

Maize polyubiquitin genes: structure, thermal perturbation of expression and transcript splicing, and promoter activity following transfer to protoplasts by electroporation

Alan H. Christensen¹, Robert A. Sharrock² and Peter H. Quail*

UC-Berkeley/USDA Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA

(*author for correspondence); ¹Department of Biology, George Mason University, Fairfax VA 22030-4444, USA; ²Department of Biology, Montana State University, Bozeman MT 59717-0001, USA

Received 13 June 1991; accepted in revised form 25 October 1991

Key words: electroporation, heat shock, maize, promoter, splicing, ubiquitin

Abstract

Two genomic clones (λ Ubi-1 and λ Ubi-2) encoding the highly conserved 76 amino acid protein ubiquitin have been isolated from maize. Sequence analysis shows that both genes contain seven contiguous direct repeats of the protein coding region in a polypeptide conformation. The deduced amino acid sequence of all 14 repeats is identical and is the same as for other plant ubiquitins. The use of transcript-specific oligonucleotide probes shows that Ubi-1 and Ubi-2 are expressed constitutively at 25 °C but are inducible to higher levels at elevated temperatures in maize seedlings. Both genes contain an intron in the 5' untranslated region which is inefficiently processed following a brief, severe heat shock. The transcription start site of Ubi-1 has been determined and a transcriptional fusion of 0.9 kb of the 5' flanking region and the entire 5' untranslated sequence of Ubi-1 with the coding sequence of the gene encoding the reporter molecule chloramphenicol acetyl transferase (CAT) has been constructed (pUBI-CAT). CAT assays of extracts of protoplasts electroporated with this construct show that the ubiquitin gene fragment confers a high level of CAT expression in maize and other monocot protoplasts but not in protoplasts of the dicot tobacco. Expression from the Ubi-1 promoter of pUBI-CAT yields more than a 10-fold higher level of CAT activity in maize protoplasts than expression from the widely used cauliflower mosaic virus 35S promoter of a 35S-CAT construct. Conversely, in tobacco protoplasts CAT activity from transcription of pUBI-CAT is less than one tenth of the level from p35S-CAT.

Introduction

Ubiquitin is one of the most highly conserved proteins known, differing at only 4 of 76 positions between organisms as diverse as yeast, maize and man [6, 32]. The protein has been implicated in many vital cellular processes, such as pro-

tein turnover [23, 29], chromatin structure [3], cell cycle control [30], DNA repair [25], and response to heat shock and other stresses [4, 10, 15]. In addition, ubiquitin itself may have intrinsic proteolytic activity [16].

Ubiquitin is highly abundant in the cytoplasm as a 'free' monomer, but can also be covalently

attached to other proteins via a multistep enzymatic process [23, 29]. The linkage occurs at specific lysine residues of the acceptor protein with formation of an isopeptide bond between the carboxyl group of the C-terminal glycine of ubiquitin and ϵ -NH₂ groups of the lysine side-chains of the target protein [29]. The conjugation of ubiquitin to proteins may alter their stability or serve as a recognition signal for proteolysis by an ATP-dependent non-lysosomal pathway. In general, 'normal' short-lived cytoplasmic proteins are degraded by this pathway, as are abnormal or denatured proteins [23, 29, 30]. In plants, conjugation of ubiquitin to the regulatory photoreceptor phytochrome correlates with its 100-fold increased rate of degradation following photo-conversion from inactive to active conformations [24, 33]. On the other hand, not all ubiquitinated proteins are recognized as substrates for proteolysis. Stable ubiquitin conjugates can be formed with histones 2A and 2B [29] and actin [2]. In these cases, ubiquitination may serve as a mechanism to reversibly modify protein function.

Polyubiquitin genes have been isolated and sequenced from a wide range of organisms comprising eight different phyla [6, 32]. These genes consist of tandemhead-to-tail repeats of 228 bp encoding the ubiquitin protein. The number of tandem repeats varies between genes within a genome and between organisms, from 3 in *Dictyostelium* to approximately 50 in *Trypanosoma cruzi* [6, 32]. A second type of ubiquitin gene encodes a single ubiquitin repeat fused to one of two other polypeptides of either 52 or 76–80 amino acids [6, 32]. Recently, these fusion polypeptides were shown to be ribosomal proteins in yeast [14], mammals [31], and plants [9].

The multiple roles of ubiquitin in many important cellular functions suggests a complex regulation of ubiquitin levels. We have begun studies on the regulation of expression of ubiquitin genes from the agronomically important crop maize. Previously, we demonstrated by Southern analysis that ubiquitin is encoded by 8 to 10 genes in maize and by northern analysis that expression of these sequences produced three different sized

mRNAs of 2.1, 1.6, and 0.8 kb [10]. We report here the isolation and sequencing of two maize polyubiquitin genes (Ubi-1 and Ubi-2), thermal regulation of transcript levels and splicing, and Ubi-1 promoter activity following electroporation of a chimeric gene into protoplasts.

Materials and methods

Plant materials

Maize Inbred B73 seedlings were grown on wet Kim-pak in the dark at 26 °C for 4 to 5 days. Seedlings were heat-shocked by immersion for 10 min in water at 45 °C and then returned to 26 °C for up to 3 h before harvesting [10]. A Black Mexican Sweet (BMS) maize suspension culture cell line was obtained from F. Morrish and M. Fromm (Plant Gene Expression Center, Albany, CA). The BMS cell line was maintained as described [18] using MS medium containing (per liter) 4.3 g Murashige-Skoog (MS) salts (Gibco/BRL, Grand Island, NY), 20 g sucrose, 2 mg 2,4-dichlorophenoxyacetic acid, 0.2 mg myo-inositol, 0.13 g L-asparagine, 1.3 mg niacin, 0.25 mg thiamin, 0.25 mg pyridoxine, and 0.25 mg calcium pantothenate. The final solution was adjusted to pH 5.8.

Tobacco TXD suspension cell culture was obtained from E. Dale and D. Ow (Plant Gene Expression Center, Albany, CA). The cell line was maintained by subculturing 7 ml into 45 ml of fresh TXD medium every three days. The TXD medium contained (per liter) 4.3 g MS salts, 30 g sucrose, 2 mg p-chlorophenoxyacetic acid, 0.005 mg kinetin, 100 mg inositol, 10 mg thiamin, 1 mg niacin, and 1 mg pyridoxine at pH 5.7.

