Plant Cell Physiol. 42(4): 404–413 (2001)

A *Brassica oleracea* Gene Expressed in a Variety-Specific Manner May Encode a Novel Plant Transmembrane Receptor

Joel E. Palmer [1], Dustin A. Dikeman [1], Toshinori Fujinuma [1], Brandon Kim [2], Joshua I. Jones [1], Miwako Denda [1], José Miguel Martínez-Zapater [3] and Marilyn Cruz-Alvarez [1, 4]

[1] *Department of Bioscience, Salem-Teikyo University, P.O. Box 500, Salem, WV 26426, U.S.A.*
[2] *Department of Biology, CSUSB, 5500 University Parkway, San Bernardino, CA 92407, U.S.A.*
[3] *Departamento de Genética Molecular de Plantas, Centro Nacional de Biotecnología, Universidad Autónoma de Madrid, Cantoblanco, E-28049 Madrid, Spain*

The species *Brassica oleracea* includes several agricultural varieties characterized by the proliferation of different types of meristems. Using a combination of subtractive hybridization and PCR (polymerase chain reaction) techniques we have identified several genes which are expressed in the reproductive meristems of the cauliflower curd (*B. oleracea* var. *botrytis*) but not in the vegetative meristems of Brussels sprouts (*B. oleracea* var. *gemmifera*) axillary buds. One of the cloned genes, termed *CCE1* (*CAULIFLOWER CURD EXPRESSION 1*) shows specific expression in the *botrytis* variety. Preferential expression takes place in this variety in the meristems of the curd and in the stem throughout the vegetative and reproductive stages of plant growth. *CCE1* transcripts are not detected in any of the organs of other *B. oleracea* varieties analyzed. Based on the nucleotide sequence of a cDNA encompassing the complete coding region, we predict that this gene encodes a transmembrane protein, with three transmembrane domains. The deduced amino acid sequence includes motifs conserved in G-protein-coupled receptors (GPCRs) from yeast and animal species. Our results suggest that the cloned gene encodes a protein belonging to a new, so far unidentified, family of transmembrane receptors in plants. The expression pattern of the gene suggests that the receptor may be involved in the control of meristem development/arrest that takes place in cauliflower.

Key words: *Brassica oleracea* — Cauliflower — G-protein-coupled receptors — Meristems — Transmembrane proteins.

Abbreviations: FMI, floral meristem identity; GPCR, G-protein-coupled-receptor; LRR, leucine-rich repeat; RACE, rapid amplification of cDNA ends; SSC, saline sodium citrate; SAM, shoot apical meristem; TM, transmembrane.

The nucleotide sequence reported in this paper has been submitted to GenBank under accession number AF227978 (*CCE1*-cDNA).

Introduction

All aerial parts of a plant develop from the shoot apical meristem (SAM), a group of undifferentiated cells at the tip of the shoot. During the vegetative growth of the plant, the SAM is involved in the formation of stem and leaves. After floral induction the SAM gives rises to bracts, floral branches, and lateral floral meristems, and in determinate species it develops into a terminal flower. Plant morphogenesis is therefore highly dependent on the patterns and numbers of cell divisions in the SAM. Cell division and differentiation in the SAM must be tightly regulated in order to maintain the meristem, and concomitantly produce different tissues and organs.

Extensive genetic analyses have led to the identification of genes expressed in the SAM which: (a) control cell divisions or (b) determine the identity of this meristem, i.e. whether it is a vegetative, inflorescence or floral meristem. *SHOOTMERISTEMLESS* (*STM*) in Arabidopsis, and *KNOTTED1* (*KN1*) in maize, for example, belong to a family of homeobox genes which activate cell division in different meristematic regions (Barton and Poethig 1993, Long et al. 1996). On the other hand, genes like *CLAVATA1* (*CLV1*) and *CLAVATA3* (*CLV3*) in Arabidopsis promote transition of meristematic cells into differentiation pathways, or repress meristematic cell division, or both (Clark et al. 1997, Fletcher et al. 1999). Mutations in the genes *CLV1* and *CLV3* give rise to enlarged meristems, which continue to increase in size throughout the vegetative and reproductive development of the plant (Clark et al. 1995).

Among the floral meristem identity (FMI) genes identified in Arabidopsis, are *LEAFY* (*LFY*), *APETALA1* (*AP1*) and *CAULIFLOWER* (*CAL*) (Weigel et al. 1992, Mandel et al. 1992, Kempin et al. 1995). These genes work cooperatively to determine the formation of lateral floral meristems. Mutations in either *LFY* or *AP1* cause reversion of these floral meristems into inflorescence meristems. A mutation in the *CAL* gene by itself does not cause any distinguishable phenotype in Arabidopsis. However, this mutation enhances the *ap1* phenotype, in such a way that in the double mutant *ap1cal*, floral meristems are transformed into inflorescence meristems that give rise to additional meristems in a spiral phyllotaxy. The resulting phenotype looks remarkably similar to that of the cauliflower curd (Kempin et al. 1995). Genes which repress the floral transition have also been identified (Chen et al. 1997, Ratcliffe et al. 1999). Expression of one of these genes, *TERMINAL FLOWER*

[4] Corresponding author: E-mail, cruzalvarezm@aol.com; Fax, +1-305-663-5843; Phone, +1-304-203-2383.

1 (*TFL1*) in the central zone of the SAM in Arabidopsis seems to be responsible for the maintenance of the inflorescence identity of this meristem (Ratcliffe et al. 1999).

