Theor Appl Genet (2005) 110: 416–424
DOI 10.1007/s00122-004-1801-y

ORIGINAL PAPER

Preeti A. Mehta · K. Sivaprakash · M. Parani
Gayatri Venkataraman · Ajay K. Parida

Generation and analysis of expressed sequence tags from the salt-tolerant mangrove species *Avicennia marina* (Forsk) Vierh.

Received: 1 June 2004 / Accepted: 19 August 2004 / Published online: 18 December 2004

Abstract Salinization poses an increasingly serious problem in coastal and agricultural areas with negative effects on plant productivity and yield. *Avicennia marina* is a pantropical mangrove species that can survive in highly saline conditions. As a first step towards the characterization of genes that contribute to combating salinity stress, the construction of a cDNA library of *A. marina* genes is reported here. Random expressed sequence tag (EST) sequencing of 1,841 clones produced 1,602 quality reads. These clones were classified into functional categories, and BLAST comparisons revealed that 113 clones were homologous to genes earlier implicated in stress responses, of which the dehydrins are the most predominant in this category. Of the ESTs analyzed, 30% showed homology to previously uncharacterized genes in the public plant databases. Of these 30%, 52 clones were selected for reverse Northern analysis: 26 were shown to be up-regulated and five shown to be down-regulated. The results obtained by reverse Northern analysis were confirmed by Northern analysis for three clones.

Introduction

Automated high-throughput DNA sequencing paved the way for the announcement of the completed version of the human genome sequence (The International Human Genome Sequencing Consortium 2001). Model organisms have also been sequenced in the plant kingdom, and currently more than 60 eukaryotic genome-sequencing projects are underway, including the *Arabidopsis*, *Oryza* and maize genome-sequencing projects (http://www.tigr.org/tdb/euk/). While genome-sequencing provides information on the total gene pool available in the organism, it does not reflect the functional and dynamic nature of gene regulation. In this regard the partial sequencing of anonymous cDNA clones (expressed sequence tags, ESTs) is a rapid and cost-effective method for generating data on the actual coding capacity of genomes. The top ten in rank order for the most abundant plant ESTs include wheat, maize, barley, soybean, rice sugarcane, *Arabidopsis*, *Medicago*, sorghum and tomato (dbEST release:032604; http://www.ncbi.nlm.nih.gov/dbEST/dbEST_summary.html). This abundance of sequence information presents opportunities to accelerate progress towards understanding genetic mechanisms that control plant growth and responses to the environment.

Environmental factors such as drought, extreme temperatures and high or fluctuating salinity can affect the plant growth and performance and, in the case of agronomically important plants, may translate to reduced yields. In particular, increasing soil salinization in irrigated areas has necessitated the identification of crop traits/genes that confer resistance to salinity, either by conventional breeding or through molecular biology techniques (Cushman and Bohnert 2000; Munns et al. 2002). Hyperosmotic stress, such as that caused by the exposure of cells to high concentrations of NaCl, causes imbalances in cellular ions, changes in turgor pressure and cell volume and alters the activity and stability of macromolecules. Although the basic cellular responses appear to be conserved among all plants, plant species employ a variety of mechanisms to cope with osmotic stress. While extensive work on salinity tolerance in *Arabidopsis* and *Mesembryanthemum* has led to the identification of candidate salinity-sensitive determinants, these plants are not true halophytes (Chauhan et al. 2000; Zhu 2002). *Avicennia* is a monotypic pantropical mangrove genus with eight species, of which *A. marina* is widely distributed both latitudinally and longitudinally. The high salt tolerance of *A. marina* is a consequence of water use efficiency, which balances the relation between carbon

Communicated by P. Langridge

P. A. Mehta · K. Sivaprakash · M. Parani · G. Venkataraman
A. K. Parida (✉)
M.S. Swaminathan Research Foundation, III Cross Street,
Institutional Area Taramani, Chennai 600 113, India
E-mail: ajay@mssrf.res.in

gain, water loss, and ion uptake with the transpiration stream on a low but constant level. *A. marina* grows in coastal regions where the salt concentration can be as high as 9% (Rao 1987). The regulation of inorganic ions occurs partially by exclusion at the roots and also by excretion via salt glands, the excretion rate for sodium and chloride ions being 0.4 μmol m^{-2} s^{-1} and 0.046 μmol m^{-2} s^{-1}, respectively (Shimony et al. 1973; Boon and Allaway 1982). It is thus an ideal candidate plant for mining genes for salt tolerance.

In the investigation reported here, single-pass sequences of randomly selected cDNA clones from a 0.5 *M* NaCl-stressed leaf cDNA library of *A. marina* were obtained and the ESTs analyzed and compared with the known database of genes. These results were used to assign putative functions to cDNAs, which were then categorized. Fifty-two cDNA clones were selected for further study based on their sequence analysis [significant homology in BLASTX results, presence of complete/near complete ORF (open reading frame) and a poly A tail]. Reverse Northern analysis data (comparing control and salt-treated samples) for the 52 genes is also presented along with Northern analysis for a select few. The relevance of the data with respect to salt tolerance is discussed.

Materials and methods

Construction of the cDNA library

Seeds of *Avicennia marina* (Forsk) Vierh. were collected from their natural mangrove habitat at Pichavaram (Tamil Nadu, India). These seeds were grown in sand-filled trays in the greenhouse at 37°C under a 12/12-h (light/dark) photoperiod (illuminated from 0600 hours to 1800 hours) in near-submergence conditions with daily watering. One-month-old *A. marina* seedlings (four-leaf stage) were acclimatized for 72 h in half-strength MS (Murashige Skoog1962) medium (no pH adjustment). The plants were subsequently transferred into half-strength MS medium supplemented with 0.5 *M* NaCl for 48 h. Total RNA was isolated from leaves of the salt-stressed plants; all leaves were harvested (Chomczynski and Sacchi 1987). cDNA prepared from poly (A+) mRNA (Superscript Lambda System; Invitrogen, Carlsbad, Calif.) was size-fractionated over a SizeSep-400 spun column (Amersham-Pharmacia Biotech, Piscataway, N.J.) and directionally cloned in the *Sal*I (5′)/*Not*I (3′) sites of pSPORT1 (Invitrogen). The ligated cDNA library was transformed into *Escherichia coli* DH5α. A library consisting of approximately 10^5 recombinants was obtained.

