



National Library  
of Canada

Acquisitions and  
Bibliographic Services Branch

395 Wellington Street  
Ottawa, Ontario  
K1A 0N4

Bibliothèque nationale  
du Canada

Direction des acquisitions et  
des services bibliographiques

395, rue Wellington  
Ottawa (Ontario)  
K1A 0N4

*Your file* *Votre référence*

*Our file* *Notre référence*

## NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

## AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

Canada

**Molecular Studies on Salt-Stress-Inducible Genes  
of the Salt-Tolerant Wild Wheatgrass  
Lophopyrum elongatum**

**Wei Shen**

**A Thesis  
in  
The Department  
of  
Biology**

**Presented in Partial Fulfilment of the Requirements  
for the Degree of Master of Science at  
Concordia University  
Montreal, Quebec, Canada**

**August 1994**

**©Wei Shen, 1994**



National Library  
of Canada

Acquisitions and  
Bibliographic Services Branch

395 Wellington Street  
Ottawa, Ontario  
K1A 0N4

Bibliothèque nationale  
du Canada

Direction des acquisitions et  
des services bibliographiques

395, rue Wellington  
Ottawa (Ontario)  
K1A 0N4

Your file    *Voire référence*

Our file    *Notre référence*

THE AUTHOR HAS GRANTED AN IRREVOCABLE NON-EXCLUSIVE LICENCE ALLOWING THE NATIONAL LIBRARY OF CANADA TO REPRODUCE, LOAN, DISTRIBUTE OR SELL COPIES OF HIS/HER THESIS BY ANY MEANS AND IN ANY FORM OR FORMAT, MAKING THIS THESIS AVAILABLE TO INTERESTED PERSONS.

L'AUTEUR A ACCORDE UNE LICENCE IRREVOCABLE ET NON EXCLUSIVE PERMETTANT A LA BIBLIOTHEQUE NATIONALE DU CANADA DE REPRODUIRE, PRETER, DISTRIBUER OU VENDRE DES COPIES DE SA THESE DE QUELQUE MANIERE ET SOUS QUELQUE FORME QUE CE SOIT POUR METTRE DES EXEMPLAIRES DE CETTE THESE A LA DISPOSITION DES PERSONNE INTERESSEES.

THE AUTHOR RETAINS OWNERSHIP OF THE COPYRIGHT IN HIS/HER THESIS. NEITHER THE THESIS NOR SUBSTANTIAL EXTRACTS FROM IT MAY BE PRINTED OR OTHERWISE REPRODUCED WITHOUT HIS/HER PERMISSION.

L'AUTEUR CONSERVE LA PROPRIETE DU DROIT D'AUTEUR QUI PROTEGE SA THESE. NI LA THESE NI DES EXTRAITS SUBSTANTIELS DE CELLE-CI NE DOIVENT ETRE IMPRIMES OU AUTREMENT REPRODUITS SANS SON AUTORISATION.

ISBN 0-315-97598-9

Canada

## ABSTRACT

### Molecular Studies on Salt-Stress-Inducible Genes of the Salt-Tolerant Wild Wheatgrass Lophopyrum elongatum

Wei Shen

Full length cDNA clones for the ESI3 gene and the ESI18 multigene family members ESI18-2, ESI18-15 and ESI18-17 were isolated from the salt-tolerant wheatgrass Lophopyrum elongatum. These genes are rapidly induced in roots by salt treatment. Nucleotide sequence of the ESI3 cDNA indicates that ESI3 encodes a 5.9-kDa hydrophobic protein containing two putative membrane-spanning segments. This protein is identical in amino acid sequence to the barley protein BLT101. Genes ESI18-2, -15 and -17 code for proteins with molecular weights of 41.7, 15.9 and 12.8 kDa respectively. These proteins belong to the plant dehydrin family, members of which are involved in stress response and embryogenesis. The ESI18-15 protein contains a serine cluster and two copies of a lysine-rich motif which is typical of dehydrins. ESI18-2 and ESI18-17 do not have the serine cluster, but contain seven and three lysine-rich repeats respectively with slight variations. The sequences between the lysine-rich repeats of the ESI18 proteins are rich in glycine. A positive correlation between the degree of salt induction of the dehydrin-like genes in four Lophopyrum species including L. elongatum and their abilities to survive salt stress was demonstrated.

## ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to all those who helped me fulfil this task, especially to:

Dr. Patrick J. Gulick, my supervisor, for giving me the opportunity to study and do research in his laboratory, for his scientific guidance, for sharing his knowledge and ideas, for his support, encouragement and friendship, and for his asking me about the results whenever he saw me.

Drs. P. Joyce and A. Tsang, committee members for my M. Sc. program, for their support, scientific guidance.

Dr. B. S. Mangat, Graduate Program Director, and his successive secretaries, Drs. R. Storms and R. K. Ibrahim and all the staff in the Department of Biology.

The Department of Biology, the Faculty of Arts and Science and Concordia University, for the academical environments, the financial supports and the administrative services.

Hongyan A. (did lots of pioneer work), Meera (we caught ESI18-17) and Ehab A., my co-workers in the project, for sharing their discoveries and for their friendship.

Other labmates, Ursula O., Sylvie A, Sirinart A., Rahim J., Antonin G. (always reminded me that old-days grads never left lab earlier than 5 p.m.), Wendy D., Susan A., etc., and people in Dr. Ibrahim's lab (finally I did not have to ask them for permission to use their equipment, sometimes) and in other labs in Departments of Biology and Biochemistry.

All my friends inside and outside Canada.

My parents, for their love and endless support of every kind and for their forgetting telling me the chinese laohua 'Never travel far away, stay with your parents'.

Finally, my wife Yingnian, for her love, understanding, support, encouragement, and everything she did for me (including cooking).

This work was supported by a Department of Biology Teaching assistance fellowship, a Concordia University Graduate fellowship, a grant from the National Science and Engineering Council of Canada and a Faculty Development Research grant from Concordia University.

## TABLE OF CONTENTS

LIST OF FIGURES	x
LIST OF TABLES	xii
ABBREVIATIONS	xiii
PART I. INTRODUCTION	1
1. Significance of genetic studies on osmotic stress response and tolerance in plants	1
2. Osmotic-stress-induced genes and their possible functions	5
2.1. <u>Lea</u> and <u>lea</u> -like genes	5
2.1.1. Group 1 <u>lea</u> genes	7
2.1.2. Group 2 <u>lea</u> genes	9
2.1.3. Group 3 <u>lea</u> genes	11
2.1.4. Other <u>lea</u> genes	13
2.2. Osmotin genes and proteins	14
2.3. Phosphoenolpyruvate carboxylases and other enzymes involved in CAM	15
2.4. Genes for proteins related to proteases and protease inhibitors	17
2.5. Genes related to biosynthesis of proline and other osmolytes	18
2.6. H <sup>+</sup> -ATPases	20
2.7. Other osmotic-stress-responsive genes	21

3.	Regulation of expression of osmotic-stress-induced genes	23
4.	Genetic response to salt stress in the salt-tolerant wheatgrass <u>Lophopyrum elongatum</u>	28
PART II. MATERIALS AND METHODS		
		32
1.	Plant materials	32
2.	Extraction of total RNA and isolation of poly(A) <sup>+</sup> RNA	33
3.	Plasmid isolation and purification	34
4.	Digestion of DNA with restriction endonucleases	35
5.	Polymerase chain reaction	35
6.	Agarose gel electrophoresis of DNA and purification of DNA fragments	36
7.	Preparation of radioactive probes	37
8.	Northern blot analysis	37
9.	cDNA library screening and plasmid rescue	38
10.	Hybridization	39
11.	Preparation of serial deletion clones for sequencing	40
12.	DNA sequencing	41
13.	Primer extension	42
14.	Inverse PCR cloning	43
15.	Computer-assisted data analysis	44



PART III. RESULTS	45
1. Characterization of the <u>ESI3</u> gene	45
1.1. Isolation of cDNA clones for the <u>ESI3</u> gene	45
1.2. Determination of transcription start site and confirmation of ESI3F to be a nearly-full length cDNA clone	47
1.3. Cloning the 5'-flanking region of the <u>ESI3</u> gene	47
1.4. Analysis of the predicted <u>ESI3</u> gene product	50
2. Isolation and characterization of cDNA clones for members of the <u>ESI18</u> multigene family	51
2.1. Isolation of full length cDNAs for members of the <u>ESI18</u> family	51
2.2. Analysis of the predicted <u>ESI18</u> gene products	56
2.2.1. ESI18-15 protein	56
2.2.2. ESI18-2 protein	56
2.2.3. ESI18-17 protein	58
3. Expression of dehydrin-like genes in wild wheatgrass species with differences in salt tolerance	59
PART IV. DISCUSSION	63
1. The <u>ESI3</u> gene	63
1.1. Database search for proteins similar to ESI3	63
1.2. Possible functions of the <u>ESI3</u> gene in salt	

resistance	64
2. The <u>ESI18</u> genes and other <u>L. elongatum</u> dehydrin-like genes	67
2.1. The predicted gene products	67
2.2. Comparison of subdomains of the <u>L. elongatum</u> dehydrin-like proteins	71
2.2.1. Lysine-rich repeats	72
2.2.2. Glycine-rich sequences	77
2.2.3. Structure and function	77
3. Expression of the <u>ESI</u> genes in <u>L. elongatum</u>	79
4. Adaptive roles of the <u>ESI</u> genes in salt tolerance	81
REFERENCES	82
APPENDICES	100
A. Nomenclature of the <u>ESI18</u> genes	100

## LIST OF FIGURES

- Figure 1. Nucleotide sequence of the ESI3 gene and deduced amino acid sequence of the gene product 46
- Figure 2. Primer extension assay to determine transcription start site for the ESI3 gene 48
- Figure 3. Strategies for the cloning of the 5'-flanking sequence of the ESI3 gene by inverse PCR 49
- Figure 4. Hydropathy index computation for the predicted ESI3 protein 52
- Figure 5. Nucleotide sequences of the cDNAs ESI18-2-28, ESI18-15-34 and ESI18-17 54, 55
- Figure 6. Amino acid sequences of the predicted products of the genes ESI18-2, ESI18-15 and ESI18-17 57
- Figure 7. Steady state mRNA levels for dehydrin-like genes of four Lophopyrum species during the time course of exposure to salt stress 62
- Figure 8. Comparison of the ESI3 protein and proteins with similar amino acid sequences 65
- Figure 9. Diagram of the primary structures of the L. elongatum dehydrin-like proteins 69
- Figure 10. Alignment of the amino acid sequences of the lysine-rich repeats of the L. elongatum dehydrin-like proteins 73
- Figure 11. Dendrogram of the alignment of the nucleotide

sequences for the lysine-rich repeats of the L.  
elongatum dehydrin-like proteins 75

Figure 12. Alignment of the amino acid sequences between the  
lysine-rich repeats of the L. elongatum dehydrin-  
like proteins 78

LIST OF TABLES

Table 1. Correlation of induction of dehydrin-like genes and  
salt tolerance in different wheatgrass Lophopyrum  
species      60

## ABBREVIATIONS

- ABA: Abscisic acid.
- ABRE: Abscisic acid responsive element.
- BADH: Betaine-aldehyde dehydrogenase.
- BLAST: Basic Local Alignment Search Tool.
- bp: Base pair.
- CAM: Crassulacean acid metabolism.
- cDNA: Complementary deoxyribonucleic acid.
- cpm: Count(s) per minute.
- dATP: 2'-Deoxyadenosine 5'-triphosphate.
- dCTP: 2'-Deoxycytidine 5'-triphosphate.
- dGTP: 2'-Deoxyguanosine 5'-triphosphate.
- dhn, DHN: dehydration-induced, dehydrin.
- DNA: Deoxyribonucleic acid.
- DRE: Dehydration-responsive element.
- DTT: Dithiothreitol.
- dTTP: 2'-Deoxythymidine 5'-triphosphate.
- EDTA: Ethylenediaminetetraacetic acid.
- EGTA: Ethylene glyco-bis( $\beta$ -aminoethyl ether)-N,N,N',N'-  
tetraacetic acid.
- ESI: Early salt-stress-induced.
- H<sup>+</sup>-ATPase: H<sup>+</sup>-adenosine triphosphatase.
- IPTG: Isopropyl  $\beta$ -D-thiogalactopyranoside.
- lea, LEA: Late-embryogenesis-abundant.

MOPS: 3-(N-morpholino)propanesulfonic acid.  
mRNA: Messenger ribonucleic acid.  
ORF: Open reading frame.  
PCR: Polymerase chain reaction.  
PEPCase: Phosphoenolpyruvate carboxylase.  
pfu: Plaque forming unit.  
P5CR:  $\Delta^1$ -Pyrroline-5-carboxylate reductase.  
P5CS:  $\Delta^1$ -Pyrroline-5-carboxylate synthetase.  
Rab, RAB: Responsive to abscisic acid.  
RNA: Ribonucleic acid.  
SDS: Sodium dodecyl sulphate.  
Ta: Annealing temperature.  
TIP: Tonoplast intrinsic protein.  
Tm: Melting temperature.  
Tris: Tris(hydroxymethyl)aminomethane.  
UV light: Ultraviolet light.  
2-D gel electrophoresis: Two-dimensional gel electrophoresis.

## PART I. INTRODUCTION

### 1. Significance of genetic studies on osmotic stress response and tolerance in plants

Plants often suffer adverse environmental conditions such as drought, saline soils and high and low temperatures. Soil salinity is a serious problem in agriculture throughout the world. Under saline conditions crops can not reach maximal yields allowed by their genetic potentials (Boyer, 1982). However, many plant species have evolved with remarkable levels of tolerance to high salinity. Understanding how these plants tolerate saline environments is potentially of great importance to agriculture. Salt sensitive crops can be possibly improved by the transfer of genetic material conferring salt tolerance from tolerant plants.

Drought and salinity may stimulate similar response from plants because both stresses cause a decrease in cellular water potential. It is reasonable to think that the mechanisms of tolerance to salinity and drought may have commonalities. Similarly, the low or freezing temperatures which lead to cellular water crystallization have stress components in common with salt stress. Therefore, in this thesis, drought and low temperature will be often discussed together with salt stress. These stresses, except low



temperature stress, will be generally referred to as osmotic or water-deficit stresses.

The mechanisms by which plants have adapted to saline environments are poorly understood. Many changes in structure, physiology and biochemistry have been observed in plants subjected to osmotic stress. These changes occur both in facultative halophytes which can grow in both saline and non-saline conditions, and in glycophytes which usually can not survive high salinity. Such responses should give us clues to the ultimate new physiological balance established in those halophytes adapted to a saline environment.

When exposed to high soil salinity, the halophyte Mesembryanthemum crystallinum switches its CO<sub>2</sub> fixation from C<sub>3</sub> photosynthesis to Crassulacean acid metabolism (CAM), a CO<sub>2</sub> fixation pathway found in plants growing in warm, dry climates (Höfner et al., 1987). Other plants increase cellular concentrations of proline, mannitol, or other compatible solutes as osmolytes to counteract environments of high osmotic pressure (Delauney and Verma, 1993). Elevated membrane-associated H<sup>+</sup>-ATPase activity, which may function in osmoregulation, has also been observed in salt-stressed plants (Braun et al., 1986).

However, neither structural nor physiological responses of

plants to osmotic stresses have been systematically studied. In recent years, people have turned their interest to changes in gene expression in plants in response to osmotic stress. A number of genes which are osmotically regulated have been isolated (Skriver and Mundy, 1990; Bray, 1993). It is apparent that salt response and salt stress tolerance are multigenic traits. Two-dimensional (2-D) gel electrophoresis of proteins translated in vitro from mRNAs from salt-treated plants revealed many genes which are up-regulated in response to salt stress and others which are down-regulated (Gulick and Dvořák, 1987; Hurkman et al., 1989). In some plants, large sets of genes induced by salt stress have been isolated and preliminarily characterized (Gulick and Dvořák, 1992; Yamaguchi-Shinozaki et al., 1992).

Although few osmotic-stress-induced genes have had their role in stress tolerance functionally defined, many of the genes showed sequence similarity to other genes coding for enzymes or defined proteins. Other stress-regulated genes encode proteins with surprisingly unique structures which may suggest their functions in osmoprotection. Moreover, the expression pattern and the mechanisms of regulation of osmotic-stress-responsive genes may also help decipher the roles of these genes in plant tolerance to stresses. Therefore, the study of the molecular biology of osmotic stress response is a powerful approach to understanding how plants survive high salinity and

other osmotic stresses. The genetic materials isolated and characterized could be of great potential in improving salt- or drought-sensitive crop plants.

Likely, not all genes induced by osmotic stress contribute to the adaptation of plants to stress. It is reasonable to think that the products of the responsive genes include important components of the plant response mechanisms that counteract the unfavourable environments. Correlation between gene expression and stress tolerance in different plant species can provide supportive evidence for the role of a gene in salt resistance.

Many osmotic-stress-induced genes also showed inducibility by other stresses such as heat shock, wounding and infections. The activation of these genes may belong to the general response of plants to environmental changes; their functions could be important in resisting osmotic stress although not specifically. Moreover, many genes developmentally activated at the desiccation stage during embryogenesis such as the lea genes are found to be induced also in vegetative growth by osmotic stress. Therefore, embryogenesis and stress adaptation may share some genetic components. In fact, the lea and lea-like genes are the largest group of genes so far found to be responsive to osmotic stress.

The internal level of the plant hormone abscisic acid (ABA) increases in plant tissues subjected to osmotic stress. Exogenously applied ABA in many cases mimics osmotic stress in stimulating similar physiological, biochemical or genetic responses from plants (Skriver and Mundy, 1990). ABA is likely a factor in signal transduction pathways of stress response, in addition to playing a regulatory role in embryogenesis (Skriver and Mundy, 1990). Studying the response of plant gene to ABA and osmotic stress will contribute to the physiological and biochemical knowledge of plant hormone action and signal transduction pathways. Therefore, ABA is also discussed together with osmotic stress in this thesis.

## 2. Osmotic-stress-induced genes and their possible functions

### 2.1. Lea and lea-like genes

A set of proteins and corresponding mRNAs were originally found to be late-embryogenesis-abundant in cotton (Dure et al., 1981; Baker et al., 1988), therefore, the genes coding for these proteins were designated leas (Galau et al., 1987). These genes appear to be transcriptionally regulated during embryogenesis because the mRNA levels increase at the same time as the increase of protein levels. The onset of lea gene expression is immediate after the abscission stage or at the

beginning of the desiccation stage. Maximal gene expression is found when the seeds are approaching maturity and the mRNA and protein levels are maintained in mature dry seeds. Both the mRNAs and proteins disappear rapidly after germination starts. ABA and desiccation seem to be involved in the developmental regulation of the lea genes (Galau et al., 1986 and 1987). The start of gene expression occurs at the same time as the internal ABA level reaches a peak value in immature embryos. The genes can also be induced by ABA or desiccation in germinating seeds or precociously in immature embryos during embryogenesis. lea-like gene expression was also found in embryogenic cells and somatic embryos in carrots during somatic embryogenesis (Zimmerman, 1993).

