

Functional analysis of the *CRINKLY4* gene family in *Arabidopsis thaliana* and *Viviparous1*  
promoter in *Zea mays*

by

**Xueyuan Cao**

A dissertation submitted to the graduate faculty  
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Molecular, Cellular and Developmental Biology

Program of Study Committee:  
Philip W. Becraft, Major Professor  
Thomas Peterson  
Jo Anne Powell-Coffman  
David J Hannapel  
Yanhai Yin

Iowa State University

Ames, Iowa

2005

Copyright © Xueyuan Cao, 2005. All rights reserved.

UMI Number: 3200404

### INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

**UMI**<sup>®</sup>

---

UMI Microform 3200404

Copyright 2006 by ProQuest Information and Learning Company.

All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

ProQuest Information and Learning Company  
300 North Zeeb Road  
P.O. Box 1346  
Ann Arbor, MI 48106-1346

Graduate College  
Iowa State University

This is to certify that the doctoral dissertation of  
  
Xueyuan Cao  
  
has met the dissertation requirements of Iowa State University

Signature was redacted for privacy.

Major Professor

Signature was redacted for privacy.

For the Major Program

*To my parents.*

## TABLE OF CONTENTS

CHAPTER 1. GENERAL INTRODUCTION .....	1
Thesis Organization .....	6
Literature cited .....	6
CHAPTER 2. MOLECULAR ANALYSIS OF THE <i>CRINKLY4</i> GENE FAMILY IN <i>ARABIDOPSIS THALIANA</i> .....	9
Abstract .....	9
Keywords .....	10
Abbreviations .....	10
Introduction .....	11
Materials and methods .....	13
Results .....	18
Discussion .....	27
Acknowledgements .....	33
References .....	34
Figure legends .....	39
Tables .....	42
CHAPTER 3. TRANSCRIPTIONAL REGULATION OF <i>VIVIPAROUS1</i> ( <i>VPI</i> ) EXPRESSION IN MAIZE .....	52
Abstract .....	52
Introduction .....	53
Results .....	57
Discussion .....	67
Materials and methods .....	73
Acknowledgements .....	77
References .....	78
Figure legends .....	86
Table .....	90
CHAPTER 4. GENERAL CONCLUSIONS.....	101
ACKNOWLEDGEMENTS .....	104

## CHAPTER 1 GENERAL INTRODUCTION

In cereal plants, seed development starts with double fertilization, in which one sperm nucleus fertilizes the egg cell in the embryo sac resulting in a diploid zygote and the other sperm nucleus fuses with two polar nuclei of the central cell to initiate the development of the triploid endosperm (Dumas and Mogensen, 1993). During embryogenesis, the diploid zygote develops into a mature embryo with a polar axis of the plant body and embryonic organ systems. The primary triploid endosperm nucleus undergoes a series of nuclear divisions without cytokinesis, resulting in a syncytial endosperm with nuclei lying on the periphery. Cellularization follows through free cell wall formation that begins at the periphery and grows centripetally, forming tube-like alveoli. The following cell divisions occur with cytokinesis, progressing centripetally toward the central vacuole until the endosperm is fully cellularized (Becraft, 2001; Olsen, 2001, 2004). In the maturation step of seed development, the embryo and aleurone layer of the endosperm acquire desiccation tolerance and enter developmental arrest.

The cereal endosperm is an excellent system to study cell differentiation and fate specification because of its structural simplicity. It consists of three major cell types (Becraft, 2001; Olsen, 2001). The transfer cells develop in the basal endosperm over the main vascular tissue of the maternal plant. Extensive inward cell wall projections increase the area of plasma membrane to facilitate nutrient uptake from maternal tissue to the endosperm. The starchy endosperm functions mainly as storage tissue, enriched in starch and prolamin storage proteins, and represents the majority of the seed. The aleurone consists of 1~3 layers

of cells covering most part of the endosperm periphery. It derives from periclinal divisions in the outer layer of cells during early endosperm cell divisions. In maize, the aleurone cells are cubic, regularly shaped with high lipid concentration and abundant in hydrolytic enzymes and proteins. It mainly serves as a digestive tissue to remobilize the storage starch and proteins in the starchy endosperm, nurturing the germinating seedling prior to the acquisition of photosynthetic competency. In *Arabidopsis*, the endosperm degenerates and is reabsorbed into the embryo. Only one layer of aleurone-like cells is left on the periphery of the embryo in mature seeds (Olsen, 2001).

In particular maize genotypes, the aleurone is pigmented due to the accumulation of anthocyanin, a phenolic pigment. The regulation of anthocyanin synthesis in aleurone cells is well studied and serves as a good model for gene regulation. The *C1* gene encodes a myb transcription activator that regulates the expression of anthocyanin synthesis structural genes in the aleurone and scutellum. It functions as a dimer with another transcription factor R1. Recessive *c1* alleles lack anthocyanin pigmentation resulting in colorless kernels (Cone et al., 1986; Paz-Ares et al., 1987; Kagaya et al., 2005). Anthocyanin also serves as an excellent genetic marker for studying cell fate specification of aleurone cells. The seeds become colorless when peripheral cells of mutant kernels lose the aleurone cell identity, such as in *cr4*, *dek1* and other cell fate mutants (Becraft and Asuncion-Crabb, 2000; Becraft et al., 2002).

Mosaic analyses showed that the aleurone and the starchy endosperm can derive from the same parental cells and are not from separate cell lineages (Becraft and Asuncion-Crabb,

2000). Gain and loss of function mosaic analyses further suggested that the aleurone and starchy endosperm cell identities are plastic and can switch cell fates throughout endosperm development. This implies that positional cues are continuously required to specify and maintain aleurone cell fate (Becraft and Asuncion-Crabb, 2000; Becraft et al., 2002). Based on the analyses of several aleurone mutants, Becraft and Asuncion-Crabb (2000) proposed a genetic hierarchy model for controlling aleurone differentiation. In this model, genes such as CR4 and Dek1 are required for the aleurone cell fate specification and mutants of these genes fail to specify aleurone cell fate resulting in starchy endosperm cell identity in the peripheral cell layer. Following that, a set of genes function in the differentiation of the cellular characteristics of aleurone. Mutants of this group of genes lose some characteristics of aleurone. Mutants of genes functioning in late steps of aleurone differentiation only show minor defects in aleurone cell morphology and organization.

Maize *crinkly4* (*cr4*) encodes a serine/threonine receptor-like kinase (RLK), playing an important role in an array of processes both in the plant and in aleurone (Becraft et al., 1996; Becraft and Asuncion-Crabb, 2000; Jin et al., 2000). In the plant, CR4 is required for cellular development throughout the shoot, regulating cell proliferation, fate, patterning, morphogenesis, and differentiation, particularly in the epidermis. In the endosperm, CR4 is required for aleurone cell fate specification (Becraft and Asuncion-Crabb, 2000) and the *cr4* mutant results in starchy endosperm in the peripheral endosperm. Because the loss of aleurone cell identity also resulted in the loss of anthocyanin pigmentation, the mutant kernels exhibited a mosaic anthocyanin pigmented aleurone phenotype. Genetic mosaic analysis showed that CR4 functions cell autonomously (Becraft et al., 2001).

In Arabidopsis, a family of 5 RLKs are closely related to maize CR4. T-DNA insertions in *ACR4*, the ortholog of *CR4*, exhibited defects in the development of the integument and seed coat (Tanaka et al., 2002; Gifford et al., 2003). The *acr4* mutant also shows disruption of cell organization in leaf epidermis and cuticle formation (Watanabe et al., 2004). Gifford et al. (2005) demonstrated that functional ACR4 turned over very fast and the 'crinkly' domain, an extracellular domain motif, was important for the receptor internalization and function. A mutation in the crinkly domain blocked ACR4 function, while the kinase activity did not appear crucial for ACR4 function in vivo.

In the late stage of cereal seed development, starchy endosperm undergoes program cell death, while both the embryo and aleurone remain alive, acquire desiccation tolerance, and enter developmental arrest. In this process, abscisic acid (ABA) plays an important regulatory role. Kernels of mutants that disrupt ABA synthesis or ABA sensitivity fail to enter developmental arrest and show a viviparous phenotype (Neil et al., 1987). In maize, loss of function mutants of *viviparous1* (*vp1*) result in ABA insensitivity in seeds. In endosperm, *vp1* mutants fail to activate *C1* expression, one of the key regulators of anthocyanin synthesis genes, resulting in colorless kernels (McCarty et al., 1989). *Vp1* was cloned by transposon-tagging and encodes a B3 domain containing transcription factor (McCarty et al., 1989; McCarty et al., 1991). It functions as both a transcriptional activator to regulate ABA-inducible gene expression and transcriptional repressor to repress germination-specific  $\alpha$ -amylase gene expression (Hoecker et al., 1995).

In aleurone, cell fate specification genes such as *Cr4* and *Dek1* are required throughout endosperm development to specify aleurone cell identity (see above). On the other hand, *Vp1* is well studied to regulate late stages of aleurone differentiation, including anthocyanin synthesis through the regulation of *Cl* expression, and the maturation process. The steps in between cell fate specification and *Vp1* function are unknown. The knowledge of the gene functioning in between will help us to better understand the molecular mechanism of aleurone cell fate specification and differentiation. This could be pursued in two directions. One is to further characterize the signal transduction pathways of cell fate specification genes such as CR4 signaling. The other is to identify regulators of *Vp1*.

The main objective of this research is to understand more of signal transduction and cell fate determination in aleurone cells. It includes two specific goals. Firstly, we address the function of CR4 family in *Arabidopsis thaliana*. Secondly, we study the regulation of *Vp1* expression in maize seed. To simplify the study of *Vp1* regulation, we use cultured embryos as the model tissue because of the availability and easier manipulation of the embryo tissue. Future experiments will determine whether the results obtained in embryos could be generalized to aleurone cells. As illustrated by two EMSA experiments in Chapter 3, where two nuclear protein binding sites identified in embryos showed similar binding activity in aleurone cells, it appears there is at least some similarity between the two tissues.

## Dissertation Organization

The dissertation is organized in the format consisting of two journal articles preceded by a General Introduction and followed by a General Conclusion. The journal articles are formatted according to the requirements of each journal. The first article “Molecular analysis of the *CRINKLY4* gene family in *Arabidopsis thaliana*” was published in *Planta* (2005, 220:645-657). The second article “Transcriptional regulation of *viviparous1* (*Vp1*) expression in maize” will be submitted for publication in *The Plant Journal*. Xueyuan Cao was the primary investigator for the work under the supervision of Dr. Philip W. Becraft and is the first author for both articles.

## Literature Cited

- Becraft, P.W.** (2001). Cell fate specification in the cereal endosperm. *Semin Cell Dev Biol* **12**, 387-394.
- Becraft, P.W., and Asuncion-Crabb, Y.** (2000). Positional cues specify and maintain aleurone cell fate in maize endosperm development. *Development* **127**, 4039-4048.
- Becraft, P.W., Stinard, P.S., and McCarty, D.R.** (1996). CRINKLY4: A TNFR-like receptor kinase involved in maize epidermal differentiation. *Science* **273**, 1406-1409.
- Becraft, P.W., Kang, S.H., and Suh, S.G.** (2001). The maize CRINKLY4 receptor kinase controls a cell-autonomous differentiation response. *Plant Physiol* **127**, 486-496.

- Becraft, P.W., Li, K., Dey, N., and Asuncion-Crabb, Y.** (2002). The maize *dek1* gene functions in embryonic pattern formation and cell fate specification. *Development* **129**, 5217-5225.
- Cone, K.C., Burr, F.A., and Burr, B.** (1986). Molecular analysis of the maize anthocyanin regulatory locus C1. *Proc Natl Acad Sci U S A* **83**, 9631-9635.
- Dumas, C., and Mogensen, H.L.** (1993). Gametes and Fertilization: Maize as a Model System for Experimental Embryogenesis in Flowering Plants. *Plant Cell* **5**, 1337-1348.
- Gifford, M.L., Dean, S., and Ingram, G.C.** (2003). The Arabidopsis *ACR4* gene plays a role in cell layer organisation during ovule integument and sepal margin development. *Development* **130**, 4249-4258.
- Hoecker, U., Vasil, I.K., and McCarty, D.R.** (1995). Integrated control of seed maturation and germination programs by activator and repressor functions of *Viviparous-1* of maize. *Genes Dev* **9**, 2459-2469.
- Jin, P., Guo, T., and Becraft, P.W.** (2000). The maize *CR4* receptor-like kinase mediates a growth factor-like differentiation response. *Genesis* **27**, 104-116.
- Kagaya, Y., Toyoshima, R., Okuda, R., Usui, H., Yamamoto, A., and Hattori, T.** (2005). *LEAFY COTYLEDON1* controls seed storage protein genes through its regulation of *FUSCA3* and *ABSCISIC ACID INSENSITIVE3*. *Plant Cell Physiol* **46**, 399-406.
- McCarty, D.R., Carson, C.B., Stinard, P.S., and Robertson, D.S.** (1989). Molecular Analysis of *viviparous-1*: An Abscisic Acid-Insensitive Mutant of Maize. *Plant Cell* **1**, 523-532.

**McCarty, D.R., Hattori, T., Carson, C.B., Vasil, V., Lazar, M., and Vasil, I.K. (1991).**

The Viviparous-1 developmental gene of maize encodes a novel transcriptional activator. *Cell* **66**, 895-905.

**Neil, S.J., Horgan, R., and Rees, A.F. (1987).** Seed development and vivipary in *Zea mays*

*L. Planta* **171**, 358-364.

**Olsen, O.A. (2001).** ENDOSPERM DEVELOPMENT: Cellularization and Cell Fate

Specification. *Annu Rev Plant Physiol Plant Mol Biol* **52**, 233-267.

**Olsen, O.A. (2004).** Nuclear endosperm development in cereals and *Arabidopsis thaliana*.

*Plant Cell* **16 Suppl**, S214-227.

**Paz-Ares, J., Ghosal, D., Wienand, U., Peterson, P.A., and Saedler, H. (1987).** The

regulatory *c1* locus of *Zea mays* encodes a protein with homology to myb proto-oncogene products and with structural similarities to transcriptional activators. *Embo J* **6**, 3553-3558.

**Tanaka, H., Watanabe, M., Watanabe, D., Tanaka, T., Machida, C., and Machida, Y.**

(2002). ACR4, a putative receptor kinase gene of *Arabidopsis thaliana*, that is expressed in the outer cell layers of embryos and plants, is involved in proper embryogenesis. *Plant Cell Physiol* **43**, 419-428.

**Watanabe, M., Tanaka, H., Watanabe, D., Machida, C., and Machida, Y. (2004).** The

ACR4 receptor-like kinase is required for surface formation of epidermis-related tissues in *Arabidopsis thaliana*. *Plant J* **39**, 298-308.

CHAPTER 2. MOLECULAR ANALYSIS OF THE *CRINKLY4* GENE FAMILY IN  
*ARABIDOPSIS THALIANA*

A paper published in *Planta*

Xueyuan Cao<sup>1,3</sup>, Kejian Li<sup>1,3</sup>, Sang-Gon Suh<sup>1,5</sup>, Tao Guo<sup>1,4</sup> and Philip W. Beecraft<sup>1,2,3</sup>

<sup>1</sup>Department of Genetics, Development and Cell Biology, Iowa State University, Ames,  
IA 50011, USA

<sup>2</sup>Department of Agronomy, Iowa State University, Ames, IA 50011, USA

<sup>3</sup>Molecular, Cellular, and Developmental Biology Program, Iowa State University, Ames,  
IA 50011, USA

<sup>4</sup>Interdepartmental Genetics Program, Iowa State University, Ames, IA 50011, USA

<sup>5</sup>Department of Horticulture, Yeungnam University, Kyongsan, Republic of Korea

**Abstract**

The maize (*Zea mays* L.) *CRINKLY4* (*CR4*) gene encodes a serine/threonine receptor-like kinase that controls an array of developmental processes in the plant and endosperm. The *Arabidopsis thaliana* (L.) Heynh. genome encodes an ortholog of CR4, ACR4, and four CRINKLY4-RELATED (CRR) proteins: AtCRR1, AtCRR2, AtCRR3 and AtCRK1. The available genome sequence of rice (*Oryza sativa* L.) encodes a CR4 ortholog, OsCR4, and four CRR proteins: OsCRR1, OsCRR2, OsCRR3 and OsCRR4, not necessarily orthologous to the *Arabidopsis* CRRs. A phylogenetic study showed that AtCRR1 and AtCRR2 form a

clade closest to the CR4 group while all the other CRRs form a separate cluster. The five *Arabidopsis* genes are differentially expressed in various tissues. A construct formed by fusion of the *ACR4* promoter and the *GUS* reporter, *ACR4::GUS*, is expressed primarily in developing tissues of the shoot. The ACR4 cytoplasmic domain functions in vitro as a serine/threonine kinase, while the AtCRR1 and AtCRR2 kinases are not active. The ability of ACR4 to phosphorylate AtCRR2 suggests that they might function in the same signal transduction pathway. T-DNA insertions were obtained in *ACR4*, *AtCRR1*, *AtCRR2*, *AtCRR3* and *AtCRK1*. Mutations in *acr4* show a phenotype restricted to the integuments and seed coat, suggesting that *Arabidopsis* might contain a redundant function that is lacking in maize. The lack of obvious mutant phenotypes in the *atcrr* mutants indicates they are not required for the hypothetical redundant function.

**Keywords** Arabidopsis · CRINKLY4 · Evolution · Oryza · Receptor-like kinase · Signal transduction

**Abbreviations** CR4: CRINKLY4 · CRR: CRINKLY4-RELATED · GST: Glutathione S-transferase · GUS:  $\beta$ -Glucuronidase · RCC1: Regulator of chromosome condensation 1 · RLK: Receptor-like kinase · SAM: Shoot apical meristem · TNFR: Tumor necrosis factor receptor

## Introduction

In higher plants, most development occurs post-embryonically, during which plant cells must sense and respond to environmental conditions and internal cues. One means of perceiving signals is through cell-surface receptors, including receptor-like kinases (RLKs). RLKs can sense signaling molecules including polypeptides, and the steroid hormone, brassinolide. Extracellular ligand binding induces RLK activation and downstream signal transduction leading to cellular responses. The *Arabidopsis* genome contains 417 predicted genes encoding RLKs, which have a monophyletic origin within the superfamily of plant kinases (Shiu and Bleeker 2001a, 2001b). RLKs are involved in diverse developmental and defense functions including shoot apical meristem (SAM) equilibrium, pollen–pistil interaction, and hormone perception (Becraft 2002). While RLKs are known to function in diverse processes, the vast majority have unknown functions.

The maize *CRINKLY4* (*CR4*) gene is important for a complex array of processes in plant and endosperm development. Loss-of-function *cr4* mutants lead to the disruption of endosperm cell fate specification, causing the peripheral endosperm cells to develop as starchy endosperm instead of aleurone (Becraft and Asuncion-Crabb 2000; Becraft et al. 1996). *cr4* mutant plants grow short with crinkled leaves showing graft-like tissue fusions. Analysis of an allelic series and genetic mosaics showed that *CR4* functions preferentially in the epidermis, but is required for cellular development throughout the shoot, regulating cell proliferation, fate, patterning, morphogenesis, and differentiation (Becraft and Asuncion-Crabb 2000; Becraft et al. 2001; Jin et al. 2000). This suggests that *CR4* functions in a growth factor-like response.

*CR4* encodes an RLK representing a subfamily in the plant RLK family (Becraft et al. 1996; Shiu and Bleecker 2001a, 2001b). *CR4* is expressed in the growing regions of the shoot, particularly in the SAM and lateral organ primordia (Becraft 2001; Jin et al. 2000). The protein of 901 amino acids encoded by *CR4* contains a functional serine/threonine kinase in the cytoplasmic domain (Jin et al. 2000). The ectodomain contains two motifs (Fig. 1b); one is similar to the ligand-binding domain of mammalian tumor necrosis factor receptor (TNFR), which suggests that the ligand of CR4 might be a polypeptide similar to the TNF (Becraft et al. 1996). The other motif contains 7 repeats of about 39 amino acids ('crinkly' repeats), that are hypothesized to form a regulator of chromosome condensation 1 (RCC1)-like propeller structure, which is also thought to participate in protein-protein interactions (Becraft et al. 1996; McCarty and Chory 2000). The predicted cytoplasmic domain has a serine/threonine kinase catalytic domain and a 116-amino-acid carboxyl domain of unknown function.

*Arabidopsis* contains a family of five RLKs related to CR4 (Shiu and Bleecker 2001a, 2001b). The ortholog of CR4, *ACR4*, is expressed in protodermal cells of the embryo and shoot (Tanaka et al. 2002; Gifford et al. 2003). Antisense *ACR4* exhibited mild defects in seed formation and embryo morphogenesis (Tanaka et al. 2002). *ACR4* T-DNA knockouts showed defects in the development of the integuments and seed coat but no defects in embryo morphology were observed (Gifford et al. 2003). The lack of obvious mutant phenotypes in leaves suggested that *Arabidopsis* may contain a gene(s) that functions redundantly with *ACR4*, and other family members were suggested as candidates. Another member, *AtCRK1*, is orthologous to the tobacco *CRK1*, which is negatively regulated at the transcript level in cell cultures by exogenous cytokinin (Schafer and Schmulling 2002). Here,

we describe the molecular analysis of the *CR4* family of RLKs in *Arabidopsis*. The phylogeny, expression, and biochemical characterization are presented. Knockouts were obtained to attempt to assign developmental functions to these genes and to test whether they function redundantly with *ACR4*. We also report on related genes in the available rice genomic sequence.

## Materials and methods

### Plant materials

The *Arabidopsis thaliana* (L.) Heynh. genotypes used in this study were Columbia (Col) and Wassilewskija (WS). Plants were grown on Sunshine LC1 (Sun Gro Horticulture Inc.) at 21°C under continuous light (120  $\mu\text{mol photons m}^{-2} \text{s}^{-1}$ ). A T-DNA insertion in *ACR4* in a WS background was obtained by screening the Wisconsin Knockout Facility collection. T-DNA insertions in *ACR4*, *AtCRR1* and *AtCRR3* in the Col background were obtained from Syngenta. T-DNA insertions in *AtCRR1*, *AtCRR2* and *AtCRK1* in the Col background were obtained from the Salk Institute collection at the Arabidopsis Biological Resource Center (ABRC), Ohio State University.