Isolation of genomic clones

Two λ Charon 35 libraries of a partial *Sau3AI* digest of maize genomic DNA were screened for clones containing ubiquitin sequences [27]. The first library was screened with nick-translated maize ubiquitin cDNA clone p6T7.2b1 [10]

in $6 \times \text{SSC}$ ($1 \times \text{SSC} = 0.15 \text{ M NaCl}$, 0.015 M sodium citrate), $1 \times \text{Denhardt's solution}$ [27], $100 \mu\text{g/ml}$ denatured salmon DNA, $100 \mu\text{g/ml}$ heparin sulfate (Sigma, St. Louis, MO), $10 \mu\text{g/ml}$ poly-adenylic acid, 1 mM EDTA , and 0.5% SDS. Filters were washed twice for 30 min at room temperature in $2 \times \text{SSC}$, 10 mM EDTA , and 0.5% SDS and once for 60 min at 65°C in $0.1 \times \text{SSC}$, 10 mM EDTA , and 0.5% SDS. A second library [35] was screened with an RNA probe generated by SP6 polymerase (Promega, Madison, WI) transcription of a $2.85 \text{ kb Eco RI-Hind III}$ fragment of $\lambda\text{Ubi-1}$. This fragment contains homology with the ubiquitin cDNA and was subcloned into pSP65 (Promega). Transcription reactions were carried out according to the manufacturer's recommendations. Hybridization was carried out in $100 \text{ mM sodium phosphate pH } 6.8$, $20 \text{ mM sodium pyrophosphate}$, 5 mM EDTA , 0.5% SDS, $500 \mu\text{g/ml}$ heparin sulfate, 10% dextran sulfate, $100 \mu\text{g/ml}$ denatured salmon DNA, $100 \mu\text{g/ml}$ yeast RNA at 65°C for 18 h. Filters were washed in $50 \text{ mM sodium phosphate pH } 6.8$, $5 \text{ mM sodium pyrophosphate}$, 5 mM EDTA , $50 \mu\text{g/ml}$ yeast RNA at 65°C for 60 min.

DNA sequencing

Pst I restriction fragments encompassing Ubi-1 and Ubi-2 were subcloned into M13mp19 in both orientations. Single-strand and replicative form phage DNAs were prepared by standard procedures [1]. Sets of unidirectional deletions were constructed by exonuclease III/exonuclease VII digestion [22]. Dideoxynucleotide-chain termination sequencing reactions were performed using Sequenase (United States Biochemical Corp.) according to the manufacturer's recommendations. The labeled nucleotide was [$\alpha\text{-}^{35}\text{S}$]-dATP ($>400 \text{ Ci/mmol}$; Amersham, Sterling Heights, IL). DNA and protein sequences were analyzed using programs from the University of Wisconsin Genetics Computer Group [13] and Textco (West Lebanon, NH).

S1 mapping of Ubi-1 gene

The transcription start site and the $3'$ border of the intron in the $5'$ untranslated region of Ubi-1 were determined by S1 nuclease protection. A $946 \text{ bp Pst I-Bgl II}$ fragment and a $643 \text{ bp Eco RI-Xho I}$ fragment labeled at the *Bgl* II and *Xho* I sites, respectively, with T4 polynucleotide kinase (New England Biolabs, Beverly, MA) and [$\gamma\text{-}^{32}\text{P}$]ATP ($>5000 \text{ Ci/mol}$; DuPont/New England Nuclear, Wilmington, DE) were isolated from polyacrylamide gels [1]. The end-labeled DNA fragment ($10\text{--}20 \text{ fmol}$) was hybridized to $2 \mu\text{g}$ of poly(A)⁺ RNA in $30 \mu\text{l}$ of S1 hybridization buffer [27]. The mixture was heated to 80°C for 15 min and then incubated at 42°C for 16 h. Unprotected nucleic acids were digested with S1 nuclease (250 U/ml ; DuPont-NEN) for 30 min at 37°C as described [27]. Protected fragments were separated on a 6% polyacrylamide-8 M urea gel and visualized by autoradiography. The end points of the protected fragments in the Ubi-1 sequence were determined by comparison with a Maxam-Gilbert [1, 27] sequence ladder generated from the end-labeled fragment used as the probe.

RNA isolation and northern analysis

RNA was isolated from 5-day old unshocked or heat-shocked maize seedlings using the guanidine thiocyanate procedure as described [10]. RNA was size-fractionated on 1.5% agarose/ 3% formaldehyde gels and transferred to GeneScreen (DuPont-NEN). Filters were hybridized with RNA probes in 50% formamide, $5 \times \text{SSC}$, $40 \text{ mM sodium phosphate pH } 6.8$, $100 \mu\text{g/ml}$ denatured salmon DNA, $20 \mu\text{g/ml}$ polyadenylic acid and 10% dextran sulfate at 65°C for 18 h. Filters were washed twice at room temperature for 30 min in $2 \times \text{SSC}$, 0.5% SDS and once at 65°C for 1 h in $0.1 \times \text{SSC}$, 0.5% SDS. Filters probed with ^{32}P -labeled oligonucleotides were hybridized in $6 \times \text{SSC}$, $10 \times \text{Denhardt's solution}$, and $200 \mu\text{g/ml}$ tRNA at 42°C for 20 h and washed for 30 min in $6 \times \text{SSC}$ at room temperature and then at

37 °C. Autoradiography was carried out at -80 °C with Kodak X-OMAT AR film and one intensifying screen. A dilution series of RNA was run on the northern blots to permit estimation of the fold increases of transcript abundance.

Preparation of northern analysis probes

RNA probes for northern analysis were generated by SP6 polymerase transcription of linearized templates. Transcription reactions were carried out according to the manufacturer's recommendations. The templates used were the sense strands of the maize ubiquitin cDNA insert [10] subcloned into pSP64 (Promega Biotec), the 0.6 kb *Eco* RI-*Pst* I fragment of Ubi-1 intron (Fig. 1A) and the 1.5 kbp *Pst* I-*Pst* I fragment of Ubi-2 intron (Fig. 1B), both subcloned into pGEM3 (Promega Biotec). Two 18-base oligonucleotides were synthesized using an Applied Biosystems Model 380B DNA synthesizer and were labeled with [γ -³²P]ATP (> 5000 Ci/mmol; DuPont/NEN) and T4 polynucleotide kinase (New England Biolabs). The oligonucleotides have the sequences 5'-GCAGATACTTTGACCAACC-3' for Ubi-1 and 5'-AGATACCCAGGTACAGCA-3' for Ubi-2 and are complementary to nucleotides 2725 to 2774 of Ubi-1 sequence (Fig. 2A) and 3273 to 3290 of Ubi-2 sequence (Fig. 2B).

Construction of pUBI-CAT

A ubiquitin promoter-CAT fusion gene was constructed by subcloning the *Bam* HI-*Hind* III restriction fragment containing the CAT gene and the nopaline synthase 3' untranslated region and polyadenylation signals of pNOS-CAT [17] into the *Bam* HI and *Hind* III sites of pUC18. This construct was termed pUC18-CAT. The 2 kb *Pst* I fragment immediately upstream of the polypeptide coding region of Ubi-1 was subcloned into M13mp19 for sequencing (see above). The fragment was removed from vector DNA by digestion with *Pst* I, blunt-ended by treating with

T4 DNA polymerase (New England Biolabs) and cloned into the *Sma* I site of pUC18-CAT. This construct was termed pUBI-CAT.