A number of genes with sequence similarity to the cotton leas have been isolated in many other plant species. They seem to be ubiquitous among both monocotyledonous and dicotyledonous plant species. Some lea genes also showed expression in non-embryonic tissues. However, the expression of most of the lea and the lea-like genes is related to ABA or osmotic stress (Skriver and Mundy, 1990; Bray, 1993). Some conserved sequence motifs have been identified in the amino acid sequences of their products (Dure et al., 1989; Dure, 1993b). The lea genes and their products are grouped according to their sequence similarity. At least six groups have been identified, three of these are primarily discussed here

because a considerably large amount of information for the genes in these groups has been accumulated. Most LEA proteins are highly hydrophilic. They usually contain conserved, sometimes repeating, amino acid sequences which indicate unique secondary structures. These properties, together with the fact that they are abundant in embryos when undergoing the desiccation process, suggest they may function as osmoprotectants for cellular structures during desiccation (Dure et al., 1989).

#### 2.1.1. Group 1 lea genes

Group 1 lea genes have been found in a number of dicots and monocots. They show relatively high conservation in nucleotide and amino acid sequence. Proteins encoded by these genes are very hydrophilic, containing high contents of glycine and charged amino acids. They contain a conserved 20-amino acid motif, which is tandemly repeated up to four times in some of the gene products (Espelund et al., 1992). Glycine and charged amino acid residues are more abundant in the 20-amino acid motif than in other regions of the protein. The proteins contain almost no cysteine or tryptophan.

Group 1 lea genes usually are present in the plant genome as a small gene family. The genes in this group identified so far seem to have their expression limited to seeds and young

seedlings. LeaD19 and LeaD132 in cotton and B19s in barley can be induced by desiccation, salt and mannitol treatments in immature embryos or germinating seeds (Galau et al., 1992; Espelund et al., 1992). A gene encoding a leucine zipper protein, EmBP-1, was isolated in wheat and found to bind specifically to an 8-bp nucleotide sequence CACGTGGC in the promoter region of the Em gene which encodes a group 1 LEA protein (Guiltinan et al., 1990). The sequence is called an ABA-responsive element (ABRE) since it apparently confers ABA responsiveness to the Em gene. A similar leucine zipper protein TAF-1 was found in tobacco (Oeda et al., 1991). The target sequence for TAF-1 is similar to ABRE, with a core sequence of ACGTGG. ABRE or similar sequences also exist in the rice Empl, cotton LeaD19 and LeaD132 genes and a number of other ABA responsive genes (Galau et al., 1992; Litts et al., 1992).

Litts et al. (1987) studied the hydrodynamic and optical properties of the wheat Em protein and concluded that the protein does not likely exist in the cytoplasm as a globular structure, but rather as a random coil. It has been suggested that the group 1 LEA proteins have very high affinity for water, can bind a large number of water molecules and create local aqueous environment so as to protect intracellular structures during water loss (Espelund et al., 1992).

### 2.1.2. Group 2 lea genes

The genes in lea group 2 have been isolated in many plant species including both monocots and dicots. Group 2 LEA proteins have been immunologically detected in cyanobacteria (Close and Lammers, 1993). More than 30 such gene sequences have been published (Dure, 1993b; Close et al., 1993b). Representative members of the group are the leaD11 in cotton, Rab16A-D (responsive to ABA) in rice and dhns (dehydration-induced, also called dehydrins) in barley (Baker et al., 1988; Mundy and Chua, 1988; Yamaguchi-Shinozaki et al., 1989; Close et al., 1989).

The gene products of this group are characterized, in most cases, by two separate lysine-rich repeats, one of which occurs at the carboxyl terminus of the proteins. The consensus core amino acid sequence of the repeat is KIKEKLPG. Some group 2 LEA proteins have more than two such repeats. Most dehydrins also have a cluster of 7 to 9 consecutive serine residues in the middle of the protein. A glycine-rich sequence located between the lysine-rich repeats occurs as a third motif. However, the glycine-rich sequences vary considerably among different genes. These proteins are highly hydrophilic. They are rich in glycine and threonine and contain no cysteine or tryptophan. The proteins are generally heat- and boiling-stable. This property facilitated the



isolation of the dehydrins in barley and maize (Close et al., 1989).

Many genes in this group are readily induced by ABA and a wide range of osmotic stresses such as NaCl, dehydration and low temperatures in mature plant tissues. Two such genes were identified in the resurrection plant, Craterostigma plantagineum, and are induced in the leaves when the plant starts to desiccate (Piatkowski et al., 1990). In plants so far studied, these genes often exist as a small gene family. In expression studies where gene-specific probes were used or high stringency was applied in Northern blot hybridization, some genes showed tissue- or developmental-stage-specific expression. In an ABA-deficient Arabidopsis thaliana mutant, the Rab18 gene lost its usual responsiveness to cold and drought (Lang and Palva, 1992), indicating that ABA is required for the expression of this gene. A putative ABRE, TACGTG, was found repeated twice in the 5'-upstream region of this gene. The rice Rab16 genes and the maize Rab17 gene also have one or two copies of similar 5'-upstream sequences (Mundy and Chua, 1988; Vilardeell et al., 1990).

Close et al. (1993a) used antibodies specific to the carboxyl terminal lysine-rich repeat, GEKKGIMDKIKEKLPGQH, to study the expression pattern of the dehydrin genes in different barley organs in response to dehydration. They failed to define any

uniform modes of regulation among the large number of dehydrin genes, and therefore, suggested that it was impossible to deduce the gene function by gene regulation studies. The common properties of these proteins are the consensus lysine-rich motif, the high content of hydroxylated amino acid residues and being highly hydrophilic. Baker et al. (1988) suggested the role of these proteins to be replacement of water during cellular dehydration in that the highly hydroxylated dehydrin molecules solvate structural surfaces to avoid their crystallization. In this way, the cytosol dielectric constraint would be maintained when water potential decreases, and possible collapse of the structural organization of cytosol, especially membranes, would be prevented. It should be pointed out that such a function can be fulfilled only when the proteins are amorphous random coils so that they span rather great surface distance to protect cytosolic structures. They are apparently better protectants than small molecules such as sucrose. The highly conserved lysine-rich motif must be important to the structure and function of the dehydrins, but so far their function has been primarily a subject of speculation.

### 2.1.3. Group 3 lea genes

Genes in this group encode proteins with a unique structure which implies their function in resisting water deficiency

(Dure, 1993a). The proteins are composed largely of an 11-amino acid motif tandemly repeated a number of times. The consensus sequence of the motif is

(T/A)(A/T)(Q/E)(A/T)(A/T)(K/R)(Q/E/D)(K/R)(A/T)(X)(E/D/Q).

This motif indicates an amphiphilic  $\alpha$ -helix structure. In every position except the tenth, the chemical nature of the amino acid residues, apolar, positively charged, or amide/negatively charged, is highly conserved. It should be noted that the charged amino acid residues make up from 33% to 43% of the 11-mer repeats. The repeats are sometimes interrupted by non-conserved sequences which may act as hinges between the  $\alpha$ -helices. Computer assisted molecular modelling studies and molecular sieve chromatography of the cotton LEA-D7 protein suggested that it might exist as coiled coils formed by the binding of two molecules through their hydrophobic stipes (Dure, 1993a). In this case, the hydrophilic face of the helices remains exposed.

Dure (1993a) suggested that the function of the group 3 LEA proteins was to sequester ions when cellular water potentials decreased as a consequence of desiccation or dehydration. In this case, the charged and amide residues on the hydrophilic surface of the dimer may form amino acid salts when binding ions in cytosol. The proteins therefore become ion carriers to prevent ion precipitation or crystallization. Roberts et al. (1993) determined the concentration of the cotton LEA-D7

protein to be 226  $\mu$ M in mature embryos. This number accounts for about 4% of the non-organelle cytosolic proteins. Moreover, they are evenly distributed in the cytosol and are present in all types of tissues in the embryo. These properties of the group 3 LEA proteins are apparently required for their proposed function of sequestering ions.

Some other lea gene products showed structures similar to the group 3 LEA proteins. For example, the cotton LEA-D29 protein and the cC3-06 gene product in C. plantagineum are very similar to each other in amino acid sequences, containing the 11-amino acid sequence motif in several repeats (Baker et al., 1988; Piatkowski et al., 1990). The chemical properties of the amino acid residues in each position are the same as those of the group 3 LEA proteins. Although the sequences are less conserved they may play a similar role in desiccation tolerance.

#### 2.1.4. Other lea genes

Cotton LeaD113, tomato Le25 and sunflower Hads11 code for proteins with similar amino acid sequences (Baker et al., 1988; Cohen et al., 1991; Almoguera and Jordano, 1992). Their products are very hydrophilic and may function in cells to solvate structural surfaces under water deficit conditions (Baker et al., 1988).

Another group of lea genes are the cotton LeaD34, maize Rab28 and carrot ECP31 (Baker et al., 1988; Pla et al., 1991; Kiyosue et al., 1992). Unlike most other lea genes, their gene products are relatively hydrophobic. A possible function for these genes has not been proposed.

## 2.2. Osmotin genes and proteins

The osmotin proteins were first found in NaCl-adapted tobacco cultured cells (Singh et al., 1985). They constituted 10 to 12% of total cellular protein. The amino acid sequence of osmotin is similar to those of the sweet tasting protein, thaumatin, maize  $\alpha$ -amylase/trypsin inhibitor and a tobacco pathogenesis-related protein (Singh et al., 1989).

The level of the osmotin protein in NaCl-adapted cells is stable during the growth cycle. Plants derived from adapted cells also accumulate osmotin in roots (LaRosa et al., 1992). Subsequent studies showed that the osmotin mRNA can be induced by a number of stimuli including salt shock, desiccation, UV light, ABA, ethylene, wounding and infection with tobacco mosaic virus and fungi in cultured cells and/or certain plant tissues. However, only adaptation to NaCl, ethylene and infection with fungi induced detectable protein (Raghothama et al., 1993). Therefore, the expression of the osmotin gene is likely regulated at two levels, transcriptional and post-

transcriptional. It should be noted that, unlike the lea genes, the osmotin gene is not expressed in developing embryos and mature seeds (LaRosa et al., 1992).

The function of osmotin in osmoregulation in cells is not known. However, induction of osmotin protein by ethylene and fungi led to the finding that osmotin inhibited in vitro the growth of fungal hyphae and caused hyphal and spore lysis. Osmotin is now considered to belong to the pathogenesis-related PR-5 protein family in tobacco. A similar antifungal protein in maize, zeamatin, is able to permeablize fungal plasma membranes. Osmotin could have the same antifungal function. Moreover, permeablization probably also plays a role in protecting cell membranes during desiccation or in a high salt environment through membrane reorganization (Kononowicz et al., 1992).

Similar genes were cloned from other plant species. NP24 is a tomato gene found to be induced in roots under salt treatment (King et al., 1988). A Solanum commersonii gene pA13 is regulated by ABA, cold temperature and water deficit in suspension cultures (Zhu et al., 1993).

**2.3. Phosphoenolpyruvate carboxylases and other enzymes involved in CAM**

The common ice plant M. crystallinum switches its CO<sub>2</sub> fixation from C<sub>3</sub> photosynthesis to CAM when exposed to high soil salinity (Höfner et al., 1987). CAM is a mechanism some plants have developed in adapting to dry climates, by which plants open their stomata at night for CO<sub>2</sub> uptake and close them during daytime to prevent water loss caused by evaporation. Phosphoenolpyruvate carboxylase (PEPCase) catalyses the primary fixation of CO<sub>2</sub>, which is released intercellularly during the day for the secondary fixation via photosynthesis by ribulose-bisphosphate carboxylase.

In the ice plant, PEPCase activity, PEPCase protein and its mRNA all increase substantially when plants switch to CAM due to high salinity or water stress (Höfner et al., 1987; Osterm et al., 1987). Nuclear run-on studies showed that the transcription rate of a PEPCase gene increased six-fold during salt stress or drought (Cushman et al., 1989). This indicates that the regulation of salt induction of the PEPCase gene occurs primarily at the transcriptional level. A number of PEPCase genes form a small gene family in the common ice plant. Ppc1 is believed to be involved in the salt-induced CAM. On the other hand, another gene, Ppc2, is largely unresponsive to salt treatment (Cushman et al., 1987).

The ABA level of the ice plants is elevated during salt stress. However, the expression of the Ppc1 gene is not

affected by external ABA. Interestingly, treatment with cytokinin can mimic the NaCl induction of Ppci (Thomas et al., 1992). It is likely that sensing of osmotic stress leading to the expression of the PEPCase gene is different from that for the lea genes.

Other ice plant genes encoding enzymes involved in CAM also showed NaCl inducibility. They include Gpd1 for glyceraldehyde-P-dehydrogenase-1, Imt1 for myo-inositol O-methyltransferase-1, Mod1 for malic enzyme-1, and Mdh1 for malate dehydrogenase-1 (DeRocher and Bohnert, 1993).

#### 2.4. Genes for proteins related to proteases and protease inhibitors

Salt stress inducible tobacco osmotin and the tomato NP24 protein are similar in amino acid sequence to the maize  $\alpha$ -amylase/trypsin inhibitor, as is the rapeseed protein BnD22 (Downing et al. 1992). BnD22 shows sequence similarity to the members of the soybean Kunitz protease inhibitor family. The expression of the gene coding for the protein is organ specific. It is highly inducible in leaves by drought, NaCl or ABA, to a lesser extent in hypocotyls, but not in roots. That it is not expressed in seeds makes it a rare case for the ABA- and water-stress-induced genes. Downing et al. (1992) suggested that protease inhibitors have a defensive role



against exogenous proteases, or more possibly, they may regulate endogenous proteases to prevent certain degradation processes caused by water deficiency.

A gene encoding a thiol protease-like protein was reported from wilted pea shoots (Guerrero et al., 1990). The gene is induced when the turgor of shoots is reduced to zero. The amino acid sequence of the gene product shares 41% identity with a cysteine protease from Dictyostelium. Proteases may degrade denatured proteins caused by stress and the resulting free amino acids could be used for synthesis of new proteins.

The increased level of certain proteases in plant cells upon stress could also be explained to actively alter metabolism by activating or inhibiting specific enzymes or proteins through digestion (Schaffer and Fischer, 1988). Proteases and protease inhibitors may function in concert in regulation when the specificity of their substrates and their specific intracellular localizations are concerned.

#### 2.5. Genes related to biosynthesis of proline and other osmolytes

Some plant species accumulate proline in cells which are under osmotic stress. As an osmolyte, proline may counteract high osmotic potential in the exterior environment. Accumulation

of other compatible organic solutes, such as betaine, mannitol, sorbitol and pinitol, was also reported (Delauney and Verma, 1993). Although there is lack of evidence that proline plays an adaptive role in plant osmoregulation, the finding by mutation studies that proline indeed is an osmoprotectant in Escherichia coli suggests that proline accumulation is a good candidate for studying mechanisms of osmotic stress tolerance in plants.

It has been found that the accumulation of proline in plants by osmotic stress is primarily due to the induction of proline biosynthesis. Proline is synthesized mainly from glutamic acid in stressed plants. Two genes coding for enzymes  $\Delta^1$ -pyrroline-5-carboxylate synthetase (P5CS) and  $\Delta^1$ -pyrroline-5-carboxylate reductase (P5CR) involved in this pathway have been cloned (Delauney and Verma, 1990; Hu et al., 1992; Williamson and Slocum, 1992; Verbruggen et al., 1993). mRNAs for both genes are induced by NaCl treatment. However, it has been shown in tobacco that elevated P5CR mRNA level and enzymatic activity induced by salt treatment does not contribute to the increased cellular proline concentration (LaRosa et al., 1991; Szoke et al., 1991). On the other hand, P5CS, a bifunctional enzyme catalyzing the first two steps of reactions in the glutamate pathway of proline biosynthesis, may be subjected to feedback inhibition by proline. Reduced sensitivity to the feedback inhibition of this enzyme in moth

bean may explain the drought tolerance of this plant (Delauney and Verma, 1993).

In spinach, a gene for betaine-aldehyde dehydrogenase (BADH) has been cloned (Weretilnyk and Hanson, 1990). The levels of the BADH activity, BADH protein and translatable mRNA increase under salt stress. The enzyme is involved in the biosynthesis of betaine, which is thought to be a nontoxic or protective cytoplasmic osmolyte accumulating during water deficit.

Moreover, a bacterial gene encoding mannitol-1-phosphate dehydrogenase expressed in transgenic tobacco led to the accumulation of mannitol, a polyol which may function similarly to proline in osmoadjustment (Tarczynski et al., 1992). As a result, the transgenic plants showed increased salt tolerance (Tarczynski et al., 1993).

## 2.6. H<sup>+</sup>-ATPases

There are two principal ways in plants to maintain normal cytoplasmic NaCl concentrations during NaCl influx in a high salinity environment. One is to exclude Na<sup>+</sup> through plasmalemma; the other is to sequester Na<sup>+</sup> into vacuoles (Flowers et al., 1977; Greenway and Munns, 1980). This active transport of NaCl seems to be facilitated by the Na<sup>+</sup>/H<sup>+</sup>-antiport which requires a proton gradient across the membrane

(Gabarino and Dupont, 1989; Staal et al., 1991). The proton gradient is likely generated by H<sup>+</sup>-ATPase (Sze, 1985). A number of studies demonstrated increases in plant H<sup>+</sup>-ATPase activity in plasmalemma or tonoplasts under NaCl stress (Braun et al., 1986; Reuveni et al., 1990). Niu et al. (1993) showed elevated H<sup>+</sup>-ATPase mRNA levels in both glycophyte tobacco and halophyte Atriplex nummularia in response to high NaCl concentrations. The inducibility in roots is substantially greater in the halophyte than in the glycophyte. This might be an indication that H<sup>+</sup>-ATPase is really a salt tolerance determinant.

## 2.7. Other osmotic-stress-responsive genes

A wheat cDNA related to serine/threonine protein kinase has been isolated (Anderberg and Walker-Simmons, 1992). The gene coding for it is expressed in developing seeds and is inducible by ABA in embryos and by dehydration in seedlings. Protein kinases can regulate gene expression via trans-acting factors which are DNA-binding proteins. Kinases can also alter the activities of other classes of proteins or enzymes. Such changes could lead to expression of novel proteins or enzymes, or, to enhanced activity of existing proteins, which are necessary for plants to adapt to stress.