### Sequence analysis

The maize *CR4* cDNA sequence was used in a BLAST search to identify genes in *A. thaliana* that are related to *CR4*. The ZmCR4, ACR4 and AtCRRs amino acid sequences were used to search rice (*Oryza sativa* L.) genome sequences to obtain *CR4*-related genes in rice. The

sequences were used to construct multiple sequence alignments using ClustalW of the OMIGA 1.1 software package with a BLOSUM matrix (Accelrys Inc.). The consensus and shading were performed using the program BOXSHADE 3.21 with a cut-off of 50% ([http://www.ch.embnet.org/software/BOX\\_form.html](http://www.ch.embnet.org/software/BOX_form.html)). Phylogenetic analysis was conducted using MEGA version 2.1 (Kumar et al. 2001). The programs SignalP (Nielsen et al. 1997), PSORT (Nakai and Horton 1999), iPSORT (Bannai et al. 2002), and DAS (Cserzo et al. 1997) were used to analyze the predicted structure and predicted localization of the CR4 family proteins.

#### Nucleic acid manipulation

Nucleic acid manipulation was according to (Sambrook et al. 1989). A bacterial artificial chromosome (BAC) clone, F20O21, containing the *ACR4* locus was obtained from the ABRC. A *SacI/XbaI* genomic fragment containing *ACR4* was subcloned from this BAC into pBluescript II SK (Stratagene) and named pBFScX. The coding region of *ACR4* was re-sequenced at the Iowa State University DNA facility.

Total RNA was isolated by grinding tissues in liquid nitrogen followed by extraction with TRIzol reagent (Invitrogen), treated with RQ1 RNase-Free DNase (Promega) for 30 min at 37°C and purified with the RNeasy Mini Kit (Qiagen) according to the manufacturer's instructions.

#### Reverse transcription–polymerase chain reaction (RT–PCR)

RT-PCR was performed to examine the expression of the *A. thaliana* *CR4-RELATED* genes in the following tissues: roots, 14 days after germination (DAG); SAM, 14 DAG; leaves, 21 DAG; flower buds, 30 DAG; siliques, 30 DAG. First-strand cDNA was synthesized using 3 µg total RNA. Reverse transcription was performed with the SuperScript First-Strand Synthesis System for RT-PCR (Invitrogen) according to the manufacturer's instructions. PCR was performed using Takara *ExTaq* DNA Polymerase (PanVera) with gene-specific primers (Table 1). For *ACR4*, *AtCRR1*, 2, 3 and *AtCRK1*, the amplification conditions were: 94°C for 2 min followed by 25 cycles of 94°C for 30 s, 65°C for 30 s and 72°C for 1 min. For *UBQ10*, the amplification conditions were: 94°C for 2 min followed by 25 cycles of 94°C for 30 s, 58°C for 30 s and 72°C for 1 min. PCR products were run in a 0.8% agarose gel, blotted to Duralon-UV membrane (Stratagene), hybridized with the corresponding gene-specific probes, and autoradiographed.

#### Construction of *ACR4::GUS* and plant transformation

A 1.38-kb promoter region upstream of the ATG start codon of *ACR4* was amplified from pBFScX by PCR. The PCR fragment was cloned into the *EcoRI/PstI* site of the binary vector pCAMBIA 1381Z (CAMBIA, Australia) to obtain a transcriptional fusion of the *ACR4* promoter and the β-glucuronidase (GUS) coding sequence. The construct was introduced into *Agrobacterium tumefaciens* strain C58C1. Plants were transformed by the flower-dip method modified from the vacuum-infiltration method (Bechtold and Pelletier 1998). Seeds from infiltrated plants were harvested, surface-sterilized with 50% bleach containing 0.02% Triton X-100, and sown in petri dishes on solid (0.8% agar) MS medium (Murashige and Skoog

1962) supplemented with  $0.5 \text{ g l}^{-1}$  MES (pH 5.7),  $10 \text{ g l}^{-1}$  sucrose and  $30 \mu\text{g ml}^{-1}$  hygromycin. Hygromycin-resistant seedlings (T1) were transplanted into soil, grown to maturity, and seeds harvested.

#### Histochemical GUS analysis

Histochemical localization of GUS activity in *ACR4::GUS* transgenic plants was carried out essentially as described by Jefferson et al. (1987). Plant tissues and plantlets were incubated in GUS-staining solution [ $50 \text{ mM}$  sodium phosphate (pH 7.0),  $0.5 \text{ mM}$  potassium ferricyanide,  $0.5 \text{ mM}$  potassium ferrocyanide,  $10 \text{ mM}$  EDTA,  $0.05\%$  (v/v) Triton X-100,  $0.35 \text{ mg ml}^{-1}$  5-bromo-4-chloro-3-indolyl- $\beta$ -D-glucuronide (X-Gluc)] for 24 h at  $37^\circ\text{C}$ . After staining, the samples were fixed in GUS-fixation solution (10% methanol, 15% acetic acid) and photographed with an Olympus SZH10 stereomicroscope or Olympus BX-60 compound microscope, both with a PM20 photography system.

Samples to be sectioned were stained as above, cleared in 70% ethanol, dehydrated through an ethanol series and embedded in paraffin. Sections ( $8 \mu\text{m}$ ) were dewaxed with xylene, mounted on glass slides with Permount and photographed with differential interference contrast (DIC) optics.

#### Kinase analysis

The presumptive cytoplasmic domain of ACR4 (Arg<sup>458</sup>-Phe<sup>895</sup>) was expressed in *Escherichia coli* as a 6 $\times$ His-tagged thioredoxin fusion protein in the pET-32a (+) vector (Novagen).

Expression of this protein, ACR4 K, was induced in a 100-ml culture with  $1 \text{ mM}$  isopropyl

$\beta$ -D-thiogalactopyranoside (IPTG; Sigma), and was purified from the cell lysate using a TALON purification kit (Clontech). As a negative control, the essential lysine<sup>540</sup> was substituted with an alanine (ACR4K-Mu). The mutation was introduced by PCR where an AAG-to-GCG codon change was incorporated by primer mismatch. The presumptive cytoplasmic domain of AtCRR2 (Ala<sup>483</sup>-Phe<sup>776</sup>) was expressed in *E. coli* as a glutathione *S*-transferase (GST) fusion in the pGEX-4T1 vector (Pharmacia Biotech). The GST fusion protein, AtCRR2 K, was purified from the cell lysate by binding to glutathione Sepharose 4B (Pharmacia Biotech), and then eluted with 10 mM reduced glutathione (pH 8.0). The eluates were dialyzed in kinase buffer [50 mM Tris-HCl (pH 7.5), 10 mM MgCl<sub>2</sub>, 10 mM MnCl<sub>2</sub>, 0.1% Triton X-100]. Twenty  $\mu$ l of each elution was incubated at room temperature for 30 min with 370 kBq  $\gamma$ -[<sup>32</sup>P]ATP (Perkin Elmer) in kinase buffer, and then subjected to SDS-PAGE and Coomassie staining followed by autoradiography.

The autophosphorylated ACR4 K was subjected to phosphoamino acid analysis. A 20- $\mu$ l autophosphorylation reaction was conducted as above, subjected to SDS-PAGE and blotted to Immobilon-P PVDF membrane (Millipore) using a Mini Trans-Blot Electrophoretic Transfer Cell (BioRad). The piece of membrane containing the phosphorylated ACR4 K band was excised based on autoradiography, hydrolyzed in 6 N HCl for 1 h at 110°C, lyophilized and loaded onto a thin-layer cellulose plate. The amino acids in the hydrolyte were separated by two-dimensional thin-layer electrophoresis (Boyle et al. 1991). O-Phospho-DL-serine, O-phospho-DL-threonine, and O-phospho-DL-tyrosine standards were visualized by ninhydrin staining, and radiolabeled amino acids resulting from ACR4 K autophosphorylation were detected by autoradiography.

## Results

### CR4-related genes in *Arabidopsis* and rice

To identify genes related to the *Zea mays* CR4 (*ZmCR4*) in *Arabidopsis*, extensive searching of the *Arabidopsis* genomic sequences was performed using several algorithms including BLAST (Altschul et al. 1990), PSI-BLAST (Altschul et al. 1997) and FASTA (Pearson and Lipman 1988), using various motifs of the CR4 protein. Searches with the TNFR domain, the crinkly repeats or the carboxyl terminus, produced informative hits. The *Arabidopsis* genome encodes five RLK proteins related to *ZmCR4*. *ACR4* (Tanaka et al. 2002; Gifford et al. 2003) was clearly most related to *ZmCR4* and is believed to represent the ortholog. The other four proteins were designated *AtCRR* for *Arabidopsis thaliana* CR4-RELATED. *AtCRR4* turned out to be similar to CRK1 in tobacco and was named *AtCRK1* (Schafer and Schmulling 2002). The protein identities (NCBI protein database) are shown in Fig. 1b. Gene identification numbers are: *ACR4*, At3g59420; *AtCRR1*, At3g09780; *AtCRR2*, At2g39180; *AtCRR3*, At3g55950; *AtCRK1*, At5g47850.

BLAST searches also identified a CR4 homolog in rice which was designated *OsCR4* (GenBank accession number AB057787). The extracellular sequences of *ACR4*, *ZmCR4*, *OsCR4*, *AtCRRs* and *AtCRK1* were used to search rice genomic sequences. Four putative genes encoding CR4-related proteins were identified and designated *OsCRR1*, 2, 3 and 4 (note: *OsCRR1*, 2, 3 are not necessarily orthologous to *AtCRR1*, 2, 3). The GenBank accession numbers for the rice genes are: *OsCRR1*, AL606452.2 (complement of bases 72163–74631); *OsCRR2*, AP004584.3 (bases 80615–83125); *OsCRR3*, AC129720.2 (bases

81114–83507); *OsCRR4*, AC123524.2 (complement of bases 90800–88365). The GenBank protein identification numbers for annotated genes are given in Fig. 1b. We also searched the Syngenta database (Goff et al. 2002), but did not obtain any sequences that were not represented in the public databases. It is noteworthy that none of the open reading frames of these 11 genes is interrupted by introns.

#### Features and phylogenetic relationships of the CR4-related proteins

To address the relationships among the CR4 family members, a multiple sequence alignment was conducted using ClustalW with a BLOSUM matrix (Thompson et al. 1994; Fig. 1a). *OsCR4* has the highest conservation to *ZmCR4* with 87% identity and 90% similarity, and possesses all the sequence motifs contained in *ZmCR4* (Fig. 1b for diagrammatic representation of the protein structure). *ACR4* has 60% amino acid identity and 74% similarity with *ZmCR4*, and also possesses all the sequence motifs contained in *ZmCR4*. The proteins encoded by the *CRRs* and *AtCRK1* have 29–32% overall amino acid identity with *ZmCR4* or *ACR4*. *AtCRR1* and 2 contain both extracellular motifs of CR4, the crinkly repeats and TNFR repeats, but are predicted to contain inactive kinases because of a conserved deletion that removes essential residues of kinase subdomain VIII (Gibbs and Zoller 1991; Fig. 1a,b). *AtCRR3*, *AtCRK1* and all four *OsCRRs* lack TNFR repeats in the extracellular domains, but contain the crinkly repeats and have kinase domains that are predicted to be functional (Fig. 1a,b). All the *CRRs* lack the carboxyl domain.

A phylogenetic analysis was conducted using MEGA version 2.1 (Kumar et al. 2001). The 11 proteins fell into 2 major clusters (Fig. 1c); the first contained 2 sub-clades, one

consisting of the 3 CR4 proteins and the other contained AtCRR1 and AtCRR2. No orthologs of AtCRR1 and AtCRR2 were found in the existing rice databases. The second major cluster was less related to CR4 and contained AtCRR3, AtCRK1, and OsCRR1, 2, 3 and 4. Within this cluster, AtCRR3 and OsCRR3 formed a clade and appeared to be potential orthologs. The rest were divergent between *Arabidopsis* and rice. OsCRR1 and OsCRR2 were closer to one another than to AtCRK1.

Several programs were used to predict the protein structure and subcellular location of the CR4-related proteins. SignalP was used to predict the signal-peptide cleavage sites (Nielsen et al. 1997) and iPSORT was used to predict sorting signals (Bannai et al. 2002). The subcellular localization was predicted using PSORT (Nakai and Horton 1999). Transmembrane sites were predicted using DAS (Cserzo et al. 1997) and the fragments with the highest score and 20–24 amino acid residues were selected. ACR4 was predicted to localize to the plasma membrane, which is consistent with the localization pattern of ACR4–GFP fusion protein in tobacco suspension-cultured cells (Tanaka et al. 2002; Gifford et al. 2003), while the mitochondrial targeting signal prediction by iPSORT was not consistent. It was also predicted to have a cleavage site between positions 29 and 30. AtCRR2 was predicted to have a plasma membrane signal peptide by iPSORT, although it was predicted to localize to endoplasmic reticulum by PSORT. All other CR4-related proteins were predicted to localize to the plasma membrane and contain a single-pass transmembrane fragment. The results are summarized in Table 2.

The protein sequences were analyzed using Superfamily 1.59 (Gough et al. 2001). All the proteins contained a kinase domain falling into the protein kinase-like superfamily, as expected. The crinkly repeats of all 11 proteins fall into the RCC1/BLIP-II superfamily,

consistent with the former report that the ZmCR4 crinkly repeats might form an RCC1-like 7-bladed propeller structure (McCarty and Chory 2000). The sequence alignment of the crinkly repeats with  $\beta$ -lactamase inhibitor, another member of this family, supports this prediction (Fig. 2). ZmCR4, OsCR4, ACR4, AtCRR1 and AtCRR2 contain domains similar to the ligand-binding domain of TNFR, while the remaining proteins of the family do not, consistent with our multiple sequence alignment results. The structures of the crinkly repeats and TNFR repeats were also predicted using 3D-PSSM (Kelley et al. 2000). The TNFR repeats of the three CR4s, AtCRR1 and AtCRR2 were predicted to form a structure similar to TNFR, consistent with the multiple sequence alignment data and superfamily prediction. Similarly, the crinkly repeats are predicted to form the expected seven-bladed propeller.

#### *ACR4*, *AtCRRs* and *AtCRK1* are differentially expressed

Fragments of each coding region were amplified by PCR from *Arabidopsis* genomic DNA. Genomic DNA gel blot analysis confirmed that all probes used for the expression analysis were gene specific (not shown). To investigate the expression sites of the five genes, RNA was isolated from roots, SAMs and leaf primordia, mature leaves, flower buds and developing siliques, subjected to semi-quantitative RT-PCR, and the products analyzed by DNA gel blot. That reactions were within the quantitative range of amplification was confirmed by performing template dilutions (not shown). *ACR4* transcripts were detected in all tissues examined (Fig. 3), but expression in SAMs and flower buds was substantially higher than in roots, which is consistent with our RNA gel blot results (data not shown) and the results previously reported (Tanaka et al. 2002).

*AtCRRs* and *AtCRK1* were also expressed at differential levels in all tissues examined, except that *AtCRK1* was not detected in the developing siliques (Fig. 3). *AtCRR1* and *AtCRR2* were most highly expressed in SAMs and flower buds, similar to *ACR4*. The expression of *AtCRR1* in roots was higher than that of *ACR4* and *AtCRR2*, and the expression of *AtCRR2* in developing siliques was relatively low. The expression patterns of *AtCRR3* and *AtCRK1* were different from those of *ACR4*, *AtCRR1* and *AtCRR2*. Both were highly expressed in roots and mature leaves. The expression of *AtCRK1* in flower buds was lower than that of the other four genes. Interestingly, the expression patterns of the five genes were correlated to their phylogeny. *ACR4*, *AtCRR1* and *AtCRR2* were expressed similarly and formed a distinct phylogenetic cluster. Likewise, *AtCRR3* and *AtCRK1* showed similar expression profiles and were phylogenetically related. This suggests that the related family members might be performing related functions.

#### Expression analysis of *ACR4::GUS*

To further investigate the spatial and temporal expression patterns of *ACR4*, a construct formed by fusion of the *ACR4* promoter with the *GUS* reporter was introduced into *A. thaliana* Columbia. The expression of GUS activity during vegetative growth was analyzed in axenically grown transgenic plants, while the reproductive growth phase was examined in soil-grown plants. The results are shown in Fig. 4. At 1 day after germination, GUS activity was detected in all tissues including the cotyledons, hypocotyls and embryonic roots (not shown). By 3 days after germination, GUS staining remained strong in the hypocotyls and cotyledons, but became weaker in the roots (not shown). By 5 days, seedlings showed the

strongest GUS activity in the two youngest leaves (Fig. 4a). Nine independent transformants gave little or no GUS activity in the roots, while one retained GUS activity in the root, especially in the vasculature and root tip (Fig. 4a). All seedlings checked showed high GUS staining in hypocotyls, especially in the vascular system. Two-week-old seedlings showed variable weak signals in the roots, and GUS activity was mainly detected in the SAM and leaf primordia (Fig. 4b). In expanded leaves, GUS activity declined, becoming localized to the trichomes and vascular systems (Fig. 4b,c). The high *ACR4* promoter activity in the SAM and trichomes was consistent with our RNA in situ hybridization results (data not shown).

*ACR4::GUS* was also expressed during reproductive development. In inflorescences, GUS levels were highest in flower buds and throughout young flowers (Fig. 4e). GUS activity was also detected in cauline branches, especially in the cauline apical meristems and trichomes (Fig. 4d). In mature flowers, GUS staining was detected in sepals and the gynoecium (stigmatic papillae, styles, ovules and ovary walls), but only weakly in petals and filaments, and not in anthers (Fig. 4f-h). In siliques, developing seeds gave a GUS signal and GUS staining was stronger in endosperms than in the developing seed coats (Fig. 4i-k). In later-stage seeds, GUS staining was strongest in the persistent aleurone-like epidermal layer of the endosperm, and in the embryo (Fig. 4k). GUS activities were heritable and consistent in three independent events observed.

The maize *CR4* transcript was reported to be light-induced in seedling leaves (Kang et al. 2002). To test whether *ACR4* transcription was also light-regulated, *ACR4::GUS* seedlings were placed in the dark for 5 days and then stained for GUS. No difference in histological staining was detected between light-grown and dark-grown seedlings, suggesting that *ACR4* is not light-regulated at the transcriptional level.

ACR4 contains a functional serine/threonine kinase, while AtCRR1 and AtCRR2 are inactive

The function of the predicted kinase domain of ACR4 was tested biochemically by an *in vitro* autophosphorylation assay. The ACR4 kinase domain was expressed in *E. coli* as a thioredoxin fusion, ACR4 K. The purified ACR4 K fusion protein was phosphorylated upon incubation with  $\gamma$ -[<sup>32</sup>P]ATP. A substitution of Lys<sup>540</sup> to Ala<sup>540</sup> (ACR4K-Mu) abolished this activity (Fig. 5a), indicating that the detected kinase activity was due to ACR4 and not a bacterial contaminant. This amino acid corresponds to the essential Lys<sup>72</sup> in subdomain II of the type- $\alpha$  cAMP-dependent protein kinase catalytic subunit (Gibbs and Zoller 1991).

Sequence analysis predicted that ACR4 contained a serine/threonine protein kinase domain. Phosphoamino acid analysis verified that autophosphorylation occurred on serine and threonine residues, but not tyrosine residues (Fig. 5b). Thus, as expected, ACR4 contains a functional serine/threonine protein kinase.

Sequence analysis predicted that the kinase domains of AtCRR1 and AtCRR2 would be inactive because of deletions in kinase subdomain VIII. This subdomain contains the activation loop and is thought to be important for substrate binding. It also contains an invariant aspartate that forms an ion pair with an arginine in subdomain X, stabilizing the large lobe of the catalytic domain (Knighton et al. 1991). Substitution of this aspartate with alanine in the type- $\alpha$  cAMP-dependent protein kinase results in a dramatic decrease in kinase activity (Gibbs and Zoller 1991). The AtCRR2 kinase domain was expressed as a GST fusion (AtCRR2 K) in *E. coli*, purified and subjected to an autophosphorylation assay. The result showed that the AtCRR2 kinase domain had very little activity (Fig. 5a). An

autophosphorylation assay showed that AtCRR1 also contains a nearly inactive kinase domain, as predicted (data not shown). A GST fusion of the AtCRK1 kinase domain also showed activity in an autophosphorylation assay (data not shown). A site-mutagenized negative control was not included, but no problems with bacterial contamination were observed in the highly related ACR4K, AtCRR1K and AtCRR2K proteins using the same expression and purification system. This argues that the AtCRK1 kinase is functional *in vitro*.

In known examples of dead-kinase receptor kinases, they typically function as heterodimers with related kinase-active receptor kinases (Kroiber et al. 2001). In those cases, the dead kinase is phosphorylated by the active member of the pair upon ligand binding. We found that ACR4K could phosphorylate AtCRR2K, while ACR4K-Mu could not (Fig. 5a). This is consistent with the possibility that AtCRR1 and AtCRR2 might act as heterodimers with ACR4.

#### T-DNA insertional mutant analyses

The *ACR4*, *AtCRRs* and *AtCRK1* transcripts are all expressed and contain conserved sequences, indicating that they probably function in the plant. To investigate potential developmental functions, T-DNA insertions in each of the five genes were obtained. One mutant allele of *ACR4* (*acr4-1*) was obtained through the Wisconsin knockout facility (Krysan et al. 1999). *acr4-2*, *atcrr1-1* and *atcrr3-1* were from the Syngenta Arabidopsis Insertion Library (SAIL) of the Torrey Mesa Research Institute. *atcrr1-2*, *atcrr2-1*, *atcrr2-2* and *atcrk1-1* were from SIGnAL (Salk Institute Genomic Analysis Laboratory). The T-DNA

insertion sites were verified by sequencing PCR-amplified flanking DNA, and are shown in Fig. 1b. *acr4-1* and *acr4-2* appear to be independent isolates of identical alleles previously reported (Gifford et al. 2003) In *atcrr2-1*, the T-DNA was inserted at 43 bp of 3' UTR (untranslated region). This could potentially interfere with mRNA stability and/or translation efficiency, but this remains to be confirmed. All other mutants contain T-DNA insertions in coding regions.

As previously reported (Gifford et al. 2003), *acr4-1* and *acr4-2* showed defects in integument and seed coat development, but no obvious phenotypic alterations in leaves or any other tissue. This suggests that *Arabidopsis* contains a redundant function not contained in maize. We hypothesized that one or more of the *CRR* genes might encode that redundant function. Under normal growth conditions (21°C, continuous light), *atcrr1-1*, *atcrr1-2*, *atcrr2-1*, *atcrr2-2*, *atcrr3-1* and *atcrk1-1* did not show any obvious mutant phenotypes. To test for redundancy, homozygous double mutants were generated between the following pairs: *acr4-2* and *atcrr1-1*; *acr4-2* and *atcrr2-1*; *atcrr1-1* and *atcrr2-1*; *atcrr3-1* and *atcrk1*. None of the double-mutant combinations showed an obvious mutant phenotype under normal growth conditions either. Because the phylogeny showed that *AtCRR1* and *AtCRR2* formed a clade in *Arabidopsis* that was most similar to the *CR4* clade, and which was absent in rice, we hypothesized that they may function redundantly with *ACR4*. However, the triple-mutant phenotype was indistinguishable from the *acr4* single-mutant (Fig. 6). Therefore, the hypothetical redundant functions that differentiate the *Arabidopsis* *acr4* and maize *cr4* mutant phenotypes do not appear to require the *CRR1* or *CRR2* genes.

