Characterization of full-length enriched expressed sequence tags of dehydration-treated white fibrous roots of sweetpotato

Sun-Hyung Kim1,3, Wan-Keun Song1,2, Yun-Hee Kim1, Suk-Yun Kwon1, Haeng-Soon Lee1, In-Chul Lee2 & Sang-Soo Kwak1,*
1Environmental Biotechnology Research Center, Korea Research Institute of Bioscience and Biotechnology (KIRIBB), Daejeon 305-333, 2Department of Biology, Daejeon University, Daejeon 300-716, 3Department of Environmental Horticulture, The University of Seoul, Seoul 130-743, Korea

Sweetpotato (Ipomoea batatas (L.) Lam.) is relatively tolerant to unfavorable growth conditions such as drought, yet has not been exploited to provide a better understanding of the molecular basis of drought stress tolerance. We obtained 983 high-quality expressed sequence tags of 100 bp or longer (average length of 700 bp) from cDNA libraries of detached white fibrous root tissues by subjecting them to dehydration for 6 h. The 431 cDNAs were each assigned a function by alignment using the BLASTX algorithm. Among them, three genes associated with various abiotic stresses and nine genes not previously associated with drought stress were selected for expression pattern analysis through detailed reverse transcription-polymerase chain reaction. The direct and indirect relationships of the 12 genes with drought tolerance mechanisms were ascertained at different developmental stages and under various stress conditions. [BMB reports 2009; 42(5): 271-276]

INTRODUCTION

Drought has become an increasingly important constraint to the stability of crop productivity and yield worldwide. It is one of the major environmental stresses in agriculture (1). To date, numerous physiological drought stress-responsive traits including stress-inducible genes have been isolated in various plants (2, 3). Hundreds of genes are involved in drought stress responses and/or tolerance from plant species such as Arabidopsis (4) and rice (5).

Sweetpotato is the seventh most important crop in the world in terms of dry matter production, providing not only staple food, but also important raw industrial materials (6). Recent sweetpotato research has mainly focused on storage root induction conditions (7, 8). However, in contrast to plants such as rice, almost no physiological or molecular information is available on drought stress in sweetpotato, even though sweetpotato is known to be a relatively drought-tolerant plant. Sweetpotato requires a period of moist conditions after planting until the root system becomes fully developed. Thereafter, it will tolerate for brief periods of drought stress, recovering quickly when soil moisture is restored. Prolonged dry conditions will restrict root development and limit photosynthetic ability. As well, storage root development may be impeded, leading to shorter or misshapen roots (9). Thus, understanding the molecular mechanism of drought tolerance is important for improving the yield and quality of sweetpotato crops.

Full-length cDNA resources are also extremely useful, not only for gene annotation and the determination of transcriptional start sites, but also for functional analyses (10). Although sweetpotato expressed sequence tags (ESTs) have been deposited into a database, they appear, to a large extent, to be ‘partial’ rather than ‘full-length’ (7, 8).

In this study, we have isolated and characterized ESTs from a full-length enriched cDNA library from white fibrous roots of sweetpotato subjected to dehydration stress.

RESULTS AND DISCUSSION

Analysis of EST sequencing

The cDNA library was constructed from pooled RNA samples isolated from white fibrous roots of sweetpotato at 6 h after drought treatment. Bacterial colonies were randomly picked and 1455 single-pass sequence reactions were performed on cDNAs present in plasmids from these clones. From these reactions, 983 high-quality ESTs of 100 bp or longer (average sequence of 900 bp) were obtained. Cluster analysis revealed the 47 most redundant transcripts having three or more EST copies in each cluster that comprised 172 ESTs. The 773 clones of the library included all of the unique clones of the cDNA fragments, which consisted of 649 singleton sequences and 124 assembled sequences. The sequences of 289 clones were assembled into 124 clusters, with a redundancy of 29%. The 431 cDNAs were assigned a function by aligning them with the translated sequences of the GenBank nucleotide sequence database using the BLASTX algorithm with an expectation value (E-value) cut off at 10^-4 or lower, a limit which sug-
ESTs from sweetpotato roots
Sun-Hyung Kim, et al.

Fig. 1. Overview of the functional classification of the sweetpotato ESTs.

suggests that the aligned genes are either the same gene or belong to the same gene family.

