AC021640	AK061556	51.0-53.5		OJ1001_B06	AP005382	9e-85
CB967148	expressed protein		AK099532	51.0-53.5		OJ1001_B06	AP005382	e-134
BI305798	membrane protein YJR151c			51.0-53.5		OSJNBa0037N01	AP004993	0.0
BI306362	zinc finger protein	AC079281		51.0-53.5		OSJNBa0037N01	AP004993	0.0
CB965818	No hit			56.3	S8172	P0456E06	AP005822	2e-72
CB966864	No hit		AK072967	58.7	R10069S	OSJNBb0026L12	AP006053	e-105
BU673002	No hit		AK073716	61.6-62.6		OSJNBa0042E12	AP005748	0.0
BI306082	hypothetical protein		AK058338	64.6	S16072	B1156C07	AP005968	e-118
CB964968	60S ribosomal protein L31-1			64.6	C999A	P0592B08	AP003543	e-178
BU673657	No hit		AK062356	64.6	C999A	P0592B08	AP003543	6e -89
CB967431	chlorophyll a/b-binding protein precursor	AF094776	AK098864	64.9	R538	P0578B12	AP003511	0.0
BI306753	unknown protein			65.8	C62815S	P0531C01	AP003763	e-129
BI305506	Nad-dependent formate dehydrogenase	AB019533	AK104788	65.8	R111	P0008F02	AP003518	e-168
BI305690	delta-type tonoplast intrinsic protein	U86763	AK104464	65.8		P0427E01	AP005449	0.0
BI305630	No hit		AK058968	65.8	S2570	P0583E12	AP004728	0.0
BU673851	No hit			65.8		OSJNBb0026F02	AP004742	0.0
CB964641	ubiquitin		AK100224	65.8	C58	OJ1294_G12	AP003952	0.0
CB965393	conserved hypothetical protein			65.8	C991	P0021H10	AP003520	0.0
CB964477	methionyl-tRNA synthetase	AF040700		66.8	Y2686R	P0561B08	AP003618	8e-79
BI305692	LEAF-SPECIFIC THIONIN PRECURSOR	M19046	AK062831	67.1	S1809, C30378SA	OJ1536_A04	AP003959	0.0
CB966739	DEAD/DEAH box helicase			67.7-68.5		OSJNBa0043B22	AP005470	0.0
BI306576	ribosomal protein S10 mitochondrial	AB035348		70.9	R32	P0551A03	AP004818	e-140
CB967331	40S ribosomal protein S24		AK059383	74.1	R437	P0458E11	AP003613	e-173
BI306541	putative protein	NM_125047	AK103977	74.1-76.5		P0456F09	AP003762	7e-78
BU673301	cysteine synthase		AK109759	76.5	C214	P0656E03	AP003714	e-133
BI306514	3-hydroxyisobutyryl-coenzyme A hydrolase	NM_117410	AK072650	82.9	R674	P0652A05	AP004571	7e-81
BI305214	Photosystem I assembly protein ycf4			84.5	C12560S	P0417D05	AP004236	1390
BI306483	No hit		AK060002	85.4	S12715	P0486H12	AP003615	e-151
BI306245	unknown protein	NM_128223	AK072067	85.4	S12715	P0486H12	AP003615	0.0
CB966242	No hit			87.5	C30378SB	P0029C06	AP005446	6e-91
BI305951	unknown protein		AK066355	87.5	C30378SB	P0029C06	AP005446	0.0
BI305317	putative protein		AK070667	87.5-90.5		P0556B08	AP004279	0.0
BI305948	putative protein		AK072533	87.5-90.5		P0556B08	AP004279	0.0
BI306343	glycine-rich protein,		AK104605	90.5	R1559	P0417G12	AP003711	0.0
BI305443	bZIP transcription factor		AK107021	91.9	R2549	OJ1008_D02	AP003938	e-156
CB967067	amino acid transporter family	NP_566854	AK073428	96.5	C12361S	P0505A04	AP004792	0.0
BI306424	glycine-rich RNA-binding protein grp1a	NM_102006		99.2	L263	P0523F01	AP003573	0.0
CB966674	protein kinase family	NP_194952	AK068750	99.2	L263	P0458E02	AP003571	0.0
CB966995	expressed protein	NP_565372	AK073270	99.2	L263	P0523F01	AP003573	e-101
BU673870	HOS59	AB007628		100.3, 100.8	R959, C882	OJ1226_A12	AP004008	3e -97
CB964529	hypothetical protein		AK062869	100.3, 100.8	R959, C882	OJ1226_A12	AP004008	0.0
CB966502	No hit		AK065143	100.8	C882, S1994	OJ1226_A12	AP004008	0.0
BU673305	No hit		AK073865	100.8-103.0		P0453H04	AP005453	e-175
BI305923	translocation complex Sec61gamma chain	AY059131	AK059084	103.0	R276	P0709F06	AP003579	e-130
CB964714	redundant			105.1	S1306S	P0514G12	AP000616	e-131
BI305954	unknown protein		AK058614	105.1	S1306S	OSJNBa0051O02	AP005769	0.0
BI305621	expressed protein	NM_113124	AK073551	107.3	G1091	P0686E06	AP003635	0.0
BI306616	hypothetical protein		AK107710	107.3	C556	P0637D03	AP003633	0.0
CB965534	protein kinase C inhibitor (Zinc-binding protein)			107.3	G1091, C556	P0686E06	AP003635	3e -51
BI305640	lipid acyl hydrolase	AC091123		109.5	C11635S	P0710B08	AP003728	0.0
BU673627	No hit			110.6	E3288S	P0547F09	AP004797	0.0
BI306328	wound induced protein	X59882	AK103599	110.6-113.1		P0462F09	AP004277	0.0
CB966858	expressed protein		AK065448	113.1	C962	P0623A10	AP005395	e-104
CB965992	hypothetical protein			113.4	C1424, C2851	P0532H03	AP005457	e-134
CB965601	thaumatin-like protein	NP_173261	AK108191	113.4	C1424, C2851	P0532H03	AP005457	0.0

BI306368	RNA polymerase I, II and III 16.5 kDa subunit	AF017074	AK068912	115.6	R11	P0028E05	AP005445	e-150
BU673245	xyloglucan endotransglycosylase	X93175	AK104912	115.6	R11	P0028E05	AP005445	e-104
BI305610	xyloglucan endo-1,4-beta-D-glucanase (EC 3.2.1.-)	X93175	AK104912	115.6	R11	P0622F03	AP003771	0.0
CB964730	aminolevulinate dehydratase	X92402	AK101836	117.0-119.2		P0018H04	AP003761	0.0
BI305766	hypersensitivity-related gene	X95343		119.2, 119.5	R1782, R1028	P0655A07	AP003634	2e-59
BI305746	cyclophilin CYP5	AF020433	AK072490	119.5	R1028	P0655A07	AP003634	0.0
BU672832	hypothetical protein		AK069642	119.5, 120.1	R1028, E21149	OJ1663_H12	AP004329	e-104
BI306471	putative protein		AK100047	121.7	R2614	P0481E08	AP003614	e-179
CB965401	putative protein		AK071639	121.7, 122.8	R2614, E12186S	OJ1540_H01	AC091774	0.0
CB964592	heat shock protein 90	AB037681	AK102478	121.7, 122.8	R2614, E12186S	OJ1540_H01	AC091774	0.0
BI306443	catalase	D26484	AK069446	124.4	R1167	P0017G10	AP004685	0.0

AccNo	Putative	ACC NO		cM	Marker	chromosome 7	BAC ACC	E-VAL
BU673817	alpha tubulin	AJ420858	AK104900			P0408B10	AP004271	e-170
BI306430	hypothetical protein	AJ012688	AK062043			OJ1513_F02	AP005244	e-161
BU672976	disease resistance response protein	NM_123616	AK065027	0.8	E2609S	B1317D11	AP006186	0.0
BI306674	alanine aminotransferase-like protein	AY042902	AK067732	1.9	S14164	P0585H11	AP004342	e-159
CB965669	No hit		AK066728	2.2	E2475S	OJ1567_G09	AP003759	e-179
BU673463	pectinesterase		AK102922	4.4	E3124S	P0022B05	AP004262	e-177
CB965839	cytoskeletal protein		AK070421	4.4	E3124S	P0022B05	AP004262	e-162
CB967125	pathogenesis-related protein 1	AF306651		6.4-7.0		OSJNBa0088O14	AP006172	0.0
CB966756	LRK1 protein	AF193835		7.0	E20128	OJ1118_D07	AP003742	0.0
BI305984	unknown protein; tRNA-Val; tRNA-Ala	AC084320	AK099414	9.6	S20268	OSJNBa0087K02	AL607095	3e-99
BI305735	23 kDa polypeptide of photosystem II	AF052203	AK104722	9.6	S20268	OJ1351_C05	AP004010	0.0
CB965482	hypothetical protein		AK062172	11.0	R1561	OJ1470_H06	AP003957	4e-92
BI305397	photosystem I PSI-K subunit	L12707	AK058788	11.0	R1561	P0022E03	AP004263	0.0
CB965311	transcription factor-like protein	BAC84333		24.2	R565	P0455F03	AP005454	1e-97
CB966765	hypothetical protein		AK060475	24.8	S12540	OJ1714_H10	AP003847	3e-72
CB965406	Sm protein		AK109573	25.4, 26.0	S11633, S10012	P0039H02	AP004267	e-141
BU673578	zinc finger-like protein	AY086347		26.0-31.0		OSJNBa0050F10	AP005840	0.0
CB964668	root hairless 1		AK103352	26.0-31.0		P0496D04	AP004670	e-126
BI305680	RNA polymerase II 13.6 kDa subunit	U28048	AK099622	26.0-31.0		P0496D04	AP004670	e-113
CB965351	casein kinase-like protein		AK100663	26.0-31.0		P0496D04	AP004670	e-146
BU673115	hypothetical protein			31.0	E50408S	P0534H07	AP004307	0.0
BU673495	histone H1-like protein	AF461814	AK106132	31.0-35.7		OJ1046_F10	AP003861	0.0
BI306248	thioredoxin h	D26547	AK059196	35.7	R2401	P0506C07	AP004384	e-146
BU672864	expressed protein	AY136359	AK066157	35.7	R2401	OJ1046_F10	AP003861	e-161
BU673869	expressed protein	NM_115079		38.4	S4774	OSJNBb0084L07	AP005179	9e -60
CB966748	No hit		AK104473	38.4-41.7		OJ1715_A07	AP003848	8e -35
BI305675	rpS28	AJ001161	AK103440	38.4-41.7		P0440B02	AP004313	e-122
CB966544	No hit		AK070547	38.4-41.7		OJ1118_E12	AP003743	0.0
BU673493	ribosomal protein S15	L27461		42.6	R1807	OJ1559_F09	AP003837	0.0
BU673561	tensile reaction protein (WAR1)	AY072932		42.6-43.8		OSJNBa0008C11	AP005098	e-138
CB966886	ribosomal protein L7Ae		AK071026	45.5	G2009	OSJNBa0016A21	AP005764	2e-98
CB966421	No hit		AK109207	45.5	G2009	OJ1203_D08	AP003819	e-139
BI305652	hypothetical protein		AK067680	46.5	C636	P0523A04	AP004340	0.0
CB966982	glutamate dehydrogenase	D49475	AK101492	46.5	C636	P0565A07	AP005194	2e-61
BI306179	No hit		AK101492	46.5	C636	OJ1008_F04	AP003939	e-132
BI306302	RNA-binding protein	NM_128511		49.7	S10674B	P0557D09	AP005260	0.0
BU673022	B12D protein	BAC84633		49.7	E1160S	OSJNBb0032G22	AP005880	e-123
CB966731	unknown protein			49.7	C12887S	P0613B07	AP005462	0.0
CB964525	methionyl aminopeptidase-like protein	NP_172785	AK068716	50.0	R658	P0673E01	AP005200	0.0
BI305816	photosystem I chain IV precursor	Y00966		50.0	R658	OJ1483_E04	AP003834	0.0
CB967266	Pridicted protein			50.9	C1464	B1157F01	AP006159	e-167
BI305644	60S ribosomal protein L144	AF398144	AK063789	53.4	C735	OJ1047_A06	AP003802	1e-91
CB965197	water channel protein		AK103970	53.4	C735	OJ1047_A06	AP003802	e-154
BU672781	Expressed protein		AK101381	53.4	C1226	OSJNBb0062P14	AP005737	e-121
CB964537	putative protein		AK067630	55.6	E3307S	OJ1634_H04	AP003916	0.0
BI306538	1-aminocyclopropane-1-carboxylate oxidase	X85747	AK059964	55.6	C60626SB	P0038F09	AP005256	e-174
BU672985	CDSP32 protein	Y09987	AK066045	57.5-60.8		P0640E12	AP005261	0.0
CB966105	anthranilate phosphoribosyltransferase-like protein	CAB67616	AK067097	60.8	S1723	OJ1136_F08	AP005465	0.0
BI305928	N64969		AK058326	60.8	G20	P0434A03	AP004299	0.0
BI306454	No hit			61.6-61.9		OSJNBa0064M11	AP005515	3e-23
BU673410	p53 binding protein	AY072931	AK101803	61.9	S20608S	P0477A12	AP005190	e-158
BI306363	nucleoside diphosphate kinase	D16292		61.9	S954	P0038F10	AP004266	7e-64
BU673601	hypothetical protein	NM_117809	AK073883	61.9-62.4		OJ2096_F11	AP003964	7e -43
BU673777	protein kinase CK2 regulatory subunit CK2B3	AF239818		61.9-62.4		OJ2096_F11	AP003964	2e -46
BU673631	small GTP-binding protein			62.4	S11250	OJ1457_D07	AP003956	0.0
BI305327	hypothetical protein	T04607	AK067791	62.4-67.0		OJ1112_F06	AP003995	1e-67
BU673338	symbiosis-related protein		AK059939	67.0-69.2		P0580A11	AP005196	e-147
CB964588	expressed protein		AK099660	71.6	C451	P0678G09	AP004573	7e-79
BU673663	Pridicted protein		AK060014	73.2	C1467	P0681F05	AP004674	0.0
BI305563	translation initiation factor (GOS2)	AF094774	AK105037	73.2	C1467	P0681F05	AP004674	e-151
BI305420	similarity to beta-1,3-glucanase-like protein	AB008265		74.0	R2286	P0711B07	AP004575	e-135
CB967282	No hit		AK099127	75.6	C1008	OJ1582_D10	AP003838	8e-31
BI305386	protein H2A	D38091	AK059228	75.6	C1008	OJ1582_D10	AP003838	0.0
BU672861	photosystem II oxygen-evolving complex protein 3 precursor	M87435	AK099127	75.6	C1008	OJ1582_D10	AP003838	e-157
BI305757	beta-amylase gene	L10346		75.6	R1477A	OJ1729_E01	AP005156	

BI306632	ribosomal protein	U86017	AK058262	76.2	C1237	P0554D11	AP004569	e-114
BU673542	No hit		AK071294	78.9	R1477B	OSJNBa0058I18	AP005125	e-153
BI306021	chlorophyll a/b-binding protein	AF058796	AK103924	78.9	R1357	OSJNBa0006F16	AC087839	e-122
BU673850	putative protein	NM_124804		78.9	R1357	OSJNBa0006F16	AC087839	0.0
BU673083	fiber protein		AK106307	78.9-80.5		OSJNBb0088N21	AC093094	0.0
BU673641	Pridicted protein			78.9-80.5	R1357	P0567H04	AP005195	0.0
BI305348	calcium-dependent protein kinase	AL133248	AK068315	80.5, 80.8	R2394, S1563	OJ1092_A07	AP003866	0.0
BI305275	root border cell-specific protein	AF139187	AK104308	81.4	S10051	OJ1699_E05	AP003845	
BI305892	alpha 1 tubulin	Z11931	AK104900	81.4	S10051	P0408B10	AP004271	e-116
CB964708	photosystem I antenna protein	AF010321	AK104651	81.4	E1186S	P0408B10	AP004271	1e-77
CB966592	Yippee-like protein	BAC82961	AK073870	81.9	E61009S	OJ1127_E01	AP003747	e-151
BU672856	unknown protein	NM_102495		81.9	E61009S	OJ1301_C12	AP004185	8e -97
BU672829	unknown protein	AY050865	AK059561	81.9	E61009S	OJ1127_E01	AP003747	0.0
BI306557	hypothetical protein	AP003223		81.9		P0453G03	AP004276	0.0
CB967154	hypothetical protein		AK062359	83.3	R2677	OJ1174_G05	AP003750	2e-86
BI305925	peroxisome assembly protein PER8		AK059637	89.8-91.7		OSJNBb0024A20	AC079038	e-128
BI305765	60S ribosomal protein L27a	AB042856		89.8-91.7		OJ1154_D08	AP003814	0.0
BI305944	small nuclear ribonucleoprotein G	AC079038		89.8-91.7		OSJNBb0024A20	AC079038	2e-58
BU672891	expressed protein	NM_106153	AK068566	91.7	C60318SB	OJ1505_A06	AP004189	e-112
BI306545	elongation factor 1 beta,	D23674	AK071736	91.7	C60318SB	P0616D06	AP005198	5e-54
CB967248	glycolate oxidase	NP_188031	AK071738	93.9	C50076S	B1056G08	AP004988	0.0
BI306295	cyclin D2.1 protein	AJ011892	AK063671	94.7	E10204S	P0594D10	AP004380	e-110
BU673638	calmodulin-like protein	AC079853	AK100302	94.7	C1579S	P0560B08	AP004309	0.0
BU673904	putative protein		AK067310	97.4, 98.2	R10215S, E20959S	P0519E12	AP004339	e-138
CB966831	Hydrophobic protein RCI2B (Low temperature and salt responsive protein LTI6B)			99.3	C60933	P0487A05	AP004383	e-143
BI305988	No hit			99.3	C60933	OJ1136_D11	AP003749	1e-30
BI306256	hypothetical protein			99.3, 99.6	R2658S, S14165	P0524G08	AP004671	e-144
BI306274	Expressed protein		AK058954	102.3	C1340	P0503D09	AP005455	e-129
BI305189	hypothetical protein	AC005662		105.7	C596	OJ1477_F01	AP003833	
CB964965	ERD15 protein		AK107202	105.7-114.5	C596	OJ1357_E01	AP004186	0.0
CB967207	alcohol dehydrogenase-like protein	BAC22440	AK107157	105.7-114.5		P0453E03	AP005452	2e-95
CB965631	CONSTANS family zinc finger protein	AC069471		105.7-114.5		P0450A04	AP004274	e-111
BI306155	putative protein	NM_120719	AK103678	114.5	C53986	P0470D12	AP004300	e-116
BU673224	No hit		AK105813	114.5	C53986	P0470D12	AP004300	8e-51
BI306408	alpha-galactosidase-like protein	NM_115489	AK099229	116.1	E60814S	OJ1205_F02	AP005243	0.0
CB967299	osr40g3	Y08988	AK072989	116.6	S11279	OSJNBa0060O17	AP005167	0.0
BU673523	hypothetical protein		AK072158	118.3	R411	P0034A04	AP004333	e-108
CB967424	Tat binding protein	D17789	AK099264	118.3	R411	P0034A04	AP004333	e-161
AccNo	**Putative**	**ACC NO**		**cM**	**Marker**	**chromosome 8**	**BAC ACC**	**E-VAL**
BI306358	ferredoxin	AF010320	AK061654	1.6	C50915S	OJ1300_E01	AP003909	0.0
BU672806	P450	AB042627	AK067200	1.6, 1.9	C50915S, R1943	P0450B04	AP004462	0.0
BI305188	cytochrome P450	X81828	AK067200	1.6, 1.9	C50915S, R1943	OJ1300_E01	AP003909	e-103
CB966041	acidic ribosomal protein P2	D29689	AK058240	3.6-12.8		P0470F10	AP004562	e-153
CB965970	avr9			3.6-12.8		P0470F10	AP004562	1e-47
CB965010	NAM (no apical meristem)-		AK072275	3.6-12.8		P0427G12	AP005657	e-103
BU673826	chloroplast RNA-binding protein cp33	NM_099990	AK099630	3.6-12.8		P0470F10	AP004562	e-170
CB966454	mitochondrial ribosomal protein S14		AK059789	12.8	R3003	OJ1005_B05	AP003925	e-141
BI305233	fructose-bisphosphate aldolase	AJ133146	AK104952	12.8	R3003, G278	P0023G04	AP004374	0.0
BI305448	manganese-binding protein PsbY precursor, photosystem II-associated	AF060198	AK068273	12.8	R3003	OJ1005_B05	AP003925	e-121
BI305464	ribosomal portein S14 mitochondrial	AB017429	AK099077	12.8	R3003	OJ1005_B05	AP003925	0.0
BU673624	glutamine phosphoribosylpyrophosphate amidotransferase	L23834	AK101577	13.4, 13.9	C390, C770	P0582D05	AP004591	e-111
BI305529	60S RIBOSOMAL PROTEIN L37	X79074	AK058280	13.9	C770	P0582D05	AP004591	e-109
CB964457	expressed protein	NP_180787		16.7	C905	OJ1119_D01	AP003876	e-151
BI306681	tryptophan decarboxylase DOPA	X67662	AK069031	20.2	R2367	OJ1368_G08	AP003911	0.0
BI306601	monodehydroascorbate reductase	AF109695	AK099848	21.6-25.2		P0443G08	AP004461	0.0
CB966660	Tyrosyl-tRNA synthetase	CAA71881	AK101530	21.6-25.2		P0443G08	AP004461	5e-71
CB966416	60S RIBOSOMAL PROTEIN L34	S60476	AK103424	27.1	R2723	P0438H08	AP004460	e-135
BU673855	caffeic acid 3-O-Methyltransferase	AJ231133	AK064768	27.1	R2723	P0438H08	AP004460	3e -48
CB966186	expressed protein		AK103973	28.4	S13883A	P0498E12	AP004698	0.0
BI305215	plastid RNA polymerase sigma factor	AB005290	AK071725	31.6	E60549S	OJ9990_A01	AP005847	0.0
BI306682	snRNP splicing factor	NM_126432		35.7, 36.0	E1133S,R2976	P0583B06	AP004619	e-168
CB964739	SET-domain transcriptional regulator family	NP_193253	AK058948	36.8	R902	P0455A11	AP004692	0.0
BI306688	No hit		AK059070	38.8	C1121	P0610E02	AP005505	e-163
BI306089	aconitase-iron regulated protein 1 (IRP1)	AF073507		38.8	C1121	P0047G03	AP004660	0.0
BU673124	photosystem II 10 kDa polypeptide	U86018	AK105055	42.9	S12665S	P0556A11	AP004589	e-158
CB966839	hypothetical protein		AK062655	44.6-45.4		P0486F07	AP004565	1e-65
BI305657	ribosomal protein S17		AK058705	44.6-45.4		OJ1734_E04	AP003919	5e-40
BI306663	F12A21.16	AL442114		48.8-49.2		P0682A06	AP004705	0.0
CB965858	plasma membrane H+-ATPase	AJ440216	AK108449	49.5, 50.5	S779, C1107	P0031C02	AP004657	1e-82
CB967300	expressed protein		AK068554	50.8	S10715S	OSJNBb0070J06	AP005478	0.0
BI306634	NADPH-cytochrome P450 oxydoreductase isoform 3	AF302498	AK101320	50.8	S10715S, R80	P0437G01	AP004690	e-127
BI306050	No hit		AK101898	50.8	S10715S, R80	P0437G01	AP004690	0.0
BU673889	chloroplast apocytochrome b6 (petB)	M35995		50.8-52.1		OSJNBa0036E18	AP005161	e-129
BU673042	hypothetical protein		AK062738	52.9	E31119S	OJ1540_G08	AP004650	5e -61
CB966395	hypothetical protein		AK101221	52.9-53.7		OSJNBa0091D16	AP005731	e-116
BI306307	hypothetical protein		AK072208	54.0-54.3		OJ1705_A03	AP003918	e-101
BI306385	XA1			54.3	C1374	OJ1127_B04	AP005484	e-118

CB966393	1,2-diacylglycerol 3-beta-galactosyltransferase	NP_588394	AK064148	54.3	R2381	OSJNBa0063H21	AP005819	0.0
BU673532	unknown protein	AB016886	AK073839	54.3	C1115, R2466	P0670E08	AP004759	e-129
BU672912	poly(A)-binding protein	U81318	AK104792	54.3		P0451H06	AP005500	0.0
BU672876	No hit			54.3-55.4		P0474F07	AP004695	1e -90
BU673639	wound inducive gene	AB009885	AK105899	58.0-58.4		P0426E02	AP005520	1e -83
BI305878	predicted protein		AK070805	58.4-58.7		B1027A11	AP005464	e-125
BU672895	actin related protein 2	AF095912		59.0	S15506S	OSJNBa0091C18	AP005389	2e -21
BI306627	RUB1 conjugating enzyme	AF202771	AK104511	60.1	E2623S	P0508B09	AP004631	0.0
CB967336	expressed protein		AK065829	60.4-61.2		B1090H08	AP005795	e-146
BI306182	RBP2 protein		AK065829	60.4-61.2		B1090H08	AP005795	6e-95
CB965960	unknown protein		AK099893	61.2-66.5		P0671F11	AP004634	0.0
CB966757	putative protein		AK060114	66.5	R1394A	OJ1198_B10	AP003947	e-156
BI306372	No hit		AK068312	67.3-70.1		P0453D01	AP004691	e-148
BI305986	No hit			70.1-72.2		P0433E10	AP004667	0.0
BI305366	unknown protein	AB005242	AK061339	72.2	E3802	OJ1136_E11	AP003885	e-130
BI306242	hypothetical protein			75.7	S815	P0479C08	AP004617	0.0
BI306480	high mobility group I/Y-2	AF291748	AK068506	75.7	S815	P0479C08	AP004617	e-119
BI306060	GF14-c protein	U65957	AK070230	76.7	E50105S	OJ1124_B05	AP003881	5e -57
CB965968	DNA repair protein RAD23 homolog		AK063768	76.7	E50105S	OJ1124_B05	AP003881	e-171
CB964925	unknown	AAM66974	AK099471	76.7-78.5		P0431A03	AP004666	e-107
BU673766	ribonuclease	AB052842	AK059802	78.5	S10631	P0413H11	AP005757	e-131
BI305356	non phosphorylating glyceraldehyde-3-phosphate dehydrogenase	X75326	AK062559	80.4, 80.7	S10622, S11114	P0528B09	AP004703	4e-81
BU672938	Phosphoglycerate dehydrogenase-like protein	AY086001		82.8	E60162SB	P0429B05	AP004665	e-101
BI306438	No hit		AK101935	85.1	E61231	OJ1117_F10	AP003871	e-103
BU673217	conserved hypothetical protein		AK067942	86.7	E2448S	P0048G02	AP004662	e-158
BU673625	hypothetical protein		AK101386	86.7	L677	P0493A04	AP004586	0.0
CB966763	germin-like protein 5	AF032975	AK104729	86.7-88.6	L677	P0690E03	AP004707	0.0
CB966259	No hit		AK098966	86.7-88.6		B1111C03	AP005405	e-115
BI305495	hypothetical protein		AK060706	90.5-92.2	R2382	P0451G12	AP004399	0.0
CB964648	hypothetical protein		AK062882	90.5-92.2		P0451G12	AP004399	3e-48
BI306299	No hit		AK062882	90.5-92.2		P0451G12	AP004399	5e-47
BI305666	mitochondrial F0 ATP synthase D chain	AJ271469	AK068050	92.2	S4036S	OJ1666_A04	AP003917	0.0
BU673354	14-3-3 protein homolog GF14-12	M96856	AK101599	92.2-96.6		OSJNBb0092C08	AP005391	2e -88
BU673035	No hit			92.2-96.6		OJ1381_H02	AP004164	5e -45
CB966745	phytocyanin		AK105053	92.2-96.6		OJ1111_H02	AP004213	e-176
BU673704	homeodomain leucine zipper protein	AF145728	AK106818	92.2-96.6		OSJNBb0092C08	AP005391	e-140
BI305664	No hit			92.2-96.6		OJ1381_H02	AP004164	0.0
BI305526	signal recognition particle receptor-like protein		AK101591	92.2-96.6		OJ1113_A10	AP004643	e-104
CB965175	No hit		AK064774	99.1-102.1		OJ1506_F01	AP004190	e-156
BI306522	aminotransferase 1	AY066012	AK099206	99.1-102.1		OJ1506_F01	AP004190	0.0
BI306320	4-coumarate-CoA ligas	X52623	AK103001	99.1-102.1		OJ1506_F01	AP004190	e-156
CB967426	unknown protein		AK069190	102.1	S6487S	OJ1134_H03	AP003883	e-153
CB964474	Pridicted protein		AK101254	103.2	S5064S	OSJNBa0016N23	AP006049	e-127
CB966872	beta-glucosidase-like protein		AK098938	103.2	S5064S	B1168A08	AP005816	3e -62
BI305387	expressed protein		AK107181 ASP	104.8, 105.7	S10655, C10122S	P0686C03	AP004761	0.0
BI305862	peptidylprolyl cis-trans isomerase	X86903	AK103172	107.7	C502	OJ1191_A10	AP003888	e-110
CB966976	No hit		AK104251	109.3	S14003	P0702E04	AP005529	8e-33
BI306507	coated vesicle membrane protein	AC009176	AK105383	109.3	S14003	P0702E04	AP005529	e-127
CB967195	unknown protein		AK061214	109.3	S14003	P0702E04	AP005529	0.0
BU673450	nuclear transport factor 2	AB011262	AK065141	109.3	S14003	P0702E04	AP005529	0.0
BU673516	NAM-like protein	AY086093	AK104766	111.2	R2027	OSJNBa0033D24	AP005439	e-127
CB967419	NAM (no apical meristem) - like protein		AK104766	111.2	R2027	OSJNBa0033D24	AP005439	e-117
CB964609	biotin synthase	AAO41898	AK059296	112.6	R2662, C12955	OJ1211_G06	AP003948	0.0
BI306103	cysteine endopeptidase precursor	AF099203	AK105612	119.9	S1088	OJ1150_A11	AP003928	e-124
CB964980	fiber protein Fb14	AAP34361		119.9	S1088	OJ1150_A11	AP003928	e-109
BU673012	histidine-containing phosphotransfer protein	AB024293	AK061111	119.9	S1088	OJ1150_A11	AP003928	3e -65
BI305523	nucleic acid-binding protein	AJ224324	AK070743	119.9	S1088	OJ1150_A11	AP003928	2e-59
BI305679	No hit		AK060208	119.9, 120.4	C50778S	P0604E01	AP005544	4e-98
BU673582	No hit		AK099629	119.9,120.4	C50778S	P0604E01	AP005544	2e -85
CB965300	NADP-dependant malate dehydrogenase	AJ512373	AK105935	120.4	C50580S, R10036S	P0543D10	AP004587	3e-85
CB964766	type 1 membrane protein		AK065186	121.2	S1433	OSJNBa0044E16	AP005411	e-159
BU673384	expressed protein	NM_126585	AK066210	121.2	S1433	OSJNBa0044E16	AP005411	0.0