Sequencing and analysis of the EST clones

Plasmid DNA from randomly selected clones was extracted by alkaline lysis (Feliciello and Chinali 1993). The DNA sequence of the selected clones was determined by single-pass sequencing of the 5′ end using M13 reverse primer and the BigDye Terminator method (ABI Prism 310 DNA Sequencer; Applied Biosystems, Foster City, Calif.). The text file resulting from each sequencing reaction was edited manually and the vector and adapter sequences trimmed. The read length obtained after vector trimming was approximately 350–400 bp. Two classes of sequences were excluded from this study: sequences without inserts and sequences with a low e-value. In addition, 52 selected clones were also sequenced from the 3′ end with the M13 forward primer. The ESTs were subsequently subjected to BLASTX comparisons with the non-redundant protein database at the NCBI (National Center for Biotechnology Information, Bethesda, Md.) website (Altschul et al. 1990). Default parameters of the program were used in all cases. A minimum *P* cutoff value of 10^{-3} (the probability that alignment would be generated randomly is 1 < 1,000) was used to determine homology of the ESTs to known proteins. ESTs were assembled in overlapping contigs and non-overlapping sequences that correspond to different parts of the same gene using the CAP3 program (Huang and Madan 1999). Each EST sequence was classified as a single hit or redundant based on sequence comparisons. ESTs were classified as redundant when two or more copies in the library exhibited more than 95% identity over aligned regions or showed homology to the same database accession. The sequences were deposited in the NCBI EST database (http://www.ncbi.nlm.nih.gov/dbEST). Full-length sequences for PR541, PC2 and PC40 were also deposited in GenBank under the accession nos. AY639950, AY639951 and AY639952, respectively.

Reverse Northern hybridization

The 52 clones were amplified by PCR [one 1-min pre-amplification at 94°C; 28 PCR cycles of 30 s at 94°C, 30 s at 61°C, 1 min and 30 s at 72°C; a final extension cycle of 3 min at 72°C] with M13F and M13R primers in a 50-μl reaction volume using 10 ng of template. PR244 (accession no. BM172899) served as a control. Aliquots (1.5 μg) of each amplified product were electrophoresed on two identical 1% agarose gels (0.5× TBE), transferred to nylon membranes (Hybond-N, Amersham, UK) and immobilized via UV crosslinking according to the manufacturer's instructions.

mRNA was enriched from the total RNA population using streptavidin paramagnetic particles (Sigma, S2415, St. Louis, Mo.) and biotin-labeled oligo $d(T)_{18}$ primer. Total RNA (250 μg) in 10 m *M* Tris-HCl, pH 7.5, 0.5 *M* KCl (Buffer A) was heated to 65°C and annealed with 200 ng biotin-labeled oligo d(T) primer. Biotin-captured mRNA was immobilized by incubation with a streptavidin paramagnetic bead suspension (equilibrated in Buffer A). The beads were washed with 10 m *M* Tris-HCl, pH 7.5, 0.25 *M* KCl and the mRNA eluted in DEPC water and concentrated by lyophilization. The

concentration and integrity of the eluted mRNA was checked on a 1.2% formaldehyde-agarose gel (Sambrook et al. 1989).

For the reverse Northern analysis, 2 μg each of mRNA from the control and salt-treated leaves was reverse-transcribed in the presence of 1,850 kBq α-[^{32}P]dCTP (BRIT, specific activity: 111,000 Gbq). The double-stranded mRNA-cDNA hybrids were converted to single-stranded cDNA by RNaseH (Invitrogen) digestion. Unincorporated dNTPs were removed by purification over Sephadex G50 columns and labeling efficiency measured by scintillation counting. The master blots were probed, one with the control radiolabeled cDNA probe and the other with the treated radiolabeled cDNA probe (equal specific activity), in a hybridization solution [5× SSC, 5% dextran sulfate, 0.05 *M* Na-phosphate pH 7.2, 5× Denhardt's solution, 0.0025 *M* EDTA, 0.4% sodium dodecyl sulfate (SDS) and 100 μg/ml salmon sperm DNA] for 12–16 h at 65°C and then washed sequentially (15 min each time) with 2× SSC, 0.1% SDS and 1× SSC, 0.1% SDS. Hybridization signals were observed on Kodak X-ray films after a 3- to 48-h exposure.

Northern analysis

One-month-old *A. marina* seedlings were first conditioned for 72 h in half-strength MS nutrient solution, then stressed in half-strength MS containing 0.5 *M* NaCl. Samples of leaf and root tissue were taken and immediately frozen at 6, 12, 24 and 48 h of NaCl treatment and at 12 h and 24 h after their removal from the salt medium. Total RNA was isolated as described above. Equal amounts of total RNA (30 μg) were electrophoresed on a 1.2% MOPS-formaldehyde gel, transferred to nylon membrane (Hybond-N, Amersham) and fixed by UV crosslinking according to the manufacturers instructions. PCR-amplified products (PR541, PC2 and PC40—accession nos. BM173113, BM173162, BM173196, respectively) were labeled by the random primer method (Rediprime; Amersham) and used as probes. The radiolabeled probes were denatured and hybridized to the membrane at 65°C in an aqueous buffer as mentioned above. Washing and exposure to X-ray films were as already described.