Despite the identification and cloning of increasing number of genes, including those mentioned above, the molecular mechanisms involved in the control of cell division and differentiation in the SAM remain, to a large extent, unknown. A molecular approach to the identification of genes expressed in meristems would complement the information available from the genetic analyses. However, this approach has been hampered by the small size of the SAM in most species. *Brassica oleracea* is a useful system for such a molecular approach. This species includes different cultivated varieties: broccoli, Brussels sprouts, cabbage, collards, cauliflower, and kohlrabi, and is therefore of great economic value. In some of these varieties the harvestable part is produced as a result of a massive proliferation of meristems. For example, in the *botrytis* variety, commonly known as cauliflower, the curd is produced by an enormous proliferation of apical meristems that are arrested at an inflorescence and early-floral stages (Sadik 1962, Carr and Irish 1997). In the *italica* variety, commonly known as broccoli, the head is produced by a proliferation of apical meristems that have developed into flowers (Carr and Irish 1997). In another variety, *gemmifera*, the little cabbages commonly known as Brussels sprouts are produced by a proliferation of axillary meristems in a vegetative stage (Gómez-Campo and Prakash 1999). The proliferation of meristems in these varieties results in ample sources of mRNAs that may be used for the identification of genes expressed in the meristems. The available varieties may also provide the opportunity to identify genes responsible for differences in meristem development in almost identical genetic backgrounds.

We have used the cauliflower curd and the axillary buds in the Brussels sprouts as sources of two different types of meristems from the same species. We used a subtractive hybridization protocol in an attempt to clone genes expressed specifically in the meristems of the cauliflower curd and that might be involved in (a) the proliferation of the SAM and/or (b) the developmental arrest at the inflorescence/early floral stage. In this paper we describe the identification of a gene expressed in the cauliflower curd meristems following this molecular approach, and its further characterization. The pattern of expression of this gene, we have designated as *CCE1* for *c*auliflower *c*urd *e*xpression *1*, suggests that it may indeed be involved in the formation of the curd.

Materials and Methods

Plant material

Brassica oleracea plants were grown in pots in a greenhouse at approximately 20°C. Tissues were frozen in liquid nitrogen and stored at –80°C for later RNA or DNA extraction. Except where indicated, the following varieties and cultivars were utilized for preparation of RNA and DNA samples: cauliflower (*B. oleracea* var. *botrytis* cv. White Magic), Brussels sprouts (*B. oleracea* var. *gemmifera* cv. Catskill), Broccoli (*B. oleracea* var. *italica* cv. Marathon) and Rbo (Rapid cycling *B. oleracea*).

Cloning of the original CCE1 cDNA by subtractive hybridization

cDNAs constructed from poly(A)$^+$ RNAs extracted from the meristems of cauliflower curds or axillary buds of Brussels sprouts were used for six cycles of subtractive hybridization and amplification. Construction of the cDNA and the subtractive hybridization protocols have been described in detail elsewhere (Wang and Brown 1991, Franco-Zorrilla et al. 1999). The cauliflower cDNAs obtained after six cycles of subtraction and amplification, cDNA-6, were cloned into the *Eco*RI site in pBSKS$^+$ plasmid (Stratagene), and JV30 competent cells transformed with the recombinant plasmids. Colonies were replica-plated onto nitrocellulose filters and screened by hybridization with [^{32}P]dCTP-labeled cauliflower cDNA-6 and [^{32}P]dCTP-labeled original (not subtracted) Brussels sprouts cDNA, cDNA-0. Clones hybridizing to the subtracted cauliflower cDNA and not to the Brussels sprouts cDNA were picked as putative positives. Colonies containing unique inserts, as shown by lack of hybridization to any of the other inserts, were picked for further analyses. One of these clones, number 153, corresponds to *CCE1*.

Screening of a cauliflower curd cDNA library

A cauliflower curd cDNA library constructed in the vector λgt-11 was kindly provided by Dr. E. Lifschitz. Approximately 5×10^5 plaques were screened in standard conditions. The *Eco*RI insert from the clone isolated by the subtractive hybridization protocol was labeled with [^{32}P]dCTP (New England Life Science) by random priming (Boehringer) and used as probe.

Rapid amplification of cDNA ends

5′-RACE (rapid amplification of cDNA ends) was used to clone the 5′-sequences of the *CCE1* cDNA. The protocol used has been described previously (Frohman et al. 1988). A specific 20-base oligonucleotide primer OPCR (5′-GCATGTGTCAGTTGAGTACC-3′) (Life Technologies) (Fig. 1A) corresponding to a sequence close to the 5′ end of the previously cloned *CCE1* cDNAs was annealed to 1 µg of poly(A)$^+$ RNA from cauliflower curd meristems and used for reverse transcription with Superscript II reverse transcriptase (Life Technologies). Excess primer was removed by column chromatography, using a cDNA Spun Column Sephacryl S-300 (Pharmacia). First strand cDNA was oligo-A tailed on its 3′ end in a total volume of 30 µl. The reaction mixture was diluted to 500 µl, and a 10 µl aliquot used for cDNA amplification with an oligo-dT-containing primer (5′-GCGGCCGCTT-TTTTTTTTTTTT-3′) and a nested primer OPC2 (5′-CGGAGAGGT-GATACTATTGAC-3′) (Fig. 1A). PCR reactions were carried out in 100 µl containing: 50 mM KCl, 10 mM Tris-HCl pH 8.3, 2 mM $MgCl_2$, 0.2 mM dNTPs, 100 pmol of each primer and 2.5 units of AmpliTaq DNA polymerase (Perkin Elmer). The PCR product was purified from a low-melting agarose gel by passing through an Elutip-D column (Schleicher & Schuell). The purified PCR product was cloned into the pCR^R II TOPO vector (Invitrogen).