AccNo	Putative	ACC NO		cM	Marker	chromosome 9	BAC ACC	E-VAL
CB967293	No hit					OJ1399_D07	AP004138	7e-29
BI306612	NADH dehydrogenase	AJ295997				P0657B04	AP005712	0.0
BI306593	pyrophosphate-fructose 6-phosphate 1-phosphotransferase alpha-subunit	M55190	AK099939			0091I24	AC091687	
BI305865	peroxidase BP 1	M73234	AK066336			OJ1399_D07	AP004138	1e-53
BU673698	No hit		AK100974	0.8		OJ1134_E08	AP005860	e-162
CB964734	protein phosphatase 2A regulatory A subunit	AJ243828	AK069757	0.8	S1535, C1232	OJ1695_A02	AP005578	2e-75
BI305794	ribosomal protein L17-1	AF264022	AK072488	0.8		OSJNBb0044M12	AP005736	e-108
BI306147	Bci-5 protein	AJ250661	AK062461	3.2		PAC0651G05	AC090055	0.0
CB965456	zinc finger-like protein	BAB02496	AK071104	4.4	R1164	PAC0645D04	AC090054	1e-37
BU673335	No hit			4.4-6.8	R1164	P0603H10	AP006527	0.0
CB966486	E2, ubiquitin-conjugating enzyme		AK104965	10.0	P0592C05f	P0592C05	AP004756	e-158

CB966177	ADP-glucose pyrophosphorylase small subunit	AY028315	AK073146	10.0	P0592C05f	P0592C05	AP004756	7e-76
BU673417	ADP-glucose pyrophosphorylase small subunit	AY028315	AK073146	10.0	P0592C05f	P0592C05	AP004756	e-159
BU673167	No hit			10.0-20.7		OSJNBb0079K11	AP005971	1e-77
BI305990	peroxidase	X98322		26.7-30.6		OSJNBb0024J13	AP005879	0.0
BI305490	apospory-associated protein C-like	AB013389	AK068681	31.3	C3096SB	P0706E03	AP005811	e-138
CB967321	hypothetical protein			33.0	E61041	OJ1253_E02	AP005566	1e-38
CB967467	hypothetical protein		AK073913	34.4	E60109SC	P0448B03	AP005704	3e-93
CB966928	histone acetyltransferase	T02064	AK101995	35.2	C1176	P0698G06	AP005400	0.0
BU672792	DRE binding factor 1	AF493800	AK058899	40.1	C397	OJ1759_F09	AP005580	0.0
CB966335	hypothetical protein			40.7, 41.0	C873, R41	OSJNBa0041C07	AP005838	e-103
CB965477	acetyl-CoA synthetase			41.9	S752	P0027G10	AP005702	0.0
BI306619	No hit		AK105382	45.2-49.3		OJ1655_B12	AP005577	1e -95
BI306528	positive element factor 1 (PF1)	L24390	AK068506	45.2-49.3		OSJNBb0014M19	AP005732	0.0
CB966275	unknown protein			50.7	C472	P0643D11	AP005397	0.0
BI306104	hypothetical protein		AK062405	50.7	C472	P0643D11	AP005397	0.0
BU673199	CCR4-associated factor	NM_128767	AK065377	50.7-55.3		P0014G10	AP005784	0.0
CB966932	copper chaperone COX17-1	AF349684	AK111264	50.7-55.3		P0668D04	AP005426	0.0
CB966319	lipid transfer protein; glossy1 homolog		AK060786	55.3-58.3		OJ1299_A11	AP005568	e-124
CB964557	expressed protein		AK099618	58.3-60.8		OJ1328_D07	AP005833	8e-61
BI305417	cytochrome P450 monooxygenase CYP92A1	AY072297	AK100321	58.3-60.8		OJ1344_B01	AP005570	0.0
BI306214	chaperonin 21 precursor	AF233745	AK098903	58.3-60.8		P0556H01	AP005424	0.0
CB964876	10-formyltetrahydrofolate synthetase		AK065164	60.8-62.4		OJ1596_C06	AP005575	0.0
BI306101	hypothetical protein		AK102328	63.0-65.1		OJ1057_F03	AP005548	e-157
BI305609	carbonate dehydratase		AK069530	63.0-65.1		OSJNBb0019B14	AP005755	0.0
BU672925	calmodulin-like protein	AC079853		63.0-65.1		P0011A05	AP005418	0.0
BU672828	cyclin-dependent kinase inhibitor protein		AK063208	63.0-65.1		B1045B05	AP005891	e-141
BU673324	queuine tRNA-ribosyltransferase		AK064033	65.1	R2638	OJ1595_D08	AP005574	0.0
CB965114	heat shock protein		AK073817	65.1-68.2		OSJNBa0026C08	AP006169	e-131
BI306721	blight-associated protein p12 precursor	AF015782	AK063132	65.1-68.2		OJ1509_C06	AP005573	e-154
CB964829	heat-shock protein	CAA82945	AK073817	65.1-68.2		OSJNBa0026C08	AP006169	2e-60
CB965347	hypothetical protein		AK107853	68.2	C846	P0463G11	AP005633	e-178
CB964497	TMV-MP30 binding protein 2C	AF326729	AK073131	68.2	C846	P0556A05	AP005759	e-120
BI306657	heat shock protein 82	Z15018	AK064780	68.2		OSJNBa0087J09	AC108761	e-170
BI306090	cold acclimation protein WCOR410b	U73210	AK073885	69.5	C670SA	OSJNBa0046G16	AC108756	8e-27
CB965815	sodium-dicarboxylate cotransporter		AK105756	69.5	C670SA	OSJNBa0046G16	AC108756	e-153
BI305963	nucleoid DNA-binding protein cnd41	D26015	AK105666	69.5	C670SA	OSJNBa0087J09	AC108761	0.0
BI305633	unknown protein	AC004521	AK063923	69.5	C670SA	OSJNBa0087J09	AC108761	0.0
BI305895	60S ribosomal protein	AF140494	AK099282	74.7	C2194	OSJNBb0004A05	AC108763	0.0
CB965569	No hit		AK102600	74.7	C2194	OSJNBb0004A05	AC108763	1e-20
BU673337	URK1_MOUSE URIDINE KINASE		AK102065	75	C662	OSJNBa0010B06	AC108753	0.0
BI306009	pyruvate dehydrogenase		AK061863	77.2	E31511S	OSJNBa0094K23	AC138007	e-178
CB966751	expressed protein		AK106007	77.2	E31511S	OSJNBa0094K23	AC138007	0.0
CB966254	beta-glucosidase		AK101420	77.7	S2655	OSJNBa0065A15	AC137594	1e-70
BU673193	protein kinase	AL356014		77.7-78.0		OSJNBa0074B03	AC108760	0.0
CB965073	glyoxalase II	AY054407	AK103563	78.0	C11861S	OSJNBb0034B12	AP005735	0.0
BI305711	expressed protein	NM_101999	AK067677	78.0	C11861S	OSJNBb0034B12	AP005735	e-128
CB965248	BAG domain containing protein	NP_200019	AK060485	78.8	C442	OSJNBa0047P18	AP005864	e-138
CB966520	RNA-binding protein		AK073499	79.1	R3330	OJ1439_F07	AP005681	8e-74
BI305980	unknown protein			79.1-82.1		OJ1531_B07	AP005682	2e-80
BI306518	timing of CAB expression 1-like protein	AF272040	AK067728	79.1-82.1		OJ1254_E07	AP005567	e-144
BI306043	RNase S-like protein ,	AY056038	AK104656	82.4, 83.2	S10578, S955	P0569E11	AP006067	0.0
BI306399	conserved hypothetical protein		AK067622	90.1	E60222S	OSJNBa0038K02	AP005862	e-147
CB965528	No hit			90.7	C632	OJ1228_C03	AP005564	2e-57
BU672900	histone H4	M12277	AK058741	90.7	C632	OJ1228_C03	AP005564	0.0
BU672975	dihydroneopterin aldolase		AK071381	91.8,	S6120S, S10879	P0635G10	AP005396	1e -74
BU672941	expressed protein	NM_128327	AK068800	93.1	C60055S	OJ1155_H10	AP005558	1e -31
AccNo	**Putative**	**ACC NO**		**cM**	**Marker**	**chromosome 10**	**BAC ACC**	**C E-VAL**
BI306301	No hit		AK058403			nbxb0049A03	AC027661	
BI305501	polygalacturonase isoenzyme 1 beta subunit	AF251069				nbxb0049A03	AC027661	0.0
BI305509	expressed protein		AK102067			nbxb0018F16	AC025905	e-103
BU673365	ORF; able to induce HR-like lesions	U66269	AK059279			nbxb0094K20	AC025907	
BI305953	No hit					OSJNBa0027L23	AC018929	5e-28
BI306066	minor allergen	Z99708				OSJNBa0041F04	AC026759	0.0
BI306716	valyl-tRNA synthetase	NM_101328	AK066973			OSJNBb0022I16	AC091238	1e-80
BU673019	UDP-glucosyltransferase	AF515727				OSJNBb0095E09	AC125385	0.0
BI306643	32 kDa protein jakalin homolog	AF021256	AK111282	3.9	R2309	OSJNBb0004A06	AC099734	0.0
BI305372	rRNA large subunit	Z11889	AK106482	5.5-6.8		OSJNBa0030B02	AC074105	0.0
BI305772	CYP18 gene	Y08273	AK058898	6.8	S21067S	OSJNBa0004E08	AC091724	0.0
BU673025	proline-rich protein	AB055842	AK104622	6.8	S2083	OSJNBa0031A07	AC084884	e-169
BI305444	proline-rich protein	AC091665	AK059665	6.8	C404	OSJNBb0016M10	AC091665	0.0
BI305891	Hypothetical protein	AC078891	AK068977	10.9	C913A	OSJNBb0081F12	AC090488	5e-71
BU673310	ribosomal protein S10p/S20e	AC074355	AK062522	11.7	E1064S	OSJNBa0071I20	AC074355	0.0
BI306221	RING zinc finger protein	AC015985		15.2	G1125	OSJNBa0004P12	AC099040	0.0
BU673678	antifreeze glycoprotein precursor			16.4-16.8		OSJNBa0073L20	AC099774	0.0
BI305181	ribulose bisphosphate carboxylase large subunit		AK107087	19.0-21.8		OSJNAb0023M11	AC112514	e-131
BI306542	Orf122	AF287482	AK059639	19.0-21.8		OSJNBb0075K12	AC092750	e-107
CB965036	atpH pseudogene	D13098		19.0-21.8		OSJNBa0034L04	AC091680	e-172
BI305844	hypothetical protein	AP003106	AK105766	19.0-21.8		OSJNBb0075K12	AC092750	0.0
CB966372	acetyl-CoA carboxylase		AK058817	19.0-21.8		OSJNBa0073L01	AC092548	e-131

BI306259	Photosystem I assembly protein ycf4		AK062797	19.0-21.8		OSJNBb0075K12	AC092750	e-157
BU673483	ovule/fiber cell elongation protein Ghfe1	AF513859		24.7-26.1		OSJNBb0058B20	AC108884	0.0
BI306410	epimerase/dehydratase	NM_122767	AK104822	31.8	S10886A	OSJNBa0061K21	AC016780	0.0
BI306024	class III chitinase	AF296279	AK104292	31.8	S10886A	OSJNBa0061K21	AC016780	0.0
BI306513	mitochondrial chaperonin-60	Z12115	AK069617	44.3	C234	OSJNBa0071K18	AC027038	e-125
BI305882	RNA binding protein	AC026815	AK059172	44.6-45.7		OSJNBa0079L16	AC026815	e-163
BI305608	No hit			44.6-45.7		OSJNBa0079L16	AC026815	0.0
BI306035	No hit		AK062758	48.4	S21707S	OSJNBa0055P24	AC037425	e-160
CB967241	No hit			48.4	S21707S	AE017105	AE017105	e-117
CB965792	malate dehydrogenase	AF353203	AK104691	48.4	S21707S	OSJNBa0055P24	AC037425	e-155
CB964576	transcription factor BTF3	AY224531		48.8-51.5		OSJNBa0012L23	AC051632	0.0
CB965012	Pridicted protein		AK070826	51.5	S11148	OSJNBa0029C15	AC087182	0.0
CB966297	cytochrome P450-like protein		AK109052	51.5	S11148	OSJNBa0029C15	AC087182	e-145
BI306141	membrane-associated protein	AC068923	AK066573	53.3-53.6		OSJNBa0041P03	AC068950	0.0
BI306020	prolyl 4-hydroxylase, alpha subunit	AC068923		53.9	S20832S	OSJNBa0017E08	AC068923	e-140
BU673351	signal recognition particle	U85037		54.3	E31548S	OSJNBa0078O01	AC079888	e-114
BI305535	No hit			54.3	E31548S	OSJNBa0078O01	AC079888	e-105
CB967325	unknown protein		AK061011	54.3	E31548S	OSJNBa0078O01	AC079888	e-128
CB965262	malate oxidoreductase	NP_191966	AK101024	55.3	S21174S	OSJNBa0062C05	AC084023	e-118
BI306032	vesicle-associated membrane protein		AK058656	55.6	G4003	OSJNBb0015K05	AC090870	0.0
BI306176	glycine decarboxylase subunit	AF022731	AK062851	58.9	C488	OSJNBa0076F20	AC025296	5e-84
CB967478	No hit		AK108675	58.9	C488	OSJNBa0076F20	AC025296	e-140
BU672858	MITOGEN-ACTIVATED PROTEIN KINASE HOMOLOG MMK2	X82268		61.7-68.6		OSJNBa0053C23	AC092389	e-119
BU673471	receptor-like protein kinase		AK070172	61.7-68.6		OSJNBa0053C23	AC092389	
CB965191	ubiquitin-conjugating-like enzyme Ahus5	CAD29823	AK105227	61.7-68.6		OSJNBb0060I05	AC092697	e-173
CB965602	hexose transporter	AJ534445	AK102640	68.6-71.4		OSJNBb0064P21	AC073166	0.0
BU672878	unknown protein	AY081355	AK103033	71.4-72.5		OSJNBa0001O14	AC025783	6e-85
BU673477	expressed protein		AK062858	72.8	C797	OSJNBa0015J15	AC026758	1e-55
BI306181	pEARLI 1-like protein		AK107851	72.8	C797	OSJNBa0015J15	AC026758	0.0
BU672857	beta-expansin	AF261271	AK100959	73.7	E61768S	AE017118	AE017118	0.0
BU673910	expressed protein	NP_176441		73.7	E61768S	OSJNBb0014I11	AC037426	e-167
BI305458	nucleoside diphosphate kinase	AF271362	AK072751	73.7-83.0		OSJNBa0027P10	AC084763	e-161
CB965938	bZIP	AF268596		73.7-83.0		AE017119	AE017119	2e-38
BI305586	carnitine/acylcarnitine translocase	AC060755		73.7-83.0		OSJNBa0003O19	AC060755	0.0
BU673508	transcription factor MYBS3		AK101062	73.7-83.0		OSJNBa0042H09	AC079874	3e-57
BU673620	unknown	AY086010	AK104818	73.7-83.0		OSJNBa0003O19	AC060755	e-132
BI306535	hypothetical protein	AC084763	AK107529	73.7-83.0		OSJNBa0027P10	AC084763	1e-91
BI306515	unknown protein			83.0	C371	OSJNBa0027L23	AC018929	e-110
BU673832	CEO protein	AC027037	AK072769	83.5	F15	OSJNBa0035H01	AC027037	
BI305345	CEO protein	AC027037	AK072769	83.5	F15	OSJNBa0035H01	AC027037	0.0
BI306142	ubiquinone oxidoreductase subunit	AC018727	AK103692	83.8	C405	OSJNBa0056G17	AC018727	e-170
BU673696	expressed protein	NP_568647		83.8	C405	OSJNBa0056G17	AC018727	e-145
AccNo	**Putative**	**ACC NO**		**cM**	**Marker**	**chromosome 11**	**BAC ACC**	**E-VAL**
BU673266	conserved hypothetical protein		AK073940			Ba0060B06	AC136149	0.0
CB964923	unknown protein		AK105759			OSJNBa0039D03	BX000499	7e-91
BI305512	lipid transfer protein precursor	U29176	AK104005			OSJNBa0025K19	CNS08CE2	e-116
BU673496	phospholipid transfer protein precursor	J04176	AK073363			OSJNBb0004B05	AC123514	5e -63
BU673016	RuBisCO activase large isoform precursor	AB034698	AK104332			OSJNBb0026K20	AC133008	e-143
BI306191	protein transport protein subunit					OSJNBa0032J07	BX000501	e-154
BI305569	alpha-tubulin	X91807	AK102560			OSJNBa0095K08	AC137924	0.0
CB966474	No hit		AK103742	6.0	S10792	OSJNBa0056E15	AC123525	4e-24
CB964899	No hit		AK067080	8.1	C718	OSJNBa0068G15	AC123527	e-146
BU673563	No hit		AK058585	8.6	R642A	OSJNBa0059J06	AC123526	1e -61
BI306392	putative protein	NM_122168	AK068167	9.2	R1938	OSJNBa0074E19	AC120307	e-128
BU673291	expressed protein	NM_122164	AK058243	9.2		OSJNBa0074E19	AC120307	0.0
CB965321	Expressed protein		AK058630	10.3	C83B	OSJNBa0032M21	AC138196	e-139
CB964845	F-box containing tubby family protein	NP_173899	AK070139	17.3	E50055S	OSJNBa0081F16	AC120533	0.0
CB964501	expressed protein		AK104405	18.1	R10571S	chromosome 11	AY257863	4e-83
BI305739	abscisic acid- and stress-inducible protein (Asr1)	AF039573	AK064892	18.1	R10571S	OSJNBa0011J22	AC120527	0.0
CB964857	ribosomal protein L3		AK099225	18.1	R10571S	chromosome 11	AY257863	3e-37
BU672886	protein kinase Xa21	U72725		19.0	S10318	OSJNBb0007B15	AC128643	7e -73
BU672814	aldolase	AF017362	AK104899	19.0	S10318	OSJNBb0007B15	AC128643	e-148
CB964902	aldolase	AF017362	AK104899	19.0	S10318, S20116S	OSJNBb0007B15	AC128643	e-167
CB966095	transcription factor PCF3	AB071804		19.8	S1559	OSJNBb0063D09	AC120539	e-113
BU672826	disease resistance response protein	NM_105259	AK063148	19.8, 20.3	S1559,S10616A	OSJNBa0019D06	AC134044	0.0
CB965307	No hit		AK111265	20.3-27.8		OSJNBa0052O08	AC134047	e-131
CB965301	No hit		AK065928	20.3-27.8		OSJNBb0084H09	AC135794	8e-58
BU673758	OsNAC5 protein	AB028184	AK063399	20.3-27.8		OSJNBa0052O08	AC134047	1e -52
BI306574	adh1-adh2 region	AF172282		30.5	C496	OSJNBa0037P04	AC123521	1e-44
BU673390	bZIP transcription factor		AK072267	32.1	R2316	OSJNBb0074K09	AC123517	0.0
BI305564	CP26	D85512	AK104824	45.3-45.6		OSJNBb0091I11	AC135460	0.0
CB965072	Expressed protein			55.9	C53961S	OSJNBa0015P05	AC112208	e-126
CB965634	transposase		AK067349	56.2	S790A	OSJNBa0086N07	AC135561	e-143
BI306554	serine carboxypeptidase	AC079632	AK059355	56.2	S20262S	OJ1115_A03	AC128642	e-147
BU673037	cytokinin binding protein CBP57	D16139	AK059712	57.3	R1891	OSJNBa0041C22	AC137752	0.0
CB965160	hydroxymethyltransferase	T05907	AK071541	57.3	R1891	OSJNBa0041C22	AC137752	0.0
BI305774	6-phosphogluconate dehydrogenase	AF061839	AK059546	64.2	R728	OSJNBb0041J18	AC137745	0.0
BI306451	11-beta-hydroxysteroid dehydrogenase-like	AB023037	AK106189	68.4	G257	OSJNBb0019E05	AC124151	e-161
BI305884	Xaa-Pro aminopeptidase 1	AJ308541	AK105673	79.1		OSJNBa0007D07	AC136481	0.0
BU672815	putative protein	NM_116900	AK064128	80.2	C1172	OSJNBa0038F07	AC108224	2e - 89

BI306741	50S ribosomal protein L4	AY072419	AK069751	85.7	S723	P0480H08	AC104847	1e -78
CB965959	hypothetical protein		AK062126	101.9-104.9		Ba0090F16	AC122143	e-126
CB966302	dormancy-asociated protein	AF467730	AK060981	112.9	R251	Bb0076M06	AC134924	0.0
CB965481	NBS-LRR-like protein	AY043283	AK064454	112.9-114.4		Ba0004D06	AC134922	0.0
BI306315	ATP-citrate-lyase	AC003970	AK061978	117.0	R3342S, Y6855RA	OSJNBb0005C17	AC112658	e-157
BI306598	No hit			117.0	Y6855RA	OSJNBb0005C17	AC112658	7e-51
CB967502	RuBisCO activase large isoform precursor	AB034698	AK104332	117.9	S10003	OSJNBb0003C04	AC137064	6e-58

AccNo	Putative	ACC NO		cM	Marker	chromosome 12	BAC ACC	E-VAL
BU673300	wali7	L28008	AK071762			OSJNBa0039D03	AL513005	e-160
BI305570	Rubisco subunit binding-protein alpha subunit	X07851	AK061410			OSJNBb0115B15	CNS08CCL	0.0
BI306099	arm repeat containing protein homolog	NM_114518	AK105835			OSJNBa0016C14	AL513403	e-114
BU673914	Pridicted protein		AK061229			OSJNBa0073H17	AL713909	e-138
BI305567	clathrin heavy chain		AK058226	0.0-5.5		OSJNBb0068K19	BX000491	0.0
CB964589	No hit		AK058279	7.1	G193	OSJNBa0024J08	BX000492	e-113
BU673536	lipid transfer protein LPT III	AF017360	AK059808	9.4-9.7		OJ1136_E08	CNS08CE1	8e-32
CB965579	hypothetical protein		AK103629	9.4-9.7		OSJNBb0027B07	BX000500	e-101
CB967024	ribosomal protein S16	L36313	AK058697	9.7		OSJNBb0119L20	BX000495	0.0
BI305549	unknown protein			9.7	G1112	OJ1085_G07	BX000509	e-111
CB967074	expressed protein		AK061862	9.7		OSJNBb0119L20	BX000495	7e-50
CB966708	glutamine-fructose-6-phosphate transaminase		AK105856	9.7		OSJNBa0009F13	BX000498	8e-53
CB966006	conserved hypothetical protein		AK105296	10.8	C1116A, R769B	OJ1575_G05	BX000508	0.0
CB967305	hypothetical protein		AK105075	12.2-26.0		OSJNBb0050K10	AL928755	e-164
BI305396	AT5g09850/MYH9_6	AY058141	AK099823	27.1	C1589S	OSJNBa0041K23	AL513404	2e-75
BI306381	hypothetical protein	NM_105056	AK068475	27.1	C1589S	OSJNBa0041K23	AL513404	0.0
BU673656	hypothetical protein		AK101515	29.2	S20454S	OJ1119_B11	AL731761	0.0
CB965019	P49107		AK059037	30.0-38.1		OSJNBa0002O20	CNS08CBM	e-136
BI305619	carboxyvinyl-carboxyphosphonate phosphorylmutase		AK068710	30.0-38.1		OSJNBa0042N11	AL954157	e-133
BI306747	alpha-tubulin		AK068710	30.0-38.1		OSJNBa0042N11	AL954157	2e -91
BI306290	Ca2+ sensitive 3'(2'),5-diphosphonucleoside 3'(2') phosphohydrolase	U33283	AK058937	30.0-38.1		OSJNBb0089D09	AL954158	e-162
BI306697	chloroplast atpB for ATP synthase beta subunit,	AB037543		39.7	R3375	OSJNBa0021D06	AL513004	e-135
BI306498	hypothetical protein		AK103802	41.8	C60015	OSJNBb0036J04	AL607103	0.0
CB966823	conserved hypothetical protein		AK107948	48.2	S10904	OSJNBa0077B14	AL831810	2e-49
BU673906	ribulose bisphosphate carboxylase/oxygenase	L22155	AK061611	49.3	S10704	OSJNBa0056I18	CNS08C7O	0.0
BI305808	cytochrome P450-like sequence	AF088221	AK071599	49.3		OJ1111_C09	AL713951	0.0
BI305598	small subunit of ribulose-1,5-bisphosphate carboxylase		AK068555	50.4-51.5		OJ1523_A01	CNS09S4U	0.0
BI306234	putative protein		AK060975	50.4-51.5		OSJNBb0085N21	CNS08CA8	0.0
BU673060	unknown protein		AK064558	50.4-51.5		OSJNBa0012G19	AL831796	0.0
CB965133	hypothetical protein		AK103895	51.5	R887	OJ1771_D09	BX000557	0.0
BI305514	photosystem I protein (PSI-L)	M61146	AK099109	51.8	S14025	OJ1112_F01	AL732381	0.0
BI306495	NB-ARC domain		AK067669	51.8-55.1		OJ1112_B07	AL731753	4e -64
BU673466	unknown protein			58.9-61.6		OSJNBa0090O14	AL731763	1e -90
CB964439	unknown protein			58.9-61.6		OSJNBa0038H18	AL732639	e-134
BU673242	one helix protein	AAM22751		61.6	E4418S	OSJNBa0005P03	AL513002	e-145
CB967093	50S ribosomal protein L4			62.2-64.4		OSJNBa0087A02	CNS08CAC	1e-29
CB967254	unknown protein			64.4-65.3		OSJNBa0018L16	CNS09S4W	e-112
BI306748	No hit		AK070613	66.6	S21024S	OSJNBa0009K11	AL731759	e-144
CB966431	hypothetical protein			66.6	S21024S	OSJNBa0009K11	AL731759	0.0
CB966584	initiation factor eIF-4D	D10399	AK099039	66.6	S21024S	OSJNBa0009K11	AL731759	e-138
CB967146	hypothetical protein			71.2	C11001SB	OJ1126_F07	CNS07YQ0	2e-97
BU673094	Fo-ATPase subunit 9	X16936	AK064564	71.2	S11679	OJ1521_B04	AL713949	0.0
BU673555	ribosomal proteins and tRNA-fMet	D32052	AK107620	71.2-71.7		OJ1521_B04	AL713949	0.0
CB966422	acyl carrier - like protein		AK060472	73.0	C751, C449	OSJNBb0094E08	AL713932	0.0
BU673076	nucloside diphosphate kinase 2	AB078008	AK070054	75.8		OJ1123_B09	AL731742	2e -30
BI306153	cytochrome c oxidase subunit 5c	AB027123	AK072527	78.9-86.5		OSJNBa0028L05	AL935072	e-177
BI305839	unknown protein		AK105895	86.5	R10289S	OSJNBb0076G11	CNS08CBQ	e-155
BU673633	unknown protein	AB026636	AK060401	86.5	R10289S	OSJNBb0076G11	AL928753	e-129
CB965421	hypothetical protein			88.6	C2808	OSJNBa0016C14	AL731784	e-115
BI305532	No hit		AK109459	88.6-91.4		OSJNBa0016C14	AL513403	e-151
CB967448	metallothionein 2a	S57768	AK103445	88.6-91.4		OSJNBa0018E22	AL845346	3e-83
BU672942	small zinc finger-like protein (TIM9)	AF150113		91.4	R1709	OJ1396_C02	AL772425	0.0
CB965674	transcription factor OsRS2	AB071600		91.4	R1709	OJ1396_C02	CNS08CA3	6e-98
CB967153	unknown protein		AK106267	91.9-94.6		OSJNBa0014A07	CNS08C8O	e-175
BI305436	diacylglycerol kinase	D63787	AK100331	91.9-94.6		OSJNBa0014A07	AL731884	0.0
CB967090	D-type cyclin		AK101602	95.4	G148	OJ1573_A06	AL954852	3e-67
BI305637	N-carbamyl-L-amino acid amidohydrolase-like protein	AB016875	AK103799	97.3	C1069	OSJNBa0001B02	AL731743	e-159
BI305840	No hit			97.3-99.7		OSJNBa0035N13	AL732535	1e -36
BI305806	auxin-regulated protein	J03919	AK068213	99.7	S861	OSJNBb0062H20	AL837528	e-129
BU673773	bZIP transcription factor	AC009917	AK109716	100.9-103.1		OJ1327_A12	AL713940	e-120
CB965473	expressed protein	NP_680193		105.1	C51116S	OSJNBb0101I10	AL845347	0.0
CB966269	unknown protein		AK072503	105.1-106.6		OSJNBa0027I07	AL732532	0.0
CB965103	RAB5A protein			108.2	R2292, S13561	OJ1559_C07	CNS08C8R	8e-29
BI306200	glycine-rich protein	AJ302060	AK073614	108.2	R2292, S13561	OJ1559_C07	CNS08C8R	e-112
BI305802	small GTP-binding protein (rab5A)	AY029301	AK066784	108.2	R2292	OJ1584_D02	AL713905	e-170

Table 4.2(b): WGS contig localization of non-redundant ESTs

WGS contig: Whole Genome Shotgun rice genomic fragment contig ID
GenBank Acc: Identical sequence accession in GenBank
FL cDNA: full-length cDNA
NL: Identical rice genomic fragment was not located in the available rice genomic sequence in the database