## Discussion

### Conservation and divergence of CR4 family members in monocots and dicots

The *Arabidopsis* genome is predicted to encode more than 400 RLKs, among which CR4 represents a subfamily (Shiu and Bleecker 2001a, 2001b). The RLK family has expanded to nearly twice that size in rice (Shiu et al. 2004). Five CR4-related genes were obtained by database searches in *A. thaliana* and five in *O. sativa*. The OsCRR amino acid sequences were predicted from existing rice genome sequences. Both ACR4 and OsCR4 share high sequence similarity with ZmCR4 and contain all the amino acid sequence features, arguing that they are orthologs of ZmCR4. This suggests that CR4 might be well conserved in monocots and dicots. As expected, ZmCR4 and OsCR4 have higher amino acid sequence similarity with each other than with AtCR4. The high conservation among CR4 proteins also suggests that they might have conserved functions in plant development, but surprisingly *acr4* mutants show phenotypic effects restricted to the integuments and seed coat, in stark contrast to maize where *cr4* mutants severely alter cellular morphology throughout the shoot tissues (Becraft et al. 1996, 2001; Jin et al. 2000).

Although CR4s show high conservation, other members of the CR4 family show significant divergence. No orthologs of AtCRR1 and AtCRR2 were found in rice. Extensive public (NCBI) and private (Syngenta) database searches only identified four CRRs in rice. All four lack the TNFR repeats, contain kinase domains predicted to be active, and cluster with AtCRR3 and AtCRK1 in the phylogenetic analysis. Since maize is more closely related to rice than to *Arabidopsis*, this argues that there are probably no orthologs of AtCRR1 and

AtCRR2 in maize either, and database searches to date bear this out. Multiple sequence alignment and phylogenetic analysis suggested that OsCRR3 might be the ortholog of AtCRR3, but AtCRK1, OsCRR1, 2, and 4 are quite divergent and orthologous relationships were not readily apparent. Thus, it appears that both rice and *Arabidopsis* contain some CRRs with no orthologs in other species.

#### Motifs and structure prediction of the CR4 family

Like other plant RLKs, members of the CR4 family contain a signal peptide, predicted plasma membrane localization, a single-pass transmembrane fragment and a cytoplasmic protein kinase domain. The CR4 family formed a clade in a phylogenetic analysis of kinase domains (Shiu and Bleecker 2001a, 2001b). This family also contains related extracellular domains. The distinguishing feature of the family is the presence of seven 'crinkly' repeats in the extracellular domain. McCarty and Chory (2000) predicted that this motif in ZmCR4 might form an RCC1-like seven-bladed propeller structure. The sequence similarity of this domain with  $\beta$ -lactamase inhibitor, another seven-bladed propeller protein (Lim et al. 2001), and 3D-PSSM analysis (Kelley et al. 2000) support this prediction. The seven-bladed  $\beta$ -propeller of  $\beta$ -lactamase inhibitor protein-II is involved in the interaction with TEM-1  $\beta$ -lactamase (Lim et al. 2001). This suggests that this domain might also be involved in protein interactions in the CR4 RLK family.

Another feature of the CR4 family is that some members contain TNFR repeats in their extracellular domains. In AtCRR1, AtCRR2, and the three CR4 proteins, there are 12 highly conserved cysteine residues in this domain, which have potential to form disulfide bonds. In

the TNFR family, the presence of repeating cysteine-rich units is a characteristic feature and disulfide-bridge formation is crucial for their functions (Banner et al. 1993; Idriss and Naismith 2000). The extracellular region of human sTNF-R55 contains 24 conserved cysteine residues in 4 repeats and connectivity of 9 disulfide bonds in the first 9 repeats has been determined by crystal structure (Banner et al. 1993). 3D-PSSM (Kelley et al. 2000) modeling results support the possibility that this domain in the CR4 proteins, AtCRR1 and AtCRR2 might form a structure similar to TNFR. In rice, and by extrapolation probably maize, CR4 is the only TNFR-containing protein. The other CRR proteins contain six cysteine residues in this region, which are conserved within this clade, but not with TNFR-containing proteins. No known structures were identified in this region of the non-TNFR CRR proteins.

#### Activities of ACR4 family kinase domains

Kinase assays demonstrated that ACR4 is an active serine/threonine kinase, while AtCRR1 and AtCRR2 were nearly inactive in autophosphorylation assays. The active ACR4 kinase domain can phosphorylate the AtCRR2 inactive kinase domain. AtCRR1 and AtCRR2 are the first examples of dead-kinase RLKs known to be expressed in plants. In other known examples of kinase-inactive receptors, they often heterodimerize with a related kinase-active receptor. For example, in mammals, the kinase-inactive ErbB3 forms a complex with kinase-active EGFR upon ligand binding, is phosphorylated by the active partner, and serves as docking sites for downstream signaling molecules (reviewed in Kroiher et al. 2001). This raises the possibility that AtCRR1 and AtCRR2 might form heterodimers or heteromultimers

with ACR4 or other members of this family. The similar expression patterns and phosphorylation of AtCRR2 by ACR4 are consistent with such a scenario in *Arabidopsis*, although the lack of mutant phenotypes argues that this hypothetical heterodimerization would not be necessary for ACR4 function.

There are also several examples in animals where the kinase activities are not required for receptor kinases to function. Upon growth hormone treatment, epidermal growth factor receptor (EGFR) is phosphorylated by the cytoplasmic kinase JAK2, providing docking sites for Grb2 and activating a MAP kinase (mitogen-activated protein kinase) and gene expression; the intrinsic kinase activity of EGFR is not essential for this response (Yamauchi et al. 1997). ILK (integrin-linked kinase), lacking the essential DFG motif and kinase activity, can regulate the S473 phosphorylation of PKB (protein kinase B), presumably by acting as an adaptor to recruit a S473 kinase or phosphatase (Lynch et al. 1999). In *Arabidopsis*, CLAVATA2 (CLV2) lacks a cytoplasmic kinase domain (Jeong et al. 1999). CLV2 is required for the accumulation of CLV1, and is hypothesized to heterodimerize with CLV1 to bind the presumed ligand, CLV3 (Jeong et al. 1999). Thus, it is possible that AtCRR1 and 2 could perform receptor functions in the absence of kinase activity.

#### Expression of the CR4 family in *Arabidopsis*

RT-PCR and the *ACR4::GUS* fusion showed that *ACR4* is expressed in most tissues examined. The expression is high in developing shoot tissues including the SAM and organ primordia, flower buds, and pistils, consistent with previous reports (Tanaka et al. 2002; Gifford et al. 2003). *ZmCR4* is also mainly expressed in developing shoot tissues (Jin et al.

2000). The expression of *ACR4::GUS* in leaves declines as tissues mature. Several aspects of the expression pattern we observed have not been previously reported. We observed expression throughout the germinating seedling, becoming restricted to apical growth zones over a period of 5–7 days, whereas Gifford et al. (2003) report that post-germination expression is restricted to apical meristems and organ primordia. As expression declined in maturing organs, GUS expression was retained in vasculature and in leaf trichomes. The potential role that *ACR4* might play in these cells is not clear. In contrast to previously reported results of *ACR4* promoter–reporter construct expression (Gifford et al. 2003), we observed clear expression of *ACR4::GUS* in the persistent aleurone-like epidermal layer of the endosperm, and while there may have been preferential expression in the embryo epidermis, it was not epidermis-specific (Fig. 4k). RT–PCR showed that *AtCRR1* and *AtCRR2* are expressed in similar organs to *ACR4*. On the other hand, the expression of *AtCRR3* and *AtCRK1* is high in roots and mature leaves, distinct from *ACR4*, *AtCRR1* and *AtCRR2*. This suggests that *AtCRR3* and *AtCRK1* perform different functions from *ACR4*, *AtCRR1* and *AtCRR2*, although the overlapping expression patterns do not exclude the possibility of overlapping functions.

#### Mutants of CR4 family members lack obvious phenotypes in *Arabidopsis*

T-DNA insertion mutants were obtained for each of the five CR4 family member genes in *Arabidopsis* (Fig. 1b). Consistent with a previous report (Gifford et al. 2003), we observed defects in integument and seed coat development, but no other obvious phenotypic effects in *acr4* mutants. It was also reported that antisense *ACR4* produced defects in embryo

morphogenesis (Tanaka et al. 2002). In our study, we obtained two independent T-DNA insertions in coding sequence, in two different ecotypes, and careful examination did not reveal any of the embryo defects reportedly induced by the antisense construct. One possible explanation for this discrepancy is that the antisense construct contained the entire cDNA, which might inhibit multiple *AtCRRs*, or possibly one or more of the three highly related cytoplasmic kinases (Shiu and Bleecker 2001a, 2001b).

The lack of phenotypic defects in leaves and other shoot organs is surprising considering the severe mutant phenotype of maize *cr4* mutants and the high sequence conservation between the maize and *Arabidopsis* CR4 proteins. This could be explained by functional redundancy in *Arabidopsis* or if the mutants are not completely null. The location of the insertions and molecular characterization of the mutants (not shown) make it unlikely that the mutant genes are still functional. One possibility is that some of the *AtCRRs*, particularly *AtCRR1* and 2, might function redundantly with *ACR4*. Several lines of evidence suggest this hypothesis. First, multiple sequence alignment showed that *ACR4*, *AtCRR1* and *AtCRR2* have similar extracellular structures. Second, phylogenetic analysis grouped *ACR4*, *AtCRR1* and *AtCRR2* as a cluster. Third, the expression patterns are similar among *ACR4*, *AtCRR1* and *AtCRR2*. Examples cited above of receptor kinases that function independently of kinase activity allow the possibility for the kinase-inactive *AtCRR1* and 2 to function redundantly with *ACR4*. If the apparent lack of *AtCRR1* and 2 orthologs in rice is a general feature of grasses, then the maize *cr4* mutant phenotype could result from the lack of these redundant functions.

The hypothesis that the *crinkly4* related genes functioned in related developmental processes was tested genetically by identifying T-DNA insertional mutants. None of the

homozygous mutants, nor the double mutants including *atcrr1-1;atcrr2-1*, showed any obvious phenotypic defects, including integument development. Furthermore the *acr4-2;atcrr1-1;atcrr2-1* triple mutant showed a mutant phenotype that was indistinguishable from the *acr4-2* single mutant. Thus, these results are not consistent with the *AtCRR1* or *AtCRR2* genes providing redundancy with *ACR4*. An alternate hypothesis is that a biologically redundant function might be provided by a gene or pathway unrelated to *CR4*. Further analysis of these mutants in different growth conditions might reveal hidden functions that the *CR4* family plays in *Arabidopsis*.

### **Acknowledgements**

The authors thank the Becraft lab for discussions and critical reading of the manuscript. We are grateful to Syngenta and Torrey Mesa Research Institute for providing T-DNA mutants and access to their rice genomic sequence, Frans Tax and the Functional Genomics of Plant Phosphorylation project for screening for a T-DNA mutant, the Salk Institute Genomic Analysis Laboratory and the Arabidopsis Biological Resource Center for providing T-DNA mutant seeds, and the Bessey Microscopy Facility at Iowa State University for technical assistance and sample preparation. This research was funded by grants DE-FG02-98ER20303 from the U.S. DOE Energy Biosciences and IBN:96-04426 from the U.S. National Science Foundation to P.W.B., and by grant No. R01-2000-000-00195-0 from the Basic Research Program of the Korea Science & Engineering Foundation to S-G.S.

**References**

- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990)** Basic local alignment search tool. *J Mol Biol* 215:403–410
- Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997)** Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25:3389–3402
- Bannai H, Tamada Y, Maruyama O, Nakai K, Miyano S (2002)** Extensive feature detection of N-terminal protein sorting signals. *Bioinformatics* 18:298–305
- Banner DW, D'Arcy A, Janes W, Gentz R, Schoenfeld H-J, Broger C, Loetscher H, Lesslauer W (1993)** Crystal structure of the soluble human 55 kd TNF receptor-human TNF $\beta$  complex: implications for TNF receptor activation. *Cell* 73:431–445
- Bechtold N, Pelletier G (1998)** In planta *Agrobacterium*-mediated transformation of adult *Arabidopsis thaliana* plants by vacuum infiltration. *Methods Mol Biol* 82:259–266
- Becraft PW (2001)** Plant steroids recognized at the cell surface. *Trends Genet* 17:60–62
- Becraft PW (2002)** Receptor kinase signaling in plant development. *Annu Rev Cell Devel Biol* 18:163–192
- Becraft PW, Asuncion-Crabb YT (2000)** Positional cues specify and maintain aleurone cell fate during maize endosperm development. *Development* 127:4039–4048
- Becraft PW, Stinard PS, McCarty DR (1996)** CRINKLY4: a TNFR-like receptor kinase involved in maize epidermal differentiation. *Science* 273:1406–1409

- Becraft PW, Kang S-H, Suh S-G** (2001) The maize CRINKLY4 receptor kinase controls a cell-autonomous differentiation response. *Plant Physiol* 127:486–496
- Boyle WJ, van der Geer P, Hunter T** (1991) Phosphopeptide mapping and phosphoamino acid analysis by two-dimensional separation on thin-layer cellulose plates. *Methods Enzymol* 201:110–149
- Cserzo M, Wallin E, Simon I, von Heijne G, Elofsson A** (1997) Prediction of transmembrane alpha-helices in prokaryotic membrane proteins: the dense alignment surface method. *Protein Eng* 10:673–676
- Gibbs CS, Zoller MJ** (1991) Rational scanning mutagenesis of a protein kinase identifies functional regions involved in catalysis and substrate interactions. *J Biol Chem* 266:8923–8931
- Gifford ML, Dean S, Ingram GC** (2003) The *Arabidopsis* ACR4 gene plays a role in cell layer organisation during ovule integument and sepal margin development. *Development* 130:4249–4258
- Goff SA, Ricke D, Lan TH et al** (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). *Science* 296:92–100
- Gough J, Karplus K, Hughey R, Chothia C** (2001) Assignment of homology to genome sequences using a library of hidden Markov models that represent all proteins of known structure. *J Mol Biol* 313:903–919
- Idriss HT, Naismith JH** (2000) TNF alpha and the TNF receptor superfamily: structure–function relationship(s). *Microsc Res Tech* 50:184–195
- Jefferson RA, Kavanagh TA, Bevan MW** (1987) GUS fusions: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *EMBO J* 6:3901–3907

- Jeong S, Trotochaud AE, Clark SE (1999)** The *Arabidopsis CLAVATA2* gene encodes a receptor-like protein required for the stability of the CLAVATA1 Receptor-like kinase. *Plant Cell* 11:1925–1933
- Jin P, Guo T, Becraft PW (2000)** The maize CR4 receptor-like kinase mediates a growth factor-like differentiation response. *Genesis* 27:104–116
- Kang S-G, Lee HJ, Suh S-G (2002)** The maize *crinkly4* gene is expressed spatially in vegetative and floral organs. *J Plant Biol* 45:219–224
- Kelley LA, MacCallum RM, Sternberg MJ (2000)** Enhanced genome annotation using structural profiles in the program 3D-PSSM. *J Mol Biol* 299:499–520
- Knighton DR, Zheng J, Ten Eyck LF, Ashford VA, Xuong N-H, Taylor SS, Sowadski JM (1991)** Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. *Science* 253:407–414
- Kroiher M, Miller MA, Steele RE (2001)** Deceiving appearances: signaling by dead and fractured receptor protein-tyrosine kinases. *Bioessays* 23:69–76
- Krysan PJ, Young JC, Sussman MR (1999)** T-DNA as an insertional mutagen in *Arabidopsis*. *Plant Cell* 11:2283–2290
- Kumar S, Tamura K, Jakobsen IB, Nei M (2001)** MEGA2: molecular evolutionary genetics analysis software. *Bioinformatics* 17:1244–1245
- Lim D, Park HU, De Castro L, Kang SG, Lee HS, Jensen S, Lee KJ, Strynadka NC (2001)** Crystal structure and kinetic analysis of beta-lactamase inhibitor protein-II in complex with TEM-1 beta-lactamase. *Nat Struct Biol* 8:848–852

- Lynch DK, Ellis CA, Edwards PA, Hiles ID** (1999) Integrin-linked kinase regulates phosphorylation of serine 473 of protein kinase B by an indirect mechanism. *Oncogene* 18:8024–8032
- McCarty DR, Chory J** (2000) Conservation and innovation in plant signaling pathways. *Cell* 103:201–209
- Murashige T, Skoog F** (1962) A revised medium for rapid growth and bioassays with tobacco tissue culture. *Physiol Plant* 15:473–497
- Nakai K, Horton P** (1999) PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. *Trends Biochem Sci* 24:34–36
- Nielsen H, Engelbrecht J, Brunak S, von Heijne G** (1997) Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng* 10:1–6
- Pearson WR, Lipman DJ** (1988) Improved tools for biological sequence comparison. *Proc Natl Acad Sci USA* 85:2444–2448
- Saitou N, Nei M** (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4:406–425
- Sambrook J, Fritsch EF, Maniatis T** (1989) *Molecular cloning: a laboratory manual*, 2nd edn. CSHL Press, Cold Spring Harbor, NY
- Schafer S, Schmulling T** (2002) The CRK1 receptor-like kinase gene of tobacco is negatively regulated by cytokinin. *Plant Mol Biol* 50:155–166
- Shiu SH, Bleeker AB** (2001a) Plant receptor-like kinase gene family: diversity, function, and signaling. *Sci Sig Trans Knowl Environ* 2001:RE22. 2

- Shiu S-H, Bleecker AB** (2001b) Receptor-like kinases from *Arabidopsis* form a monophyletic gene family related to animal receptor kinases. *Proc Natl Acad Sci USA* 98:10763–10768
- Shiu S-H, Karlowski WM, Pan R, Tzeng, Y-H, Mayer KFX, Li W-H** (2004) Comparative analysis of the receptor-like kinase family in *Arabidopsis* and rice. *Plant Cell* 16:1220–1234
- Smyth DR, Bowman JL, Meyerowitz EM** (1990) Early flower development in *Arabidopsis*. *Plant Cell* 2:755–767
- Tanaka H, Watanabe M, Watanabe D, Tanaka T, Machida C, Machida Y** (2002) *ACR4*, a putative receptor kinase gene of *Arabidopsis thaliana*, that is expressed in the outer cell layers of embryos and plants, is involved in proper embryogenesis. *Plant Cell Physiol* 43:419–428
- Thompson JD, Higgins DG, Gibson TJ** (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22:4673–4680
- Yamauchi T, Ueki K, Tobe K, Tamemoto H, Sekine N, Wada M, Honjo M, Takahashi M, Takahashi T, Hirai H, Tushima T, Akanuma Y, Fujita T, Komuro I, Yazaki Y, Kadowaki T** (1997) Tyrosine phosphorylation of the EGF receptor by the kinase Jak2 is induced by growth hormone. *Nature* 390:91–96

## Figure legends

**Fig. 1 a** Sequence alignments of CRINKLY4-related proteins from *Arabidopsis thaliana* and *Oryza sativa*. The following are indicated: conserved (*black shading*) and similar (*gray shading*) residues; the seven 'crinkly' repeats (*solid bars*); the three TNFR repeats (*solid bars with round ends*); the predicted transmembrane domain of ZmCR4 (*dotted line labeled TM*); the twelve conserved kinase subdomains (*solid bars with diamond ends*); the region of subdomain VIII deleted in AtCRR1 and AtCRR2 (*box*). **b** Diagrammatic representation of CR4-related proteins. The numbers on the diagram represent identity to maize (*Zea mays*) ZmCR4 of each domain. ACR4, AtCRR1, 2, 3 and AtCRK1 are labeled with *arrows* to show T-DNA insertion sites and left border directions in the corresponding T-DNA insertion alleles. **c** Phylogenetic analysis of CR4-related genes from *A. thaliana* and *O. sativa*. The multiple sequence alignment used is the same as in **a**. The phylogenetic analysis was carried out by the neighbor-joining method (Saitou and Nei, 1987) using MEGA version 2.1 (Kumar et al., 2001). 1,000 bootstrap replicates were calculated and bootstrap values above 50% are indicated at each node

**Fig. 2** Sequence alignments of the crinkly repeats of ACR4, AtCRR1 and AtCRK1 with  $\beta$ -lactamase inhibitor. Conserved and similar residues to  $\beta$ -lactamase inhibitor are marked in *black* and *gray*, respectively. *b\_lac\_in*  $\beta$ -Lactamase inhibitor. The seven conserved WG motifs are marked by *asterisks*

**Fig. 3** Expression of CR4-related genes in *A. thaliana*. Semi-quantitative RT-PCR was used to study the expression levels in different organs. RT-PCR products were blotted, hybridized with corresponding radiolabeled probes and detected by a phosphorimager. *UBQ10* was used as a control for equal loading. *F-buds* Flower buds; *-RT* pool of all five RNA samples without reverse transcriptase; *G-DNA* 10 ng genomic DNA, allowing normalization of the signal across the different genes.

**Fig. 4a–k** Histochemical analysis of GUS activity in transgenic *A. thaliana* plants transformed with chimeric *ACR4::GUS* including 1.38 kb of the *ACR4* 5' region fused to the GUS coding sequence. **a** Seedlings 5 days after germination. The *arrow* shows GUS staining in the root tip. **b** Ten-day old seedling. **c** Close-up of an expanded leaf from **b**, showing trichome staining. **d** A cauline branch from a 30-day old plant. **e** Primary inflorescence apex with open flower. **f** Stage-12 flower (Smyth et al., 1990). **g** Stage-11 pistil. **h** An ovule from the stage-11 pistil in **g**. **i** A seed at 4 days after anthesis, showing GUS expression in the endosperm. **j** Stage-16 silique. **k** A sectioned seed showing expression in the aleurone-like layer of the endosperm (*arrow*), and in the embryo. **g–i**, and **k** were taken under a transmitted-light compound microscope; the others were photographed with a stereo microscope.