Sequencing and sequence analysis

The following restriction fragments were used for DNA sequencing: the *Eco*RI insert in the original cDNA cloned by subtractive hybridization (*cauli 153*), the *Eco*RI–*Hin*dIII, *Hin*dIII–*Bam*HI and *Bam*HI–*Eco*RI fragments from the cDNA obtained from the cDNA library (*10–2–1*) and the 5′end-*Xho*I, *Xho*I–*Bam*HI, 5′end-*Bgl*II, *Bgl*II–*Nhe*I and *Nhe*I–*Eco*RI fragments from the clone obtained by 5′-RACE (*5′-RACE#7*) (Fig. 1A). All fragments were sequenced in both orientations using T7 Sequenase version 2.0 (United States Biochemicals).

Analyses of DNA sequences were done using the DNASTAR

software (DNASTAR, Inc., Madison, WI, U.S.A.). The BLAST program (Altschul et al. 1997) at the National Library of Medicine web site of the National Institutes of Health (www.ncbi.nlm.nih.gov) was used for sequence homology search against all nucleotide and protein sequences in GenBank. Analysis of the protein sequence was done using the following proteomic tools: ProtScale and TMpred (Hofmann and Stoffel 1993) at the Expert Protein Analysis System web site of the Swiss Institute of Bioinformatics (www.expasy.ch), and PSORT (Nakai and Kanehisa 1992) and MOTIF at the GenomeNet web server (www.genome.ad.jp). MOTIF was used against the PROSITE (Hofmann et al. 1999), BLOCKS (Henikoff and Henikoff 1994) and PRINTS (Attwood et al. 1999) databases. The CLUSTAL method of MegAlign (DNASTAR) was used for amino acid sequence comparisons between specific proteins.

Slot, Northern and Southern blot analyses

RNA was extracted from the tissues indicated following the procedure described by Das et al. (1990). The RNA concentration and purity were determined spectrophotometrically. For slot blot analysis, poly(A)$^+$ RNAs were purified using oligo-dT affinity column chromatography. Double-stranded cDNAs were constructed and an *Eco*RI linker was ligated to both ends of the cDNAs. An oligonucleotide complementary to the linker was used as primer for PCR amplification. Same amounts of amplified cDNA from different tissues, as indicated in Fig. 2A were bound to a nitrocellulose filter, and this was hybridized to a 571-bp *Eco*RI–*Bam*HI restriction fragment from the *CCE1* cDNA, labeled with [^{32}P]dCTP by random priming. Hybridization was carried out for 16–20 h in 5× Denhardt's, 5× SSC (1× SSC is 3 M sodium chloride, 0.3 M sodium citrate), 50% formamide, 20 mM EDTA, 0.1% SDS at 42°C. The filter was washed twice at room temperature in 2× SSC, 0.1% SDS, twice in the same solution at 65°C, then twice in 1× SSC, 0.05% SDS also at 65°C, then once in 0.2× SSC, 0.01% SDS at 65°C.

For Northern blot analysis, 10 μg of total RNA from each tissue were fractionated on a 1% agarose-formaldehyde gel. The RNA was transferred to a nitrocellulose filter by capillarity. The blot was hybridized to an 838-bp 5′end-*Eco*RI fragment from the *CCE1* cDNA, labeled with [^{32}P]dCTP by random priming. Hybridization and washing stringencies were the same as for the slot blot.

DNA was isolated from leaves or seedlings of the different varieties and cultivars of *B. oleracea* plants as described previously (Das et al. 1990). DNA samples were digested with the enzymes indicated. Digestion products were subjected to electrophoresis on a 0.8% agarose gel and transferred to a nitrocellulose filter. The filters were hybridized to the same probe used for Northern analysis. Hybridization and washes were done in the same stringency conditions indicated above.

Results

Cloning of CCE1 DNA

By using a combination of subtractive hybridization and PCR we identified several genes expressed in the meristems of the cauliflower curd but not in the meristems of Brussels sprouts axillary buds. We designated one of these genes as *CCE1* for *c*auliflower *c*urd *e*xpression *1*. The original *CCE1* cDNA fragment cloned by subtractive hybridization, *cauli153*, was short (270 bp). This was expected, since the protocol followed for the subtractive hybridization involved digestion of the double-stranded cDNAs with restriction enzymes *Rsa*I and/or *Alu*I (Wang and Brown 1991). Screening of a cDNA library and 5′-RACE were performed in order to isolate overlapping cDNAs corresponding to the full-length transcript, approximately 1.5–1.6 kb in size as determined from Northern blot analysis. Fig. 1A shows the three cDNA clones obtained for *CCE1*, and their overlapping regions. Screening of a λgt11 cauliflower curd cDNA library using the original *cauli153* cDNA as a probe, resulted in the isolation of a second cDNA clone, *10–2–1*, which was 647 bp long. This cDNA included the entire length of the original clone *cauli153*, except for 31 bp at its 5′ end. The *10–2–1* cDNA included a short polyA tail, 14 bases long, and a putative polyadenylation signal AAUAUA (AATATA in the DNA) 13 bases upstream from this polyA tail. These results indicated that the *10–2–1* cDNA included sequences complementary to the 3′-end of the transcript. Based on the available cDNA sequence, two primers were designed to conduct 5′-RACE in order to obtain a complementary cDNA corresponding to the 5′ part of the transcript. The sequences chosen to construct the primers were close to the 5′ end of the cDNAs and were in the overlapping region of *cauli153* and *10–2–1* (Fig. 1A). Primer OPCR was used for first strand cDNA synthesis using as template poly(A)$^+$ RNA from cauliflower curd meristems. The single-stranded cDNA was oligo-dA tailed, and primer OPC2 and a primer containing an oligo-dT tail were used for amplification. The amplified cDNA, *5′-RACE #7*, was 907 bp long. Its sequence included that of the primer used for amplification, OPC2, and overlapped with those of clones *cauli153* and *10–2–1* over 71 and 40 bp, respectively. The sum of the lengths of the overlapping cDNAs is 1,515 bp. The complete nucleotide sequence of these cDNAs is shown in Fig. 1B. From the size of the transcript detected by Northern blotting analysis, it is reasonable to conclude that these overlapping cDNAs correspond to a full-length transcript.

