

A Maize Embryo-Specific Gene Encodes a Proline-Rich and Hydrophobic Protein

Matilde Josè-Estanyol, Luis Ruiz-Avila, and Pere Puigdomènech¹

Departament de Genètica Molecular, Centre d'Investigació i Desenvolupament–Consejo Superior de Investigaciones Científicas, Jordi Girona 18, 08034 Barcelona, Spain

A gene from maize that encodes a hybrid proline-rich protein (HyPRP) formed by two well-defined domains, proline-rich and hydrophobic, respectively, has been characterized at the level of its structure and expression. The proline-rich domain is composed of elements PPYV and PPTPRPS, similar to those found in PRP proteins from soybean. The hydrophobic domain is rich in cysteine and is similar to seed proteins, mainly to a soybean hydrophobic seed protein. In maize, HyPRP is encoded by a single gene, and its mRNA accumulates in immature maize zygotic embryos, with a maximum accumulation between 12 and 18 days after pollination. The HyPRP mRNA can also be detected in ovary prior to pollination. In situ hybridization experiments on embryo sections show an expression of the gene in scutellum and in nonvascular cells from the embryo axis. Functional hypotheses related to HyPRP are discussed.

INTRODUCTION

Different classes of proteins highly abundant in proline residues have been described in plants. The best studied of these proteins at both the protein and DNA levels are hydroxyproline-rich glycoproteins (HRGPs) and proline-rich proteins (PRPs). HRGPs are one of the main extractable components of both dicotyledonous (where they are also called extensins) and monocotyledonous species. PRPs have been better characterized in soybean (Hong et al., 1987; Averyhart-Fullard et al., 1988). In general, these proteins are supposed to have a structural role in the cell wall (Cassab and Varner, 1988). Immunological methods have shown that the maize HRGP (Ludevid et al., 1990) and soybean PRP (Ye et al., 1991) are located in the cell wall.

The availability of DNA probes corresponding to different cell wall proteins has allowed the study of how their mRNAs are accumulated in different parts of the plant or under different physiological conditions. Extensins are developmentally regulated (Memelink et al., 1987), and they are induced by fungal infection and wounding (Corbin et al., 1987). In some cases, genes encoding HRGPs present very specific patterns of expression as is the case for a gene mainly expressed in lateral root initiation in tobacco (Keller and Lamb, 1989). PRPs are encoded by a number of different genes whose mRNAs are accumulated preferentially in distinct plant tissues (Hong et al., 1989).

A different class of PRP proteins with proline-rich repetitive stretches followed by a hydrophobic C terminus has recently been described. These proteins include a partial cDNA sequence corresponding to an mRNA accumulated in young

tomato fruit (Salts et al., 1991) and a full-length cDNA for an ubiquitous bean PRP whose mRNA is rapidly downregulated by fungal elicitor and wounding (Sheng et al., 1991). No data on the occurrence of such proteins in monocots or on the structure of genes encoding them have been reported so far. Other plant proteins also contain partial proline-rich stretches: for instance, proteins having specific functions such as nodulins (Franssen et al., 1987) and storage proteins (Prat et al., 1987; Kreis and Shewry, 1989). At least in the latter case, no functional relation to the cell wall has been proposed.

In dicotyledonous species, multigenic families encode extensins and PRPs, and each particular gene has a defined pattern of expression. In contrast, cereal HRGPs are part of a much simpler system. They are probably encoded by a single gene in maize (Stiefel et al., 1990), sorghum (Raz et al., 1991), and rice (Caelles et al., 1992). In maize, an HRGP has been studied at the protein (Kieliszewski and Lampion, 1987; Hood et al., 1988), cDNA (Stiefel et al., 1988), and genomic (Stiefel et al., 1990) levels. Maize HRGP mRNA is accumulated during cell division and upon wounding (Ludevid et al., 1990), and it has tissue-specific expression in different parts of the immature embryo. In particular, neither the HRGP nor its mRNA can be detected in the scutellum (Ruiz-Avila et al., 1991). In view of these results, it was of interest to look for other related maize genes that could have a distinct pattern of expression in the maize embryo.

This article describes the identification of a gene from maize that encodes a protein having a proline-rich domain. It is expressed only in immature embryos and ovaries and more specifically in nonvascular cells having a pattern of expression complementary to that reported for HRGP in the embryo.

¹ To whom correspondence should be addressed.

The sequence of the protein is rich in proline, and it has two main domains that are well defined. The N-terminal half is repetitive and proline rich, and it is similar to PRPs; however, its C terminus is hydrophobic and similar to a group of seed proteins from soybean (Odani et al., 1987).

RESULTS

Cloning of a Maize Gene Encoding a Protein with Proline-Rich and Hydrophobic Domains: Nucleotide and Protein Sequences

One of the main extractable proteins from the maize cell wall is an HRGP that has been characterized at the protein (Kieliszewski and Lamport, 1987; Hood et al., 1988), cDNA (Stiefel et al., 1988), and genomic (Stiefel et al., 1990) levels. The cDNA and genomic probes have been used to study the accumulation of HRGP mRNA in different organs of maize and under different conditions. In particular, tissue specificity has been observed in immature embryos because neither the protein nor the corresponding mRNA has been detected in the scutellum (Ruiz-Avila et al., 1991); this raises the question of whether the existence of other proteins could fulfill functions similar to the HRGP in defined tissues. It was therefore of interest to try to identify other genes having sequence similarity to those encoding the HRGP. To this end, different maize genomic libraries were screened with a probe corresponding to the coding sequence of the HRGP cDNA.

One of the clones detected in the screening of one of the maize genomic libraries, clone *mIP9*, appeared interesting due both to its sequence and to the features of its expression. One fragment of clone *mIP9* hybridized to a labeled cDNA synthesized from the maize embryo poly(A)⁺ RNA. The sequence of this zone is shown in Figure 1. In this region, a single open reading frame was present. Its translation into a protein is also shown in Figure 1. The codon usage of the protein is biased toward the use of a G or C residue in the third base of the codons, giving a final G+C content of 72% in the coding sequence; this is the case for many cereal genes as well. The open reading frame is continuous, suggesting that no intron is present in this gene. The presence of introns in the 3' untranslated region, a typical feature of many genes encoding HRGPs from plants (Chen and Varner, 1985; Stiefel et al., 1990), cannot be discarded, although only an imperfect intron consensus sequence can be found in this zone. In the sequence, the putative TATA-box and polyadenylation signals can be observed, and they are underlined in Figure 1. The TATA-box is surrounded by complementary sequences that could give rise to a hairpin structure producing a protruding TATA sequence. A CACGTGGC motif that has been related to abscisic acid (ABA) regulation is present at -188 bp from the TATA sequence (data not shown in Figure 1).