Protoplast preparation and electroporation

Maize protoplasts were prepared from 2-day old BMS cell cultures (40 ml). Cells were pelleted at 200 × g, washed once with digestion buffer (3 mM MES pH 5.7, 10 mM CaCl₂, 0.5 M mannitol, and 2 mg arginine) and incubated for 2 h with 40 ml of digestion buffer containing 2% cellulase (Onozuka RS, Yakult Honsha Co., Tokyo, Japan). After cell wall digestion, the protoplasts were filtered through a 62 µm nylon mesh (Small Parts, Inc., Miami, FL) washed once with digestion buffer, and once with electroporation buffer (10 mM HEPES pH 7.2, 150 mM NaCl, 4 mM CaCl₂, and 0.5 M mannitol).

Maize protoplasts (ca. 4 × 10⁶/ml) were mixed with 40 µg of plasmid and 100 µg of carrier DNA and electroporated on ice with a 250 V, 70 ms pulse. After incubation for 10 min on ice, the protoplasts were diluted into 4 ml of MS medium containing 0.5 M mannitol and incubated at 26 °C in the dark for 24 h.

Tobacco protoplasts were prepared from 3- to 4-day old cultures (40 ml) of TXD cells. Cells were pelleted at 200 × g and washed once with protoplast isolation medium (PIM) containing 0.2 M mannitol, 50 mM CaCl₂, 10 mM sodium acetate pH 5.8. Cell wall digestion was carried out for 4 h at 26 °C in medium containing: 0.5% cellulase, 0.5% hemicellulase (Sigma), 0.02% Pectinase Y-23 (Seishim Pharmaceutical, Tokyo, Japan), 0.5% bovine serum albumin, and 7 mM 2-mercaptoethanol. Protoplasts were pelleted by centrifugation at 100–200 × g, washed once in electroporation buffer (as for BMS protoplasts, except with 0.2 M mannitol) and then resuspended in electroporation buffer.

Tobacco protoplasts (1–2 × 10⁶/ml) were mixed with 20 µg of plasmid and 150 µg of carrier DNA and electroporated with a 200V, 50 ms pulse. The electroporated protoplasts were incubated on ice for 10 min, diluted into 10 ml TXD

medium plus 0.2 M mannitol and incubated as for the BMS protoplasts.

Extracts of electroporated protoplasts were prepared essentially as described [18]. Protoplasts were pelleted at 50–100 × *g*, washed once with MS medium plus 0.5 M mannitol, resuspended in CAT assay buffer (0.25 M Tris pH 7.8, 1 mM EDTA, 1 mM 2-mercaptoethanol, and 2 mM phenylmethylsulfonyl fluoride) and disrupted by sonication. Debris was pelleted by centrifugation and the supernatant was heated at 65 °C for 10 min and stored at –80 °C. Protein concentration in the extracts was determined by the Bradford method [1] (BioRad, Richmond, CA).

Assay of CAT activity

The CAT assay mixture contained 100 µl of protoplast extract, 1.6 µl [¹⁴C]-chloramphenicol (40–60 Ci/mol; Amersham) and 20 mg/ml of acetyl coenzyme A (Sigma) in CAT assay buffer in a final volume of 200 µl. Reactions were incubated for 90 min at 37 °C. Chloramphenicol and its acetylated derivatives were extracted with cold ethyl acetate and analyzed by thin-layer chromatography run with chloroform/methanol (95:5) as described [18]. After autoradiography of the chromatogram, regions of the chromatogram corresponding to chloramphenicol and its acetylated derivatives were excised and the amount of radioactivity in each was determined by liquid scintillation counting. Activity is expressed as per cent of chloramphenicol converted to its acetylated forms per equivalent amount of protein.

Results

Isolation of polyubiquitin genomic clones

The previously isolated maize ubiquitin cDNA clone p6T7.2b1 [10] was used to screen a maize genomic library (about 1 genome equivalent) in λCharon 35. One clone (λUbi-1) containing a 17 kb insert homologous with the ubiquitin cDNA

was isolated (Fig. 1A). An SP6 polymerase-generated RNA probe containing ubiquitin sequences was then used to screen a second genomic library (ca. 8 genome equivalents) [35]. The template for preparation of this probe was made by subcloning a 2.85 kb *Eco* RI–*Hind* III fragment of λUbi-1 (Fig. 1A) containing the ubiquitin coding sequences into pSP65. Of an additional 20 ubiquitin-containing genomic clones that were isolated, three overlapped with λUbi-1, as determined by restriction analysis (data not shown). Characterization of several of the remaining clones revealed a different polyubiquitin genomic clone (λUbi-2; Fig. 1B) and an overlapping clone (data not shown). The remaining 15 ubiquitin genomic clones may contain DNA

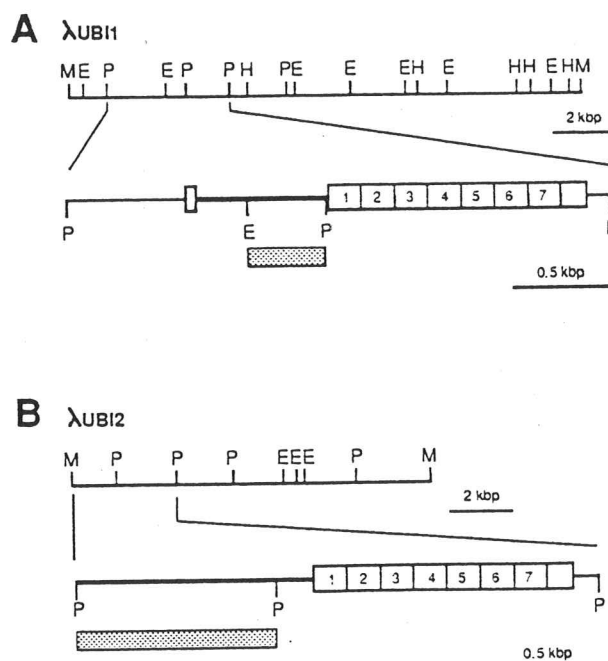


Fig. 1. Physical map and organization of two maize polyubiquitin genes. A. λUbi-1. B. λUbi-2. Restriction map of the insert of each genomic clone is shown at the top. Below the maps are the two *Pst* I fragments of each gene which were subcloned and sequenced and a schematic of the gene structure. Numbered boxes represent ubiquitin coding regions; other boxes are 5' and 3' untranslated regions. Introns are indicated by thick lines and flanking DNA is indicated by thin lines. Hatched boxes are restriction fragments used to generate intron-specific probes. E, *Eco* RI; P, *Pst* I; H, *Hind* III; M, multiple cloning site of λCharon 35.

sequences overlapping only a small portion of λ Ubi-1 or λ Ubi-2. Alternatively, the clones may encode genes for the 1.6 kb transcript (see Fig. 4A) [10] or maize homologues of ubiquitin-ribosomal protein fusion genes as found in other organisms [6, 32], or possibly are pseudogenes.

