Identification of genes possibly related to storage root induction in sweetpotato

Min Kyoung You[a], Cheol Goo Hur[b], Young Sup Ahn[c], Mi Chung Suh[a], Byeong Choon Jeong[c], Jeong Sheop Shin[a], Jung Myung Bae[a,*]

[a] *Graduate School of Biotechnology, Korea University, Seoul 136-701, South Korea*
[b] *National Center for Genome Information, Korea Research Institute of Bioscience and Biotechnology, Daejon 305-333, South Korea*
[c] *Mokpo Experiment Station, National Honam Agricultural Experiment Station, RDA, Muan, Jeollanam-do 534-833, South Korea*

Received 7 November 2002; revised 2 January 2003; accepted 3 January 2003

First published online 17 January 2003

Edited by Ulf-Ingo Flügge

Abstract To identify genes related to initiation of storage root development in sweetpotato, a cDNA library was constructed with early stage storage roots (0.3–1 cm in diameter). Single-pass sequences of the 5′ ends of 2859 sweetpotato cDNA clones were assembled into 483 clusters and 442 singletons. Comparison of sweetpotato expressed sequence tags (ESTs) to nodulation/tumorigenesis-related sequence databases (nodule-, tumor-, potato tuber- and development-related sequences) revealed that homologs of 39 sweetpotato EST sequences potentially involved in gene regulation, signal transduction and development were present in at least one of the nodulation/tumorigenesis-related sequence databases. Northern blot analyses of these 39 sequences identified 22 differentially expressed genes in early stage storage root and fibrous root. These differentially expressed genes will be potential candidates for research to elucidate the molecular processes related to sweetpotato storage root induction.

Key words: Comparative analysis; Expressed sequence tag; Storage root; Storage root induction; Sweetpotato; *Ipomoea batatas*

1. Introduction

Sweetpotato (*Ipomoea batatas*) is one of the world's important food crops. With more than 133 million tons in annual production, sweetpotato currently ranks as the fifth most important food crop, on a fresh-weight basis, in developing countries after rice, wheat, maize, and cassava. Sweetpotato storage root is a commercially valuable organ that provides a high level of biomass and nutrients per hectare.

Sweetpotato forms colorless fibrous roots in the early stage of root development. Some of the colorless fibrous roots thicken, produce skin colors and develop into storage roots. Research has been focused on the effects of soil temperature, humidity, light intensity, photoperiod and carbon dioxide on the thickening growth and/or yield of sweetpotato storage roots [1–7] and studies on the shape and/or color of the storage root can also be found [4,8]. Nevertheless, almost no physiological or molecular information is available on the storage root-induction conditions in sweetpotato. In the case of the potato tuber, another well-known storage organ, tuber initiation is induced by short days, high light intensity, and high sucrose levels, and inhibited by high nitrogen levels, high temperatures and gibberellic acid [9]. However, the fact that the potato tuber develops from the stolon, underground stem, not from the root suggests that distinct induction mechanisms are possibly involved in potato tuber and sweetpotato storage root.

Large-scale single-pass sequencing of cDNAs generated from specific tissue has efficiently aided discovery of genes related to specific metabolic or signaling pathways in specific tissues. To facilitate the gene discovery process, a specific library was often enriched with specific transcripts by eliminating abundant non-specific sequences and/or the library was compared against an expressed sequence tag (EST) database to mine putative target genes [10].

As the first step to gain insight into the molecular processes occurring in the initiation of storage root development in sweetpotato, large-scale single-pass sequencing was employed with cDNA clones of early stage storage root. The 2859 partial sweetpotato sequences were obtained. Among them, 39 potential gene regulation-, signal transduction- and development-related ESTs were found to be present in at least one of the nodulation/tumorigenesis-related sequence data sets. These 39 sequences were used to identify the genes transcriptionally regulated during the early stage of storage root development.

2. Materials and methods

2.1. Plant material

Sweetpotato cultivar (*I. batatas* cv. Jinhongmi) was obtained from the Mokpo Experiment Station, National Honam Agricultural Experiment Station, RDA, Korea. Plants were grown in the greenhouse at 28 ± 3°C.