EST Acc	Putative	GenBank Acc	FL cDNA	WGS contig	VGS e-valu
BU673549	28S ribosomal RNA	X90411		AAAA01042579	
BI306717	ALANYL-TRNA SYNTHETASE, MITOCHONDRIAL PRECURSOR		AK066368	AAAA01003899	2e-25
BI305898	aldolase C-1	D50307	AK104719	AAAA01028973	0.0
CB965544	aspartic proteinase		AK068348	AAAA01001103	e-144
BI305258	ATP-dependent Clp protease proteolytic subunit	NM_103884	AK071333	AAAA01002736	e-116
BU672830	Caffeoyl CoA O-methyltransferase	AJ242980	AK104801	AAAA01005061	
BI305850	calmodulin (CaM1)	AF042840	AK070090	AAAA01015948	e-158
BU672831	cap-binding protein p28	AC022457		AAAA01013971	
BI305248	dehydrin	U60097		AAAA01012244	e-160
BU673228	DRE-binding protein 1B	AF300972	AK062422	AAAA01001957	
BI306163	early nodulin	AB018377	AK072428	AAAA01003189	e-146
BI306052	enolase	U09450	AK099292	AAAA01004378	1e-52
BU673121	extensin-like protein	AB073166	AK106204	AAAA01001290	
BI306257	F-box protein family, AtFBL6		AK067705	AAAA01003642	5e-79
BU673649	glutathione S-transferase	AF309381	AK069105	AAAA01000768	
BI306250	glycine-rich protein	AF010579	AK070016	AAAA01009071	e-133
BU673251	histone H2A	AY062171		AAAA01043039	
CB965826	inorganic pyrophosphatase		AK104659	AAAA01002086	8e-44
BI306605	metalloprotease			AAAA01000216	e-104
BI306252	metallothionein-like protein	U18404		AAAA01023086	2e-68
BI306310	NDR1/HIN1-like protein	AL589883	AK104600	AAAA01002765	3e-80
BU673475	peroxidase	AB027752	AK069456	AAAA01004379	
BI305589	phenylalanine ammonia-lyase	X16099	AK058306	AAAA01007822	e-167
BU672866	chlorophyll synthase	AC087599	AK071899	AAAA01000620	
BU672863	GTP-BINDING PROTEIN YPTV3	AC079843	AK071529	AAAA01009076	
BU673036	phospholipid cytidylyltransferase	NM_129424	AK068868	AAAA01006697	
CB965512	putative protein		AK067663	AAAA01002393	2e-41
BI306279	senescence-associated protein	AB049723		AAAA01039016	e-124
BI305744	water channel protein		AK072531	AAAA01004842	1e-73
BI306047	RNA binding protein		AK065531	AAAA01000411	2e-65
BI305658	serine threonine kinase, putative	NM_099996		AAAA01000865	e-124
BU673788	syringolide-induced protein 14-1-1	AB083026	AK073400	AAAA01016038	
BI305736	transaldolase	U95923	AK105800	AAAA01000199	e-164
BI305424	WOUND-INDUCED BASIC PROTEIN	D30015	AK073415	AAAA01033026	2e-71
BI305452	zinc finger protein			AAAA01002232	0.0
BU673135	unknown cold induced protein	AY090535	AK067043	AAAA01005444	
BU673413	expressed protein	NM_103640	AK067180	AAAA01006696	
BI306547	expressed protein			AAAA01004943	e-121
BU673562	expressed protein	NM_100294	AK103688	AAAA01000569	0.0
CB965184	expressed protein	NP_565392	AK110558	AAAA01012766	
BI306158	hypothetical protein	AP002855		AAAA01004936	0.0
BI305848	hypothetical protein		AK064766	AAAA01016109	0.0
BI306085	hypothetical protein	AF114171	AK058813	AAAA01004390	e-103
CB967189	unknown mRNA		AK070257	AAAA01019605	6e-47
BU673285	unknown protein	AC013258	AK099407	AAAA01003242	
BI306278	unknown protein	AB018121		AAAA01003064	e-146
BI305375	unknown protein	AC004667	AK068177	AAAA01020019	3e-84
CB966989	unknown protein		AK105686	AAAA01012855	
BI305937	unknown protein	AC008261	AK099511	AAAA01006847	2e-93
BI306212	No hit			AAAA01013895	2e-53
BI305430	No hit			AAAA01001070	2e-51
BI306450	No hit			AAAA01008643	1e-16
BI306287	hypothetical protein			AAAA01001705	e-101
BI305536	No hit			AAAA01004275	2e-25
BI305616	unknown protein		AK072656	AAAA01007495	e-109
BI306243	RE45246p			AAAA01009117	1e-30
BI306112	unknown protein		AK065458	AAAA01002339	0.0
BI305759	No hit			AAAA01007890	5e-85
BI305805	C-type U1 snRNP		AK060957	AAAA01000425	e-108

BI305827	conserved hypothetical protein			AAAA01000134	2e-44
BI306119	No hit		AK099523	AAAA01017256	e-150
BI306118	No hit			AAAA01005499	6e-72
BI305915	unknown protein		AK060419	AAAA01000560	4e-88
BI306055	No hit			AAAA01001826	1e-33
BI306578	No hit			AAAA01002858	4e-57
BI306241	unknown protein		AK103473	AAAA01005941	1e-74
BU673177	No hit			AAAA01018808	
BU673441	No hit		AK061903	AAAA01000850	
BU673437	kinetochore protein			AAAA01005863	
CB964559	No hit		AK068288	AAAA01000323	
BI306550	No hit			AAAA01009797	1e-63
BU673640	No hit		AK059903	AAAA01005523	
BI306552	No hit			AAAA01003608	4e-63
BU672937	No hit		AK099158	AAAA01004451	
BU672906	No hit		AK100953	AAAA01016474	
BU673614	No hit		AK071025	AAAA01038708	
BU673287	No hit		AK102217	AAAA01001987	
BI306698	No hit			AAAA01004980	3e-43
BI306610	No hit			AAAA01004571	1e-63
BI306696	No hit			AAAA01006526	9e-49
BU673662	No hit			AAAA01045789	
BI306703	glycine decarboxylase subunit			AAAA01003782	2e-16
BI306715	No hit			AAAA01015087	1e-33
BU672801	hypothetical protein			AAAA01011477	
BI306622	No hit			AAAA01001980	2e-86
BI306746	BARE-1 copia-like retroelement	Z17327		AF474373	e-179
BI306594	No hit			AF474373	0.0
BU673140	14-3-3 protein	AY029473		NL	
BI306744	21kd polypeptide	D12626		NL	
BU673436	28S ribosomal RNA	AY049041		NL	
BI305868	40S ribosomal protein S23	AJ291613		NL	
CB966175	60S ribosomal protein L6	CAB57309		NL	
CB965527	60S ribosomal protein L7	NP_850411		NL	
BI305413	60S RIBOSOMAL PROTEIN L9 induced by GA3	D83527		NL	
CB967130	ABC transporter system integral membrane protein	T31089	AK106112	NL	
CB967253	acid phosphatase-like protein	CAB71336		NL	
CB964896	aconitate hydratase	NP_178634		NL	
BU673430	Actophorin			NL	
BI305832	adenosine kinase	AJ012281		NL	
BI306679	adenylate cyclase	AP003583		NL	
CB967186	apyrase		AK104429	NL	
BI306614	aspartic proteinase	D12777		NL	
BI306571	ATP/ADP translocator	D12637	AK100143	NL	
CB965615	calmodulin TaCaM1-3	U48689.	AK071852	NL	
BU672981	catalase			NL	
BI305491	CG6214 gene product (MDR)	XM_079251		NL	
BI305355	choline kinase	U43838		NL	
CB965599	conserved hypothetical protein	NP_706553		NL	
BU673319	cyclophilin	AF017993		NL	
BI305887	cytochrome oxidase subunit 2	AF179912		NL	
BI306357	cytochrome P450 like_TBP	D64052		NL	
BI305377	DNA ligase	AJ414141		NL	
CB964731	Elongation factor 1-alpha	CAC10565		NL	
BI305336	ERD1 protein	D17582	AK068727	NL	
CB966971	fumarylacetoacetate hydrolase	AAP36709		NL	
CB967472	glucose-repressible protein grg-1	CAC28672		NL	
BI306655	glyoxalase I	AB017042	AK103694	NL	
CB965947	heat shock protein, Hsp20 family	AAM73188		NL	
BI305879	homogentisate 1,2-dioxygenase	AF149017		NL	

BU673298	ISRSO14-TRANSPOSASE ORFA PROTEIN	AL646061		NL
CB966415	ketopantoate hydroxymethyltransferase	AF134703		NL
BI305238	leucine-responsive regulatory protein	AL627268		NL
BI306667	lipid transfer protein LPT IV	AF017361	AK070414	NL
BI306045	lipid transfer protein precursor	U29176		NL
BI305982	magnaporin	AF126872		NL
BI305284	metalloprotease	NM_112804	AK060417	NL
CB966340	metallothionein			NL
BI306678	metallothionein-like protein	U57638		NL
CB965852	metallothionein-like type 1			NL
CB966931	mitochondrial processing peptidase		AK060718	NL
BI305365	NAD-DEPENDENT MALIC ENZYME	AE000245		NL
CB964594	NADH dehydrogenase I chain F	AE004692		NL
BI305457	NADH dehydrogenase I chain J	D91024		NL
BI305230	nicotianamine aminotransferase A	D88273	AK060557	NL
CB966039	nucleoside diphosphate kinase	AF271362		NL
BI306313	oligopeptide transporter	AP003235		NL
BI306493	photosystem II 10 kDa polypeptide	U86018		NL
BI306262	photosystem II 10 kDa polypeptide	U86018		NL
BI305432	probable membrane protein yibP	D86037		NL
CB966750	probable RIC1 protein	CAD70889		NL
CB965851	auxin-induced protein			NL
BI305993	S-adenosylmethionine:2-demethylmenaquinone methyltransferase	AC018363	AK068833	NL
BI305193	respiratory burst oxidase homolog	AB050661		NL
BI305748	reversibly glycosylated polypeptide	Y18624		NL
BU673461	ribosomal protein L27	AF400191		NL
BU672848	ribosomal protein L38	U86017	AK058262	NL
CB966919	ribosomal protein S17	AAN52389		NL
BU672838	Ribosomal protein S25	AAL48698		NL
BI305454	ribulose 1,5-bisphosphate carboxylase small subunit	D00644		NL
BI305519	ribulose bisphosphate carboxylase/ oxygenase	L22155	AK061611	NL
CB965581	RicMT		AK104075	NL
CB965846	salT gene	Z25811	AK062520	NL
BI305307	signal peptidase I	AE004511		NL
CB966534	Skp1	AY050559		NL
CB967465	stromal cell-derived factor 2-like 1	NP_071719		NL
BI305587	symbiosis-related like protein	NM_117751		NL
BI306465	thioredoxin h	D21836		NL
CB965166	translationally controlled tumor protein homolog	NP_594328		NL
BU673611	transposon-like sequence	AJ309824	AK106431	NL
CB965085	type I light-harvesting chlorophyll a/b binding protein of photosystem II	D00641	AK104281	NL
BI305439	ubiquinone/menaquinone biosynthesis methyltransferase	AAP18852		NL
CB965120	ubiquitin fusion protein	AAA70104		NL
CB965545	expressed protein	NP_849563		NL
BU673735	Expressed protein			NL
CB965365	expressed protein	NP_563782		NL
CB967122	expressed protein	NP_565590		NL
BI305286	hypothetical protein	NP_418545		NL
BU673531	hypothetical protein		AK068177	NL
CB966228	hypothetical protein	XP_324816		NL
BI305392	hypothetical protein	T38669		NL
BI305513	predicted protein			NL
CB965359	Pridicted protein		AK065107	NL
CB967242	Pridicted protein			NL
CB967060	Pridicted protein			NL
CB965609	Pridicted protein			NL
CB964957	Pridicted protein			NL

CB964921	unknown	AAM63926		NL
BU673581	unknown protein		AK069867	NL
BU672797	unknown protein	AB017068		NL
CB964544	hypothetical protein			NL
CB964544	hypothetical protein			NL
BU673661	conserved hypothetical protein			NL
CB964498	No hit		AK101444	NL
CB964487	hypothetical protein		AK067400	NL
BU673825	No hit		AK103440	NL
CB964506	fumarylacetoacetase		AK105018	NL
CB966223	Chlorophyll a/b binding protein			NL
BU673469	No hit		AK066594	NL
CB964853	No hit		AK058313	NL
CB964867	No hit		AK072445	NL
CB964918	hypothetical protein			NL
CB964992	40S ribosomal protein S30 homolog			NL
CB966790	hypothetical protein			NL
CB966398	light-regulated unknown 11 kDa protein			NL
CB966401	hypothetical protein			NL
CB967320	hypothetical protein			NL
CB967315	ribosomal protein L37			NL
CB965478	CEO protein			NL
CB965402	conserved hypothetical protein		AK102098	NL
CB965489	No hit		AK072650	NL
CB965221	No hit		AK068659	NL
CB965228	guanylate kinase			NL
CB965718	myosin heavy chain-like protein			NL
CB965719	hypothetical protein			NL
CB965687	arginine/serine-rich splicing factor			NL
CB965685	unknown protein			NL
BI305280	No hit		AK059070	NL
BI306227	hypothetical protein			NL
CB965927	unknown protein		AK103942	NL
CB965931	cell wall protein			NL
CB966017	hypothetical protein			NL
CB965926	light-regulated unknown 11 kDa protein			NL
BU673063	osERF3			NL
BU672954	hypothetical protein			NL
CB965771	TMV induced protein 1-2			NL
BU673349	HMGI/Y protein			NL
CB964570	OSMYB1			NL
BI305231	unknown protein			NL
CB964795	No hit		AK059353	NL
CB965758	No hit		AK070645	NL

4.6 Putative known candidate genes in the Quantitative trait loci

Target QTL regions were extracted from reference genetic maps (Harushima et al. 1998) and QTL studies associated with drought and yield traits (Price et al. 2002; Zhang et al. 2001; Babu et al. 2003). The known putative stress responsive genes identified in the QTL locations (Table 4.3).

Table 4.3: Known stress responsive genes at QTL locations

Signaling	
Mitogen-activated protein kinase homolog MMK 2	Chromosome 10
Putative receptor-like protein kinase	”
Small GTP-binding protein (rab5a)	Chromosome 12
14-3-3 protein homolog GF14-12	Chromosome 8
Signal recognition particle receptor-like protein	”
Calcium dependent protein kinase	Chromosome 5
1-aminocyclopropane-1-carboxylate oxidase	Chromosome 7
Transcription Factors	
~~EREBP~~-like protein	Chromosome 3
AP2 domain containing protein	”
Ethylene responsive protein	”
Helicase-like transcription factor	Chromosome 4
OSMYB1	Chromosome 5
Homeodomain leucine zipper protein	Chromosome 8
Metabolism	
Sucrose-6F-phosphate phosphohydrolase SPP3	Chromosome 1
Beta-oxyacyl-[acyl-carrier protein] reductase	Chromosome 3
Putative anthocyanidin reductase	”
HMG protein	Chromosome 1
RNase S-like protein	Chromosome 9
Membrane proteins	
Photosystem I chain IV precursor	Chromosome 7
Mitochondria F0 ATP synthase D chain	Chromosome 8
Water channel protein	Chromosome 8

4.7 Identification of putative abiotic stress responsive genes

The putative candidate drought stress responsive genes have been identified based on experimental data (Bray. 2002), and comparing the EST dataset to the documented gene sequences and possible candidate gene sequences identified from micoarray gene expression profiles of rice, *Arabidopsis*, barley plants under different abiotic stresses. The stress responsive genes identified from the above expression profiles include 650 from *Arabidopsis*, 150 from barley and 100 from rice. These were compared to the EST data set using TBLASTX with an E-value >1e-20. Interestingly the distribution of the 583 putative stress responsive ESTs among the functional categories showed that transcription factors were efficiently captured next only to the cell metabolism class.

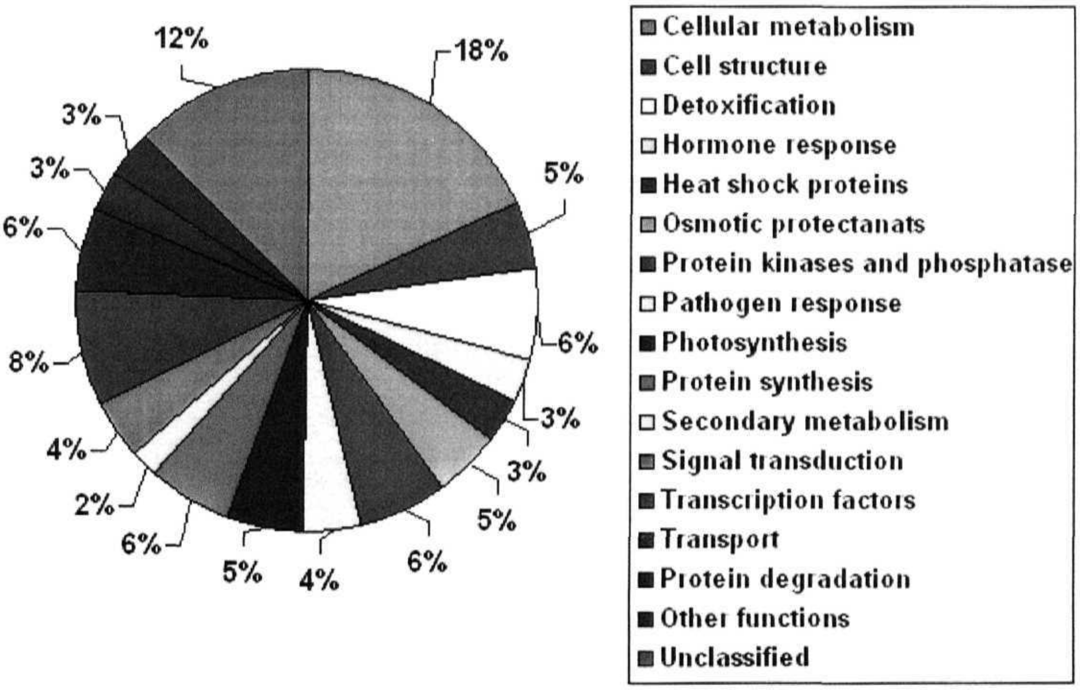


Fig. 4.14: Functional classification of Putative Stress Response Genes (SRG)

The putative stress responsive genes identified from global gene expression profiles using cDNA microarrays and experimentally documented genes along with the source of identification and type of stress are given in (Table 4.4).

Table 4.4: List of identified putative stress responsive genes

ASRG paralouges: *Arabidopsis* stress responsive gene family members
Arabidopsis HL: *Arabidopsis* high light stress
Arabidopsis TF: Two fold difference in the expression profiles of *Arabidopsis* stress responsive genes under cold, drought, and Mannitol.
SUR: salt up regulated genes in rice or barley
ADR: Down regulated *Arabidopsis* stress responsive genes under cold, drought, and salinity.
SDR: salt down regulated genes in rice or barley
DUR: Drought upregulated genes in barley
DDR: Drought down regulated genes in barley

EST Acc	FL-cDNA	Putative Function	GenBank Acc	Source Reference	Organism and stress
		Cellular Metabolism			
BU673417	AK073146	ADP-glucose pyrophosphorylase small subunit	AY028315	Kimura et al.(2003)	Arabidopsis HL
BI305757		beta-amylase	L10346	Kreps et al. (2002)	Arabidopsis TF
BI305259	AK060557	nicotianamine aminotransferase A	D88273	Seki et al.(2002)a	SUR
BI305391		beta-oxyacyl-[acyl-carrier protein] reductase	AJ243091	ASRG paralouges	
BI306140	AK059546	6-phosphogluconate dehydrogenase	AF061839		
BI305887		cytochrome oxidase subunit 2 (cox-2)	AF179912	ASRG paralouges	
BI305862	AK103172	peptidylprolyl cis-trans isomerase	X86903	Kreps et al. (2002)	Arabidopsis TF
BI306264	AK104987	shoot GS1 for cytosolic glutamine synthetase	X14245	Kreps et al. (2002)	Arabidopsis TF
BU673457	AK104842	P40-like protein	AB012702	Ozturk et al. (2002)	SUR
BI306288	AK059867	copper amine oxidase	NM_129810	Kreps et al. (2002)	Arabidopsis TF
BI306176	AK062851	glycine decarboxylase subunit	AF022731	Seki et al.(2002)a	ADR
BI305879		homogentisate 1,2-dioxygenase	AF149017	ASRG paralouges	
BI306517	AK098938	beta-glucosidase-like protein	AY056828	Seki et al.(2002)a	ADR
BI305808	AK071599	cytochrome P450-like sequence	AF088221	Seki et al.(2002)b	Arabidopsis
BI305831	AK098923	sucrose synthase	Z15028	Kawasaki et al. (2001)	SDR
BU672938		Phosphoglycerate dehydrogenase-like protein	AY086001	Kreps et al. (2002)	Arabidopsis TF
BU673621	AK101577	glutamine phosphoribosylpyrophosphate amidotransferase	L23834	Kreps et al. (2002)	Arabidopsis TF
BI305885	AK065295	acetohydroxy acid isomeroreductase	AJ251333	Ozturk et al. (2002)	
BI306752	AK063764	cytochrome P450	AF321867	Seki et al.(2002)a	SUR
BI305557		ubiquinol--cytochrome c reductase	X79275	ASRG paralouges	
BI306026	AK104719	aldolase C-1	D50307	Kreps et al. (2002)	Arabidopsis TF
BI306674	AK067732	alanine aminotransferase-like protein	AY042902	ASRG paralouges	
BI305253	AK064960	cytosolic glyceraldehyde3phosphate dehydrogenase GAPDH	AF251217	Seki et al.(2002)a	ADR
BI306514	AK072650	3-hydroxyisobutyryl-coenzyme A hydrolase	NM_117410	Kreps et al. (2002)	Arabidopsis TF
BU672779	AK062745	ubiquinol--cytochrome-c reductase-like protein	AL353912	ASRG paralouges	
BI306500	AK070623	dTDP-glucose 4-6-dehydratase	NM_128345	ASRG paralouges	
BI305681		cytochrome P450	AB023038	Seki et al.(2002)b	Arabidopsis
BI306478	AK105036	phosphoshikimae1carboxyvinyltransferase	AB052962	Kreps et al. (2002)	Arabidopsis TF
BI306425	AK067934	beta-glucosidase homolog	AP003272	Seki et al.(2002)a	ADR
BI305736	AK105800	transaldolase	U95923	ASRG paralouges	
BI306357		cytochrome P450 like_TBP	D64052	Kreps et al. (2002)	Arabidopsis TF
BI306315	AK061978	Similar to ATP-citrate-lyase	AC003970	ASRG paralouges	
BI306290	AK058937	Ca2+ sensitive 3'(2'),5-diphosphonucleoside 3'(2')phosphohydrolase	U33283	Kreps et al. (2002)	Arabidopsis TF
BI306501	AK067029	ADP-ribosylation factor	D17760	ASRG paralouges	
BU673110	AK071279	cysteine synthase	AF073697	Kreps et al. (2002)	Arabidopsis TF
BI306726	AK068241	enoyl CoA hydratase	AJ275305	ASRG paralouges	
BU672806	AK067200	P450	AB042627	Seki et al.(2002)b	Arabidopsis
BU672814	AK104899	aldolase	AF017362	Ozturk et al. (2002)	SUR
BU672850	AK060831	succinic semialdehyde dehydrogenase	AF117335	Seki et al.(2002)b	Arabidopsis
BI305374	AK060423	ornithine/acetylornithine aminotransferase	AF022915	Ozturk et al. (2002)	SDR
BI305272	AK105255	cytochrome P450	AB023038	Seki et al.(2002)b	Arabidopsis
BI305756	AK071695	lysine decarboxylase-like protein	AB006700	ASRG paralouges	
BI306593	AK099939	pyrophosphate-fructose 6-phosphate 1-phosphotransferase alpha-subunit	M55190	Kreps et al. (2002)	Arabidopsis TF
BI305761	AK071525	sucrose-6F-phosphate phosphohydrolase SPP3	AY029159	ASRG paralouges	
BU673116	AK060890	chloroplast carbonic anhydrase	U08404	Kreps et al. (2002)	Arabidopsis TF
BI305212		respiratory burst oxidase homolog	AB050661	Kreps et al. (2002)	Arabidopsis TF
BI305211	AK067642	hydroxypyruvate reductase	AB060810	Seki et al.(2002)a	ADR
BI305362	AK067200	cytochrome P450	X81828	Kreps et al. (2002)	Arabidopsis TF
BI305631	AK064821	cytochrome B5	X75670	Kawasaki et al. (2001)	SDR
BI306720	AK104533	arabinogalactan-like protein	AP004194	Kreps et al. (2002)	SUR
BI305318	AK062559	non phosphorylating glyceraldehyde-3-phosphate dehydrogenase	X75326	Seki et al.(2002)b	Arabidopsis
CB965408	AK069137	phosphoethanolamine methyltransferase	AAL40895	Ozturk et al. (2002)	DDR
BI305298		cytochrome P450 monooxygenase	AJ004810	ASRG paralouges	
BI306458	AK103963	ubiquinol--cytochrome c reductase	X79276	ASRG paralouges	
BI305543	AK073698	mitochondrial malate dehydrogenase	X78800	Seki et al.(2002)a	ADR
BU673612		cytochrome P-450	U32624	Seki et al.(2002)a	SUR
CB964947	AK062214	chloroplast phosphoglycerate kinase	X15233	Kreps et al. (2002)	Arabidopsis TF
CB964525	AK068716	methionyl aminopeptidase-like protein	NP_172785.1|	Kreps et al. (2002)	Arabidopsis TF
BI305673		prolyl 4-hydroxylase, alpha subunit	AC068923	Kreps et al. (2002)	Arabidopsis TF
BI305417	AK100321	cytochrome P450 monooxygenase CYP92A1	AY072297	Seki et al.(2002)a	SUR
BI305524	AK070516	fructose-1,6-bisphosphatase (cytosolic)	AB007193	Seki et al.(2002)a	ADR
BI306408	AK099229	alpha-galactosidase-like protein	NM_115489	ASRG paralouges	

CB966658	AK109382	quinone oxidoreductase -like protein	NM_121703	Kreps et al. (2002)	Arabidopsis TF
CB967383		phosphoribosylanthranilate transferase	AY224452	Kreps et al. (2002)	Arabidopsis TF
BI305771	AK103027	beta-glucosidase	U28047	Kawasaki et al. (2001)	SUR,ADR
BI306681	AK069031	tryptophan decarboxylase	X67662	Kreps et al. (2002)	Arabidopsis TF
BI306523	AK099342	enolase	U09450	Seki et al.(2002)b	Arabidopsis
CB965160	AK071541	hydroxymethyltransferase	T05907	Seki et al.(2002)a	ADR
BI305704	AK104875	S-adenosylmethionine synthetase	AJ296743	Seki et al.(2002)a	
BI305434	AK104484	acyl-CoA:1-acylglycerol-3-phosphate acyltransferase	AP002039	ASRG paralouges	
BI305755	AK103260	succinate dehydrogenase subunit 3 (sdh3)	AF362741	ASRG paralouges	
CB967219	AK066861	cytochrome b5	NP_200168	Kreps et al. (2002)	Arabidopsis TF
CB964609	AK059296	biotin synthase	AAO41898	Kreps et al. (2002)	Arabidopsis TF
BI305202	AK099387	aldolase (T25),	NM_126176	Seki et al.(2002)a	ADR
CB964504	AK067066	trehalose-6-phosphate synthase	NM_106505	Kreps et al. (2002)	Arabidopsis TF
BI306052	AK099292	enolase	U09450	Kreps et al. (2002)	Arabidopsis TF
BI305740		sterol-C5(6)-desaturase	AF099969	Kreps et al. (2002)	Arabidopsis TF
CB965665	AK061229	GDSL-motif lipase/hydrolase protein	NP_199404	Kreps et al. (2002)	Arabidopsis TF
BI305365		NAD-DEPENDENT MALIC ENZYME (NAD-ME)	AE000245	Seki et al.(2002)a	SUR
CB966982	AK101492	glutamate dehydrogenase	D49475	Kreps et al. (2002)	Arabidopsis TF
BI305233	AK104952	fructose-bisphosphate aldolase	AJ133146	Seki et al.(2002)b	Arabidopsis
CB965262	AK101024	malate oxidoreductase	NP_191966	Kreps et al. (2002)	Arabidopsis TF
BI305402	AK059712	cytokinin binding protein CBP57	D16139	Kawasaki et al. (2001)	SUR
CB966346		nicotianamine synthase 2	AB023818	Kawasaki et al. (2001)	SDR
CB966304	AK098950	Pyruvate dehydrogenase E1 component alpha subunit, mitochondrial precursor (PDHE1-A)	BAC57468	Kreps et al. (2002)	Arabidopsis TF
CB966601	AK060559	beta-carotene hydroxylase	CAB55625	Kreps et al. (2002)	Arabidopsis TF
BU673356	AK099252	precursor chloroplastic glutamine synthetase	AF480497	Seki et al.(2002)b	Arabidopsis
BI305440	AK100908	UDP-glucuronic acid decarboxylase	AB079064	Kreps et al. (2002)	Arabidopsis TF
BI305501		polygalacturonase isoenzyme 1 beta subunit	AF251069	Kreps et al. (2002)	Arabidopsis TF
BI305352	AK059503	S-adenosylmethionine decarboxylase 2	AJ251899	Ozturk et al. (2002)	SUR
BU673346	AK072414	amine oxidase	NM_129863	Kreps et al. (2002)	Arabidopsis TF
BI305568		fatty acid elongase-like protein	X88779	Kreps et al. (2002)	Arabidopsis TF
BI305640		lipid acyl hydrolase	AC091123	Kreps et al. (2002)	Arabidopsis TF
BI306651	AK059808	lipid transfer protein LPT III	AF017360	ASRG paralouges	
BI305582	AK106129	lipase	AF026480	Kreps et al. (2002)	Arabidopsis TF
BI306097	AK104005	lipid transfer protein precursor	U29176	Ozturk et al. (2002)	DDR
BI306352	AK070414	lipid transfer protein LPT IV	AF017361	Ozturk et al. (2002)	SUR
BI306045		lipid transfer protein precursor	U29176	Kreps et al. (2002)	Arabidopsis TF
BI306183	AK061118	lipase	U38916	Kreps et al. (2002)	Arabidopsis TF
BI305263	AK067352	lipase	AC007508	Ozturk et al. (2002)	DDR
BU673036	AK068868	phospholipid cytidylyltransferase	NM_129424	Kreps et al. (2002)	
BI305484	AK071598	lipid transfer protein	AF109195	Seki et al.(2002)b	Arabidopsis
BI305994		Similar to lipase	U38916	Kreps et al. (2002)	Arabidopsis TF
CB966380	AK062463	nonspecific lipid transfer protein	U88090	Seki et al.(2002)a	SUR
BU673496	AK073363	phospholipid transfer protein precursor	J04176	Seki et al.(2002)b	Arabidopsis
BI306700		fatty acid longase (fae2)	AJ292770	Kreps et al. (2002)	SUR
		Structural proteins			
BI305985	AK061423	beta-expansin (EXPB7)	AF261275	Kreps et al. (2002)	Arabidopsis TF
BI306497	AK061988	ubiquitin (mub1)	M60175	Ozturk et al. (2002)	SUR
BU673240		pectinesterase	AL132956	Kreps et al. (2002)	Arabidopsis TF
BU673900	AK100267	actin	AJ012685	Kreps et al. (2002)	Arabidopsis TF
BI306341	AK069053	endosomal protein-like	AP000421	Kreps et al. (2002)	Arabidopsis TF
BI305569	AK102560	alpha-tubulin	X91807	Kawasaki et al. (2001)	SDR
BI305581		beta-tubulin 1	U76744	Seki et al.(2002)b	Arabidopsis
CB964785	AK101915	xyloglucan endo-transglycosylase	S71223	Kreps et al. (2002)	Arabidopsis TF
BU673557		histone H2A.F/Z	NP_191019	Kreps et al. (2002)	Arabidopsis TF
CB966661		histone H2B	X82362	Ozturk et al. (2002)	SUR
BI305970	AK104912	xyloglucan endo-1,4-beta-D-glucanase	X93175	Seki et al.(2002)b	Arabidopsis
BI306389	AK104912	xyloglucan endotransglycosylase	X93175	Seki et al.(2002)a	ADR
BI306132		beta-D-glucan exohydrolase, isoenzyme ExoII	A	ASRG paralouges	
BU673300	AK071762	wali7	L28008	Kimura et al.(2003)	Arabidopsis HL
BI306217	AK058741	histone H4	M12277	Kreps et al. (2002)	Arabidopsis TF
CB965683		ovule/fiber cell elongation protein Ghfe1	AF513859	Kreps et al. (2002)	Arabidopsis TF
BU673438	AK065323	beta-tubulin 2	X78142	Seki et al.(2002)a	ADR
BI305824	AK071291	fibrillarin 2	NM_118695	Kreps et al. (2002)	Arabidopsis TF
BI305420		similarity to beta-1,3-glucanase-like protein	AB008265	Kreps et al. (2002)	Arabidopsis TF
BU673121	AK106204	extensin-like protein	AB073166	Kreps et al. (2002)	Arabidopsis TF
BU673405	AK099211	histone-like protein	AF394115	Seki et al.(2002)a	ADR