Some osmotic-stress-responsive genes code for putative

transmembrane proteins. The desiccation-upregulated A. thaliana protein RD28 and the turgor-loss-upregulated pea protein 7A share sequence similarity and both contain six membrane-spanning domains (Guerrero et al., 1990; Yamaguchi-Shinozaki et al., 1992). Proteins with similar structure have been widely found in bacteria, plants and mammals. Some of them have been identified to be channels for transmembrane transport of small molecules, for example, the 27-kDa tonoplast intrinsic protein ( $\gamma$ -TIP) in A. thaliana was shown to create water channels in Xenopus oocytes (Maurel et al., 1993). Such a function suggests this class of proteins may have a role in osmoregulation of cytoplasm.

As mentioned above, the wheat EmBP-1 protein binds to a putative ABRE of the ABA- and osmotic-stress-regulated genes and may function in regulating expression of these genes. Another example is an A. thaliana gene, Atmyb2, which is a homolog of the myb genes (Urao et al., 1993). Some myb gene products have been defined as specific transcription activators in humans and plants. Atmyb2 is induced by ABA, dehydration and salt and is expressed during seed maturation like the lea genes. The gene product ATMYB2 was demonstrated in vitro to bind specifically oligonucleotides containing the sequence TAACTG which is a consensus MYB protein recognition sequence. Therefore, this gene may activate other stress-regulated genes.

There are reports of an ABA- and water-stress-induced maize gene encoding an RNA-binding protein (Gómez *et al.*, 1988; Ludevid *et al.*, 1992). The gene, MAH9, is also expressed in developing and mature embryos. The protein is rich in glycine and has a consensus sequence type RNA-binding domain. In vitro studies showed that it bound preferentially to poly(G) and, to a lesser extent, poly(U). Because of its RNA-binding property, this protein may have a role in posttranscriptional regulation of stress-related genes.

There are a number of other osmotic-stress-induced genes reported, for example, the rice salt-stress- and drought-responsive gene salt (Claes *et al.*, 1990), the alfalfa glycine-rich gene SM2075 (Luo *et al.*, 1991), and the A. thaliana desiccation-responsive rd29 gene (Yamaguchi-Shinozaki and Shinozaki, 1993). Their products do not have sequence similarity to any functionally known enzymes or protein factors. Neither could any possible function be speculated from their amino acid sequences.

### 3. Regulation of expression of osmotic-stress-induced genes

The new physiological balance a plant establishes when adapting to salinity or water deficit should start with the sensing of changes in its environment. It is widely accepted that loss of turgor pressure caused by osmotic-stress triggers

the cellular response from plants. Thereafter, through a single, or, more likely, multiple signal transduction pathways these environmental signals ultimately lead to production of functional proteins. Little is known about the signal transduction pathways. However, it seems that ABA plays an important role in transmitting the environmental signals. Many stress-inducible genes can also be induced by ABA. Cis-acting regulatory elements responsible for an ABA-specific response have been found in the promoter regions of a number of osmotic-stress-induced genes. Bray (1993) classified these genes into three categories according to their relationships to ABA: Genes non-responsive to ABA; genes responsive to ABA but not requiring ABA for expression; and genes requiring ABA for expression.

Seed morphogenesis includes the synthesis of storage proteins, the development of desiccation resistance, and the onset and maintenance of dormancy (Skriver and Mundy, 1990). During embryogenesis, the ABA level peaks almost at the same time as the onset of seed desiccation. The lea genes start to be expressed shortly afterwards. Many lea genes are also responsive to ABA in vegetative tissues. In vegetative tissues, the ABA level itself is often elevated upon osmotic stress. Therefore, it is likely that after the loss of turgor pressure is sensed by cells, ABA biosynthesis is initiated, and then some genes are activated by the increased level of

ABA. Drought may regulate a certain step in ABA biosynthesis (Parry, 1993). Gene regulation by ABA is possibly at the transcriptional level since the ABREs are found in the promoters of some stress-induced genes. The G-box-like ABREs containing a core sequence of C/TACGTGG in the rice Rab16s, wheat Em and many other genes have been shown to confer ABA responsiveness in transgenic plants (Guiltinan et al., 1990). There are likely other yet unidentified ABRE sequences in plants.

Genes requiring ABA for expression can be identified in ABA-deficient A. thaliana aba or maize viviparous2 mutants, for example, the cold-acclimation-related Rab18 gene (Lang and Palva, 1992) and the water-stress-inducible Rab28 gene (Pla et al., 1991). The ABA-deficient mutants can also be used to define genes which are responsive to ABA but do not require ABA for their expression. The A. thaliana desiccation-responsive gene rd29A is an example (Yamaguchi-Shinozaki and Shinozaki, 1993). This gene contains at least two kinds of cis-acting elements, one responsive to ABA, the other, DRE, is responsive to dehydration (Yamaguchi-Shinozaki and Shinozaki, 1994). The 9-bp DRE, TACCGACAT, is quite different from the G-box-like ABRE and is not responsive to ABA. This indicates that there are multiple signal transduction pathways which can regulate rd29A. However, many other ABA-responsive genes are not yet characterized in terms of their ABA requirement.



Some genes are inducible by osmotic stresses but are not responsive to ABA, for example, the cold specific Wcs120 gene in wheat (Houde et al., 1992) and the NaCl inducible Ppc1 gene coding for PEPCase (Thomas et al., 1992). The environmental signals might act directly on these genes, or the genes might be controlled by signal transduction pathways involving factors other than ABA.

Regulation at the post-transcriptional level of stress-responsive genes has been reported, but the significance of such regulation is not fully understood. In young tobacco leaves, both ABA and NaCl stimulate accumulation of osmotin mRNA. However, only NaCl but not ABA stimulates accumulation of osmotin proteins. This indicates there is a mechanism, by which NaCl affects gene expression, either directly or indirectly, that affects the stability or the processing of osmotin mRNA, or the translation of mRNA, or the stability of synthesized osmotin protein (LaRosa et al., 1992). Detection of protein is largely dependent on the availability of specific antibodies. Most osmotic-stress-induced genes have not been investigated with immunodetection methods.

The maize Rab17 gene products have been shown to be phosphorylated in vivo at the site of a serine cluster, probably by protein kinase similar to the casein kinase 2 (Vilardell et al., 1990). The phosphorylation is restricted

to the developing embryos and ABA-treated germinating seeds. Some osmotic-stress-inducible genes encode proteins containing leader sequences which are cleaved in conjunction with transport to their cellular destination. Such a putative signal peptide is found in osmotin by comparison of the N-terminal part of the amino acid sequence deduced from its cDNA and the sequence determined on the purified protein from plant tissues (Singh et al., 1989). Many other proteins showed a discrepancy between their predicted molecular weights and those determined by electrophoresis under denaturing conditions. This likely indicates the posttranslational processes of cleavage, phosphorylation, or glycosylation of the gene products.

The spatial expression of the stress-induced-genes is also important. It may imply or support the speculated role of these genes in response to and resistance against the stresses. The cotton LEA proteins D7 and D113 are found to be distributed evenly in cytosol, at high cellular concentrations, in all cell types in maturing seeds (Roberts et al., 1993). This is in accordance with the proposed function as osmoprotectants for these LEA proteins.  $H^+$ -ATPase is thought to play an important role in cellular osmoregulation to counteract high salinity environments. In the halophyte A. nummularia, induction of this gene by NaCl is much higher in roots than in leaves (Niu et al., 1993). It is

probably due to the fact that roots are the primary barrier to the uptake of  $\text{Na}^+$  and  $\text{Cl}^-$ .

The mechanisms for the regulation of the osmotic-stress-induced genes are complex; it seems that there is not a universal model for the regulation of expression of these genes. Extensive studies are needed to elucidate these mechanisms.

#### 4. Genetic response to salt stress in the salt-tolerant wheatgrass Lophopyrum elongatum

A number of wheatgrass species, which are phylogenetically related to cultivated wheats, are highly salt tolerant (McGuire and Dvořák, 1981). Among them, the diploid Lophopyrum elongatum (Host) A. Love (syn. Elytrigia elongata [Host] Nevski, Agropyrum elongatum Host,  $2n = 2x = 14$ ) has been extensively studied for its salt tolerance. This species is found in salt marshes in the Mediterranean region. An amphiploid from a cross between the salt-tolerant L. elongatum and salt-sensitive bread wheat cultivar Triticum aestivum L. cv. Chinese Spring ( $2n = 6x = 42$ ) showed elevated salt tolerance compared to Chinese Spring (Omielan et al., 1991). This indicates that the salt tolerance trait is dominant over sensitivity and the genetic determinants for salt tolerance in the wheatgrass can be expressed in the wheat background.

Therefore, L. elongatum is an excellent material to study how plants tolerate saline environments, and such studies could be of great significance for the improvement of salt tolerance of cultivated cereals.

Gene expression has been characterized by 2-D gel electrophoresis of proteins from in vitro translation of mRNAs and shown to be altered in the roots of the amphiploid when stressed during a time course of seven days with 250 mM NaCl (Gulick and Dvořák, 1987). No changes were observed in the translatable mRNA population from leaves. Since L. elongatum is a salt excluder, the subsequent studies on its genetic response to salt stress were mainly focused on the roots. Eleven partial cDNA clones were obtained from mRNAs from L. elongatum roots stressed for 6 h with 250 mM NaCl by differential screening of an enriched cDNA library (Gulick and Dvořák, 1990). Nucleotide sequence analysis showed that the clones represent different genes. There is no or very low expression of the genes corresponding to these clones in plants grown in normal conditions. When stressed with 250 mM NaCl, steady state mRNA levels in roots can be elevated within 2 h, reach peak values at about 6 h, and decline afterwards (Gulick and Dvořák, 1992). All the genes are induced only in roots, except for ESI15 which is also induced in leaves. In this way, these genes have similar expression patterns, and they are considered to represent the early stage of genetic

response to salt stress in L. elongatum. Therefore, the genes are designated ESI for 'early salt-stress-induced'.

These genes are also induced by salt in Chinese Spring wheat and the amphiploid derived from L. elongatum and Chinese Spring. For most of the ESI genes, their induction and expression is higher in the wheatgrass than in the amphiploid, and higher in the amphiploid than in Chinese Spring (Galvez et al., 1993). The response in roots of the wheatgrass of these genes to other stimuli were investigated (Galvez et al., 1993). There is no difference in response to NaCl and KCl. The genes can also be induced by mannitol and ABA. Therefore, the induction of the genes may represent a general response to osmotic stress, and such response may be mediated by ABA.

The ESI3 gene is one of these salt-induced genes. The induction of this gene in the roots of the salt-tolerant L. elongatum by 250 mM NaCl is 4.02 fold; whereas in the salt-sensitive wheat it is only 1.56 fold (Galvez et al., 1993). The original cloned cDNA is 199 bp in length, apparently a partial cDNA clone (Gulick and Dvořák, 1992). Subsequently, a 432-bp cDNA clone, ESI3-11, was isolated and sequenced. ESI3-11 and the originally isolated cDNA are derived from the same gene because the sequences of the overlapping region of the two are identical.

Among the 11 salt-induced genes, ESI18 is unique in that it shows four bands in the Northern blot analysis. Indeed, subsequent cloning work showed that it represents a multigene family. Four different cDNAs which cross-hybridize to each other have been identified. Another salt-induced gene, ESI35, shows sequence similarity to the ESI18 clones but did not hybridize to them. However, all the genes are considered to encode dehydrin-like proteins since the predicted amino acid sequences obtained from their cDNA clones, partial or full length, contain the typical lysine-rich consensus motif found in the dehydrin proteins. The full length cDNA clones for the ESI18-10 and ESI35 genes have already been obtained (Gulick and An, 1993; unpublished data). Only partial cDNAs were available for ESI18-1, -2 and -15.

Cloning and characterization of the ESI3 gene and the ESI18 multigene family are the subject of this work. The objectives are to obtain full or near-full length cDNA clones of these genes and to predict the amino acid sequences of their gene products. The polymorphism of the sequences and primary structures of the dehydrin-like genes and their products will be discussed. The significance and possible functions of these genes in relation to the salt stress response and salt tolerance will also be discussed.

## PART II. MATERIALS AND METHODS

All the procedures used in the course of this work will be described below, at least briefly. However, the standard methods for DNA and RNA manipulation were according to a laboratory manual by Sambrook et al. (1989). The reagents used in this study were of analytical or biochemical grades. For enzymes and other special proteins and chemicals, their manufacturers or suppliers are indicated. In case reagents supplied by manufacturers were used in conjunction with kits, the methods involved will be described whenever possible.

### 1. Plant materials

Hydroponically grown plants were used in this study. Seeds of L. elongatum were soaked in 5% (w/v) bleach for 15 min and rinsed extensively with sterile distilled water. The sterilized seeds were placed on wet filter paper on slant boards and imbibed at 4°C in the dark for three days before being transferred to the greenhouse for germination. The bottom edge of the filter paper was immersed in distilled water to keep it wet. When the roots of the seedlings reached 5 to 10 cm, the seedlings were transferred to culture tanks with modified Hoagland solution containing 3 mM KNO<sub>3</sub>, 1 mM NH<sub>4</sub>H<sub>2</sub>PO<sub>4</sub>, 0.5 mM MgSO<sub>4</sub>, 5.5 mM Ca(NO<sub>3</sub>)<sub>2</sub>, 90 µM FeEDTA, 25 µM KCl, 12.5 µM H<sub>3</sub>BO<sub>3</sub>, 1 µM MnSO<sub>4</sub>, 1 µM ZnSO<sub>4</sub>, 0.25 µM CuSO<sub>4</sub>, and

2  $\mu\text{M}$   $\text{H}_2\text{MoO}_4$  (Gulick and Dvořák, 1987).

For salt treatment, plants were moved to fresh Hoagland solution 24 h prior to treatment with 250 mM NaCl. After being treated for 2, 6, 12, or 24 h respectively, the roots were removed and immediately frozen in liquid nitrogen and stored at  $-80^\circ\text{C}$ . Plants used for untreated controls were maintained in non-salinized nutrient solution for 6 h before the roots were collected.

## 2. Extraction of total RNA and isolation of poly(A)<sup>+</sup>RNA

Frozen root tissues were ground to a very fine power in liquid nitrogen with pestle and mortar. Five millilitres of extraction mixture preheated to  $65^\circ\text{C}$  was immediately added for each gram of ground tissue. The extraction mixture was composed of equal volumes of chloroform and phenol saturated with a solution containing 0.2 M sodium borate, 30 mM EGTA, 2% (w/v) SDS, and 1% sodium lauryl sarcosinate. The tissues and the extraction mixture were blended for 3 to 4 min with a Polytron homogenizer and subsequently incubated at  $22^\circ\text{C}$  for 10 min. The extract was centrifuged at  $10,000 \times g$  for 15 min, and the supernatants were collected and extracted once with the extraction mixture and twice with chloroform. The final supernatants were mixed with equal volumes of 4 M LiCl and incubated at  $4^\circ\text{C}$  for 24 h. After centrifugation, the pellets



were washed in 2 M LiCl, dried, dissolved in 10 mM Tris-HCl, pH 7.5, and centrifuged again. RNA in the supernatants was precipitated with sodium acetate and ethanol at  $-20^{\circ}\text{C}$  for 24 h. Finally, the total RNA was collected by centrifugation and dissolved in distilled water. Samples were used directly or reprecipitated with sodium acetate and ethanol for storage.

To isolate poly(A)<sup>+</sup>RNA, total RNA was equilibrated with 20 mM Tris-HCl, pH 7.5, 0.5 M NaCl, 1 mM EDTA, and 0.1% sodium lauryl sarcosinate, and passed through an oligo(dT) column (Type 7, Pharmacia). The column was rinsed with 10 volumes of loading buffer, and the poly(A)<sup>+</sup>RNA was subsequently eluted with a solution containing 10 mM Tris-HCl, pH 7.5, 1 mM EDTA, and 0.05% SDS.

### 3. Plasmid isolation and purification

Plasmids were normally isolated from E. coli strain XL1-Blue by the boiling method as recommended by Stratagene. Briefly, bacteria cultured in liquid LB medium (1% NaCl, 1% tryptone, 0.5% yeast extract) were collected by centrifugation and resuspended in 8% (w/v) sucrose, 0.5% Triton X-100, 50 mM Tris-HCl, pH 8.0, 50 mM EDTA, and 0.5 mg/ml lysozyme (from egg white, Sigma). The suspension was incubated in boiling water for 30 sec and centrifuged to remove the pellet. An equal volume of isopropanol was added to the supernatant and

centrifuged to precipitate DNA. The DNA pellet was dissolved in TE buffer (10mM Tris-HCl, pH 8.0, 1 mM EDTA) containing 0.1 mg/ml RNaseA and incubated at 22<sup>o</sup>C for 30 min before phenol/chloroform extraction and ethanol precipitation. The DNA was pelleted and washed in 70% (v/v) ethanol, dried, and finally dissolved in TE buffer for further use.

Alternatively, QIAprep-spin Plasmid Kits (QIAGEN) were used to prepare plasmids. This involved the alkaline lysis of bacteria and the adsorption of DNA onto silica membrane at high salt. For larger scale plasmid preparation, the QIAGEN Plasmid Midi Kit (QIAGEN) was used.

#### 4. Digestion of DNA with restriction endonucleases

The restriction endonucleases were from BIO/CAN or Pharmacia. The conditions for the enzymatic reactions were recommended by the suppliers. Usually, digestion was done with between 2 to 10 units of enzyme per µg of DNA.

#### 5. Polymerase chain reaction

Polymerase chain reactions (PCR) were based on the methods in the manual by Innis et al. (1990). The melting temperature (T<sub>m</sub>) for a primer 17 to 28 nucleotides in length was estimated according to the formula  $T_m (^{\circ}C) = (\text{number of G and C}) \times 4 +$

(number of A and T) x 2. The annealing temperature ( $T_a$ ) was selected as  $5^{\circ}\text{C}$  below the lower  $T_m$  of the two primers. In a typical reaction, there were  $10^4$  to  $10^8$  copies of target DNA,  $0.2\ \mu\text{M}$  of each primer,  $0.2\ \text{mM}$  each of dATP, dCTP, dGTP and dTTP,  $0.1\ \text{mg/ml}$  bovine serum albumin,  $0.025\ \text{u}/\mu\text{l}$  Taq polymerase (BIO/CAN),  $50\ \text{mM}$  Tris-HCl, pH 9.0,  $1.5\ \text{mM}$   $\text{MgCl}_2$ ,  $0.1\%$  Triton X-100. The reaction mixtures were covered with a layer of mineral oil. The thermal cycles were carried out in an automatic thermal reactor (HYBAID). In the initial cycle, the reaction was at  $95^{\circ}\text{C}$  for 5 min for denaturing, then at the  $T_a$  for 30 sec, and at  $72^{\circ}\text{C}$  for 30 to 60 sec for extension of DNA. This was followed by 29 to 34 cycles of  $95^{\circ}\text{C}/30\ \text{sec}$ ,  $T_a/30\ \text{sec}$ ,  $72^{\circ}\text{C}/30\ \text{to}\ 60\ \text{sec}$ . Finally, the reaction mixture was incubated at  $72^{\circ}\text{C}$  for a further 10 min to complete the synthesis of DNA molecules.