2.2. cDNA library construction and sequencing

Total RNA was extracted from young storage roots (SRs) (0.3–1.0 cm in diameter) using the modified method with guanidinium–SDS lysis buffer [11] and the CsCl gradient method [12]. Poly(A)$^+$ RNA was prepared with Poly(A) Track mRNA isolation system (Promega, Madison, WI, USA). A cDNA library was constructed using the ZAP-cDNA synthesis kit and the ZAP-cDNA Gigapack®III Gold packaging extract (Stratagene, La Jolla, CA, USA) according to the manufacturer's instructions. After in vivo mass-excision of a cDNA library, more than 5000 colonies were randomly picked and cultivated

*Corresponding author. Fax: (82)-2-927 9028.
E-mail address: jmbae@korea.ac.kr (J.M. Bae).

Abbreviations: EST, expressed sequence tag; MTD1, methylenetetrahydrofolate dehydrogenase1

0014-5793/03/$22.00 © 2003 Published by Elsevier Science B.V. on behalf of the Federation of European Biochemical Societies.
doi:10.1016/S0014-5793(03)00035-8

for storage of the colonies and the isolation of plasmid DNA. Plasmid DNA was purified with an AccuPrep™ Plasmid Extraction kit (Bioneer, Nami, Cheongwon-gun, Korea). cDNA inserts were amplified by PCR with T3 primer (5′-AATTAACCCTCACTAAAGGG-3′). PCR products were cleaned with Sephadex G50 (DNA grade, Amersham Pharmacia, Piscataway, NJ, USA) in filter plates (MultiScreen-HV plates, Millipore, Bedford, MA, USA). Sequence reactions were run on an automated DNA sequencer RISA-384 (Shimadzu, Nakagyo, Kyoto, Japan).

2.3. Sequence processing and analysis

Sequence data was analyzed using the SGI Origin 3200 Unix machine (SGI, Mountain View, CA, USA). The ABI formatted chromatogram sequences were processed using PHRED [13]. Sequences that were less than 200 bp or >4% ambiguous were not further processed. Before clustering of the ESTs, sequences were analyzed using RepeatMasker software (http://ftp.genome.washington.edu/RM/RepeatMasker.html) with default parameters to screen out the interspersed repeats, vectors and low-complexity DNA sequences. Possible non-sweetpotato and non-nuclear sequences were eliminated by searching for BLASTN matches for the ESTs with strong homology (below *E*-value at 10^{-10}) to mitochondrial, chloroplastidic, and ribosomal RNA genes using the Mendel DB (www.mendel.ac.uk.genomedb.html). To remove redundant ESTs, sequences were clustered and consensus sequence into groups that contained more than 100 bp of core sequence using StackPack S/W (provided by SANBI, http://www.sanbi.ac.za). Groups that contained only one sequence were classified as singleton.

2.4. Functional classification and comparative analysis

To assign the functions of ESTs, sweetpotato sequences were aligned to the GenBank nucleotide sequence database using the BLASTX algorithm with an *E*-value cut off at 10^{-10} or lower. Based on the BLASTX comparison results, sweetpotato ESTs were classified according to their predicted function.

For comparative analysis, we performed a BLASTN search against 5719 potato tuber ESTs (extracted from NCBI Solanum EST) and 41 558 development-related ESTs (extracted from NCBI EST others) and a BLASTX search against 66 nodule-related protein sequences (extracted from NCBI nodule-related database) and tumor-related protein sequences (provided by Weizmann, GeneCard).

2.5. Northern analysis

25 μg total RNA was denatured, electrophoresed and then transferred onto Tropilon-Plus™ (Tropix, Bedford, MA, USA) nylon membrane using the downward alkaline capillary method [14]. Biotin-labeled probes were prepared by PCR amplification. PCR conditions included 95°C 3 min for predenaturation and then followed by 30 cycles of 95°C 30 s, 60°C 30 s, 72°C 60 s using dNTP mixed with biotin-labeled dCTP (Invitrogen, Carlsbad, CA, USA). SK-primer as 5′ primer and T7-primer as 3′ primer were used. Labeled probes were purified using a PCR purification kit (Qiagen, Valencia, CA, USA) according to the manufacturer's instructions. The membranes were hybridized at 60°C for 16 h and washed twice with 0.1×SSC/1% SDS at 60°C for 15 min, then hybridized bands were detected using the Southern-Star™ system (Tropix, Bedford, MA, USA).