Sequence and structure of Ubi-1 and Ubi-2

For both Ubi-1 and Ubi-2, the *Pst* I fragments containing the region of homology with the ubiquitin cDNA clone [10] and the *Pst* I fragment immediately upstream were subcloned in both orientations into M13mp19. Nested sets of unidirectional deletions were constructed [22] for each strand and sequenced. The resulting nucleotide and derived amino acid sequences are shown in Fig. 2. Comparison of the sequence of a previously characterized maize ubiquitin cDNA [10] with the sequence of Ubi-1 and Ubi-2 revealed that the cDNA is derived from a transcript of Ubi-2. The cDNA sequence corresponds to nucleotides 2433 to 3388 (Fig. 2B).

Both genes contain an open reading frame of 1599 bp arranged as seven tandem, head-to-tail repeats of 228 bp encoding ubiquitin (Fig. 1; Fig. 2). The terminal repeat of both genes contains an additional, 77th amino acid, glutamine (Q). This additional amino acid is not conserved among ubiquitin genes, nor is it found in mature ubiquitin. It is presumably removed during processing of the polyubiquitin protein to the mature monomers. The sequence of all 14 protein coding repeats of Ubi-1 and Ubi-2 is identical and is the same as for barley [19] and *Arabidopsis* [5], but differs from yeast and animal ubiquitin [6, 32].

The nucleotide sequences of the coding regions of Ubi-1 and Ubi-2 genes reflects the conserva-

tion of the protein sequence. A comparison of the two sequences reveals greater than 90% identity in the protein coding region. The sequence conservation extends into the 3' untranslated region also. The two sequences are greater than 80% similar for more than 150 bp extending past the stop codon. Polyadenylation signals are present in these regions for both genes: at 2808 in the Ubi-1 sequence (Fig. 2A), and at 3339 in the Ubi-2 sequence (Fig. 2B). About 200 nucleotides immediately upstream of the ubiquitin coding region are also highly conserved between the two genes. In addition, a short region of Ubi-1 is repeated 11 times in Ubi-2. These 11 repeats are more than 90% identical, are about 140 bp long, and are tandemly arranged, like the protein coding region. This sequence is present at low copy number in the maize genome (data not shown). The Ubi-1 gene has a unique tandem repeat of about 220 bp beginning at 429 and 649 (Fig. 2A). These two repeats are also highly conserved (>85%). The functional significance of these tandem repeats in ubiquitin gene expression is unclear.

S1 mapping of the Ubi-1 gene was used to define the transcription start site. Initially, a DNA fragment end-labeled in the first repeat of the coding region +7 nucleotides downstream of the ATG (Fig. 3C) was hybridized to poly(A)⁺ RNA. Digestion with S1 nuclease resulted in a protected fragment of about 52 bases, extending just past the translation start (Fig. 3A). Examination of the sequence in this region revealed a consensus 3' splice junction, TGCAG, suggesting an intron in the 5' untranslated region. Further analysis using Ubi-1 fragments as probes for northern blot hybridization localized an exon about 1 kb upstream of the translation start (data not shown). S1 nuclease protection experiments were performed with a DNA fragment end-labeled at the *Bgl* II

Fig. 2. Nucleotide sequence of maize Ubi-1 (A) and Ubi-2 (B). A. The deduced amino acid sequence is shown underneath in single letter code. Methionines at the start of each of the seven ubiquitin protein repeats are boxed. The transcription start site is designated +1. The TATA box at -30 and a polyadenylation signal at +2808 are underlined. Two overlapping sequences at -214 and -204 which are similar to the consensus heat shock element are also underlined. The sequence at 2757 to 2774 complementary to the Ubi-1 gene-specific oligonucleotide probe is overlined. Restriction sites used to generate DNA fragments for production of intron-specific and S1 analysis probes are indicated. B. Sequence is numbered starting with the 5'-most base of Ubi-2 intron. Polyadenylation signal at 3339 is underlined and the sequence (3273-3290) complementary to the Ubi-2 gene-specific oligonucleotide probe is overlined.

2901 TATGTTCTGTCTTTCAGTTGTCTCCTAATATTTCCTGCAG 2941

[illegible]

Fig. 2B. See Fig. 2 for legend.