#### 6. Agarose gel electrophoresis of DNA and purification of DNA fragments

Agarose gel electrophoresis was carried out in  $1 \times$  TAE buffer ( $0.04\ \text{M}$  Tris-acetate, pH 8.0,  $1\ \text{mM}$  EDTA,). The concentrations of agarose ranged from  $0.7\%$  to  $1.5\%$  (w/v) according to the size of DNA fragments of interest. The gels contained ethidium bromide at  $0.5\ \mu\text{g/ml}$ . The electrophoresed DNA fragments were observed by ultraviolet (UV) light.

To isolate a DNA fragment from an agarose gel, the region of the DNA band was cut from the gel and the QIAEX Extraction Kit (QIAGEN) was used to extract DNA from the gel. The purification was achieved by the binding of DNA to the QIAEX beads and the subsequent elution of DNA with TE buffer.

#### 7. Preparation of radioactive probes

Gel-purified DNA fragments were radio-labelled and used as probes in hybridization. The fragments were obtained from restriction enzyme digestion of plasmids or were PCR products from amplification of a certain region of a clone. The Amersham Multiprime Labelling System was used for labelling. Briefly, 25 ng of DNA were denatured by boiling in a water bath and mixed with dATP, dGTP, dTTP, [ $\alpha$ - $^{32}$ P]dCTP (3,000 Ci/mole, ICN), random primers, and Klenow enzyme, and incubated at 37°C for 30 min. The unincorporated radioactivity was separated from the labelled DNA products by passing the reaction mixture through a Sephadex G-50 (Pharmacia) column packed in a Pasteur pipet. The probe was collected in the void fraction. The total activity of the probes derived from 25 ng of template DNA was usually from  $1 \times 10^7$  to  $4 \times 10^7$  cpm.

#### 8. Northern blot analysis

The Northern blot analysis was performed according to the protocol of Fourney *et al.* (1988) with some modifications. Ten micrograms of total RNA or 1  $\mu$ g of poly(A)<sup>+</sup>RNA were precipitated with ethanol and redissolved in 1 x MOPS buffer (0.05 M sodium acetate, 0.02 M MOPS, pH 7.0, 1 mM EDTA), 6% (w/v) formaldehyde, 50% (v/v) formamide, 5 mg/ml bromophenol blue, 65  $\mu$ g/ml ethidium bromide. The samples were loaded onto a 1% agarose gel made with 1 x MOPS buffer and 6% (w/v) formaldehyde. The electrophoresis buffer was 1 x MOPS. The electrophoresis was carried out at 60 to 80 volts for 3 to 4 h. After electrophoresis, the gel was briefly washed in 20 x SSC (3 M NaCl, 0.3 M sodium citrate). The RNA was transferred to a Hybond-N nylon filter (Amersham) by capillary blotting in 20 x SSC for 16 h. The filter was then baked at 80°C for 2 h and used subsequently for hybridization.

#### 9. cDNA library screening and plasmid rescue

The cDNA libraries used in this study were constructed with Stratagene's ZAP cDNA synthesis system. The  $\lambda$  phage vector Uni-ZAP XR contains pBluescript plasmid vector sequence, which can be excised *in vivo* in the presence of f1 helper phage. *E. coli* strain XL1-Blue (recA1, endA1, gyrA96, thi-1, hsdR17, supE44, relA1, lac, [F' proAB, lacI<sup>q</sup>ZAM15, Tn10(tet<sup>r</sup>)]) was used in screening for the genes of interest. Amplified libraries were plated at 10,000 to 50,000 pfu per 150-mm plate

in NZY medium (0.5% NaCl, 0.2% MgSO<sub>4</sub>·H<sub>2</sub>O, 0.5% yeast extract, 1% casein hydrolysate, pH 7.5). After growth at 37°C for 16 h, the plaques were lifted onto Hybond-N filters. The filters were then denatured in 1.5 M NaOH, 0.5 M NaCl for 7 min, neutralized twice in 1.5 M NaCl, 0.5 M Tris-HCl, pH 7.2, for 3 min, washed briefly in 2 x SSC, and baked at 80°C for 2 h according to Amersham's instructions.

After hybridization (described below) with gene-specific probes, the plaques corresponding to signals on the autoradiograph were taken for a second round of plating and hybridization to purify the positive clones. Plasmids were rescued from λ phages by in vivo excision with the f1 helper phage R408. λ phage and the helper phage were mixed with XL1-Blue cells in 2 x YT medium (1% NaCl, 1% yeast extract, 1.6% [w/v] tryptone) and incubated at 37°C for 3 h and then heated at 70°C for 20 min. The cells were pelleted and the supernatants containing phagemids were used to infect fresh XL1-Blue cells. The infected bacterial cells were plated onto LB agar plates containing 50 µg/ml ampicillin and grown at 42°C (to inhibit growth of cells co-infected with the helper phage) for 16 h. The single colonies were streaked on LB-ampicillin plates and grown at 37°C for 16 h. Thus, the cDNAs were recovered in the pBluescript vector.

## 10. Hybridization

Hybridization was performed according to the instructions from the manufacturer of the Hybond-N nylon filter (Amersham). The filters containing fixed DNA or RNA were prehybridized at 65°C for 1 h in the hybridization solution containing 5 x SSPE (20 x SSPE = 3.6 M NaCl, 0.2 M sodium phosphate buffer, pH 7.7, 0.02 M EDTA), 5 x Denhardt's solution (1 x = 0.02% bovine serum albumin, 0.02% Ficoll, 0.02% polyvinylpyrrolidone), 0.5% SDS, and 100 µg/ml sonicated denatured salmon sperm DNA. After prehybridization, the radio-labelled probe was denatured and added to the filters for further incubation at 65°C for 12 to 24 h. Then, the filters were washed 2 to 6 times in 1 x SSPE, 0.1% SDS at 22°C for 15 min. Afterwards, for low stringency conditions, the filters were washed twice further in 1 x SSPE, 0.1% SDS at 50°C for 1 h each; or, for high stringency conditions, they were washed in 0.1% SSPE, 0.1% SDS at 65°C for 1 h each. Fuji medical X-ray films were used for autoradiography which was carried out at -80°C with various durations of exposure. DUPOND Cronex Lightning Plus intensifying screens were often used to enhance signals.

#### 11. Preparation of serial deletion clones for sequencing

Plasmid DNA was digested to completion with a pair of restriction enzymes for the purpose of generating serial deletion clones by using the Promega Erase-A-Base system. One of the enzymes resulting 5'-overhang ends was proximal to the

cdNA insert; the other resulting 3'-overhang ends was distal to the cdNA insert. For a 5'-deletion series, the two restriction sites were at the 5'-side of the insert. For a 3'-deletion series, two sites at the 3'-side of the insert were selected for digestion. The restriction-enzyme-digested plasmid DNA was then treated with Exo III nuclease at 30°C. Aliquots were taken every 1 min and the reaction was stopped, treated with S1 nuclease and Klenow enzyme. In this way, DNA fragments deleted from one end for approximately 200 bp at each time point were obtained. These fragments were circularized with the T4 DNA ligase and used to transform competent XL1-Blue cells. The insert size of the deletion clones were checked by digestion with PvuII and proper clones were selected for sequencing.

## 12. DNA sequencing

To determine the nucleotide sequence of DNA, 3 µg of plasmid DNA were denatured in 0.2 M NaOH, 0.2 mM EDTA at 37°C for 30 min and neutralized with the addition of sodium acetate, pH 5.2, to 0.3 M. The denatured DNA was precipitated with ethanol, pelleted, washed in 70% ethanol and dissolved in distilled water for sequencing reaction with the Sequenase Kit (Version 1.0, United States Biochemicals). [ $\alpha$ -<sup>35</sup>S]dATP (600 Ci/mmol, ICN) was used for labelling DNA. The final reaction mixtures were heated at 80°C for 2 min before being loaded



onto a 6% (w/v) polyacrylamide sequencing gel (containing 8 M urea, acrylamide:N,N'-methylene-bis-acrylamide = 19:1) for electrophoresis in TBE buffer (0.045 Tris-borate, pH 8.0, 1 mM EDTA). Then the gel was fixed in 10% (v/v) acetic acid, 10% (v/v) methanol 3 times for 2 min each and dried. Fuji X-ray films were exposed to the dried gel at 22°C for autoradiography.

When the full length cDNA clones of the ESI18 genes were to be sequenced, sets of plasmids from the serial deletion clones were prepared and sent to Dr. L. Pelcher's laboratory at the Plant Biotechnology Institute, National Research Council, Saskatoon, Saskatchewan. In this case, high temperature cycle sequencing was employed, and gel electrophoresis and reading was performed with an automated Sequencer. Sequences were compiled and proof-read by myself, and remaining ambiguities were resolved by further sequencing by myself as described above.

### 13. Primer extension

An oligonucleotide complementary to a sequence of ESI3F close to the 5'-end was end-labelled in a 20- $\mu$ l reaction mixture containing 10 pmole oligonucleotide, 50  $\mu$ Ci [ $\gamma$ -<sup>32</sup>P]ATP (4,500 Ci/mmole, ICN), 8 u T4 polynucleotide kinase (BIO/CAN), 20 mM Tris-HCl, pH 7.6, 5 mM MgCl<sub>2</sub> and 5 mM DTT. The mixture was

incubated at 37°C for 45 min and then heated at 65°C to inactivate the kinase. After phenol/chloroform extraction, the labelled primer was precipitated with ethanol and redissolved in water. Poly(A)<sup>+</sup>RNA from roots of salt-treated plants was used as template in the extension of the primer. One microgram of RNA was mixed with 4 pmole labelled primer, 16 u RNAGuard (Pharmacia) and 25 u Moloney murine leukemia virus (M-MuLV) reverse transcriptase (Pharmacia) in a volume of 30 µl containing 50 mM Tris-HCl, pH 7.5, 5 mM MgSO<sub>4</sub>, 5 mM DTT, 100 mM KCl, and 1 mM of each dATP, dCTP, dGTP and dTTP. After incubation at 37°C for 1 h, the synthesized cDNA was precipitated with ethanol and redissolved in TE buffer. The size of the extension products was determined by electrophoresis on a sequencing gel. A sequencing reaction with the ESI3F clone with the same primer was used for size comparison.

#### 14. Inverse PCR cloning

Inverse PCR cloning procedures (Ochman et al., 1990) were used to isolate genomic DNA corresponding to the region 5' to a nearly full length cDNA clone of ESI3. Two micrograms of L. elongatum DNA (from H. An, see An [1994]) were digested with MspI and circularized by T4 DNA ligase (BIO/CAN) at a concentration of 5 µg DNA per ml. A pair of primers with opposite orientations were used for amplification by PCR.

There was an XhoI site in one of the primers. The PCR product which was displayed as a sharp band by agarose gel electrophoresis was purified from the gel, digested with MspI and XhoI. The resulting MspI-XhoI fragment was inserted into the ClalI and XhoI sites in the pBluescript SK- vector. Competent E. coli XL1-Blue cells were heat-shocked in the presence of the ligated DNA and plated on LB agar plates containing 50 µg/ml ampicillin, 0.04 mg/ml X-gal, 0.2 mg/ml IPTG. After incubation at 37°C for 16 h, white colonies were selected for further characterization.

#### 15. Computer-assisted data analysis

Analysis of nucleotide sequences of genes and amino acid sequences of proteins was mainly achieved through IntelliGenetics' PC/Gene computer programs (Release 6.70, by A. Bairoch, University of Geneva, Switzerland). The software was used for data entry, manipulation, comparison and identity search, and other applications. The access to gene databases was through the NCBI (National Center for Biotechnology Information) BLAST (Basic Local Alignment Search Tool) E-mail Server system (Altschul et al., 1990). Sequences were sent for similarity comparison. Also, sequences and related information were obtained from databases through the NCBI Retrieve E-mail Server.

## PART III. RESULTS

### 1. Characterization of the ESI3 gene

#### 1.1. Isolation of cDNA clones for the ESI3 gene

A cDNA library derived from the mRNAs isolated from the roots of L. elongatum treated with 250 mM NaCl for 6 h and constructed in the  $\lambda$  Uni-ZAP XR vector (Stratagene) was available for this work. To isolate clones corresponding to the full length cDNA sequence for the ESI3 gene, the cDNA insert of the ESI3-11 clone was radio-labelled and used as probe in screening the cDNA library. Approximately 250,000 pfu were screened at high stringency and 20 ESI3-positive clones were selected for further identification. The ESI3F clone contains a cDNA insert of 530 bp. In the 5'-region, ESI3F has overlapping sequence relative to ESI3-11, as well as 53 bp extra at the 5'-end; the 3'-region contains sequence identical to ESI3-11 as well as additional 45 bp (excluding the poly[A] tail) at the 3'-end. Therefore, ESI3F is the longest cDNA clone for the ESI3 gene obtained (Figure 1). There is no sequence near the polyadenylation site of ESI3F similar to the consensus plant polyadenylation signal sequence, A/TAATAAPu (Heidecker and Messing, 1986). The additional 45 bp in this region of ESI3F, compared to ESI3-11, indicates putative multiple polyadenylation sites on the ESI3



gene.

### 1.2. Determination of transcription start site and confirmation of ESI3F to be a nearly-full length cDNA clone

An oligonucleotide complementary to a sequence near the 5'-end of the cDNA ESI3F (corresponding to the nucleotide positions 81 to 65 in Figure 1) was used in the reverse-transcription primer extension experiment. As shown in Figure 2, the extension product was a single band in the autoradiograph corresponding to a position 16 bases upstream from the first base of cDNA ESI3F. It indicates that ESI3F is a nearly-full length cDNA lacking only 16 bp at the 5'-end. The transcription start site for the ESI3 gene was thus determined. The size of transcript predicted from this sequence, when the poly(A) tail is considered, is in accordance with the data from the Northern blot analysis, which revealed a single band at about 700 bp (not shown).

### 1.3. Cloning the 5'-flanking region of the ESI3 gene

The inverse PCR method (Ochman et al., 1990) was used for the cloning of the 5'-flanking sequence of the ESI3 gene. A pair of ESI3F-specific primers with opposite orientations were used in PCR amplification of the circularized, MspI-digested genomic DNA fragments of L. elongatum (Figure 3). The primer

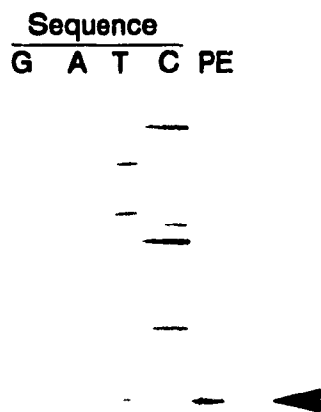


Figure 2. Primer extension assay to determine transcription start site for the ESI3 gene. A primer complementary to the positions +81 to +65 (as numbered in Figure 1) was used in reverse transcription of poly(A)<sup>+</sup>RNA from salt-treated roots of L. elongatum. Lane PE shows the extension product indicated by the arrow. The nucleotide sequence shown in the sequence lanes was from a sequencing reaction on the ESI3F cDNA clone with the oligonucleotide used for the primer extension.

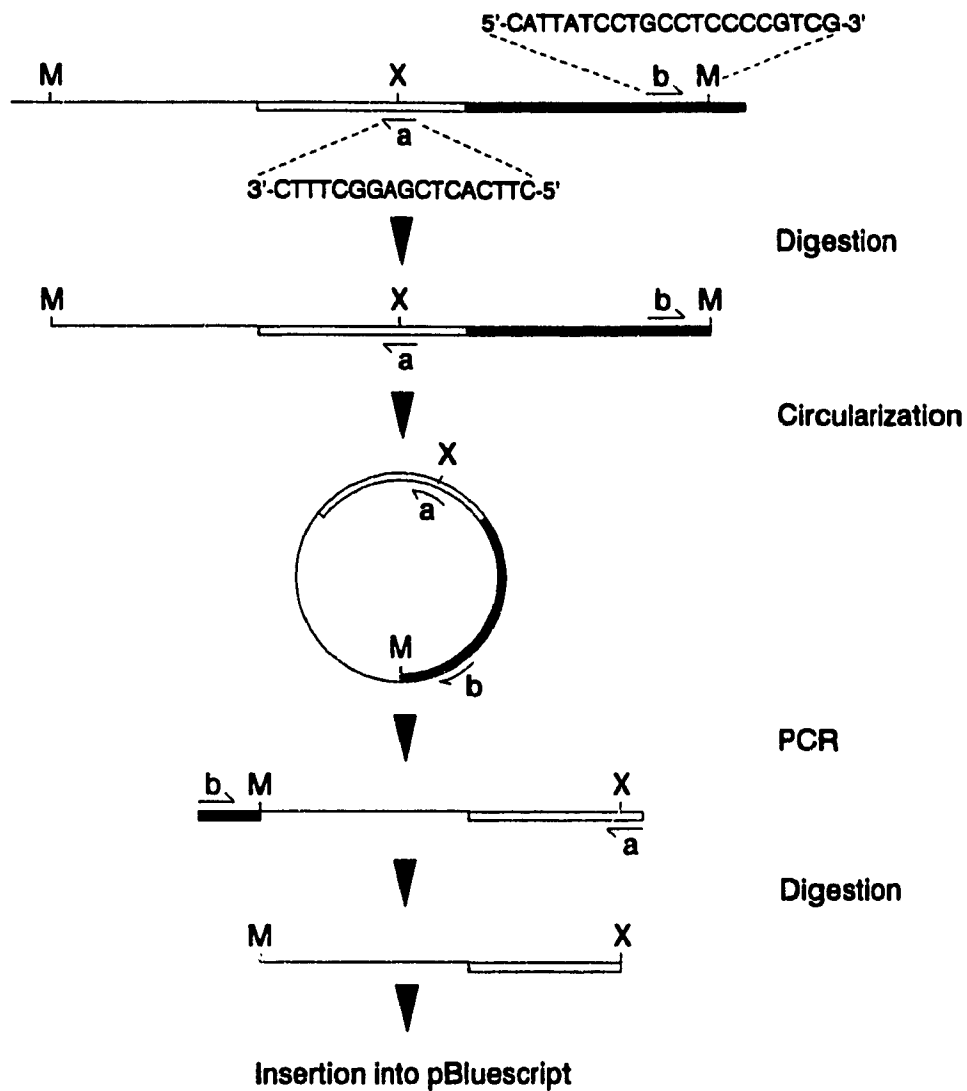


Figure 3. Strategies for the cloning of the 5'-flanking sequence of the *ESI3* gene by inverse PCR. The primers, *a* and *b*, and their orientations are indicated by the arrows. Line, 5'-flanking sequence; open box, 5'-non-coding region; darkened box, coding region. *M*, *Msp*I; *X*, *Xho*I.



toward the 5'-end corresponds to the nucleotide positions 81 to 65; and the primer toward the 3'-end corresponds to positions 130 to 149, as shown in Figure 1. The PCR product was digested with XhoI and MspI, and the 182-bp fragment was cloned in pBluescript as clone ESI3F2. This clone was sequenced in both directions. The sequence of ESI3F2 overlaps with that of ESI3F for a fully identical 60 bp, so that both clones can be considered to represent the same gene, that is, ESI3. Thus, the ESI3F2 clone includes the 5'-end sequence of the ESI3 gene transcript and 106 bp of genomic DNA immediately upstream of the transcribed region. Figure 1 shows the combined 546 bp of cDNA and 106 bp of genomic DNA sequences of the ESI3 gene.