3. Results and discussion

3.1. Sequencing and clustering of sweetpotato ESTs

A sweetpotato cDNA library was constructed with the mRNAs extracted from early stage storage roots (0.3–1.0 cm in diameter). Initially, 5′-end sequences of 3159 cDNA clones were obtained and the sequence process gave rise to 2859 high quality ESTs. The average read length of these ESTs, after vector trimming and removal of low quality sequences, was 400 bp. The average insert size of the corresponding ESTs was 800 bp. The 2859 sweetpotato EST sequences of early stage storage root have been submitted to the dbEST and GenBank databases (accession numbers BU690119 to BU692977).

To identify ESTs that belong to the same gene, 2859 ESTs were clustered into groups. The storage root ESTs represent up to 925 independent genes, 483 clusters assembled from more than two ESTs and 442 singletons. Relatively high (84.5%) redundancy (number of ESTs assembled in clusters/total number of ESTs) and the low number of active genes in the storage root probably reflect the characteristics of non-photosynthetic storage organs and/or the characteristics of genome organization in autohexaploids. The ESTs of mature potato tuber, the storage organ of the potato, also showed high levels of redundancy (74.8%) [15].

3.2. Functional categorization of sweetpotato ESTs

Approximately 71.7% of the sweetpotato ESTs were assigned a function by aligning them with the translated sequences of the GenBank nucleotide sequence database using the BLASTX algorithm with an *E*-value cut off at 10^{-10} or lower. A relatively large number of sequences (28.3%) were similar to proteins of unknown functions or produced no match in the database. When these sweetpotato ESTs were classified using the functional classification of the Munich Information Center for Protein Sequences for *Arabidopsis thaliana*, only 32% of ESTs were classified and 68% of ESTs were revealed as unclassifiable. Thus, the functional classification of sweetpotato ESTs was carried out according to the previously reported classification [16–21] with the predicted function (Fig. 1). Sweetpotato storage root consisted of a relatively high percentage of genes involved in metabolism (20.5%) and protein metabolism (9.9%) and a relatively low percentage of sequences in cellular organization (2.0%) and development (1.9%). A relatively large fraction of sweetpotato sequences were classified as genes involved in cellular gene expression (9.6%) and signal transduction (5.4%). This probably reflects the transitional growth phase of fibrous root to storage root. It is noteworthy that 1% of ESTs showed similarities with genes of nodulation- or tumorigenesis-related proteins such as early nodulin 93 (ENOD93), early nodulin 18 (ENOD18), early nodulin binding protein 1, methylenetetrahydrofolate dehydrogenase1 (MTD1), tumor-related protein and translationally controlled tumor protein. These were classified as nodule or tumor development. The ESTs in this category may possess a high potential to be involved in sweetpotato storage root formation.

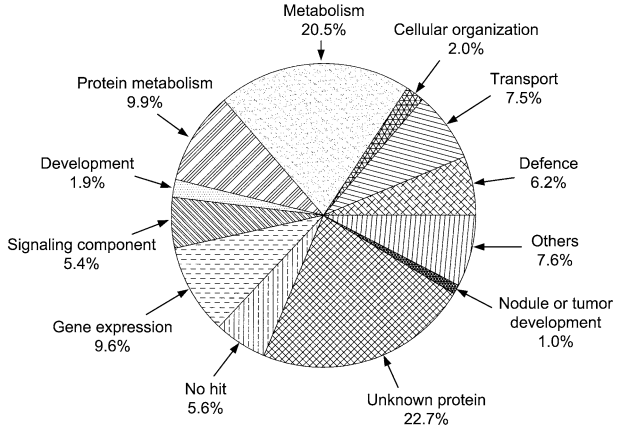

Fig. 1. Functional classification of the sweetpotato storage root ESTs.