Results

Putative identification and classification of EST sequences

All 1,841 ESTs were compared against sequences in the non-redundant database (nr) at the NCBI using the BLASTX program. The result of each comparison was screened manually, and 1,602 "readable" ESTs (87%) were retained after screening. Assignment of putative functions to the ESTs was based on the numerical cutoff values obtained in the BLAST comparisons and supplemented with information from PubMed (NCBI). Of 1,602 ESTs, 1,155 had significant homology to previously identified genes, and these were grouped into 12 functional categories. The remaining "unknown" genes (447) refer to that subset of the ESTs that show a significant similarity (high e-value) to genes reported in the public database but do not have an assigned function to date. The distribution of the ESTs in the 13 categories is shown in Fig. 1a. Unknown genes form the largest category (30%), followed by genes required for primary metabolism (13%). Genes involved in transcription and chromatin organization, protein synthesis and processing each represent 10% of the sequenced ESTs, while those involved in membrane transport and intracellular trafficking represent 9% of the ESTs. Eight percent of the ESTs relate to signal transduction, while 7% are similar to previously reported stress-induced genes.

Clustering of the 1,602 ESTs using the CAP3 program produced a total of 292 contigs encompassing 814 ESTs, resulting in a 51% redundancy. The remaining 788 ESTs were singletons; that is, they did not cluster with any of the other ESTs. A summary for the EST clustering is shown in Table 1. The largest contig contained 18 ESTs, while 186 contigs had two ESTs each. The 447 "unknown" ESTs could be assembled into 80 contigs and 247 singletons.

The availability of the full-genome sequences of *Arabidopsis thaliana* and *Oryza sativa* allows comparison with other plant databases (Yamamoto and Sasaki 1997; The Arabidopsis Genome Initiative 2000). An analysis revealed that 88% of *A. marina* sequences in the unknown category matched the *A. thaliana* genome (Fig. 1b). In the other categories, 25–65% of the ESTs showed the highest homology to *A. thaliana* genes, while a very small percentage, less than 10% in most classes, matched the *O. sativa* database.

Putative stress-regulated genes

ESTs showing homology to stress-tolerant genes reported in literature represent the sixth most abundant category. Dehydrins (34 clones/12 genes) predominate in this class and include late-embryogenesis-related proteins (accession no. BM172740) and dessication or drought-induced proteins (accession nos. BM172744 and BM172999). Other stress-induced genes present in this category include heat shock proteins, thioredoxin, osmotin and genes for osmolyte production such as betaine aldehyde dehydrogenase (BADH) and pyrroline-5-carboxylate synthase. The analyzed EST pool contained transcripts coding for enzymes involved in the oxidative stress response such as catalase, superoxide dismutase, peroxidases, glutathione S-transferase and epoxide hydrolase. In addition, genes reported to be induced by heavy metal stress, such as metallothioneins, aluminium-induced protein(s), truncated copper-binding