BU673293	AK105078	cellulose synthase CesA-1	AF323039	Kreps et al. (2002)	Arabidopsis TF
BI305752	AK111242	xyloglucan endo-transglycosylase	AF163820	Seki et al.(2002)b	Arabidopsis
BU673817	AK104900	alpha tubulin	AJ420858	Kreps et al. (2002)	Arabidopsis TF
BI306079	AK073162	actin depolymerizing factor	AF112887	Kreps et al. (2002)	Arabidopsis TF
BI305650	AK065044	beta-D-glucan exohydrolase, isoenzyme ExoII	U46003	Kreps et al. (2002)	Arabidopsis TF
BI305933	AK068686	cell division protein FtsH-like protein	NM_111112	Kreps et al. (2002)	Arabidopsis TF
		Cell rescue and defence			
BI306369	AK059845	peroxiredoxin Q	AB037598	Kreps et al. (2002)	Arabidopsis TF
BI306058	AK065962	glutaredoxin	D86744	Kreps et al. (2002)	Arabidopsis TF
BI305865	AK066336	peroxidase BP 1	M73234	Kreps et al. (2002)	Arabidopsis TF
BU673649	AK069105	glutathione S-transferase OsGSTZ1	AF309381	Kimura et al.(2003)	Arabidopsis HL
BU673487		glutaredoxin	AP001305	ASRG paralouges	
BU673288	AK069838	ascorbate peroxidase (TL29)	AJ251882	Kimura et al.(2003)	Arabidopsis HL
BI306443	AK069446	catalase	D26484	Kimura et al.(2003)	Arabidopsis HL
BI306573	AK070471	glutathione dependent dehydroascorbate reductase precursor	AF301597	Kimura et al.(2003)	Arabidopsis HL
BI306602	AK103694	glyoxalase I	AB017042	Kawasaki et al. (2001)	
BI305199	AK107087	No hit	D45423	Kawasaki et al. (2001)	SUR
BI306648	AK058509	peroxiredoxin	AF203879	Kreps et al. (2002)	Arabidopsis TF
BU673475	AK069456	peroxidase	AB027752	Seki et al.(2002)a	SUR
BI305325	AK061610	lipoxygenase	L23968	Seki et al. (2002)a	SUR
BI305692	AK062831	LEAF-SPECIFIC THIONIN PRECURSOR	M19046	Kreps et al. (2002)	Arabidopsis TF
BI305703		metallothionein-like protein	U18404	Kawasaki et al. (2001)	SUR
BI305398	AK058313	metallothionein-like protein	AF009959	Kawasaki et al. (2001)	
BI305786	AK106758	thioredoxin h	AB053294	Ozturk et al. (2002)	DDR
CB967248	AK071738	glycolate oxidase	NP_188031	Kreps et al. (2002)	Arabidopsis TF
BI305983	AK067891	thioredoxin-like protein	AB020610	Kreps et al. (2002)	Arabidopsis TF
BI305481	AK058313	metallothionein-like protein	AF001396	Ozturk et al. (2002)	DDR
CB967448	AK103445	metallothionein 2a	S57768	Seki et al.(2002)a	SUR
CB965399	AK073938	phospholipid hydroperoxide glutathione peroxidase	AAB61598	Seki et al.(2002)a	ADR
BI306379	AK104420	peroxidase BP 1	M73234	Seki et al.(2002)b	Arabidopsis
BI305617	AK058529	metallothionein-like protein	U57638	ASRG paralouges	
BI305796	AK062796	RicMT	AB002820	Kreps et al. (2002)	
BU673812	AK059226	glutathione S-transferase GST 13	AF244678	Seki et al.(2002)a	ADR
BI306678		metallothionein-like protein	U57638	ASRG paralouges	
BI305990		peroxidase	X98322	Kreps et al. (2002)	SUR
BU672800	AK069318	Metallothionein-like protein type 2	D89931	Kawasaki et al. (2001)	
BU673476	AK104160	manganese superoxide dismutase	L34039	Kimura et al.(2003)	Arabidopsis HL
BI306411	AK099142	glutathione S-transferase II	AF062403	Seki et al.(2002)a	ADR
CB965073	AK103563	glyoxalase II	AY054407	Kreps et al. (2002)	Arabidopsis TF
BI306659	AK058529	metallothionein-like protein	U77294	Seki et al.(2002)b	Arabidopsis
BI306601	AK099848	monodehydroascorbate reductase	AF109695	Kimura et al.(2003)	Arabidopsis HL
CB965340		thylakoid-bound ascorbate peroxidase	BAC79363	Kreps et al. (2002)	Arabidopsis TF
BI305614	AK059196	thioredoxin h	D26547	Kreps et al. (2002)	Arabidopsis TF
BI306473	AK074018	histone H2A	D38090	Kreps et al. (2002)	Arabidopsis TF
CB965405	AK073938	phospholipid hydroperoxide glutathione peroxidase	AJ270955	Kreps et al. (2002)	Arabidopsis TF
BU672826	AK063148	disease resistance response protein	NM_105259	Kreps et al. (2002)	Arabidopsis TF
BI305468	AK098919	cyclophilin 2 (Cyp2)	L29469	Kawasaki et al. (2001)	SUR
BI306525	AK071076	cyclophilin	AJ132763	Seki et al.(2002)b	Arabidopsis
BU673432	AK105899	wound inducive	AB009885	ASRG paralouges	
BI306376	AK059279	ORF; able to induce HR-like lesions	U66269	ASRG paralouges	
BI305424	AK073415	WOUND-INDUCED BASIC PROTEIN	D30015	ASRG paralouges	
BI306328	AK103599	wound induced protein	X59882	ASRG paralouges	
BI305772	AK058898	CYP18	Y08273	Seki et al.(2002)b	Arabidopsis
BI306276	AK066771	pathogenesis-related protein	U20347	ASRG paralouges	
BI305746	AK072490	cyclophilin CYP5	AF020433	Seki et al.(2002)b	Arabidopsis
CB965601	AK108191	thaumatin-like protein	NP_173261	Kreps et al. (2002)	SUR
BU673319		cyclophilin	AF017993	Seki et al.(2002)a	ADR
BI305750	AK065866	class III chitinase homologue (OsChib3H-h)	AB027426	Seki et al.(2002)b	Arabidopsis
BI306420	AK104699	cyc07	D26060	ASRG paralouges	
CB967125		pathogenesis-related protein 1	AF306651	Kreps et al. (2002)	Arabidopsis TF
BU672976	AK065027	disease resistance response protein	NM_123616	Kreps et al. (2002)	Arabidopsis TF
BI306721	AK063132	blight-associated protein p12 precursor	AF015782	ASRG paralouges	
BI305708	AK062114	chitinase	D55708	Kreps et al. (2002)	Arabidopsis TF
BU673901	AK107139	gt-2	X68261	Kreps et al. (2002)	Arabidopsis TF
BU672803		chitinase-B	AF402939	Kreps et al. (2002)	Arabidopsis TF
BI306024	AK104292	class III chitinase	AF296279	ASRG paralouges	

BU673411		defensin	AB089942	Kreps et al. (2002)	Arabidopsis TF
BU672887	AK059268	Sgt1	AF192467	Kreps et al. (2002)	Arabidopsis TF
BI306073	AK059239	wound induced protein homolog	AB059238	Che et al. (2002)	
BU673702	AK063671	cyclin D2.1 protein	AJ011892	Kreps et al. (2002)	Arabidopsis TF
		Harmone Response			
BI305806	AK068213	auxin-regulated protein (Aux28)	J03919	Seki et al.(2002)a	SUR
CB965518	AK104524	No hit	AAB71969	Kreps et al. (2002)	SUR
BI306388		jasmonate-induced protein	X98124	Ozturk et al. (2002)	DUR
BI305497		ethylene responsive protein (ebp-89)	AJ304840	Kreps et al. (2002)	Arabidopsis TF
BI305642	AK100314	phytochrome-associated protein	AB059238	Kreps et al. (2002)	Arabidopsis TF
BI305739	AK064892	abscisic acid- and stress-inducible protein (Asr1)	AF039573	Kreps et al. (2002)	SUR
BI305913	AK106163	EREBP-like protein	AC079633	Seki et al.(2002)b	Arabidopsis
BI306631	AK071149	Chain A, Inositol Monophosphatase	U39059	ASRG paralouges	
BI306538	AK059964	1-aminocyclopropane-1-carboxylate oxidase	X85747	Kreps et al. (2002)	Arabidopsis TF
BI306077	AK065289	inosine monophosphate dehydrogenase	AJ010201	Kreps et al. (2002)	Arabidopsis TF
BU673267		SR3 sucrose-regulated mRNA, 3'-end sequence	U16257	ASRG paralouges	
BI305323		osERF3	AB036883	Kreps et al. (2002)	SUR
BU673158	AK108268	auxin-regulated protein	NM_148656	ASRG paralouges	
BI305728	AK101835	indole-3-glycerol phosphate synthase	AB023039	Kreps et al. (2002)	Arabidopsis TF
BI306643	AK111282	32 kDa protein jakalin homolog	AF021256	Ozturk et al. (2002)	
BI305758	AK070498	SR3 sucrose-regulated , 3'-end sequence	U16257	ASRG paralouges	
BU673848		gibberellin-20 oxidase (Sd-1)	AF465256	Kreps et al. (2002)	Arabidopsis TF
BU673190	AK059838	IAA1 protein	AJ251791	Kreps et al. (2002)	Arabidopsis TF
		Heat shock proteins			
BU673034	AK060145	copper chaperone	NM_121751	Kreps et al. (2002)	Arabidopsis TF
BI305213	AK101934	heat stress transcription factor Spl7	AB050095	Seki et al.(2002)a	SUR
BI305595	AK064780	heat shock protein 82	Z15018	Kreps et al. (2002)	Arabidopsis TF
BI305618		GrpE protein	AJ010819	ASRG paralouges	
BI306591	AK066658	HMG protein	Y08807	Kawasaki et al. (2001)	SDR
BI306555	AK104352	chaperonin 10	D29698	Seki et al.(2002)b	Arabidopsis
BI306480	AK068506	high mobility group I/Y-2	AF291748	Kreps et al. (2002)	Arabidopsis TF
BI305855	AK105145	hmgc1	AJ131374	Kreps et al. (2002)	Arabidopsis TF
BI306513	AK069617	mitochondrial chaperonin-60	Z12115	Seki et al.(2002)b	Arabidopsis
BI305726	AK074003	low molecular weight heat shock protein precursor (hsp22)	AF035460	Kreps et al. (2002)	Arabidopsis TF
BI306486	AK104048	Luminal binding protein 2 precursor (BiP2)	U08382	Seki et al.(2002)a	SUR
BU673667	AK101334	chaperonin	Z68903	Seki et al.(2002)b	Arabidopsis
BI305606	AK058833	Acyl-CoA-binding protein	A	ASRG paralouges	
BI306560		expressed protein	A	Seki et al.(2002)a	ADR
BI306214	AK098903	chaperonin 21 precursor	AF233745	Kreps et al. (2002)	Arabidopsis TF
BI306548	AK065690	16.9 kDa heat shock protein	AP003250	Kreps et al. (2002)	Arabidopsis TF
BI306208	AK101334	60 KD Chaperonin Beta subuni	Z68903	Seki et al.(2002)a	SUR
		Osmatic prtotectants			
BI306190		reversibly glycosylated polypeptide	Y18624	Seki et al.(2002)b	Arabidopsis
BI305444	AK059665	proline-rich protein	AC091665	Seki et al.(2002)b	Arabidopsis
BI305336	AK068727	ERD1 protein	D17582	Seki et al.(2002)a	SUR
BI305566	AK106205	RD22	AY072821	Kreps et al. (2002)	Arabidopsis TF, HL
BI306254	AK073614	glycine-rich protein	AJ302060	Seki et al.(2002)a	ADR
BI306627	AK104511	RUB1 conjugating enzyme	AF202771	Kreps et al. (2002)	Arabidopsis TF
BI305248		dehydrin	U60097	Ozturk et al. (2002)	SUR
BI305828	AK099576	salt-induced protein,lectin	AC084295	Seki et al.(2002)b	Arabidopsis
BI306033	AK102919	hydroxyproline-rich glycoprotein	X61280	ASRG paralouges	
BI305591	AK100217	proline-rich protein RiP-15	AF221552	ASRG paralouges	
CB966302	AK060981	dormancy-asociated protein	AF467730	Seki et al.(2002)a	ADR
BI305899	AK104006	VIP2 protein	AFA251051	ASRG paralouges	
BI305977	AK106743	rab28 protein	X59138	ASRG paralouges	
CB966763	AK104729	germin-like protein 5	AF032975	Ozturk et al. (2002)	DUR
BU673448	AK099497	submergence induced protein 2A	AF068332	Kimura et al.(2003)	Arabidopsis HL
BU673292	AK065226	dehydrin-like protein	AF314251	Kreps et al. (2002)	SUR
BI306310	AK104600	NDR1/HIN1-like protein	AL589883	Kreps et al. (2002)	Arabidopsis TF
BI306279		senescence-associated protein	AB049723	ASRG paralouges	
BI306250	AK070016	glycine-rich protein (OSGRP1)	AF010579	Kreps et al. (2002)	Arabidopsis TF
BU673470			AF314810	Seki et al.(2002)b	Arabidopsis
BI305659	AK069815	novel protein, osr40c1	X95402	Kawasaki et al. (2001)	SUR,DDR
BI306402	AK101787	annexin p35	X98245	Kreps et al. (2002)	Arabidopsis TF
BI306210	AK098933	reversibly glycosylated polypeptide	Y18624	Kreps et al. (2002)	Arabidopsis TF
BI306424		glycine-rich RNA-binding protein grp1	NM_102006	Kreps et al. (2002)	Arabidopsis TF
CB965087		Hydrophobic protein RCI2B	BAC23051	Kawasaki et al. (2001)	SUR

BI306090	AK073885	cold acclimation protein WCOR410b	U73210	Seki et al.(2002)a	ADR
BI306343	AK104605	glycine-rich protein,	A	Kawasaki et al. (2001)	SDR
BI305981	AK061818	Group 4 late embryogenesis-abundant protein	M88321	Ozturk et al. (2002)	SDR,SUR
BI306107		Similar to DREB1A	AB007787	Seki et al.(2002)a	
		Protein degradation			
BI305677	AK059011	ubiquitin protein fused to a ribosomal protein	D12629	Kreps et al. (2002)	Arabidopsis TF
CB964882	AK065206	aspartic proteinase	D12777	Seki et al.(2002)a	ADR
BI305583	AK064782	ATP-dependent RNA helicase-like protein	AC079022	Kreps et al. (2002)	SUR
BU673736	AK059355	serine carboxypeptidase	AC079632	Kreps et al. (2002)	Arabidopsis TF
BI306152	AK058779	26S proteasome regulatory particle triple- A ATPase subunit2b	AB037154	Seki et al.(2002)b	Arabidopsis
BI306349	AK070376	serine proteinase	AP003106	ASRG paralouges	
BI306614		aspartic proteinase	D12777	Kreps et al. (2002)	Arabidopsis TF
BI305600	AK098858	possible apospory-associated protein	U13149	Kreps et al. (2002)	Arabidopsis TF
BI306103	AK105612	cysteine endopeptidase precursor	AF099203	Kreps et al. (2002)	Arabidopsis TF
BI305258	AK071333	ATP-dependent Clp protease proteolytic subunit	NM_103884	ASRG paralouges	
BI306360	AK066232	E2, ubiquitin-conjugating enzyme	NM_105097	Kawasaki et al. (2001)	SUR
BI305958	AK067838	unknown protein	AL161578	ASRG paralouges	
CB965510		senescence-associated protein	BAB33421	Seki et al.(2002)a	ADR
BI305284	AK060417	metalloprotease	NM_112804	ASRG paralouges	
BI306695	AK062495	subtilisin-chymotrypsin inhibitor 2	Y08625	Kreps et al. (2002)	
BI305490	AK068681	apospory-associated protein C-like	AB013389	Kreps et al. (2002)	Arabidopsis TF
		Protein Kinases and Phosphtases			
BU672886		protein kinase Xa21	U72725	Kreps et al. (2002)	Arabidopsis TF
BU673308	AK070054	nucloside diphosphate kinase 2	AB078008	Kreps et al. (2002)	Arabidopsis TF
BI305201		MAP3K beta 1 protein kinase	AJ010093	Kreps et al. (2002)	Arabidopsis TF
BI305832		adenosine kinase	AJ012281	ASRG paralouges	
BI306416		nucleoside diphosphate kinase	D16292	Seki et al.(2002)b	Arabidopsis
CB966828	AK070255	adenylate kinase	AAN76661	Kreps et al. (2002)	Arabidopsis TF
BI305482	AK069537	PKF1	X97547	Kreps et al. (2002)	Arabidopsis TF
BI305734		phosphoenolpyruvate carboxylase kinase	AF399915	Kreps et al. (2002)	Arabidopsis TF
BU672925		calmodulin-like protein	AC079853	ASRG paralouges	
BI305458	AK072751	nucleoside diphosphate kinase	AF271362	Kawasaki et al. (2001)	SUR
BI306067	AK059786	OsCDPK7	AB042550	Kreps et al. (2002)	Arabidopsis TF
CB967004		protein kinase	NP_193628	Kreps et al. (2002)	Arabidopsis TF
BI305315		serine/threonine kinase	Y12465	Kreps et al. (2002)	SUR
BI306125	AK073725	shaggy-like kinase etha	Y13437	Kimura et al.(2003)	Arabidopsis HL
CB964545		LRK1 protein	AF193835	Kreps et al. (2002)	Arabidopsis TF
BI306130	AK100780	protein kinase, putative	AC027135	Kreps et al. (2002)	SUR
CB966792	AK066295	leucine rich repeat receptor-like kinase	Y07748	Kawasaki et al. (2001)	SDR
BI306714	AK099461	phosphoribulokinase	X51608	ASRG paralouges	
BI305658		serine threonine kinase, putative	NM_099996	Seki et al.(2002)b	Arabidopsis
BU672858		MITOGEN-ACTIVATED PROTEIN KINASE HOMOLOG MMK2	X82268	Seki et al.(2002)b	Arabidopsis
BI305355		choline kinase	U43838	Kimura et al.(2003)	Arabidopsis HL
BI305279	AK068315	calcium-dependent protein kinase	AL133248	Kreps et al. (2002)	Arabidopsis TF
CB966430		mitogen-activated protein kinase	AJ535841	Kreps et al. (2002)	Arabidopsis TF
BU673193		protein kinase	AL356014	Seki et al.(2002)b	Arabidopsis
BI305334	AK070065	GmCK2p	U43839	Kreps et al. (2002)	Arabidopsis TF
BU673337	AK102065	uracil phosphoribosyltransferase 1	AF116860	Kreps et al. (2002)	Arabidopsis TF
CB966116	AK106842	protein kinase family	NP_175879	Kreps et al. (2002)	Arabidopsis TF
BI305304	AK066495	OsCDPK7	AB042550	Kawasaki et al. (2001)	SUR
CB967253	AK106112	acid phosphatase-like protein	CAB71336	Kreps et al. (2002)	Arabidopsis TF
BI306580	AK060035	protein phosphatase	AY065090	ASRG paralouges	
BU672963	AK071996	protein phosphatase 2C	AP003251	Kreps et al. (2002)	Arabidopsis TF
CB964933	AK061618	protein phosphatase 2C-like protein	NP_195564	Kreps et al. (2002)	Arabidopsis TF
BI306327	AK072534	Protein phosphatase 2C-like protein	NM_122403	Ozturk et al. (2002)	DUR, SUR
BI305224		contains similarity to protein phosphatase-2c~	AB022217	Seki et al.(2002)b	Arabidopsis
BI306474	AK060314	Iron(III)-zinc(II) purple acid phosphatase precursor	AJ006224	Kreps et al. (2002)	Arabidopsis TF
		Photosynthesis			
CB965038	AK098864	chlorophyll a/b-binding protein precursor	AF094776	Seki et al.(2002)a	ADR
BI305821	AK058284	photosystem II subunit (22KDa) precursor	AP003235	Seki et al.(2002)a	ADR
BI306390	AK105055	photosystem II 10 kDa polypeptide	U86018	Ozturk et al. (2002)	SUR
BI305816		photosystem I chain IV precursor	Y00966	Kreps et al. (2002)	Arabidopsis TF
BI305843	AK060920	triosephosphate isomerase (Rictipi2)	L04967	ASRG paralouges	
BI305422	AK104176	chlorophyll a/b-binding protein	U74295	Seki et al.(2002)a	ADR
BI306046	AK061611	ribulose bisphosphate carboxylase/oxygenase	L22155	Ozturk et al. (2002)	DDR

BI306054	AK104760	33kDa oxygen evolving protein of photosystem II	X57408	Seki et al.(2002)a	ADR
BI306021	AK103924	chlorophyll a/b-binding protein	AF058796	Kreps et al. (2002)	Arabidopsis TF
BI305735	AK104722	23 kDa polypeptide of photosystem II	AF052203	Seki et al.(2002)a	ADR
BI306036	AK104332	RuBisCO activase large isoform precursor	AB034698	Ozturk et al. (2002)	DDR
BI305945	AK098931	disulfide isomerase A6 precursor (P5)	A	ASRG paralouges	
BI305663		chloroplast apocytochrome b6 (petB)	M35995	ASRG paralouges	
BU672861	AK099127	photosystem II oxygen-evolving complex protein 3 precursor	M87435	Seki et al.(2002)a	ADR
BI305564	AK104824	CP26 , partial sequence	D85512	Ozturk et al. (2002)	DDR
BI305570	AK061410	Rubisco subunit binding-protein alpha subunit	X07851	Kreps et al. (2002)	Arabidopsis TF
BI305397	AK058788	photosystem I PSI-K subunit	L12707	Kreps et al. (2002)	Arabidopsis TF
BI305445	AK059247	3-deoxy-D-arabino-heptulosonate 7-phosphate synthase	Y14797	Kreps et al. (2002)	Arabidopsis TF
BI306262		photosystem II 10 kDa polypeptide	U86018	Seki et al.(2002)b	Arabidopsis
BU673610	AK104260	phytocyanin	T06555	Ozturk et al. (2002)	DUR
BI305551	AK059037	P49107	X66428	Seki et al.(2002)a	
BI305247		chloroplast RNA helicase VDL isoform 1	AF261020	Kreps et al. (2002)	SUR
BU672866	AK071899	chlorophyll synthase	AC087599	ASRG paralouges	
BI305514	AK099109	photosystem I protein (PSI-L)	M61146	Seki et al.(2002)a	ADR
BU673575	AK105600	ribulose 1,5-bisphosphate carboxylase/oxygenase large subunit	AB084766	ASRG paralouges	
BI305454		ribulose 1,5-bisphosphate carboxylase small subunit	D00644	ASRG paralouges	
BI306166	AK070447	plastocyanin precursor	Y00704	Seki et al.(2002)a	ADR
BI305574	AK069860	ferredoxin	AB001386	Kreps et al. (2002)	Arabidopsis TF
BI306736	AK103503	photosystem II D1 protein	D21291	ASRG paralouges	
BI305448	AK068273	manganese-binding protein PsbY precursor, photosystem II-associated	AF060198	Seki et al.(2002)a	ADR
BI306493		photosystem II 10 kDa polypeptide	U86018	Seki et al.(2002)b	Arabidopsis
BI305598	AK068555	small subunit of ribulose-1,5-bisphosphate carboxylase	AF017364	ASRG paralouges	
		Protein Synthesis			
CB967086	AK063879	RSZp22 splicing factor	T52628	Kreps et al. (2002)	Arabidopsis TF
BI306583	AK060429	RNA binding protein	NM_121073	ASRG paralouges	
CB964819		60S RIBOSOMAL PROTEIN L9 induced by GA3	D83527	Kreps et al. (2002)	Arabidopsis TF
CB965467	AK058881	S18.A ribosomal protein	AF370463	Kreps et al. (2002)	Arabidopsis TF
BI305521	AK102423	ribosomal protein S4	Y15009	Kawasaki et al. (2001)	SUR
CB967158	AK058918	60S RIBOSOMAL PROTEIN L36	AL132960	Kreps et al. (2002)	Arabidopsis TF
BI305675	AK103440	rpS28	AJ001161	Kreps et al. (2002)	Arabidopsis TF
BI306740	AK110621	60S ribosomal protein L17	AY054508	Kreps et al. (2002)	Arabidopsis TF
CB965342		ribosomal protein L25	AF061508	Kreps et al. (2002)	Arabidopsis TF
~~BI305765~~		60S ribosomal protein L27a	AB042856	ASRG paralouges	
BU673172	AK073196	elongation factor 1 alpha	AF136826	Kreps et al. (2002)	Arabidopsis TF
BU673030	AK098982	ribosomal protein S31	D38011	Kreps et al. (2002)	Arabidopsis TF
CB966416	AK103424	60S RIBOSOMAL PROTEIN L34	S60476	Kreps et al. (2002)	Arabidopsis TF
BI305964	AK101927	L24 ribosomal protein	X94296	Kreps et al. (2002)	Arabidopsis TF
BI306038	AK103497	translation initiation factor 3, subunit g (eIF3g)	AJ293728	Seki et al.(2002)b	Arabidopsis
BI306106	AK071736	elongation factor 1 beta,	D23674	ASRG paralouges	
BI305919	AK073640	translation initiation factor 4A	X61206	Kreps et al. (2002)	Arabidopsis TF
BI306741	AK069751	50S ribosomal protein L4	AY072419	Kreps et al. (2002)	Arabidopsis TF
BI305946	AK060387	translation initiation factor, eIF-5A	AJ312906	ASRG paralouges	
BI306117	AK103035	elongation factor EF-2	AY054461	ASRG paralouges	
CB967045		60S ribosomal protein L31-1	AAA80638	Kreps et al. (2002)	Arabidopsis TF
CB967196	AK058262	ribosomal protein	U86017	Kreps et al. (2002)	Arabidopsis TF
BI305450	AK105037	translation initiation factor (GOS2)	AF094774	Kreps et al. (2002)	Arabidopsis TF
BU672831		cap-binding protein p28	AC022457	Kreps et al. (2002)	Arabidopsis TF
CB965656		ribosomal protein L27	AF400191	Kreps et al. (2002)	Arabidopsis TF
CB965835	AK106339	S-ribonuclease binding protein SBP1	AF223395	Kreps et al. (2002)	Arabidopsis TF
BU672810		RNA-binding glycine rich protein	D26182	Kreps et al. (2002)	Arabidopsis TF
BI305215	AK071725	plastid RNA polymerase sigma factor	AB005290	Kimura et al.(2003)	Arabidopsis HL
BI306312	AK059802	ribonuclease	AB052842	Kreps et al. (2002)	SUR
CB967287	AK062184	40S subunit ribosomal protein	D12632	Kreps et al. (2002)	Arabidopsis TF
CB964857	AK099225	ribosomal protein	S38359	Kreps et al. (2002)	Arabidopsis TF
BU672912	AK104792	poly(A)-binding protein	U81318	Kreps et al. (2002)	Arabidopsis TF
BI306102	AK072648	EF-1 alpha	D63580	Kawasaki et al. (2001)	SUR
		Secondary metabolism			
BI306467	AK099001	strictosidine synthase-like	AL589883	Seki et al.(2002)a	SUR
BI305713		UDP-glucosyltransferase	AF515727	Kreps et al. (2002)	Arabidopsis TF
BI306691	AK107665	cinnamoyl CoA reductase	AJ428493	Seki et al.(2002)a	SUR
CB965041	AK061390	chalcone isomerase	AF474922	Kreps et al. (2002)	Arabidopsis TF