**Fig. 5a,b** Kinase and phosphoamino acid analyses. **a** Autophosphorylation assays of ACR4 and AtCRR2 and transphosphorylation of AtCRR2 by ACR4. *Lanes: 1*, ACR4 kinase domain fusion protein (ACR4 K); *2*, inactive ACR4 kinase fusion protein with a site-directed mutation of lysine<sup>540</sup> to alanine (ACR4K-Mu); *3*, CRR2 kinase domain GST fusion protein

(AtCRR2 K); 4, ACR4 K + AtCRR2 K; 5, ACR4K-Mu + AtCRR2 K. *Filled arrowheads* ACR4 K; *open arrowheads* AtCRR2 K. **b** ACR4 autophosphorylates on serine and threonine residues. Autophosphorylated ACR4 K was acid-hydrolyzed and subjected to two-dimensional thin-layer electrophoresis. The autoradiograph revealed labeling of serine and threonine but not tyrosine residues. *P-Ser* Phospho-o-serine, *P-Thr* phospho-o-threonine, *P-Tyr* phospho-o-tyrosine

**Fig. 6a–d** Phenotypes of seeds from homozygous *acr4* and *crr* mutant *A. thaliana* plants. **a** Wild-type seeds of the Columbia ecotype. **b** *acr4-2* mutant seeds. **c** *crr1-1*; *crr2-1* double-mutant seeds. **d** *acr4-2*; *crr1-1*; *crr2-1* triple-mutant seeds. All mutants are in the Columbia genetic background. Bars = 1 mm

Table 1. Primers used in RT-PCR analysis. *FP* Forward primer, *RP* reverse primer

<i>ACR4</i>	FP	5'-ATAGCGGCTCTCCTTTGGCAGAACT-3',
	RP	5'-CTCCTTCAGATGCCGATGATTCCTCCT-3';
<i>AtCRR1</i>	FP	5'-CCGAAATGGGAGAGAGGCTTCTTGTT-3'
	RP	5'-CGCAAGTTCCGACATAGTTGGTTGTTGA-3';
<i>AtCRR2</i>	FP	5'-CGGGAGAGGTAGCTTCGGGTTTGTCTAT-3'
	RP	5'-TTGCGCACATTTACCGTCTAACAAGAT-3'
<i>AtCRR3</i>	FP	5'-CGCAGTGCCACCGATTATTCATAGA-3'
	RP	5'-CAAATGCCACTAGAGATACTCCCATGAC-3'
<i>AtCRK1</i>	FP	5'-CTAGCCGATCACCTTCACAATCCTCAAT-3'
	RP	5'-CGGCCAAAGCACTCTCTAGTTTACTGAC-3'
<i>UBQ10</i>	FP	5'-GATCTTTGCCGAAAACAATTGGAGGATGGT-3'
	RP	5'-CGACTTGTCATTAGAAAGAAAGAGATAACAGG-3'

**Table 2.** Predicted signal-peptide cleavage sites and transmembrane sites. *TM*:

Transmembrane, *Num. of AA*: number of amino acid residues

	Cleavage Site <sup>a</sup>	TM site <sup>b</sup>	Num of AA
ACR4	29-30	434-454	895
AtCRR1	23-24	440-460	775
AtCRR2	22-23	433-452	776
AtCRR3	32-33	395-419	814
AtCRK1	31-32	367-386	751
OsCR4	34-35	425-443	901
OsCRR1	22-23	400-420	822
OsCRR2	30-31	399-418	836
OsCRR3	20-21	362-385	798
OsCRR4	30-31	396-417	811

<sup>a</sup>Cleavage sites were predicted by the program SignalP (Nielsen et al. 1997)

<sup>b</sup>Transmembrane sites were predicted by the program DAS (Cserzo et al. 1997) and the fragments with the highest score and 20–24 amino acid residues were selected.

I

```

ZmCR4 1 DHVPAIAGCCGAEIPWVCGLGS...
OsCR4 1 DIVPVAACCIPSWAYLGMASIA...
ACR4 1 RMFETRAWRVIRKPTSIWLASAL...
AtCR1 1 ETRCSLHSLVYKPKSGFGSGPI...
AtCR2 1 CNSHIFVYSSLSLSSVYSGS...
OsCR1 1 SPHILHFLASPEVSSSFP...
OsCR4 1 PSRRHGCGARVSLPULLLRTAA...
OsCR2 1 PMSRHGALITSEFAAVLPGHAA...
OsCR3 1 LLLIITPIAPASAST...
AtCR3 1 KRFINSTPTSTHATTEFFLL...
AtCR1 1 ALTISCFESYFMSV...

```

II III

```

ZmCR4 88 DLRPEVGSNSYVK...
OsCR4 86 DLRPEVGSNSYVK...
ACR4 93 DLRPEVGNATFIQ...
AtCR1 102 CNSHIFVYSSLSLSSVYSGS...
AtCR2 92 CNSHIFVYSSLSLSSVYSGS...
OsCR1 111 PMSRHGALITSEFAAVLPGHAA...
OsCR4 102 RAGDVAHMDLQKRTHTPTNGTKDGGSRK...
OsCR2 104 CHAAHMSFDEAAM...
OsCR3 81 AAGAPAPQGLRRVYVNGQLS...
AtCR3 96 SLDNII...
AtCR1 97 SNSTELRFRVDGTM...

```

IV V VI

```

ZmCR4 173 FARNR...
OsCR4 171 FARNR...
ACR4 180 GRS...
AtCR1 199 GIR...
AtCR2 186 VTK...
OsCR1 182 VVG...
OsCR4 188 HOARF...
OsCR2 181 VVGNASAT...
OsCR3 151 AETS...
AtCR3 170 GIE...
AtCR1 170 SQP...

```

VII VIII

```

ZmCR4 287 IKMY...
OsCR4 285 IRVY...
ACR4 296 IES...
AtCR1 297 SSK...
AtCR2 286 NTS...
OsCR1 282 LA...
OsCR4 290 LA...
AtCR2 277 VGRAS...
OsCR3 251 AVS...
AtCR3 277 TDS...
AtCR1 261 LDH...

```

IX TM

```

ZmCR4 394 LK...
OsCR4 393 K...
ACR4 394 K...
AtCR1 409 L...
AtCR2 399 L...
OsCR1 385 S...
OsCR4 387 K...
OsCR2 381 A...
OsCR3 349 T...
AtCR3 382 S...
AtCR1 359 TW...

```

I II III IV V

```

ZmCR4 490 ...
OsCR4 489 ...
ACR4 497 ...
AtCR1 505 ...
AtCR2 504 ...
OsCR1 464 ...
OsCR4 455 ...
OsCR2 466 ...
OsCR3 446 ...
AtCR3 481 ...
AtCR1 428 ...

```

VIA VIB VII VIII

```

ZmCR4 587 ...
OsCR4 586 ...
ACR4 594 ...
AtCR1 602 ...
AtCR2 601 ...
OsCR1 573 ...
OsCR4 562 ...
OsCR2 572 ...
OsCR3 556 ...
AtCR3 581 ...
AtCR1 535 ...

```

IX X XI

```

ZmCR4 687 ...
OsCR4 686 ...
ACR4 694 ...
AtCR1 676 ...
AtCR2 677 ...
OsCR1 692 ...
OsCR4 675 ...
OsCR2 680 ...
OsCR3 663 ...
AtCR3 689 ...
AtCR1 634 ...

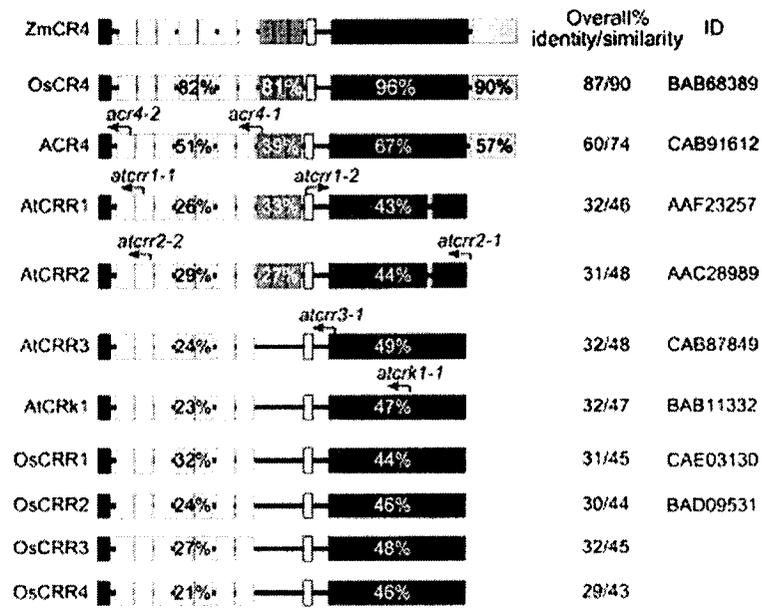
```

```

ZmCR4 796 LVQNS...VRA...
OsCR4 795 VQR...V...R...C...
ACR4 803 SWRIG...KR...G...T...C...A...E...E...E...R...K...Q...G...S...
ALCRR1 .....
ALCRR2 .....
OsCRR1 803 AGAG...EP...L...C...D...G...T...A...
OsCRR4 789 VTGA...RR...L...G...D...Q...F...D...L...D...T...D...
OsCRR2 795 AQQNS...SR...H...G...D...V...S...R...E...S...D...D...P...S...P...H...T...R...E...F...A...G...V...G...H...G...R...S...N...S...A...M...
OsCRR3 775 AAGH...N...S...S...A...L...S...I...T...M...E...L...R...M...D...
ALCRR3 803 S...I...S...I...C...L...V...D...
ALCRK1 747 RSNTY...

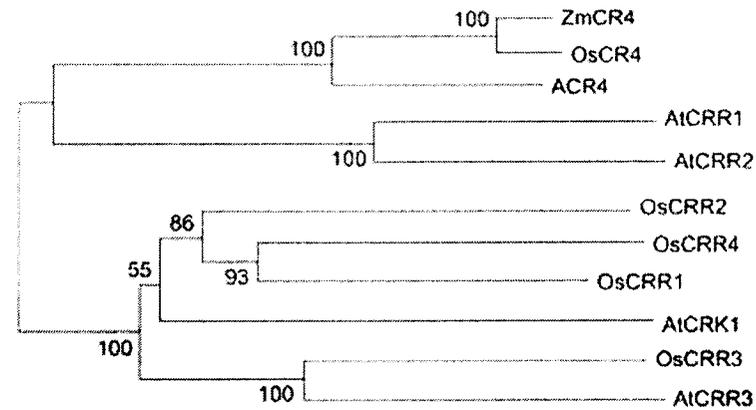
```

Fig. 1a



**b**

- Signal peptide
- ▨ Repeats
- ▤ TNFR
- ▩ Transmembrane
- Kinase
- ▨ Carboxyl tail



**c** 0.1

Fig. 1b,c

```

**
b_lac_in 1 .....KPNGIVRLHREKAL...TMSLLEQLLAAPGVAAEPSSVSVAAATSEAAACSMWGDVAVVEE...G...
ACR4 1 WQLASALGSMSSIAISY...E...GG...F...K...K...D...G... ..HL...O...SS...AIL...YGT...HL...FI...
AtCRR1 1 .....AS...G...A...F...C...D...G... ..R...D...H...K...Y...S...S...P...S...S...S...H...A...S...T...S...A...S...
AtCRK1 1 .....I...S...H...I...S...N...O...T...L...Y...C...A...N...N...H...S... ..Y...E...O...C...S...S...F...L...A...S...I...P...F...S...T...E...N...L...R...N...

**
b_lac_in 69 .....E...G...G...I...P...H...G...I...A...G...S... ..G...G...K...I...G...G...E...N... ..I...T...Y...A...A...T...Q...S...V... ..A...A...F...A...G...E...S...I...A...L...D...G... ..
ACR4 56 .....GLTGGD...P...M...C...H...L...L... ..S...H...O...P...Y...C...S...N...S...A...F...I... ..M...O...V...I...Q...P...M...T...K...A...E...Y...L...E...V...S... ..D...E...L...C...G...I...R...K...P...I...V...G...R...R...E...N...S...N...
AtCRR1 50 YNIPKMAVLSG...G...S...F...L...C...H...L...S... ..N...T...S...Q...A...F...C...S...S...L...G...S...S...S...O...D...U...P...L...A...Y...R...T...T...A...Y...S...O...L...A...N...S... ..V...C...A...V...R...G...Y...Y...S...D...H...D...S...G...I...
AtCRK1 40 ...R...F...S...G...V...V...S... ..C...N...E...Y...C...H...I...S...R...L...C...S...N...T...S...T... ..C...A...R...F...S...V...D... ..T... ..N...M...H...K...R...I...Y...H... ..P...E...I... ..E...L...E...E...N...E...R...I...C...G...V...E...R...V...S...R... ..

**
b_lac_in 122 .....E...I...A...G...G...E...D...G... ..T...T...V...P...A...E...R...S...Q...V...D...A...I...A...G... ..S...A...S...Y...A...L...K... ..D...G... ..F... ..A...A...G...S...S... ..D...G... ..P... ..P...A...E...D...G... ..
ACR4 123 IISSSL...D...G... ..C...H...N...T...R...N...F...V...F... .. ..D...K... ..H...S...L...S... ..S...E...F...N... ..O... ..S...S...K... ..I...S... ..F... ..C... ..A... ..S... ..N...S...S...Q...V...I...S...I... ..P...K...E...K...
AtCRR1 124 .....T...I...D...O...E...I...T...R...A...T...N...N...S...L...I...A...K...E...N...P...N...F...Y...D...I...V... ..N...L...V...P... ..N... ..I...S... ..G... ..G... ..F...S... ..C...G... ..G... ..I...R... ..D...G... ..M... ..L... ..C... ..E... ..P... ..N... ..S...S... ..N... ..L... ..G... ..F... ..N... ..T... ..S... ..D...
AtCRK1 108 .....R...I...R... ..C... ..O... ..P... ..Y... ..L... ..P... ..R... .. ..P... ..D... ..N... ..Y... ..R... ..S... ..L... ..L... ..D... ..N... ..E... ..F... ..C... ..G... ..S... ..O... ..P... ..P... ..G... ..M... ..I... ..S... ..C... ..E... .. ..I... ..A... ..K... ..V... ..P... ..S... ..C... ..D...

**
b_lac_in 181 SGVTA...D...V...Y...T...N...L...A...V...R... ..G... ..V... ..R... ..A... ..D... ..N... ..Y... .. ..G... ..G... ..O... ..T... ..T... ..F... ..A... ..R... .. .. ..A... ..C... ..S... ..G... ..U... ..D... ..E... ..V... ..A... ..S... ..S... ..I... ..E... ..E... ..S... ..L... ..A... ..E... .. ..D... ..E... ..N... ..G... ..A... ..K... ..
ACR4 186 K...F...Q...I...A...A...S... ..C... ..H... ..V... ..C... ..G... ..I... ..D... ..E... ..L... ..E... ..S... ..R... ..L... ..O... ..K... ..S... ..L... ..E... ..F... ..E... ..R... ..E... ..V... ..G... ..T... ..S... ..T... ..E... ..K... ..I... ..L... ..D... ..L... ..P... ..P... ..K... ..E... ..P... ..L... ..A... ..V... ..S... ..K... ..E... ..Y... ..A... ..C... ..G... ..I... ..P... ..R... ..Y... ..H... ..S... ..A... ..V... ..C... ..N... ..E... ..F... ..
AtCRR1 194 N...F...O...V... ..A... ..K... ..N... ..E... ..V... ..C... ..A... ..L... ..L... .. ..S... ..R... ..E... ..K... ..O... ..S... ..E... ..E... .. ..S... .. ..V... ..S... ..E... ..M... ..N... ..E... .. .. ..S... ..R... ..F... ..V... ..S... ..L... ..A... ..P... ..R... ..F... ..C... ..G... ..I... ..R... ..E... ..D... ..H... ..E... ..V... ..E... ..C... ..A... ..S... ..
AtCRK1 156 H...Y...I... ..I... ..A... ..G... ..S... ..R... ..C... ..O... ..T... ..V... ..D... .. ..N... ..D... ..E... ..C... ..O... ..T... .. .. ..C... ..S... ..L... ..E... ..R... .. .. .. ..R... ..K... ..F... ..L... ..A... ..V... ..E... ..D... ..E... ..G... ..G... ..V... ..R... ..W... ..S... ..N... ..E... ..V... ..C... ..A... ..S... ..

**
b_lac_in 245 ..L...N... ..K... ..O... ..T... ..Y... ..T... ..E... ..A... ..I... ..S... ..V... ..S... ..A... ..T... ..A... ..G... ..E... ..R... ..T... ..L... ..A... ..N... ..K... .. .. ..L... ..A... ..G... ..S... ..S... ..P... ..S... ..V... ..G... ..S... ..C... ..V... ..S... ..G... ..I... ..E... ..A... ..G... ..P... ..N... ..A... ..A... ..Y... ..A... ..L... ..K... ..G... ..
ACR4 265 F...V... ..S... ..T... ..P... ..A... ..P... ..T... .. .. ..G... ..I... ..F... ..Y... ..D... ..L... ..A... ..N... ..Y... ..E... ..I... ..C... ..G... ..L... ..T... ..S... ..M... ..S... ..P... ..Y... ..O... ..S... ..L... ..G... ..F... ..E... ..A... ..S... ..I... ..P... ..L... ..A... ..V... ..S... ..P... ..
AtCRR1 259 .. ..N... ..E... ..N... ..F... ..S... ..L... ..I... ..K... ..E... ..S... .. ..F... ..P... ..K... ..A... ..S... ..S... ..E... ..I... ..V... ..C... ..G... ..I... ..R... ..E... ..D... .. ..L... ..M... ..D... ..C... ..M... ..V... ..N... ..G... ..S... ..T... ..H... ..Y... ..D... ..P... ..
AtCRK1 215 N...S... ..M... ..S... ..L... ..P... ..Q... ..T... ..L... ..K... ..D... ..I... .. ..H... ..F... ..T... ..S... ..I... ..V... ..A... ..K... ..G... ..P... ..M... ..F... ..C... ..G... ..V... ..A... ..T... ..R... ..N... ..L... ..Y... ..T... ..L... ..C... ..K... ..N... ..E... ..N... ..F... ..K... ..E... ..G... ..V... ..T... ..P... ..F... ..Q... ..

```

Fig. 2

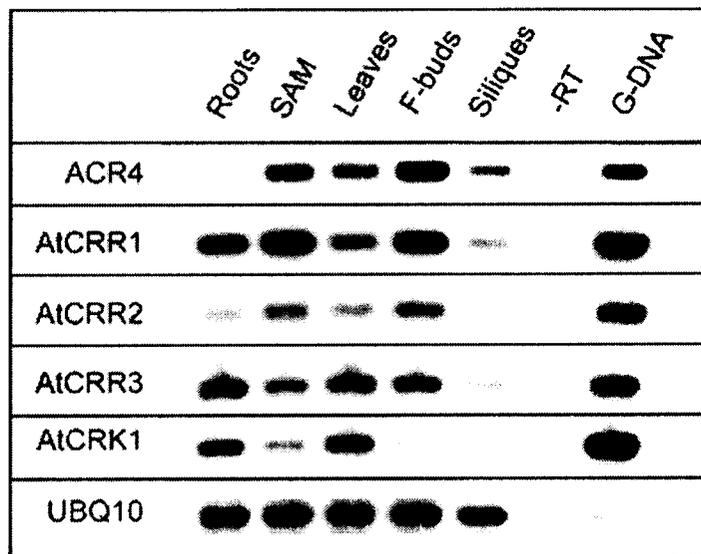


Fig. 3

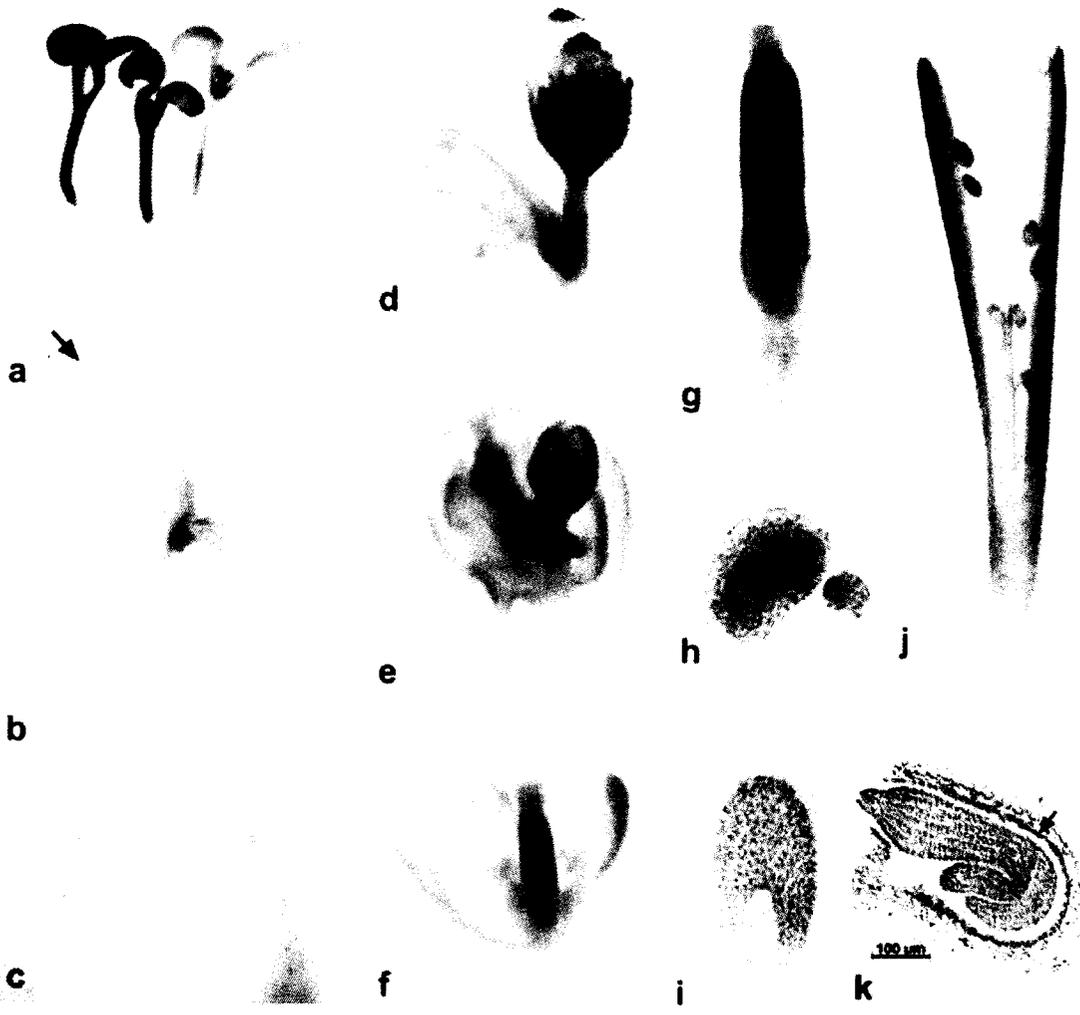


Fig. 4a-k

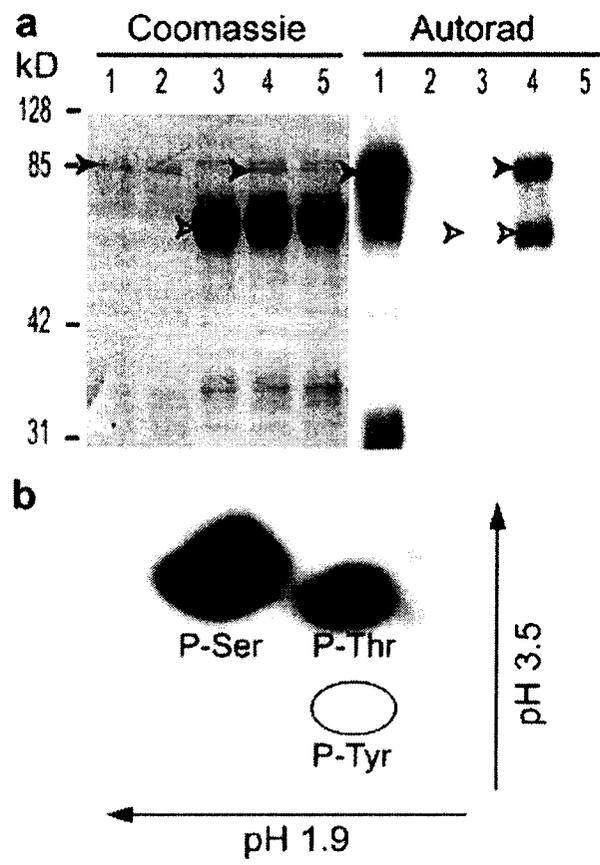
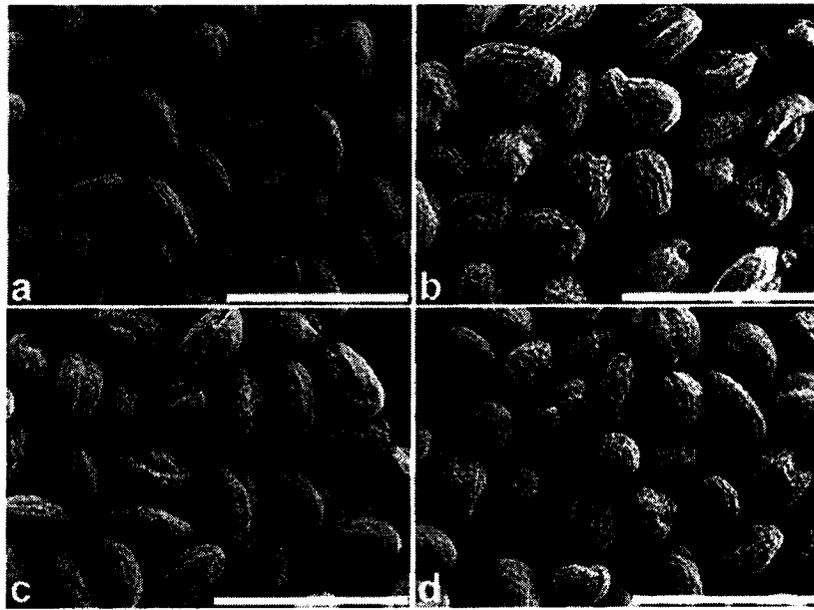