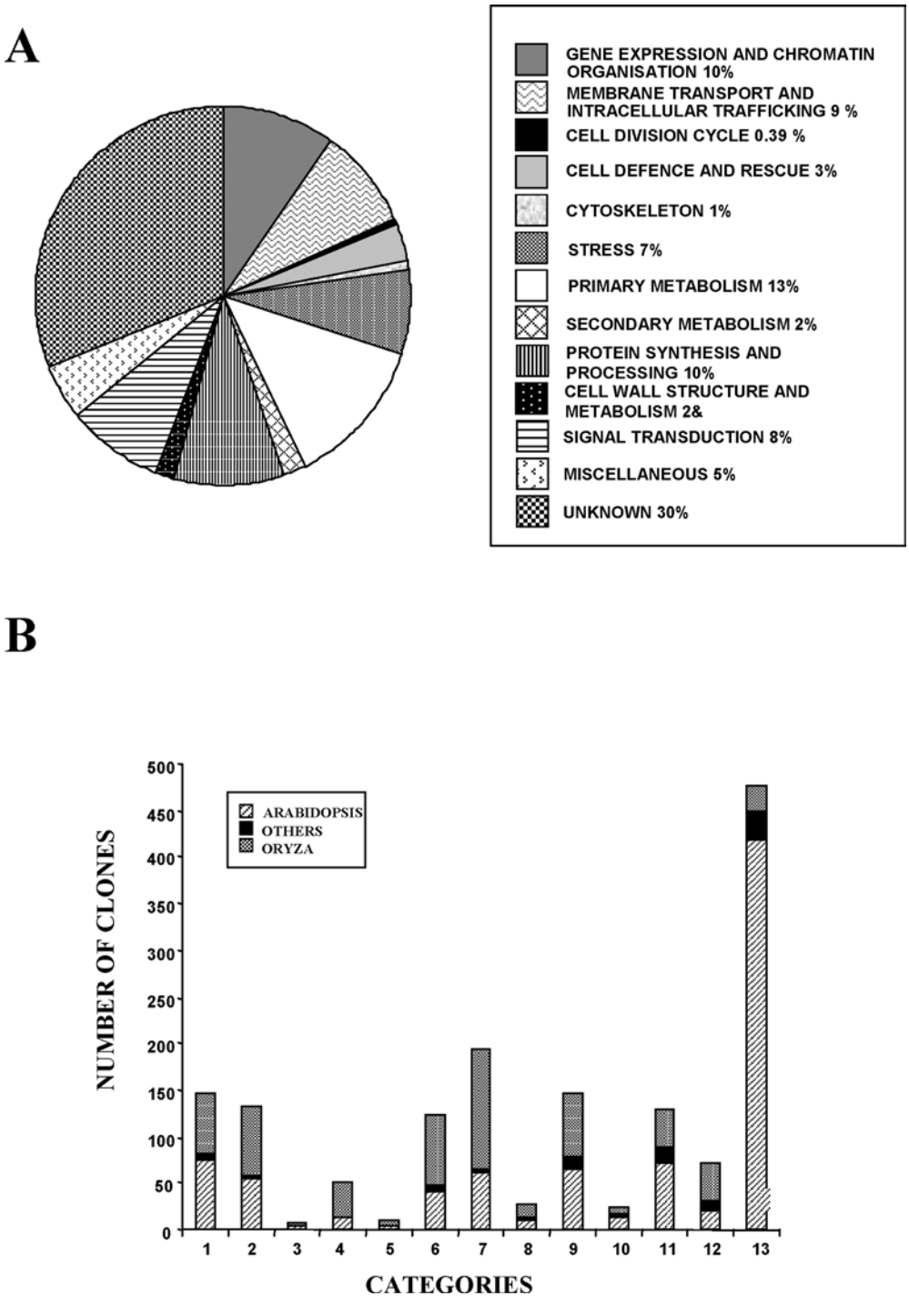


Fig. 1 **a** Distribution of *Avicennia marina* ESTs into functional categories. **b** Comparison of the distribution of *A. marina* classified ESTs with the genomes of *A. thaliana*, *O. sativa* and other plant species. The categories are: *1* gene expression and chromatin organization, *2* membrane transport and intracellular trafficking, *3* cell division cycle, *4* cell defence and rescue, *5* cytoskeleton, *6* stress, *7* primary metabolism, *8* secondary metabolism, *9* protein synthesis and processing, *10* cell-wall structure and metabolism, *11* signal transduction, *12* miscellaneous, *13* unknown

protein CUTA and divalent cation-tolerant gene, and those induced by anaerobic stress, such as alcohol dehydrogenases and submergence-induced genes, were also found in the ESTs sequenced.

Transcripts coding for proteins involved in membrane transport and processing play an integral role in the response to water deficit such as that imposed by salt stress and include aquaporins, proline/glycine betaine transporter (BM173094) and Na^+/H^+ antiporter (BM173059, BM172811). In addition, protein factors involved in the regulation of signal transduction events, such as receptors, protein and lipid kinases, calmodulins and protein phosphatases, which may have a role in stress signaling pathways, have been categorized separately. Genes for a variety of transcription factors that contain typical DNA binding motifs, such as MYB, bZIP, ERF/AP2, have been demonstrated to be stress-inducible (Zhu 2002). Transcription factors containing similar domains are present in the *A. marina* ESTs and may have a role in regulating the response to salt stress

Table 1 *Avicennia marina* EST assembly—summary

Contig assembly statistics	
Number of ESTs assembled	1,602
Number of contigs	292
Number of assembled sequences (contigs)	814
Number of singletons	788
Redundancy (%)	50.9
Contig size:	
2–4 ESTs	268
5–7 ESTs	17
8–10 ESTs	1
11–13 ESTs	4
14–16 ESTs	1
17–18 ESTs	1

Reverse Northerns

Of 447 ''unknown'' ESTs, 52 were selected on the basis of the presence of complete or near-complete ORFs; additionally, 3′ single-pass sequencing revealed the presence of a poly A tail in these 52 ESTs. Reverse Northern conducted on these 52 genes revealed up-regulation for 26 and down-regulation for five. Reverse Northern for PR541 showed up-regulation at 48 h, while the data for the remaining two clones, PC2 and PC40, suggested down-regulation at 48 h. Northern kinetic analysis (salt induction) for PR541 showed maximum expression in the leaf at 48 h of salt stress. While the amount of transcript increased gradually in the leaf, peaking at 48 h of stress, up-regulation in the root was more rapid, peaking at 6 h and remaining constant thereafter until withdrawal of the stress. Northern data for PC2 and PC40 for leaf tissues showed a decrease in the level of the transcript at 48 h and elevated expression as compared to the control root tissues, with a gradual decline. Reverse Northern data and Northern kinetic analysis for these three ESTs is shown in Fig. 2.