The ESTs represent a large range of functional categories, which include signal transduction (16.17%), metabolism (15.80%), protein destination (12.55%), transport facilitation (10.59%), protein synthesis (9.94%), and other functions (Fig. 1). This probably reflects the stress response of sweetpotato. For unique cDNA fragments, database comparisons of the cDNA GenBank non-redundant database using BLAST revealed that 431 of the 773 cDNAs showed a high degree of sequence similarity to genes from other organisms. The remaining 342 clones did not meet the criteria for a match (E-value cut-off at e^-3); these clones might represent either previously uncharacterized sequences or fragments that are too short to reveal any significant identity.

EST analysis implicated in drought stress

Resistance to different stresses is accomplished through cell rescue, cell defenses, and signal transduction (11). Components of abiotic stress signal recognition and transduction pathways and transcription activators have recently been discovered (12, 13). To analyze the most abundant ESTs from the drought-induced cDNA library, we grouped the sequence using the National Center for Biotechnology Information (NCBI) database. The classification was based on the best homology match of BLASTX searches against Arabidopsis and other plant protein sequences; an E-value threshold of 1 x 10^-10 was used. In the functional classification sequences belonging to four categories, "signal transduction", "metabolism", "transport and transport facilitation", and "protein destination" were considered to be of great potential interest for drought acclimation. Together, these categories were represented by 593 entries, corresponding to more than 54.9% of all genes (Fig. 1). Interestingly, several drought-related proteins were among the most abundant of the ESTs (Supplemental Table 2).

The selected clones, each with a query start position upstream of the subject start position in the aligned region, were defined as 'full-length'. Based on the analysis of 12 selected clones, the average insert size was 1.5 kb. The quality of the cDNA library was evaluated with 21 selected clones; 70% contained a full-length insert.

Sequence comparison of these clones using BLASTX led to the identification of three known abiotic stress-related genes (contig 10, 35, and 99) and nine low-abundant transcripts encoding gene products (contig 17, 33, 89, 121, 16, 53, 58, 103, and 123) not previously associated with drought stress. These were selected for expression analysis under various stress conditions.

Metallothionein-like protein (contig #10) was the most abundantly expressed EST. Its expression (especially of type 1 metallothionein) is usually elevated in senescent leaves and under conditions of abiotic and biotic stress (14, 15). Such metallothionein-like proteins can selectively interact with toxic heavy metal ions (e.g., cadmium) in Arabidopsis (16).

Cluster #35 encodes an AP2/EREBP-type transcription factor, which functions as a transactivator (17, 18) when fused to the DNA-binding domain. AP2/EREBP transcription factors are involved in the stress response in the dehydration-responsive element-binding (DREB) subfamily (18) and are regulated by wounding and abiotic stresses (19, 20), suggesting the role of putative AP2/EREBP-type transcription factor from sweetpotato in plant stress mechanisms.

Another abundant cluster represented by cluster #17 encoded a small GTPase that displays high amino acid identity (> 95%) with other small GTPase proteins from rice (21) and tomato (22). Small GTPases have emerged as important molecular switches in plant signal transduction (21, 23, 24). They regulate a diverse array of cellular responses such as cell polarity/tip growth and the actin cytoskeleton (23, 24). Further interest in small GTPase has been fuelled by recent studies that have demonstrated a decisive role of small GTPase in the stress and defense mechanisms of plants (21).

S-adenosylmethionine synthetases (SAM) may function to ethylene synthesis in mustard (Brassica juncea) (25). Curiously, a cluster of three ESTs (#33) encodes SAM. In tomato seedlings, NaCl stress causes the accumulation of several SAMs. Putative SAM may possibly assume related roles during the dehydration process in sweetpotato.

Two clusters (#89, #121) encoded cyclophilins (Cyp), which are ubiquitous proteins with peptidylprolyl cis-trans isomerase activity (26). The abundance of steady-state Cyp mRNA increased in bean leaves infected with alfalfa mosaic virus, and under ethephon and salicylic acid treatments.

Other proteins of interest are the Na⁺/H⁺ antiporter proteins. A putative Na⁺/H⁺ antiporter was encoded by the Arabidopsis SOS1 gene. The activity of the Na⁺/H⁺ antiporter is affected by the action of a calcium sensor on a protein kin-
ase (27, 28). Na⁺/H⁺ antiporter proteins are either up- or down-regulated in response to various abiotic stresses (29).