Analysis of the nucleotide and deduced amino acid sequences

The sequence contains a single open reading frame of 1,373 bp encoding a predicted protein of 457 amino acids (Fig. 1B). The first ATG of this open reading frame is preceded by an in-frame stop codon. It is interesting to note that the start codon is not in the context considered optimal for translation initiation in plants (Lütcke et al. 1987); there is a pyrimidine, C, in both –3 and +4 positions. However, in vitro transcription/translation experiments performed with the complete cDNA (1,515 bp) cloned in a pBluescript vector have shown that translation can start relatively efficiently from this AUG (results not shown). The mobility of the in vitro translated protein in SDS-polyacrylamide gel electrophoresis is consistent with the size of the protein calculated from the deduced amino acid sequence (50.8 kDa), confirming the existence of the predicted open reading frame within the complete cDNA.

Analysis of the deduced amino acid sequence for the putative protein encoded by *CCE1* revealed that the protein is rich in serines (14.9%) and leucines (14.4%). There are some potential post-translational modification sites in the sequence. Sev-

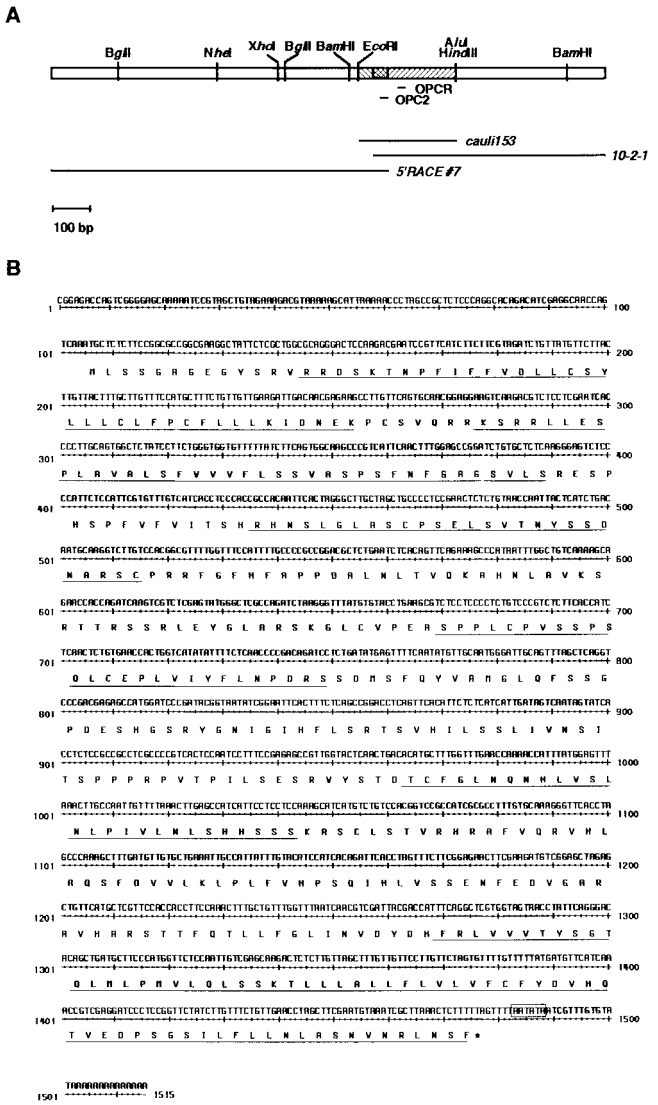

Fig. 1 Nucleotide sequence of *CCE1* cDNA and deduced amino acid sequence of encoded protein. Panel A shows a restriction map of the complete *CCE1* cDNA (1,515 bp), indicating the enzymes used to obtain restriction fragments for sequencing and generation of probes. The slashed areas correspond to regions of overlap between the partial cDNAs isolated. The relative position of the cDNAs isolated by subtractive hybridization (*cauli153*), screening of a λgt11-cDNA library (*10–2–1*), and 5′-RACE (*5′-RACE#7*) is shown below. The positions of the two sequence-specific primers (OPCR and OPC2) used for 5′-RACE are also shown. Panel B shows the complete nucleotide sequence of *CCE1* cDNA (GenBank accession number AF227978). The amino acid sequence deduced from the single open reading frame is shown below. A putative polyadenylation signal is boxed. Regions containing conserved motifs characteristic of G-protein coupled receptors are underlined.