In the 5'-flanking region, there is a TATA-Box-like sequence, TATATAAA, at position -32. At -101, the sequence GACACGTACAC is similar to the ABREs found in many ABA-responsive genes, for example, the one described in the wheat Em1b gene at -94, CACACGTCCGC (Litts et al., 1991).

#### 1.4. Analysis of the predicted ESI3 gene product

The cDNA sequence for the ESI3 gene contains a single plausible open reading frame (ORF) which is 162 bp in length. The predicted polypeptide from translation of this ORF contains 54 amino acid residues and has a molecular weight of

5.9 kDa. The amino acid sequence of the predicted ESI3 protein is shown in Figure 1. Notably, 50% of the amino acid residues are leucine, valine, or isoleucine. It contains no histidine, glutamine, asparagine, or aspartic acid. The overall hydrophobicity of the protein is considerably high. Figure 4 shows the hydrophobic analysis of the ESI3 protein by the method of Kyte and Doolittle (1982). It revealed two hydrophobic regions. These regions correspond to the two possible membrane-spanning segments predicted by the method of Klein et al. (1985). One is from amino acid positions 6 to 22; the other is from positions 38 to 54 (Figure 1). Therefore, the ESI3 protein is likely an integral membrane protein.

## 2. Isolation and characterization of cDNA clones for members of the ESI18 multigene family

### 2.1. Isolation of full length cDNAs for members of the ESI18 family

The cDNA library mentioned in the beginning of this part was also used in the isolation of additional members of the ESI18 family as well as the isolation of full or near-full length cDNAs for the ESI18-2 and ESI18-15 genes. As a result, cDNA clone ESI18-17, for the ESI18-17 gene, a novel member of the family, was obtained when the cDNA insert of the ESI18-10 was

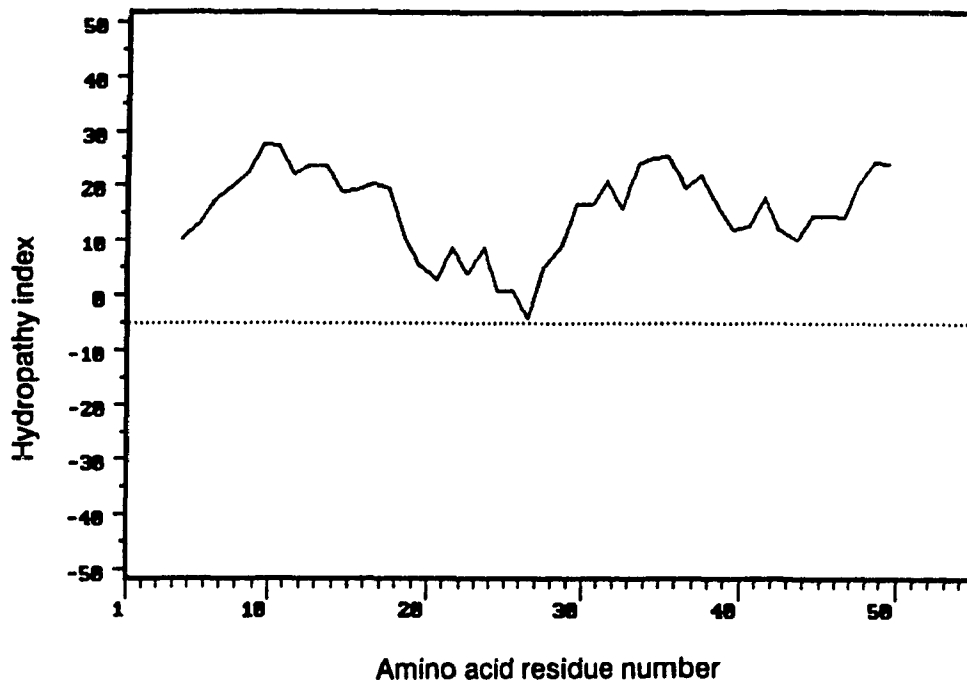


Figure 4. Hydropathy index computation for the predicted ESI3 protein. The method for the analysis was after Kyte and Doolittle (1982) with a window of nine amino acid residues. The values greater than -5 are considered to be hydrophobic; those smaller than -5 are hydrophilic.

used as probe in screening. Clones ESI18-2-28 and ESI18-15-34, for the ESI18-2 and ESI18-15 genes respectively, were obtained through probing the libraries with PCR amplified 3'-non-coding regions of the genes. Subclones from the restriction fragments or from serial deletions of cDNA inserts were constructed for the purpose of nucleotide sequence determination. Both strands of the cDNAs were sequenced. Figure 5 shows the nucleotide sequences of the three cDNAs. The cDNA inserts excluding the poly(A) tails are 1518, 761 and 680 bp for clones ESI18-2-28, ESI18-15-34 and ESI18-17 respectively. The transcript sizes are 1600, 1100 and 700 bases for the ESI18-2, -15 and -17 genes respectively, measured in the Northern blot analysis with their 3'-non-coding regions as probes (not shown). Therefore, these clones correspond to nearly full length cDNAs if the poly(A) tails of the mRNAs are considered. The plant polyadenylation signal, A/TAATAAPu, which is usually 15 to 23 nucleotides upstream to the end of mRNA (Heidecker and Messing, 1986), is not found near the polyadenylation sites of all these cDNAs.

These members of the ESI18 multigene family, including ESI18-2, -15 and -17, as well as previously identified members ESI18-1 (An, 1994) and ESI18-10 (unpublished data) are salt-inducible since the steady state mRNA levels for the genes in roots of L. elongatum were elevated when treated with 250 mM NaCl for 6 h. This was demonstrated by Northern blot analysis

ESI18-2-28

```

GGGAAGTACACAGCAGCACTTGTAGATTTC60CGGGTGAGAAGCTCAGCGCAAGATGGAGC
ACCAGGCACACGGCGCCGGCGAGAAGAAGGGCATCATGGAGAAGATCAAGGAGAAGCTCC 120
CCGGCGGCCACGGTGACCACAAGCAGACCGCTGGCACCCACGGGCATGCCGGCACGGCGA 180
CGCACGGGCCCCCGGCCACTGGTGGTGCCTACGGGCAGGAGGGTCACACCGGAACCACTG 240
GCACGGGGTTGCATGGCGCCGACGCCATGGAGAACATCAAGGACAAGCTCCCCGGTGGCC 300
ACGGTGACCACCAGCAGACTGCTGGCGCCTATGGGCAGCAGGGACATACCGGCACGGCGA 360
CGCATGGCACCCCGGCCACCGGTGGCACCTATGGTGAGCAGGGACACACCGGAGTGACCG 420
GCACGGGGACGCACGACGCCGACACCGGCGAGAAGAAGGGCGTCATGGAGAACATCAAGG 480
ACAAGCTTCCTGGTGGCCATGGGGACCACCAGCAGACCGGTGGTACCTACGCACAGCAGG 540
GACACACCGGCACGGCGACACATGGCACCCAGCCGGCGGTGGCACCTATGACTATGAGC 600
AGCAGGGACATACCGGGATGACCGGCACGGGGACACACGGCACAGGCGAGAAGAAGGGTG 660
TCATGGAGAACATCAAGGAGAAGCTCCCCGGTGGCCATGCTGACCACCAACAAACCACTG 720
GCACCTACGGACAGCAGGGTCATGTCGGCACGGGGACGCATGGCACCCACGCCACCGGAG 780
GCGCCTACGGGCAGCATGAACACACCGGAGTGACCGGCACGGGAACCCACGGCACCGGCG 840
AGAAGAAAGGGCGTCATGGAGAACATCAAGGACAAGCTCCCCGGTGGCCACGGTGACCACC 900
AACAGACCGGTGGCACCTACGGGCAGCAGGCACACATCAGCACGGGGACGCATGACACGC 960
CGGCCACCGGCGACACCTACGGGCAGCAGGCACACACCGGAGTGACCAGCACGGGGACGC 1020
ACGGCACCGGCGAGAAGAAGGGCGTCATGGAAAACATCAAGGAGAAGCTCCCCGGTGGCC 1080
ACAGTGACCACCAGCAGACCACATGACACCTACGGGCAGCATGGACATGCTGGAGTGACCG 1140
GCACGGAGACGCATGGCACACCGGCCACCGGCGGCACCTACGGGCAGCAGGGACACACCG 1200
GCACGACCGGCACTGGGACTCACGGCACCGGACCGGCACCGGCGAGAAGAAGAGCCTCATGG 1260
ACAAGATCAAGGACAAGCTGCCTGGACAGCACTAAGTCCGGTGCCTAGCGGCGGCTACC 1320
TCTGAAGAATAATAAGATGGCGAACTTCCACCGTATATATACATTGATTGAGTCTAGTTC 1380
ACCTAGCTCACGCACTTGATCGTTGGAGGAGCGAATGTATCTGGTTTAAGTTTTCACTGG 1440
CAACAGTGTGTTTACAGTTTGGAAATCACTTGT1500TTTTGCCCGACTGATAAGAGCGATCCG
TAAGCCGCCTCTCTTTC(A)n 1518

```

Figure 5. Nucleotide sequences of the cDNAs ESI18-2-28, ESI18-15-34 and ESI18-17. Both strands of the cDNAs were sequenced in entirety, but only those of the sense strands are shown. The start and stop codons of the longest open reading frames for the predicted proteins are underlined. (A)<sub>n</sub> indicates the 3'-terminal poly(A) tail.

ESI18-15-34

CACCTGTGCAAGATGGAGTACCAGGGACAGCAGCAGCATGATCAGGGCAGGACCAACCGC	60
GTCGACGAGTACGGTAACCCGGTGGCCGGACATGGCGTCGGCACCGGCATGGGCGCGCAC	120
GGCGGCGTGGGCACCGGCGCGGCCGCTGGTGGGCATTTCCAGCCCACGAGGGAGGAGCAC	180
AAGGCCGGAGGGATCCTGCAGCGCTCCGGCAGCTCTAGTAGCTCCAGCTCGTCTGAGGAT	240
GATGGCATGGGCGGGAGGAGGAAGAAGGGAATCAAGGATAAGATCAAGGAGAAGCTTCCG	300
GGTGGCCACGGCGATCAGCAGCAGACCGCTGGTACCTACGGACAGCAAGGTCATACTGGA	360
ATGGCCGGCACCGGCGGCAACTACGGCCAGCCGGGACACACCGGAATGGCTGGCACCGAC	420
GGCACCGGTGAGAAGAAGGGCATCATGGACAAGATCAAGGAGAAGCTGCCCGGACAGCAC	480
TGAGCCC GGCCCGCGGCAGCTACTCGAGAGTTGAGGCGCCGATCTGGCCACCTTTGCAGA	540
ATAATAA.GATGGAGATGCAGTATAACTTCCCGAAATAAAGTGAGCTAGCTCACTTGTAAT	600
GTCTGAGTTCGAGTCTGGTGGACTTGAATTTGGGCTGCTTGTATGTACCCGGGATGTTT	660
TGTGCTCTGTACTTCCATGTGTGAATTTCTTTGGTTGTACTGTGTATGTATTATATGG	720
GGATGTATATCATGGTATAATGCAATATATTTTGCCTTTT(A) <sub>n</sub>	761

ESI18-17

GGGACAAGTTGAGCGCAAGATGGAGCACCAGGGGCACCGCGCAGGCACCGGCGAGAAGAA	60
GGGCGTCGTGGAGAGCATCACGGAGAAGCTCCCCGTTGGCCACGGTGATCACCAACAGAC	120
CACCGGTGGCACGTACGGGCAGCAGGGACATGCCGGAGTTACCAGCACAGGCACCGGCGA	180
GAAGAAGGGTGTTCGTGAGAACATCAAGGAGAAGCTCCCCGGTGGCCACGGTGGCCACCA	240
GCACACCACTGGAATGACCGGCTCGGAGACGCATGCCACCACGGCCACCAGCGATGGCAA	300
CTACGGGCAGTCGGGACATAACCGCACTGACGGCACCGGTGAGAA/CAAGAGCATCATGGA	360
CAAGATCAAGGACAAGCTGCCTGGACAGCACTAAGCCCAACCGGTCTGCCCGCGCCCGGC	420
CCGCTACCTTTGCAGAAATATAATAAGAAGGCCAACTTCCACCGTGTATACATTAATTG	480
AGTCTAGTTCAACTAGCTCACTTGGTAGTTGGAGGAGCGCGTGATTTTCATCTGGATTTAA	540
GCTTTCATGGACAACAGTGTGTGAGGGTTTTCGTTTGTTTACAATTTGCATACTTTGAT	600
GTGGAAATTTCTTTTGGTTCAACTCGTGGATGGTATGCTATGGCATCTAATATATATAT	660
AAGCAAAATGTTAAGTTTGC(A) <sub>n</sub>	680

Figure 5. (continued.)

with the 3'-non-coding regions, which were specific for each gene, as probes (not shown).

## 2.2. Analysis of the predicted ESI18 gene products

### 2.2.1. The ESI18-15 protein

Figure 6 shows the amino acid sequences of the predicted ESI18-2, ESI18-15 and ESI18-17 proteins. ESI18-15 is composed of 156 amino acid residues, among which 24.3% are glycine, 16.7% positively charged, and 10.3% negatively charged. It contains no cysteine or tryptophan. The calculated molecular weight is 15.9 kDa. The protein is very hydrophilic according to the hydrophobic analysis (not shown). The typical dehydrin lysine-rich motif occurs twice in ESI18-15. One of them, RKKGIKDKIKEKLPGGH, is in the middle of the protein; the other, EKKGIMDKIKEKLPQGH, is at the carboxyl terminus. A cluster of 8 serine residues, which is also typical in most dehydrins, is found just before the first lysine-rich repeat. Between the two lysine-rich sequences is a glycine-rich region with two repeats: TAGTYGQQGHTGMAG and TGGNYGQPGHTGMAG (Figure 9).

### 2.2.2. The ESI18-2 protein

The predicted ESI18-2 protein is composed of 413 amino acid residues with a molecular weight of 41.7 kDa (Figure 6). It

ESI18-2

MEHQAHGAGEKKGIMEKIKEKLPGGHGDHKQTAGTHGHAGTATHGAPATGGAYGQEGHTG 60  
TTGTGLHGADAMENIKDKLPGGHGDHQQTAGAYGQQGHTGTATHGTPATGGTYGEQGHTG 120  
VTGTGTHDADTGEKKGVMENIKDKLPGGHGDHQQTGGTYAQQGHTGTATHGTPAGGGTYD 180  
YEQQGHTGMTGTGTHGTGEKKGVMENIKEKLPGGHADHQQTGTGTYGQQGHVGTGTHGTPA 240  
TGGAYGQHEHTGVTGTGTHGTGEKKGVMENIKDKLPGGHGDHQQTGGTYGQQAHISTGTH 300  
DTPATGDTYGQHGHTGVTSTGTHGTGEKKGVMENIKEKLPGGHSDHQQTDTYGQHGHTG 360  
VTGTETHGTTATGGTYGQQGHTGTTGTGTHGTDGTGEKKSMDKIKDKLPGQH 413

ESI18-15

MEYQQQQQHDQATTNRVDEYGNPVAGHGVGTGMGAHGGVGTGAAAGGHFQPTREEHKAGG 60  
ILQRSGSSSSSSSEDDGMGRRRKKGIKDKIKEKLPGGHGDQQQTAGTYGQQGHTGMAGT 120  
GGNYGQPGHTGMAGTDGTGEKKGIMDKIKEKLPGQH 156

ESI18-17

MEHQHRACTGEKKGVVESITEKLPVGHGDHQQTGGTYGQQGHAGVTSTGTGEKKGVVE 60  
NIKEKLPGGHGGHQHTTGMTGSETHATTATSDGNYGQSGHTGTDGTGENKSIMDKIKDKL 120  
PGQH 124

Figure 6. Amino acid sequences of the predicted products of the genes ESI18-2, ESI18-15 and ESI18-17. The lysine-rich repeats are underlined.



is also the largest dehydrin-like gene for which a full length cDNA clone has been identified in L. elongatum. It is rich in glycine (24.9%) and threonine (17.9%), but contains no arginine, phenylalanine, cysteine, or tryptophan. The protein is also highly hydrophilic. There are seven lysine-rich repeats in ESI18-2. The first lysine-rich repeat is 9 amino acid residues away from the amino terminus, the last one is at the carboxyl terminus, and the rest are distributed within the polypeptide at almost equal intervals. The second lysine repeat is imperfect, missing the first few amino acid residues. The sequences between the lysine-rich repeats contain two copies of a glycine-rich motif, TGPTYGQQGHTG in consensus, which is similar to that found in the ESI18-15 protein. ESI18-2 does not have the serine cluster (Figure 9).

### 2.2.3. The ESI18-17 protein

The deduced amino acid sequence of the ESI18-17 protein is shown in Figure 6. This protein is 124 amino acid residues in length and 12.8 kDa in molecular weight. It has similar amino acid composition to that of ESI18-2, with glycine at 21.7% and threonine at 14.5%, and without phenylalanine, cysteine, or tryptophan. Like ESI18-2, it does not contain the serine cluster sequence. There are three lysine-rich repeats. One is close to the amino terminus, and another at the carboxyl terminus. The first lysine repeat is imperfect in amino acid

sequence. The glycine-rich motif can also be found between the lysine repeats. However, in the first interval of the lysine sequences there is only one copy of the glycine motif, and in the second interval there are two imperfect copies of this motif (Figure 9).