The protein deduced from the DNA sequence of clone *mIP9* has 302 amino acids, giving an apparent molecular mass of

```

cgcaaacctcacttgcccggcccaaccgtacgtgacgctcctcctctataaaaggagacg -60
cggtgctgctcagtggtcacagaacagccgctcgccgctgctgagccaaggcgca -1
ATGGCAACCTCCACCTCAGTTCCTCCCTCTCTCTCTCGCCGCGCCGCGCCGCTCTC 60
M A T S T S V P V L L L L A A A A G L L 20
CTGCTCCCGGCTCGTGGCGTGGCAGAGCAGCAACTGTGCACGGCGCCGCGCGGAGT 120
L L P G S L A W Q S S N C A P G A P G S 40
GGCTCGCCGACGTGGCTTCACCCGGCTCCGGCTCCGGCGGGCCACAGCAAGCCGCC 180
G S P T W P S P G S G S G G G H S K P P 60
AAGCACCCATGGCAAGCCGCCAAGTGGCCCGCTGCAACCTCCGTACGTGCTCCG 240
K H H H G K P P K C P P C N P P Y V P P 80
ACGCCTCGGCGCTCACCCGGTACGTGCCACCGTATGTCCTCCGCGGACGCTCCGCGCTCA 300
T P R P S P P Y V P P Y V P P T P R P S 100
CGGCCCTAGCGCCACCATACGTCCCGTCCCGCCGACGCTCCGCGCTCACCCCGTAC 360
P P Y V P P Y V P V P P P T P R P S P P Y 120
GTGCCACCATACGTCCCGCTCCCGCGGACGCTCCGCGCTCACCCCGCTACGTGCCACCA 420
V P P Y V P V P P T P R P S P P Y V P P 140
TAGTCCCGTCCCGCGGCTCCGCGTCCGCGCTACGCTACCCCGGACGGCAGCG 480
Y V P V P P T P R P S P P Y V P P T P R 160
CGCCGACTCCCGCTACGTCCACCGACCGCCGCTACGTCCACCGACCGCTCCGCGCA 540
P P T P P Y V P P T P P Y V P P T P R P 180
TCACCACCGCTACGTGCCACCTACGTCCCGCGGACGCTCCGCGCTCACCCCGCTAC 600
S P P P Y V P P P Y V P P T P R P S P P Y 200
GTGCCACCGTACGTCCCGCGGACCGCGCGGCGTGGGACGTGCCGATCGACACGCTG 660
V P P Y V P P T P P A V R T C P I D T L 220
AAGTGAAACCGTCCGCGGACGCTGAGCGGGCTGATCCACTGTGATCGGGCAGGAG 720
K L N A C V D V L S G L I H L V I G Q E 240
GCCAGTCCAAAGTCTCCCGCTGGTGCAGGGGGTGGCGGACCTGGACGGCGCTGTGC 780
A R S K C C P L V Q G V A D L D A A L C 260
CTCTGCACCCATCCCGCGCCGCTGCTCAACATCAACATCTACTCCCCATCGGCTC 840
L C T T I R A R L L N I N I Y L P I A L 280
AACCTGCTCATCCTCCGCGCAAGCAGCGCCGCTCCGCGCTCCAGTCCCGCGCCCTAC 900
N L L I T C G K H A P S G F Q C P P L Y 300
GACTGATgaaacccaccacagccaactctctatacgcgcccagctgcatgcatgatgt 960
D U 302
gatggactcctgtgcaacgtagcgtagctagtagtagcatgtcaaccatttgctccgtg 1020
attctctagttcttctgctcctcggtatgctgtaagtaagctgcttctctgtaggact 1080
cgggaaactttccaactccgatgtaagcgtgtagcttgccttaatacagtcgatgaccat 1140
gttaatacaagtgtagcccttctctgctctgctgtagtcttctcgttcttcaactct 1200
tgatctgcagatgatataatgaaaacgtttatacgaattgaaatataattctctctgtt 1260
gacaacgactgactctgctcact 1282

```

Figure 1. Nucleotide Sequence of the Maize *HyPRP* Gene and the Corresponding Protein Sequence.

Nucleotides are numbered from the A of the start codon for translation (+1). Putative TATA-box and polyadenylation signals are underlined. The deduced amino acid sequence for HRGP is depicted using the single-letter code.

31.6 kD. It has 11 cysteines, more basic amino acids than acidic ones, 31% proline, and 27% hydrophobic residues. However, the residues with a specific character are not evenly distributed along its sequence. The polypeptide has a clear domain structure, and, as shown in Figure 2A, the repetitive nature of its N-terminal region can be easily emphasized. From the N terminus of the protein, the sequence begins with a stretch of amino acids typical of a signal peptide (von Heijne, 1983); this feature has been already observed in proteins that are secreted from the cell, such as the components of the cell wall. This is followed by a nonrepetitive glycine- and serine-rich zone. A basic proline-rich region containing two cysteines and three consecutive histidines follows. The central and C-terminal part of the protein sequence deduced from *mIP9* is formed by two

regions with distinct characters. The first one is a proline-rich repetitive region formed mainly by the sequences PPTPRPS and PPYV with high conservation between the repeats at both the amino acid and nucleotide levels. The second one is 91 residues long, nonrepetitive, and hydrophobic (44% I, L, A, V, or F residues), and it contains eight cysteine residues. The domain structure of the protein can also clearly be seen in the hydrophilicity plot shown in Figure 2B. The plot illustrates the putative hydrophobic signal peptide, the basic hydrophilic stretch, the repetitive proline-rich region, and the final hydrophobic domain.

After searching the Protein Identification Resource data base compiled by the National Biomedical Research Foundation for sequence homology to the 91-residue-long C-terminal region, a seed hydrophobic protein from soybean appeared to be the most similar, as shown in Figure 3. The zones around

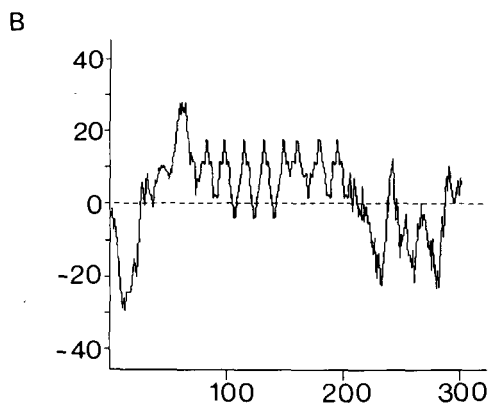
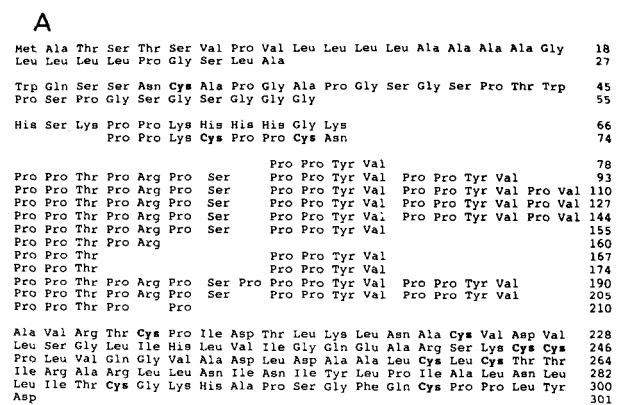


Figure 2. HyPRP Protein Sequence.

(A) The protein sequence deduced from the genomic clone of maize *HyPRP*. The sequence has been presented to emphasize the different domains of the protein and the homology of the repetitive elements in the protein sequence.

(B) Hydropathy profile of the amino acid sequence deduced from the genomic clone of the maize *HyPRP*. The mean hydropathy plot of a window of nine consecutive residues (Kyte and Doolittle, 1982) is plotted against the residue number.

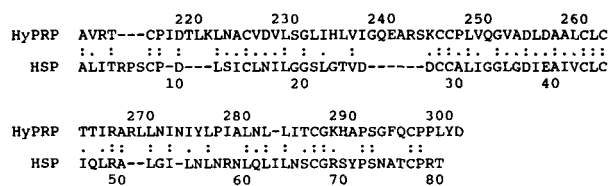


Figure 3. Homology of the HyPRP Hydrophobic Domain with the Soybean Hydrophobic Seed Protein.

Numbers indicate the corresponding residue number for each protein from the ATG of the HyPRP and from the beginning of the sequence in the soybean hydrophobic seed protein (HSP). HyPRP is a fragment of the sequence of the maize HyPRP; HSP is the whole soybean hydrophobic seed protein (Odani et al., 1987). Total identity and conservative replacements are indicated by two or one points, respectively.

the cysteine residues are the most conserved between the two proteins. Due to the features of its sequence, this gene is known as the hybrid-sequence proline-rich protein gene (*HyPRP*).

HyPRP Genomic Structure and mRNA Accumulation

The *HyPRP* probe has been used to find how many copies of the corresponding gene exist in the maize genome. The result of the DNA gel blot analysis is shown in Figure 4 and indicates that probably only one copy of this gene is present in the maize genome. The restriction sites observed in the blot match those deduced from the nucleotide sequence.