Table 1
Each sweetpotato root used in Northern analysis

Root	Description
FRN (*F*ibrous *R*oots of *N*on-storage root stage)	Colorless fibrous roots of sweetpotato plants on which no fibrous root has begun to develop into storage root
SR (*S*torage *R*oots)	SRs whose diameters range from 0.3 to 1.0 cm
FRS (*F*ibrous *R*oots of *S*torage root stage)	Colorless fibrous roots of sweetpotato plants on which some of fibrous roots have begun to develop into storage roots

3.3. Identification of genes possibly related to storage root induction

During the early stages of storage root formation, the fibrous roots having the latency to develop into the storage roots alter their growth pattern. They display a cessation of elongation growth and the initiation of radial growth. Then, cells are rapidly divided and expanded. These characteristics of storage root formation are also common features found in nodulation/tumorigenesis-related tissues such as nodule formation in legumes, tumor development in animals and tuber development in potato. Thus, sweetpotato EST data was compared to nodule-, tumor-, potato tuber- and development-related sequence data sets to screen genes involved in storage root induction.

By comparison of sweetpotato ESTs to nodulation/tumorigenesis-related sequence databases, it was found that homologs of 39 sweetpotato EST sequences potentially involved in gene regulation, signal transduction and development, including nodule or tumor development, were present in at least one of the nodulation/tumorigenesis-related sequence databases (10 were in nodule-, 12 tumor-, seven potato tuber- and 21 development-related databases). Northern analyses of these 39 sequences identified three distinct groups based on the differential expression patterns in fibrous root of non-storage root stage (Table 1 and Fig. 2D), fibrous root of storage root stage (FRS) (Table 1 and Fig. 2D) and SR (Table 1 and Fig. 2D). In the first group are genes whose transcription was up-regulated in SR (Fig. 2A), transcription of genes in the second group was down-regulated in SR (Fig. 2B) and the transcription of genes of the third group was increased in FRS (Fig. 2C). Predicted functions of genes in each group are listed in Table 2.

Transcription of seven sweetpotato genes was elevated in SR. Transcription of the putative *GIGANTEA* gene (BU690683) and J8-like protein (BU692118) was strongly induced in SR. The *Arabidopsis GIGANTEA* gene was reported to be related to the controlling of the circadian rhythms and photoperiodic flowering time [22,23]. More recently, it was found that it is a nuclear protein and involved in phytochrome signaling [24]. Although at present, it cannot be determined if the sweetpotato *GIGANTEA* homolog is possibly involved in controlling the induction time of storage root in the non-photosynthetic tissue, it will be one of the promising candidates for future research to elucidate the regulatory process occurring upon the initiation of storage root development. J8-like protein, known as calmodulin antagonist, was reported to inhibit human ocular melanoma cell invasion by reducing the attachment of melanoma cells to matrix proteins [25,26]. However, the function of the J8 protein homolog in plants has not been elucidated. The putative nodulin-like protein (ENOD93) (BU690467), late embryogenesis abundant protein 5 (LEA5) (BU690505), calmodulin-like protein (BU690910), dermal glycoprotein precursor (BU692109) and transcription factor-like protein (BU692562) were also up-regulated in SR. This result suggests that calmodulin-dependent signal transduction and nodulation-related genes are possibly related to storage root induction processes in sweetpotato. The roles of these genes in storage root induction remain to be investigated.