Fig. 5a,b



**Fig. 6a-d**

CHAPTER 3. TRANSCRIPTIONAL REGULATION OF *VIVIPAROUS1* (*VP1*)  
EXPRESSION IN MAIZE

A paper to be submitted to Plant Journal

Xueyuan Cao<sup>1,3</sup>, Donald R. McCarty<sup>4</sup>, Philip W. Beecraft<sup>1,2,3</sup>

<sup>1</sup>Department of Genetics, Development and Cell Biology, Iowa State University, Ames,  
IA 50011, USA

<sup>2</sup>Department of Agronomy, Iowa State University, Ames, IA 50011, USA

<sup>3</sup>Molecular, Cellular, and Developmental Biology Program, Iowa State University, Ames,  
IA 50011, USA

<sup>4</sup>Department of Horticultural Sciences, University of Florida, Gainesville, FL 32611, USA

**Abstract**

*Viviparous1* (*Vp1*), a B3 domain containing transcription factor, is a key regulator of late embryogenesis and seed maturation in *Zea Mays*. Mutant kernels of *vp1* show reduced ABA sensitivity, and cannot accomplish the maturation phase of development and enter dormancy, but germinate on the ear viviparously. The expression of *Vp1* was monitored in cultured embryos at 20 days after pollination (DAP) by semi-quantitative PCR. *Vp1* expression

decreased after 8-48 hours in hormone free liquid MS medium and could be induced by exogenous abscisic acid (ABA), high salinity or osmotic stress. *Vp1* expression could be induced within one hour by exogenous ABA and showed a dosage response to ABA concentrations between 0.1 to 50  $\mu$ M. The sequence analysis of the *Vp1* promoter identified a potential ABRC (ABA response complex), which consists of an ACGT-containing ABRE (ABA response element) and a CE1L (coupling element 1-like) element. Electrophoretic mobility shift assay (EMSA) confirmed that the ABRE and CE1L specifically bound proteins in 20 DAP embryo nuclear protein extracts. Treatment of 20 DAP embryos for 20 hours in hormone free MS medium blocked the specific ABRE-protein interaction, while exogenous ABA or mannitol treatment restored this interaction. The EMSA experiments also revealed a 64-bp fragment (VPP1) and a predicted bZIP910 binding site, which specifically bound to proteins in the 20 DAP embryo nuclear extracts. Transient gene expression assays of the *Vp1* promoter fused to GUS showed that ABA treatment increased the *Vp1* promoter activity while a mutation in the ABRE had a negative effect on GUS expression upon ABA treatment. The EMSA and transient gene expression data suggest that the ABRE is important for the regulation of *Vp1* expression by ABA.

## **Introduction**

Angiosperm embryo development can be divided into three overlapping phases (West and Harada, 1993). Embryo pattern formation occurs in the early morphogenesis phase, which is followed by the maturation phase. During the maturation phase, macromolecular reserves such as storage proteins, lipids and sugars are synthesized and accumulated to a high level.

The tissue growth and development is arrested and the embryos acquire the ability to withstand desiccation and go into dormancy during the final stage of embryogenesis. The mature embryos remain metabolically quiescent until favorable conditions trigger germination. The proper control of dormancy and germination transition is crucial for plants to survive in nature and is also of great agronomic interest (Bewley, 1997; Carrari et al., 2001b).

Plant hormones play crucial regulatory roles during embryogenesis and the dormancy-germination transition. Abscisic acid (ABA) is well known to trigger seed dormancy and repress seed germination, while gibberellic acid functions antagonistically to trigger seed germination (McCarty, 1995; Bewley, 1997; Koornneef et al., 2002). During the maturation phase of seed development, ABA is involved in inducing the expression of many maturation-associated genes such as *Lea* (late embryogenesis abundant) genes (Marcotte et al., 1989). Mutants deficient in ABA production or sensitivity fail to go into dormancy and result in viviparous or precocious germination. This is accompanied by the reduction of maturation-associated gene expression. VP1 also plays an important regulatory role in late embryogenesis and seed maturation. In *Zea Mays*, mutant *vp1* kernels show reduced sensitivity to ABA, fail to acquire desiccation tolerance and enter developmental arrest, and germinate on the ear viviparously (McCarty et al., 1989). In aleurone cells, *vp1* mutants fail to activate the anthocyanin regulatory gene *C1*, resulting in an unpigmented aleurone cell layer (McCarty et al., 1989). *Vp1* encodes a B3 domain containing transcription factor (McCarty et al., 1991). In *Arabidopsis*, *abi3* mutants affect various aspects of seed development. *ABI3* was cloned and shown to be the homologue of maize VP1 (Giraudat et al., 1992). Maize *Vp1* can complement the *Arabidopsis abi3* mutant phenotype and form a

VP1/ABI3 class of transcription factors (Giraudat et al., 1992; Suzuki et al., 2001).

VP1/ABI3 homologues are present in a variety of plant species such as barley (HvVP1) (Hollung et al., 1997), carrot (C-ABI3) (Shiota et al., 1998), *Phaseolus vulgaris* (PvALF) (Bobb et al., 1995), poplar (PtABI3) (Rohde et al., 1998), rice (OsVP1) (Hattori et al., 1995), Sorghum (SbVP1) (Carrari et al., 2001b), yellow-cedar (CnABI3) (Lazarova et al., 2002) and more. This indicates that the regulation through VP1/ABI3 is well conserved and is a general theme in plant kingdom.

Detailed studies of the promoter elements required for the ABA induction have been performed in various plants and genes such as barley *HVA1* (Straub et al., 1994; Shen et al., 1996) and *HVA22* (Shen et al., 1993; Shen and Ho, 1995), maize *Rab17* (Busk et al., 1997) and *Rab28* (Pla et al., 1993; Busk and Pages, 1997), rice *Em (Osem)* (Hattori et al., 1995), and wheat *Em* (Marcotte et al., 1989). These studies revealed a cis-acting DNA element, ABRE (abscisic acid responsive element) with an ACGT core sequence, which is required for the induction of transcription by ABA in transient gene assay experiments. Proteins that bind to ABREs belong to a group of bZIP transcription factors such as wheat EmBP1 (Guiltinan et al., 1990), rice OSBZ8 (Nakagawa et al., 1996) and TRAB1 (Hobo et al., 1999a), and *Arabidopsis* ABI5 (Finkelstein and Lynch, 2000). Hobo et al. (1999a) showed that TRAB1 is a transcriptional activator acting synergistically with ABA and OsVP1, the rice VP1 homologue, to induce *Osem* expression in a protoplast transient expression system. Kagaya (2002) further demonstrated that TRAB1 is phosphorylated in response to ABA and this phosphorylation is essential for TRAB1 to induce the reporter gene expression.

Although a multimerized ACGT-containing ABRE could respond to ABA induction in a heterologous minimal promoter, a single copy of ABRE is not sufficient to respond to ABA

induction. A second cis-acting element called a coupling element (CE) is required to form an ABA responsive complex (ABRC) in a natural promoter (Busk and Pages, 1998; Singh, 1998). Two types of coupling element (CE) were identified: CE1 and CE3 (Shen and Ho, 1995; Shen et al., 1996). With their corresponding ABRE, they were proposed to form ABRC1 and ABRC3 respectively. Further analysis of ABRC3 showed that CE3 and the ACGT-containing ABRE in the *Osem* promoter function equivalently in ABA responsiveness and transactivation by VP1, a B3 domain containing transcription factor, and that both elements could bind to TRAB1 (Hobo et al., 1999b). Hobo et al (1999b) concluded that the ACGT-containing ABRE and CE3 are the same class of cis-acting element and named them G-ABRE. The difference between ABRC1 and ABRC3 is also reflected by the spacing between the ABRE and the coupling element (Shen et al., 2004) and by the transactivation by VP1. ABRC3, such as in maize *rab28* (Busk and Pages, 1997) and in barley *HVA1* (Shen et al., 1996) could be transactivated by VP1, while ABRC1 did not respond to VP1 activation in barley *HVA22* (Shen et al., 1996).

Considerable work has been done studying the mechanisms of VP1 function. Transient expression of *CI*-GUS and *Em*-GUS could be transactivated by co-expression of VP1. A chimeric protein containing the amino-terminal acidic domain of VP1 fused to the GAL4 DNA binding domain could activate the CaMV 35S minimal promoter fused to GAL4-binding sites (McCarty et al., 1991). Structure-function analyses of ABI3 and PvALF in transient expression assays further support that the amino-terminal acidic domain of VP1 functions as transcriptional activator and acts synergistically with ABA (Bobb et al., 1995; Rojas et al., 1999). On the other hand, the functional VP1 was required for the full repression of germination specific  $\alpha$ -amylase expression in developing aleurone cells and the repression

was independent of the amino terminal acidic activation domain, indicating that VP1 also functions as a transcriptional repressor (Hoecker et al., 1995, 1999). These suggest that VP1 functions as both transcriptional activator and repressor during seed development to promote maturation- and repress germination- associated gene expression. VP1 activates downstream gene expression by two distinct mechanisms depending on the target cis-elements. One is through the direct binding to Sph/RV elements via the B3 domain. The activation of the *Cl* gene is through this direct binding mechanism (Suzuki et al., 1997). The activation of ABA inducible genes such as *Em* genes is mediated by indirect interaction with ABREs via bZIP transcription factors such as TRAB1 (Hobo et al., 1999a).

Although much has been done on the mechanisms by which VP1 regulates downstream gene expression, the regulation of *Vp1* itself remains largely unknown. Here, we show that *Vp1* is regulated by ABA, high salinity and osmoticum in mid-maturation phase maize embryogenesis. We also report the identification of four potential transcription factor binding sites and the regulation of *Vp1* expression through an ABRE.

## Results

### *Features of the VP1 promoter*

VP1 plays a critical role in seed maturation and dormancy in maize and was shown to be expressed in mid to late phase embryos (McCarty et al., 1989). To better understand how *Vp1* is regulated in developing maize embryos, we analyzed 958 bp (base pairs) of *Vp1* promoter region isolated from the W22 inbred line. This fragment mimicked the normal *Vp1* expression pattern in a *Vp1::GUS* transgenic plant indicating it contained all the cis

regulatory elements necessary for normal *Vpl* expression (Becraft, unpublished). Like sorghum, the maize *Vpl* promoter lacks a canonical TATA box (Carrari et al., 2001a). A 5'-RACE PCR was performed on mRNA from 20 DAP embryos to locate the transcription start site(s). *Vpl* transcription showed variable start sites within a 30 base region, which is typical for a TATA-less promoter (Sehgal et al., 1988). Nine of 13 cloned RACE PCR fragments identified the same major transcription site, which was set as position +1 (Figure 1). The 958 bases includes 891 bases of 5' promoter sequence and 67 bases of transcribed 5' UTR.

We obtained the *OsVpl* and *SbVpl* promoter sequences from NCBI (<http://www.ncbi.nlm.nih.gov/>). The *Vpl*, *OsVpl* and *SbVpl* promoter sequences were subjected to ConSite analysis to predict cis-regulatory elements (Sandelin et al., 2004). With default searching parameters, 74% conservation cutoff, window size 20 and 80% TF score threshold, five conserved regions were identified, four of which were labeled as PE1 to PE4 (Figure 1). The fifth conserved region was found to be where an ACGT-containing ABRE resides. An AtMYB2 binding site was predicted in PE1 at position +11 to +16 with core sequence: TAACTG (Urao et al., 1993) and a bZIP910 site was predicted in PE2 at position -110 to -116. The *Vpl* promoter sequence was also used to search the PLACE database to identify potential transcription factor binding sites (Higo et al., 1999). An ABRE was also identified at position -77 to -84 with sequence: GCCACGTG, which is similar to the ABRE identified from rice *OsEm* promoter (ACGTGKC) (Hattori et al., 2002). The minus strand of this fragment was also predicted to be E-box with core sequence: CANNTG (Stalberg et al., 1996). Two PAL boxes (putative cis-elements: boxes P, A, L in parsley phenylalanine ammonia-lyase promoter) were predicted at position -517 to -522 (complementary strand) and -546 to -552 respectively with the core sequence: CCGTCC (Logemann et al., 1995). At

position –494 to –505, several potential binding sites were predicted, such as: (CA)<sub>n</sub> element with CNAACAC (Ellerstrom et al., 1996); DPBF binding site with ACACNNG (Kim et al., 1997) and RAV1 binding site with core sequence CAACA (Kagaya et al., 1999). A DRE/LTRE was predicted at the complement strand from –485 to –491 with (R)CCGTC (Baker et al., 1994; Jiang et al., 1996; Dubouzet et al., 2003). Consistent with the ConSite prediction, an AtMYB2 binding site and a bZIP910 site were also predicted in the PLACE analysis. No potential protein-binding site was predicted to reside in the PE3 and PE4 regions, which were identified in the ConSite analysis.

It is noteworthy that there is no obvious binding site between –141 and –441. In rice, there is a predicted open reading frame (ORF) 2.4 Kb (kilo bases) upstream of OsVp1 in the opposite orientation. The predicted PE4 and PAL boxes are 1.4 Kb upstream of the predicted ORF, which is less likely controlled by the PE4 and PAL boxes, suggesting that the predicted PAL boxes and PE4 are likely to be part of the *Vp1* promoter.

*ABA, high salt and osmotic stress induce Vp1 expression in developing embryos*

To better understand how *Vp1* is regulated, we analyzed *Vp1* expression using semi-quantitative RT-PCR on cultured embryos subjected to several physiological conditions. Under our growth conditions, *Vp1* expression reached peak levels at around 20 DAP. 20 DAP embryos were cultured in liquid MS medium with 1% sucrose under continuous shaking in dark (Figure 2a). A reduction in *Vp1* expression became apparent after eight hours and continued to decline to barely detectable levels after 48 hours. By 24 hours, expression was greatly reduced and the embryos were released from dormancy and started to germinate.

As VP1 functions in ABA regulated processes, we checked whether *Vp1* expression could be regulated by ABA. After culturing for 24 hours in hormone free medium, *Vp1* transcripts dropped to a low level. However, the *Vp1* expression level in the embryos cultured with 10  $\mu\text{M}$  ABA or 100  $\mu\text{M}$  ABA was comparable to the expression level in the uncultured fresh embryos, indicating that exogenous ABA could maintain *Vp1* expression (Figure 2b). As 10  $\mu\text{M}$  ABA is sufficient to maintain *Vp1* expression in embryos, this concentration was used in subsequent experiments. It was possible that ABA could maintain *Vp1* expression but could not induce its expression after the expression dropped. To test this possibility, embryos were cultured in hormone free medium for 24 hours to drop *Vp1* expression to a low level, then ABA was added to the medium for various times. One hour of ABA treatment was sufficient to induce *Vp1* expression to levels comparable to that of fresh embryos (Figure 2c). As one-hour ABA treatment was sufficient to induce *Vp1* expression to a high level, the dosage response of *Vp1* expression to ABA was checked after one-hour treatment. *Vp1* expression was slightly induced by 0.1  $\mu\text{M}$  ABA. When treated with higher concentration of ABA, *Vp1* expression increased and reached plateau at 10 to 50  $\mu\text{M}$  (Figure 2d). These data indicate that ABA can induce and maintain *Vp1* expression in maturation phase maize embryos. To address whether the ABA induction of *Vp1* expression requires *de novo* protein synthesis, 20 DAP maize embryos were cultured for 24 hours, then 20  $\mu\text{M}$  cycloheximide, a protein synthesis inhibitor, was added into the medium to inhibit protein synthesis. After 20 minutes of cycloheximide treatments, 10  $\mu\text{M}$  ABA was added and the embryos further cultured for one hour. The cycloheximide treatment did not inhibit the

induction of *Vp1* expression by ABA (Figure 2e), indicating that the ABA induction of *Vp1* expression in cultured embryos does not require *de novo* protein synthesis.

During the maturation phase, the embryos accumulate storage reserves and develop desiccation tolerance. As such, embryos are subjected to a milieu with high concentrations of sugars, metabolites and osmoticum. We hypothesized that *Vp1* expression might be regulated by these physiological conditions and so we checked the effects of high salt, sucrose and mannitol on *Vp1* expression. As shown in Figure 2b, *VP1* expression could be maintained or induced in the embryos cultured in liquid MS medium containing 500 mM NaCl, 20% sucrose and/or 20% mannitol. *Vp1* expression also could be maintained or induced with 200 mM NaCl, 7% sucrose and or 7% mannitol (data not shown). Sucrose could be regulating *Vp1* expression through sugar signaling or osmotic stress, which could not be separated here. However, the effect of mannitol on *Vp1* expression indicates that osmotic stress is capable of inducing *Vp1* expression. These data indicate that salt and osmotic stress can maintain or induce *Vp1* expression in maturation phase embryos. We also checked the effects of exogenous GA3, IAA and ACC (a precursor for ethylene synthesis) on *Vp1* expression. No obvious effects could be detected with these treatments and no interaction of IAA or ACC with ABA was observed (data not shown).

#### *EMSA analysis of the potential ABRC1 in the Vp1 promoter*

Sequence analysis showed that the *Vp1* promoter contains a conserved 8-bp element with the sequence GCCACGTG, which is predicted to be an ABRE. To test whether this fragment is a transcription factor-binding site, a 29 bp oligonucleotide (VPO5, Table 1) containing the putative ABRE was synthesized as a probe for an EMSA experiment. A major retarded band

was observed using nuclear protein extract from 20 DAP embryos, indicating that VPO5 bound to nuclear protein(s) (Figure 3a). The unlabeled VPO5 competed away the retarded band, while an unrelated DNA fragment (VPO4) or poly dI-dC could not, indicating that the binding to the nuclear protein(s) was specific (Figure 3a and 3c). To determine which bases are important for the binding of VPO5 to the nuclear protein(s), serial mutations were synthesized (VPO5M1 to VPO5M7) and used as competitors in an EMSA experiment. The mutations in TGGCCACGTG, which includes the ABRE, impaired the ability to compete away the shifted band, and the CGT to AAG mutation in VPO5M3 completely abolished the competition capacity (Figure 4). Radioactively labeled VPO5M3 did not produce shifted bands (data not shown). This suggests that the 8-bp ABRE is crucial for the *in vitro* binding of VPO5 to nuclear protein(s) from developing embryos.

VP1 is expressed in aleurone cells in maize (McCarty et al., 1989; Becraft unpublished). An EMSA experiment was performed to test whether VPO5 could bind to aleurone nuclear proteins extracted from 20 DAP aleurone cells. As shown in figure 3b, two retarded bands of comparable intensity resulted indicating that VPO5 could bind to nuclear proteins from aleurone. This result differs slightly from that obtained in embryos, where one major retarded band was observed. This suggests that VPO5 might bind to different transcription factor(s) in aleurone cells or bind to additional transcription factor(s) in aleurone cells.