When the probe was used to detect the corresponding mRNA in RNA extractions from the maize embryo, a single band of ~1300 nucleotides was observed, as shown in Figure 5. This length corresponds to the one predicted from the genomic sequence, assuming that transcription starts 28 nucleotides downstream of the putative TATA box shown in Figure 1. At this point, an adenine residue surrounded by two pyrimidines could act as the transcription initiation site. In Figure 5A, the level of *HyPRP* mRNA accumulation in immature maize embryos at different stages of maturation is presented. The mRNA was detected from 12 days after pollination (DAP) until 40 DAP, with a similar level of expression from 12 to 18 DAP. At 20 DAP, its expression began to decrease and was barely detectable at 40 DAP. At stages earlier than 12 DAP, it is difficult to separate the embryo from the endosperm, but expression in whole kernel could be observed with a maximum accumulation at 8 DAP, as is shown in Figure 5E. At 12 DAP, a separate analysis of mRNA from whole kernel, endosperm, and embryo indicated that the kernel signal came mainly from the embryo at this stage of development (Figure 5C). In embryos 20 DAP, the mRNA was observed in both embryo axis and scutellum (Figure 5C). When RNA extracted from different organs of the mature plant was hybridized with the *HyPRP* probe (Figures 5C and 5D), no mRNA was observed except for immature ovaries, although expression in this tissue disappeared if no pollination had occurred.

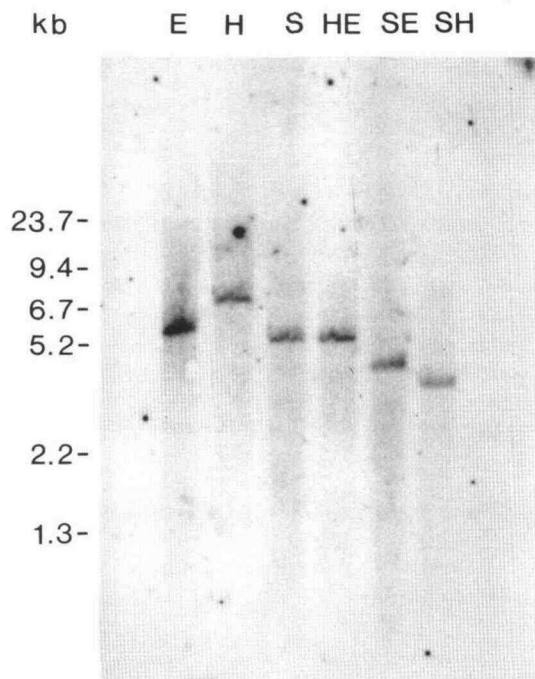


Figure 4. Genomic DNA Blot Analysis of the *HyPRP* Gene.

DNA (10 μ g) from the maize W64A variety was digested with EcoRI (E), HindIII (H), SacI (S), HindIII+EcoRI (HE), SacI+EcoRI (SE), and SacI+HindIII (SH). A 1600-bp *AccI* fragment from the 5' end and coding region of the *HyPRP* gene was used as the probe for hybridization. The positions of length markers (in kilobases) are shown.

HyPRP mRNA levels were also measured in two genotypes known to affect maize embryogenesis. To this end, mRNA steady-state levels were measured in embryos carrying the *viviparous-2* (*vp2*) mutation, which retards the onset of desiccation due to a low level of ABA, and in somatic embryos of a highly embryogenic tissue culture of the maize variety *Cat100-1*. This variety has been used to regenerate fertile plants from transformed protoplasts (Prioli and Söndahl, 1989). The result of the experiment with the *vp2* mutant is shown in Figure 5B, and it can be observed that in this case the *HyPRP* mRNA level is maintained during a longer period in immature embryos than in a normal variety (compare with Figure 5A) and decreases between 50 and 60 DAP. When the level of *HyPRP* mRNA was measured in calli from the highly embryogenic *Cat100-1* maize variety, no expression could be observed under any of the conditions assayed. After treatment with the synthetic auxin 2,4-D, the progressive organization of somatic embryos (globular, heart, and torpedo stages) could be observed under the microscope (data not shown), but no *HyPRP* mRNA could be detected (Figure 5F). Similar experiments in calli from the W64A maize variety (Torné et al., 1984; Ludevid et al., 1990) showed no evidence of *HyPRP* induction (data not shown). Therefore, the expression of *HyPRP* seems to be

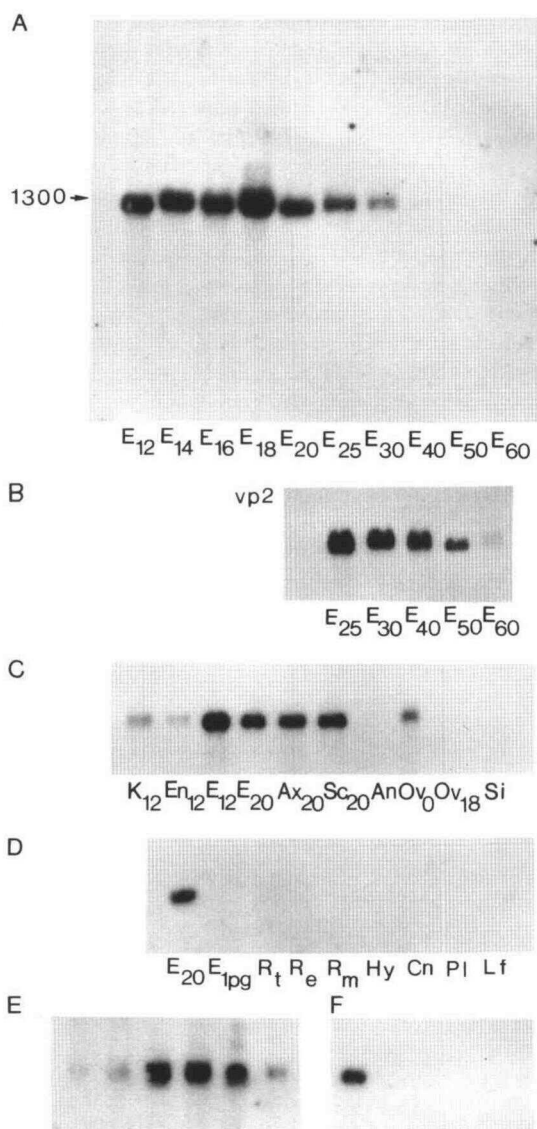


Figure 5. RNA Gel Blot of mRNA Extracted from Different Organs of Maize.

Total RNAs (10 μ g for kernel and immature embryos, and 30 μ g for the other organs) were subjected to electrophoresis on agarose-formaldehyde gels, transferred to nylon membranes, and hybridized with a 600-bp *DdeI* fragment from the coding region of the *HyPRP* genomic clone. Band length is given in nucleotides.

(A) Immature embryos from 12 (E_{12}) to 60 (E_{60}) DAP.

(B) *vp2* immature embryos from 25 (E_{25}) to 60 (E_{60}) DAP.

(C) K_{12} , kernel 12 DAP; En_{12} , endosperm 12 DAP; E_{12} , embryo 12 DAP; E_{20} , embryo 20 DAP; Ax_{20} , axis from embryos 20 DAP; Sc_{20} , scutellum from embryos 20 DAP; An, anthers; Ov_0 , ovary from the female flower prior to pollination; Ov_{18} , unpollinated ovary 18 days after the pollination of the rest of the flowers; Si, silks.