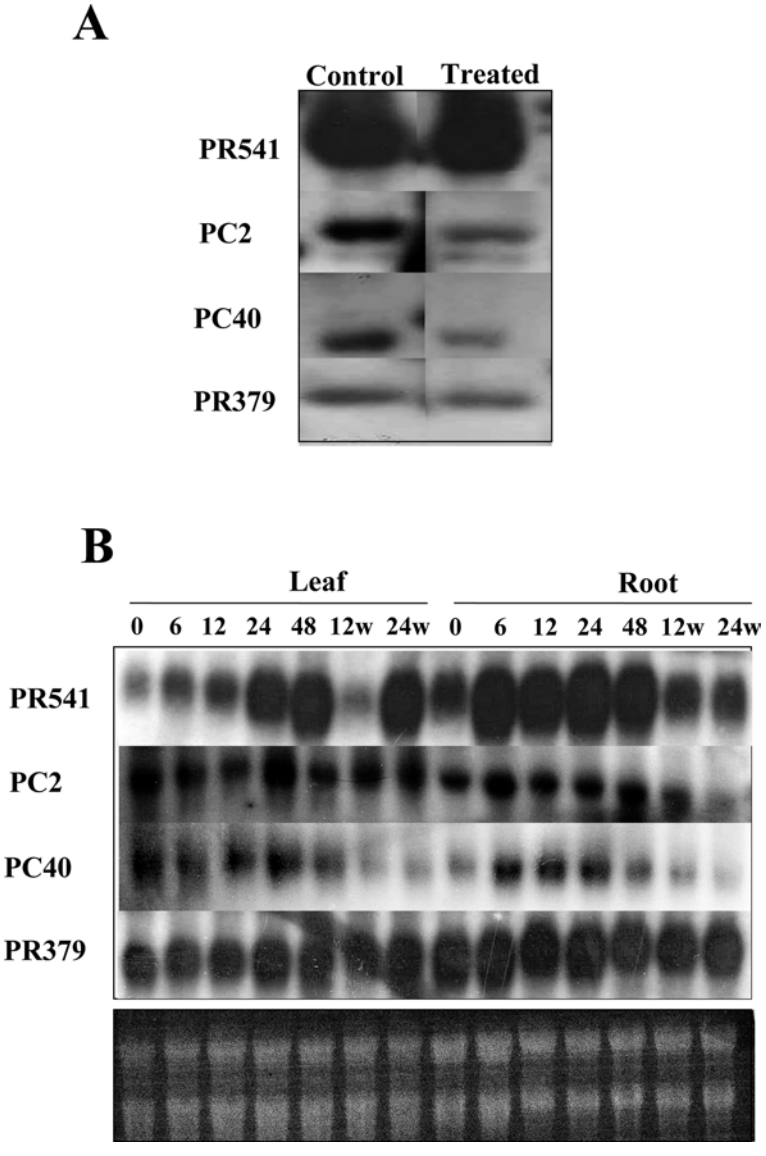


Fig. 2 **a** Comparison of transcript levels of PR541, PC2 and PC40 in leaves of *A. marina* by Reverse Northern blots. Duplicate filters were hybridized with [^{32}P]-labeled cDNA of equal specific activity under stringent conditions. The *panel on the left* shows control cDNA-hybridized samples, while the *panel on the right* shows salt-stressed cDNA hybridized to the twin blot. PR379 is included as control and did not consistently change with salt stress upon repetition. **b** Northern blot analysis. The *top three panels* show the transcript detected for PR541, PC2 and PC40 after a 0-, 6-, 12-, 24- and 48-h exposure to 0.5 *M* NaCl and 12 h and 24 h (*12w*, *24w*, respectively) after withdrawal of the salt stress to the leaves and roots of *A. marina*, while the *bottom panel* shows the ethidium bromide-stained gel before transfer. The Northern is representative of two individual experiments. The EST name is on the *left*

Discussion

Mangroves are facultative halophytes and exclude most of the salt in seawater. In addition, some species such as *A. marina* actively secrete salt from their cells. *A. marina* inhabits sediments, the salinity of which is a function of local precipitation, subterranean seepage, terrestrial run-off, evaporation and flushing. As these factors vary considerably, salt concentration in mangrove soils fluctuate markedly from 1% to more than 4% (Parani 1999).

The abundance of dehydrin clones in the *A. marina* ESTs and up-regulation of BM173212 (not included in the reverse Northern data) in response to salt stress (unpublished results) correlates with their role in dessication tolerance. Dehydrins are a specific subset of intrinsically unstructured proteins characterized by high hydrophilicity that are found only in plants. They are induced by dehydration, low temperature, osmotic stress, seed drying and/or exposure to abscisic acid (Garay-Arroyo et al. 1999; Tompa 2002). More recently, it has been suggested that dehydrins may undergo function-related conformational changes at the water/membrane interface that are perhaps related to the stabilization of vesicles or other endomembrane structures under stress conditions (Koag et al. 2003). In the ESTs analyzed here, putative low-molecular-weight (LMW) heat shock proteins (HSPs), Hsp70, and heat shock factor 7 homologs were identified. These may contribute to the stress response associated with high tropical temperatures. Metallothioneins (MT) are a group of LMW metal-binding proteins with a high cysteine content that are thought to be involved in metal ion metabolism and detoxification (Hall 2002). In higher plants, they are induced by ethylene, viral infection, wounding and cold treatment (Bausher et al. 2003). MTs were found to be abundantly expressed in the EST data collected from *Citrus sinensis* seedlings, and MT-like transcripts have been reported to be highly up-regulated in response to salt stress in barley (Ozturk et al. 2002; Bausher et al. 2003).

The accumulation of osmoprotectants by either altering metabolism or increasing transport is an important process for the adaptation to environmental stress (Waditee et al. 2002). Hibino et al. (2001) reported that under controlled conditions leaves of *A. marina* have 67 μmol g^{-1} dry weight of glycinebetaine, which increases twofold upon salt stress (0.4 *M* NaCl). The BADH gene isolated from our library shows 97% identity (amino acid) with the clone reported by Hibino et al. (2001; accession no. AB043539). The EST for the proline/betaine transporter (BM173094/partial clone) shows maximum identity with AmT1, the betaine/proline transporter reported by Waditee et al. (2002).

The unknown category of genes forms 30% of the ESTs analyzed and is a useful starting point from which to isolate new genes that govern salt tolerance. Thus, 52 genes from this pool were selected for further analysis. Of these, 26 showed up-regulation and five showed down-regulation. Reverse Northern hybridization is a rapid method used to identify differentially expressed cDNAs and this is often supplemented with Northern analysis (Gyorgyey et al. 2000). Of the 52 clones, three were randomly selected and Northern kinetic profiles for salt stress in leaf and root tissues examined.

PR541, the up-regulated clone, revealed an increase in the level of the transcript in the leaf tissue at 48 h of salt stress; this subsequently declined to normal levels 12 h following withdrawal of the salt stress. PR541 shows a strong homology to the precursor form of the polypeptide hormone phytosulphokine. The full-length cDNA is 681 bp and codes for an ORF of 77 amino acids (Fig. 3a). Phytosulphokine-α (PSK-α) is a tyrosine-sulfated mitogenic pentapeptide first isolated from *Asparagus* mesophyll cell cultures (Matsubayashi et al. 1996). It was secreted into the medium and found to induce cell proliferation at low density. PSKs with identical structures were subsequently identified in the conditioned medium of rice, *Zinnia* and carrot cultures (Ryan et al. 2002). In rice, OsPSK is single-copy gene that codes for a PSK precursor of 89 amino acids. The precursor form is inactive and upon processing gives rise to the pentapeptide YIYTQ, which is sulfated at the tyrosine residues (active form). In *Arabidopsis*, four genes that code for PSK-α precursors have been found (Yang et al. 2001). A comparison of PSK precursor sequences deposited in Genbank reveals absolute conservation of the pentapeptide and the amino acid preceding it; all other regions of the precursor showed limited homology. This suggests a new level of diversity among polypeptides that are processed into the same signaling molecules in plants. A comparison of the OsPSK ORF with that of PR541 shows the presence of a hydrophobic leader sequence and the conserved DYIYTQ stretch. Matsubayashi et al. (2002) reported the presence of a 120-kDa receptor kinase protein from carrot cells that binds to the sulfated pentapeptide with high affinity. The receptor contains leucine-rich repeats (LRR), and it has been suggested that this stretch of LRRs plays a key role in protein-protein interactions. The in vivo role of PSK has not yet been established. While the cumulative data reveal that PSK-α has a role related to cell division and development, its mechanism of action, its signaling pathway, its relationship to the other hormones and its role in regulation developmental processes remain to be determined.