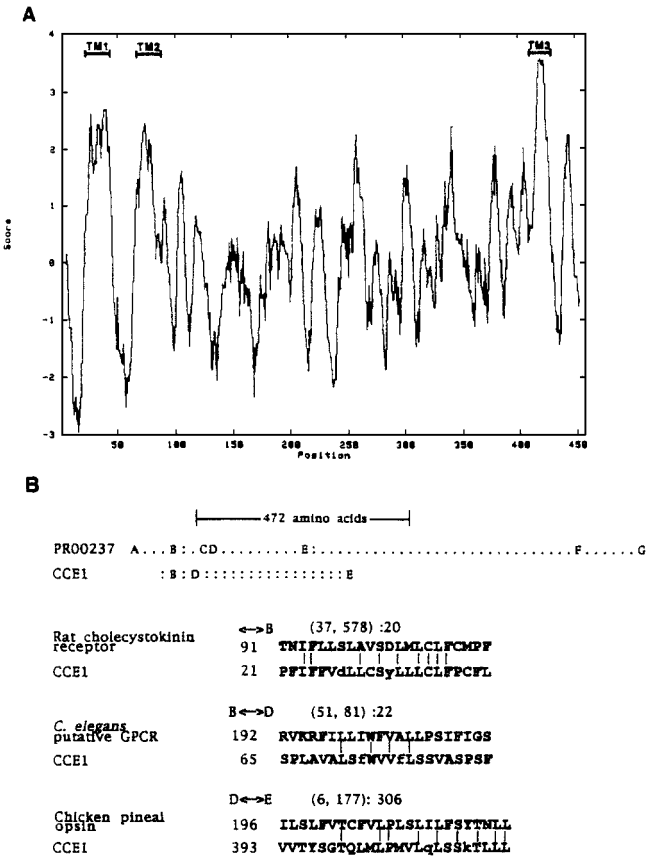

Fig. 2 Analysis of the deduced amino acid sequence for CCE1. Panel A shows a hydrophobicity plot for the putative protein encoded by *CCE1*. The sequence positions corresponding to three putative transmembrane helices are indicated above the graph. Panel B shows the distribution of the seven elements or blocks (A through G) within the characteristic fingerprint of the rhodopsin-like GPCR family (PR00237 in the PRINTS database), and the distribution of these motifs within the deduced amino acid sequence of CCE1. The minimum distance between blocks in the database is represented by a colon; the maximum distance between blocks in the database is represented by periods. The scale in number of amino acids is shown above. The alignments of the CCE1 sequence with the three sequences in the database closest to it for each of the three blocks: B, D and E are shown below. The minimum and maximum distances between those particular blocks in the database are listed in parentheses (minimum, maximum). These are followed by the distance between the corresponding blocks for CCE1. An upper case amino acid in the CCE1 sequence indicates at least one occurrence of that residue in that column in the block among the sequences in the database.

eral serines (amino acid positions 17, 59, 109, 170, 171, 231, 312 and 410) and three threonines (amino acid positions 167, 320 and 433) are in contexts that are optimal for phosphorylation by either protein kinase C, cAMP/cGMP-dependent protein kinases or casein kinase II (Feramisco et al. 1980, Kishimoto et al. 1985, Pinna 1990). In addition, three asparagine residues (amino acid positions 128, 152 and 306) are in a consensus sequence for N-glycosylation (Marshall 1972).

A hydrophobicity plot for the protein is shown in Fig. 2A (Kyte and Doolittle 1982). As indicated in this figure, the deduced amino acid sequence clearly contains three highly hydrophobic segments. These segments correspond to amino

acids 21–48, 65–85 and 405–434, and constitute strong potential transmembrane (TM) regions. Both TMpred (Hofmann and Stoffel 1993) and PSORT (Nakai and Kanehisa 1992) predict that the protein is a transmembrane protein of type IIIb. The strongly preferred model for transmembrane topology is one with the N-terminus end of the protein in the exo-cytoplasmic side and the C-terminus end in the cytosolic side. An alternative model would include a fourth transmembrane domain which would correspond to residues 437–457. The PSORT algorithm was used to further analyze the possible cellular localization of the protein encoded by *CCE1*. Both the McGeoch's (McGeoch 1985) and von Heijne's (von Heijne 1986) methods used by PSORT failed to detect a typical N-terminal signal sequence in the protein. Although type IIIb transmembrane proteins are mostly localized in the endoplasmic reticulum, the putative protein encoded by *CCE1* lacks the consensus sequences involved in retention in this membrane compartment (Denecke et al. 1992). No sequences that would suggest targeting into any of the other cell organelles were detected either. Since trafficking of proteins to the plasma membrane is believed to be the default pathway in the absence of any specific targeting signals (Sanderfoot and Raikhel 1999), the plasma membrane is the most probable subcellular location for this protein.

Comparison of the cDNA nucleotide sequence and deduced amino acid sequence (GenBank accession number AF227978) with sequences in databases at the National Library of Medicine, including *Arabidopsis thaliana* ESTs, showed no significant similarity to any other known genes or proteins. The nucleotide sequences showing highest homology to the *CCE1* cDNA were genomic clones from Arabidopsis, and corresponded to BAC clone F24H14 from chromosome 2 (68% homology in 166 bp), BAC clone T5K18 from chromosome 4 (64% homology in 197 bp) and TAC clone K17022 from chromosome 5 (62% homology in 166 bp) (data not shown). The highest score for the amino acid sequence was obtained with the mouse DNA replication licensing factor MCM5 (GenBank accession number P49718). The deduced amino acid sequence for CCE1 showed only 26% identity with this protein in 65 amino acids. These results indicate that *CCE1* is a novel gene for which no orthologs have been identified.