(D) E_{20} , embryos 20 DAP; E_{1pg} , embryos 1 day postgermination; R_t , 2-mm segments at the root tip; R_e , 5-mm segments at 1 cm from the tip; R_m , 20-mm segments at 3 cm from the root tip; Hy, hypocotyl; Cn, coleoptile node; Pl, plumule; Lf, adult leaf.

restricted to the early immature zygotic embryo, but it has a low but detectable level of expression in the ovary prior to pollination.

mRNA Accumulation in Immature Embryos: In Situ Hybridization Experiments

Given the overall pattern of accumulation described above, it was important to investigate in what embryo cell types *HyPRP* mRNA is present. To this end, in situ hybridization experiments were carried out. Figure 6 shows two longitudinal sections of an embryo 16 DAP and two transverse sections hybridized with the *HyPRP* probe. As expected from the RNA gel blot results (Figure 5), the hybridization appears both in the embryo axis and the scutellum (Figure 6A). In the scutellum (Figures 6A and 6B), it appears as a gradient of expression from the cells beneath the single layer of secretory epidermal cells located in the abaxial side of the embryo toward the inner zones. Cells where hybridization is more intense have a higher overall level of RNA, as seen by acridine orange staining (data not shown). Hybridization is observed in neither the procambium of the scutellum nor in the secretory epidermal cells. In the axis (upper coleoptile and coleorhiza), *HyPRP* mRNA is present in most of the cells; however, it is absent or at a very low level in epidermal and vascular tissues from the coleoptile (Figure 6C) or the coleorhiza (Figure 6D).

The expression of the *HyPRP* gene in immature embryos can also be observed in two areas, having a characteristic distribution of cell types, namely in the coleoptile and the radicle, as shown in Figure 7. In the coleoptile, it is distributed in most of the tissues; however, it cannot be detected in either the two provascular bundles of the coleoptile and the leaf primordia, or the epidermal tissues (Figure 7A). Sections adjacent to those presented in Figures 6 and 7 have been hybridized with a cDNA probe corresponding to a maize lipid transfer protein (Sossountzov et al., 1991). In this case, we observed a very specific concentration of the mRNA in the epidermis of the axis and scutellum and particularly in the abaxial epidermal layer. *HyPRP* mRNA is not detected in these cells, and it shows more concentrated expression in the inner tissues of both the scutellum and axis. In the same figure, a control with the sense probe is shown (Figure 7B) as well as the complementary pattern of expression obtained in a similar section using the HRGP

probe (Figure 7C). The same comparison can be made in the radicular region where the procambium and the prominent proylem cells hybridize strongly with the HRGP probe (Figure 7E) but very weakly with the *HyPRP* probe (Figure 7D). The pattern is maintained through early embryogenesis at least between 14 and 20 DAP (data not shown).

The cell types showing hybridization to the *HyPRP* probe in immature maize embryos can be better observed at higher magnification. Figure 8 shows details of both the axis and scutellum. *HyPRP* mRNA from the outer and inner epidermis of the coleoptile and from the single layer of secretory epidermal cells located in the abaxial zone of the scutellum was absent (Figures 8A, 8B, and 8C). *HyPRP* mRNA is either absent or at a very low level in the provascular tissues and in the procambium cells. Hybridization with the *HyPRP* probe occurs mainly in cells containing starch grains (parenchyma) in the axis, whereas in the scutellum it is mainly associated with cells with active transcription and less with differentiated cells of the inner zone.

DISCUSSION

In this study, we report the sequence and expression of a maize gene encoding a hybrid protein having proline-rich and C-terminal hydrophobic domains (*HyPRP*). The proline-rich region is formed by two main repeating units: PPTPRPS and PPYV. Sequences similar to the PPTPRPS can be found in the maize HRGP, which contains the PPTYTP repeat. However, a remarkable similarity to the PPYV sequence is found in soybean PRPs that are formed by PPVYK motifs (Hong et al., 1989). In maize, this protein domain is composed of 12 elements, not all containing the entire set of repetitive motifs. An interesting feature is that the conservation of the sequences between the repeated elements occurs not only at the protein level but also at the nucleotide level. This suggests that mechanisms of production of sequence variability already proposed for HRGP sequences (Raz et al., 1992) may also be active in other repetitive maize sequences, such as *HyPRP*. Another characteristic feature of *HyPRP* is that the main basic amino acid of its sequence is arginine instead of lysine, which is the residue most often appearing in HRGP or PRP sequences described previously. The presence of proline and serine may also indicate the possibility of post-translational modifications such as hydroxylation of prolines and glycosylation of serines. Finally, the C terminus is hydrophobic and it contains eight cysteine residues, which raises the possibility of intramolecular (or intermolecular) bridges that could be important for the formation of the secondary structure in this part of the protein.

As indicated before, the *HyPRP*-encoded protein appears as the hybrid of two types of sequences, a proline-rich domain followed by a cysteine-rich one. The cDNAs encoding two other proteins with a proline-rich domain and a nonrepetitive C-terminal sequence have recently been described. The first one (Sheng et al., 1991) has a C-terminal sequence with no

Figure 5. (continued).

(E) Whole immature kernels from 2 (K_2) to 12 (K_{12}) DAP.

(F) *Cat100-1* embryogenic calli: E₂₀, embryos 20 DAP from maize W64A; 1, calli cultured in the presence of 2,4-D (1.65 mg/mL); 2, calli suspension culture in the presence of a reduced concentration of 2,4-D (1 mg/mL); 3, differentiated (globular stage) calli suspension culture after 2 months in the presence of the same reduced 2,4-D concentration; 4, differentiated calli (heart- and torpedo-stage embryos) cultured in the absence of 2,4-D.

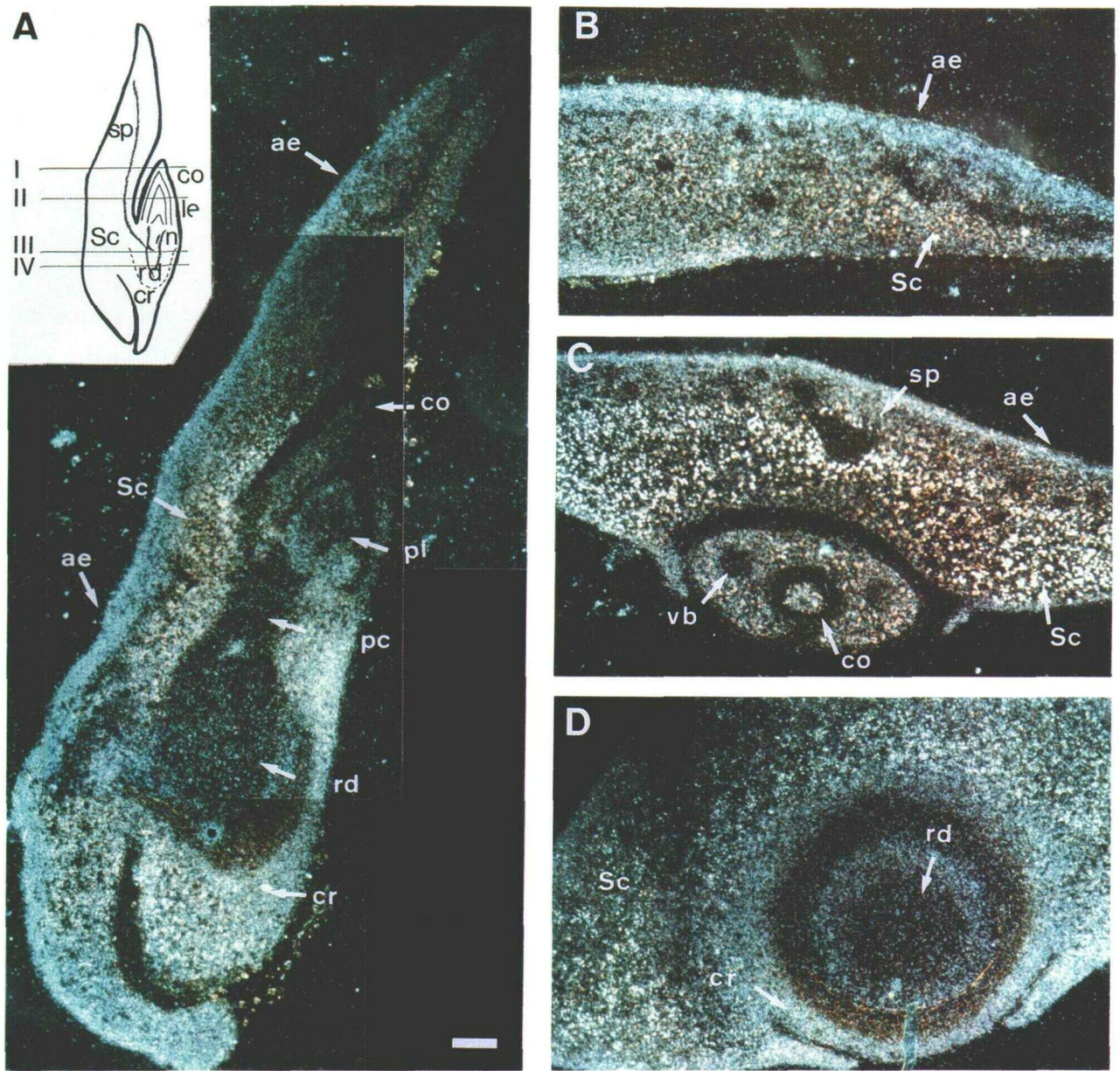


Figure 6. Localization of *HyPRP* mRNA in Sections from Embryos 16 and 18 DAP.