PC2 and PC40 were both down-regulated in control leaves relative to the levels observed in 48-h stressed leaves. The PC2 cDNA consists of 1,295 bp and encodes a putative protein of 336 amino acids (Fig. 3c). BLASTX (Altschul et al. 1990) results reveal significant similarity to the hypothetical *Arabidopsis* protein At1g16520. Domain analysis using SMART (Schultz et al. 1998; Letunic et al. 2004) identified a coiled-coil region closer to the N-terminus of the putative PC2 ORF. The hypothetical At1g16520 protein also shows a coiled-coil domain at the N-terminus. The coiled-coil protein

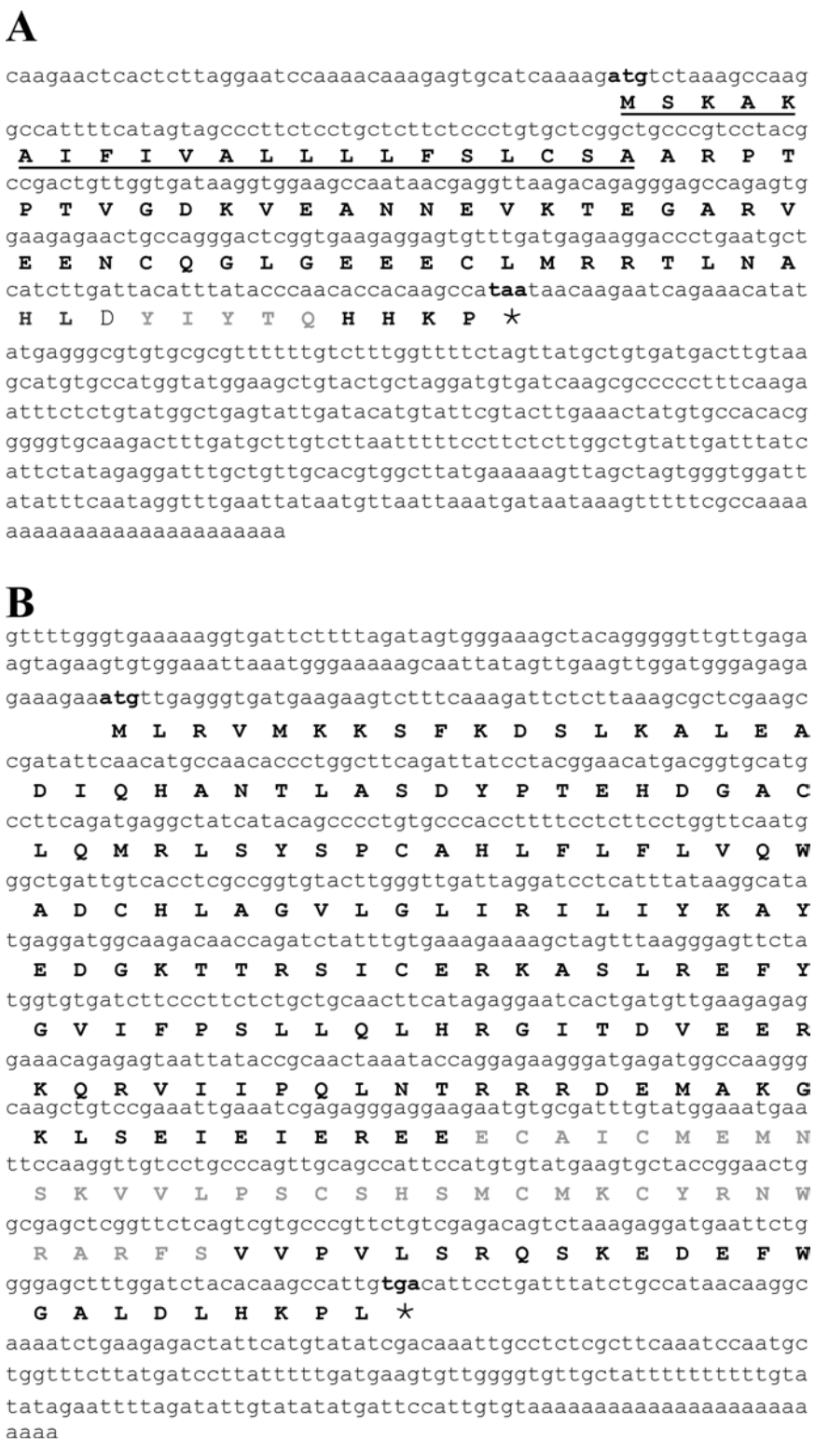
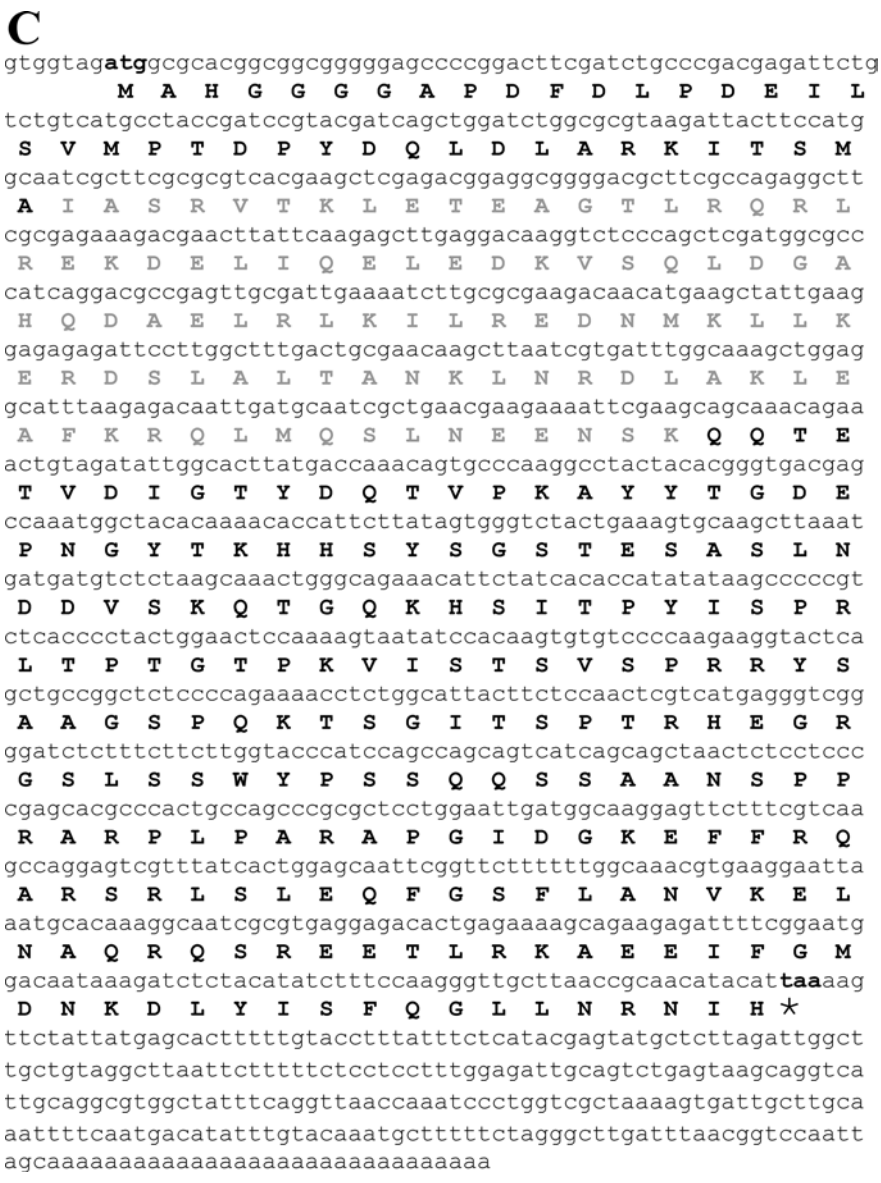