Two clones (cluster #16) exhibited significant homology to P-protein, the expression of which usually occurs in response to arachidonic acid, which induces programmed cell death (30). That gene is known to be involved in the defense against pathogen infections. However, responses to abiotic stresses have not been reported.

Cluster #53 encoded a PnFL-2, which is found in cotyledons and leaves, but is not found in stems and roots of morning glory (*Pharbitis nil*). PnFL-2 in Pharbitis may be photoperiodically regulated and associated with photoperiodic events (31). However, there is no evidence that any of the PnFL-2 play roles in the stress response.

A putative tonoplast intrinsic protein (TIP) belonging to the aquaporin gene family with the MipK gene from *Mesembryanthemum crystallinum* was encoded by cluster #58, which is comprised of two ESTs. The expression level of that gene increased transiently in response to water deficit stress and ABA treatment (32). Similarly, a closely related TIP gene from *Olea europaea* was induced by water stress (Genbank accession number ABB76813). In contrast, mRNA expression of the TIP gene family is down-regulated under water deficit stress in Arabidopsis (33). However, details of the functional roles of these genes have not been reported in terms of their responses to various stress conditions.

The functions of 32.5% of the ESTs were uncharacterized. The many undefined clones (i.e., no match and no hit sequence in the NCBI database) may be unknown stress-related genes present in plants. These unknown genes offer an opportunity to study the stress mechanisms of the plants and, eventually, develop a new pathway for the investigation of plant stresses.

Reverse transcription-polymerase chain reaction (RT-PCR) confirmation for differentially expressed genes in various tissues

To understand the expression profile of drought-induced genes, the expression of 12 clusters of interest were further investigated by RT-PCR (Supplemental Table 1). All 12 genes were highly expressed in white fibrous roots, but transcripts of metallothionein-like type 1 protein (contig #10) were detected in storage roots. A very low level of this protein was evident in white fibrous roots using RT-PCR, thus confirming the sensitivity of the technique. Five of the 12 genes, namely contig #17 (ARF-like small GTPase), #33 (S-adenosyl-L-methionine synthetase), #99 (Na⁺/H⁺ antiporter), #16 (P-rich protein), and #103 (stress-related protein), also reached high levels in flowers, but were weakly expressed or undetectable in tuberous roots. Interestingly, tonoplast intrinsic protein (contig #58) could not be detected in flower, leaf, or stem tissue (Fig. 2A), demonstrating that expression is root-specific rather than meristem-specific. The RT-PCR analysis indicate that each gene is differently regulated in a tissue-specific manner.

If candidate genes have a direct role in leaf development, one might expect manifestation of different expression profiles based on leaf size. Therefore, we compared accumulation patterns for candidate gene transcripts, finding that a developmental gradient existed, in which younger leaves near the shoot tips differentiated later than older leaves located near the shoot apical meristems. *Tubulin* was used as a control.
shoot base. Transcript levels of four genes—contig #35 (AP-2 domain-containing protein), #17 (ARF-like small GTPase), #53 (PnFL-2), and #33 (S-adenosyl-L-methionine synthetase)—were tightly coordinated with this gradient, with levels declining near the basal leaf and peaking toward the tip (Fig. 2B). The expressions of contig #10 and #16 progressively increased during leaf growth. During leaf development, the expression profiles of these genes fell into two groups: (1) contig #10, #16, #35, #17, #33, #53, and #58, which were expressed in a stage-specific manner, and (2) contig #89, #99, #121, #103, and #123, which were constitutively expressed throughout the growth period. These results suggest that expression of each cluster is differentially regulated by leaf growth stage.