(A) In situ hybridization of an *HyPRP* antisense RNA probe with an embryo 16-DAP longitudinal section. (Schematic described below.)
 (B) In situ hybridization of an *HyPRP* antisense RNA probe with an embryo 16-DAP longitudinal section through the upper scutellar region.
 (C) In situ hybridization of an *HyPRP* antisense RNA probe with an embryo 18-DAP transverse section through the upper coleoptile.
 (D) In situ hybridization of an *HyPRP* antisense RNA probe with an embryo 18-DAP transverse section through the coleorhiza.
 Bar = 100 μ m. All the sections were examined under a dark-field microscope. Hybridization was observed as blue areas. A schematic drawing of immature maize embryo is also included showing the location of sections shown in the photographs: I corresponds to (C); II to Figures 7A, 7B, and 7C; III to Figures 7D and 7E; and IV to (D). pc, procambium; co, coleoptile; le, leaf primordia; n, scutellar node; rd, radicle; cr, coleorhiza; Sc, scutellum; ae, abaxial epidermis; sp, scutellar procambium; vb, vascular bundle.

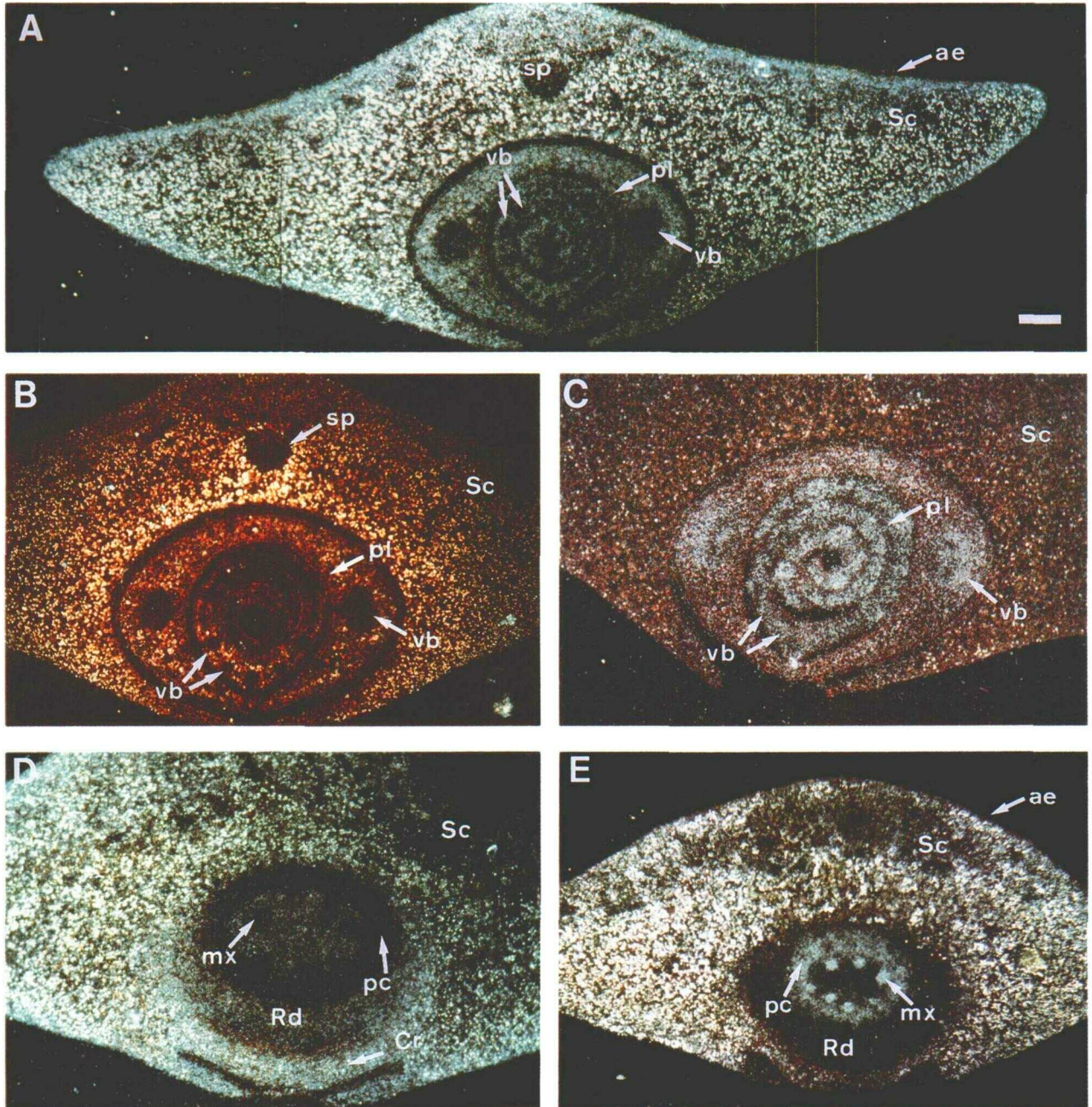


Figure 7. Comparison of the Patterns of *HyPRP* and *HRGP* sense RNA Accumulation in Immature Maize Embryos by in Situ Hybridization.

(A) In situ hybridization of an *HyPRP* antisense RNA probe with a coleoptile transverse section from an embryo 18 DAP.
 (B) In situ hybridization of an *HyPRP* sense RNA probe with a coleoptile transverse section from an embryo 18 DAP.
 (C) In situ hybridization of an *HRGP* antisense RNA probe with a coleoptile transverse section from an embryo 20 DAP.
 (D) In situ hybridization of an *HyPRP* antisense RNA probe with a radicle transverse section from an embryo 18 DAP.
 (E) In situ hybridization of an *HRGP* antisense RNA probe with a radicle transverse section from an embryo 16 DAP.
 Bar = 100 μ m. All the sections were examined under a dark-field microscope with a blue filter, with the exception of the section shown in (B), where a yellow filter was used. Sc, Scutellum; sp, scutellar procambium; vb, vascular bundle; pl, plumule; ae, abaxial epidermis; pc, procambium; Rd, radicle; Cr, coleorhiza; mx, metaxylem precursor cells.