Fig. 3 **a** Nucleotide and deduced amino acid sequence of PR541. The nucleotide sequence is indicated by the *top sequence* and the deduced amino acid sequence by the *underlying sequence*. The pentapeptide sequence YIYTQ is highlighted in *gray*. The predicted signal peptide sequence is *underlined in the underlying sequence*. The nucleotide sequence of PR541 has been deposited in GenBank with the accession no. AY639950. **b** The nucleotide and deduced amino acid sequence of PC40. The nucleotide sequence is indicated by the *top sequence* and the deduced amino acid sequence by the *underlying sequence*. The putative RING-HC domain is highlighted in *grey*. The nucleotide sequence of PC40 has been deposited in GenBank with accession no. AY639952. **c** Nucleotide and deduced amino acid sequence of PC2. The nucleotide sequence is indicated by the *top sequence* and the deduced amino acid sequence by the *underlying sequence*. The putative coiled-coil domain is highlighted in *grey*. The nucleotide sequence of PC2 has been deposited in GenBank with accession no. AY639951. *Asterisk* indicates the stop codon

oligomerization motif consists of two or more amphipathic α-helices that twist around each other in a supercoil (Rose et al. 2004). These are found in a number of functionally distinct proteins that are often involved in attaching protein complexes to larger cellular structures such as the Golgi, centrosome, centromere or the nuclear envelope. In plants, coiled-coil domains are not as large as in animal proteins, and long coiled-coil domains with fewer than 100 amino acids form 50% of the coiled-coil domain-containing proteins in *Arabidopsis* (Rose et al. 2004).

The PC40 cDNA (964 bp) was also sequenced completely and found to encode a 207-amino acid protein showing high homology to the putative

Fig. 3 (Contd.)



Arabidopsis C3HC4 RING zinc finger protein AT5g01520 (Fig. 3b). Domain analysis with SMART suggested the presence of a RING finger domain at the C-terminus of the protein, while at the N-terminus there was a potential hydrophobic region. Interestingly, *BRH1*, a brassinosteroid down-regulated gene in *Arabidopsis*, encodes a C-terminal ring-H2 finger protein and also has an N-terminal hydrophobic region (Molnar et al. 2002). RING finger domain proteins are defined by a conserved pattern of cysteine and histidine residues capable of coordinating two atoms of zinc within a characteristic cross-brace structure. Numerous members of this protein family occur in eukaryotic organisms and participate in diverse functions, such as signal transduction, vesicular transport, cell proliferation and embryonal patterning (Molnar et al. 2002). Ring finger domains thus help in the assembly of multi-protein complexes mediating diverse cellular processes (Xu and Li 2003). The RING finger domain of PC40, unlike *BRH1*, belongs to the RING-HC class of RING finger domains and is characterized by a histidine-metal residue at the fourth ligand position (Kosarev et al. 2002).

The EST data reported herein are the first available data for *A. marina* and can be used to establish this mangrove species as a model system for the molecular genetic studies of plant salinity tolerance. This data set contains previously reported as well as uncharacterized genes. Some of the previously reported genes have well-characterized roles in stress tolerance. *A. marina* grows in anoxic soils in coastal areas with high salinity and often under conditions of high temperature and light. The subset of genes in the 'stress category' reflects genes induced by all of the above-mentioned abiotic stresses. Expression patterns for the 'unknown' genes provide a starting point for the isolation of salt tolerance candidate genes, and further functional analysis will elucidate their role in salt tolerance.

Acknowledgements We are grateful to the Department of Biotechnology and Department of Atomic Energy, Government of India for financial support. The help and support provided by Ms. Y. Dilhara Begam in EST analysis is gratefully acknowledged.