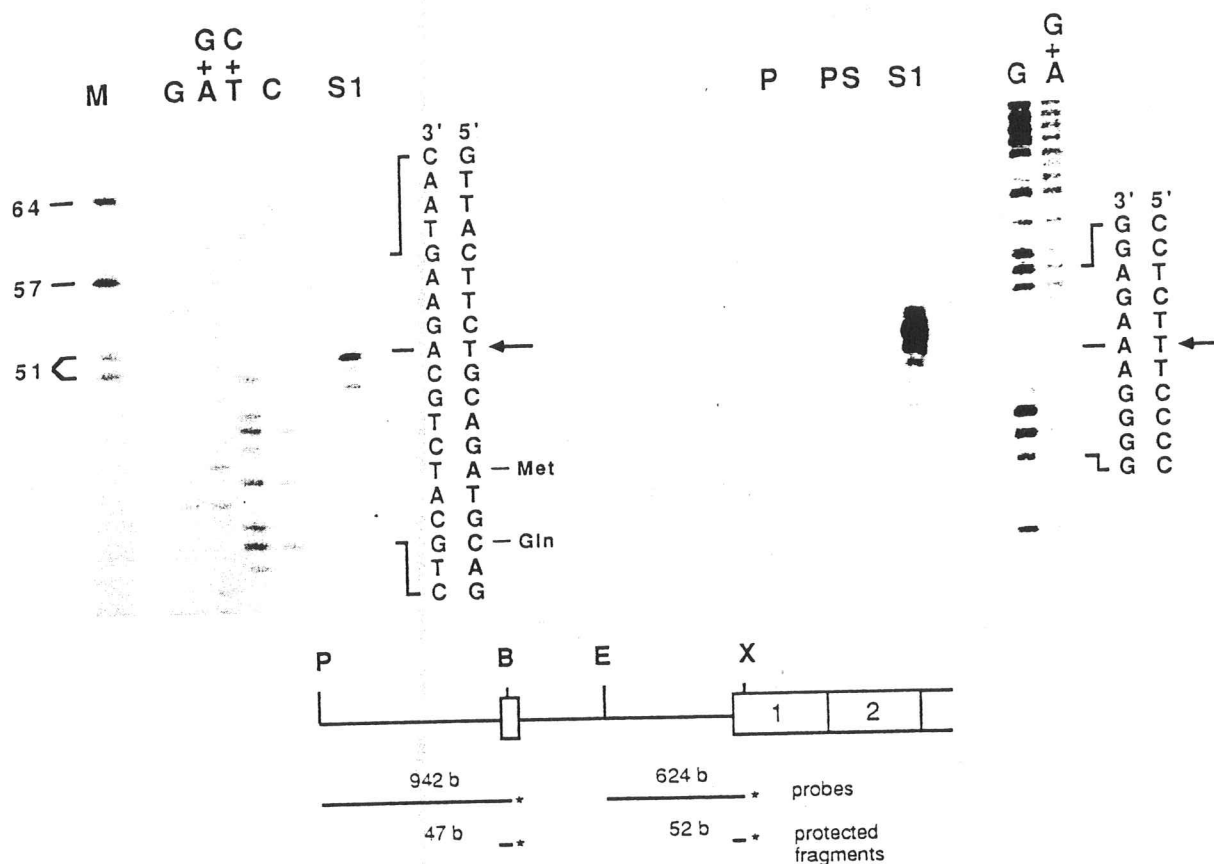


Fig. 3. S1 nuclease protection analysis of Ubi-1 mRNA. A (top left). Mapping the 3' end of the intron in the 5' untranslated region. The probe used was a 642 base *Eco* RI-*Xho* I fragment 5' end-labeled at the *Xho* I site 43 bp downstream of the ATG of the first ubiquitin-coding repeat. Only the portion of the gel with protected fragments is shown. M, markers (32 P-labeled Φ X174 *Hae* III fragments); G, G + A, C + T, and C lanes are Maxam and Gilbert sequencing reactions of the labeled probe fragment. S1, protected fragments from 2 μ g of poly(A)⁻ RNA hybridized with the probe and digested with S1 nuclease. B (top right). Mapping the transcription start point of Ubi-1 gene. The probe used was a 942 base *Pst* I-*Bgl* II fragment 5' end labeled at the *Bgl* II site 1061 bp 5' of the ATG of the first ubiquitin coding repeat. Only the portion of the gel with protected fragments is shown. P, probe only; PS, probe plus S1 nuclease; S1, as for A. C (bottom). Schematic representation of Ubi-1 gene structure showing relevant restriction enzyme sites, and the end-labeled probes and protected fragments. P, *Pst* I; B, *Bgl* II; E, *Eco* RI; X, *Xho* I.

site and extending 5' to the *Pst* I site (Fig. 3C). Fragments of about 47 bases were protected and the T residue 47 bases upstream of the end-label in the *Bgl* II site was designated as the transcription start site (+1; Fig. 3B). A bestfit consensus 5' splice junction (AGGT) is located at +82. This, plus a size estimate from northern analysis of about 1 kb for the intron, yields a Ubi-1 gene structure consisting of a 5' exon of about 82 bp, an intron of about 1010 bp, a coding region of 1599 bp and about 140 bp of 3' untranslated region. The structure of Ubi-2 gene is less clear. The

coding region and 3' untranslated regions are similar in size to Ubi-1. However, the Ubi-2 gene apparently contains a larger intron in the 5' untranslated region than is found in Ubi-1. This intron is about 3.1 kb (Fig. 4 and below) about half of which is located on λ Ubi-2. The high degree of sequence similarity immediately upstream of the polyubiquitin coding regions in Ubi-1 and Ubi-2 (ca. 80% for about 200 bp) suggests that the Ubi-2 intron is located in a position analogous to the Ubi-1 intron with a 3' splice junction immediately upstream of the translation start site.

The size of the 5' exon of Ubi-2 is unknown at present.

Several sequence elements common to promoters of many other eukaryotic genes are also found in the Ubi-1 5' flanking sequence. A TATA box sequence is found at -30 and two overlapping sequences at -214 and -204 are similar to the consensus heat shock element found in heat inducible genes [28]. This implies that Ubi-1 may be one of the ubiquitin genes transcriptionally induced by heat shock of maize seedlings [10]. Conversely, the Ubi-1 5' flanking sequence does not contain a canonical CCAAT sequence.

There are two other interesting sequence motifs present in the Ubi-1 promoter. A 15 base stretch of pyrimidines is located at -110 to -96 and a second, shorter segment of pyrimidines at -51 to -39. Regions of alternating C and T residues are present in promoters of several *Drosophila his* and *hsp* genes and specifically bind a 66 kDA nuclear protein [20]. The sequence CACGGCA is present 4 times, at -236, -122, -96, and -91. Whether these sequence motifs are functional in Ubi-1 gene expression is currently being investigated.

Response of Ubi-1 and Ubi-2 transcripts to heat shock

A brief severe heat shock of maize seedlings produces a 4- to 5-fold increase in the level of 2.1 kb ubiquitin transcripts [10]. In addition, two higher-molecular-weight transcripts of 3.1 kb and 5.2 kb are detected in RNA from heat-shocked tissue and are most abundant 30 min after heat shock (Fig. 4A), decreasing to undetectable levels within 3 to 4 h [10].

The presence of introns in the 5' untranslated region of both Ubi-1 and Ubi-2, along with reports of partial inhibition of splicing by a brief, severe heat shock [37] suggests that the 3.1 kb and 5.2 kb ubiquitin RNAs may not have arisen from newly transcribed genes, but rather are unprocessed transcripts. To test this hypothesis, DNA fragments specific for the intron of Ubi-1 and Ubi-2 (see Fig. 1A, B) were subcloned and

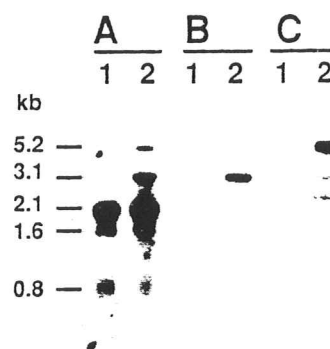


Fig. 4. Northern analysis of total RNA from unshocked maize seedlings (lanes 1) or heat-shocked seedlings allowed to recover for 30 min (lanes 2). The RNAs (2 µg per lane) were size-fractionated on a 1.5% agarose/3% formaldehyde gel, transferred to GeneScreen and hybridized with antisense RNA probes generated by SP6 polymerase transcription of the maize ubiquitin cDNA clone (A), the *Eco* RI-*Pst* I fragment of the Ubi-1 intron (B), and the *Pst* I-*Pst* I fragment of the Ubi-2 intron (C). The filters were washed at high stringency as described in Materials and methods.

used to generate probes for northern analysis. Hybridization with the Ubi-1 intron probe did not detect any transcripts in control (unshocked) RNA (Fig. 4B, lane 1), but hybridized with a transcript in heat-shocked RNA (Fig. 4B, lane 2) which comigrated with the 3.1 kb RNA detected in heat-shocked RNA probed with the ubiquitin cDNA (Fig. 4A, lane 2). Similarly, the Ubi-2 intron probe did not detect any transcripts in control RNA (Fig. 4C, lane 1), but hybridized primarily with a 5.2 kb transcript and to a lesser extent with two smaller RNAs. These results clearly demonstrate that the 3.1 kb and 5.1 kb RNAs represent unspliced transcripts of the Ubi-1 and Ubi-2 genes.