A single copy of ABRE is not sufficient to initiate ABA responsiveness. In natural promoter contexts, ABREs usually function in multiple copies or form ABA response complexes (ABRC) with another element called a coupling element (CE) (Shen et al., 1993; Shen et al., 1996). Examination of the *Vp1* promoter identified a CE1-like core sequence CCACC lying 50 bp downstream of the ABRE. To determine whether this core sequence is a

potential nuclear protein-binding site, an EMSA experiment was performed on 20 DAP embryo nuclear protein extract using a synthesized 28-bp oligonucleotide (CE1L) as a probe. CE1L bound to nuclear protein(s) resulting in a shifted band. The binding was competed away by unlabeled CE1L but not by the unrelated VPO4 oligonucleotide, indicating that the interaction was specific (Figure 5a and b). This is consistent with CE1L functioning as a true CE1. To determine the critical binding sites of CE1L, scanning mutagenesis was conducted and substituted versions of CE1L (CE1LM1-CE1LM8) were used as competitors in an EMSA experiment. Niu et al. showed that ZmABI4, a maize homologue of *Arabidopsis* ABI4, could bind to CE1 in a number of ABA-related genes and *adh1* (Niu et al., 2002). An A to T substitution in CCACC completely abolished the binding capacity of CE1 to ZmABI4. Surprisingly, in our EMSA experiment, this A to T (data not shown) or A to C substitution (CE1LM4) could compete away the retarded band indicating that it still allowed binding. Similarly, CE1LM3, which mutates the first two bases of CCACC could partially compete away the retarded band. Thus, it appears that the CCACC is not crucial for CE1L binding to nuclear protein(s) (Figure 5b). However, other substitutions of CE1L failed to compete away the shifted band, indicating that these bases are important for the protein binding. Taken together, CE1L does not appear to have typical CE1 binding activity. It might function as a new class of coupling element or represent a different regulatory site. The function of CE1L needs further investigation.

#### *Identification of a 64 bp nuclear protein-binding site.*

Several conserved fragments identified in the ConSite and PLACE analyses did not bind to 20 DAP embryo nuclear proteins in EMSA experiments (data now shown). To search for

additional promoter elements, EMSA experiments were performed using 100 bp PCR fragments with 50 bp overlaps to scan throughout the *Vp1* promoter. A 64 bp oligonucleotide fragment (VPP1) was derived and shown to bind to nuclear protein(s) from 20 DAP embryos (Figure 6a). The unlabeled VPP1 competed away the shifted band, while VPO5, VPO4 and poly dI-dC could not, indicating that VPP1 binding was specific (Figure 6a and 6c). This also indicates that VPP1 and VPO5 bind to different nuclear proteins. We also checked whether VPP1 could bind to nuclear protein(s) from 20 DAP aleurone cells. As shown in Figure 5b, a retarded band was obtained.

We performed a series of experiments to try to define the sequences required for nuclear protein interactions. There are two predicted PAL boxes in reverse orientation and a (CA)<sub>n</sub> / DPBF in this 64- bp VPP1 (Figure 1). The oligonucleotide fragments corresponding to PAL boxes or (CA)<sub>n</sub> / DPBF alone did not bind to nuclear protein(s) to produce a shifted band in EMSA experiments (data not shown). The VPP1 fragment was scanned with 12 base substitutions. As shown in figure 7, all the mutant versions of VPP1 lost or partially lost the ability to compete away the shifted band. Furthermore, overlapping 32-mers were synthesized and failed to produce a retarded band (not shown). Thus, it appears that only an intact 64-base VPP1 fragment binds to nuclear proteins and that bases throughout the fragment are required for this binding.

#### *Validation of other predicted elements in ConSite and PLACE analysis*

To further address whether the other predicted elements in ConSite and PLACE analyses have nuclear protein binding activities, oligonucleotides spanning the PAL boxes, (CA)<sub>n</sub>/DPBF, DRE/LTRE, PE3, DRE1 and AtMYB2 were synthesized and used as probes in

EMSA experiments (Table 1). Only the AtMYB2 probe produced an obvious retarded band with 20 DAP embryo nuclear protein(s) (data not shown). This suggests that these other predicted elements are not nuclear protein binding sites under the assay conditions. This does not rule out the possibility that these elements might bind to the nuclear proteins of other tissues such as aleurone cells or of embryos from other developmental stages.

An oligonucleotide including the predicted bZIP910-binding site in PE2 (ZIP1, Table 1) was synthesized and used as the probe in an EMSA experiment. The radioactively labeled ZIP1 bound to nuclear proteins in 20 DAP embryo nuclear protein extract and produced a retarded band (Figure 8). The unlabeled ZIP1 could compete away the retarded band while VPO4, a non-related oligonucleotide, could not, indicating that ZIP1 bound specifically to nuclear protein(s) in 20 DAP embryos.

#### *The ABRE functions in the induction of *Vp1* expression by ABA*

The semi-quantitative RT-PCR showed that exogenous ABA could induce *Vp1* expression. Sequence analyses predicted an ABRE in the *Vp1* promoter and EMSA experiments showed that this element could interact with nuclear proteins from 20 DAP embryos. To investigate whether this ABRE is important for ABA regulation of *Vp1* expression *in vivo*, a transient gene expression experiment was performed by biolistically introducing *VP1* promoter::GUS fusion constructs into 20 DAP embryos. As shown in Figure 9, the wild type *Vp1* promoter (pVD1), containing bases -437 to +67, responded to ABA treatment resulting in a significant increase of GUS expression (p-value = 0.0195), compared to embryos cultured without exogenous ABA. Surprisingly, a 6-bp mutation in the ABRE CACGTG to GGATCC (pVD5) increased the basal level of *GUS* expression in cultured embryos and expression was

suppressed by ABA treatment (p-value = 0.0133). To test whether the mutated ABRE in pVD5 might result in a new *cis* regulatory site, which increased the basal transcription activity in cultured embryos and could be suppressed by ABA treatment, database searches were performed and EMSA was conducted with a mutated oligonucleotide. No predicted transcription factor binding site was created by the substituted bases and no interaction was detected with nuclear proteins from 20 DAP cultured embryos (not shown). Thus the molecular nature of the modulatory activity of this mutation remains unclear. However, this result indicates that the ABRE is important for the regulation of *Vp1* expression by ABA.

*ABA or osmotic stress induce nuclear protein interaction with the ABRE*

The ABA or osmotic stress treatment induced *Vp1* expression and mutation in the ABRE blocked this induction in a transient gene assay (see above). To further explore the basis for the regulation of *Vp1* by ABA and osmoticum, an EMSA experiment was performed on nuclear protein extracts from treated or untreated 20 DAP embryos using VPO5 as the probe. As shown in Figure 10, 20 hours of culture in hormone free medium greatly reduced the binding capacity of the embryo nuclear extract to VPO5, compared with nuclear extract from freshly harvested embryos. Applying 100 uM ABA or 20% mannitol in the medium restored the binding capacity comparable to the fresh embryos. A similar EMSA experiment was also performed on the CE1L. There was no difference observed among all the treatments in terms of the binding capacity (data not shown). This indicates again that the ABRE is crucial for the regulation of *Vp1* expression by ABA or osmotic stress.

## Discussion

During later stages of seed development, plant embryos undergo maturation, acquire desiccation tolerance and achieve dormancy. VP1, a B3 domain containing transcription factor, functions in the regulation of these processes with the plant hormone ABA and other factors. The knowledge of how *Vp1* is regulated in seed development not only helps us to better understand the regulation of seed maturation, but also is of great agronomic interest. Inadequate establishment of dormancy can result in pre-harvest sprouting, which is serious problem in grains such as wheat and sorghum (Maiti et al., 1985; Sharma et al., 1994). We analyzed a *Vp1* promoter region that confers the normal *Vp1* expression pattern in *Vp1::GUS* transgenic plants (Becraft, unpublished). Our data suggest that ABA functions in regulating *Vp1* expression in developing embryos via an ABRC.

In previous work, Carrari et al (2001) isolated the sorghum *Vp1* (*SbVp1*) promoter and showed that the *SbVp1* promoter does not have canonical TATA and CAAT boxes. The maize *Vp1* promoter also lacks these canonical boxes. One common feature of TATA-less promoters is that the transcription start sites are variable (Sehgal et al., 1988; Chow and Knudson, 2005). Using 5'-RACE PCR, we identified a major transcription starting site and several ancillary ones. The sequence comparison in a ConSite analysis (Sandelin et al., 2004) of *Vp1* promoters from maize, rice and sorghum revealed five highly conserved fragments including one with a predicted ABRE. Consistent with the ConSite analyses, a PLACE database (Higo et al., 1999) search of the four conserved regions only predicted a potential binding site in PE1 and PE2. There was no potential binding site predicted in PE3 or PE4, and EMSA experiments did not detect any obvious protein DNA interaction in PE3 and PE4

either (data not shown). The predicted AtMYB2 binding site in PE1 and the bZIP910 binding site in PE2 were confirmed to have nuclear protein binding activities by EMSA experiments. The PLACE analysis showed several drought or cold response elements, which might be relevant considering the physiological conditions that maize embryos confront in the maturation phase. However, the EMSA experiments did not detect any obvious protein binding in these predicted sites. This does not rule out the possibility that there might be some factors binding to these sites under certain developmental and physiological conditions as we only used the 20 DAP fresh embryo nuclear extracts in the EMSA experiments. In the *Arabidopsis ABI3* promoter, there are three ORFs in the 5' UTR, which negatively regulate *ABI3* expression. Removing these ORFs dramatically increased the promoter activity and expanded the expression domain to roots (Ng et al., 2004). We did not find similar ORFs in the 5' UTR of the maize *Vp1* promoter. This might suggest a difference in mechanisms of *Vp/ABI3* regulation between monocots and dicots.

It was shown that maize *Vp1* is expressed from as early as 10 DAP to very late in seed development (McCarty et al., 1989; Niu et al., 2002). Miyoshi et al. (2002) also demonstrated that *OsVp1* and *OsEm* expression is correlated in developing rice embryos as early as 6 DAP. We used 20 DAP embryos to address which physiological factors could regulate *Vp1* expression. *Vp1* transcript levels decline by eight hours of culturing in hormone free medium and can hardly be detected after 48 hours. The cultured embryos start to germinate within 24 hours. This inverse correlation between *Vp1* expression and germination transition is consistent with the function of VP1 to promote seed dormancy and repress germination (McCarty, 1995). This decrease in *Vp1* expression could be reversed by exogenous ABA, which also arrests germination. Thus, *Vp1* mediates and activates the ABA induction of

maturation-associated gene expressions and itself is positively regulated by ABA. Barley *Vp1* mRNA was only slightly increased by ABA compared to untreated embryos and it was concluded that Barley *Vp1* was not significantly modulated by ABA (Hollung et al., 1997). This difference is possibly due to the saturated effect of endogenous ABA on *Vp1* expression.

ABA induced *Vp1* transcript accumulation at concentrations as low as 0.1  $\mu\text{M}$ , indicating that the induction of ABA on *Vp1* expression could occur at physiological levels. This induction does not require *de novo* protein synthesis, as the protein synthesis inhibitor, cycloheximide, did not block the inducibility by ABA. However, cycloheximide alone caused a slight increase in *Vp1* expression, similar to what has been observed in the barley *HVA22* gene, where ABA and cycloheximide have a synergistic effect on expression (Shen et al., 1993). The ABA induction of *Vp1* expression was further confirmed by the transient gene expression assay. The ABA treatment roughly doubled the *Vp1::GUS* expression, indicating that the ABA induced transcript increases occurred at least partly through transcriptional mechanisms. The decrease of *Vp1* expression in cultured embryos could also be rescued by high salt and/or mannitol treatments. Mannitol is a non-penetrating solute that can mimic drought stress. The concentration of ABA was shown to increase when plant tissues are stressed by desiccation or high osmoticum (Close et al., 1989), so it is possible that high salinity and osmotic stress increases *Vp1* expression via ABA mediated processes.

Since ABREs are present in many ABA inducible genes, it is not surprising that a predicted ABRE was found to reside in the *Vp1* promoter (Figure 2). The EMSA and transient gene expression experiments confirmed the prediction and identified the ABRE core CCACGTGT critical for protein binding. Treatment in hormone free medium greatly

reduces the binding capacity of the embryo nuclear extract to VPO5. In contrast, ABA and osmotic stress could restore the ABRE protein interaction. A similar result was also observed in barley, where ABA, salt and mannitol increased the protein binding to an ABRE-containing fragment of *Lea* B19.1 (Hollung et al., 1997). This could be explained by turning over of the potential transcription factor(s) in hormone free culture and blocking the turnover upon ABA or mannitol treatment. Considering that the ABA induction of *Vpl* expression does not require *de novo* protein synthesis, the ABA or mannitol treatments more likely regulate the ABRE binding activity of transcription factors through post-translational modifications, such as phosphorylation by ABA.

The two-fold induction of *Vpl* promoter expression is not as high as in other ABA responsive promoters such as wheat *Em* (Marcotte et al., 1989) or *HVA1* and *HVA22* (Shen and Ho, 1995; Shen et al., 1996; Shen et al., 2004). This difference might be due to the sequence of the ABRE, the context and/or relative position to the coupling element, or to the presence of multiple ABREs or other sequences in the other promoters. The mutation of the ABRE (CACGTG to GGATCC) blocked ABA induction. Surprisingly, the mutated promoter had elevated expression compared to the wild type promoter. This elevation of promoter activity was repressed by ABA. This could be due to the creation of a new site for an ABA repressible binding activity, however no binding activity was detected by EMSA, nor were any predicted sites identified through bioinformatic means. This phenomenon is worthy of further study, and we are currently generating another mutated ABRE site to test functionally in transient expression assays.

In natural promoter contexts, ABREs are usually associated with coupling elements to form ABRC complexes, which specify ABA responsiveness (Shen and Ho, 1995; Shen et al.,

1996). Two kinds of coupling elements have been identified, referred to as CE1 and CE3. In the *Osem* promoter, CE3 was shown to function equivalently and interchangeably with the ACGT-containing ABRE in ABA induction and VP1 activation (Hobo et al., 1999b). They proposed that ACGT-containing ABREs, CE3, and non-ACGT-containing ABREs with a strong half-site of the palindromic G-box sequence should be categorized as a single class of *cis*-elements: G-ABRE. ACGT-containing ABREs could form two distinct types of ABRCs. The ABRC3 type (composed of 2 G-ABREs) can be induced by both VP1 and ABA (Guiltinan et al., 1990; Vasil et al., 1995; Busk et al., 1997). The ABRC1 type found in *HVA22* is distinct and did not respond to VP1 activation (Shen et al., 1996), although VP1 appeared to be required for ABA activation (Casaretto and Ho, 2003). Using transient gene assay, Shen et al. (2004) further demonstrated that the distance between CE3 and ACGT-containing ABRE is crucial for the ABRC3 activity, while the multiple of 10 bases between CE1 and the ABRE is important for ABRC1 function. In our case, we identified a CE1-like *cis*-element (CE1L) with CCACC core residing 50 bp downstream of the ABRE and showed that the CE1L had specific protein binding activity in an EMSA experiment (Figure 7). However, this CE1L is distinct from previously identified CE1s as binding was still observed when the adenine in the core was mutated to thiamine or cytosine. This contrasts with what Niu et al (2002) observed on CE1 of maize ABA-responsive promoters using *E. coli* expressed maize ABI4 protein. It is also noteworthy that the two cytosines on both sides are not symmetrical in respect to the requirement for the protein DNA interaction. The two cytosines on the right side are more important for the DNA protein interaction. Experiments are in progress to test the effect of CE1L on *Vpl* expression.

One interesting and striking observation in the EMSA experiments is the identification of the potential binding site VPP1. VPP1 is a 64-bp fragment that specifically bound to proteins in the embryo and aleurone nuclear extracts (Figure 5). The PLACE analysis of the sequence spanning VPP1 predicted several potential binding sites such as two PAL boxes in opposite orientations, and a (CA)<sub>n</sub>/DPBF site. In parsley, the PAL boxes in phenylalanine ammonia-lyase promoter are necessary but not sufficient for elicitor- or light-mediated PAL gene expression (Logemann et al., 1995). However, no binding activity could be detected on these sites alone. Binding of VPP1 to the nuclear protein(s) appears to require nearly the entire 64 bp fragment intact (Figure 6). This suggests that VPP1 is a complex binding site, possibly requiring several factors to bind to it cooperatively. In human, the virus-inducible enhancer of the interferon- $\beta$  (IFN- $\beta$ ) contains four positive regulatory domains and requires cooperative binding of several transcription factors and HMG I(Y) to form an enhanceosome in order to activate IFN- $\beta$  expression (Kim and Maniatis, 1997). There are also several other examples of cooperative binding of transcription factors to enhancers to form enhanceosomes (see review by Merika and Thanos, 2001). It raises the possibility that VPP1 might form an enhanceosome upon cooperative binding to activate *Vp1* expression during development. In the future, it would be very interesting to determine the function of VPP1 in the regulation of *Vp1* expression and to identify the VPP1 binding proteins. A bZIP910 binding site (ZIP1) was predicted and also shown to bind proteins specifically in an EMSA experiment. In *Antirrhinum*, bZIP910 represents a distinct family of bZIP transcription factors containing a single basic region without other discernable functional domain, and was shown preferentially to bind a hybrid C-box/G-box motif containing ACGT *in vitro* (Martinez-Garcia et al., 1998). It is possible that transcription factors similar to bZIP910 of *Antirrhinum*

bind to ZIP1 and function as adapters to form a complex. It is also possible that some other bZIP transcription factors bind to ZIP1 *in vivo* in maize.

In barley, HvVP1 and HvABI5 were sufficient to transactivate *HvABI5* expression and HvVP1 and HvABI5 together regulated ABA-inducible gene expression during late embryogenesis (Casaretto and Ho, 2003; 2005). HvABI5 itself is ABA inducible (Casaretto and Ho, 2003; 2005) and forms part of a complex feed-forward mechanism for ABA regulation of gene expression (Finkelstein et al., 2002). Our results demonstrate that Vp1 is also ABA regulated, forming another key link in this regulatory system. With the identification of other potential binding sites, this work also provides a framework for further studying how *Vp1* is regulated during seed development.

## **Materials and methods**

### *Plant materials*

The B73/W22 hybrid plants used in *Vp1* expression study and transient expression assay were grown in a green house at 28 °C with a 16 hours day and 8 hours dark cycle. The aleurone cells used in electrophoretic mobility shift assay (EMSA) were collected from *thick\** mutants (unpublished) grown in the field. The embryos used in EMSA were collected from B73/W22 hybrids grown in the field, frozen and stored in –50°C freezer.

### *5' RACE-PCR*

Total RNA was isolated from 2 grams of B73/W22 hybrid embryos 20 days after pollination (DAP) using a hot phenol method (Schmitt et al., 1990). mRNA was prepared using PolyATtract<sup>®</sup> mRNA Isolation System II (Promega) according to the manufacturer's protocol. 5' RACE cDNA was synthesized from 200 ng mRNA using SMART<sup>™</sup> RACE cDNA Amplification Kit (Clontech). PCR amplification was carried out with 2.5 µl 5'-RACE-ready cDNA in a 50 µl reaction system. The amplification products were purified with QIAquick<sup>®</sup> PCR Purification Kit (Qiagen) and cloned into a T-vector. Plasmid DNAs were prepared from thirteen individual clones and sequenced at the Iowa State University DNA facility.

### *Bioinformatic analysis of the Vp1 promoter*

The maize *Vp1* promoter sequence was obtained by sequencing a genomic lambda clone from the W22 inbred line. The *OsVp1* promoter sequence was obtained from NCBI (AP003436: base 71033 ~ 72060 complement strand) and *SbVp1* promoter sequence was also obtained from NCBI (AF317201) (Carrari et al., 2001a). The *Vp1*, *OsVp1* and *SbVp1* promoter sequences were applied to ConSite analysis with 74% conservation cutoff, window size 20 and 80% TF score threshold (Sandelin et al., 2004). The *Vp1* promoter sequence was also applied to PLACE database to search for potential transcription factor binding sites (<http://www.dna.affrc.go.jp/PLACE/>) (Higo et al., 1999).

*Transient gene activity assay*

The embryos of 20 DAP B73/W22 hybrids were excised from the seeds and incubated on MS agar plates (Murashige and Skoog, 1962) with 1% sucrose (all the following MS media contain 1% sucrose). The embryos were co-bombarded with 4 ug of the pVD1 (the wild type *Vp1* promoter driving GUS) or pVD5 (a 6-bp mutation, CACGTG to GGATCC, in the ABRE) plasmid DNA and 0.2 ug of ubi-LUC as an internal standard using a PDS-1000/He Biolistic Particle Delivery System (BioRad). Each sample consisted of twelve embryos. After bombardment, each sample was divided in two and incubated for 20 hours in MS liquid medium or 20 hours in MS liquid medium with 100 uM ABA. The samples were frozen in liquid nitrogen after the treatments.

The co-bombarded samples were homogenized in 250 ul 1X CCLR (Cell Culture Lysis Reagent) (Promega) and centrifuged for 10 min at 10,000 g at 4 °C. A luciferase assay was performed on 20 ul of the supernatant with the Luciferase Assay System (Promega) according to the manufacturer's instructions and read on a Sirius Luminometer (Berthold Detection Systems GmbH). Fluorometric MUG (4-methylumbelliferyl-b-D-glucuronide) assays were carried out as follows: 100 ul protein extracts were added to pre-incubated 600 ul 1mM MUG in 1X CCLR with 20% methanol; 100 ul reaction mixtures were added into 900 ul 0.2M Na<sub>2</sub>CO<sub>3</sub> at various time points. The 4MU were read on a Synergy HT plate reader (Bio-Tek inc.).

*Electrophoretic mobility shift assay*

The nuclear protein extracts were prepared from 20 DAP embryos or aleurones by isolation of nuclei using a glycerol based method (Dorweiler et al., 2000) followed by isolation nuclear proteins in protein isolation buffer (10mM Tris-HCl, 1 mM DTT, 0.4 M NaCl (pH 7.5)) (Escobar et al., 2001). DNA probes were generated by annealing the complementary oligonucleotides (synthesized at the Iowa State University DNA Facility) and by filling in the single-strand overhangs with  $\alpha$ -[<sup>32</sup>P]dCTP (PerkinElmer) using the Klenow fragment of DNA polymerase I (Promega) or by polymerase chain reaction (PCR) amplification (Sambrook et al., 1989). The labeled probes were purified with QIAquick® Nucleotide removal kit or PCR Purification Kit (Qiagen).

The protein/DNA binding reactions were performed in a 25 ul reaction system containing 20ng/100bp PCR-labeled probes or 0.8 picomole oligonucleotides, 1 ug sonicated pBlueScript II SK (Stratagene) and 10  $\mu$ g nuclear proteins in 1x binding buffer (10 mM Tris-HCl pH 7.4, 50mM NaCl, 2 mM EDTA, 2.5 mM DTT, 1.25  $\mu$ g BSA, 0.05% v/v NP-40, 10% glycerol). The binding reactions were incubated at room temperature for 20 mins and separated on a native 4.5% polyacrylamide gel. Electrophoresis was carried out at 70 V constant voltages for 4 hours with 0.5x TBE (45 mM Tris-borate, 0.5 mM EDTA pH 8.0) buffer in a 4°C cold room. The gels were vacuum dried and autoradiographed. In the competition assays, the nuclear protein extract was incubated with unlabeled oligonucleotides in 1x binding buffer on ice for 10 mins prior to adding the radioactive probe and continuing with incubation for a further 20 mins at room temperature.