CB965712	AK067801	GP28	Z15085	Kimura et al.(2003)	Arabidopsis HL
BU673749	AK059518	putative anthocyanidin reductase	CAD91910	Seki et al.(2002)a	SUR
BU672830	AK104801	Caffeoyl CoA O-methyltransferase	AJ242980	Kimura et al.(2003)	Arabidopsis HL
CB966105	AK067097	anthranilate phosphoribosyltransferase-like protein	CAB67616	Kreps et al. (2002)	Arabidopsis TF
BI305578	AK071763	gamma-tocopherol methyltransferase	AF213481	ASRG paralouges	
BI305613	AK103001	4-coumarate-CoA ligas	X52623	Kreps et al. (2002)	Arabidopsis TF
BI305589	AK058306	phenylalanine ammonia-lyase	X16099	Seki et al.(2002)b	Arabidopsis
BU673855	AK064768	caffeic acid 3-O-Methyltransferase	AJ231133	Kimura et al.(2003)	Arabidopsis HL
		Signal Transduction			
BI305802	AK066784	small GTP-binding protein (rab5A)	AY029301	Seki et al.(2002)b	Arabidopsis
BI305436	AK100331	diacylglycerol kinase	D63787	ASRG paralouges	
BI305605	AK064875	vesicle soluble NSF attachment protein receptor	AC082644	Kreps et al. (2002)	Arabidopsis TF
BI305297		signal peptidase I	AE004511	ASRG paralouges	
BI306060	AK070230	GF14-c protein	U65957	Kreps et al. (2002)	Arabidopsis TF
BU673351		signal recognition particle	U85037	ASRG paralouges	
BI306556	AK103090	signal recognition particle receptor alpha	AF360125	Kreps et al. (2002)	Arabidopsis TF
BI305572	AK067504	small GTP binding protein RACDP (RACD)	AF218381	Kreps et al. (2002)	Arabidopsis TF
BI306675		ras-related GTP binding protein	S66160	Kreps et al. (2002)	Arabidopsis TF
BI306329	AK101721	response regulator 5	AB042267	Kreps et al. (2002)	Arabidopsis TF
BU672863	AK071529	GTP-BINDING PROTEIN YPTV3	AC079843	ASRG paralouges	
BI306475	AK070090	calmodulin (CaM1)	AF042840	Kawasaki et al. (2001)	SUR
BI305502	AK103065	brain specific protein	D16140	Kreps et al. (2002)	Arabidopsis TF
BI305667	AK069903	GTP-binding protein	AJ307662	Seki et al.(2002)b	Arabidopsis
CB967365	AK067670	light-regulated protein	X68807	Kimura et al.(2003)	Arabidopsis HL
BI306652		EF-hand Ca2+-binding protein CCD1	AF181661	Kawasaki et al. (2001)	
BI305778		photoreceptor-interacting protein-like	AB013389	Kreps et al. (2002)	Arabidopsis TF
BU673354	AK101599	14-3-3 protein homolog GF14-12	M96856	Seki et al.(2002)b	Arabidopsis
BU673088		14-3-3 protein	AY029473	Seki et al.(2002)b	Arabidopsis
BI306395	AK062179	GUANINE NUCLEOTIDE-BINDING PROTEIN BETA SUBUNIT	D38231	Kreps et al. (2002)	Arabidopsis TF
BI305487	AK068609	Ras-like GTP-binding protein	NM_111825.1|	Seki et al.(2002)b	Arabidopsis
BI305561	AK061099	small GTP-binding protein (ORRab-2)	L35845	Kreps et al. (2002)	Arabidopsis TF
BI305552	AK072731	small GTP-binding protein (Ran1)	AB015971	Kreps et al. (2002)	Arabidopsis TF
BI305837	AK059970	small GTP-binding protein OsRac3	AB029510	Seki et al.(2002)b	Arabidopsis
		Transcription Factors			
BU672783		transcription factor Hap5a-like protein	AY072402	Kreps et al. (2002)	Arabidopsis TF
BU673773	AK109716	bZIP transcription factor PF|00170 domain.	AC009917	Kreps et al. (2002)	Arabidopsis TF
CB966397	AK063621	small nuclear ribonucleoprotein	AL161576	Kreps et al. (2002)	Arabidopsis TF
BI305762	AK065182	transcription factor IIA small subunit	AC079022	ASRG paralouges	
CB967092		one helix protein	AAM22751	Kreps et al. (2002)	Arabidopsis TF
CB967424	AK099264	Tat binding protein	D17789	Seki et al.(2002)a	SUR
BI306078		RING-H2 finger protein RHA1a	AF078683	Kreps et al. (2002)	SUR
BI306362		zinc finger protein	AC079281	Kreps et al. (2002)	Arabidopsis TF
BI306209	AK073342	RING finger protein	Y09539	Kreps et al. (2002)	
BI306221		RING zinc finger protein	AC015985	ASRG paralouges	
BI305731	AK059172	RNA binding protein	AC026815	Kreps et al. (2002)	Arabidopsis TF
BU672792	AK058899	DRE binding factor 1	AF493800	Kreps et al. (2002)	Arabidopsis TF
BI306418	AK099223	R2R3MYB-domain protein	AF099432	Seki et al.(2002)a	SUR
BI305874		AP2 domain protein homolog	U77655	Seki et al.(2002)b	Arabidopsis
CB965937	AK059146	RING-H2 finger protein RHA2a	NP_172962	Seki et al.(2002)a	SUR
BI305268	AK102284	helicase-like transcription factor	AP003224	Seki et al.(2002)b	Arabidopsis
BU672942		small zinc finger-like protein (TIM9)	AF150113	ASRG paralouges	
BI305238		leucine-responsive regulatory protein	AL627268	ASRG paralouges	
BI305249	AK067456	C3HC4-type RING zinc finger protein	AC004238	Kreps et al. (2002)	Arabidopsis TF
BI306249		transcription factor BTF3	AY224531	ASRG paralouges	
CB965488	AK065107	nuclear RNA binding protein A	AF110228	Kreps et al. (2002)	Arabidopsis TF
BU673061	AK061664	zinc finger protein	AB001883	Seki et al.(2002)a	SUR
BI305764	AK068762	zinc finger protein	AF466199	Kreps et al. (2002)	Arabidopsis TF
BU673053		zinc finger-like protein	AY086347	Seki et al.(2002)a	SUR
BI305935	AK069926	DNA binding protein, putative	NM_103643	Kreps et al. (2002)	Arabidopsis TF
BI306302		RNA-binding protein	NM_128511	Kreps et al. (2002)	Arabidopsis TF
BU673870		HOS59	AB007628	Kreps et al. (2002)	Arabidopsis TF
BU673703		dnaJ-like protein	NM_115594	Kreps et al. (2002)	Arabidopsis TF
BI306687		zinc-finger-like protein	X60700	Kreps et al. (2002)	Arabidopsis TF
BI305963	AK105666	nucleoid DNA-binding protein cnd41, chloroplast	D26015	Kreps et al. (2002)	Arabidopsis TF
BU673228	AK062422	DRE-binding protein 1B	AF300972	Kreps et al. (2002)	SUR
BI305518	AK071730	transcriptional regulator	NM_118306	Kreps et al. (2002)	Arabidopsis TF
BI306059	AK066834	OSMYB1	D88617	Kreps et al. (2002)	Arabidopsis TF
BI305523	AK070743	nucleic acid-binding protein	AJ224324	Seki et al.(2002)a	ADR

BI305625		zinc finger protein, putative, 5' partial	AC069474	Kreps et al. (2002)	Arabidopsis TF
BI305955	AK102849	similar to RING-H2 finger protein RHA1a	AF078683	Kreps et al. (2002)	Arabidopsis TF
BU673704	AK106818	homeodomain leucine zipper protein	AF145728	Seki et al.(2002)a	SUR
CB965631		CONSTANS family zinc finger protein	AC069471	Kreps et al. (2002)	SUR
BI305714	AK109719	TGA-type basic leucine zipper protein	AF402608	Kreps et al. (2002)	Arabidopsis TF
BI305867	AK063880	RING3-like bromodomain protein	AC004238	ASRG paralouges	
CB967252	AK072938	WRKY family transcription factor	NP_849559	Seki et al.(2002)a	SUR
BU673758	AK063399	OsNAC5 protein	AB028184	Kreps et al. (2002)	Arabidopsis TF
CB965560	AK105720	transcription factor GT-3b	AAP13348	Kreps et al. (2002)	Arabidopsis TF
BI305705	AK061000	Dof zinc finger protein	AB028132	Kreps et al. (2002)	Arabidopsis TF
CB967107	AK059311	nucleic acid binding protein	T02745	Kreps et al. (2002)	Arabidopsis TF
BU673410	AK101803	p53 binding protein	AY072931	Ozturk et al. (2002)	SDR
CB966697	AK062882	AP2 domain transcription factor	NP_195167	Kreps et al. (2002)	SUR
BU673403	AK104280	RING finger-like protein	AL132971	ASRG paralouges	
		Transport			
CB964997	AK066019	H+-transporting ATP synthase chain 9	AY224460	Seki et al.(2002)a	ADR
BI305505	AK104123	gamma-Tip	D25534	Seki et al.(2002)a	ADR
BI306491	AK073967	integral membrane protein	NM_105398	Kreps et al. (2002)	Arabidopsis TF
BI306213	AK103970	water channel protein	AB058678	Kreps et al. (2002)	SUR
BI306589	AK103611	glucose 6 phosphate/phosphate translocator, putative	NM_102035	Kreps et al. (2002)	Arabidopsis TF
BI306649	AK061443	membrane protein	L13655	Kawasaki et al. (2001)	
BU673507	AK059032	vacuolar membrane ATPase subunit G	AF181688	Kreps et al. (2002)	Arabidopsis TF
BU673226		amino acid permease	AB022783	Kreps et al. (2002)	Arabidopsis TF
BU673203	AK099203	ABC transporter family protein	NM_100271	Kreps et al. (2002)	SUR
BI305690	AK104464	delta-type tonoplast intrinsic protein	U86763	Seki et al.(2002)b	Arabidopsis
BI305682	AK062070	Sec61 alpha subunit	AY044237	ASRG paralouges	
BI305534	AK063469	transporter-like protein	AL356014	ASRG paralouges	
BI305835	AK066933	vacuolar H+-pyrophosphatase	D45384	ASRG paralouges	
CB965602	AK102640	hexose transporter	AJ534445	Kreps et al. (2002)	SUR
BU673450	AK065141	nuclear transport factor 2	AB011262	ASRG paralouges	
CB967178	AK062513	plasma membrane associated protein	NP_197398	Ozturk et al. (2002)	DUR
BI305712	AK101693	amino acid selective channel protein	AJ011921	Kreps et al. (2002)	Arabidopsis TF
BI306571	AK100143	ATP/ADP translocator	D12637	Kreps et al. (2002)	Arabidopsis TF
BI305433	AK106868 ASP	Scl1 protein	AF149807	Kreps et al. (2002)	Arabidopsis TF
BI305586		carnitine/acylcarnitine translocase	AC060755	Kreps et al. (2002)	Arabidopsis TF
BI306260	AK069611	mitochondrial phosphate transporter	AB016065	ASRG paralouges	
BI305222	AK063891	ABC transporter	AC069158	Kreps et al. (2002)	Arabidopsis TF
BI306124	AK071200	VIP2 protein	AJ251051	ASRG paralouges	
BI305209	AK071314	amino acid permease, putative	AC079041	ASRG paralouges	
BI306339		protein transport protein subunit	AB011483	Kreps et al. (2002)	Arabidopsis TF
BI306282	AK068840	oligopeptide transporter	AP003235	Seki et al.(2002)a	SUR
BU673012	AK061111	histidine-containing phosphotransfer protein	AB024293	Kreps et al. (2002)	Arabidopsis TF
BI306386	AK105044	vacuolar H+-ATPase (vatp-P1)	U27098	Kreps et al. (2002)	Arabidopsis TF
BU672768	AK068806	major intrinsic protein	D17443	Kreps et al. (2002)	SUR
BI305870	AK102155	plasma membrane major intrinsic protein 2	U60148	Seki et al.(2002)b	Arabidopsis
BI305853	AK072064	mitochondrial carrier protein	AC002535	Kreps et al. (2002)	Arabidopsis TF
BI306522	AK099206	aminotransferase 1	AY066012	Kreps et al. (2002)	Arabidopsis TF
BI306141	AK066573	membrane-associated protein	AC068923	ASRG paralouges	
BI306313		oligopeptide transporter	AP003235	Kreps et al. (2002)	Arabidopsis TF
BI305923	AK059084	translocation complex Sec61gamma chain	AY059131	Kreps et al. (2002)	Arabidopsis TF
		Other and Unclassified			
CB967012	AK059231	pollen allergen-like protein	NP_173813	Kreps et al. (2002)	Arabidopsis TF
CB965572		CG6214 product (MDR)	XM_079251	Kreps et al. (2002)	Arabidopsis TF
BI305275	AK104308	root border cell-specific protein	AF139187	ASRG paralouges	
CB966764	AK104429	apyrase	NP_187058	Kreps et al. (2002)	Arabidopsis TF
BI306163	AK072428	early nodulin	AB018377	ASRG paralouges	
BI306519		pumilio/Mpt5 family RNA-binding protein	NM_128471	Kreps et al. (2002)	Arabidopsis TF
BI305683		HSP90-like protein	AY077617	Kreps et al. (2002)	SUR
BI305624	AK072275	NAM (no apical meristem)-	A	Seki et al.(2002)a	SUR
BI305699	AK106182	remorin 1	AF123265	Seki et al.(2002)a	SUR
BU673516	AK104766	NAM-like protein	AY086093	Seki et al.(2002)a	SUR
CB964440		nonphototrophic hypocotyl 1b	AB018443	Seki et al.(2002)a	SUR
BI306089		aconitase-iron regulated protein 1 (IRP1)	AF073507	Kreps et al. (2002)	Arabidopsis TF
BI306066		minor allergen	Z99708	Kreps et al. (2002)	Arabidopsis TF
BI305256	AK101642	vegetative storage protein	L20233	Kreps et al. (2002)	Arabidopsis TF
CB966827		dermal glycoprotein precursor	BAB89709	Kreps et al. (2002)	Arabidopsis TF
BI305947	AK072166	gigantea-like protein	AJ133787	Kawasaki et al. (2001)	SUR
CB965002	AK098887	nifU-like protein	NM_118347	Kreps et al. (2002)	Arabidopsis TF

BI305817	AK067728	timing of CAB expression 1-like protein	AF272040	Kreps et al. (2002)	SUR
CB964540		Argonaute protein	NP_849784	Kreps et al. (2002)	Arabidopsis TF
BI305721	AK062573	symbiosis-related protein	NM_104884	Kreps et al. (2002)	Arabidopsis TF
CB967400	AK058786	unknown protein	AF336926	Kreps et al. (2002)	Arabidopsis TF
BU673869		expressed protein	NM_115079	Kreps et al. (2002)	Arabidopsis TF
CB967407	AK104712	putative protein		Ozturk et al. (2002)	DUR
CB967051	AK064894	MutT/nudix family protein	AC012396	Kreps et al. (2002)	Arabidopsis TF
CB966912	AK058684	expressed protein	NP_565300	Kreps et al. (2002)	SUR
CB964493	AK060099	expressed protein	NP_568659	Kimura et al.(2003)	Arabidopsis HL
CB967465		No hit		Kreps et al. (2002)	SUR
CB965319		hypothetical protein		Seki et al.(2002)a	SUR
CB964951	AK065178	expressed protein	NM_127785	Kreps et al. (2002)	Arabidopsis TF
BU673291	AK058243	expressed protein	NM_122164	Seki et al.(2002)a	ADR
CB967415		hypothetical protein	AAO16698	Kreps et al. (2002)	Arabidopsis TF
CB967222	AK066597	hypothetical protein	NP_565685	Kreps et al. (2002)	Arabidopsis TF
CB967195	AK061214	unknown protein	AAM62634	Seki et al.(2002)a	SUR
CB966814	AK100804	hypothetical protein	AAM67097	Kreps et al. (2002)	Arabidopsis TF
BU673383	AK102404	Unknown protein	AY054590	Kreps et al. (2002)	SUR
BI305245	AK068177	unknown protein	AC004667	Kreps et al. (2002)	Arabidopsis TF
BI305350	AK070437	No hit		Kawasaki et al. (2001)	SUR
BI306476	AK110485	putative protein	NM_120678	Seki et al.(2002)a	SUR
BI305327	AK067791	hypothetical protein	T04607	Kreps et al. (2002)	Arabidopsis TF
BI306484	AK101293	expressed protein	NM_100442	Kreps et al. (2002)	Arabidopsis TF
CB966274	AK071246	expressed protein	NM_111122	Kreps et al. (2002)	SUR
BI306541	AK103977	putative protein	NM_125047	Kreps et al. (2002)	Arabidopsis TF
BU673730	AK107758	unknown protein	NM_126499	Kreps et al. (2002)	Arabidopsis TF
CB967177	AK060823	hypothetical protein	AAC78508	Kimura et al.(2003)	Arabidopsis HL
CB966314	AK068566	expressed protein	NM_106153	Kreps et al. (2002)	Arabidopsis TF
BU672967	AK058980	putative protein	NM_116040	Kreps et al. (2002)	Arabidopsis TF
BU672832	AK069642	No hit		Kreps et al. (2002)	Arabidopsis TF
BI306384		Expressed protein	NM_119249	Kreps et al. (2002)	Arabidopsis TF
BI306623		putative protein	NM_114287	Kreps et al. (2002)	Arabidopsis TF
BI306338	AK106979	hypothetical protein	AJ271079	Ozturk et al. (2002)	DUR
BI306324	AK072161	unknown protein	AP003213	Seki et al.(2002)a	ADR
BI306238	AK105690	putative protein	NM_124534	Kreps et al. (2002)	Arabidopsis TF
CB965390		expressed protein	C96587	Kreps et al. (2002)	Arabidopsis TF
CB965440	AK067058	unknown	AY085105	Kreps et al. (2002)	Arabidopsis TF
CB964980		fiber protein Fb14	AAP34361	Kreps et al. (2002)	Arabidopsis TF
BU672931		putative protein	NM_124804	Kreps et al. (2002)	Arabidopsis TF
BU673063		No hit		ASRG paralouges	
CB964679	AK100542	No hit		Kreps et al. (2002)	Arabidopsis TF
CB964692	AK071397	expressed protein	NP_191814	Kreps et al. (2002)	Arabidopsis TF
BU673135	AK067043	unknown cold induced protein	AY090535	ASRG paralouges	
CB967406	AK070095	expressed protein	NM_129622	Kreps et al. (2002)	Arabidopsis TF
CB965039	AK105948	unknown protein	AAP46641	Kreps et al. (2002)	Arabidopsis TF
BI305429		putative protein	AL391143	Kreps et al. (2002)	Arabidopsis TF
CB965568	AK071818	hypothetical protein	NM_117949	Kreps et al. (2002)	Arabidopsis TF
CB965473		expressed protein	NP_680193	Kreps et al. (2002)	Arabidopsis TF
BI305915	AK060419	expressed protein		Kreps et al. (2002)	Arabidopsis TF
BU673735		Expressed protein	AF303134	Seki et al.(2002)a	SUR
CB964922	AK106819	expressed protein	NP_567894	Seki et al.(2002)a	ADR
BU673285	AK099407	unknown protein	AC013258	Kreps et al. (2002)	Arabidopsis TF
BI306670	AK067677	expressed protein	NM_101999	Seki et al.(2002)a	ADR
BU672829	AK059561	unknown protein	AY050865	Kreps et al. (2002)	Arabidopsis TF
BU673620	AK104818	unknown	AY086010	Kreps et al. (2002)	Arabidopsis TF
CB964499	AK066210	expressed protein	NM_126585	Kreps et al. (2002)	Arabidopsis TF
CB964814	AK065860	unkown protein	NM_100379	Kreps et al. (2002)	Arabidopsis TF
CB967067	AK073428	amino acid transporter family	NP_566854	Kreps et al. (2002)	Arabidopsis TF
CB964865	AK073157	Expressed protein	NP_567805	Kreps et al. (2002)	Arabidopsis TF
BI306017	AK106330	heavy-metal-associated domain-containing protein	A	Kreps et al. (2002)	Arabidopsis TF
CB965957	AK066157	expressed protein	AY136359	Kreps et al. (2002)	Arabidopsis TF
CB966839	AK062655	No hit		Kreps et al. (2002)	SUR
CB965122	AK062405	No hit		Ozturk et al. (2002)	SUR
BI305633	AK063923	unknown protein	AC004521	Kreps et al. (2002)	Arabidopsis TF
CB965143	AK065832	expressed protein	NP_200876	Kreps et al. (2002)	Arabidopsis TF
CB965868	AK062628	unknown protein	AF360123_1	Seki et al.(2002)a	ADR
CB965454	AK100277	expressed protein	NP_190772	Kreps et al. (2002)	Arabidopsis TF
BI305999	AK058271	expressed protein	NM_116345	Seki et al.(2002)a	ADR
BU673062		expressed protein	NM_129142	Kreps et al. (2002)	Arabidopsis TF
CB965947		No hit		Kreps et al. (2002)	Arabidopsis TF

CB965933	AK099963	putative protein	AAL33781	Kreps et al. (2002)	Arabidopsis TF
BI305760	AK066261	unknown protein	AP003076	Kreps et al. (2002)	Arabidopsis TF
CB965629	AK104682	GDSL-motif lipase/hydrolase protein	NP_199379	Seki et al.(2002)a	ADR
BU673070	AK073839	unknown protein	AB016886	Kreps et al. (2002)	Arabidopsis TF
CB964647		unknown protein	NM_102495	Kreps et al. (2002)	Arabidopsis TF
BU672770	AK100371	No hit		Kawasaki et al. (2001)	SDR

4.8 Analysis of gene organization of stress responsive genes

Gene organization analysis by aligning the cDNA sequences onto the genomic sequences revealed genes with single exon to seventeen exons (Table 4.5). Single exon genes were mainly those having small coding sequence and organellar gene products, which are integrated into the nuclear genome. The observed exon size of minimum length 21 bp, maximum length 1600 bp and intron sequence lengths of minimum 37 bp and maximum 2000 bp reflect the complexity of gene organization and may indicate a highly regulated pattern of gene expression during transcription.

The compact nature of plant gene promoters in which <1000 bp of promoter sequence often is sufficient to drive proper regulated patterns of transcription and regulatory elements tend to be clustered near the transcription start site. Promoter sequences of 130 stress responsive genes were analyzed for the known motifs associated with stress response. Of these, eighty of them showed *cis-* acting elements in the promoter region associated with abiotic stress (Table 5.1).

Table 4.5: Gene organization of stress responsive genes

S. NO	Putative Function	GenBank Acc	No of Exons	Exon Size	Intron Size
1.	Lipid transfer protein	Z23271	1	1495	0
2.	Mitochondrial Ribosomal protein s10,	AB035348	1	327	0
3.	Dof zinc finger protein	AB028132	1	1305	0
4.	Ferredoxin	AF010320	1	829	0
5.	Apocytochrome b6	M35995	1	648	0
6.	Root-specific rcc3	BI305683	1	773	0
7.	Cyclophilin 2	BI305468	1	824	0
8.	Chloroplast atpb ATP synthase beta subunit	AB037543	1	1543	0
9.	Cytochrome P450-like sequence	AF088221	1	769	0
10.	Early nodulin	AB018377	1	617	0
11.	Heat shock protein 82	BI305595	1	665	0
12.	Amylogenin	Y18623	1	1151	0
13.	Plastocyanin precursor	AF093636	1	676	0
14.	Chlorophyll a/b binding protein of photosystem II	D00641	1	1022	0
15.	DREB1A	AB007787	1	645	0
16.	4-coumarate-coa ligase	X52623	1	1058	0

17.	Mitochondrion atp9 gene for Fo-atpase subunit 9	X16936	1	475	
18.	Nuclear transport factor 2	AB011262	2	360,301	1410
19.	Calmodulin (cam2)	AF042839	2	166,484	1443
20.	Glycine-rich protein	AF010579	2	121,568	248
21.	Glycine rich RNA binding protein	AJ302060	2	215,653,	109
22.	Abscisic acid- and stress-inducible protein	AF039573	2	301,512	119
23.	Metallothionein-like protein	AF017366	2	127,406	141
24.	Putative tensile reaction protein (WAR1)	AY072932	2	51, 531	105
25.	Ribulose 1,5-bisphosphate carboxylase small subunit	D00644	2	213, 684	111
26.	Putative DRE-binding 1B protein	AY166833	2	383, 484	105
27.	OSMYB1	D88617	2	388, 791	100
28.	S-adenosylmethionine synthetase	AJ296743	2	97,1483	1325
29.	IAI2 wound induced protein homolog,	AB059238	2	171, 501	192
30.	Cytochrome c oxidase subunit 5c	AB027123	2	62, 445	2385
31.	Ribonuclease	AB052842	2	328, 837	961
32.	Lipid transfer protein	BU673284	2	216, 225	90
33.	Chlorophyll a/b binding protein	AF061577	2	179, 821	89
34.	Ribulose bisphosphate carboxylase/ oxygenase	L22155	2	182, 625	103
35.	Gamma-tip	D25534	2	211,864	754
36.	Succinate dehydrogenase subunit 3	AF362741	2	64, 326	679
37.	Lipid transfer protein LPT III	AF017360	2	412, 374	138
38.	RIP1 ribosome inactivating protein 1	AB051107	2	130, 1168	141
39.	Vacuolar H+-atpase (vatp-P1)	U27098	3	191,286,426	314,946
40.	Glutathione S-transferase II	AF062403	3		
41.	Metallothionein-like protein	BI305617	3	156,75,428	109,115
42.	Beta tubulin	X78142	3	220, 270, 952	713, 534
43.	Cytochrome B5	X75670	3	241,67,456	950,119
44.	Rice MT	AB002820	3	138,66,295	64,110
45.	EF-1 alpha	BI306125, D63583	3	53,484,1062	659,595
46.	EF-1 alpha	D63581	3	43,484, 1109	532, 702
47.	Thioredoxin	AB053294	3	226,123,338	1108,110
48.	Thioredoxin h	D21836	3	146,123,418	1100,81
49.	Ribosomal protein	U86017	3	52, 184, 267	128, 873
50.	Peroxiredoxin	AF203879	3	194,80,352	1004, 287
51.	23 kda polypeptide of photosystem II	AF052203	3	537, 51, 502	100, 246
52.	Aquaporin (PIP2a)	AF062393	3	715,141,426	1886, 124
53.	Beta-expansin (EXPB7)	AF261275	3	494, 185, 522	223, 324
54.	Positive element factor 1 (PF1)	L24390	3	115, 186, 653	78, 566
55.	Osnac5 protein	AB028184	3	251, 296, 1165	89, 818
56.	S-adenosylmethionine decarboxylase 2	A251899	3	200,157,1614	961,91
57.	Gt-2 gene	X68261	3	458, 1561, 931	894, 112
58.	Metallothionein-like protein	AF001396	3	101, 48,375	139,501
59.	Hydroxyproline-rich glycoprotein	X61280	3	430, 299,259	37,62,
60.	Cytochrome b5	X75670	3	231, 67, 455	950, 119

61.	Thionin Osthi1,	AB072337	3	319, 69, 401	181, 113
62.	Chloroplast rubisco large subunit (rbcl)	L24073	3	606,	
63.	Zinc finger protein,	AB001883	3	208, 348, 431	132, 95
64.	Plastid RNA polymerase sigma factor	AB005290	3	902, 199, 93	333, 138
65.	Ubiquitin/ribosomal polyprotein	D12629	4	130, 190, 67, 228	91,1508, 275
66.	Alanine:glyoxylate aminotransferase	BI305374	4	231, 175, 144, 59	113, 150,109
67.	Cytochrome P450	AB038597	4	531, 245, 385, 426	105, 120, 107
68.	Ribosomal protein S15	D10962	4	21, 155, 58, 465	82, 132, 1043
69.	Alpha-tubulin	X91807	4	174,235,371,883	950,86,112
70.	Reversibly glycosylated polypeptide	Y18624	4	415,251,169,549	837,93,127
71.	Rnase S-like protein	AY056038	4	128, 162, 193, 516	715, 83,113
72.	Glutaredoxin	X77150	4	77, 89, 99,294	993,1007,174
73.	Beta-expansin (EXPB3)	AF261271	4	280, 110, 185, 720	103, 136, 345
74.	H protein subunit of glycine decarboxylase	AF022731	4	170,89,171,286	360,528,330
75.	Novel protein, osr40c1	X95402	4	371,507,229,341	109,109,249
76.	Major intrinsic protein,	D17443	4	212, 422, 62, 645	94, 183,1363
77.	Hos59	AB007628	4	185, 191, 132, 670	462, 74, 84
78.	Dehydrin	U60097	4	206,91,21,532	54,87,68
79.	Alpha 1 tubulin	Z11931	5	204, 215, 200, 509, 594	910,81, 557,119
80.	Ribosomal protein S4	Y15009	5	68, 78,181,98,645	841,94,89,249
81.	Photosystem II 10 KDa polypeptide	U86018	5	159, 53, 56, 53, 249	125, 245, 113, 671
82.	Translation initiation factor (GOS2)	AF094774	5	120, 75, 212, 54, 272	1001, 85, 93, 106
83.	Cap-binding protein p28	U34598	5	346, 169, 126, 66, 293	969, 129, 853, 118
84.	Cinnamoyl coa reductase	AJ428493	5	145,155,183,353,178	94,93,101,110
85.	Elicitor and UV light related transcription factor	AY083611	5	115, 80, 128, 105,154	110, 997, 343, 81
86.	Phosphoribulokinase precursor nuclear gene for chloroplast product	AF529237	5	710, 85, 85, 245, 521	293, 89, 170, 230
87.	Cyc07	D26060	6	111, 106, 138, 186, 201, 270	118, 807, 70, 95, 86
88.	SOD a gene manganese -superoxide dismutase		6	295, 47, 126, 57, 78, 93	2101,108, 96, 764,653
89.	Small GTP-binding protein	L35845	6	146,140, 83, 205, 69, 378	162,721, 84, 85506