Previously, we observed a 4- to 5-fold increase in steady-state level of the 2.1 kb transcript by 3 h after a 10 min 45 °C heat shock [10]. To test whether the increased abundance was due to higher levels of transcripts of Ubi-1, Ubi-2 or both genes, northern blots of heat-shocked RNA were hybridized with transcript-specific probes. The ubiquitin coding sequence is >85% conserved between the two genes; the 3' untranslated regions of the two genes are nearly as highly conserved (data not shown). However, a region of

18 bases in the 3' untranslated region of each gene was found which differed by at least 8 bases with any other 18 base sequence in either gene. Oligonucleotides complementary to the sense strand of this region will hybridize to the 2.1 kb mature transcript, but will also hybridize to the unprocessed transcript from the respective gene. The results of the hybridization are shown in Fig. 5. The 2.1 kb transcript detected by the Ubi-1 oligonucleotide increases about 4-fold within 3 h after heat shock (Fig. 5A, lane 3). In addition the Ubi-1 oligonucleotide probe hybridizes with a 3.1 kb transcript in heat-shocked RNA, but not in control RNA. Similarly, a 4-fold increase in a 2.1 kb transcript is observed in heat-shocked RNA hybridized with the Ubi-2 oligonucleotide (Fig. 5B). In addition to the 2.1 kb transcript, this oligonucleotide hybridizes to the 5.2 kb unspliced RNA of Ubi-2. The results demonstrate that each oligonucleotide probe was specific for its respective transcript, and that both Ubi-1 and Ubi-2 transcripts are coordinately induced by heat shock.

Expression of a chimeric ubiquitin-CAT gene in electroporated protoplasts

A chimeric gene was constructed from the Ubi-1 promoter and 5' untranslated sequence (including intron) by fusion to the coding region of a gene

encoding CAT and the 3' untranslated and polyadenylation signals of a nopaline synthase gene. Expression of CAT activity from the Ubi-1 promoter was measured in protoplasts electroporated with this plasmid (pUBI-CAT). High levels of CAT activity were measured in BMS protoplasts electroporated with the pUBI-CAT construct (Fig. 6A, lane 1). CAT activity was detectable within 3 h after electroporation and increased over 15 h (data not shown). Negligible CAT activity was observed in protoplasts electroporated with either the promoter-less CAT construct pGP229 (lane 3), or with no plasmid DNA (lane 4), demonstrating that the measured CAT activity arises due to transcription from the Ubi-1 promoter. We have also measured CAT activity arising from pUBI-CAT in electroporated mesophyll protoplasts derived from maize, barley, and oat leaves (data not shown).

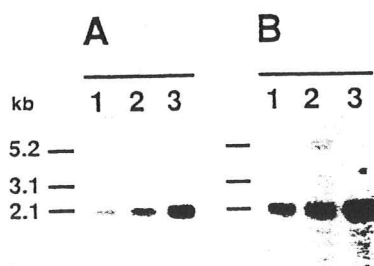


Fig. 5. Northern analysis of Ubi-1 and Ubi-2 transcripts following heat shock of maize seedlings. Total RNA (20 μ g per lane) was treated as in Fig. 4. Filters were hybridized with 32 P-labeled oligonucleotides specific for Ubi-1 (A) or Ubi-2 (B). RNA was isolated from unshocked seedlings (lane 1) or heat-shocked seedlings allowed to recover for 0.5 h (lane 2) or 3 h (lane 3).

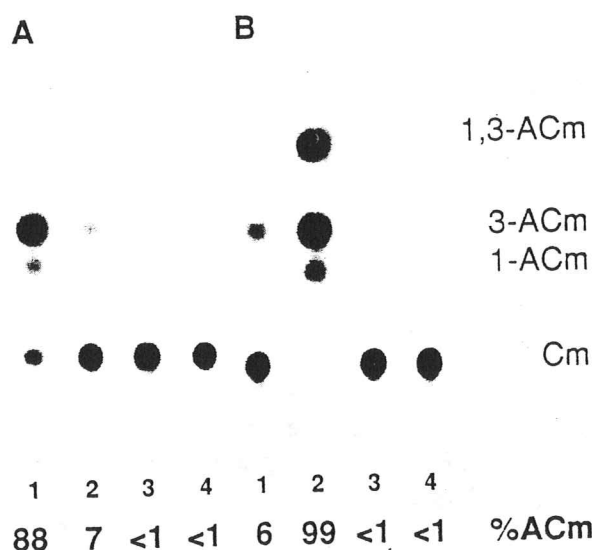


Fig. 6. CAT gene expression in electroporated plant protoplasts. Protoplasts generated from maize BMS (A) or tobacco TXD (B) suspension cells were electroporated with plasmid and carrier DNA as described in Materials and methods. Protoplasts were electroporated with pUBI-CAT (lanes 1); p35S-CAT (lanes 2); pGP229, a promoterless CAT construct (lanes 3); and no DNA (lanes 4). Cm, chloramphenicol; 1-ACm, 3-ACm, and 1,3-ACm are mono- and di-acetylated forms of chloramphenicol. %ACm is the percent conversion of chloramphenicol to its acetylated forms. Data are averages of assays of protein extracts from replicate electroporations of protoplasts from two separate preparations. The standard error of the mean for all values was less than 2.

The level of CAT activity in extracts of pUBI-CAT electroporated BMS protoplasts was compared with levels in extracts of protoplasts electroporated with p35S-CAT (Fig. 6A, tracks 1, 2). A more than 10-fold higher level of CAT activity was measured in the former than in the latter. On the other hand, in extracts of electroporated tobacco protoplasts, the CAT activity derived from transcription of pUBI-CAT was less than one-tenth of that from p35S-CAT (Fig. 6B, tracks 1, 2). Thus, the maize Ubi-1 promoter drives more than 10-fold higher expression of the CAT gene in maize protoplasts than the cauliflower mosaic virus 35S promoter, whereas the converse is true in tobacco protoplasts.