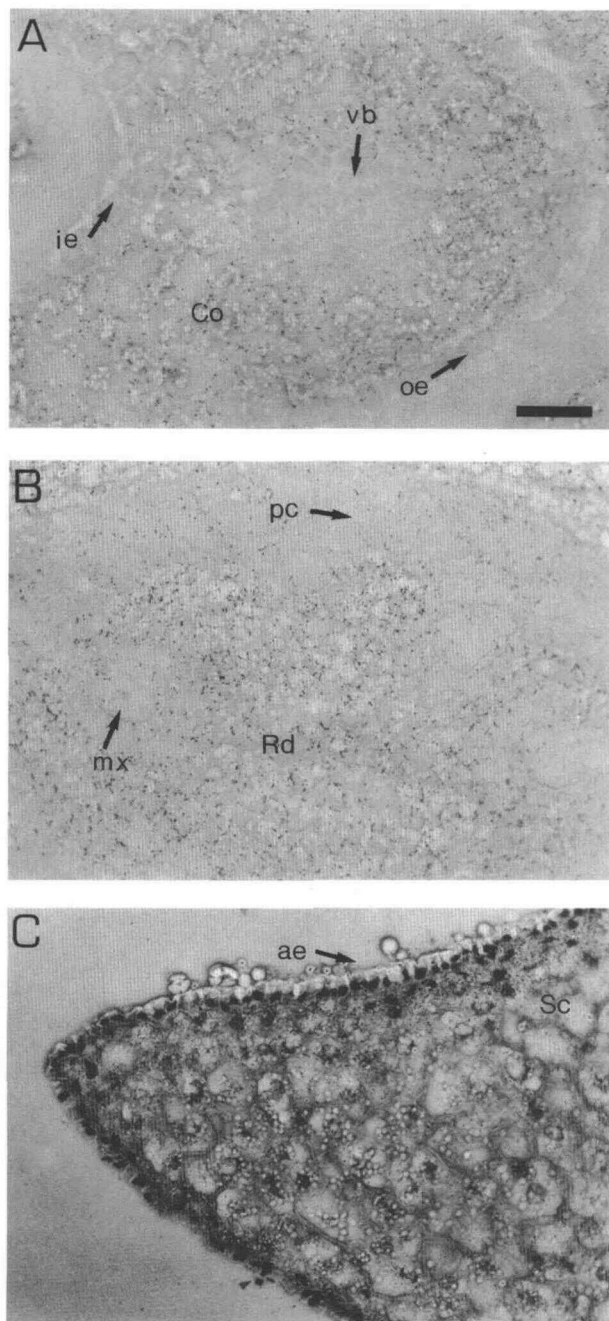


Figure 8. Details of Maize *HyPRP* mRNA Localization in Sections from Embryos at 18 DAP by in Situ Hybridization.

(A) In situ hybridization of an *HyPRP* antisense RNA probe with a coleoptile vascular bundles transverse section stained with fast green.

(B) In situ hybridization of an *HyPRP* antisense RNA probe with a radicle transverse section stained with fast green.

(C) In situ hybridization of an *HyPRP* antisense RNA probe with a scutellum transverse section stained with haematoxylin to better observe the morphology of the scutellum.

Bar = 40 μ m. Co, coleoptile; Rd, radicle; Sc, scutellum; ie, inner epidermis; oe, outer epidermis; vb, vascular bundle; pc, procambium; mx, metaxylem precursor cells; ae, abaxial epidermis.

homology with the one described here; it is poor in cysteines. The pattern of expression is also different for the maize *HyPRP* because its mRNA appears to be ubiquitous in different organs of the plant. The second one is a partial cDNA for a tomato PRP that is apparently accumulated in immature fruit (Salts et al., 1991). Its partial cDNA sequence shows a proline-rich domain formed by repeating units that are more variable than in the maize *HyPRP*, although the PPYV motif appears at least two times. The hydrophobic C terminus of tomato PRP has a higher proportion of glycine than the maize protein, but both sequences can easily be aligned with 63% homology. The regions around the cysteine residues are the most similar between the two proteins. This sequence homology is interesting given the evolutionary distance between monocot and dicot plants. The C termini of both proteins are similar to that of a seed protein, the soybean hydrophobic protein (Odani et al., 1987). Both the *HyPRP*-encoded protein and tomato PRP can be aligned with it, and a 36% amino acid homology exists between the maize and soybean proteins (Figure 3). One of the main storage proteins from maize, γ -zein (Prat et al., 1987), also presents a repetitive proline-rich domain followed by another one rich in cysteines. Both *HyPRP* and γ -zein C-terminal domains have an amino acid distribution shared by a large family of proteins having a storage or hydrolase inhibitor function (Henrissat et al., 1988; Kreis and Shewry, 1989). The C terminus of maize *HyPRP* differs from γ -zein in that it lacks some of the amino acids proposed to be essential for nitrogen storage such as glutamine but is similar in the distribution of cysteine residues. Thionin is another cysteine-rich protein from cereal endosperm (Ponz et al., 1986) that is synthesized as a two-domain protein; in this case, one is cysteine-rich and the other is acidic. This protein is processed to give the mature thionin. It is not known whether such a mechanism occurs in the *HyPRP*-encoded protein or in any other of these proteins.

HyPRP has a well-defined pattern of expression. Its mRNA is accumulated in the embryo, and it cannot be detected in any organ of the adult plant with the exception of the ovary. This result indicates that transcription factors responsible for *HyPRP* gene expression are activated in an ovary that is ready to be fertilized and are present until the embryo finishes its morphogenetic process. These factors are not present in somatic embryos because no expression of the *HyPRP* gene can be detected in embryogenic calli or suspension cultures. It appears that the developmental program controlling the expression of *HyPRP* has to be initiated in the ovary, and it is activated only in embryos that are going to pass through the maturation stage. The decrease in the *HyPRP* mRNA level is particularly noticeable at 20 DAP when morphogenetic and cell division stages have stopped in the embryo, and therefore just before the onset of storage and desiccation stages in embryogenesis. These processes have been correlated with the increase in the level of ABA in the maize embryo (Jones and Brenner, 1987). Although the temporal pattern of expression of *HyPRP* is distinct from that of genes induced by ABA (Quatrano, 1987), a sequence consensus with the ABA-responsive element (Guilting et al., 1990) can be found in the promoter region of the maize *HyPRP* at -188 bp from the

TATA-box. This suggests the interesting possibility that in some cases this sequence might also be correlated with repression by ABA. The accumulation of *HyPRP* mRNA in *vp2* maize embryos for periods longer than under normal conditions would favor this possibility. In contrast, proteins encoded by *rab* (*responsive to abscisic acid*) genes, which are under ABA regulation, are expressed in a lower amount in these mutants as a result of the reduced ABA concentration in *vp2* embryos (Pla et al., 1989).

The in situ hybridization experiments show a pattern of *HyPRP* mRNA accumulation in embryo sections complementary to the one observed for the maize HRGP. It is present in the scutellum where HRGP mRNA is absent (Ruiz-Avila et al., 1991, 1992). In this organ, it accumulates in a gradient from the abaxial epidermis of the scutellum toward the inner zone (Figure 6B), correlated with cells where transcription and division are very active. In the axis, the *HyPRP* mRNA appears to be accumulated in nonprovascular cells and associated with cells having a high proportion of storage material, i.e., parenchyma cells or related ones (Figure 8). In both the axis and scutellum, *HyPRP* mRNA is always absent from epidermal cells. These cells show a strong hybridization with probes from cDNAs encoding other proteins such as the lipid transfer protein (Sossountzov et al., 1991).

Finally, although nothing is known about the HyPRP protein and its role in the immature embryo, its sequence may indicate a double function. On the one hand, PRPs have been shown to be located in the cell wall (Ye et al., 1991), and, on the other, the C-terminal half is similar to proteins having either a storage or defense function. Assuming the maize HyPRP protein is a wall protein, the complementary pattern of expression between maize *HyPRP* and HRGP indicates that in the embryo there are at least two types of cells that differ in their wall protein composition. The presence of HyPRP could modify the structure of the wall protecting the cell during later embryogenic development. HyPRP could contribute temporally to the stability and defense of the developing embryo.