### 3. Expression of dehydrin-like genes in wild wheatgrass species with differences in salt tolerance

To study the correlation between the expression or induction of the dehydrin-like genes in different wild wheatgrass species and their tolerance against salt stress, steady state mRNA levels for these genes during salt treatment were investigated. Four species whose sensitivity to salt stress had previously been determined were chosen on the basis of high or low sensitivity (Table 1, data from McGuire and Dvořák [1981]). Lophopyrum pontica is highly salt-tolerant, Lophopyrum elongatum is less but still very tolerant, Lophopyrum caespitosa is rather sensitive to salt stress, and Lophopyrum intermedia is the most salt-sensitive. Total RNAs were extracted from the roots of these plants treated with 250 mM NaCl for 2, 6, 12, or 24 h, and subjected to Northern blot assay. The probe was a PCR-amplified 121-bp DNA fragment coding for the carboxyl terminus of the barley DHN3 protein including the conservative lysine-rich repeat (Close et al., 1989). The 121-bp barley sequence is 92.5% similar to the

Table 1. Correlation of induction of dehydrin-like genes and salt tolerance in different wheatgrass Lophopyrum species. The survival data of the wild wheatgrass species are from McGuire and Dvořák (1981), which indicate the percentages of hydroponically grown plants surviving after treated with 500 mM NaCl for 26 days. The gene induction was determined with Northern blot analysis of RNA from plant roots treated with 250 mM NaCl in a time course of 24 h, with a DNA probe derived from the barley DHN3 gene. + indicates weak induction; +++ indicates very strong induction.

Species	Accession	Survival rate at 500 mM NaCl	Induction of dehydrin-like genes
<u>L. elongatum</u>	2E54	90%	+++
<u>L. pontica</u>	PI315312	100%	+++
<u>L. intermedia</u>	PI238222	0%	+
<u>L. caespitosa</u>	PI229914	4%	+

corresponding sequence of the L. elongatum ESI18-15 gene, 89.3% to ESI18-2, 87.6% to ESI18-1, 86.6% to ESI18-17, 85.6% to ESI18-10, and 70.2% to ESI35. Therefore, this heterologous barley probe is able to detect most of the L. elongatum dehydrin-like genes in low stringency hybridization; moreover, with this probe, the effect of divergence of sequences among the wheatgrass species can be diminished. The results are shown in Figure 7. The untreated roots of the four species showed rather weak mRNA levels for the dehydrin-like genes. When treated with salt, the mRNA levels were elevated in all the species, but to different extents. All the four species demonstrated multiple RNA bands with similar, although not identical, patterns of induction during the time course of the salt treatment. Overall, salt-tolerant L. pontica and L. elongatum demonstrated much higher steady state mRNA levels than salt-sensitive L. caespitosa and L. intermedia throughout the duration of the salt treatment. The correlation between the salt stress induction of dehydrin genes and salt tolerance among these four wheatgrass species is apparent (Table 1).

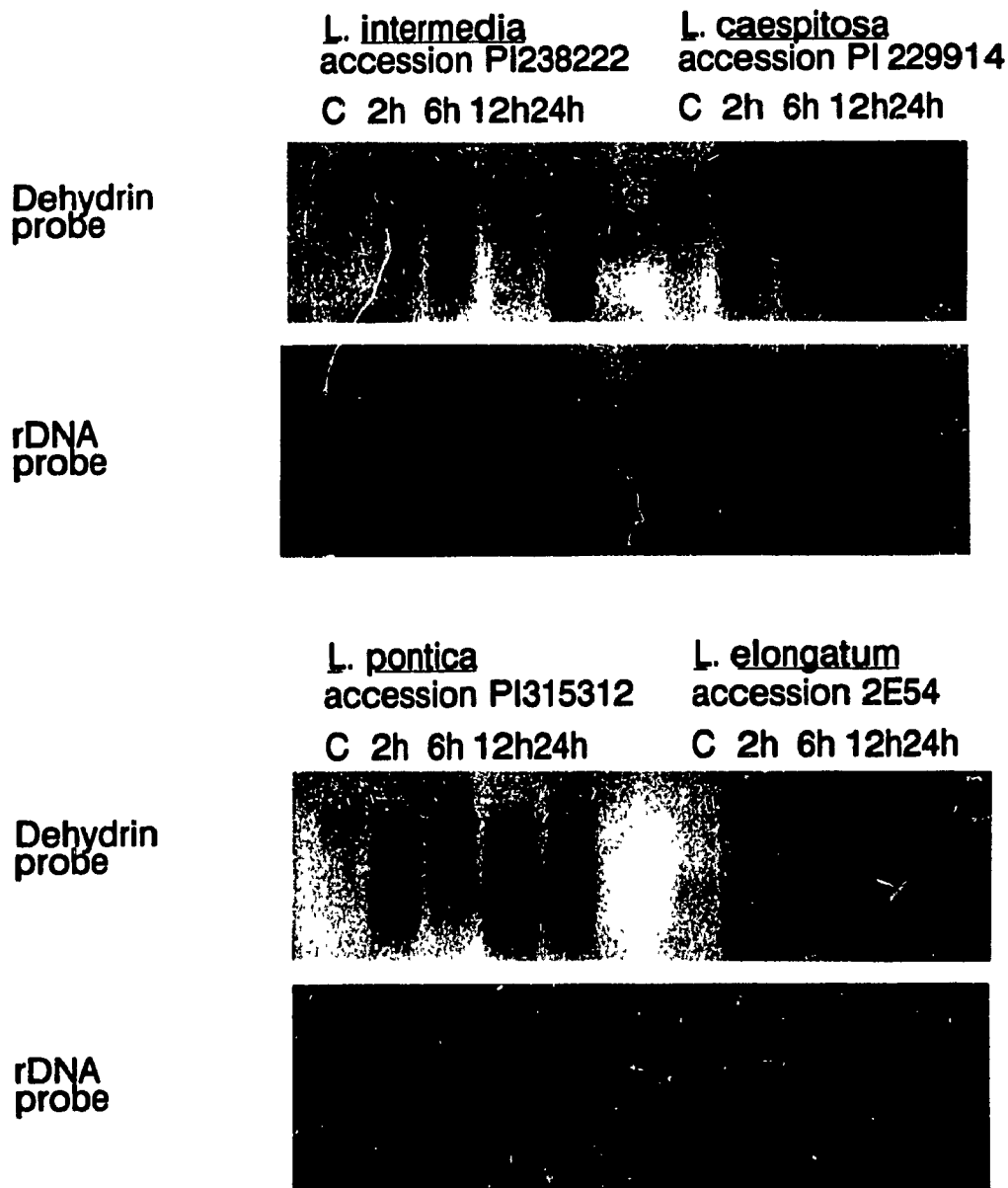


Figure 7. Steady state mRNA levels for dehydrin-like genes of four Lophopyrum species during the time course of exposure to salt stress. RNAs were extracted from the plant roots treated with 250 mM NaCl for the various time periods as indicated and subjected to Northern blot hybridization probed with the 121-bp barley DHN3 cDNA fragment corresponding to the carboxyl terminus of the gene product. Control RNAs shown in Lanes C were from plant roots which were not salt-treated. The same blots were stripped of the dehydrin probe and hybridized again with an rDNA fragment probe for the purpose of internal control.

## PART IV. DISCUSSION

### 1. The ESI3 gene

#### 1.1. Database search for proteins similar to ESI3

In this study, the nucleotide sequence of the cDNA for the L. elongatum ESI3 gene was determined and the predicted amino acid sequence of the putative gene product was analyzed. The ESI3 gene encodes a small hydrophobic protein with two possible membrane-spanning segments. Protein databases were searched for proteins with similar amino acid sequences to that of the predicted product of the ESI3 gene. The Basic Local Alignment Search Tool (BLAST, Altschul et al., 1990) was used for the search (done on April 7, 1994). The results indicate that three (putative) proteins have high scores of sequence similarity to the ESI3 gene product. Remarkably, the barley low-temperature-inducible gene blt101 product has a predicted amino acid sequence identical to ESI3, whereas its cDNA nucleotide sequence of the coding region has 92.0% similarity to that for ESI3 (Goddard et al., 1993). Since barley and wheatgrass are phylogenetically close, it is reasonable to conclude that these two genes are homologous. The ESI3 protein also shows 47% similarity in amino acid sequence to ZK632.10 from Caenorhabditis elegans and 34% similarity in amino acid sequence to a yeast hypothetical

protein (Dean-Johnson and Henry, 1989). Both putative proteins from C. elegans and yeast have not been ascribed any functions. However, it should be noted that the identical amino acid residues occur mostly in the putative membrane-spanning regions in ESI3 (Figure 8).

## 1.2. Possible functions of the ESI3 gene in salt resistance

The barley blt101 gene was up-regulated by low temperatures, but not by ABA or drought in shoot meristems; its expression in roots in response to stress was not investigated. However, the identical amino acid sequences of the wheatgrass ESI3 and barley blt101 gene products suggest the functional importance of such plant proteins in stress response or resistance. They are likely specifically active in response to water deficiency caused by high salinity or low temperatures.

Most of the salt-stress- or water-stress-induced genes in plants, for example, the lea and lea-like genes, encode hydrophilic proteins. The salt-stress-induced L. elongatum ESI3 gene is quite different in this aspect since it encodes a small and very hydrophobic polypeptide. Such a unique primary structure does not likely allow the ESI3 protein to be an enzyme. However, there are two regions with amino acid sequence conservation among the ESI3 protein and two other proteins from yeast and C. elegans (figure 8). Notably, these

		*   * * * *   * * *   * * *	
ZK632.10	(1)	MCQILLAILAIFLPPIAVLLDVGC-N	(25)
ESI3	(1)	<u>MGSATVLEVILAILPPVGVFLRYKL</u> -G	(27)
BLT101	(1)	<u>MGSATVLEVILAILPPVGVFLRYKL</u> -G	(27)
Y15.2	(17)	-INKDDL <del>LL</del> MVLAVFIPPVAVWKRKGMFN	(44)
		*   * *   * * * *   * * *	
		*   * * *   * *   * * * *   * * *	
ZK632.10	(26)	CDLLINILLTCLGIIPGIIHAWYIILC-	(52)
ESI3	(28)	VEFWICLLLTILGYIPGIIYAVYVLVV	(54)
BLT101	(28)	VEFWICLLLTILGYIPGIIYAVYVLVV	(54)
Y15.2	(45)	RD <del>LLN</del> LLLFLLFFPAIIHACYVVYE-	(71)
		* * *   *   * * * *   * * *	

Figure 8. Comparison of the ESI3 protein and proteins with similar amino acid sequences. BLT101, a barley protein (Goddard et al., 1993); ZK632.10, a hypothetical protein from *Caenorhabditis elegans*; Y15.2, a 15.2-kDa hypothetical yeast protein from (Dean-Johnson and Henry, 1989). The asterisks on the top of the sequences indicate the consensus among the proteins from *L. elongatum*, barley and *C. elegans*; the asterisks below the sequences indicate the consensus among those from *L. elongatum*, barley and yeast. The putative membrane-spanning sequences in the ESI3 protein are underlined. Only partial sequences of the yeast and *C. elegans* proteins are shown.



two regions fall within the two predicted membrane-spanning segments of the ESI3 protein. It implies that these proteins are possibly integral membrane polypeptides.

A possible function of the ESI3 protein is that it forms membrane channels for transporting small molecules. Yamaguchi-Shinozaki et al. (1992) reported the isolation of the desiccation-induced gene, RD28, which codes for a putative transmembrane channel protein. The RD28 protein containing six membrane-spanning segments could theoretically form a molecular channel by itself. However, there are only two such domains in the ESI3 molecule. If the ESI3 protein functions in a transmembrane channel, oligomers of the ESI3 protein molecules must be formed to fulfil such an activity. It is reasonable to speculate that in a high NaCl concentration environment, the influx and exclusion of ions or osmotica through membranes must be very active in plant cells in order to maintain normal cellular osmotic balance for the survival of the cells. Such speculation can also be applied to the more general response to water deficiency stress, in that plant cells need to adjust their contents by adsorbing or excluding specific molecules to counteract water loss. Thus, transmembrane channels for transportation of ions or other molecules are needed for these purposes. ESI3 might play a role in transportation of specific molecules.

Another possibility is that ESI3, as an integral membrane protein, may function as a structural protein to maintain the membrane integrity since cellular water loss caused by dehydration or high salinity often results in the disintegration of the membrane structures.

The location of the ESI3 proteins in cells should be determined to support one or the other possible function of this protein. Such an investigation usually requires that specific antibodies be raised against the protein and used in immunocytochemical assays. Although the ESI3 protein is generally hydrophobic the region between the two membrane-spanning domains is relatively hydrophilic and possibly antigenic (Figure 1).

## 2. The ESI18 genes and other L. elongatum dehydrin-like genes

### 2.1. The predicted gene products

ESI18, a multigene family coding for proteins with conserved lysine-rich repeats, as well as ESI35, belongs to a group of plant genes usually referred to as dehydrin-like genes which are widely found in monocots and dicots. Most of these genes are responsive to water-deficiency-related stress and ABA (see Part I. Introduction). Full length cDNAs for the three L. elongatum dehydrin-like genes, ESI18-2, ESI18-15 and ESI18-17,

were isolated and characterized in this study. Previously, full length cDNAs were already available for the ESI18-10 (unpublished data) and ESI35 (Gulick and An, 1993) genes. A partial length cDNA clone was available for another gene, ESI18-1 (An, 1994). Compared to the data published so far, the numbers of the dehydrin-like genes isolated in L. elongatum and barley are the largest in a single plant species (Close et al., 1989 and 1993b); four have been described in rice (Yamaguchi-Shinozaki et al., 1989).

The proteins encoded by the L. elongatum dehydrin-like genes display extensive variations in their primary structures, including the number of lysine-rich repeats, the occurrence of a serine cluster and the usually glycine-rich sequence between the lysine-rich repeats. Figure 9 shows a diagram of the structures of these proteins. There is a lysine-rich domain at the carboxyl terminus in all these dehydrins except ESI35 where there are 19 additional amino acid residues. Serine clusters are found only in ESI18-15 and ESI35. The serine clusters are located immediately upstream of the lysine-rich repeat most distal to the carboxyl termini. There are two lysine-rich repeats in ESI18-15 and three in ESI35. ESI18-2, -10 and -17 are relatively simple: Seven lysine-rich repeats occur in ESI18-2 and three in ESI18-10 and -17. The sequences between the lysine-rich repeats of all these proteins except ESI35 are rich in glycine and threonine. The amino acid

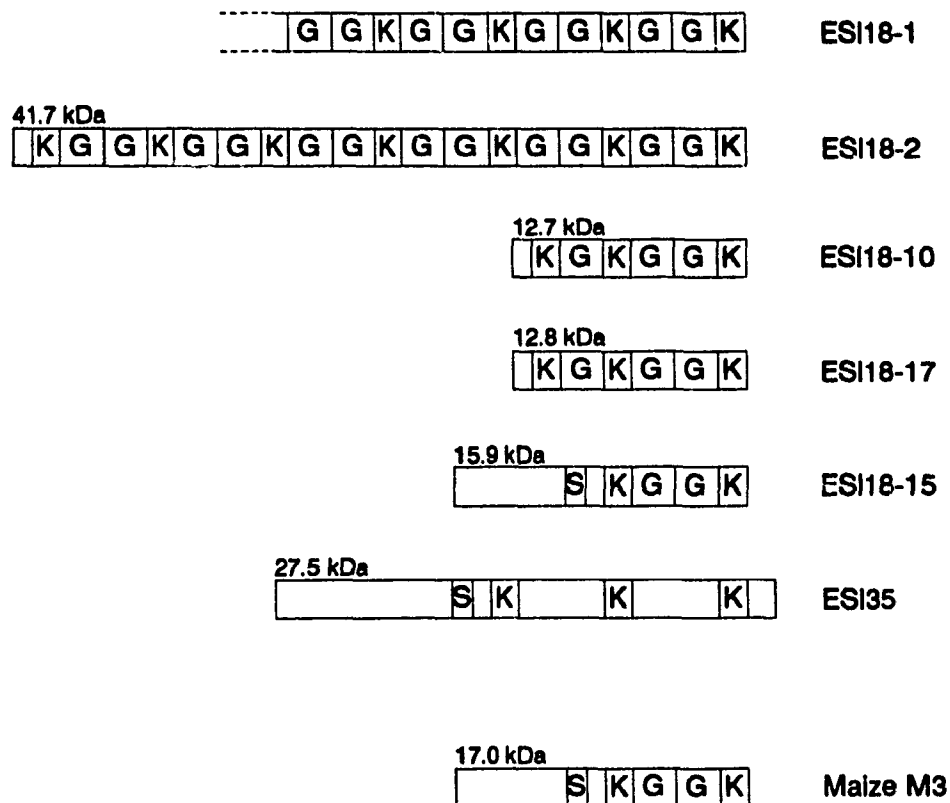


Figure 9. Diagram of the primary structures of the *L. elongatum* dehydrin-like proteins. S, serine cluster; K, lysine-rich repeat; G, glycine-rich repeat. The molecular weights of the proteins are indicated. The ESI18-1 sequence is not complete. The maize dehydrin protein M3 is shown here as a reference sequence (Close *et al.*, 1989).

sequence of ESI18-1 is not complete; only the carboxyl portion of this protein deduced from a 1.0-kb cDNA clone is available. It contains four lysine-rich repeats and four glycine-rich domains. (The size of the transcript of ESI18-1 determined by Northern blot analysis is approximately 2.6 kb.)

Notably, ESI18-10 and -17 are very closely related; 90.3% of their amino acid sequences are identical. However, the degrees of similarity in amino acid sequence among the other dehydrin-like proteins can not be readily compared by linear alignments and single values of 'per cent similarity' because they are not strictly collinear due to the various degrees of internal duplications of the lysine- and glycine-rich motifs. In ESI18-2, -10 and -17, over 90% of the sequences are composed of such repeats. Only ESI18-15 and ESI35 have larger parts of non-repeating sequences, which occur between the amino termini and the serine clusters, and which are 20.0% identical to each other. Therefore, it is more meaningful to compare the conserved subdomains, namely, the lysine- and glycine-rich repeats, of these proteins. Such comparisons will be discussed in the following paragraphs.

A search of the protein databases revealed that the barley dehydration-inducible DHN3 protein (Close et al., 1989) is the most similar to ESI18-15 among all the identified dehydrin-like proteins. Between the two proteins, 89.1% of the amino

acid sequences are identical. The majority of the more than 30 dehydrin-like proteins so far identified have structures similar to that of the barley DHN3, with two lysine-rich repeats, a glycine-rich motif repeated one to five times between the lysine domain, and a serine cluster (see reviews by Dure [1993b] and Close et al [1993b]). The structures of ESI18-2, -10 and -17 are similar to the cold-regulated wheat WCS120 and COR39 (Houde et al., 1992; Guo et al., 1992) and the barley DHN5 (Close et al., 1993b) in that they do not contain serine clusters but do contain more than two lysine-rich repeats intervened by glycine-rich repeats. Such a structure is found among several other identified plant dehydrins (Dure, 1993b; Close et al., 1993b). ESI35 is not closely related to any other reported dehydrin-like proteins because the sequence similarity between ESI35 and other dehydrins are limited to the serine and lysine domains (Gulick and An, 1993).

## 2.2. Comparison of subdomains of the L. elongatum dehydrin-like proteins

Dehydrin-like proteins contain typically the lysine-rich motif and, in some cases, glycine-rich sequences and/or a serine cluster. The occurrence and conservation of these motifs in this family must be of structural or functional importance. The six isolated L. elongatum dehydrins demonstrate extensive

amino acid sequence variations, therefore, they are excellent materials for the comparative studies to find clues to the gene function as well as the evolution of this multigene family.