Comparison of the protein sequence with the PRINTS database showed the presence of several motifs (underlined in the sequence in Fig. 1B) characteristic of the rhodopsin-like GPCR superfamily. Although rhodopsin-like GPCRs show considerable diversity at the sequence level, they contain a 7-element fingerprint that corresponds to seven transmembrane (7-TM) helices and constitutes the signature for this superfamily (Attwood and Findlay 1993). The deduced amino acid sequence for CCE1 contains motifs present in 6 (A, B, D, E, F and G in Fig. 2B) out of the seven elements that compose the fingerprint. Analysis of the distribution of these motifs (Henikoff and Henikoff 1994), showed that three of the elements or blocks (B, D, and E in Fig. 2B) are in the same order

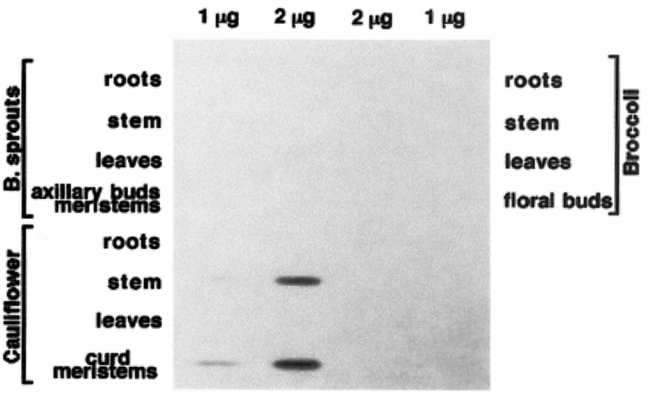

Fig. 3 Slot blot analysis of *CCE1* expression. The indicated amounts of amplified cDNAs synthesized from poly(A)$^+$ RNA from cauliflower, Brussels sprouts and broccoli tissues were bound to a nitrocellulose filter. The filter was hybridized to a 571-bp *Eco*RI–*Bam*HI restriction fragment from the *CCE1* cDNA, labeled with [^{32}P]dCTP by random priming.

within the deduced amino acid sequence of CCE1 and in the characteristic fingerprint. Fig. 2B shows the alignment of the CCE1 sequence with the sequences closest to it in the PRINTS database for each of these three blocks: rat cholecystokinin type A receptor (GenBank accession number P30551), a putative GPCR from *C. elegans* (GenBank accession number U53340) and chicken pineal opsin (GenBank accession number P51475), respectively. These blocks overlap with the three predicted transmembrane regions.

Expression pattern of CCE1

An initial screening of the expression pattern of *CCE1* was performed by slot blot analysis. Double stranded cDNAs were constructed from poly(A)$^+$ RNA isolated from cauliflower, Brussels sprouts and broccoli tissues. Linkers were ligated to both ends of the cDNA and an oligonucleotide complementary to the linkers used as primer for PCR amplification. Different amounts of amplified cDNAs were bound to a nitrocellulose filter in a slot blot, and hybridized to a *CCE1* probe. It is unlikely that the efficiency of reverse transcription of the same mRNAs would be different in samples from different tissues, resulting in different lengths of the corresponding cDNAs. If the average length for cDNAs from different tissues is the same, for any given gene the relative amounts of amplified cDNAs in those tissues would be equivalent to the relative amounts of mRNAs. The results show, as it could be expected from the isolation protocol, that *CCE1* is expressed in the meristems of the cauliflower curd but not in the meristems of the Brussels sprouts axillary buds (Fig. 3). Expression was also detected, at comparable, although somewhat lower levels in the inflorescence stem from cauliflower. After prolonged exposures, weak hybridization bands could be observed in the film on the slots corresponding to cauliflower leaves and roots,