The transcription levels of 10 sweetpotato genes were decreased in SR in that expression of the putative MADS-box protein (BU691821) was severely suppressed. Recently, two MADS-box protein genes (*ibMADS3* and *ibMADS4*) were isolated from sweetpotato and shown to express preferentially in vegetative tissues, especially root tissues: white fibrous

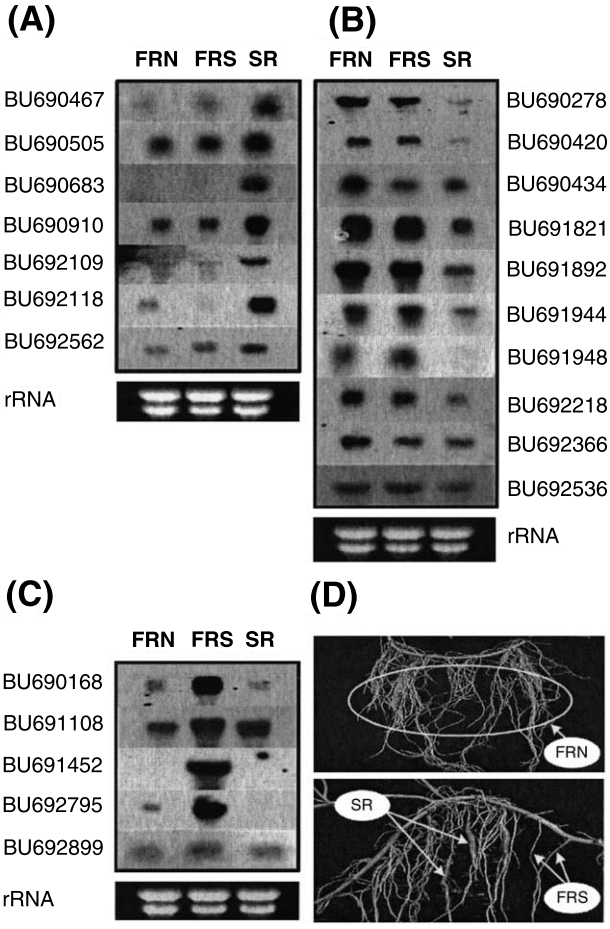

Fig. 2. RNA blot analysis of the expression patterns of selected EST clones. Total RNA was isolated from fibrous root of non-storage root stage (FRN), FRS and SR. The loading of an equal amount of total RNA in each lane was verified by ethidium bromide staining. A: Expression patterns of genes whose transcription was up-regulated in SR. B: Expression patterns of genes whose transcription was down-regulated in SR. C: Expression patterns of genes whose transcription was up-regulated in FRS. D: Illustration of each sample used in Northern analysis.

Table 2
Sweetpotato storage root ESTs transcriptionally regulated during early stage of storage root development

Accession number	Putative function	Organism[a]	*E*-value[b]
BU690467	ENOD93	*Glycine max*	1E-25
BU690505	Late embryogenesis abundant protein 5	*Nicotiana tabacum*	2E-10
BU690683	GIGANTEA	*A. thaliana*	1E-71
BU690910	Calmodulin-like protein	*A. thaliana*	2E-17
BU692109	Dermal glycoprotein precursor	Carrot	9E-22
BU692118	J8-like protein	*A. thaliana*	1E-26
BU692562	Transcription factor-like protein	*A. thaliana*	2E-30
BU690278	NAM-like protein	*A. thaliana*	1E-19
BU690420	Translationally controlled tumor protein	*Hordeum vulgare*	5E-30
BU690434	Hypothetical protein	*Cicer arietinum*	5E-36
BU691821	MADS-box protein	*I. batatas*	3E-93
BU691892	G10-like protein	*A. thaliana*	2E-16
BU691944	Glycine rich protein	*Nicotiana tabacum*	2E-18
BU691948	No match	-	-
BU692218	Developmental protein DG1118	*A. thaliana*	2E-16
BU692366	MTD1	*Medicago truncatula*	2E-19
BU692536	Putative proteasome regulatory subunit	*A. thaliana*	2E-87
BU690168	ENOD18	*Vicia faba*	4E-36
BU691108	MAP kinase kinase 4	*A. thaliana*	3E-49
BU691452	Expansin precursor	*Lycopersicon esculentum*	9E-67
BU692795	Tumor-related protein	*Nicotiana tabacum*	3E-36
BU692899	Extensin-like protein	*Populus nigra*	5E-21

[a]Organism that showed the highest similarity with corresponding EST clone.
[b]*E*-value released by the end of September, 2002.