METHODS

Biological Materials

Unless otherwise stated, plant material used in this study was derived from seeds of a *Zea mays* cv W64A pure inbred line grown in a greenhouse in Barcelona, Spain. When necessary, seeds were germinated in the dark at 25°C. Immature kernels between 2 and 20 DAP were stored at -80°C before RNA extraction. Heterozygous caryopses of the *viviparous-2* (*vp2*) mutant of *Zea mays* were obtained from Dr. R. J. Lambert, Maize Genetic Stock Center, University of Illinois (Urbana, IL). *vp2* kernels were selected at different stages of development on the basis of white colorless endosperm for homozygous *vp2* mutants. Embryos were manually dissected from the endosperm. Embryogenic calli from Cateto maize inbred line *Cat100-1* were obtained from Dr. Paulo Arruda, Centro de Biologia Molecular e Engenharia Genética UNICAMP (Campinas, Brazil). Undifferentiated or differentiated *Cat100-1* calli were maintained in CM-1 and RM-1 media, respectively (Prioli and Söndahl, 1989). *Cat100-1* friable calli were maintained in suspension cultures; differentiation was induced in a modified MS-1

culture (Torné et al., 1984). Embryogenic and meristematic calli from maize inbred line W64A were obtained and cultured as described previously (Torné et al., 1984; Ludevid et al., 1990).

Genomic Cloning

Genomic DNA was prepared from adult leaves of maize inbred line W64A as described previously (Burr and Burr, 1981) and purified through a CsCl gradient (Sambrook et al., 1989). Maize genomic DNA was partially digested with HindIII and EcoRI. After continuous sucrose gradient purification, fragments of 5- to 10-kb were cloned into λ NM1150 (Murray, 1983), following standard procedures (Sambrook et al., 1989). The genomic library was screened with the insert of the pMC56 HRGP cDNA clone (Stiefel et al., 1988) by standard methods (Sambrook et al., 1989). A positive clone (*mIP9*) was isolated, and the region of homology to the probe was detected by blot hybridization analysis. The region was subcloned into the pUC19 vector and sequenced on both strands over a 4-kb region covering the coding region of the hybrid-sequence proline-rich protein (*HyPRP*) gene. DNA sequences were determined by the dideoxy chain sequencing method of Sanger et al. (1977) after subcloning in M13mp18 and M13mp19 using a DNA polymerase T7 sequencing kit (Pharmacia, Sweden). Some regions of high G+C content were sequenced by the chemical degradation method (Maxam and Gilbert, 1980) or with a Deaza T7 sequencing mixes kit (Pharmacia, Sweden). The nucleotide sequence data reported has been submitted to EMBL, GenBank, and DDBJ genetic sequence data banks as accession number X60432.

Genomic Blot Analysis

Genomic DNA was prepared from leaves of maize inbred line W64A as described by Burr and Burr (1981). For DNA gel blot analysis (Southern, 1975), the DNA was digested with restriction enzymes, separated by size in a 0.8% agarose gel, and blotted onto nylon membranes (Zeta-Probe; Bio-Rad) as recommended by the manufacturer. Hybridization was carried out in 180 mM NaCl, 10 mM sodium phosphate, 10 mM EDTA, 1% SDS, 0.5% nonfat dried milk, 0.5 mg/mL sonicated salmon sperm DNA at 68°C. A 1600-bp *AccI* fragment from the *HyPRP* 5' end and coding region of the *mIP9* genomic clone was used as the DNA probe, and it was labeled with phosphorous-32 to a specific activity of 0.5 to 2 $\times 10^8$ cpm/ μ g using a random priming kit (Boehringer Mannheim). Final washes were done in 0.1 \times SSC (1 \times SSC is 0.15 M NaCl, 0.015 M sodium citrate), 0.1 SDS at 65°C, and the membrane was exposed to Curix RP2 films (AGFA Corp., Orangeburg, NY) with intensifying screens (Lightning Plus; Du Pont) for 4 days at -70°C.

RNA Preparation and Gel Blot Analysis

RNA was extracted by the guanidinium-hydrochloride procedure (Logemann et al., 1987). All RNAs were checked by nondenaturing electrophoresis on 1.5% agarose gels and by ethidium-bromide staining, and the concentrations were adjusted by reading the absorbance at 260 nm. For the blotting, total RNAs were separated on gels containing 1.5% agarose-formaldehyde (Lehrach et al., 1977). The gels were blotted onto nylon membranes (Hybond-N; Amersham International), treated as described by the manufacturer, and hybridized at 42°C for at least 12 hr in a 50% formamide-containing solution with a specific probe (a 670-bp *DdeI* fragment from the *HyPRP* coding region of the *mIP9* genomic clone) that was labeled by random priming

to a specific activity of 10^8 cpm/ μ g. After hybridization, filters were washed to 0.02 M NaCl stringency and autoradiographed.

In Situ Hybridization

Embryos were collected from hand-dissected kernels at 16, 18, and 20 DAP and immediately submerged in fixing solution. The whole procedure of in situ hybridization was performed essentially according to Langdale et al. (1988). A 670-bp Ddel fragment of the 5' coding region of the *HyPRP* gene was cloned in a pBluescript SK+ vector (Stratagene) and used as template for synthesis of sense and antisense riboprobes. Transcripts from T3 and T7 promoters were produced following the instructions of the manufacturers, using 35 S-CTP (37 TBq mmol $^{-1}$; Amersham International) as a radioactive precursor. Final concentration of the probes was 0.1 μ g/mL, and each slide was hybridized with \sim 0.5 kBq of labeled probe. Hybridization was then performed as previously described (Langdale et al., 1988). The slides were exposed for 10 days using Kodak NTB-2 emulsion and stained after developing with 0.5% (w/v) fast green in 95% ethanol. Some sections were stained with haematoxylin to check the morphology of the tissues. The photographs were taken using an automated camera on a microscope (Zeiss, Oberkochen, Germany).

ACKNOWLEDGMENTS

We are indebted for technical assistance to Pilar Fontanet in calli culture and to Patricio Gómez in DNA sequencing. We also thank Dr. Paulo Arruda (Unicamp, Brazil) for furnishing *Cat100-1* embryogenic calli. M.J.E. was a recipient of Postdoctoral Fellowships de Reincorporación a España del Plan de Formación de Personal Investigador and d'Investigació Científica a Catalunya, and L.R.A. was the recipient of a fellowship from Plan de Formación de Personal Investigador. The work has been supported by a grant from Plan Nacional de Investigación Científica y Técnica (Bio88-0242).

Received November 20, 1991; accepted January 28, 1992.