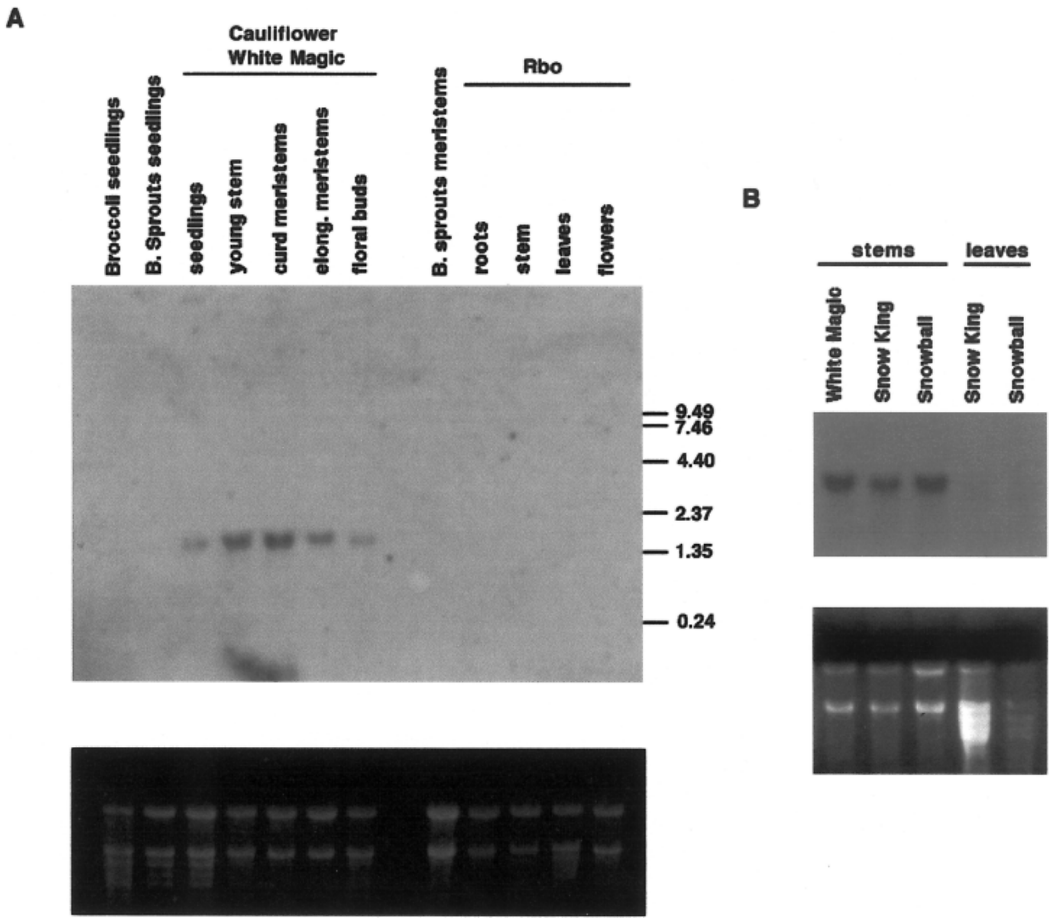

Fig. 4 Northern blotting analysis of *CCE1* expression. RNA was isolated from plant tissues as described. Approximately equal amounts (10 μg) of total RNA from the cauliflower, Brussels sprouts or Rbo tissues indicated (panel A) or from stem or leaf tissue of different cauliflower cultivars (panel B) were loaded on a 1% agarose-formaldehyde gel. RNA was transferred to nitrocellulose filters by capillarity. The top panels show the results of hybridization to an 838 bp 5′-*Eco*RI fragment from *CCE1* cDNA, labeled with [^{32}P]dCTP by random priming. Mobilities of size markers in kilobases are shown on the right margin for panel A. The bottom panels show ethidium bromide staining of the RNA on the gels.

revealing much lower levels of expression of *CCE1* in these tissues. Expression was not detected in any tissues of the other *B. oleracea* varieties analyzed.

A more detailed analysis of *CCE1* expression was conducted by Northern blotting. Expression was analyzed in the vegetative stage of plant development, in seedlings, and young stem, prior to the formation of the curd. Expression was also analyzed in cauliflower after the flowering transition resumes in the apical meristems, at the time when the branches within the curd elongate and a later stage, when flowers develop. As shown in Fig. 4A, expression of *CCE1* takes place throughout the vegetative and reproductive stages of plant growth in cauliflower. *CCE1* transcripts were detected in seedlings and in the vegetative stem prior to the appearance of the inflorescence. *CCE1* transcripts could also be detected in the meristems after curd elongation and in the flower buds. However, transcript levels seem to decrease after resumption of floral development.

CCE1 transcripts could be detected neither in Brussels sprouts or broccoli seedlings, nor in roots, stem, leaves or flowers of Rbo, a rapid cycling variety of *B. oleracea*, considered to be wild type with respect to plant morphology (Fig. 4A). Lack of detection of *CCE1* transcripts in these *B. oleracea* varieties could be due to highly restricted expression of this gene, spatially or temporally. However, the relatively high levels of *CCE1* transcripts in cauliflower seedlings and the absence of detection of *CCE1* transcripts in Brussels sprouts or broccoli seedlings strongly suggest that *CCE1* is specifically expressed in cauliflower. Differences in expression do not seem to be simply due to differences in cultivars either. *CCE1* expression has been detected in all cauliflower cultivars analyzed: White Magic, Snowball, Snow King, Snow Crown, Candid Charm and Incline. Either stems (Fig. 4B) or seedlings (results not shown) of these cultivars were used for Northern blotting analysis.