roots, pigmented roots, and developing storage roots [27]. Transcripts of these two genes in roots were found in the vascular cambium region where the most active cell proliferation occurs during storage root development. Thus it was suggested that these two genes may be involved in initiation of storage root development. The sequence comparison of BU691821 to these MADS-box genes revealed that BU691821 is a new MADS-box gene rather than a homolog of *ibMADS3* or *ibMADS4*. The transcription of putative NAM (no apical meristem) -like protein (BU690278), translationally controlled tumor protein (BU690420), G10-like protein (BU691892), glycine-rich protein (BU691944), developmental protein DG1118 (BU692218), MTD1 (BU692366) and putative proteasome regulatory subunit (BU692536) was down-regulated in SR. The transcription levels of a hypothetical protein (BU690434) and a no-hit protein (BU691948) were also decreased in SR. The significance of transcriptional down-regulation of these genes in the early stage storage roots needs to be further examined.

Transcription of five sweetpotato genes was elevated in FRS. Transcription of the putative tumor-related protein (BU692795), ENOD18 (BU690168) and expansin precursor (BU691452) was strongly induced and the transcription levels of putative MAP kinase kinase 4 (BU691108) and extensin-like protein (BU692899) were moderately increased in FRS. The tumor-related protein (TID91) was isolated from tobacco genetic tumors, and it was highly expressed in tobacco callus and moderately expressed in genetic tumors of tobacco [28]. The functional role of TID91 in genetic tumor and callus formation has not been elucidated. The cDNA and gene of nodulin ENOD18 (*VfENOD18*) were isolated from broad bean [29]. The corresponding transcripts were detected in early and late stages of nodule development and localized exclusively in the nitrogen-fixing zone III. Expansins were first reported as cell wall proteins that mediate pH-dependent extension of the plant cell wall and growth of the cell. More recently, these proteins were found to be involved in a variety of plant processes such as morphogenesis, softening of fruits and growth of the pollen tube in grasses [30]. The structure, function and suggested mode of action of expansins have recently been reviewed [31]. The fact that transcription of tumor-related protein, ENOD18 and expansin precursor homologs is highly induced in FRS suggests the possibility that these genes are involved in the initiation of storage root development and molecular initiation of storage root development may already be triggered in FRS.

At present, the functional roles of these differentially expressed genes in the storage root induction process cannot be determined. These 22 differentially expressed genes, however, possibly play certain roles in storage root induction by regulating the processes, including cessation of elongation growth of fibrous root, initiation of radial growth of storage root and promotion of thickening growth by rapid cell division and expansion.

4. Conclusion

The EST data presented here is the first overview of genes that are expressed at the initiation stage of storage root development. The genes possibly related to the storage root induction were also identified. These genes can be exploited to unravel regulatory networks involved in the storage root induction process in sweetpotato and possibly in other rootcrops. This EST data will make it feasible for molecular breeders to develop new varieties of rootcrops with higher or more efficient production capacities.

Acknowledgements: This research was supported by a Grant (PF003202-04) from Plant Diversity Research Center of 21st Century Frontier Research Program funded by Ministry of Science and Technology of Korean Government.

References

[1] Eguchi, T., Kitano, M. and Eguchi, H. (1997) Biotronics 26, 67–72.
[2] Eguchi, T., Kitano, M. and Eguchi, H. (1998) Biotronics 27, 93–96.