Discussion

We report here the sequence and structure of two of the 8 to 10 ubiquitin genes of maize as estimated from Southern blot analysis [10]. Both genes are polyubiquitin genes containing seven repeats encoding an identical 76 amino acid protein (Fig. 1, 2). Similar polyubiquitin genes are found in other eukaryotic organisms, although the number of tandem repeats varies widely [6, 32]. The tandem repeats of ubiquitin coding regions in polyubiquitin genes have been proposed to arise from unequal cross-over events [34]. The 2 and 11 tandem repeats found in the introns of both Ubi-1 and Ubi-2, respectively, may have been produced by the same (or similar) events. Whether additional repeats are present in the region of the Ubi-2 intron yet to be characterized remains to be determined. At present, about half of the Ubi-2 intron sequence has been determined, based on the estimate of the intron size obtained from the difference in size of the unspliced and mature Ubi-2 transcripts. The possible functional significance of these repeats in stabilizing ubiquitin transcripts or altering gene expression is at present unclear.

Previously, a 4- to 5-fold increase in ubiquitin gene expression following heat shock of maize seedlings was demonstrated by run-on transcription [10]. Although a cDNA for the Ubi-2 gene

was used as the probe in these experiments, the nucleotide sequences of ubiquitin genes are highly conserved. Thus, run-on transcription analysis measures expression of all ubiquitin genes, and not of an individual gene. Northern analysis using transcript-specific oligonucleotide probes for Ubi-1 and Ubi-2 genes demonstrates that mRNAs for both genes increase in abundance following heat shock. Thus, at least part of the increased ubiquitin gene transcription observed in the run-on transcription experiments may be a result of increased expression of both Ubi-1 and Ubi-2 genes.

Analysis of RNA from tissue harvested after a brief, severe heat shock using probes specific for the Ubi-1 and Ubi-2 introns indicates that in maize, as in *Drosophila*, splicing of precursor RNAs can be disrupted by heat shock. This contrasts to previous reports of inhibition of intron processing in plants. Intron processing of primary transcripts of a soybean heat shock gene and a petunia *hsp70* gene was unaffected by heat shock, but was disrupted by treatment of plants with CdCl₂ [11, 36]. However, in these experiments, the conditions of heat shock may not have been severe enough to affect splicing. The effect of CdCl₂ on the processing of Ubi-1 and Ubi-2 transcripts in maize has not yet been investigated. An alternative, albeit less likely, explanation for the appearance of unspliced transcripts following heat shock is that increased ubiquitin transcription overloads the splicing machinery. Further experiments are required to resolve which explanation is accurate.

High levels of CAT activity were detected in extracts of maize protoplasts electroporated with the pUBI-CAT construct. This indicates that sufficient sequence elements required for high-level expression of coding sequences located downstream are present in the 900 bp of Ubi-1 5' flanking DNA contained in this construct. Callis *et al.* [9] have reported expression of ubiquitin promoters in a heterologous system. β -Glucuronidase was expressed in transgenic tobacco from promoters of *Arabidopsis* ubiquitin-extension protein genes in contrast to the polyubiquitin gene promoter as described here. Expression from the

ubiquitin-extension protein gene promoters was constitutive in the transgenic tobacco, but these promoters were not tested in a monocot.

We obtained CAT activity at least 10-fold higher in maize protoplasts electroporated with UBI-CAT as compared with similar protoplasts treated with p35S-CAT (Fig. 6A). This is a considerably higher relative activity than was observed with several other promoters recently tested in transient gene expression assays in rice [12]. However, the converse was observed in electroporated tobacco protoplasts (Fig. 6B). There are several possible explanations for this different relative activity in maize and tobacco protoplasts. The maize Ubi-1 promoter may be monocot-specific, that is, it is highly expressed in maize and other monocots, but requires a *trans*-acting factor(s) not present in a dicot such as tobacco. Alternatively, the intron in the 5' untranslated region of the transcription unit of UBI-CAT may have contrasting effects on expression in the monocot and dicot cells. It may increase CAT expression in electroporated BMS protoplasts similar to results obtained with the *adh1* intron [7], whereas in tobacco protoplasts, it may not be effectively spliced from the chimeric ubiquitin-CAT transcript and may result in the lower levels of CAT activity. Intron-exon splice junctions are different between monocots and dicots. Monocot junctions are predominately pyrimidine rich in the 3' splice junction, while dicot 3' splice sequences are higher in purine bases [21]. In one case where processing of a monocot intron in a dicotyledonous plant was examined, the first intron of a wheat *rbcS* gene was inefficiently processed in transgenic tobacco [26]. The sizes of the chimeric ubiquitin-CAT transcripts have not been determined in either electroporated maize BMS or tobacco TXD protoplasts. Also, a chimeric construct lacking the Ubi-1 intron has not been tested in the transient expression assay system.

Callis *et al.* [8] have demonstrated that BMS protoplasts respond to thermal stress. Synthesis of heat shock proteins as well as expression of a chimeric maize *hsp70*-CAT gene are induced by incubation of protoplasts at 40 to 42 °C for 17 h.

We did not observe a heat shock-induced increase in CAT activity in electroporated protoplasts. However, the presence of heat shock elements at -214 and -204 in the Ubi-1 promoter is consistent with increased expression of this gene by heat shock. Previously, we reported a transient 4- to 5-fold increase in ubiquitin gene transcription in heat shocked maize B73 seedlings and in the steady-state level of the 2.1 kb mRNA [10]. Northern analysis using an oligonucleotide specific for Ubi-1 mRNA supports the conclusion that Ubi-1 is a heat-inducible gene (Fig. 5). The heat-induced increase in Ubi-1 mRNA levels may be masked in protoplasts because levels of the 2.1 kb mRNA appear to represent a greater fraction of total RNA in protoplasts than in suspension cells. Thus, the ubiquitin genes may already be fully induced due to cellular stresses such as oxidative damage, osmotic stress, or anaerobiosis occurring during cell wall digestion, electroporation and/or incubation.

Results presented here demonstrate that the maize Ubi-1 5' flanking sequence contains an efficient promoter for constitutive expression in maize protoplasts. Currently, the CaMV 35S promoter is widely used in monocot transformation studies. The Ubi-1 promoter described here expresses CAT activity at a 10-fold higher level than a 35S-CAT construct in electroporated maize protoplasts (Fig. 6). Thus, this promoter might be beneficial for generating a high level of expression of selectable marker genes necessary for generation of transgenic monocots. Also, since ubiquitin has a role in many important cellular functions, it is likely to be expressed in all tissues, although this remains to be determined for the Ubi-1 gene. Therefore, the Ubi-1 promoter system described here should prove useful for facilitating transformation of maize and other agronomically important cereals.

Acknowledgements

We thank F. Morrish and M. Fromm for the BMS suspension cell culture, E. Dale and D. Ow for the tobacco TXD cells and help with the to-

bacco protoplast electroporations, J. Slightom for a maize genomic library, W. Bruce and C. Gatz for the plasmids pGP229 and p35S-CAT, respectively, and other members of the laboratory for many valuable discussions. This work was supported by grants from Agrigenetics Research Associates and the NIH (GM36381).