*Expression studies by semi-quantitative RT-PCR analysis*

The embryos, collected from B73/W22 hybrids at 20 DAP, were cultured in petri dishes with 30 ml MS liquid medium containing 1% sucrose and corresponding treatments under continuous gentle shaking in dark. Total RNA was isolated by grinding tissues in liquid nitrogen followed by extraction with TRIzol reagent (Invitrogen) and purified with RNeasy Mini kit (Qiagen) according to the manufacture's instructions. First strand cDNA was synthesized using 3 µg total RNA. Reverse transcription was performed using the SuperscriptII First Strand Synthesis System for RT-PCR (Invitrogen) with Oligo(dT) and 10 picomole gene specific primer of maize 18S ribosomal RNA gene (ZMU42796). PCR was performed using Taq DNA polymerase (Promega) with gene specific primers: *Vp1*, CAAGAGCAAGGCAGTGGTTCCAG (Forward) and CAAATTTAGCGTCACACAGCGGGTAG (Reverse); 18S rRNA: AGCAGCCGCGGTAATTCCAGCTC (Forward) and CTGGTGGTGCCCTTCCGTCAATTC (Reverse). For *Vp1*, the amplification conditions were 94 °C for 2 mins followed by 25 cycles of 94 °C for 30 s, 58 °C for 30 s, and 72 °C for 1 min. For 18S rRNA, the amplification conditions were 94 °C for 2 mins followed by 22 cycles of 94 °C for 30 s, 58 °C for 30 s, and 72 °C for 1 min. PCR products were run in a 0.8% agarose gel, blotted to Duralon-UV membrane (Stratagene), hybridized with the corresponding gene-specific probes, and autoradiographed.

**Acknowledgments**

The authors thank the Becraft's lab members for critical reading of the manuscript.

**References**

- Baker, S.S., Wilhelm, K.S., and Thomashow, M.F.** (1994). The 5'-region of *Arabidopsis thaliana* cor15a has cis-acting elements that confer cold-, drought- and ABA-regulated gene expression. *Plant Mol Biol* **24**, 701-713.
- Bewley, J.D.** (1997). Seed Germination and Dormancy. *Plant Cell* **9**, 1055-1066.
- Bobb, A.J., Eiben, H.G., and Bustos, M.M.** (1995). PvAlf, an embryo-specific acidic transcriptional activator enhances gene expression from phaseolin and phytohemagglutinin promoters. *Plant J* **8**, 331-343.
- Busk, P.K., and Pages, M.** (1997). Protein binding to the abscisic acid-responsive element is independent of VIVIPAROUS1 in vivo. *Plant Cell* **9**, 2261-2270.
- Busk, P.K., and Pages, M.** (1998). Regulation of abscisic acid-induced transcription. *Plant Mol Biol* **37**, 425-435.
- Busk, P.K., Jensen, A.B., and Pages, M.** (1997). Regulatory elements in vivo in the promoter of the abscisic acid responsive gene rab17 from maize. *Plant J* **11**, 1285-1295.
- Carrari, F., Frankel, N., Lijavetzky, D., Benech-Arnold, R., Sanchez, R., and Iusem, N.D.** (2001a). The TATA-less promoter of VP1, a plant gene controlling seed germination. *DNA Seq* **12**, 107-114.
- Carrari, F., Perez-Flore, L., Lijavetzky, D., Enciso, S., Sanchez, R., Benech-Arnold, R., and Iusem, N.** (2001b). Cloning and expression of a sorghum gene with homology to maize vp1. Its potential involvement in pre-harvest sprouting resistance. *Plant Mol Biol* **45**, 631-640.

- Casaretto, J., and Ho, T.H.** (2003). The transcription factors HvABI5 and HvVP1 are required for the abscisic acid induction of gene expression in barley aleurone cells. *Plant Cell* **15**, 271-284.
- Casaretto, J.A., and Ho, T.H.** (2005). Transcriptional regulation by abscisic acid in barley (*Hordeum vulgare* L.) seeds involves autoregulation of the transcription factor HvABI5. *Plant Mol Biol* **57**, 21-34.
- Chow, G., and Knudson, W.** (2005). Characterization of promoter elements of the human HYAL-2 gene. *J Biol Chem* **280**, 26904-26912.
- Close, T.J., Kortt, A.A., and Chandler, P.M.** (1989). A cDNA-based comparison of dehydration-induced proteins (dehydrins) in barley and corn. *Plant Mol Biol* **13**, 95-108.
- Dorweiler, J.E., Carey, C.C., Kubo, K.M., Hollick, J.B., Kermicle, J.L., and Chandler, V.L.** (2000). mediator of paramutation1 is required for establishment and maintenance of paramutation at multiple maize loci. *Plant Cell* **12**, 2101-2118.
- Dubouzet, J.G., Sakuma, Y., Ito, Y., Kasuga, M., Dubouzet, E.G., Miura, S., Seki, M., Shinozaki, K., and Yamaguchi-Shinozaki, K.** (2003). OsDREB genes in rice, *Oryza sativa* L., encode transcription activators that function in drought-, high-salt- and cold-responsive gene expression. *Plant J* **33**, 751-763.
- Ellerstrom, M., Stalberg, K., Ezcurra, I., and Rask, L.** (1996). Functional dissection of a napin gene promoter: identification of promoter elements required for embryo and endosperm-specific transcription. *Plant Mol Biol* **32**, 1019-1027.
- Escobar, C., Aristizábal, F., Navas, A., Campo, F.F.D., and Fenoll, C.** (2001). Isolation of Active DNA-Binding Nuclear Proteins From Tomato Galls Induced by Root-Knot Nematodes. *Plant Mol Biol Rep.* **19**, 375a-375h.

- Finkelstein, R.R., and Lynch, T.J.** (2000). The Arabidopsis abscisic acid response gene ABI5 encodes a basic leucine zipper transcription factor. *Plant Cell* **12**, 599-609.
- Finkelstein, R.R., Gampala, S.S., and Rock, C.D.** (2002). Abscisic acid signaling in seeds and seedlings. *Plant Cell* **14 Suppl**, S15-45.
- Giraudat, J., Hauge, B.M., Valon, C., Smalle, J., Parcy, F., and Goodman, H.M.** (1992). Isolation of the Arabidopsis ABI3 gene by positional cloning. *Plant Cell* **4**, 1251-1261.
- Guiltinan, M.J., Marcotte, W.R., Jr., and Quatrano, R.S.** (1990). A plant leucine zipper protein that recognizes an abscisic acid response element. *Science* **250**, 267-271.
- Hattori, T., Terada, T., and Hamasuna, S.** (1995). Regulation of the Osem gene by abscisic acid and the transcriptional activator VP1: analysis of cis-acting promoter elements required for regulation by abscisic acid and VP1. *Plant J* **7**, 913-925.
- Hattori, T., Totsuka, M., Hobo, T., Kagaya, Y., and Yamamoto-Toyoda, A.** (2002). Experimentally determined sequence requirement of ACGT-containing abscisic acid response element. *Plant Cell Physiol* **43**, 136-140.
- Higo, K., Ugawa, Y., Iwamoto, M., and Korenaga, T.** (1999). Plant cis-acting regulatory DNA elements (PLACE) database: 1999. *Nucleic Acids Res* **27**, 297-300.
- Hobo, T., Kowyama, Y., and Hattori, T.** (1999a). A bZIP factor, TRAB1, interacts with VP1 and mediates abscisic acid-induced transcription. *Proc Natl Acad Sci U S A* **96**, 15348-15353.
- Hobo, T., Asada, M., Kowyama, Y., and Hattori, T.** (1999b). ACGT-containing abscisic acid response element (ABRE) and coupling element 3 (CE3) are functionally equivalent. *Plant J* **19**, 679-689.

- Hoecker, U., Vasil, I.K., and McCarty, D.R.** (1995). Integrated control of seed maturation and germination programs by activator and repressor functions of Viviparous-1 of maize. *Genes Dev* **9**, 2459-2469.
- Hoecker, U., Vasil, I.K., and McCarty, D.R.** (1999). Signaling from the embryo conditions Vp1-mediated repression of alpha-amylase genes in the aleurone of developing maize seeds. *Plant J* **19**, 371-377.
- Hollung, K., Espelund, M., Schou, K., and Jakobsen, K.S.** (1997). Developmental, stress and ABA modulation of mRNA levels for bZip transcription factors and Vp1 in barley embryos and embryo-derived suspension cultures. *Plant Mol Biol* **35**, 561-571.
- Jiang, C., Iu, B., and Singh, J.** (1996). Requirement of a CCGAC cis-acting element for cold induction of the BN115 gene from winter Brassica napus. *Plant Mol Biol* **30**, 679-684.
- Kagaya, Y., Ohmiya, K., and Hattori, T.** (1999). RAV1, a novel DNA-binding protein, binds to bipartite recognition sequence through two distinct DNA-binding domains uniquely found in higher plants. *Nucleic Acids Res* **27**, 470-478.
- Kim, S.Y., Chung, H.J., and Thomas, T.L.** (1997). Isolation of a novel class of bZIP transcription factors that interact with ABA-responsive and embryo-specification elements in the Dc3 promoter using a modified yeast one-hybrid system. *Plant J* **11**, 1237-1251.
- Kim, T.K., and Maniatis, T.** (1997). The mechanism of transcriptional synergy of an in vitro assembled interferon-beta enhanceosome. *Mol Cell* **1**, 119-129.
- Koornneef, M., Bentsink, L., and Hilhorst, H.** (2002). Seed dormancy and germination. *Curr Opin Plant Biol* **5**, 33-36.

- Lazarova, G., Zeng, Y., and Kermode, A.R.** (2002). Cloning and expression of an ABSCISIC ACID-INSENSITIVE 3 (ABI3) gene homologue of yellow-cedar (*Chamaecyparis nootkatensis*). *J Exp Bot* **53**, 1219-1221.
- Logemann, E., Parniske, M., and Hahlbrock, K.** (1995). Modes of expression and common structural features of the complete phenylalanine ammonia-lyase gene family in parsley. *Proc Natl Acad Sci U S A* **92**, 5905-5909.
- Maiti, R.J., Raju, P.S., and Bindinger, R.R.** (1985). Studies on germinability and some aspects of pre-harvest physiology of sorghum grain. *Seed Sci. Technol.* **13**: 27–35. *Seed Sci. Technol.* **13**: 27–35. **13**, 27–35.
- Marcotte, W.R., Jr., Russell, S.H., and Quatrano, R.S.** (1989). Abscisic acid-responsive sequences from the em gene of wheat. *Plant Cell* **1**, 969-976.
- Martinez-Garcia, J.F., Moyano, E., Alcocer, M.J., and Martin, C.** (1998). Two bZIP proteins from *Antirrhinum* flowers preferentially bind a hybrid C-box/G-box motif and help to define a new sub-family of bZIP transcription factors. *Plant J* **13**, 489-505.
- McCarty, D.R.** (1995). Genetic control and integration of maturation and germination pathways in seed development. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **46**, 71-93.
- McCarty, D.R., Carson, C.B., Stinard, P.S., and Robertson, D.S.** (1989). Molecular Analysis of viviparous-1: An Abscisic Acid-Insensitive Mutant of Maize. *Plant Cell* **1**, 523-532.
- McCarty, D.R., Hattori, T., Carson, C.B., Vasil, V., Lazar, M., and Vasil, I.K.** (1991). The Viviparous-1 developmental gene of maize encodes a novel transcriptional activator. *Cell* **66**, 895-905.
- Merika, M., and Thanos, D.** (2001). Enhanceosomes. *Curr Opin Genet Dev* **11**, 205-208.

- Murashige, T., and Skoog, F.** (1962). A revised medium for rapid growth and bioassays with tobacco tissue culture. *Physiol Plant* **15**, 473-497.
- Nakagawa, H., Ohmiya, K., and Hattori, T.** (1996). A rice bZIP protein, designated OSBZ8, is rapidly induced by abscisic acid. *Plant J* **9**, 217-227.
- Ng, D.W., Chandrasekharan, M.B., and Hall, T.C.** (2004). The 5' UTR negatively regulates quantitative and spatial expression from the ABI3 promoter. *Plant Mol Biol* **54**, 25-38.
- Niu, X., Helentjaris, T., and Bate, N.J.** (2002). Maize ABI4 binds coupling element1 in abscisic acid and sugar response genes. *Plant Cell* **14**, 2565-2575.
- Pla, M., Vilardell, J., Gultinan, M.J., Marcotte, W.R., Niogret, M.F., Quatrano, R.S., and Pages, M.** (1993). The cis-regulatory element CCACGTGG is involved in ABA and water-stress responses of the maize gene *rab28*. *Plant Mol Biol* **21**, 259-266.
- Rohde, A., Ardiles-Diaz, W., Van Montagu, M., and Boerjan, W.** (1998). Gene note. Isolation and expression analysis of an ABSCISIC ACID-INSENSITIVE 3 (ABI3) homologue from *Populus trichocarpa*. *Journal of Experimental Botany* **49**, 1059-1060.
- Rojas, A., Almoguera, C., and Jordano, J.** (1999). Transcriptional activation of a heat shock gene promoter in sunflower embryos: synergism between ABI3 and heat shock factors. *Plant J* **20**, 601-610.
- Sambrook, J., Fritsch, E.F., and Maniatis, T.** (1989). *Molecular Cloning: A Laboratory Manual*.
- Sandelin, A., Wasserman, W.W., and Lenhard, B.** (2004). ConSite: web-based prediction of regulatory elements using cross-species comparison. *Nucleic Acids Res* **32**, W249-252.

- Schmitt, M.E., Brown, T.A., and Trumppower, B.L.** (1990). A rapid and simple method for preparation of RNA from *Saccharomyces cerevisiae*. *Nucleic Acids Res* **18**, 3091-3092.
- Sehgal, A., Patil, N., and Chao, M.** (1988). A constitutive promoter directs expression of the nerve growth factor receptor gene. *Mol Cell Biol* **8**, 3160-3167.
- Sharma, S.K., Dhaliwal, H.S., Multani, D.S., and Bains, S.S.** (1994). Inheritance of preharvest sprouting tolerance in *Triticum aestivum* and its transfer to an amber-grained cultivar. *J Hered* **85**, 312–314.
- Shen, Q., and Ho, T.H.** (1995). Functional dissection of an abscisic acid (ABA)-inducible gene reveals two independent ABA-responsive complexes each containing a G-box and a novel cis-acting element. *Plant Cell* **7**, 295-307.
- Shen, Q., Uknes, S.J., and Ho, T.H.** (1993). Hormone response complex in a novel abscisic acid and cycloheximide-inducible barley gene. *J Biol Chem* **268**, 23652-23660.
- Shen, Q., Zhang, P., and Ho, T.H.** (1996). Modular nature of abscisic acid (ABA) response complexes: composite promoter units that are necessary and sufficient for ABA induction of gene expression in barley. *Plant Cell* **8**, 1107-1119.
- Shen, Q.J., Casaretto, J.A., Zhang, P., and Ho, T.H.** (2004). Functional definition of ABA-response complexes: the promoter units necessary and sufficient for ABA induction of gene expression in barley (*Hordeum vulgare* L.). *Plant Mol Biol* **54**, 111-124.
- Shiota, H., Satoh, R., Watabe, K., Harada, H., and Kamada, H.** (1998). C-ABI3, the carrot homologue of the Arabidopsis ABI3, is expressed during both zygotic and somatic embryogenesis and functions in the regulation of embryo-specific ABA-inducible genes. *Plant Cell Physiol* **39**, 1184-1193.

- Singh, K.B.** (1998). Transcriptional regulation in plants: the importance of combinatorial control. *Plant Physiol* **118**, 1111-1120.
- Stalberg, K., Ellerstom, M., Ezcurrea, I., Ablov, S., and Rask, L.** (1996). Disruption of an overlapping E-box/ABRE motif abolished high transcription of the napA storage-protein promoter in transgenic *Brassica napus* seeds. *Planta* **199**, 515-519.
- Straub, P.F., Shen, Q., and Ho, T.D.** (1994). Structure and promoter analysis of an ABA- and stress-regulated barley gene, HVA1. *Plant Mol Biol* **26**, 617-630.
- Suzuki, M., Kao, C.Y., and McCarty, D.R.** (1997). The conserved B3 domain of VIVIPAROUS1 has a cooperative DNA binding activity. *Plant Cell* **9**, 799-807.
- Suzuki, M., Kao, C.Y., Cocciolone, S., and McCarty, D.R.** (2001). Maize VP1 complements *Arabidopsis* *abi3* and confers a novel ABA/auxin interaction in roots. *Plant J* **28**, 409-418.
- Urao, T., Yamaguchi-Shinozaki, K., Urao, S., and 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.
- Vasil, V., Marcotte, W.R., Jr., Rosenkrans, L., Cocciolone, S.M., Vasil, I.K., Quatrano, R.S., and McCarty, D.R.** (1995). Overlap of Viviparous1 (VP1) and abscisic acid response elements in the Em promoter: G-box elements are sufficient but not necessary for VP1 transactivation. *Plant Cell* **7**, 1511-1518.
- West, M., and Harada, J.J.** (1993). Embryogenesis in Higher Plants: An Overview. *Plant Cell* **5**, 1361-1369.

## Figure legends

**Fig. 1** The predicted features of the *Vp1* promoter. The *Vp1* and *OsVP1* promoter sequences were subjected to ConSite analysis and the *Vp1* promoter sequence was also analyzed by a PLACE database search. The conserved regions between *Vp1* and *OsVP1* are gray shaded and labeled as PE1 to PE4 and ABRE. The predicted binding sites by PLACE search are in bold italics. The DNA sequences used in EMSA are underlined.

**Fig. 2** *Vp1* expression is regulated by ABA and stress signals. Semi-quantitative RT-PCR was performed to study the expression of *Vp1*. All treatments were carried out in the dark and all MS liquid medium contained 1% sucrose, unless otherwise indicated. **a** The expression of *Vp1* decreases during hormone free culture. B73/W22 hybrid embryos were cultured in MS liquid medium for 1 hour, 4 hours, 8 hours, 24 hours, 48 hours respectively. **b** *Vp1* expression is maintained or induced by ABA, salt, and osmotic stress. B73/W22 hybrid embryos were cultured for 24 hours in MS liquid medium or in MS liquid medium with 500 mM NaCl, 10  $\mu$ M ABA, 100  $\mu$ M ABA, 20% sucrose or 20% mannitol in the dark or in MS liquid medium under light. **c** *Vp1* expression shows a dosage response to ABA. B73/W22 hybrid embryos were cultured in MS liquid medium or in MS liquid medium with 0.075% ethanol, 0.1  $\mu$ M ABA, 1  $\mu$ M ABA, 5  $\mu$ M ABA, 10  $\mu$ M ABA and 50  $\mu$ M ABA. **d** *Vp1* expression can respond to ABA treatment within one hour. B73/W22 hybrid embryos were cultured in MS liquid medium for 24 hours, then 10  $\mu$ M ABA was added and treated for 1 hour, 2 hours, 4 hours, 8 hours or 12 hours. **e** Induction of *Vp1* expression by ABA does not require de novo protein synthesis. B73/W22 hybrid embryos were cultured in MS medium

for 24 hours followed by adding 20  $\mu\text{M}$  cycloheximide (CYH). After 20 minutes, one sample was treated with 10  $\mu\text{M}$  ABA and all samples were incubated for another one hour.

untreated: freshly harvested embryos, hr: hour;

**Figure 3** VPO5 specifically binds to embryo and aleurone nuclear protein(s). Each DNA-protein binding reaction contained 0.4 picomole  $^{32}\text{P}$  labeled VPO5 with or without 10  $\mu\text{g}$  nuclear proteins. **a** VPO5 bound to nuclear protein(s) from 20 DAP embryos causing one major retarded band. The retarded band could be competed away with increasing unlabeled VPO5. **b** VPO5 bound to nuclear protein(s) from 20 DAP aleurone cells resulting in two retarded bands, both of which could be competed away by unlabeled VPO5. **c** VPO5 bound specifically to nuclear proteins in 20 DAP embryos. The retarded band could be competed away by unlabeled VPO5, but could not be competed away by VPO4 (an unrelated oligonucleotide) and poly dI-dC.  $\blacktriangleleft$  : retarded band;  $\blacktriangleleft$  : free probe.

**Figure 4.** Mutation scanning analyses identified the critical binding sites for VPO5. To identify the critical binding sites of VPO5, three-base substitutions were introduced and used as unlabeled competitor in an EMSA experiment on 20 DAP embryo nuclear proteins. The substitutions in the GCCACGTG lost or partially lost the binding activity, failing to compete away the retarded band, especially VPO5M3. Each DNA-protein binding reaction contained 0.4 picomole  $^{32}\text{P}$  labeled VPO5 with or without 10  $\mu\text{g}$  embryo nuclear proteins. The unlabeled competitors were added as 10 times of the labeled VPO5.  $\blacktriangleleft$  : retarded band;  $\blacktriangleleft$  : free probe.

**Figure 5.** CE1L bound specifically to embryo nuclear protein(s). **a** CE1L bound to 20 DAP embryo nuclear proteins causing a retarded band, which could be competed away by increasing unlabeled CE1L. **b** Mutation analysis of CE1L binding site. One, two or three base substitutions of CE1L were used as unlabeled competitors in EMSA. Mutations in the left side of the canonical CE1 core binding site CCACC (CE1LM3 and CE1LM4) competed away the retarded band, indicating they retained their binding activity. Each lane contained 0.4 picomole  $^{32}\text{P}$  labeled CE1L with or without 10  $\mu\text{g}$  embryo nuclear proteins. The unlabeled competitors were added as 25 times of radioactively labeled CE1L in **b**.  $\blacktriangleleft$ : retarded band;  $\blacktriangleleft$ : free probes.

**Figure 6.** VPP1 specifically bound to nuclear protein(s). Each DNA-protein binding reaction contained 12ng  $^{32}\text{P}$  labeled VPP1 with or without 10  $\mu\text{g}$  nuclear proteins. **a** VPP1 bound to nuclear protein(s) from 20 DAP embryos causing a retarded band. The retarded band was competed away by increasing unlabeled VPP1. **b** VPP1 bound to nuclear protein(s) from 20 DAP aleurone cells resulting in a retarded band, which could be competed away by unlabeled VPP1. **c** VPP1 bound specifically to nuclear proteins from 20 DAP embryos. The retarded band could be competed away by unlabeled VPP1, but not by unlabeled VPO5, VPO4 or poly dI-dC.  $\blacktriangleleft$ : retarded band;  $\blacktriangleleft$ : free probe.