90.	Mitochondrial phosphate transporter,	AB016065	6	435,362, 44, 215, 125,371	808, 100, 856, 475, 96
91.	Cytoplasmic aldolase,		6	147, 270, 110, 91, 270, 279	631, 176, 104, 101, 338
92.	S-phase-specific ribosomal protein (RSPSP94)	AF052503	6	91, 106, 138, 186, 201, 295	140, 612, 95, 108,904
93.	Nad-dependent formate dehydrogenase	AB019533	6	166, 172, 541, 125, 113, 321	412, 639, 86, 444,464
94.	GF14-c protein	U65957	6	42,412,79,123, 117,407	1037,88,183,452,93
95.	Translation initiation factor eif-5A.	AJ252135	6	46, 150, 123, 57, 75, 271	107, 982, 85, 284,91
96.	Malate dehydrogenase (MDH)	AF444195	6	253, 174, 169, 62, 125,120	105, 483, 103, 581,104
97.	26S Proteasome regulatory particle triple-A atpase subunit2b	AB037154	6	517, 246, 170, 217, 141, 297	1452, 212, 133, 75, 187
98.	Brain specific protein	D16140	6	193,456,79,123,117,487	1184,93,362,711,112
99.	Beta-amylase gene	L10346	7	51,399,165,264,210,253,125	145,139,394,260,127,643
100.	RAB5A protein	AJ292320	7	221, 78, 168, 49, 74, 94, 284	263, 95, 859, 201, 534, 138
101.	OSCDPK7	AB042550	7	900,144,153,116,168,225,420	526,88,145,475,479,420
102.	Submergence induced protein 2A	AF068332	7	144, 112, 95, 115, 125, 71, 291	941, 120, 73, 119, 229, 92
103.	ADP-ribosylation factor	D17760	7	154,161,71,120,119,82,357	1010,527,327,341,307,123
104.	Carbonic anhydrase 3	AF182806	7		
105.	Small GTP binding protein	AF218381	7	210,88,110,64, 65,66,	
106.	Rubisco activase small isoform	AB034748	7	73, 299, 192, 473, 282, 85, 260	99, 113, 85, 92, 345, 602
107.	Beta-glucosidase	U28047	8	61,88,244,116, 218,32,103,560	119,371,253,139,92,133,112
108.	Vacuolar H+-pyrophosphatase	D45384	8	347, 575, 380, 87, 331,411, 60, 519	1394, 141, 102, 85,132, 83, 96
109.	Catalase	D26484	8	145,97,278,777,90,68,94,266	116,911,351,221,93,112
110.	Small GTP-binding protein	AB029510	8	235, 88, 110, 64, 65, 66, 104, 319	354, 147, 109,859, 169, 118,716
111.	Ca2+ sensitive 3'(2'),5-diphosphonucleoside 3'(2') phosphohydrolase	U33283	8	130, 76, 77, 203, 211, 99, 154,416	394, 240, 90, 454, 101, 77, 162
112.	Triosephosphate isomerase	M87064	8	84, 124, 85, 133, 95, 79, 83, 289	133, 354, 270, 481, 129, 117, 92

113.	Ran 1,	AB015288	8	70, 70,57, 118, 90, 180,78, 389	110, 508, 88, 578, 75, 85, 90
114.	Ras-related GTP binding protein	S66160	8	117, 73, 48, 48, 72, 156, 104, 329	453, 836, 93, 88, 87, 106, 190
115.	LLS1 protein (Lls1)	AF284781	9	153, 243, 105, 279, 148, 185, 113, 295,589	539, 174, 81, 95, 140,112, 120,152
116.	Ribulose-5-phosphate-3-epimerase (RPE)	AF047444	9	188, 109, 141, 114,79, 110, 57, 70, 160	537, 148, 99, 458,526, 71, 119,125
117.	Ascorbate peroxidase	D45423	9	158, 175, 66, 49, 86, 80, 103, 59, 208	95, 94, 1356, 333, 86, 95, 117, 86
118.	Glyoxalase I	AB017042	9	22,121,125,65, 216,48,108,153,48,	864,72,218,111,131,140,80,535
119.	Glyceraldehyde-3-phosphate dehydrogenase	AF357884	9	154, 194, 129, 120, 137, 156, 378, 117, 273	467, 106, 1654,79, 81, 440, 97, 100
120.	Oxidase (IM1) plastid gene for plastid product	AF085174	9	125, 172, 94, 116, 109, 69, 72, 48, 731	84, 78, 129, 232, 111,394, 97, 304
121.	Gigantea-like protein	AJ133787	10	86,59,241,151, 210, 1497,69,219,93,306	422,866,77,133,798,104,200,117,88
122.	Sgt1 (sgt1)	AF192467	10	225, 70, 125, 139, 97, 66, 139, 83, 105,172	2003, 87, 222, 248, 82, 307, 559, 96, 119
123.	Cytosolic glutamine synthetase	X14245	11	128, 40, 104, 49, 107,88, 129, 75, 92,160, 581	690, 202, 108,420,114, 116, 127, 87, 93183
124.	Cysteine synthase (rcs3)	AF073697	11	94, 72, 110, 58, 264, 138, 52, 80, 60, 81, 309	443, 106, 82, 424, 84,323, 209, 324, 95,134
125.	Fructose-1,6-bisphosphatase	AB007193	12	239, 48, 75, 62, 67, 114, 87, 71, 66, 87, 76,552	104, 92, 88, 108, 404, 82, 127, 76, 157, 99,118
126.	Glyceraldehyde-3-phosphate dehydrogenase	U31676	12	32, 24, 101, 116, 100, 147, 61, 98, 143, 84, 51	133, 115, 287, 92, 600, 83, 82,424, 229,127
127.	UDP-glucuronic acid decarboxylase	AB079064	12	96,87,88,92,65, 89,78,94,6764, 41,192	
128.	1-deoxy-D-xylulose 5-phosphatereductoisomeraseprecursor	AF367205	12	132, 182, 87, 126, 63, 150, 68, 112, 105, 122, 147, 402	1010,278, 99, 82, 704, 235, 79, 80, 169, 497, 854

129.	Shaggy-like kinase etha	BI306125	13	349,41,90,60,273,75,57,141,51,96,117,168,57	129,499,89,76,94,90,414,95,99,61,565,93
130.	Sucrose synthase	Z15028	15	105,121,152,193,119,217,96,174,117,167,225,319,245,139,238	411,89,83,149,95,82,96,84,82,88,339,130,111,97
131.	Enolase	U09450	17	149,64,65,48,87,81,105,78,75,63,81,195,51,89,52,57,347	655,120,553,79,98,88,76,75,227,99,116,83,80,90,111,65

4.9 Analysis of selected gene families associated with drought stress response

Transcription factors play important roles in plant development and during adverse environmental conditions. This class of regulatory genes is induced as an early response to adverse conditions towards adaptation by manipulating the flux of different metabolic pathways. In *Arabidopsis* different families of transcription factors each containing a distinct type of DNA binding domain have been implicated in plant stress responses. These transcription activators execute their function directly through binding to specific motifs in the promoter region of stress responsive genes. Over expression of the regulatory proteins such as DREB1A and DREB1B has resulted in an enhanced tolerance to drought, salt and freezing (Jaglo-Ottosen *et al.*, 1998; Kasuga *et al.*, 1999) in *Arabidopsis*. Efficient identification and isolation of such TFs will improve our understanding of stress response mechanisms in rice. The analysis uncovered the transcripts encoding proteins similar to DREB like genes (Table 4.6).

In silico analysis identified eight paralogues of DREB like sequences in the rice genome with distinct AP2/REEBP DNA binding domain (Fig. 4.15 & Fig. 4.16). The DREB genes are organized into clusters to exhibit cumulative response under different abiotic stresses (Fig: 4.17). Similar organization of DREB genes was reported in *Arabidopsis* (Shinwari *et al.*, 1998).

The expression pattern of DREB1A and DREB1B in N22 *indica* rice seedlings showed that, DREB1A is induced under abiotic stress conditions (ABA 100um, PEG 20%, and NaCl 150 mM), (Fig. 4.18) these treatments showed to mimic stress responses in rice but the transcript levels of DREB1B was not altered.

4.9.1 DREB Gene Family of Rice:

DREB Genes	Chromosome	Accession	location (cM or bp)
CBF	chr 6, clone:P0425F02	AP001168	6 cM
CRT/DRE	chr 6, clone P0680A03	AB023482	12 cM
CRT/DRE 1	chr 8, clone P0623F08	AP004632	13,300- 14,062
CRT/DRE 2	chr 8, clone P0623F08	AP004632	17,190- 17,902
CBF2	chr 2 clone P0135D07	AP006060	114 cM
CBF3	chr 4 OSJNBa0019K04	AL606640	87.1 cM
DREB 1A	WGS	AAAA01001957	14469-13510
DREB1 C	WGS	AAAA01001957	10003- 9264
DREB 1B	WGS	AAAA01001957	5992 - 4988
DREB2	chr1, clone P0701D05	AP003301	16.1 cM

Table 4.6: Members of DREB gene family and their locations

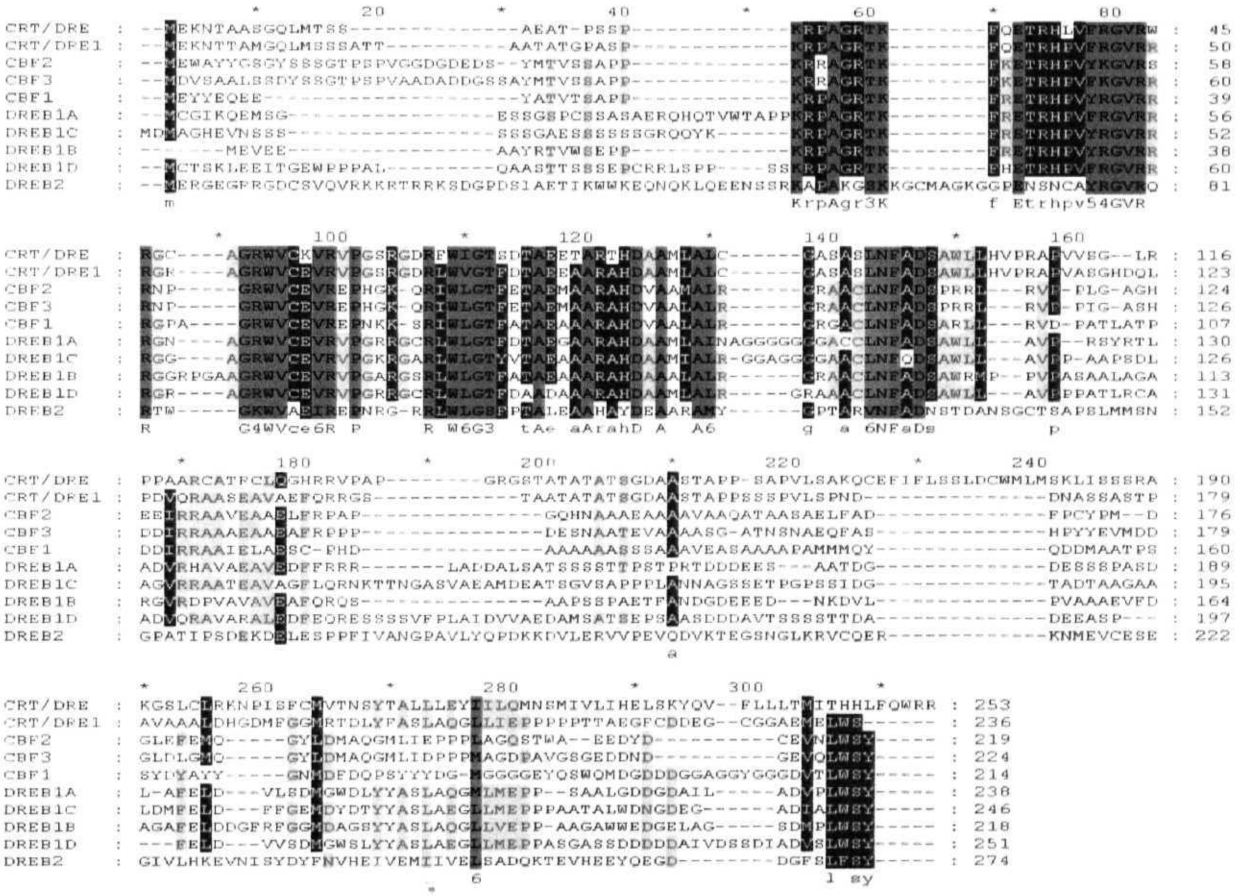


Fig. 4.15: Multiple alignment of a.a seqs of DREB family members of rice

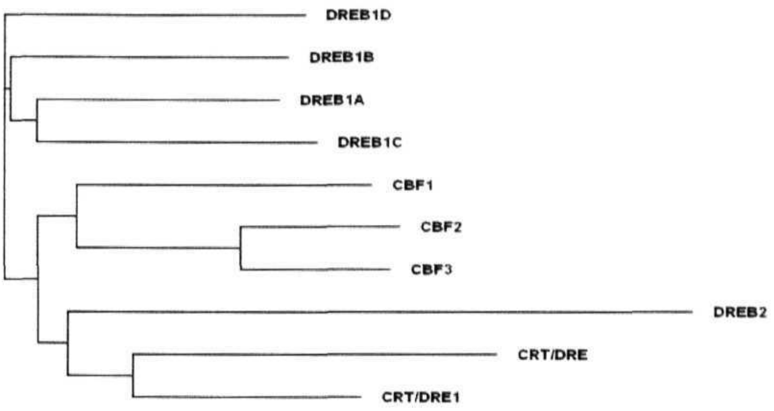


Fig. 4.16: Dendrogram showing the relation of different members of the DREB genes

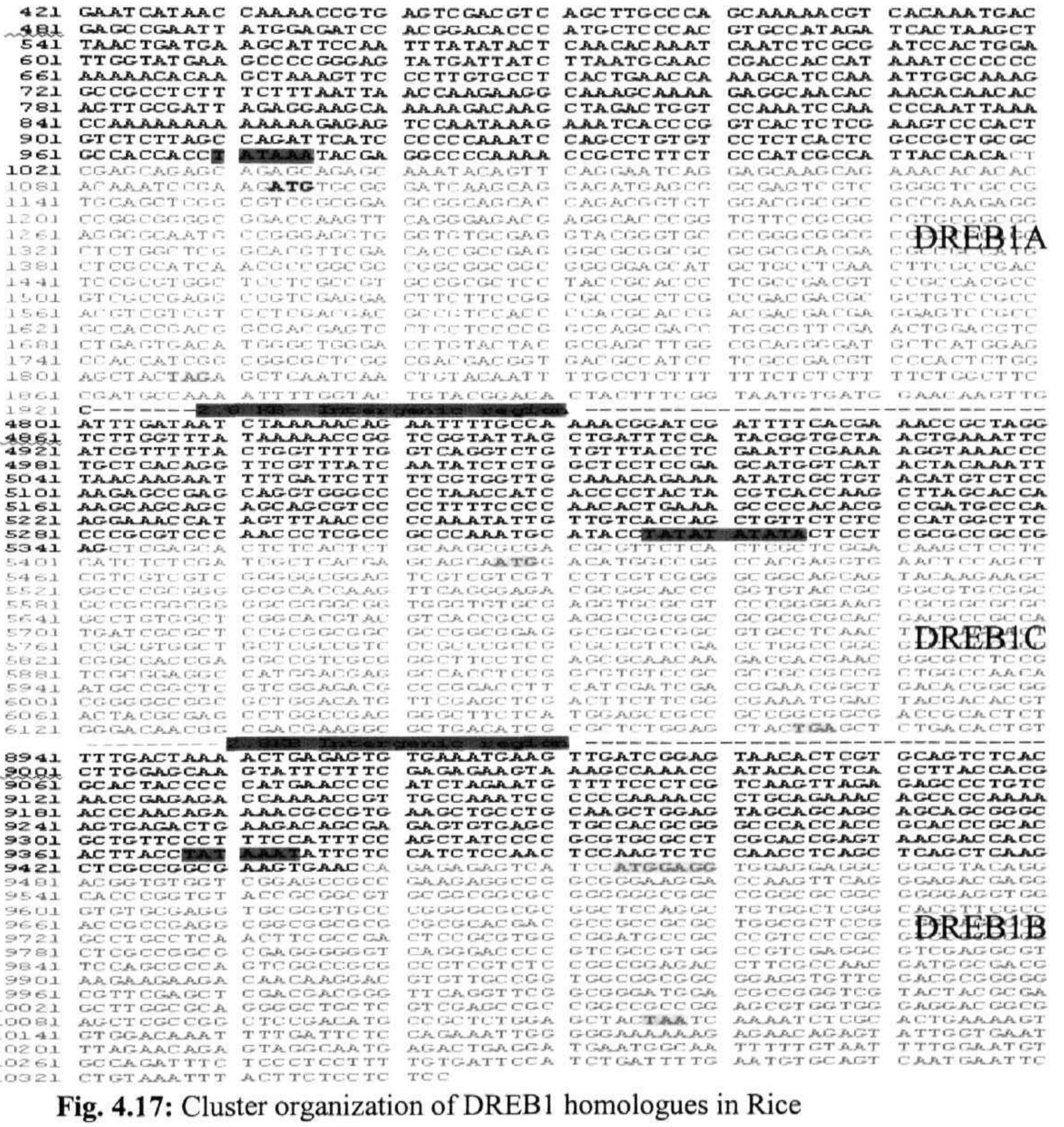


Fig. 4.17: Cluster organization of DREB1 homologues in Rice

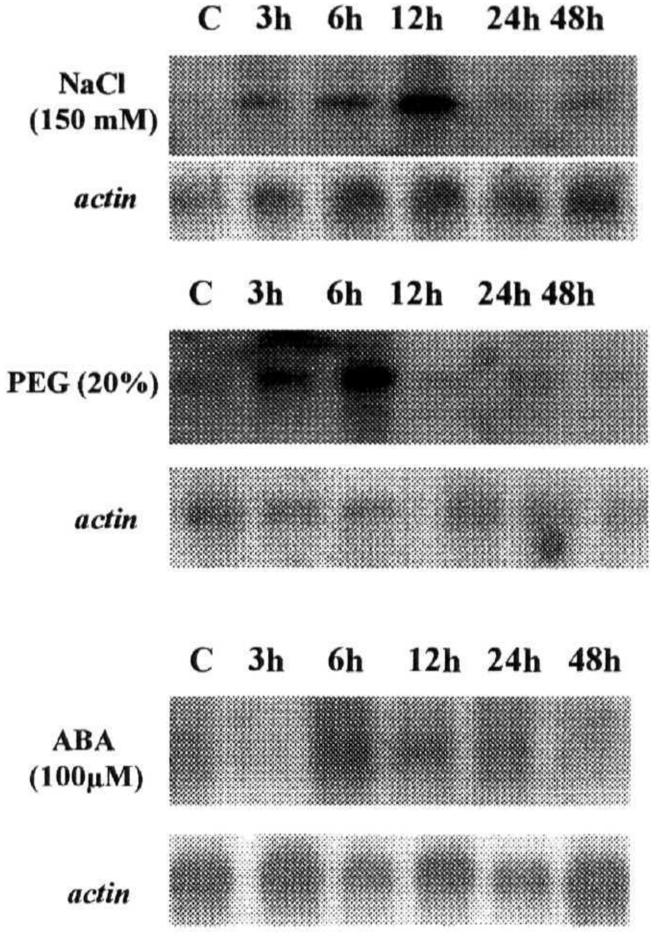


Fig 4.18: Northern analysis of EST similar to DREB 1A

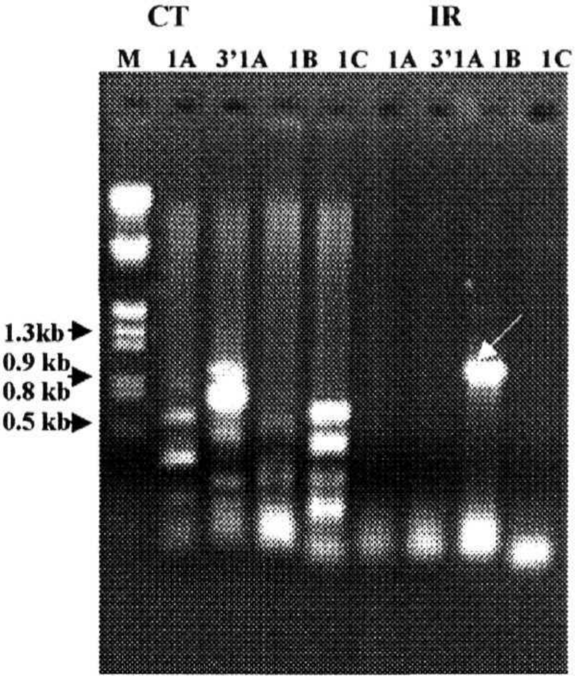


Fig. 4.19: DREBs PCR

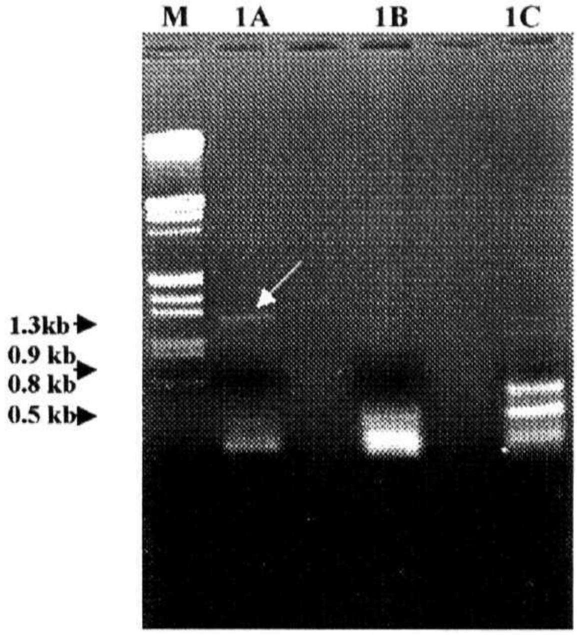


Fig. 4.20:DREBs RT- PCR

To isolate the other members of the DREB gene family specific primers were designed from rice genomic sequence. Two members of this family were isolated (Fig. 4.19 and Fig. 4.20) and sequenced from both 5' and 3'ends.

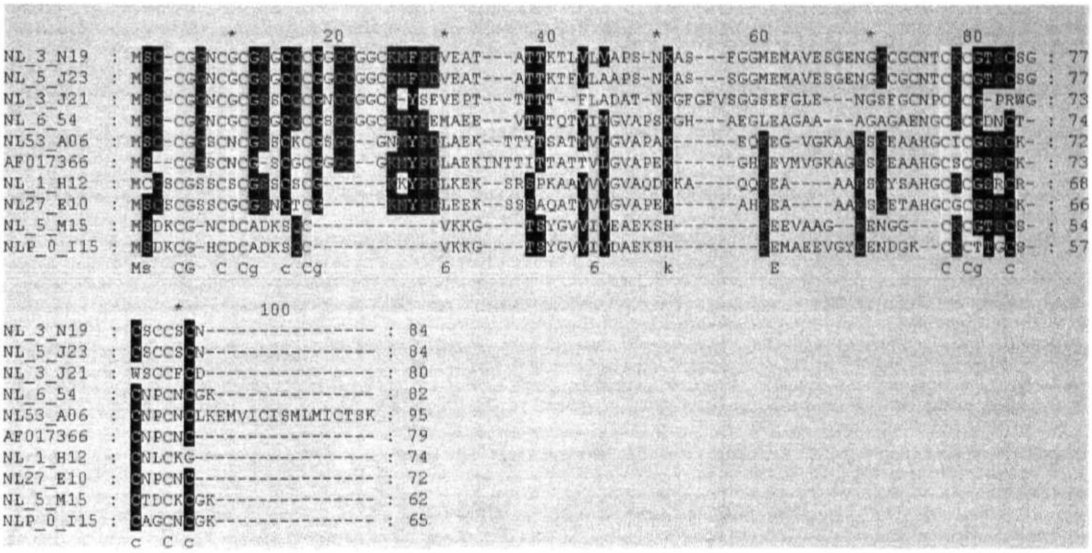
4.9.2. Metallothionein gene family in rice

Analysis of the ESTs uncovered 10 different metallothionein-like protein coding genes in rice. Of these, five were represented only once where as the other five were represented abundantly (Table 4.7). Genomic analysis rice metallothionein gene family resulted in understanding the structure and function in rice. The repeats and small coding regions of these genes posed a significant difficulty in the genome sequence assembly and annotation. The distribution of cysteine residues in the translated aminoacid sequences identified the presence of two type 1and nine type 2 metallothionein like proteins (Fig.4.21 & Fig.4.22). All type1 MTs have two exons and one intron, the type 2 MTs have three exons and two introns. Two type2 MTs show tandem duplications on chr1 and chr12 on rice. The duplication region on 12th chromosome might have resulted in a truncated protein due to misassembly. Earlier it was reported that the MT genes form small families of clustered gene copies in mammals (Palmiter *et al.*, 1992) and tomato (Giritch *et al.*, 1998).

The biological significance of metallothionein was studied recently in mammalian systems and assigned several functions such as heavy metal tolerance, antioxidant properties and in maintenance of redox and regulation of intracellular signaling besides other undefined functions in plants. These genes were reported to play a major role in abiotic stress tolerance. The available mutant screens in rice, *Arabidopsis* and transgenics along with the biochemical analysis will help to uncover the functions of these genes.

Table 4.7: Members of metallothionein gene family, redundancy and location

Clone	**Type**	**Abundance**	**Exons**	**BAC/PAC CLONE**
NL_1_H12	Type I variant	11	2	chromosome 12 OSJNBb0078B20
NL27_E10	Type 1	U	2	chromosome 3 OSJNBb0022M22
NL_3_J21	Type 2	101	3	chromosome 1 (P0459B04)
NL_3_N19	Type 2	47	3	chromosome 5 P0016H04
NL_5_J23	Type 2 variant	U	3	
NL_6_54	Type 2	U	3	chromosome 1 P0434B04 (Duplication)
NL_5_M15	Type2	49	3	chromosome 1 B1015E06
NLP_0_I15	Type2	U	3	chromosome 5 P0015G13
NL53_A06	Type 2	2	3	chromosome 12 OSJNBa0018E22 (Duplication)



```
                   *        20         *        40         *        60         *        80
NL_3_N19   : MSC-CGCNCGCGSGCCCCGGCGGCKMFPDVEAT---ATTKTLVLVAPS-NKAS---FGGMEMAVESGENGCCGCNTCKCGTSCSG : 77
NL_5_J23   : MSC-CGCNCGCGSGCCCCGGCGGCKMFPDVEAT---ATTKTFVLAAPS-NKAS---SGGMEMAVESGENGCCGCNTCKCGTSCSG : 77
NL_3_J21   : MSC-CGCNCGCGSSCCCCGNCGGCK-YSEVEPT---TTTT--FLADAT-NKGFGFVSGGSEFGLE---NGSFGCNPCKCG-PRWG : 73
NL_6_54    : MSC-CGCNCGCGSGCCCCGSCGGCKNYPEMAEE---VTTTQTVIMGVAPSKGH---AEGLEAGAA---AGAGAENGCKCGDNCT- : 74
NL53_A06   : MSC-CGGSCNCGSSCKCGSGC--GNMYPDLAEK--TTYTSATMVLGVAPAK------EQKEG-VGKAAKSKEAAHGCICGSSCK- : 72
AF017366   : MS--CGGSCNCG-SCGCGGC--GKMYPDLAEKINTTITTATTVLGVAPEK------GHKEVMVGKAGKSKEAAHGCSCGSSCK- : 73
NL_1_H12   : MCCSCGSSCSCGSSCSCG------KKYPDLKEK--SRSPKAAVVVGVAQDKKA----QQKEA----AAKSKYSAHGCKCGSRCR- : 68
NL27_E10   : MSCSCGSSCGCGSNCTCG------KNYPDLEEK--SSSAQATVVLGVAPEK------AHKEA----AAKSKETAHGCGCGSSCK- : 66
NL_5_M15   : MSDKCG-NCDCADKSQC------------VKKG----TSYGVVIVEAEKSH---------EEEVAAG--EENGG---CKCGTNCS- : 54
NLP_0_I15  : MSDKCG-HCDCADKSQC------------VKKG----TSYGVMIMDAEKSH---------EEMAEEVGYEENDGK--CKCTTGCS- : 57
              Ms   CG  C Cg  c Cg            6        6     k          E                C Cg  c

                 *         100
NL_3_N19   : CSCCSCN----------------- : 84
NL_5_J23   : CSCCSCN----------------- : 84
NL_3_J21   : WSCCFCD----------------- : 80
NL_6_54    : CNPCNCGK---------------- : 82
NL53_A06   : CNPCNCLKEMVICISMLMICTSK : 95
AF017366   : CNPCNC------------------ : 79
NL_1_H12   : CNLCKG------------------ : 74
NL27_E10   : CNPCNC------------------ : 72
NL_5_M15   : CTDCKCGK---------------- : 62
NLP_0_I15  : CAGCNCGK---------------- : 65
              c  C c
```

Fig. 4.21: Multiple alignment of amino acid seqs of metallothionein genes in Rice

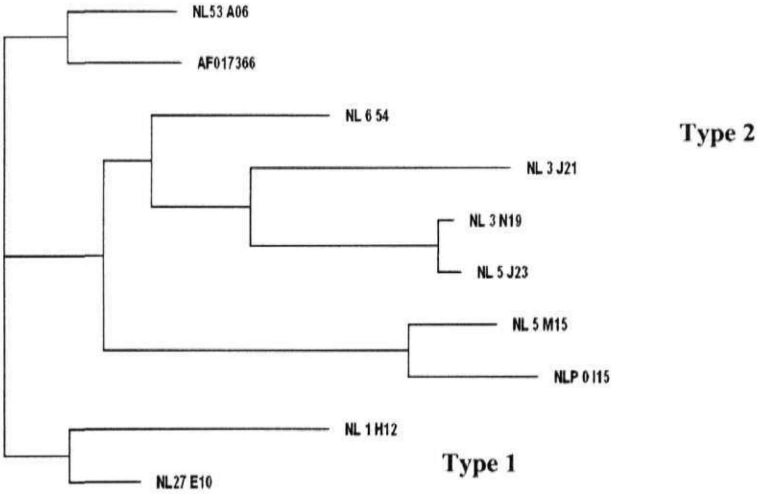


Fig 4.22: Dendrogram showing the Type I and Type 2 metallothionein gene members of identified ESTs

5. Discussion

Global analysis of transcription profiles associated with abiotic stresses provided the leads towards understanding these complex traits and has become prerequisite for identifying the genetic determinants, to aid the plant breeding programs in improving the crop performance under adverse conditions. These approaches are therefore highly suitable to the analysis of genome wide changes due to drought stress rather than a narrow window representing expression of a single or a few genes.