REFERENCES

- Averyhart-Fullard, V., Datta, K., and Marcus, A. (1988). A hydroxyproline-rich protein in the soybean cell wall. *Proc. Natl. Acad. Sci. USA* **85**, 1082–1085.
- Burr, B., and Burr, F.A. (1981). *Controlling-element events at the shrunken locus* in maize. *Genetics* **98**, 143–156.
- Caelles, C., Delseny, M., and Puigdomènech, P. (1992). The hydroxyproline-rich glycoprotein gene from *Oryza sativa*. *Plant Mol. Biol.* **18**, 617–619.
- Cassab, G.I., and Varner, J.E. (1988). Cell wall proteins. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **39**, 321–353.
- Chen, J., and Varner, J.E. (1985). An extracellular matrix protein in plants: Characterization of a genomic clone for carrot extensin. *EMBO J.* **4**, 2145–2151.
- Corbin, D.R., Sauer, N., and Lamb, C.J. (1987). Differential regulation of a hydroxyproline-rich glycoprotein gene family in wounded and infected plants. *Mol. Cell. Biol.* **7**, 4337–4344.
- Franssen, H.J., Nap, J.P., Gloudemans, T., Stiekema, W., van Dam, H., Govers, F., Louwerse, J., von Kammen, A., and Bisseling, T. (1987). Characterization of cDNA for nodulin-75 of soybean: A gene product involved in early stages of root nodule development. *Proc. Natl. Acad. Sci. USA* **84**, 4495–4499.
- Guiltinan, M.J., Marcotte, W.R., and Quatrano, R.S. (1990). A plant leucine zipper protein that recognizes an abscisic acid response element. *Science* **250**, 267–271.
- Henrissat, B., Popineau, Y., and Kader, J.C. (1988). Hydrophobic cluster analysis and lipid-transfer proteins. A domain homology between storage and lipid-transfer proteins. *Biochem. J.* **255**, 901–905.
- Hong, J.C., Nagao, R.T., and Key, J.L. (1987). Characterization and sequence analysis of a developmentally regulated putative cell wall protein gene isolated from soybean. *J. Biol. Chem.* **262**, 8367–8376.
- Hong, J.C., Nagao, R.T., and Key, J.L. (1989). Developmentally regulated expression of soybean proline-rich cell wall protein genes. *Plant Cell* **1**, 937–943.
- Hood, E.E., Shen, Q.X., and Varner, J.E. (1988). A developmentally regulated hydroxyproline-rich glycoprotein in maize pericarp cell walls. *Plant Physiol.* **87**, 138–142.
- Jones, R.J., and Brenner, M.L. (1987). Distribution of abscisic acid in maize kernel during grain filling. *Plant Physiol.* **83**, 905–909.
- Keller, B., and Lamb, C.J. (1989). Specific expression of a novel cell wall hydroxyproline-rich glycoprotein gene in lateral root initiation. *Genes Dev.* **3**, 1639–1646.
- Kieliszewski, M., and Lampport, D.T.A. (1987). Purification and partial characterization of a hydroxyproline-rich glycoprotein in a graminaceous monocot, *Zea mays*. *Plant Physiol.* **85**, 823–827.
- Kreis, M., and Shewry, P.R. (1989). Unusual features of cereal seed protein structure and evolution. *BioEssays* **10**, 201–207.
- Kyte, J., and Doolittle, R.F. (1982). A simple method for displaying the hydrophobic character of a protein. *J. Mol. Biol.* **157**, 105–132.
- Langdale, J.A., Rothermel, B.A., and Nelson, T. (1988). Cellular pattern of photosynthetic gene expression in developing maize leaves. *Genes Dev.* **2**, 106–115.
- Lehrach, H., Diamond, D., Wozney, J.M., and Boedtker, H. (1977). RNA molecular weight determination by gel electrophoresis under denaturing conditions, a critical reexamination. *Biochemistry* **16**, 4743–4751.
- Logemann, J., Schell, J., and Willmitzer, L. (1987). Improved method for the isolation of RNA from plant tissues. *Anal. Biochem.* **163**, 16–20.
- Ludevid, M.D., Ruiz-Avila, L., Vallés, M.P., Stiefel, V., Torrent, M., Torné, J.P., and Puigdomènech, P. (1990). Expression of genes for cell-wall proteins in dividing and wounded tissues of *Zea mays* L. *Planta* **180**, 524–529.
- Maxam, A.M., and Gilbert, W. (1980). Sequencing end-labelled DNA with base-specific chemical cleavages. *Methods Enzymol.* **65**, 499–560.
- Memelink, J., Hoge, J.H.C., and Schilperoort, R.A. (1987). Cytokinin stress changes the developmental regulation of several defence-related genes in tobacco. *EMBO J.* **6**, 3579–3583.
- Murray, N.E. (1983). Phage lambda and molecular cloning. In *Lambda II*, R.W. Hendrix, J.W. Roberts, F.W. Stahl, and R.A. Weisberg, eds (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory), pp. 395–432.

- Odani, S., Koide, T., Ono, T., Seto, Y., and Tanaka, T.** (1987). Soybean hydrophobic protein. Isolation, partial characterization and the complete primary structure. *Eur. J. Biochem.* **162**, 485–491.
- Pla, M., Goday, A., Vilardell, J., Gómez, J., and Pagès, M.** (1989). Differential regulation of ABA-induced 23–25 kDa proteins in embryo and vegetative tissues of the viviparous mutants of maize. *Plant Mol. Biol.* **13**, 385–394.
- Ponz, F., Paz-Ares, J., Hernández-Lucas, C., García-Olmedo, F., and Carbonero, P.** (1986). Cloning and nucleotide sequence of a cDNA encoding the precursor of the barley toxin α -hordothionin. *Eur. J. Biochem.* **156**, 131–135.
- Prat, S., Pérez-Grau, L., and Puigdomènech, P.** (1987). Multiple variability in the sequence of a family of maize endosperm proteins. *Gene* **52**, 41–49.
- Prioli, L.M., and Söndahl, M.R.** (1989). Plant regeneration and recovery of fertile plants from protoplasts of maize (*Zea mays* L.). *Biol Technology* **7**, 589–594.
- Quatrano, R.S.** (1987). The role of hormones during seed development. In *Plant Hormones and Their Role in Plant Growth Development*, P. Davies, ed (Boston: Martinus Nijhoff), pp. 494–514.
- Raz, R., Crétin, C., Puigdomènech, P., and Martínez-Izquierdo, J.A.** (1991). The sequence of a hydroxyproline-rich glycoprotein gene from *Sorghum vulgare*. *Plant Mol. Biol.* **16**, 365–367.
- Raz, R., José-Estanyol, M., Moya, A., Martínez-Izquierdo, J.A., and Puigdomènech, P.** (1992). Different mechanisms generating sequence variability are revealed in distinct regions of the hydroxyproline-rich glycoprotein gene from maize and related species. *Mol. Gen. Genet.* (in press).
- Ruiz-Avila, L., Ludevid, M.D., and Puigdomènech, P.** (1991). Differential expression of a hydroxyproline-rich cell-wall protein gene in embryonic tissues of *Zea mays* L. *Planta* **184**, 130–136.
- Ruiz-Avila, L., Burgess, S.R., Stiefel, V., Ludevid, M.D., and Puigdomènech, P.** (1992). Accumulation of cell wall hydroxyproline-rich glycoprotein mRNA is an early event in maize embryo cell differentiation. *Proc. Natl. Acad. Sci. USA* (in press).
- Salts, Y., Wachs, R., Gruissem, W., and Barg, R.** (1991). Sequence coding for a novel proline-rich protein preferentially expressed in young tomato fruit. *Plant Mol. Biol.* **17**, 149–150.
- Sambrook, J., Fritsch, E.F., and Maniatis, T.** (1989). *Molecular Cloning: A Laboratory Manual*. 2nd ed. (Cold Spring Harbor, NY: Cold Spring Harbor Press).
- Sanger, F., Nicklen, S., and Coulson, A.R.** (1977). DNA sequencing with chain-termination inhibitors. *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
- Sheng, J., D'Ovidio, R., and Mehdy, M.C.** (1991). Negative and positive regulation of a novel proline-rich protein mRNA by fungal elicitor and wounding. *Plant J.* **1**, 345–354.
- Sossountzov, L., Ruiz-Avila, L., Vignols, F., Jolliot, A., Arondel, V., Tchong, F., Grosbois, M., Guerbette, F., Miginiac, E., Delseny, M., Puigdomènech, P., and Kader, J.-C.** (1991). Spatial and temporal expression of a maize lipid transfer protein gene. *Plant Cell* **3**, 923–933.
- Southern, E.M.** (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **98**, 503–517.
- Stiefel, V., Pérez-Grau, L., Albericio, F., Giralt, E., Ruiz-Avila, L., Ludevid, M.D., and Puigdomènech, P.** (1988). Molecular cloning of cDNAs encoding a putative cell wall protein from *Zea mays* and immunological identification of related polypeptides. *Plant Mol. Biol.* **11**, 483–493.
- Stiefel, V., Ruiz-Avila, L., Raz, R., Vallés, M.P., Gómez, J., Pagès, M., Martínez-Izquierdo, J.A., Ludevid, M.D., Langdale, J.A., Nelson, T., and Puigdomènech, P.** (1990). Expression of a maize cell wall hydroxyproline-rich glycoprotein gene in early leaf and root vascular differentiation. *Plant Cell* **2**, 785–793.
- Torné, J.M., Santos, M.D., and Blanco, J.L.** (1984). Methods of obtaining maize totipotent tissues. II. Atrophic tissue culture. *Plant Sci. Lett.* **33**, 317–325.
- von Heijne, G.** (1983). Patterns of amino acids near signal-sequence cleavage sites. *Eur. J. Biochem.* **133**, 17–21.
- Ye, Z.-H., Song, Y.-R., Marcus, A., and Varner, J.E.** (1991). Comparative localization of three classes of cell wall proteins. *Plant J.* **1**, 175–183.