**RT-PCR under different stress conditions**

The expression profiles of 12 clusters of genes were determined under abiotic stresses such as chilling, NaCl exposure, heating, and drought in leaves or root tissues from whole plants. The expression pattern against drought stress was investigated in leaves and roots 0, 1, 2, 4, 8, 16, and 24 h after drought treatment (Fig. 3A). Three of the candidate genes are known and have been reported previously as responsive to abiotic or biotic stress in plants. These genes were metallothionein-like type 1 protein, AP-2 domain-containing protein, and Na+/H+ antiporter. The other nine genes have not been reported previously as stress-inducible genes in plants. All candidate genes responded to drought stress, but the pattern and intensity of induction were slightly different for each gene. The transcript levels of clusters #35, #89, #99, #103, and #123 were progressively increased in leaf tissues from 4 h after exposure to drought conditions, implying that this gene responds to drought stress. When exposed to drought stress, the expressions of clusters #35 and #89 increased continuously in leaves, but their expressions increased from 1 h after drought treatment and decreased in root tissues. These findings indicate that those genes may respond to drought stress in quite a complex manner in leaf and root tissues. The expression levels of clusters #33 and #53 were not changed significantly after drought treatment in leaf and root tissues. Even though all candidate genes responded to drought stress, their induction times and patterns varied indicating that the expression pattern of these genes is differently regulated in sweetpotato.

Drought stress-related genes are also highly induced by other abiotic stresses (29). Sweetpotato plants were subjected to chilling (6 h at 4°C), heating (6 h at 42°C), and NaCl stress (3 h at 300 mM) to examine the effect of other abiotic stresses on the expression patterns in leaf tissues. The candidate genes responded differently to chilling, heating, and NaCl stress in leaf tissues (Fig. 3B). Clusters #10, #99, #53, and #103 were strongly expressed under NaCl stress conditions. The transcript level of clusters #10, #35, #17, #99, #121, and #123 decreased upon exposure to heat stress, suggesting that those genes may be negatively regulated by high temperature. Interestingly, the expression level of cluster #121 was significantly decreased by temperature and NaCl stress, indicating that it may have an important function in the stress response (34).

In conclusion, the EST comparative analysis described here was successful in the identification of drought stress-related genes in sweetpotato. Genes that might possibly be related to the induction of other abiotic stresses (temperature and NaCl)
were also identified. Of the sweetpotato ESTs, 32.5% share no significant similarity to known sequences or show significant similarity to known genes of unknown function. Such unknown genes will also be important in understanding their mechanisms, and consequently, to develop new strategies to protect against drought stress. Gene discovery from ESTs can also serve as a rich source for genetic information regarding regulatory networks involved in the stress induction process in sweetpotato (35).

MATERIALS AND METHODS

Plant materials and cDNA library construction
Sweetpotato plants (Ipomoea batatas (L.) Lam. cv. White star) were grown in a growth chamber at 25 ± 3°C under a photoperiod cycle of 16 h dark/8 h light (600-700 μmol photons m⁻² s⁻¹). Detached white fibrous roots of sweetpotato dehydrated for 6 h were used as the plant material for cDNA library construction. The extraction of sweetpotato RNA and construction of a cDNA library were described previously (35).

DNA sequencing and analysis
Colonies were randomly selected and cultured, and plasmid DNAs were subsequently purified using the AccuPrep Plasmid Extraction Kit (Bioneer, Republic of Korea). Sequence reactions were separated using a RISA 384 multi-capillary sequencer (Shimadzu, Japan). The ESTs were edited for removal of contaminating vector sequences and poor quality data. Sequences that were shorter than 100 bp were discarded. The CAP3 program was used to cluster the individual ESTs so that they represented unique transcripts. Consensus sequences of all clusters were generated based on 75% homology over a minimum of 30 bp. The individual ESTs were compared to the sequences in the GenBank nr database using a BLASTX algorithm with a cut-off E-value of 10⁻¹⁴. The functions of the ESTs were assigned by comparing them to the NCBI database.

RT-PCR
Sweetpotatoes were grown in a growth chamber at 25 ± 3°C under 16 h dark/8 h light (600-700 μmol photons m⁻² s⁻¹). For stress treatment, sweetpotato leaf and root tissues were collected after 0 (untreated control), 1, 2, 4, 8, 16, and 24 h of dehydration. For treatment with 300 mM NaCl, the plants were incubated in bottle wrapping with aluminum foil at 25°C for 3 h. For temperature treatment, the plants were kept at 4°C or 42°C for 6 h. All treated samples were immediately frozen in liquid nitrogen and stored at -70°C until further use. Quantitative RT-PCR was performed as described previously (35). The sequences of the forward and reverse primers and expected size of PCR products are listed in supplemental Table 1.

Acknowledgment
This work was supported by the Korea Foundation for International Cooperation of Science & Technology (KICOS), Biogreen 21 Program (20070301034015), Bioenergy Program (20070201030040) and the KIRIBB initiative program.

REFERENCES