**Figure 7.** Mutation analysis of the VPP1 binding site. To identify the critical binding sites of VPP1, twelve-base substitutions were introduced and used as unlabeled competitors in an EMSA experiment on 20 DAP embryo nuclear proteins. All substitutions lost or partially lost the binding activity, failing to compete away the retarded band, especially VPP1M3 and

VPP1M4. Each lane contained 0.4 picomole  $^{32}\text{P}$  labeled VPP1 with or without 10  $\mu\text{g}$  embryo nuclear proteins. The unlabeled competitors were added as 10 times that of radioactively labeled VPP1.  $\blacktriangleleft$ : retarded band;  $\blacktriangleleft$ : free probes.

**Figure 8.** ZIP1 specifically bound to embryo nuclear proteins. The radioactively labeled ZIP1 could bind to 20 DAP embryo nuclear protein(s), resulting in a retarded band. The unlabeled ZIP1 competitor competed away the shifted band, while VPO4 did not. Each lane contained 0.4 picomole  $^{32}\text{P}$  labeled ZIP1 with or without 10  $\mu\text{g}$  embryo nuclear proteins.  $\blacktriangleleft$ : retarded band;  $\blacktriangleleft$ : free probes.

**Figure 9.** The ABRE is crucial for ABA induction of *VPI* expression. 4  $\mu\text{g}$  pVD1 (wild type) or pVD5 (6 base mutation in ABRE: CACGTG to GGATCC) were co-bombarded with 0.2  $\mu\text{g}$  *ubi::LUC* into 20 DAP embryos, which were treated with or without ABA for 20 hours afterwards. The GUS and luciferase activity were measured separately. The relative ratios were obtained by dividing GUS activity by Luciferase activity. The average and standard deviation of 8 replicates are reported.

**Figure 10.** VPO5 (containing ABRE) responds to ABA and mannitol treatment in an EMSA. Nuclear extracts from 20 DAP embryos treated in MS liquid medium, MS liquid medium with 100 $\mu\text{M}$  ABA or 20% mannitol for 20 hours were applied to an EMSA with fresh embryo nuclear extract as a positive control. The MS treatment abolished the retarded band, while ABA or mannitol treatment restored the retarded band to the level of fresh embryos.

Each DNA-protein binding reaction contained 0.4 picomole  $^{32}\text{P}$  labeled VPO5.  $\blacktriangleleft$ : retarded band;  $\blacktriangleleft$ : free probes.

Table 1. The oligonucleotides used in EMSA experiments.

Name	Oligonucleotide sequence
(CA) <sub>n</sub> /DPBF	AGGAATCGTCCCAACACTCGGGATTTCAGT
AtMYB2	CTCCTCAGTAACTGCACCCACCTCCACACCTA
CE1L	TTCTCCCTCTCCACCTCCTGCTTGCACT
DRE/LTRE	TCGGGATTTCAGTCGGCGCGCCGGC
DRE1	CAGCCGCATCTCGGTTCCCTCCTC
PAL Boxes	GCTTCCGTCCGTGGGATCACGCAGAGACTGCATCGGACGGTCAGC
PE3	GGCGACACGGACATCCTACGGCGCC
VPO4	CCGGCGACACGGACATCCTACGGCGCCGAAGCAGC
VPO5	TCGGTCGCCACGTGTGTCACGCAGCCGCA
VPP1	GAGCTTCCGTCCGTGGGATCACGCAGAGACTGCATCGGACGGTCA
	GGAATCGTCCCAACACTCG
ZIP1	AGGAGAGGCTGGCGTGGGCGGTGG



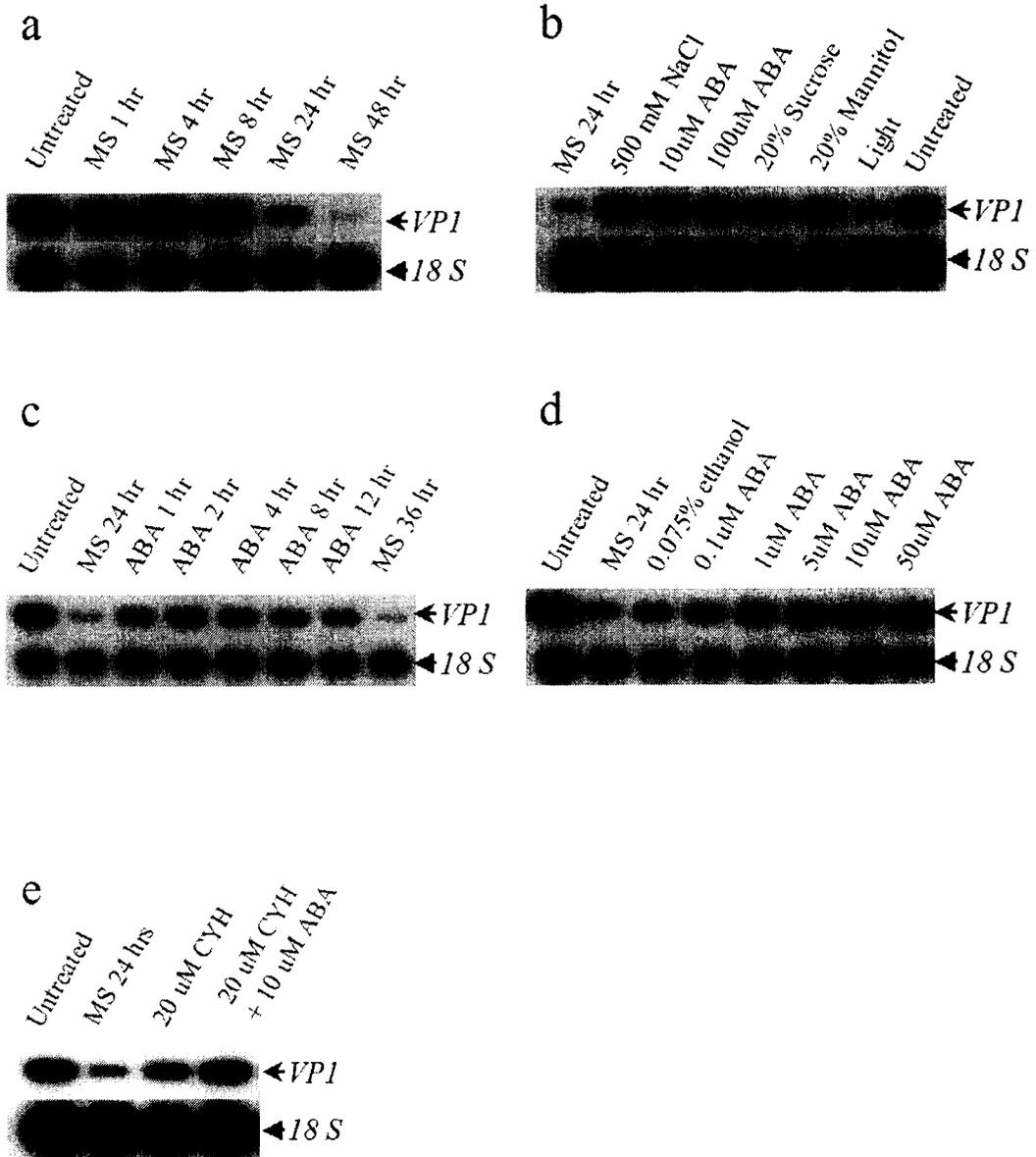
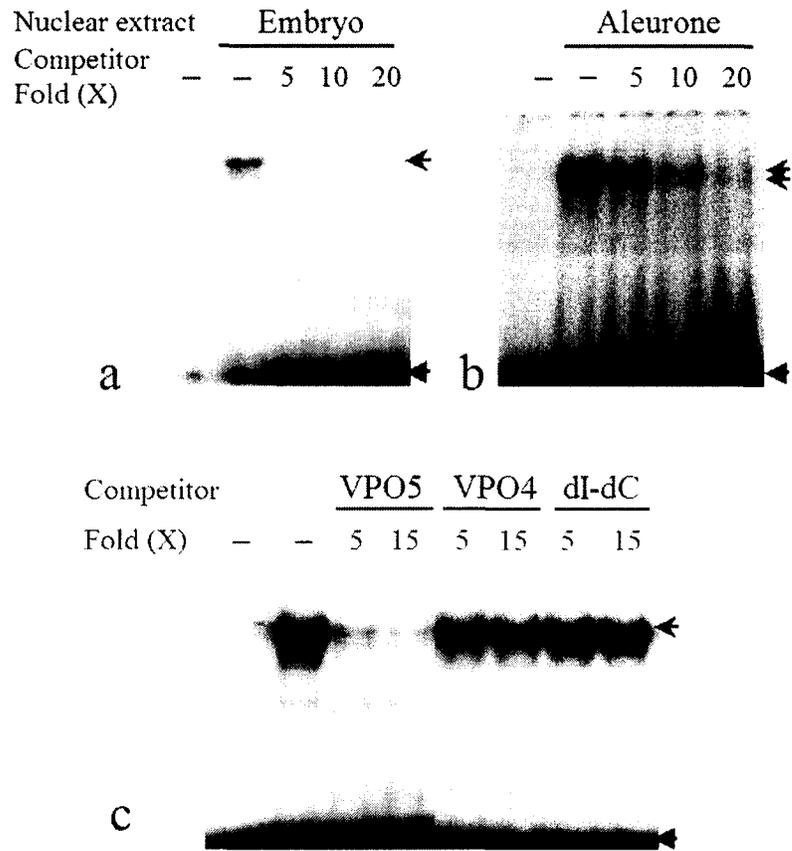


Figure 2



**Figure 3**

```

          *****
VPO5      TCGGTCGCCCACGTGTGTGCACGCAGCCGCA
VPO5M1    .....GAA.....
VPO5M2    .....AAG.....
VPO5M3    .....AAG.....
VPO5M4    .....AGA.....
VPO5M5    .....GAG.....
VPO5M6    .....AAA.....
VPO5M7    .....GAA.....
    
```

Red: ABRE, Underlined: EBOX

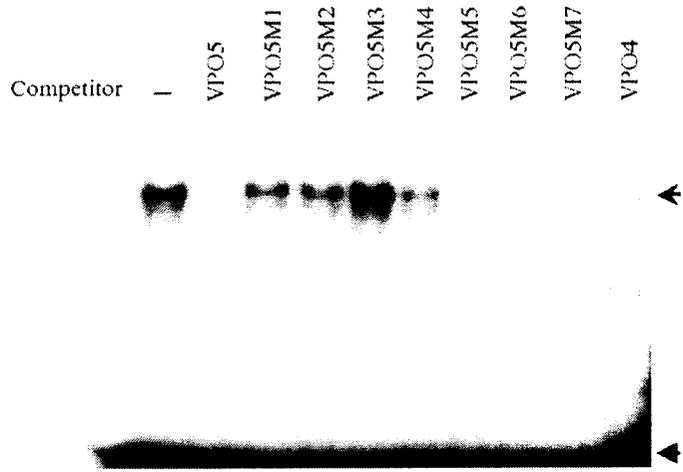
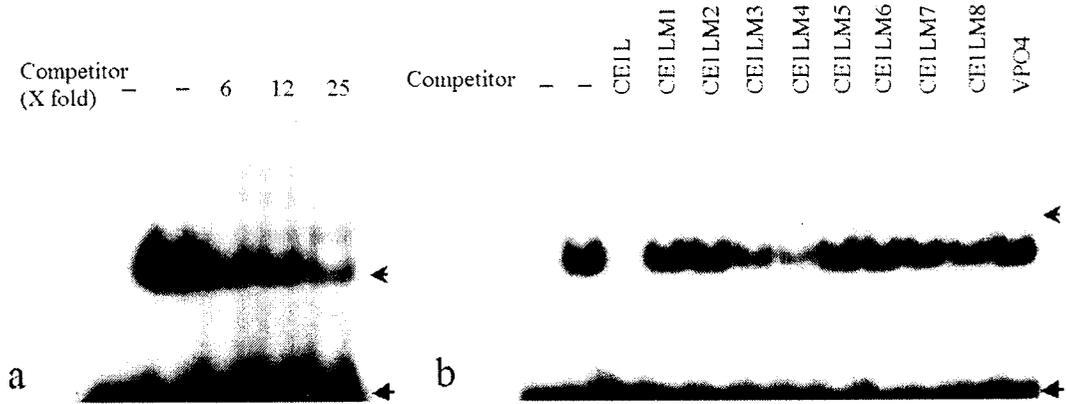


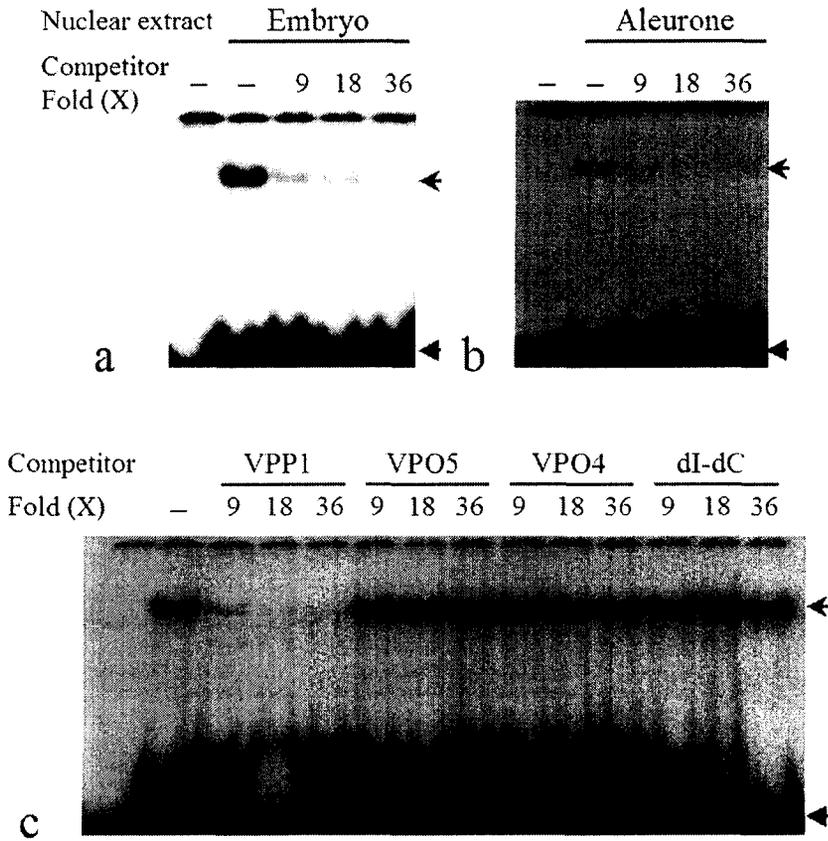
Figure 4

```

CEIL      TTCTCCCTCTCCACCTCC TGCTTGC ACT
CEILM1   .....AAA.....
CEILM2   .....GAG.....
CEILM3   .....AA.....
CEILM4   .....C.....
CEILM5   .....AA.....
CEILM6   .....GAA.....
CEILM7   .....GT A.....
CEILM8   .....GGT.....
    
```

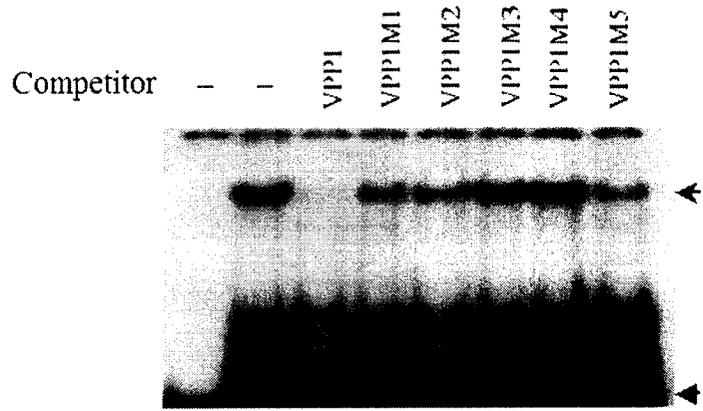


**Figure 5**



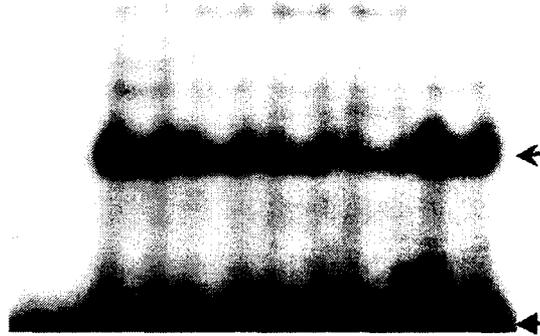
**Figure 6**

VPP1 GAGCTICCGICCGTGGGATCAUGCAGAGACTGCATCGGACGGFCAGGAATCGTCCCAACACTG  
 VPP1M1 ..... CCTFACTTACA .....  
 VPP1M2 ..... AAGCTGTAIGA .....  
 VPP1M3 ..... GAGTCATGCTA .....  
 VPP1M4 ..... AGTAACTGAAG .....  
 VPP1M5 ..... GCTACGTGGTG .....



**Figure 7**

Competitor		ZIP1			VPO4	
Fold (X)	-	-	6	12	25	25

**Figure 8**

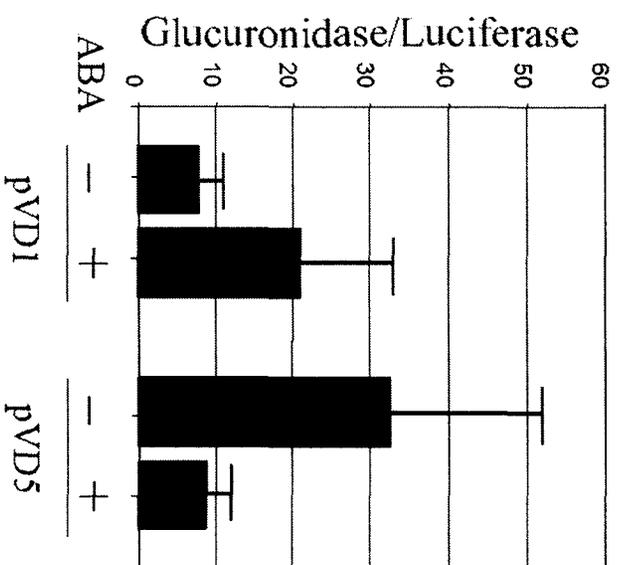


Figure 9

Treatment	<u>Fresh</u>		<u>MS</u>		<u>ABA</u>		<u>Mannitol</u>	
E.NE.(ug) –	24	12	24	12	24	12	24	12

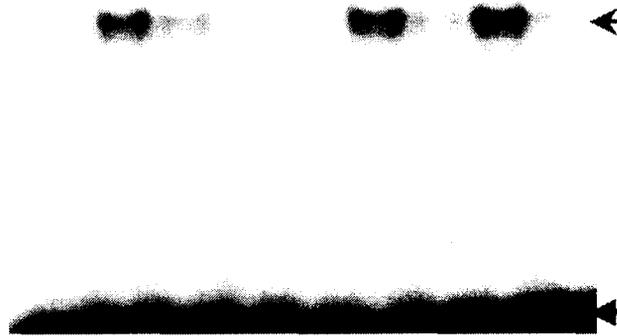


Figure 10

## CHAPTER 4. GENERAL CONCLUSIONS

In this thesis, we reported two studies, which will help to understand endosperm cell fate specification. CR4 is critical for aleurone cell fate specification in maize, and the analysis of the CR4 gene family in Arabidopsis confirmed the conservation of CR4 among plant species. The functional study of the regulation of *Vp1* expression sheds new light on how physiological conditions regulate this important developmental regulator in maize embryos, and possibly in aleurone cells.

CR4 is well conserved among maize, rice and Arabidopsis. Other members of the CR4 gene family exhibit conservation and divergence. Two AtCRRs, AtCRR1 and AtCRR2, appear not to be present in rice, and possibly not in the maize genome either. In Arabidopsis, the five members of CR4-related genes are differentially expressed and the expression patterns are coincident with the phylogenetic data. The *ACR4* promoter:GUS fusion result is consistent with expression data and gains more insight into the expression pattern. In developing seed, the *ACR4* promoter is active in the aleurone-like peripheral cell layer of the endosperm, mimicking *Cr4* expression in maize aleurone cells. *ACR4* plays an important role in seed coat development. The function of the other four AtCRRs could not be assigned in this study because gene knockouts had no apparent mutant phenotype.

*VP1* plays an important regulatory role in late embryogenesis and seed maturation in *Zea mays*. It also regulates the expression of *CI*, a transcription factor regulating anthocyanin genes in aleurone cells. In this study, we showed that physiological conditions such as ABA,

high salinity and osmotic stress could induce *Vp1* expression in 20 DAP embryos. This induction of *Vp1* expression by ABA did not require *de novo* protein synthesis. EMSA experiments identified a potential ABRC1 in the *Vp1* promoter and showed that the protein-ABRE interaction was regulated by ABA and high osmoticum. The ABA induction of *Vp1* expression was also confirmed by transient gene assay, where ABA could significantly increase *Vp1::GUS* expression. All these data suggest the importance of ABA in the regulation of *Vp1* expression in maturation phase embryos.

In this study, we also showed that the ABRE (VPO5) could also interact with nuclear proteins in aleurone cells. The EMSA data suggest that the protein-ABRE interaction in aleurones is not identical to that in embryos. The difference could be due to differential modification of transcription factors or due to regulation by a different set of transcription factors. The EMSA data of VPP1 also showed interactions with nuclear proteins from both embryos and aleurone cells. It will be interesting to test the identified ZIP1 and CE1L binding sites and the functional importance of these binding sites in aleurone cells, and to determine the similarities and differences in *Vp1* regulation between embryos and aleurone cells.

Taken together, the *Vp1* promoter work provides a framework for studying the regulation of *Vp1* expression both in embryos and aleurone cells. It gives a starting point to isolate developmental regulators of *Vp1* such as the transcription factors binding to ABRE, ZIP1, CE1L and VPP1. The transcription regulators will provide one step further up the genetic hierarchy that regulates aleurone cell development. The functional analysis of these

transcription factors will help to better understand the regulation of *VP1* expression and aleurone cell differentiation.

## ACKNOWLEDGEMENTS

I am grateful to Dr. Philip W. Becraft for his guidance and generosity throughout my graduate studies. I would like to give my special thanks to my committee members, Dr. Thomas Peterson, Dr. Joanne Powell-Coffman, Dr. David Hannapel and Dr. Yanhai Yin, I also want to thank the former committee members, Dr. Steven Rodermel and Dr. Dan Nettleton.

During these years, the whole Becraft lab, especially Kejian Li, Dr. Nrisingha Dey and Antony Chettoor gave continuous support. The discussion with Chuan Shen from Dr. Powell-Coffman's Lab and Dr. Ping Che are very helpful.

I would like to give special thanks to my wife, Mingjuan Wang, and my son, Kevin Cao, for their love, support and encouragements. I also want to thank my parents and sisters for their understanding and supports in various ways, which cannot be forgotten.

With assistance from these people and those not mentioned here, it becomes possible for me to finish this thesis.