Large-scale Expressed Sequence Tags have been analyzed from rice in view of their utility as powerful tools to catalogue all the genes (Uchimiya *et al.*, 1992; Yamamoto and Sasaki, 1997) and deciphering the roles of transcriptionally regulated genes in different tissues (Ewing *et al.*, 1999). Only a few studies have focused on the analysis of transcriptome profiles of rice seedlings subjected to stress (Umeda *et al.*, 1994; Matsumura *et al.*, 1999; Kawasaki *et al.*, 2001) or drought.

Towards this end, we have utilized the ESTs from normalized cDNA libraries constructed from rice seedlings subjected to progressive drought stress (Reddy *et al.*, 2002), and analyzed to understand and characterize the cellular activities under drought stress response. To uncover the genetic determinants of drought tolerance and also understand the similarity between coding portion of *indica* and *japonica* sub species of rice, we have carried out the EST analysis using a range of comparative genomic approaches. A total of 5, 500 ESTs were obtained with 70% of sequencing efficiency. Among these, 200 ESTs were from root tissue and 5, 300 ESTs were from leaf tissue. Of the remainng 30% clones, about a half represented ribosomal RNA and the other half are low sequence quality read length and vector contamination.

Expressed Sequence Tag resource generated from normalized cDNA libraries of leaf and root tissues of drought stressed seedlings of *indica* sub species rice cultivar Nagina 22 have become the central resources to uncover novel genes and to identify genes associated with drought stress response in rice (Fig. 3.1.). These ESTs, generated from the 3'end of cDNAs, represents the 3'sequence therefore useful in identification of gene paralogues and diversity among rice genotypes. Further, the 3' UTR sequences harbor SSRs or other potential marker sequences that are useful in molecular mapping and dissecting the QTLs and candidate genes.

The *indica* sub species rice cultivar Nagina 22 (N22), which is grown under upland conditions, has several morphological and physiological characters such as early maturity, heat tolerance, two-point root system, accumulation and mobilization of carbohydrates, high regeneration and recovery, all associated with drought tolerance mechanisms in plants.

Initially this EST analysis was utilized to characterize the efficiency of normalization to uncover novel genes for in depth sequencing of the library. These results clearly revealed that normalization had effectively reduced the representation of highly abundant transcripts such as Rubisco to a narrow range compared to non-normalized library.

5.1 Development of non-redundant set of ESTs

EST clustering, contig assembly and BLAST analyses to identify the unigene transcripts resulted in 2069 non-redundant ESTs. The representation of 1241 ESTs as singletons in the unigene set shows the utility of deep coverage of transcriptome (Fig. 4.6).

5.2 *In silico* transcription profiles

The transcription profiles clearly revealed the presence of potential stress related genes as shown in (Table 4.1) Among the 90 highly represented transcripts, 25% were novel with unknown functions. The genes RicMT, Thioredoxin h, Metallothionein-like protein Rd22, Glycine-rich protein, Lipid transfer protein LPT IV, EF-hand Ca2+-binding protein, Calmodulin, Cytochrome P450 monooxygenase, Fructose-1,6-bisphosphatase, Glutaredoxin, Catalase, HSP90-like protein, Dof zinc finger protein, Class III chitinase, and Serine/threonine kinase associated with different metabolic pathways were highly represented in the analysis (Table 4.1)

5.3 Gene annotation

The annotation of non-redundant set of ESTs allowed us to assign putative functions for 50% of ESTs through homology search in nr database. The annotation was further improved by utilizing rice genome annotation in rice GAAS and large catalogue of full-length cDNA sequences. This resulted in the assignment of gene functions to 78% of the ESTs and 22% of ESTs remained novel. Full-length cDNA clones were assigned to

1151 ESTs through BLAST N homology search with 95% similarity cutoff. This forms a formidable resource towards characterizing the gene structure and identification of mutants in the rice mutant database. Homology search in dbEST division of rice sequences revealed a significant number of ESTs (912) are novel at the 3' end of gene sequences. This higher sequence diversity in the 3-prime UTRs of the transcripts is best used to distinguish between gene family members and also for accurate gene prediction in automated rice genome annotation.

5.4 Identification of putative stress responsive genes

Based on the putative functions, the sequences were further characterized by grouping stress responsive genes into 15 distinct functional classes (Bevan *et al.*, 1998) associated with different cellular processes. The putative candidate stress responsive genes have been identified based on reported experimental data (Bray. 2002), and comparing the EST dataset to the documented gene sequences and possible candidate gene sequences identified from micoarray gene expression profiles of rice, *Arabidopsis*, barley, and maize plants under different abiotic stresses. The stress responsive genes identified from the above expression profiles include 650 from *Arabidopsis*, 150 from barley and 120 from maize, and 100 from rice. These were compared to the EST data set using TBLASTX with an E-value >1e-20. Genes with stress-related and regulatory functions comprised 25% of the total ESTs. These categories included proteins with known function in cellular defenses against abiotic (drought, cold and salinity) and biotic (pathogen) stresses, and proteins involved in developmental and stress response signaling and transcription. Based on the types of genes represented, tolerance mechanisms rely on precise integration of developmental processes with stress-related responses. 12% of identified stress responsive ESTs are with unknown functions forms a large proportion as unclassified proteins. Interestingly the distribution of the 583 putative stress responsive ESTs among the functional categories showed that transcription factors (8%) were efficiently captured next only to the cell metabolism class (18%) (Fig. 4.14). The data imply that transcription factors seem to be an abundantly expressed class of genes during stress in rice. This also agrees with the earlier reports on the role of transcription activators in stress response associated changes in gene expression (Chen *et al.*, 2002).

5.5 Candidate genes at QTL associated with drought tolerance

The features of chromosome structure (gene structure) and function have significant implications for studies in crop improvement. The coding and non-coding DNA sequences linked to specific traits in genetic studies through QTL analyses provided a source of molecular markers to increase the efficiency of breeding activities. Rice genome sequencing opened new opportunities towards delineating the QTL regions into Mendilian factors and QTL cloning. By localizing ESTs onto genomic sequences provide a powerful route for gene prediction and understanding gene structure (Kan *et al.*, 2001). This simplifies the cloning of agronomically important genes in the QTL s through physical map of the rice genome (Chen *et al.*, 2002). Genomic regions assigned to 1500 ESTs (Babu *et al.*, 2002) based on above 95% identity along 90% of the length of the sequence in the aligned region using BLAST N program showed that 1400 ESTs had defined genomic region in IRGSP rice genome. Genetically anchored BAC/PAC clones were identified for 1326 ESTs (Table 4.2).

Fine mapping of QTLs has sometimes revealed the presence of tightly linked loci affecting the same trait (Fridman *et al.*, 2002; Steinmetz *et al.*, 2002). Eight genes responsible for QTLs have been identified in plants so far (Doebley *et al.*, 1997; Frary *et al.*, 2000; Yano *et al.*, 2000; Takahashi *et al.*, 2001; El-Din El-Assal *et al.*, 2001; Liu *et al.*, 2002; Kojima *et al.*, 2002). Some of the genes encode for transcription factors, whereas others encode proteins involved in metabolism or active in signal perception/transduction pathways. The emerging concept is to exploit the possibility of looking at variation directly in genes and not at anonymous markers (candidate gene association studies), as well as to saturate the genome with markers (whole genome scan) (Remington *et al.*, 2001; Rafalski 2002).

Candidates can also be identified from expression profiling experiments, under the assumption that genes that show genotype-specific differences in their level of expression could be the causative agents for the variation in a trait. This approach is considered only when a QTL for a trait of interest is shown to co-map with a QTL controlling the expression of a candidate gene. Physical map locations of ESTs observed by mapping EST sequences on to genetically anchored BAC/ PAC clones of rice genomic sequences revealed known stress responsive genes in the QTLs associated with drought tolerance (Table 4.3)

5.6 Analysis of organization of stress responsive genes in rice

Gene organization analysis by aligning the cDNA sequences onto the genomic sequences revealed genes with single exon to seventeen exons. Single exon genes were mainly those having small coding sequence and organellar gene products, which are integrated into the nuclear genome (Shahmuradov *et al.*, 2003). The genes involved in cellular metabolism glyceraldehyde-3-phosphate dehydrogenase, fructose-1,6-bisphosphatase, glutamine synthetase, UDP-glucuronic acid decarboxylase, sucrose synthase, enolase, have showed large number of exons 12, 11, 12, 15, and 17 respectively in the coding region. This highlights the stringent regulation of metabolic pathways. The observed exon size of minimum length 21 bp, maximum length 1600 bp and intron sequence lengths of minimum 37 bp and maximum 2000 bp reflect the complexity of gene organization.

Promoter sequences of stress responsive genes were analyzed for the known motifs of some target genes associated with abiotic stress (Table 5.1). The relevance of the phenotype to emerging new levels of gene regulation is difficult to evaluate. This type of regulatory variation, in fact, frequently concerns both regulatory and structural genomic regions — like promoters, introns, silencers and other non-coding sequences mapping away from transcriptional units — all known to be more variable than protein-coding DNA sequences. Such regulatory variation affecting both the level and pattern of expression of the genes appears to be fairly common.

The survey of conserved noncoding sequences (CNS) among cereal gene promoters supported the hypothesis of rapid evolution of noncoding sequences between closely related species (Guo and Moose 2003).

Table 5.1: *cis*- acting elements present in promoter regions (1000bp) of the putative stress responsive genes

Putative function	**Acc No**	**ABRE**	**LTRE**	**DRE**
Metallothionein-like protein	BI305617	- 187		-543
Metallothionein-like protein	AF017366	- 104		
Cytochrome P450	AF088221	- 665	- 243, - 363	-223
Thioredoxin	AB053294		- 424	
Small GTP binding protein	AF218381	- 96		-102
HSP90-like protein	AY077617	- 390		
Nuclear transport factor 2	AB011262	- 270, -376	- 531	-552
Ras-related GTP binding protein	S66160		- 216	
Small GTP-binding protein osrac3	AB029510		- 188	-188
Photosystem II 10 kda polypeptide	U86018	- 187		
23 kda polypeptide of photosystem II	AF052203	-139	- 142	- 417
Chlorophyll a/b binding protein of photosystem II	D00641	- 89		
Chlorophyll a/b binding protein	AF061577		- 241	
Triosephosphate isomerase	M87064		- 406	
Ubiquitin/ribosomal polyprotein	D12629		- 152	
Ribosomal protein	U86017		- 306	
Glutaredoxin	X77150	- 181		
Ribonuclease	AB052842		- 335	
Cyclophilin 2 (Cyp2)	BI305468	- 143		
Root-specific rcc3	BI305683		- 662	
Shaggy-like kinase etha	BI306125		-239	-260
Brain specific protein	D16140	- 180	- 80, -329	-329
Beta-glucosidase	U28047		-71	-72, -569
Reversibly glycosylated polypeptide	Y18624		-284	-285
Alpha 1 tubulin	Z11931		-412	-413
Beta tubulin	X78142	- 67	- 55	-65
EF-1 alpha	D63583		-344	-157
Oscdpk7	AB042550		-302	-322, -447
S-adenosylmethionine synthetase	AJ296743		-100	
Translation initiation factor eif-5A.	AJ252135		139	-140, - 403
GF14-c protein	U65957	- 105	- 221	-222, -309

Enolase	U09450	- 161	-270	
Calmodulin (cam2)	AF042839		-310	- 311
S-adenosylmethionine decarboxylase 2	A251899		454	- 297 - 392
Ascorbate peroxidase	D45423	- 328	-188	- 297
Dehydrin	U60097	- 202	211	
ADP-ribosylation	D17760	- 96	-177	-178
Glycine-rich protein	AF010579		-304	
Glycine rich RNA binding protein	AJ302060		- 84	
Cinnamoyl coAreductase	AJ428493		-538	- 538
UDP-glucuronic acid decarboxylase	AB079064		- 355	- 355
H protein subunit of glycine decarboxylase	AF022731		-445	- 471 -558
Plastocyanin precursor	AF093636	- 267	- 267	- 527
Ribulose 1,5-bisphosphate carboxylase small subunit	D00644	- 170	- 161	- 162 - 176
Thionin	AB072337	- 40		
Nad-dependent formate dehydrogenase	AB019533	- 206		
Ribulose-5-phosphate-3-epimerase (RPE)	AF047444		- 402	
Lipid transfer protein	Z23271		- 53	- 621
Lipid transfer protein	BU673284	- 44	-63	- 63
Glyceraldehyde-3-phosphate dehydrogenase	U31676		- 313	- 137
Fructose-1,6-bisphosphatase	AB007193	- 362	-420	-426 -432
Zinc finger protein,	AB001883		- 137	- 487
Beta-expansin (EXPB7)	AF261275		-101	
Rubisco activase small isoform	AB034748	- 83		
Ca2+ sensitive 3'(2'),5-diphosphonucleoside 3'(2') phosphohydrolase	U33283	- 512	-195	-198
Ferredoxin	AF010320	- 70		
Ribosomal protein S15	D10962		-176	
Positive element factor 1 (PF1)	L24390	- 359	-130	- 131 - 159
Chloroplast atpb ATP synthase beta subunit	AB037543		-214	
RIP1 ribosome inactivating protein	AB051107	- 678		
Phosphoribulokinase	AF529237	-429		
Plastid RNA polymerase sigma factor	AB005290		-123	
Cap-binding protein p28	U34598	- 341	-178	-713
LLS1 protein (Lls1)	AF284781	- 92		

Sgtl (sgtl)	AF192467		-112	-112
Elicitor and UV light related transcription factor	AY083611		-117	- 117
ADP-glucose pyrophosphorylase small subunit	AY028315		- 323	
OsNAC5 protein	AB028184	- 81	- 124	
Hos59	AB007628		- 471	
DRE-binding 1B protein	AY166833			- 326
Mitochondrial phosphate transporter	AB016065			- 375
OSMYB1	D88617			- 487
abscisic acid- and stress-inducible protein (Asr1)	AF039573			- 586
Sucrose synthase	Z15028			- 664
Peroxiredoxin	AF203879			- 774
small GTP-binding protein(ORRab-2)	L35845			- 373 - 502

5.7 Analysis of gene families associated with stress response

5.7.1 Metallothionein gene family

Analysis of the ESTs uncovered 10 different metallothionein-like protein coding genes in rice. Of these, five were represented only once where as the other five were represented abundantly. Genomic analysis was carried out to understand the structure and function of metallothionein gene family in rice. The repeats and small coding regions of these genes posed a significant difficulty in the genome sequence assembly and annotation. The distribution of cysteine residues in the translated aminoacid sequences identified the presence of two type 1and nine type 2 metallothionein-like proteins. All type1 MTs have two exons and one intron, the type 2 MTs have three exons and two introns. Two type2 MTs show tendem duplications on chr1 and chr12 on rice. The duplication region on 12th chr might have resulted in a truncated protein due to misassembly. Earlier it was reported that the MT genes form small families of clustered gene copies in mammals (Palmiter *et al.*, 1992) and tomato (Giritch *et al.*, 1998).

The biological significance of metallothioneins was studied recently in mammalian systems and assigned several functions such as heavy metal tolerance, antioxidant properties and in maintenance of redox and regulation of intracellular signaling besides other undefined functions in plants. These genes were reported to play a major role in abiotic stress tolerance. The available mutant screens in rice, *Arabidopsis* and transgenics along with the biochemical analysis will help to uncover the functions of these genes.

5.7.2 DREB gene family

Transcription factors play important role in plant development and during adverse environmental conditions. In *Arabdopsis*, different families of transcription factors each containing a distinct type of DNA binding domain have been implicated in plant stress responses. Stress-inducible transcription factors were identified from a wide range of protein families: DREB family, zinc finger proteins, WRKY family members, MYBs, basic helix-loop-helix proteins, members of the bZIP family, NAC family members, and homeodomain transcription factors (Table 5.2). These transcription factors probably regulate various stress-inducible genes either cooperatively or separately. Functional analysis of these stress-inducible transcription factors should provide valuable information on signal transduction in response to drought, cold and high-salinity. Recently, Fowler and Thomashow identified 306 cold-regulated genes and 41 DREB/CBF-regulated genes using Affymetrix Gene Chips (Fowler and Thomashow 2002). Over expression of the regulatory proteins such as DREB1A and DREB1B has resulted in an enhanced tolerance to drought, salt and freezing (Jaglo-Ottosen *et al.*, 1998; Kasuga *et al.*, 1999) in *Arabidosis*.

Table: 5.2. Different classes of transcription factors associated with stress response.

BI306107	Similar to DREB1A	BI305762	Transcription factor IIA small subunit
BU672792	DRE binding factor 1	BI306249	Transcription factor BTF3
BU673228	DRE-binding protein 1B	CB967424	Tat binding protein
CB966697	AP2 domain transcription factor	CB965560	Transcription factor GT-3b
BI305874	AP2 domain protein homolog	BU672783	Transcription factor Hap5a protein
BI306078	RING-H2 finger protein RHA1a		
BI306209	RING finger protein		
BU673403	RING finger-like protein		
BI306221	RING zinc finger protein		
BI305625	zinc finger protein, putative		
BI305955	similar to RING-H2 finger protein RHA1a		
BI305867	RING3-like bromodomain protein		
CB965937	RING-H2 finger protein RHA2a		
BU672942	small zinc finger-like protein (TIM9)		
BI305705	Dof zinc finger protein		
BI305249	C3HC4-type RING zinc finger protein		
BU673773	bZIP transcription factor		
CB965631	CONSTANS family zinc finger protein		
BI305714	TGA-type basic leucine zipper protein		
BI305238	leucine-responsive regulatory protein		
BU673704	homeodomain leucine zipper protein		
CB967252	WRKY family transcription factor		
BU673758	OsNAC5 protein		
BU673061	zinc finger protein		
BI305764	zinc finger protein		
BI306362	zinc finger protein		
BU673053	zinc finger-like protein		
BI306687	zinc-finger-like protein		
BI305935	DNA binding protein		
BI306302	RNA-binding protein		
CB965488	nuclear RNA binding protein A		
BI305731	RNA binding protein		
CB966397	small nuclear ribonucleoprotein		
BU673870	HOS59		
BU673703	dnaJ-like protein		
BI305268	helicase-like transcription factor		
CB967092	one helix protein		
BI305963	nucleoid DNA-binding protein cnd41		
CB967107	nucleic acid binding protein		
BI306418	R2R3MYB-domain protein		
BI305518	transcriptional regulator		
BI306059	OSMYB1		
BI305523	nucleic acid-binding protein		
BU673410	p53 binding protein		

Efficient identification and isolation of such TFs will improve our understanding of stress response mechanisms in rice. The analysis uncovered the transcripts encoding proteins similar to DREB like genes. In silico analysis identified eight paralogues of DREB like sequences with distinct AP2/EREBP DNA binding domain in the rice genome. The DREB genes are organized into clusters to exhibit cumulative response under different abiotic stresses (Fig. 4.17). Similar organization of DREB genes were reported in *Arabidopsis* (Shinwari *et al.*, 1998).

In the present study the expression pattern of DREB1A and DREB1B showed that, DREB1A is induced under abiotic stress conditions (ABA, PEG, and NaCl), but the transcript levels of DREB1B was not altered. Further full-length DREB gene was cloned and sequenced.

5.8. Analysis of novel sequences

The ESTs which do not have significant homology in the present nucleotide, protein, and EST databases were considered as novel sequence and the accession numbers of these are given in (Table 5.3).

Table 5.3: Accession numbers of novel sequences.

BU673127	BU672884	CB965922	BI305455	BI305277	CB965680	CB965420	CB966405
BU673117	BU672944	CB965916	BI305460	BU673163	CB965675	CB965417	CB966394
CB964547	CB966013	CB965912	BI305462	BU673151	CB965660	CB965413	CB965326
BU673092	CB965932	CB965869	BI305466	BI305270	CB965657	CB965398	CB966457
CB964591	CB965934	CB965910	BI305475	BU672816	CB965653	CB967001	CB966801
BU673086	CB965935	CB965881	BI305385	BU672813	CB965183	CB966369	CB966383
BU673084	CB965939	CB965903	BI305494	BI306568	CB965688	CB967133	CB966870
BU673082	BI305194	CB965900	BI305546	BU673207	CB965734	CB967069	CB966863
BU673095	BI305185	BI305197	BI305880	BI305294	CB966368	CB967068	CB966861
CB964718	CB965961	CB965896	BI305813	BU673317	CB965756	CB967062	CB966857
CB964788	CB965963	CB965889	BI306280	BI306284	CB965744	CB967057	CB966822
CB964783	CB965964	CB965911	BI305842	BI306297	CB965743	CB967173	CB966597
CB964759	CB965965	BI306133	BI305834	BI306304	CB965739	CB967003	CB966808
CB964548	CB965967	BI305311	BI305723	BI306308	CB965737	CB967185	CB966626
BI305227	CB965980	BI305910	BI305722	BI306317	CB965708	CB966980	CB966787
BU673075	CB965788	BI305922	BI305627	BI306334	CB965735	CB966965	CB966779
CB964672	CB966009	BI305938	BI305823	BI306335	CB965619	CB966952	CB966680
CB964671	BI305183	BI305941	BI305635	BI305309	CB965733	CB966930	CB966656
CB964613	CB966014	BI305952	BI305789	BI306340	CB965730	CB966929	CB966655
CB964612	CB966016	BI305956	BI305747	BI305305	CB965724	CB966907	CB966885
BI305204	CB966025	BI305961	BI305741	BI306347	CB965722	CB967050	CB966811

BI305264	CB966026	BI305975	BI305700	BI306348	CB965736	CB967482	CB964926
BU673032	CB966033	BI305992	BI305660	BI306607	CB965266	CB967468	BU673538
BU672874	CB966034	BI306069	BI305335	BI306437	CB965641	CB967446	CB965032
BU672869	CB966035	BI306080	BI305716	BI306567	CB965368	CB967425	CB965020
BU672865	CB966036	BU673280	BU672786	BI306562	CB965360	CB967405	CB965015
BU672852	CB966048	BI306098	BI306247	BI306509	CB966053	CB967370	CB964988
BU672879	CB966049	BI305349	BI305282	BI306490	CB965328	CB967134	CB965062
BU672822	CB966052	BI306139	BI306618	BI306462	CB966057	CB966871	CB964949
BU672880	CB965991	BI306143	BI306640	BI305295	CB965380	CB967314	CB965078
BU673426	CB965884	BI306150	CB966058	BI306439	CB965273	CB967288	CB964907
BU673422	CB965784	BI306193	BI305281	BI306351	CB965392	CB967263	CB964870
BU673339	BU673318	BI306197	BI305182	BI306422	CB965252	CB967240	CB964864
BU673333	CB965790	BI306199	BI306672	BI306391	CB965226	CB967239	CB964979
BI305266	BU673249	BI306216	BI306692	BI306361	CB965225	CB967201	CB965177
BU672842	CB965794	BI306219	BI306704	BI305290	CB965211	CB967348	BU673512
BU672919	CB965803	BI306225	BI306706	BI306354	CB965276	CB966442	BU673492
BU673024	CB965813	BI306226	BI306708	BI306266	CB965430	CB966557	BU673465
CB965793	CB965820	BI306232	BI305288	BI306456	CB965617	CB966548	BU673460
BU672988	CB965829	BI306236	BU673165	CB965626	CB965582	CB966506	CB965152
BU672877	CB965830	BI306087	BI305267	CB965707	CB965530	CB966487	CB965048
BU672933	CB965838	BI305643	BI306669	CB965705	CB965501	CB966481	CB965140
BU672899	CB965861	BI305407	BU673254	CB965704	CB965494	CB966466	CB966338
BI305236	CB965866	BI305394	BI305399	CB965700	CB965376	CB966891	CB965110
BU672898	CB965909	BI305431	BU673210	CB965696	CB965468	CB966447	CB965337
BU672894	CB965776	BI305446	BI306728	CB965714	CB966391	CB966642	CB965108
BU672893	CB965923	BI305451	BU673169	CB965716	CB965428	CB966440	CB965091

CB965090
CB965080
CB965142
CB966072
CB966196
CB966194
CB966162
CB966152
CB966147
CB966123
CB966359
CB966082
CB966206
CB966071
CB966070
CB966065
CB966062
CB966061
CB967488
CB966122
CB966247
CB966334
CB966328
CB966306
CB966292
CB966290
CB966200
CB966261
CB966202
CB966232
CB966221
CB966219
CB966218
CB966208
CB966350
CB966264
CB964484
BU673724
BU673857
CB964519
BU673903
BU673821
CB964532
BU673744
CB964448
BU673858
CB964421
BU673865
BU673874
BU673877
BU673911
BU673880
BU673885
BU673891
BU673893
BU673899
BU673545
BU673565
BU673566
BU673568
CB965170
CB965116
BU673852
BU673571
BU673577
BU673579
CB965158
BU673586
BU673606
CB964513
BU673636
CB964534
BU673721
BU673700
CB964538
BU673619
BU673694
CB965155
BU673690
BU673689
BU673673
BU673670
BU673831
BU673842

The initial comparison of IRGSP rice genomic sequences with the draft rice genomes revealed that truncated genes and large deletions. The availability of the accurate complete rice genome in near future will reveal the true gene transcript among the novel ESTs. The novel gene identification through tag to gene assignment was observed in recent SAGE analysis in *Arabidopsis* (Fizames et al., 2004).

6. Summary

- ESTs are very valuable genetic resources for gene discovery in rice. Particularly, ESTs generated from normalized cDNA libraries from drought stressed rice seedlings are found to be useful to identify drought stress responsive genes besides novel genes.
- Annotation and mapping of the ESTs onto the genomic sequences resulted, in defining putative functions and corresponding genomic regions for large number of ESTs.
- Physical map locations of the ESTs revealed the presence of putative stress responsive genes in the target Quantitative Trait Loci (QTLs) associated with drought tolerance.
- The ESTs which were mapped onto *indica* whole genome shotgun contigs (WGS) but not localized on rice BAC/PAC clones in Nipponbare genome, will form a useful resource for screening the genomic libraries for identification of corresponding clones and eventually leads to gap filling in contig generation.
- The novel ESTs constituting as much as 22 % of the total ESTs have not showed significant homology in the nucleotide and protein databases. Some of them may represent true gene transcripts derived from complex regions that might have been missed in earlier sequencing programs.
- Structural organization of a large number of genes associated with drought stress response in rice has been analyzed. The preliminary analysis revealed the complexity in rice gene organization and formed an initial training set of sequences for different gene prediction programs.
- Promoter analysis of a few targeted stress responsive genes revealed the presence of known stress responsive motifs. Interestingly, such motifs are missing in few genes, which were categorized earlier as stress responsive genes. This probably points to the existence of other novel *cis-* acting elements associated with stress response in rice.

- A comprehensive database has been developed for this EST clone resource, and the information related to putative functions, redundancy, chromosomal location, and trait association. The database offered valuable information source towards utilization of the rice drought EST data set for further analysis.
- This EST library formed a rich source of drought stress-related genes represented in GenBank for the first time from *indica* rice seedlings subjected to progressive drought, and used for comprehensive analysis of the stress-response transcriptome by comparing with the expression profiles from cDNA microarrays of different plants subjected abiotic stress treatments.
- A large number of transcription factors have been uncovered with suggested role in stress response in rice. As an example DREB family transcription factors, which are involved in the regulation of multiple stress response pathways, were analysed in detail. DREB genes from rice exhibited significant sequence similarity with that of *Arabidopsis* DREB gene family in their organization and response to abiotic stress. Two members of this family were cloned and characterized.
- Annotated EST set provides a useful resource for microarray and functional genomic studies in rice.
- The present study identified a large number of potentially interesting genes towards improving drought tolerance in rice.
- These resources will be useful in discovery of candidate genes, SNP analysis and eventual deployment in crop breeding for genetic improvement of rice for water limited environments.

References

1. Adams MD, Kelley JM, Gocayne JD, *et al.*, (1991) Complementary DNA sequencing: expressed sequence tags and human genome project. **Science. 21; 252:1651-1656**.
2. Adams MD, Kerlavage AR, Fleischmann RD, Fuldner RA, Bult CJ, *et al.*, (1995) Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence. **Nature 377: 3–174**.
3. Altschul, Stephen, F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. **Nucleic Acids Res. 25: 3389–3402**.
4. Apse MP, Aharon GS, Snedden WA, Blumwald E (1999) Salt tolerance conferred by overexpression of a vacuolar Na+/H+ antiport in Arabidopsis. **Science 285:1256–1258**.
5. Babu PR, Sekhar AC, Ithal N, Markandeya G, Reddy AR. (2002) Annotation and BAC/PAC localization of nonredundant ESTs from drought-stressed seedlings of an *indica* rice. **J. Genet. 81:25-44**.
6. Babu R. C et al. (2003) Genetic Analysis of Drought Resistance in Rice by Molecular Markers: Association between Secondary Traits and Field Performance. **Crop Sci. 43:1457–1469**.
7. Bailey, L.C. Jr, Searls, D.B., Overton, G.C., (1998). Analysis of EST-driven gene annotation in human genomic sequence. **Genome Res. 8: 362–376**.
8. Barker R, Herdt RW. (1985) The rice economy of Asia. Washington, D.C. (USA): Resources for the Future, Inc. 324 p.
9. Bevan *et al.*, (1998) Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of *Arabidopsis thaliana*. **Nature: 391, 485 – 488**.
10. Bhuiyan SI. (1992). Water management in relation to crop production: case study on rice. **Outlook Agric. 21(4): 293-299**.
11. Boguski, M.S., Lowe, T.M., Tolstoshev, C.M., (1993). dbEST — database for 'expressed sequence tags'. **Nat. Genet. 4:332–333**.
12. Boheler KR, Stern MD (2003) The new role of SAGE in gene discovery. **Trends Biotechnol 21: 55–57**.
13. Bohnert HJ, Ayoubi P, Borchert C, Bresson RA, Burnap RL, Cushman JC, Cushman MA, Deyholos M, Fischer R, Galbraith DW *et al.*, (2001) A genomics approach towards salt stress tolerance. **Plant Physiol Biochem, 39:295-311**.
14. Bohnert, H., Fischer, R., Kawasaki, S., Michalowski, C., Wang, H., Yale, J. and Zepeda, G. (2000). Cataloging stress-inducible genes and pathways leading to stress tolerance. In: Ribaut, J.-M. and Poland, D. (eds.) Molecular Approaches for the Genetic Improvement of Cereals for Stable Production in Water-Limited Environments. A Strategic Planning Workshop held at CIMMYT, El Batan, Mexico, 21-25 June 1999, CIMMYT, Mexico, D.F., Mexico, pp. 156-161.
15. Bohnert, H.J., Nelson, D.E. and Jensen, R.G. (1995). Adaptations to environmental stresses. **Plant Cell 7: 1099-1111**.
16. Bonaldo, M.F. *et al.*, (1996) Normalization and subtraction: two approaches to facilitate gene discovery. **Genome Res 6: 791–806**.
17. Boon K, Osorio EC, Greenhut SF, Schaefer CF, Shoemaker J, Polyak K, Morin PJ, Buetow KH, Strausberg RL, De Souza SJ et al. (2002) An anatomy of normal and malignant gene expression. **Proc Natl Acad Sci USA 99: 11287–11292**.