References

1. Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K: Current Protocols in Molecular Biology. John Wiley, New York (1988).
2. Ball E, Karlik CC, Beall CJ, Saville DL, Sparrow JC, Bullard B, Fyrberg EA: Arthrin, a myofibrillar protein of insect flight muscle is an actin-ubiquitin conjugate. *Cell* 51: 221-228 (1987).
3. Barsoum J, Varshavsky A: Preferential localization of variant nucleosomes near the 5' end of the mouse dihydrofolate reductase gene. *J Biol Chem* 260: 7688-7697 (1985).
4. Bond U, Schlesinger MJ: The chicken ubiquitin gene contains a heat shock promoter and expresses an unstable mRNA in heat-shocked cells. *Mol Cell Biol* 6: 4602-4610 (1986).
5. Burke TJ, Callis J, Vierstra RD: Characterization of a polyubiquitin gene from *Arabidopsis thaliana*. *Mol Gen Genet* 213: 435-443 (1988).
6. Callis J, Vierstra RD: Ubiquitin and ubiquitin genes in higher plants. *Oxford Surv Plant Mol Cell Biol* 6: 1-30 (1989).
7. Callis J, Fromm M, Walbot V: Introns increase gene expression in cultured maize cells. *Genes Devel* 1: 1183-1200 (1987).
8. Callis J, Fromm M, Walbot V: Heat inducible expression of a chimeric maize hsp70-CAT gene in maize protoplasts. *Plant Physiol* 88: 965-968 (1988).
9. Callis J, Raasch JA, Vierstra RD: Ubiquitin extension proteins of *Arabidopsis thaliana*. *J Biol Chem* 265: 12486-12493 (1990).
10. Christensen AH, Quail PH: Sequence analysis and transcriptional regulation by heat shock of polyubiquitin transcripts from maize. *Plant Mol Biol* 12: 619-632 (1989).
11. Czarnecka E, Nagao RT, Key JL, Gurley WB: Characterization of Gmhsp26-A, a stress gene encoding a divergent heat shock protein of soybean: heavy metal induced inhibition of intron processing. *Mol Cell Biol* 8: 1113-1122 (1988).
12. Dekeyser R, Claes B, Marichal M, Van Montagu M, Caplan A: Evaluation of selectable markers for rice transformation. *Plant Physiol* 90: 217-223 (1989).
13. Devereux J, Haerberli P, Smithies O: A comprehensive set of sequence analysis programs for the VAX: *Nucl Acids Res* 12: 387-395 (1984).
14. Finley D, Bartel B, Varshavsky A: The tails of ubiquitin precursors are ribosomal proteins whose fusion to ubiquitin facilitates ribosome biogenesis. *Nature* 338: 394-401 (1989).
15. Finley D, Ozkaynak E, Varshavsky A: The yeast polyubiquitin gene is essential for resistance at high temperatures, starvation and other stresses. *Cell* 48: 1035-1046 (1987).
16. Fried VA, Smith HT, Hildebrandt E, Weiner K: Ubiquitin has intrinsic proteolytic activity: Implications for cellular regulation. *Proc Natl Acad Sci USA* 84: 3685-3689 (1987).
17. Fromm M, Taylor LP, Walbot V: Expression of genes transferred into monocot and dicot plant cells by electroporation. *Proc Natl Acad Sci USA* 82: 5824-5828 (1985).
18. Fromm M, Callis J, Taylor LP, Walbot V: Electroporation of DNA and RNA into plant protoplasts. *Meth Enzymol* 153: 351-366 (1987).
19. Gausing K, Barkardottir R: Structure and expression of ubiquitin genes in higher plants. *Eur J Biochem* 158: 57-62 (1986).
20. Gilmour DS, Thomas GH, Elgin SCR: *Drosophila* nuclear proteins bind to regions of alternating C and T residues in gene promoters. *Science* 245: 1487-1490 (1989).
21. Hanley BA, Schuler MA: Plant intron sequences: evidence for distinct groups of introns. *Nucl Acids Res* 16: 7159-7176 (1988).
22. Henikoff, S: Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. *Gene* 28: 351-359 (1984).
23. Hershko A: Ubiquitin-mediated protein degradation. *J Biol Chem* 263: 15237-15240 (1988).
24. Jabben M, Shanklin J, Vierstra RD: Red light-induced accumulation of ubiquitin-phytochrome conjugated in both monocots and dicots. *Plant Physiol* 90: 380-384 (1989).
25. Jentsch S, McGrath JP, Varshavsky A: The yeast DNA repair gene RAD 6 encodes a ubiquitin-conjugating enzyme. *Nature* 329: 131-134 (1987).
26. Keith B, Chua N-H: Monocot and dicot pre-mRNAs are processed with different efficiencies in transgenic tobacco. *EMBO J* 5: 2149-2155 (1986).
27. Maniatis T, Fritsch EF, Sambrook J: Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1982).
28. Pelham HRB: A regulatory upstream promoter element in *Drosophila* HSP70 heat-shock gene. *Cell* 30: 517-528 (1982).
29. Rechsteiner M: Ubiquitin-mediated pathways for intracellular proteolysis. *Annu Rev Cell Biol* 3: 1-30 (1987).
30. Rechsteiner M: Natural substrates of the ubiquitin proteolytic pathway. *Cell* 66: 615-618 (1991).
31. Redman KL, Rechsteiner M: Identification of the long ubiquitin extension as ribosomal protein S27a. *Nature* 338: 438-440 (1989).

32. Schlesinger M, Bond U: Ubiquitin genes. *Oxford Surv Eukaryotic Genes* 4: 77-89 (1987).
33. Shanklin J, Jabben M, Vierstra RD: Red light-induced formation of ubiquitin-phytochrome conjugates: Identification of possible intermediates of phytochrome degradation. *Proc Natl Acad Sci USA* 84: 359-363 (1987).
34. Sharp PM, Li W-H: Ubiquitin genes as a paradigm of concerted evolution of tandem repeats. *J Mol Evol* 25: 58-64 (1987).
35. Sullivan TD, Christensen AH, Quail PH: Isolation and characterization of a maize chlorophyll a/b binding protein gene that produces high levels of mRNA in the dark. *Mol Gen Genet* 215: 431-440 (1989).
36. Winter J, Wright R, Duck N, Gasser C, Fraley R, Shah D: The inhibition of petunia hsp70 mRNA processing during CdCl₂ stress. *Mol Gen Genet* 211: 315-319 (1988).
37. Yost HJ, Lindquist S: RNA splicing is interrupted by heat shock and is rescued by heat shock protein synthesis. *Cell* 45: 185-193 (1986).