### 2.2.1. Lysine-rich repeats

Figure 10 shows the alignment of the amino acid sequences of all the lysine-rich repeats from the isolated L. elongatum dehydrins. The maize dehydrin M3 is used as a reference sequence (Close et al., 1989). A consensus sequence for this motif can be reached according to the predominant amino acid residue at each particular position:

EKKGV(I)ME(D)N(K)IKE(D)KLPGG(Q)H.

Nine of the 22 lysine-rich repeats found in the L. elongatum dehydrins match the consensus sequence, and among the others (excluding the truncated second repeat from the amino terminus of ESI18-2), the match ranges as low as 50% as in the case of the repeat most distal to the carboxyl terminus in ESI35.

Since several amino acid sequences of this domain are identical, comparison with the nucleotide sequences includes a greater range of polymorphism. Such comparison was carried out with the PC/GENE computer program Multiple Sequence Alignment Version 1.20 which was based on the method of Higgins and Sharp (1988). A dendrogram based on the

K1A	GKKSLMDKIKEKLPGQH
K1B	EKKGVMENIKEKLPGGH
K1C	EKKGVMENIKEKLPGGH
K1D	EKKGVMENIKEKLPGGH
K2A	EKKSLMDKIKDKLPGQH
K2B	EKKGVMENIKEKLPGGH
K2C	EKKGVMENIKDKLPGGH
K2D	EKKGVMENIKEKLPGGH
K2E	EKKGVMENIKDKLPGGH
K2F	MENIKDKLPGGH
K2G	EKKGIMEKIKEKLPGGH
K10A	ENKSIMDKIKDKLPGQH
K10B	EKKGVVENIKEKLPGGH
K10C	ENKGVVESITEKLPGGH
K17A	ENKSIMDKIKDKLPGQH
K17B	EKKGVVENIKEKLPGGH
K17C	EKKGVVESITEKLPVGH
K15A	EKKGIMDKIKEKLPGQH
K15B	RKKGIKDKIKEKLPGGH
K35A	EKKGLLGKIMDKLPGYH
K35B	EKKGFMEKIKEKLPGGH
K35C	IKRKKKKGLKEKLPG-H
KM3A	EKKGIMDKIKEKLPGQH
KM3B	RKKGIKIKEKLPGGH
Consensus	EKKGVMENIKEKLPGGH
	I DK D Q

Figure 10. Alignment of the amino acid sequences of the lysine-rich repeats of the *L. elongatum* dehydrin-like proteins. The repeats are named as K (for lysine-rich)-a number (for a particular protein)-a letter (for the position of the repeat). 1, ESI18-1; 2, ESI18-2; 10, ESI18-10; 15, ESI18-15; 17, ESI18-17; 35, ESI35; M3, the maize M3 protein as a reference (Close et al., 1989). A, B, C, D, E, F and G are in the order from the position closest to the carboxyl terminus of the protein to the position most distal to the carboxyl terminus.



comparisons of the nucleotide sequences coding for these lysine-rich repeats is shown in Figure 11. According to the nucleotide sequences, these lysine-rich repeats can be categorized into three groups, except for the repeat most distal to the carboxyl terminus in ESI35 and the sixth repeat from the carboxyl terminus in ESI18-2. The former shows extensive differences in both amino acid and nucleotide sequences from the others, and latter is truncated at its amino terminus. One group includes repeats at the carboxyl termini of dehydrins including the one from the maize M3; these repeats are more similar to each other than to other repeats within the same genes. Notably, the repeats at this position of ESI18-10 and -17 are also identical in nucleotide sequences (not shown). The second group includes the lysine-rich repeats other than the carboxyl terminal repeats of ESI18-15 and the maize M3 dehydrin; both proteins contain a serine cluster. The third group encompasses repeats other than the carboxyl terminal repeats in ESI35, ESI18-1, -2, -10 and -17; the latter three proteins do not have serine clusters, and it has not yet been determined whether ESI18-1 has a serine cluster. It should be noted that the sequences of the lysine-rich repeats of ESI35 are very different from the others.

The high conservation of the lysine-rich repeats at the carboxyl termini among all the identified L. elongatum

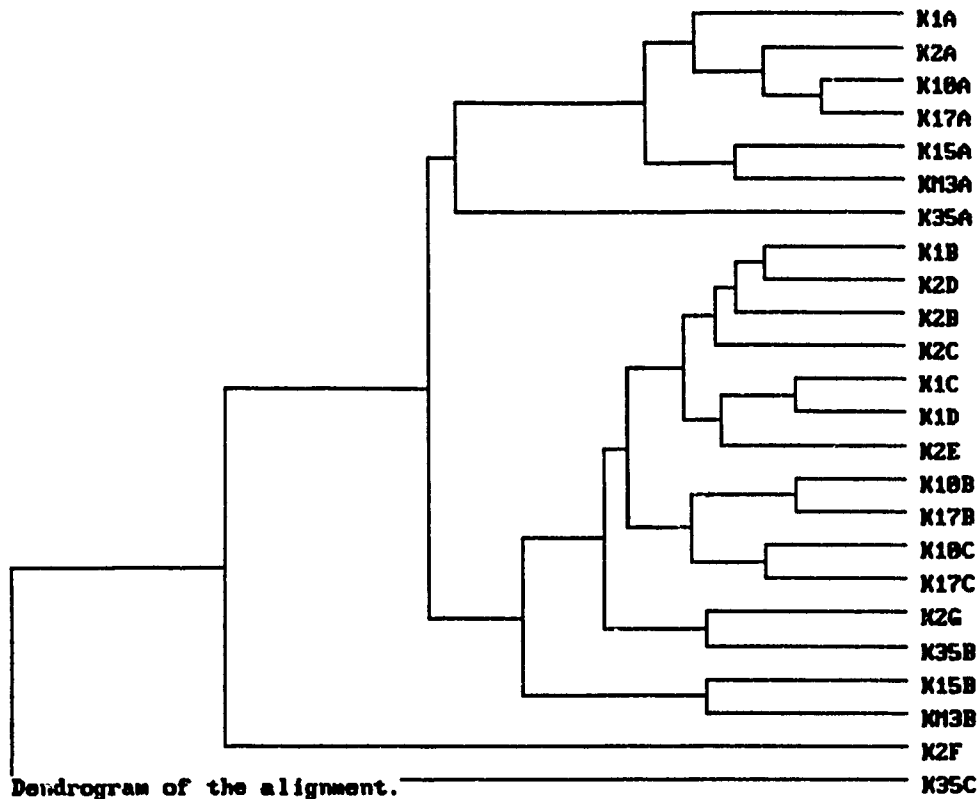


Figure 11. Dendrogram of the alignment of the nucleotide sequences for the lysine-rich repeats of the *L. elongatum* dehydrin-like proteins. The nucleotide sequences are named in the same way as described in Figure 10 for the amino acid sequences. The comparisons were carried out with the PC/GENE computer program Multiple Sequence Alignment Version 1.20, which was based on the method of Higgins and Sharp (1988). The maize dehydrin gene M3 sequence was used as a reference sequence (Close *et al.*, 1989).

dehydrin-like proteins and the heterologous maize dehydrin M3 indicates that a repeat at this position is of functional importance, directly or through the structure it assumes, if all dehydrin-like proteins have common functions. It also reveals that the lysine-rich sequence at the carboxyl terminus is probably the ancestor of the other lysine-rich repeats. It should be noted that such a high degree of amino acid sequence conservation is rare even among putative catalytic sites in similar enzymes from different species.

Proteins containing more than two lysine-rich repeats, that is, ESI18-2, -10 and -17 do not have the serine clusters; whereas in those with two repeats, that is, ESI18-15 and the maize M3, the serine clusters exist. On the other hand, the occurrence of multiple lysine-rich sequences seems to be a result of duplication of the second repeat from the carboxyl terminus since they are more similar to each other than to the one at the carboxyl terminus. Therefore, it is reasonable to speculate that dehydrins with serine clusters and dehydrins with more than two lysine-rich repeats descend through separate evolutionary processes from an ancestral gene with two lysine-rich repeats. ESI35, containing three lysine-rich repeats as well as a serine cluster, is so different in sequence from all other proteins that the evolution of this protein must be unique relative to the other two groups.

### 2.2.2. Glycine-rich sequences

The amino acid sequences between the lysine-rich repeats are rich in glycine and, to a lesser extent, threonine, in all the identified L. elongatum dehydrin-like proteins except ESI35. These interval sequences are shown in Figure 12 as an alignment. The sequence from the maize dehydrin M3 is again included as a reference sequence. The glycine-rich sequences from different proteins are similar to each other, although not as conserved as those among the lysine-rich repeats (Figure 10). The most similar of all these sequences are the ones closest to the carboxyl termini of ESI18-1 and -2, with 90.0% of the amino acid residues identical. It is interesting that most of these sequences contain two copies of a glycine-rich motif with a consensus sequence of TGGTYGQQGHTG, with different variations. In ESI18-10 and -17, the interval sequences of the second and the third lysine-rich repeats from the carboxyl termini have only one copy of the glycine-rich repeat.

### 2.2.3. Structure and function

The possible roles that the dehydrin-like (Group 2 LEA) proteins play in stress response and resistance have been discussed in Section 2 of Part I. Introduction. The suggested function for the dehydrins, for example, as osmoprotectants,

```

G1A  NDHHL----TTDTYGOHGHTGVTGTETHG----TGCTYG--QQAHTGTTGTGTHGTDGTC
G1B  ADHHQ----TAGSYGOQGDV---DTGTHDTPA-TDNTNG--QHGHTGVTGTGTHGT---G
G1C  VDHQQ----TAGSYGOQGHV---DTGTHGTPA-TGCTYG--QHGHTGVTGTGMHGT---G
G1D  GDHQQ----SGGTYGHQEDTEITGMGMHSTPA-TGGAYR--QHEHTTVTGTGMHGT---G
G2A  SDHQQ----TTDTYGOHGAGVTGTETHGTTA-TGCTYG--QOGHTGTTGTGTHGTDGTC
G2B  GDHQQ----TGCTYGOQAHI---STGTHDTPA-TGDTYG--QHGHTGVTSTGTHGT---G
G2C  ADHQQ----TTGTYGOQGHV---GTGTHGTPA-TGGAYG--QHEHTGVTGTGTHGT---G
G2D  GDHQQ----TGCTYAOQGH---GTATHGTPA-GGCTYDYEQOGHTGMTGTGTHGT---G
G2E  GDHQQ----TAGAYGOQGH---GTATHGTPA-TGCTYG--EQGHTGVTGTGTHDAD-TG
G2F  GDHKQ----TAGTHG---HA---GTATHGAPA-TGGAYG--QEGHTGTTGTGLHGADA--
G10A GDHQH----TTGMTG-----SETHATTATTDGNYG--KSGHTG-----IDGTC
G10B GDHQQ-----TTGDTYG--QQAAGVVTGTGT-----GTC
G17A GGHQH----TTGMTG-----SETHATTATSDGNYG--QSGHTG-----TDGTC
G17B GDHQQ-----TTGCTYG--QOGHAGVVTSTGT-----G
G15A GDQQQ----TAGTYGOQGHTGMAG-----TGGNYG--QPGHTGMAG-----TDGTC
GM3A KDDQHATATTGGAYGOQGH-----TGSAYG--QOGHTGGAYAT--GTEGTC

```

```

G35A KKPEDVAAPVTHAAPVHAPAAEEVSSPDAK
G35B KDNEGEHVTGLPAPAAPASVQTHHDTDVVVEKIDGDVKTEATPPVPEE

```

Consensus of the glycine-rich motif: TGCTYGOQGHTG

Figure 12. Alignment of the amino acid sequences between the lysine-rich repeats of the *L. elongatum* dehydrin-like proteins. The sequences are named as G (for glycine-rich in most cases)-a number (for a protein)-a letter (for the position of the sequence). 1, ESI18-1; 2, ESI18-2; 10, ESI18-10; 15, ESI18-15; 17, ESI18-17; 35, ESI35; M3, the maize M3 dehydrin as a reference (Close *et al.*, 1989). A, B, C, D, E, F and G are in the order from the position closest to the carboxyl terminus of the protein to the position most distal to the carboxyl terminus. The glycine-rich repeats are underlined.

relies on the structures they might assume. In this study, a set of six dehydrin-like proteins in a single species, L. elongatum, were available and their primary structures were compared. Although these proteins are different in the organization of their primary structures, the genes coding for these proteins are all responsive to salt treatment. So far there has not been any indication of a correlation between a particular subdomain, that is, serine cluster, lysine-rich repeat, or glycine-rich sequence, and a defined function related to stress response or resistance in L. elongatum or in other species. However, a segment at the carboxyl terminus with two lysine-rich repeats usually separated by a glycine-rich sequence is present in almost all dehydrin-like proteins. Such structural conservation must be essential for the dehydrin function. The existence of the serine cluster may affect the targeting of the protein. The maize dehydrin-like protein RAB17 was found to be located in nuclei when the serine residues were phosphorylated, whereas those unphosphorylated were found in the cytoplasm (Goday et al., 1994).

### 3. Expression of the ESI genes in L. elongatum

All the L. elongatum ESI genes showed early response in roots to salt treatment (Gulick and Dvořák, 1992). These genes are also responsive to ABA, KCl and mannitol (Galvez et al.,

1993). They might represent the general response to stress which causes cellular water deficiency. Although ABA can induce gene expression, the signal transduction pathway(s) leading to the expression of these genes are not known. The coincident expression of this set of genes may indicate that they share some common mechanisms regulating gene expression. The elevated steady state mRNA levels upon salt stress indicate either activated transcription, or reduced degradation of existing mRNA, or both. Regulation at the translational and post-translational levels is also possible. Nuclear run-on studies and immunological assays may help to clarify the level of regulation which controls gene expression.

The promoter region of a gene contains information concerning the regulation of its expression, that is, sequences related to responsiveness to certain stimuli or activation factors. The 106-bp 5'-flanking fragment of the ESI3 gene is the only 5'-flanking region isolated so far from the ESI genes (Figure 1). Although, it contains an ABRE-like motif, GACACGTACAC, at position -101, the information it offers is not sufficient, and the use of this fragment in expression studies is limited. Therefore, cloning and sequencing of the promoter regions of the ESI genes needs to be done for more extensive analysis of the ESI gene promoters.

#### 4. Adaptive roles of the ESI genes in salt tolerance

The salt tolerance of a plant species is likely determined by multiple genes. It may prove difficult to demonstrate that a single gene is able to confer salt tolerance to a plant, for instance, in transgenic experiments. However, it is possible to establish correlative data concerning a gene's relation to tolerance by comparing the gene expression or the activity of the gene product in closely related genomes which differ in their degree of tolerance. A positive correlation would strongly support, although not prove, the adaptive role of a gene in salt tolerance. In this study, a positive correlation was found between the salt-induced steady state levels of mRNA for the dehydrin-like genes in different wild wheatgrass species and their ability to survive high salinity (Figure 7; Table 1). It will be more interesting to extend such assays to other ESI genes. Moreover, the protein levels of the gene products in salt-treated plants should also be investigated since the proteins are presumably the active components of the salt-stress-regulated genes.



## REFERENCES

- Almoguera C, Jordano J (1992). Developmental and environmental concurrent expression of sunflower dry-seed-stored low-molecular-weight heat-shock protein and *Lea* mRNAs. *Plant Mol Biol* 19, 781-792.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990). Basic local alignment search tool. *J Mol Biol* 215, 403-410.
- An H (1994). Characterization of two members of a multigene family in the wheatgrass, *Lophopyrum elongatum*. A thesis for the Degree of M. Sc. at Concordia University, Montreal, Quebec.
- Anderberg RJ, Walker-Simmons MK (1992). Isolation of a wheat cDNA clone for an abscisic acid-inducible transcript with homology to protein kinases. *Proc Natl Acad Sci USA* 89, 10183-10187.
- Baker J, Steele C, Dure L III (1988). Sequence and characterization of 6 *Lea* proteins and their genes from cotton. *Plant Mol Biol* 11, 277-291.
- Bray EA (1993). Molecular response to water deficit. *Plant Physiol* 103, 1035-1040.

Braun Y, Hassidim M, Lerner HR, Reinhold L (1986). Studies on H<sup>+</sup>-ATPases in plants of varying resistance to salinity. I. Salinity during growth modulates the proton pump in the holophyte Atriplex nummularia. Plant Physiol 81, 1050-1056.

Boyer JS (1982). Plant productivity and environment. Science 218, 443-448.

Claes B, Dekeyser R, Villarroel R, Van den Bulcke M, Bauw G, Van Montagu M, Caplan A (1990). Characterization of a rice gene showing organ-specific expression in response to salt stress and drought. Plant Cell 2, 19-27.

Close TJ, Kortt AA, Chandler PM (1989). A cDNA-based comparison of dehydration-induced proteins (dehydrins) in barley and corn. Plant Mol Biol 13, 95-108.

Close TJ, Fenton RD, Moonan F (1993a). A view of plant dehydrins using antibodies specific to the carboxyl terminal peptide. Plant Mol Biol 23, 279-286.

Close TJ, Fenton RD, Yang A, Asghar R, DeMason DA, Crone DE, Meyer NC, Moonan F (1993b). Dehydrin: The protein. In Close TJ, Bray EA, eds, Plant Responses to Cellular Dehydration During Environmental Stress. Department of Botany and plant Sciences, University of California, Riverside, California,

U.S.A., pp 104-116.

Close TJ, Lammers PJ (1993). An osmotic stress protein of cyanobacteria is immunologically related to plant dehydrins. *Plant Physiol* 101, 773-779.

Cohen A, Plant AL, Moses MS, Bray EA (1991). Organ-specific and environmentally regulated expression of two abscisic acid-induced genes of tomato - Nucleotide sequence and analysis of the corresponding cDNAs. *Plant Physiol* 97, 1367-1374.

Cushman JC, Meyer G, Michalowski CB, Schmitt JM, Bohnert HJ (1989). Salt stress leads to differential expression of two isogenes of phosphoenolpyruvate carboxylase during Crassulacean acid metabolism induction in the common ice plant. *Plant Cell* 1, 715-725.

Dean-Johnson M, Henry SA (1989). Biosynthesis of inositol in yeast - Primary structure of myo-inositol-1-phosphate synthase (EC 5.5.1.4) and functional analysis of its structural gene, the INO1 locus. *J Biol Chem* 264, 1274-1283.

Delauney AJ, Verma DPS (1990). A soybean gene encoding  $\Delta^1$ -pyrroline-5-carboxylate reductase was isolated by functional complementation in Escherichia coli and is found to be osmoregulated. *Mol Gen Genet* 221, 299-305.