18. Bray E, Bailey-Serres J, Weretilnyk E; (2000): Responses to abiotic stresses. In Biochemistry and Molecular Biology of Plants. Edited by Buchanan B, Gruissem W, Jones R Rockville: American Society of Plant Biologists pp.1158-1203.
19. Bray EA. (2002) Classification of genes differentially expressed during water-deficit stress in *Arabidopsis thaliana*: an analysis using microarray and differential expression data. **Ann Bot. 89:803-11**.
20. Bray, E.A. (1993). Molecular response to water deficit. **Plant Physiologist 103: 1035-1040**.
21. Bray, E.A. (1997). Plant responses to water deficit. **Trends in Plant Sci 2: 48-54**.
22. Brendel V, Kurtz S, Walbot V, (2002) Comparative genomics of Arabidopsis and maize: prospects and limitations **Genome Biology, 3(3): 1005.1- 1005.6**.
23. Byrne, J.A., Tomasetto, C., Garnier, J.M., Rouyer, N., Mattei, M.G. et al., (1995) A screening method to identify genes commonly overexpressed in carcinomas and the identification of a novel complementary DNA sequence. **Cancer Res. 55: 2896–2903**.
24. Cassman KG (1999) Ecological intensification of cereal production systems: yield potential, soil quality, and precision agriculture. **Proc Natl Acad Sci USA 96: 5952–5959.**
25. Chen M, Presting G, Barbazuk WB, Goicoechea JL, Blackmon B. et al., (2002) An integrated physical and genetic map of the rice genome. **Plant Cell. 14: 537-45**.
26. Chen W, Provart NJ, Glazebrook J, Katagiri F, Chang HS, Eulgem T, Mauch F, Luan S, Zou G, Whitham SA, Budworth PR, Tao Y, Xie Z, Chen X, Lam S, Kreps JA, Harper JF, Si-Ammour A, Mauch-Mani B, Heinlein M, Kobayashi K, Hohn T, Dangl JL, Wang X, Zhu T (2002) Expression profile matrix of Arabidopsis transcription factor genes suggests their putative functions in response to environmental stresses. **Plant Cell 14:559–574**.
27. Cho RJ, Mindrinos M, Richards DR, Sapolsky RJ, Anderson M, Drenkhard E, Dewdney J, Reuber TL, Stammers M, Federspiel N, Theologis A, Yang W-H, Hubbell E, Au M, Chung EY, Lashkari D, Lemieux B, Dean C, Lipshutz RJ, Ausubel FM, Davis RW, Oefner PJ (1999) Genome-wide mapping with biallelic markers in Arabidopsis thaliana. **Nature Genet 23:203–207**.
28. Chomezynski, P. and Sacchi, N. (1987) Single-step method of RNA isolation by acid guanidium thiocyanate-phenol-chloroform extraction. **Anal.Biochem.162: 156-159**.
29. Chrast R, Scott HS, Papasavvas MP, Rossier C, Antonarakis ES, Barras C, Davisson MT, Schmidt C, Estivill X, Dierssen M et al. (2000) The mouse brain transcriptome by SAGE: differences in gene expression between P30 brains of the partial trisomy 16 mouse model of Down syndrome (Ts65Dn) and normals. **Genome Res 10: 2006–2021**.
30. Cooke R, Raynal M, Laudi ¨A M, Grellet F, Delseny M, *et al.*, (1996) Further progress towards a catalogue of all Arabidopsis genes:analysis of a set of 5000 non-redundant ESTs. **Plant J 9: 101–124**.
31. Cooke R, Raynal M, Laudie M, Delseny M (1997) Identification of members of gene families in Arabidopsis thaliana by contig construction from partial cDNA sequences: 106 genes encoding 50 cytoplasmic ribosomal proteins. **Plant J 11:1127–1140**.
32. Courtois B., et al., (2000) Mapping QTLs associated with drought avoidance in upland rice **Molecular Breeding 6: 55- 66**.
33. Cushman JC, Bohnert HJ (2000) Genomic approaches to plant stress tolerance. **Curr Opin Plant Biol 3:117–124**.

34. Davis GL, McMullen MD, Baysdorfer C, Musket T, Grant D, Staebell M, Xu G, Polacco M, Koster L, Melia-Hancock S, Houchins K, Chao S, Coe EH (1999) A maize map standard with sequenced core markers, grass genome reference points and 932 expressed sequence tagged site ESTs in a 1,736-locus map. **Genetics 152:1137–1172**.
35. de los Reyes BG, Morsy M, Gibbons J, Varma TS, Antoine W, McGrath JM, Halgren R, Redus M. (2003) A snapshot of the low temperature stress transcriptome of developing rice seedlings (*Oryza sativa* L.) via ESTs from subtracted cDNA library. **Theor Appl Genet. 107: 1071-82**.
36. Desikan R, Mackerness S, Hancock JT, Neill SJ (2001) Regulation of the Arabidopsis transcriptome by oxidative stress. **Plant Physiol 127:159–172**.
37. Devos KM, Gale MD (1997) Comparative genetics in the grasses **Plant Mol Biol. 35(1-2): 3-15**.
38. Dey, M.M., and Upadhyaya, H.K. (1996) Yield loss due to drought, cold, and submergence in Asia. *In* Rice research in Asia, progress and priorities. *Edited by* R.E. Evenson, R.W. Herdt, and M. Hossain. Oxford University Press, Cary, N.C. pp. 231–242.
39. Doebley J, Stec A, Hubbard L (1997). The evolution of apical dominance in maize. **Nature 386:485-488**.
40. Duvick DN, Cassman KG (1999) Post-green revolution trends in yield potential of temperate maize in the North-Central United States. **Crop Sci 39:1622–1630**.
41. Edmeades, G.O., Cooper, M., Lafitte, R., Zinselmeier, C., Ribaut, J.-M., Habben, J.E., Löffler, C. and Bänziger, M. (2001). Abiotic stresses and staple crops. In: J. Nosberger, H.H. Geiger, and P.C. Struik (eds). Crop Science: Progress and Prospects. Proceedings of the Third International Crops Science Congress, 17-21 August, 2000.
42. El-Din El-Assal S, Alonso-Blanco C, Peeters AJ, Raz V, Koornneef M (2001). A QTL for flowering time in Arabidopsis reveals a novel allele of CRY2. **Nat Genet, 29:435-440**.
43. Epple P, Apel K, Bohlmann H (1997) ESTs reveal a multigene family for plant defensins in Arabidopsis thaliana. **FEBS Lett 400:168–172**.
44. Ewing RM, Ben Kahla A, Poirot O, Lopez F, Audic S, Claverie JM (1999) Large-scale statistical analyses of rice ESTs reveal correlated patterns of gene expression. **Genome Res 9:950– 959**.
45. Ewing, B., and Green, P. (1998) Base calling of automated sequencer traces using phred. II. Error probabilities. **Genome Res. 8: 186–194**.
46. Ewing, B., Hillier, L., Wendl, M.C., Green, P., 1998. Basecalling of automated sequencer traces using phred. I. Accuracy assessment. **Genome Res. 8:175–185**.
47. Finkelstein, R.R. and Lynch, T.J. (2000). The Arabidopsis abscisic acid response gene ABI5 encodes a basic leucine zipper transcription factor. **Plant Cell 12: 599-610**.
48. Fizames C, Munos S, Cazettes C, Nacry P, Boucherez J. et al. (2004) The Arabidopsis Root Transcriptome by Serial Analysis of Gene Expression. Gene Identification Using the Genome Sequence. **Plant Physiol. 134:67-80**.
49. Florea, L., Hartzell, G., Zhang, Z., Rubin, G.M., Miller, W. (1998) A computer program for aligning a cDNA sequence with a genomic DNA sequence. **Genome Res. 8, 967-974**.
50. Fowler S, Thomashow MF (2002) Arabidopsis transcriptome profiling indicates that multiple regulatory pathways are activated during cold acclimation in addition to the CBF response pathway. **Plant Cell 14:1675–1690**.

51. Frary A, Nesbitt TC, Grandillo S, Knaap E, Cong B, Liu J, Meller J, Elber R, Alpert KB, Tanksley SD (2000). fw2.2: a quantitative trait locus key to the evolution of tomato fruit size. **Science 289:85-88**.
52. Fridman E, Liu YS, Carmel-Goren L, Gur A, Shoresh M, Pleban T, Eshed Y, Zamir D (2002). Two tightly linked QTLs modify tomato sugar content via different physiological pathways. **Mol Genet Genomics 266:821-826**.
53. Gai, X.W., Lal, S., Xing, L.Q., Brendel, V. and Walbot, V.(2000). Gene discovery using the maize genome database ZmDB. **Nucl. Acids Res. 28: 94–96**.
54. Giritch A, Ganal M, Stephan UW, Baumlein H. (1998) Structure, expression and chromosomal localisation of the metallothionein like gene family of tomato **Plant Mol Biol.** 37: 701–714.
55. Girke T, Todd J, Ruuska S, White J, Benning C, Ohlrogge J (2000) Microarray analysis of developing Arabidopsis seeds. **Plant Physiol 124:1570–1581**.
56. Goff SA, Ricke D, Lan TH, Presting G, Wang R, Dunn M, Glazebrook J, Sessions A, Oeller P, Varma H *et al.*, (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. japonica). **Science, 296:92-100**.
57. Gotoh O. (1982) An improved algorithm for matching biological sequences. **J Mol Biol. 162(3): 705-708**.
58. Guiltinan, M.J., Marcotte, W.R. Jr., Quatrano, R.S. (1990). A plant leucine zipper protein that recognizes an abscisic acid response element. **Science 250(4978): 267-271**.
59. Guo, H., and Moose, S.P., (2003) Conserved Noncoding Sequences among Cultivated Cereal Genomes Identify Candidate Regulatory Sequence Elements and Patterns of Promoter Evolution. **Plant Cell 15: 1143-1158**.
60. Harushima, et al. (1998). A high-density rice genetic linkage map with 2275 markers using a single F2 population. **Genetics 148: 479–494**.
61. Hasegawa, P.M., Bressan, R.A., Zhu, J.K. and Bohnert, H.J. (2000). Plant cellular and molecular responses to high salinity. **Annu.Rev. Plant Physiol. Plant Mol. Biol. 51: 463–499**.
62. Higo, K., Y. Ugawa, M. Iwamoto and T. Korenaga (1999) Plant cis-acting regulatory DNA elements (PLACE) database. **Nucl. Acids Res. 27: 297-300**.
63. Huang, X. and Madan, A. (1999) CAP3: A DNA Sequence Assembly Program. **Genome Res. 9: 868-877**.
64. Hugouvieux V, Kwak JM, Schroeder JI (2001) An mRNA cap binding protein, ABH1, modulates early abscisic acid signal transduction in Arabidopsis. **Cell 106:477–487**.
65. Ingram, J. and Bartels, D. 1996. The molecular basis of dehydration tolerance in plants. Annual Review Plant Physiology **Plant Mol Biol.** 47:377-403.
66. IRRI (International Rice Research Institute). (1989). IRRI toward 2000 and beyond. Manila (Philippines): IRRI.
67. Ismail, A.M., Hall, A.E. and Close, T.J. (1999). Allelic variation of a dehydrin gene cosegregates with chilling tolerance during seedling emergence. **Proc. Natl. Acad. Sci. USA 96: 13566–13570.**
68. Jaglo-Ottosen, K.R., Gilmour, S.J., Zarka, D.G., Schabenberger, O., and Thomashow, M.F. (1998). Arabidopsis CBF1 overexpression induces COR genes and enhances freezing tolerance. **Science, 280: 104–106**.
69. Jonak C, Okresz L, Bogre L, Hirt H (2002) Complexity, cross talk and integration of plant MAP kinase signalling. **Curr Opin Plant Biol 5:415–424**.

70. Jung SH, Lee JY, Lee DH (2003) Use of SAGE technology to reveal changes in gene expression in Arabidopsis leaves undergoing cold stress. **Plant Mol Biol 52: 553–567**.
71. Kasuga, M., Liu, Q., Miura, S., Yamaguchi-Shinozaki, K., and Shinozaki, K. (1999). Improving plant drought, salt, and freezing tolerance by gene transfer of a single stress-inducible transcription factor. **Nat. Biotechnol. 17: 287–291**.
72. Kato K., Miura H. and Sawada, S. (1999). Comparative mapping of the wheat Vrn-A1 region with the rice Hd-6 region. **Genome 42: 204–209**.
73. Kawasaki S, Borchert C, Deyholos M, Wang H, Brazille S, Kawai K, Galbraith D, Bohnert H (2001) Gene expression profiles during the initial phase of salt stress in rice. **Plant Cell, 13:889-905**.
74. Khush GS Origin, dispersal, cultivation and variation of rice **Plant Mol Biol. 35(1-2): 25-34**.
75. Kim H, Snesrud EC, Haas B, Cheung F, Town CD, Quackenbush J (2003) Gene expression analyses of Arabidopsis chromosome 2 using a genomic DNA amplicon microarray. **Genome Res 13:327–340**.
76. Klok EJ, Wilson IW, Wilson D, Chapman SC, Ewing RM, Somerville SC, Peacock WJ, Dolferus R, Dennis ES (2002) Expression profile analysis of the low-oxygen response in Arabidopsis root cultures. **Plant Cell 14:2481–2494**.
77. Kojima S, Takahashi Y, Kobayashi Y, Monna L, Sasaki T, Araki T, Yano M (2002). Hd3a, a rice ortholog of the Arabidopsis FT gene, promotes transition to flowering downstream of Hd1 under short-day conditions. **Plant Cell Physiol, 43:1096-1105**.
78. Kreps JA, Wu Y, Chang HS, Zhu T, Wang X, Harper J (2002) Transcriptome changes for Arabidopsis in response to salt, osmotic, and cold stress. **Plant Physiol, 130:2129-2141**.
79. Kurata N, Umehara Y, Tanoue H, Sasaki T (1997) Physical mapping of the rice genome with YAC clones. **Plant Mol Biol 35:101–113**.
80. Lash AE, Tolstoshev CM, Wagner L, Schuler GD, Strausberg RL, Riggins GJ, Altschul SF (2000) SAGE map: a public gene expression resource. **Genome Res 10: 1051–1060**.
81. Leung J, Giraudat J (1998) Abscisic acid signal transduction. **Annu Rev Plant Biol 49:199–222**.
82. Liang P (2002) SAGE Genie: a suite with panoramic view of gene expression. **Proc Natl Acad Sci USA 99: 11547–11548**.
83. Liang, P., Pardee, A.B., 1992. Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. **Science 257: 967–971**.
84. Liu J, Van Eck J, Cong B, Tanksley SD (2002). A new class of regulatory genes underlying the cause of pear-shaped tomato fruit. **Proc Natl Acad Sci USA, 99:13302-13306**.
85. Liu, Q., M. Kasuga, Y. Sakuma, H. Abe, S. Miura, K Yamaguchi-Shinozaki and 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 lowtemperature- responsive gene expression, respectively, in Arabidopsis. **Plant Cell 10:1391-1406**.
86. Marcotte, W.R. Jr, Russel, S.H. and Quatrano, R.S. (1989). Abscisic acid response sequences from the Em gene of wheat. **Plant Cell 1: 969–976**.
87. Matsumura H, Nirasawa S, Kiba A, Urasaki N, Saitoh H, Ito M, Kawai-Yamada M, Uchimiya H, Terauchi R (2003) Overexpression of Bax inhibitor suppresses the fungal elicitor-induced cell death in rice (*Oryza sativa* L.) cells. **Plant J 33: 425–434**.

88. Matsumura H, Nirasawa S, Terauchi R (1999) Transcript profiling in rice (Oryza sativa L.) seedlings using serial analysis of gene expression (SAGE). **Plant J 20: 719–726**.
89. McKersie BD, Bowley SR, Harjanto E, Leprince O (1996) Water-deficit tolerance and field performance of transgenic alfalfa overexpressing superoxide dismutase. **Plant Physiol. 111:1177–1181**.
90. Mekhedov, S., de Ilarduya, O.M. and Ohlrogge, J. (2000). Toward a functional catalogue of the plant genome. A survey of genes for lipid biosynthesis. **Plant Physiol. 122: 389–401**.
91. Misawa, S., Mori, N., Takumi, S., Yoshida, S. and Nakamura, C. (2000). Mapping of QTLs for low temperature response in seedling of rice (Otyza sativa L.). **Cereal Res. Comm. 28: 33–41**.
92. Mundy, J., Yamaguchi-Shinozaki, K. and Chua, N-H. (1990). Nuclear proteins bind conserved elements in the abscisic acid-responsive promoter of a rice rab gene. **Proc. Natl. Acad. Sci. USA 87: 1406–1410**.
93. Munnik T, Meijer HJ (2001) Osmotic stress activates distinct lipid and MAPK signalling pathways in plants. **FEBS Lett 498:172–178**.
94. Nguyen, H.T., Babu, R.C. and Blum, A. (1997). Breeding for drought resistance in rice: physiology and molecular genetics considerations. **Crop Sci. 37: 1426-1434**.
95. Ozturk ZN, Talame V, Deyholos M, Michalowski CB, Galbraith DW, Gozukirmizi N, Tuberosa R, Bohnert HJ (2002) Monitoring large-scale changes in transcript abundance in drought- and salt-stressed barley. **Plant Mol Biol 48:551–573**.
96. Palmiter RD, Findley SD, Whitmore TE, Durnam DM: (1992) MT III, a brainspecific member of the metallothionein gene family. **Proc Natl Acad Sci USA 89: 6333–6337**.
97. Peng S, Laza RC, Visperas RM, Sanico AL, Cassman KG, Khush GS (2000) Grain yield of rice cultivars and lines developed in the Philippines since 1966. **Crop Sci 40:307–314**.
98. Piquemal D, Commes T, Manchon L, Lejeune M, Ferraz C, Pugnere D, Demaille J, Elalouf J, Marti J (2002) Transcriptome analysis of monocytic leukemia cell differentiation. **Genomics 80: 361–371**.
99. Pleasance ED, Marra MA, Jones SJM (2003) Assessment of SAGE in transcript identification. **Genome Res 13: 1203–1215**.
100. Price et al. (2002) Mapping QTLs associated with drought avoidance in upland rice grown in the Philippines and West Africa. **Plant Mol Biol; 48: 683–695**.
101. Price, A.H. and Courtois, B. (2000.) Mapping QTLs associated with drought resistance in rice: progress, problems and prospects. **Plant Growth Regul. 29: 123–133**.
102. Rabbani MA, Maruyama K, Abe H, Khan MA, Katsura K. et al., (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**.
103. Rafalski A (2002). Applications of single nucleotide polymorphisms in crop genetics. **Curr Opin Plant Biol, 5:94-100**.
104. Reddy AR, Ramakrishna W, Sekhar AC, Ithal N, Babu PR, Bonaldo MF, Soares MB, Bennetzen JL (2002) Novel genes are enriched in normalized cDNA libraries from drought stressed seedlings of rice (*Oryza sativa* L. subsp. Indica cv Nagina 22). **Genome 45:204–211**.
105. Remington DL, Thornsberry JM, Matsuoka Y, Wilson LM, Whitt SR, Doebley J, Kresovich S, Goodman MM, Buckler ES iv (2001). Structure of linkage disequilibrium

and phenotypic associations in the maize genome. **Proc Natl Acad Sci USA, 98:11479- 11484**.

106. Rizhsky L, Liang H, Mittler R (2002) The combined effect of drought stress and heat shock on gene expression in tobacco. **Plant Physiol 130:1143–1151**.
107. Romero C, Belles JM, Vaya JL, Serrano R, Culianez-Macia FA (1997) Expression of the yeast trehalose-6-phosphate synthase gene in transgenic plants: pleiotropic phenotypes include drought tolerance. **Plant Physiol 201:293–297**.
108. Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press).
109. Schena M, Shalon D, Davis RW, Brown PO (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. **Science 270:467–470**.
110. Schenk PM, Kazan K, Wilson I, Anderson JP, Richmond T, Somerville SC, Manners JM (2000) Coordinated plant defense responses in Arabidopsis revealed by microarray analysis. **Proc Natl Acad Sci USA 97:11655–11660**.
111. Seki M, Narusaka M, Abe H, Kasuga M, Yamaguchi-Shinozaki K, Carninci P, Hayashizaki Y, Shinozaki K (2001) Monitoring the expression pattern of 1300 Arabidopsis genes under drought and cold stresses by using a full-length cDNA microarray. **Plant Cell 13:61–72**.
112. Seki M, Narusaka M, Ishida J, Nanjo T, Fujita M, Oono Y, Kamiya A, Nakajima M, Enju A, Sakurai T *et al.*, (2002) Monitoring the expression profiles of 7000 Arabidopsis genes under drought, cold, and high-salinity stresses using a full-length cDNA microarray **Plant J, 31:279-292**.
113. Shahmuradov IA, Akbarova YY, Solovyev VV, Aliyev JA. (2003) Abundance of plastid DNA insertions in nuclear genomes of rice and Arabidopsis. **Plant Mol Biol.; 52: 923-34**.
114. Sharp RE, Silk WK, Hsiao TC (1988) Growth of the maize primary root at low water potentials. I.spatial distribution of expansive growth. **Plant Physiol 87:50–57**.
115. Shen B, Carneiro N, Torres-Jerez I, Stevenson R, McGreery T, Helentjaris T, Baysdorfer C, Almira E, Ferl R, Habben J, Larkins B (1994) Partial sequencing and mapping of clones from two maize cDNA libraries. **Plant Mol Biol 26: 1085–1101**.
116. Shinozaki, K. and Yamaguchi-Shinozaki, K. (1996). Molecular responses to drought and cold stress. **Curr. Opin. Biotech. 7: 161-167**.
117. Shinwari ZK, Nakashima K, Miura S, Kasuga M, Seki M, Yamaguchi-Shinozaki K, Shinozaki K, (1998) An Arabidopsis gene family encoding DRE/CRT binding proteins involved in low-temperature-responsive gene expression. **Biochem Biophys Res Commun: 250(1): 161-70**.
118. Skriver, K. and Mundy, J. (1990). Gene expression in response to abscisic acid and osmotic stress. **Plant Cell 2:503-512**.
119. Smith, T.F., Waterman, M.S., 1981. Identification of common molecular subsequences. **J. Mol. Biol. 147: 195–197**.
120. Steinmetz LM, Sinha H, Richards DR, Spiegelman JI, Oefner PJ, McCusker JH, Davis RW (2002). Dissecting the architecture of a quantitative trait locus in yeast. **Nature 416:326-330**.
121. Takahashi Y, Shomura A, Sasaki T, Yano M (2001). Hd6, a rice quantitative trait locus involved in photoperiod sensitivity, encodes the subunit of protein kinase CK2. **Proc Natl Acad Sci USA, 98:7922-7927**.

122. Tarczynski MC, Jensen RG, Bohnert HJ (1993) Stress protection of transgenic tobacco by production of osmolyte mannitol. **Science 259:508–510**.

123. Thompson J. D., Gibson T. J., Plewniak F., Jeanmougin F. and Higgins D. G. (1997) The ClustalX Windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. **Nucl. Acids Res. 24, 4876–4882**.

124. Tollenaar M, Wu J (1999) Yield improvement in temperate maize is attributable to greater stress tolerance. **Crop Sci 39:1597–1604**.

125. Uchimiya, H., S. Kidou, T. Shimazaki, S. Aotsuka, S. Takamatsu, R. Nishi, H. Hashimoto, Y. Matsubayashi, N. Kidou, M. Umeda, and Kata. (1992) Random sequencing of cDNA libraries reveals a variety of expressed genes in cultured-cells of rice (*Oryza sativa*). **Plant J. 2: 1005–1009**.

126. Umeda, M., C. Hara, Y. Matsubayashi, H. Li, Q. Liu, F. Tadokoro, S.Aotsuka, and H. Uchimiya. (1994) Expressed sequence tags from cultured-cells of rice (*Oryza sativa*) under stressed conditions—analysis of transcripts of genes engaged in ATP-generating pathways. **Plant Mol. Biol. 25: 469–478**.

127. Van Deynze, A.E., Nelson, J.C., Yglesias, E.S., Harrington, S.E.,Braga, D.P., McCouch, S.R. and Sorrells, M.E. (1995) Comparative mapping in grasses: wheat relationships. **Mol. Gen. Genet. 248: 748–754**.

128. Velculescu VE, Zhang L, Vogelstein B, Kinzler KW (1995) Serial analysis of gene expression. **Science 270: 484–487**.

129. Velculescu VE, Zhang L, Zhou W, Vogelstein J, Basrai MA, Bassett DE Jr, Hieter P, Vogelstein B, Kinzler KW (1997) Characterization of the yeast transcriptome. **Cell 88: 243–251**.

130. Xu, Y. (2002) Global View of QTL: Rice as a Model. In Quantitative Genetics, Genomics and Plant Breeding Edited by M S Kang. CABI Publishing pp: 109-134.

131. Yamaguchi-Shinozaki, K. and Shinozaki, K. (1994). A novel cisacting element in an Arabidopsis gene is involved in responsiveness to drought, low temperature or high-salt stress. **Plant Cell 6: 251–264**.

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

133. Yano M, Katayose Y, Ashikari M, Yamanouchi U, Monna L, Fuse T, Baba T, Yamamoto K, Umehara Y, Nagamura Y, Sasaki T. (2000). Hd1, amajor photoperiod sensitivity quantitative trait locus in rice, is closely related to the Arabidopsis flowering time gene CONSTANS. **Plant Cell 12:2473-2484**.

134. Yu J, Hu S, Wang J, Wong GK, Li S, Liu B, Deng Y, Dai L, Zhou Y, Zhang X *et al.*, (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). **Science, 296:79-92**.

135. Yu LX, Setter TL (2003) Comparative transcriptional profiling of placenta and endosperm in developing maize kernels in response to water deficit. **Plant Physiol 131:568–582**.

136. Zhang, J., H. G. Zheng, A. Aarti G. Pantuwan, T.T. Nguyen, J.N. Tripathy, A.K. Sarial, S. Robin, R.C. Babu, B.D. Nguyen, S. Sarkarung, A. Blum, and H.T. Nguyen. (2001). Locating genomic regions associated with components of drought resistance in rice: Comparative mapping within and across species. **Theor. Appl. Genet. 103: 19–29**.

137. Zhu J.K, (2002) Salt and drought stress signal transduction in plants. **Annu Rev Plant Physiol Plant Mol Biol 53:247–273**.

138. Zhu T, Budworth P, Han B, Brown D, Chang HS, Zou G, Wang X (2001) Toward elucidating the global expression patterns of developing Arabidopsis: parallel analysis of 8300 genes by ahigh-density oligonucleotide probe array. **Plant Physiol Biochem, 39:221-242**.

Publications:

- **Babu, P.R**., Sekhar, A.C., Ithal, N., Markandeya, G., Reddy, A. R, (2002) Annotation and BAC/PAC localization of non-redundant ESTs from drought stressed seedlings of an *indica* rice J. Genet 81, 25-44.
- Reddy A. R., Ramakrishna W., Sekhar A. C., Ithal N., **Babu P. R**, Bonaldo M. F., Soares M. B. and Bennetzen J. L. (2002) Novel genes are enriched in normalized cDNA libraries from drought stressed seedlings of *indica* rice (*Oryza sativa* L subsp. *indica* cv. Nagina 22) Genome 45: 204-211.

Poster presented:

- **P. Ravindra Babu**, W. Ramakrishna, G. Markandeya, A.Chandra Sekhar, Nagabhushana Ithal, Jeffrey L. Bennetzen and Arjula R. Reddy Putative functions, BAC clone identification, and expression analysis of ESTs generated from cDNA libraries of drought stressed seedlings of *indica* rice (Oryza sativa L.): **Int Rice Congress** 16-20 September (2002) Beijing, China.

Genbank Submisssions:

- **Ravindra Babu,P.,** Markandeya,G., Chandra Sekhar,A. and Reddy,A.R. (2003) Isolation and characterization of a CRT/DRE-binding protein gene
  Accession No: *AY502052*.
- **Ravindra Babu,P**., Markandeya,G., Ramakrishna,W., Nagabhushana,I., Chandra Sekhar,A., Madana Mohan Reddy,A., Bennetzen,J.L. and Reddy,A.R. (2002). Isolation and characterization of mRNA of a DRE-binding protein 1B (DREB 1B) of indica rice. Accession No: *AY166833*.
- Markandeya, G.; **Ravindra Babu,P**.; Venkat Reddy, B.; Nagabhushana Ithal.; Chandra Sekhar, A.; Bennetzen,J.L.; Ramakrishna,W.; Reddy, A.R. (2003) ESTs from a normalized cDNA library of drought stressed rice seedlings (Oryza sativa L.cv Nagina 22) **www.ncbi.nlm.nih.gov/dbest.**
  Accession Numbers: *CB964418 to CB967504*.
- Reddy,A.R., Markandeya,G., Ramakrishna,W., Nagabhushana,I.,**Ravindra Babu,P**., Madana Mohan Reddy,A., Chandra Sekhar,A.,Bennetzen,J.L. (2002) Novel EST enrichment with normalized cDNA libraries from drought stressed rice (Oryza sativa L.cv Nagina 22) **www.ncbi.nlm.nih.gov/dbest.**
  Accession Numbers: *BU672765 - BU673915*.
- Reddy,A.R., Ramakrishna,W., Chandrasekhar,A., Nagabhushan,I., **Ravindrababu,P.,** Bennetzen,J.L. (2001) Novel EST enrichment with normalized cDNA libraries from drought stressed rice (Oryza sativa L.cv Nagina 22) **www.ncbi.nlm.nih.gov/dbest.** Accession Numbers: *BI305180 - BI306756*.