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410

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

Boon PI, Allaway WG (1982) Assessment of leaf-washing techniques for measuring salt secretion in *Avicennia marina* (Forsk.) Vierh. Aust J Plant Physiol 9:725–734

Chauhan S, Forsthoefel N, Ran Y, Quigley F, Nelson DE, Bohnert HJ (2000) Na + /myo-inositol symporters and Na + /H + -antiporter in *Mesembryanthemum crystallinum*. Plant J 24: 511–522

Chomczynski P, Sacchi N (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal Biochem 162:156–159

Cushman JC, Bohnert HJ (2000) Genomic approaches to plant stress tolerance. Curr Opin Plant Biol 3:117–124

Feliciello I, Chinali G (1993) A modified alkaline lysis method for the preparation of highly purified plasmid DNA from *Escherichia coli*. Anal Biochem 212:394–401

Garay-Arroyo A, Colmenero-Flores JM, Garciarrubio A, Covarrubias AA (1999) Highly hydrophilic proteins in prokaryotes and eukaryotes are common during conditions of water deficit. J Biol Chem 275:5668–5674

Gyorgyey J, Vaubert D, Jimenez-Zurdo JI, Charon C, Troussard L, Kondorosi A, Kondorosi E (2000) Analysis of *Medicago truncatula* nodule expressed sequence tags. Mol Plant Microbe Interact 13:62–71

Hall JL (2002) Cellular mechanisms for heavy metal detoxification and tolerance. J Exp Bot 53:1–11

Hibino T, Meng YL, Kawamitsu Y, Uehara N, Matsuda N, Tanaka Y, Ishikawa H, Baba S, Takabe T, Wada K, Ishii T, Takabe T (2001) Molecular cloning and functional characterization of two kinds of betaine-aldehyde dehydrogenase in betaine-accumulating mangrove *Avicennia marina* (Forsk.) Vierh. Plant Mol Biol 45:353–363

Huang X, Madan A (1999) CAP3: a DNA sequence assembly program. Genome Res 9:868–77

Koag MC, Fenton RD, Wilkens S, Close TJ (2003) The binding of maize DHN 1 to lipid vesicles. Gain of structure and lipid specificity. Plant Physiol 131:309–316

Kosarev P, Mayer KFX, Hardtke CS (2002) Evaluation and classification of RING-finger domains encoded by the *Arabidopsis* genome. Genome Biol 3:r0016.1–0016.12

Letunic I, Copley RR, Schmidt S, Ciccarelli FD, Doerks T, Schultz J, Ponting CP, Bork P (2004) SMART 4.0: towards genomic data integration. Nucleic Acids Res 32:D142–D144

Matsubayashi Y, Sakagami Y (1996) Phytosulfokine, sulfated peptides that induce the proliferation of single mesophyll cells of *Asparagus officinalis* L. Proc Natl Acad Sci USA 93:7623–7627

Matsubayashi Y, Ogawa M, Morita A, Sakagami Y (2002) An LRR receptor kinase involved in perception of a peptide plant hormone, phytosulfokine. Science 24:1470–1472

Molnar G, Bancos S, Nagy F, Szekeres M (2002) Characterisation of *BRH1*, a brassinosteroid-responsive RING-H2 gene from *Arabidopsis thaliana*. Planta 215:127–133

Munns R, Husain S, Rivelli AR, James RA, Condon AG, Lindsay MP, Lagudah ES, Schachtman DP, Hare RA (2002) Avenues for increasing salt tolerance of crops, and the role of physiologically based selection traits. Plant Soil 247:93–105

Murashige T, Skoog F (1962) A revised method for rapid growth and bioassays with tobacco tissue cultures. Physiol Plant 15:473–497

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

Parani M (1999) Isolation of salt tolerance genes from mangroves: characterization of betaine aldehyde dehydrogenase (BADH) gene from *Avicennia marina* and its expression in tobacco transgenic system. PhD thesis, Anna University, India

Rao AN (1987) Mangrove ecosystems of Asia and the Pacific. In: Umali RM, Zamora PM, Gotera RR, Jara RS, Camacho AS (eds) Mangroves of Asia and the Pacific: status and management. UNDP-UNESCO Technical Report, pp 1–41

Rose A, Manikantan S, Schraegle SJ, Maloy MA, Stahlberg EA, Meier I (2004) Genome-wide identification of *Arabidopsis* coiled-coil proteins and establishment of the ARABI-COIL database. Plant Physiol 134:927–939

Ryan CA, Pearce G, Scheer J, Moura DS (2002) Polypeptide hormones. Plant Cell 14:S251–S264

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning. Cold Spring Harbor Laboratory Press, Plainview

Schultz J, Milpetz F, Bork P, Ponting CP (1998) SMART, a simple modular architecture research tool: identification of signaling domains. Proc Natl Acad Sci USA 95:5857–5864

Shimony C, Fahn A, Reinhold L (1973) Ultrastructure and ion gradients in the salt glands of *Avicennia marina* (Forsk.) Vierh. New Phytol 72:27–36

The Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408:796–815

The International Human Sequencing Consortium (2001) Initial sequencing and analysis of the human genome. Nature 409:860–921

Tompa P (2002) Intrinsically unstructured proteins. Trends Biochem Sci 27:527–533

Waditee R, Hibino T, Tanaka Y, Nakamura T, Incharoensakdi A, Hayakawa S, Suzuki S, Futsuhara Y, Kawamitsu Y, Takabe T, Takabe T (2002) Functional characterization of betaine/proline transporters in betaine-accumulating mangrove. J Biol Chem 277:18373–18382

Xu R, Li QQ (2003) A RING-H2 zinc finger protein gene *RIE1* is essential for seed development in *Arabidopsis*. Plant Mol Biol 53:37–50

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

Yang H, Matsubayashi Y, Nakamura K, Sakagami Y (2001) Diversity of *Arabidopsis* genes encoding precursors for phytosulfokine, a peptide growth factor. Plant Physiol 127:842–851

Zhu JK (2002) Salt and drought stress signal transduction in plants. Annu Rev Plant Biol 53:247–273