Delauney AL, Verma DPS (1993). Proline biosynthesis and osmoregulation in plants. *Plant J* 4, 215-223.

DeRocher EJ, Bohnert HJ (1993). Developmental stress employ different mechanisms in the expression of a plant gene family. *Plant Cell* 5, 1611-1625.

Downing WL, Mauxion F, Fauvarque MO, Reviron MP, de Vienne D, Vartanian N, Giraudat J (1992). A Brassica napus transcript encoding a protein related to the Kunitz protease inhibitor family accumulates upon water stress in leaves, not in seeds. *Plant J* 2, 685-693.

Dure L III, Greenway SC, Galau GA (1981). Developmental biochemistry of cottonseed embryogenesis and germination: Changing mRNA populations as shown by in vitro and in vivo protein synthesis. *Biochem* 20, 4162-4168.

Dure L III, Crouch M, Harada J, Ho THD, Mundy J, Quatrano R, Thomas T, Sung ZR (1989). Common amino acid sequence domains among the LEA proteins of high plants. *Plant Mol Biol* 12, 475-486.

Dure L III (1993a). A repeating 11-mer amino acid motif and plant desiccation. *Plant J* 3, 363-369.

Dure L III (1993b). Structural motifs in *Lea* proteins. In Close TJ, Bray EA, eds, Plant Responses to Cellular Dehydration During Environmental Stress. Department of Botany and plant Sciences, University of California, Riverside, California, U.S.A., pp 91-103.

Espelund M, Sæbøe-Larssen S, Hughes DW, Galau GA, Larsen F, Jacobsen KS (1992). Late embryogenesis-abundant genes encoding proteins with different numbers of hydrophilic repeats are regulated differentially by abscisic acid and osmotic stress. *Plant J* 2, 241-252.

Flowers TJ, Troke PF, Yeo AR (1977). The mechanism of salt tolerance in halophytes. *Annu Rev Plant Physiol* 28, 89-121.

Fourney RM, Miyakoshi J, Day RS III, Paterson MC (1988). Northern blotting: Efficient RNA staining and transfer. *Focus* 10, 5-7.

Gabarino J, Dupont FM (1989). Rapid induction of  $\text{Na}^+/\text{H}^+$  exchange activity in barley root tonoplast. *Plant Physiol* 89, 1-4.

Galau GA, Hughes DW, Dure L III (1986). Abscisic acid induction of cloned cotton late embryogenesis abundant (*Lea*) mRNAs. *Plant Mol Biol* 7, 155-170.

Galau GA, Bijaisoradat n, Hughes DW (1987). Accumulation kinetics of cotton late embryogenesis-abundant (Lea) mRNAs and storage protein mRNAs: coordinate regulation during embryogenesis and the role of abscisic acid. Dev Biol 123, 198-212.

Galau GA, Wang HYC, Hughes DW (1992). Cotton Lea4 (D19) and LeaA2 (D132) group 1 Lea genes encoding water stress-related proteins containing a 20-amino acid motif. Plant Physiol 99, 783-788.

Galvez AF, Gulick PJ, Dvořák J (1993). Characterization of the early stage of genetic salt-stress response in salt-tolerant Lophopyrum elongatum, salt-sensitive wheat, and their amphiploid. Plant Physiol 103, 257-265.

Goddard NJ, Dunn MA, Zhang L, White AJ, Jack PL, Hughes MA (1993). Molecular analysis and spatial expression pattern of a low-temperature-specific barley gene, blt101. Plant Mol Biol 23, 871-879.

Goday A, Jensen AB, Cullánez-Macià FA, Mar Albà M, Figueras M, Serratos J, Torrent M, Pagès (1994). The maize abscisic acid-responsive protein Rab17 is located in the nucleus and interacts with nuclear localization signals. Plant Cell 6, 351-360.

Greenway H, Munns R (1980). Mechanisms of salt tolerance in nonhalophytes. *Annu Rev plant Physiol* 31, 149-190.

Guerrero FD, Jones JT, Mullet JE (1990). Turgor-responsive gene transcription and RNA levels increase rapidly when pea shoots are wilted. Sequence and expression of three inducible genes. *Plant Mol Biol* 15, 11-26.

Guiltinan MJ, Marcotte WR Jr, Quatrano RS (1990). A plant leucine zipper protein that recognizes an abscisic acid response element. *Science* 250, 267-271.

Gulick P, Dvořák J (1987). Gene induction and repression by salt treatment in roots of the salinity-sensitive Chinese Spring wheat and the salinity-tolerant Chinese Spring x Elytrigia elongata amphiploid. *Proc Natl Acad Sci USA* 84, 99-103.

Gulick PJ, Dvořák J (1990). Selective enrichment of cDNAs from salt-stress-induced genes in the wheatgrass, Lophopyrum elongatum, by the formamide-phenol emulsion reassociation technique. *Gene* 95, 173-177.

Gulick PJ, Dvořák J (1992). Coordinate gene response to salt stress in Lophopyrum elongatum. *Plant Physiol* 100, 1384-1388.

Gulick PJ, An H (1993). Stress-induced gene ESI35 from Lophopyrum elongatum. Plant Physiol 103, 1031-1032.

Gulick PJ, Shen W, An H (1994). ESI3, a stress-induced gene from Lophopyrum elongatum. Plant Physiol 104, 799-800.

Guo W, Ward RW, Thomashow MF (1992). Characterization of a cold-regulated wheat gene related to Arabidopsis cor47. Plant Physiol 100, 915-922.

Gómez J, Sánchez-Matínez D, Stiefel V, Rigau J, Puigdomènech P, Pagès M (1988). A gene induced by plant hormone abscisic acid in response to water stress encodes a glycine-rich protein. Nature 334, 262-264.

Heidecker G, Messing J (1986). Structural analysis of plant genes. Ann Rev Plant Physiol 37, 439-466.

Higgins DG, Sharp PM (1988). CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. Gene 73, 237-244.

Houde M, Danyluk J, Laliberté JF, Rassart E, Dhindsa RS, Sarhan F (1992). Cloning, characterization, and expression of a cDNA encoding a 50-kilodalton protein specifically induced by cold acclimation in wheat. Plant Physiol 99, 1381-1387.



Hu CA, Delauney AJ, Verma DPS (1992). A bifunctional enzyme ( $\Delta^1$ -pyrroline-5-carboxylate synthetase) catalyses the first two steps in proline biosynthesis in plants. Proc Natl Acad Sci USA 89, 9354-9358.

Hurkman WJ, Fornari CS, Tanaka CK (1989). A comparison of the effect of salt on polypeptides and translatable mRNAs in roots of a salt-tolerant and a salt-sensitive cultivar of barley. Plant Physiol 90, 1444-1456.

Höfner R, Vazquez-Moreno L, Winter K, Bohnert HJ, Schmitt JM (1987). Induction of Crassulacean acid metabolism in Mesembryanthemum crystallinum by high salinity: Mass increase and de novo synthesis of PEP-carboxylase. Plant Physiol 83, 915-919.

Innis MA, Gelfand DH, Sninsky JJ, White TJ (1990). PCR Protocols, a Guide to Methods and Applications. Academic Press, San Diego, California, U.S.A.

King GJ, Turner VA, Hussey CE Jr, Wurtele ES, Lee SM (1988). Isolation and characterization of a tomato cDNA clone which codes for a salt-induced protein. Plant Mol Biol 10, 401-412.

Klein P, Kanehisa M, DeLisi C (1985). The detection and classification of membrane-spanning proteins. Biochem Biophys

Acta 815, 468-476.

Kiyosue T, Yamaguchi-Shinozaki K, Shinozaki K, Higashi K, Satoh S, Kamada H, Harada H (1992). Isolation and characterization of a cDNA that encodes ECP31, an embryogenic-cell protein from carrot. *Plant Mol Biol* 19, 239-249.

Knowicz AK, Nelson DE, Singh NK, Hasegawa PM, Bressan RA (1992). Regulation of the osmotin promoter. *Plant Cell* 4, 513-524.

Kyte J, Doolittle RF (1982). A simple method for displaying the hydropathic character of a protein. *J Mol Biol* 157, 105-132.

LaRosa PC, Rhodes D, Rhodes JC, Bressan RA, Csonka LN (1991). Elevated accumulation of proline in NaCl-adapted tobacco cells is not due to altered  $\Delta^1$ -pyrroline-5-carboxylate reductase. *Plant Physiol* 96, 245-250.

LaRosa PC, Chen Z, Nelson DE, Singh NK, Hasegawa PM, Bressan RA (1992). Osmotin gene expression is posttranscriptionally regulated. *Plant Physiol* 100, 409-415.

Litts JC, Colwell GW, Chakerian RL, Quatrano RS (1987). The nucleotide sequence of a cDNA clone encoding the Wheat Em

protein. Nucl Acids Res 15, 3607-3618.

Litts JC, Colwell W, Chakerian RL, Quatrano RS (1991). Sequence analysis of a functional member of the Em gene family from wheat. J DNA Sequencing and Mapping 1, 263-274.

Litts JC, Erdman MB, Huang N, Karrer EE, Noueirry A, Quatrano RS, Rodriguez RL (1992). Nucleotide sequence of the rice (Oryza sativa) Em protein gene (Empl). Plant Mol Biol 19, 335-337.

Ludevid MD, Freire MA, Gómez J, Burd CG, Alberricio F, Giralt E, Dreyfuss G, Pagès M (1992). RNA binding characteristics of a 16 kDa glycine-rich protein from maize. Plant J 2, 999-1003.

Luo M, Lin L, Hill RD, Mohapatra SS (1991). Primary structure of an environmental stress and abscisic acid-inducible alfalfa protein. Plant Mol Biol 17, 1267-1269.

Liang V, Palva ET (1992). The expression of a rab-related gene, rab18, is induced by abscisic acid during the cold acclimation process of Arabidopsis thaliana (L.) Heynh. Plant Mol. Biol. 20, 951-962.

McGuire PE, Dvořák J (1981). High salt-tolerance potential in

wheatgrass. Crop Sci 21, 702-705.

Maurel C, Reizer J, Schroeder JI, Chrispeels MJ (1993). The vacuolar membrane protein  $\gamma$ -TIP creates water specific channels in Xenopus oocytes. EMBO J 12, 2241-2247.

Mundy J, Chua NH (1988). Abscisic acid and water-stress induce the expression of a novel rice gene. EMBO J 8, 2279-2286.

Mundy J, Yamaguchi-Shinozaki K, Chua NH (1990). Nuclear proteins bind conserved elements in the abscisic acid-responsive promoter of a rice gene. Proc Natl Acad Sci USA 87, 1406-1410.

Niu X, Narasimhan ML, Salzman RA, Bressan RA, Hasegawa PM (1993). NaCl regulation of plasma membrane  $H^+$ -ATPase gene expression in a glycophyte and a halophyte. Plant Physiol 103, 713-718.

Ochman H, Medhora MM, Garza D, Hartl DL (1990). Amplification of flanking sequences by inverse PCR. In MA Innis, DH Gelfand, JJ Sninsky, TJ White, eds, PCR Protocols, a Guide to Methods and Applications. Academic Press, San Diego, California, U.S.A., pp 28-38.

Oeda K, Salinas J, Chua NH (1991). A tobacco bzip transcription activator (TAF-1) binds to a G-box-like motif conserved in plants. EMBO J 10, 1793-1802.

Omielan JA, Epstein E, Dvořák J (1991). Salt tolerance and ionic relations of wheat as affected by individual chromosomes of salt-tolerant Lophopyrum elongatum. Genome 34, 961-974.

Osterm JA, Olson SW, Schmitt JM, Bohnert HJ (1987). Salt stress increase the level of translatable mRNA for phosphoenolpyruvate carboxylase in Mesembryanthemum crystallinum. Plant Physiol 84, 1270-1275.

Parry AD (1993). Abscisic acid metabolism. Methods Plant Biochem 9, 381-402.

Piatkowski D, Schneider K, Salamini F, Bartels D (1990). Characterization of five abscisic acid-responsive cDNA clones isolated from the desiccation-tolerant plant Craterostigma plantagineum and their relationship to other water-stress genes. Plant Physiol 94, 1682-1688.

Pla M, Gómez J, Goday A, Pagès M (1991). Regulation of the abscisic acid-responsive gene rab28 in maize viviparous mutants. Mol Gen Genet 230, 394-400.

Plana M, Itarte E, Eritja R, Goday A, Pagès M, Martínez MC (1991). Phosphorylation of maize RAB-17 protein by casein kinase 2. *J Biol Chem* 266, 22510-22514.

Raghothama KG, Liu D, Nelson DE, Hasegawa PM, Bressan RA (1993). Analysis of an osmotically regulated pathogenesis-related osmotin gene promoter. *Plant Mol Biol* 23, 1117-1128.

Reuveni M, Bennet AB, Bressan RA, Hasegawa PM (1990). Enhanced activity of the tonoplast ATPase in salt adapted cells of Nicotiana tabacum. *Plant Physiol* 94, 524-530.

Roberts JK, DeSimone NA, Lingle WL, Dure L III (1993). Cellular concentrations and uniformity of cell-type accumulation of two Lea proteins in cotton embryos. *Plant Cell* 5, 769-780.

Sambrook J, Fritsch EF, Maniatis T (1989). *Molecular Cloning: A Laboratory Manual*, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, U.S.A.

Singh NK, Handa AK, Hasegawa PM, Bressan RA (1985). Proteins associated with adaptation of cultured tobacco cells to NaCl. *Plant Physiol* 79, 126-137.

Singh NK, Bracker CA, Hasegawa PM, Handa AK, Buckel S,

Hermodson MA, Pfankoch E, Regnier FE, Bressan RA (1987).  
Characterization of osmotin. *Plant Physiol* 85, 529-536.

Singh NK, Nelson DE, Kuhn D, Hasegawa PM, Bressan RA (1989).  
Molecular cloning of osmotin and regulation of its expression  
by ABA and adaptation to low water potential. *Plant Physiol*  
90, 1096-1101.

Schaffer MA, Fischer RL (1988). Analysis of mRNAs that  
accumulate in response to low temperature identifies a thiol  
protease gene in tomato. *Plant Physiol* 87, 431-436.

Skriver K, Mundy J (1990). Gene expression in response to  
abscisic acid and osmotic stress. *Plant Cell* 2, 503-512.

Skriver K, Olsen FL, Rogers JC, Mundy J (1991). Cis-acting  
DNA elements responsive to gibberellin and its antagonist  
abscisic acid. *Proc Natl Acad Sci USA* 88, 7266-7270.

Staal M, Maathius FJM, Elzenga JTM, Overbeek JHM, Prins HBA  
(1991).  $\text{Na}^+/\text{H}^+$  antiport activity in tonoplast vesicles from  
roots of the salt-tolerant Plantago maritima and the salt  
sensitive Plantago media. *Physiol Plant* 82, 179-184.

Sze H (1985).  $\text{H}^+$ -translocating ATPase. Advances using  
membrane vesicles. *Annu Rev Plant Physiol* 36, 175-208.

Szoke A, Miao GH, Hong Z, Verma DPS (1992). Subcellular location of  $\Delta^1$ -pyrroline-5-carboxylate reductase in root/nodule and leaf of soybean. *Plant Physiol* 99, 1642-1649.

Tarczyński MC, Jensen RG, Bohnert HJ (1992). Expression of a bacterial mt1D gene in transgenic tobacco leads to production and accumulation of mannitol. *Proc Natl Acad Sci USA* 89, 2600-2604.

Tarczyński MC, Jensen RG, Bohnert HJ (1993). Stress protection of transgenic tobacco by production of the osmolyte mannitol. *Science* 259, 508-510.

Thomas JC, McElwain EF, Bohnert HJ (1992). Convergent induction of osmotic stress-response - Abscisic acid, cytokinin, and the effect of NaCl. *Plant Physiol* 100, 416-423.

Urao T, Yamaguchi-Shinozaki K, Urao S, Shinozaki K (1993). An *Arabidopsis* myb homolog is induced by dehydration stress and its gene product binds to the conserved MYB recognition sequence. *Plant Cell* 5, 1529-1539.

Vergruggen N, Villarreal R, Montagu MV (1993). Osmoregulation of a pyrroline-5-carboxylate reductase gene in Arabidopsis thaliana. *Plant Physiol* 103, 771-181.



Vilardell J, Goday A, Freire MA, Torrent M, Martinez MC, Torné JM, Pagès M (1990). Gene sequence, developmental expression, and protein phosphorylation of RAB-17 in maize. *Plant Mol Biol* 14, 423-432.

Weretilnyk EA, Hanson AD (1990). Molecular cloning of a plant betaine-aldehyde dehydrogenase, an enzyme implicated in adaptation to salinity and drought. *Proc Natl Acad Sci USA* 87, 2745-2749.

Williamson CL, Slocum R (1992). Molecular cloning and evidence for osmoregulation of the  $\Delta^1$ -pyrroline-5-carboxylate reductase (proC) gene in pea (Pisum sativum L.). *Plant Physiol* 100, 1464-1470.

Wolfrain LA, Langis R, Tyson H, Dhindsa RS (1993). cDNA sequence, expression, and transcript stability of a cold acclimation-specific gene, cas18, of alfalfa (Medicago falcata) cells. *Plant Physiol* 101, 1275-1282.

Yamaguchi-Shinozaki K, Mundy J, Chua NH (1989). Four tightly linked rab genes are differentially expressed in rice. *Plant Mol Biol* 14, 29-39.

Yamaguchi-Shinozaki K, Koizumi M, Urao S, Shinozaki K (1992). Molecular cloning and characterization of 9 cDNAs for genes

that are responsive to desiccation in Arabidopsis thaliana: Sequence analysis of one cDNA clone that encodes a putative transmembrane channel protein. Plant Cell Physiol 33, 217-224.

Yamaguchi-Shinozaki K, Shinozaki K (1993). Characterization of the expression of a desiccation-responsive rd29 gene of Arabidopsis thaliana and analysis of its promoter in transgenic plants. Mol Gen Genet 236, 331-340.

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

Zhu B, Chen THH, Li PH (1993). Expression of an ABA-responsive osmotin-like gene during the induction of freezing tolerance in Solanum commersonii. Plant Mol Biol 21, 729-735.

Zimmerman JL (1993). Somatic embryogenesis: a model for early development in higher plants. Plant Cell 5, 1411-1423.

## APPENDICES

### A. Nomenclature of the ESI18 genes

The designations of the ESI18 genes will be changed for future publications. They will be numbered consecutively according to the molecular weights of their gene products. The changes of the names are listed below:

---

Designation in this thesis	Future designation
<u>ESI18-1</u>	<u>ESI18-1</u>
<u>ESI18-2</u>	<u>ESI18-2</u>
<u>ESI18-15</u>	<u>ESI18-3</u>
<u>ESI18-17</u>	<u>ESI18-4</u>
<u>ESI18-10</u>	<u>ESI18-5</u>

---

The names of the clones for these genes and the gene products will also be changed accordingly.