[3] Pardales Jr., J.R., Bañoc, D.M., Yamauchi, A., Iijima, M. and Kono, Y. (1999) Plant Prod. Sci. 2, 247–251.
[4] Kano, Y. and Ming, Z.J. (2000) Environ. Control Biol. 38, 113–120.
[5] Loretan, P.A., Bonsi, C.K., Mortley, D.G., Wheeler, R.M., Mackowiak, C.L., Hill, W.A., Morris, C.E., Trotman, A.A. and David, P.P. (1994) Adv. Space Res. 14, 277–280.
[6] Hill, J., Douglas, D., David, P., Mortley, D., Trotman, A. and Bonsi, C. (1996) Acta Hortic. 440, 25–30.
[7] Mortley, D., Hill, J., Loretan, P., Bonsi, C., Hill, W., Hileman, D. and Terse, A. (1996) Acta Hortic. 440, 31–36.
[8] Yoshinaga, M., Tanaka, M. and Nakatani, M. (2000) Breed. Sci. 50, 59–64.
[9] Jackson, S.D. (1999) Plant Physiol. 119, 1–8.
[10] Ohlrogge, J. and Benning, C. (2000) Curr. Opin. Plant Biol. 3, 224–228.
[11] Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J. and Rutter, W.J. (1979) Biochemistry 18, 5294–5299.
[12] Glisin, V., Crkvenjakov, R. and Byus, C. (1974) Biochemistry 13, 2633–2637.
[13] Ewing, B., Hillier, L., Wendl, M.C. and Green, P. (1998) Genome Res. 8, 175–185.
[14] Chomczynski, P. (1992) Anal. Biochem. 201, 134–139.
[15] Crookshanks, M., Emmersen, J., Welinder, K.G. and Nielsen, K.L. (2001) FEBS lett. 506, 123–126.
[16] Ablett, E., Seaton, G., Scott, K., Shelton, D., Graham, M.W., Baverstock, P., Lee, L.S. and Henry, R. (2000) Plant Sci. 159, 87–95.
[17] Chim, S.S., Cheung, S.S.F. and Tsui, S.K.W. (2000) J. Cell. Biochem. 80, 24–36.
[18] Crépineau, F., Roscoe, T., Kaas, R., Kloareg, B. and Boyen, C. (2000) Plant Mol. Biol. 43, 503–513.
[19] Ujino-Ihara, T., Yoshimura, K., Ugawa, Y., Yoshimaru, H., Nagasaka, K. and Tsumura, Y. (2000) Plant Mol. Biol. 43, 451–457.
[20] Hegde, P., Qi, R., Gaspard, R., Abernathy, K., Dharap, S., Earle-Hughes, J., Gay, C., Nwokekeh, N.U., Chen, T., Saeed, A.I., Sharov, V., Lee, N.H., Yeatman, T.J. and Quackenbush, J. (2001) Cancer Res. 61, 7792–7797.
[21] Zhang, L., Ma, X.L., Zhang, Q., Ma, C.L., Wang, P.P., Sun, Y.F., Zhao, Y.X. and Zhang, H. (2001) Gene 267, 193–200.
[22] Park, D.H., Somers, D.E., Kim, Y.S., Choy, Y.H., Lim, H.K., Soh, M.S., Kim, H.J., Kay, S.A. and Nam, H.G. (1999) Science 285, 1579–1582.
[23] Fowler, S., Lee, K., Onouchi, H., Samach, A., Richardson, K., Morris, B., Coupland, G. and Putterill, J. (1999) EMBO J. 18, 4679–4688.
[24] Huq, E., Tepperman, J.M. and Quail, P.H. (2000) Proc. Natl. Acad. Sci. USA 97, 9789–9794.
[25] MacNeil, S., Wagner, M. and Rennie, I.G. (1994) Clin. Exp. Metastasis 12, 375–384.
[26] Dewhurst, L.O., Gee, J.W., Rennie, I.G. and MacNeil, S. (1997) Br. J. Cancer 75, 860–868.
[27] Kim, S.H., Mizuno, K. and Fujimura, T. (2002) Plant Cell Physiol. 43, 314–322.
[28] Fujita, T., Kouchi, H., Ichikawa, T. and Syono, K. (1994) Plant J. 5, 645–654.
[29] Hohnjec, N., Küster, H., Albus, U., Frosch, S.C., Becker, J.D., Pühler, A., Perlick, A.M. and Frühling, M. (2000) Mol. Gen. Genet. 264, 241–250.
[30] Lee, Y., Choi, D. and Kende, H. (2001) Curr. Opin. Plant Biol. 4, 527–532.
[31] Cosgrove, D.J. (2000) Nature 407, 321–326.