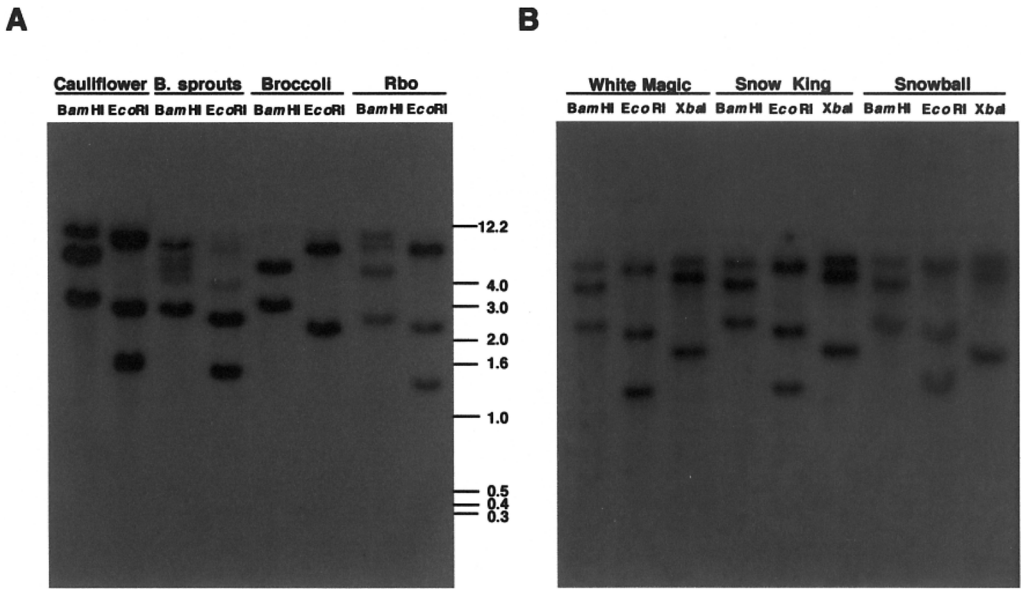

Fig. 5 Southern blot analysis of genomic DNA. DNA was isolated from leaves of the different varieties of *B. oleracea* (panel A) or different cauliflower cultivars (panel B) as described. Ten micrograms of each DNA sample were digested with *Bam*HI (B) or *Eco*RI (E). Digestion products were fractionated on a 0.8% agarose gel and transferred to a nitrocellulose filter by capillarity. The filters were hybridized to a [^{32}P]dCTP-labeled fragment (838 bp 5′ end-*Eco*RI) from *CCE1* cDNA. Mobilities of size markers in kilobases are shown on the right margin for panel A.

Southern blotting analysis of CCE1 genomic sequences

Southern blotting analysis performed under stringent conditions detected several fragments with homology to the *CCE1* cDNA in the cauliflower genome, as well as in the genome of other *B. oleracea* varieties (Fig. 5A). Analyses of genomic clones isolated from a cauliflower genomic library using *CCE1* cDNA as probe, have shown that the *CCE1* gene does not contain any introns (unpublished observations). The fact that more than one restriction fragment from the *B. oleracea* genome shows hybridization to the *CCE1* specific probe could therefore be due to (a) allelic polymorphism, since hybrid cultivars were used for DNA analysis, and/or (b) the presence of several gene copies in the genome. Analyses of DNA from the six different cauliflower cultivars mentioned above with restriction enzymes *Bam*HI, *Eco*RI and *Xba*I showed the same pattern of restriction fragments hybridizing with the *CCE1* probe. The results for three of the cultivars are shown in Fig. 5B. These results suggest that each hybridizing fragment corresponds to a different gene copy for *CCE1*. Digestion of cauliflower genomic DNA with restriction enzymes *Eco*RI or *Xba*I, and hybridization to a 5′ end-*Eco*RI (position 838) probe containing no recognition sites for these enzymes yielded three strongly hybridizing bands (Fig. 5A, B). However, digestion of the genomic DNA with *Bam*HI and hybridization to an *Eco*RI–*Bam*HI restriction fragment from positions 838 to 1,409, yielded one prominently hybridizing band (result not shown). This band corresponds to the 593 bp *Bam*HI restriction fragment expected from the cDNA sequence. These results agree with those obtained from the characterization of cauliflower genomic clones and suggest that the genome of the *botrytis* variety includes three distinct genes with a high degree of homology within the coding region. Based on the results of the Southern blot analysis, other *B. oleracea* varieties seem to contain at least two highly homologous copies of *CCE1*. In some instances weaker bands were also observed, due most likely to the presence of gene copies with a lower degree of homology in their genomes. Southern blot analysis performed under lower stringency conditions as well as preliminary characterization of the cauliflower genomic clones isolated have also revealed the presence of such gene copies in the *botrytis* variety.

Discussion

The *botrytis* variety of *B. oleracea* is characterized by the formation of a curd: a proliferation of apical meristems arrested at an inflorescence or early floral stages. Little is known about the molecular basis of formation of this curd. The excessive cell division that gives rise to the proliferation of meristems in this case can be distinguished from the uncontrolled cell division in the SAM of the *clavata* mutants in Arabidopsis (Clark et al. 1995). In cauliflower the number of cells within each meristem seems to remain the same, since upon resuming the floral transition and after elongation of the branches in the curd, the meristems give rise to normal flowers.

The molecular basis for the developmental arrest of the meristems in the curd is not completely understood, either. The

FMI genes *LFY*, *AP1* and *CAL* have been cloned in *B. oleracea* (Kempin et al. 1995, Anthony et al. 1996, Carr and Irish 1997). Although the *LFY* and *AP1* orthologs appear to be functional in cauliflower, the *CAL* gene contains a non-sense mutation in this variety. It has been proposed that this mutation is therefore responsible for the cauliflower phenotype (Kempin et al. 1995). However, arrest of the inflorescence at an early stage has been shown to be dominant or semi-dominant in crosses between cauliflower and other *B. oleracea* varieties (Bowman et al. 1993). This suggests that if a mutation in a single gene were responsible for the phenotype, the mutation should be a gain of function mutation and not the loss of function mutation expected from the production of a truncated protein. So far complementation of the mutant phenotype of cauliflower with the wild-type gene has not been reported, and therefore, there is no direct evidence that the *CAL* mutation is solely responsible for the formation of the curd.

We have used a molecular approach for the identification of genes involved in the formation of the cauliflower curd. By combining subtractive hybridization and PCR we have been able to isolate several genes expressed in the meristems of this curd, but not in other types of meristems of the same species, such as the meristems in the Brussels sprouts axillary buds. One of the cloned genes, *CCE1* appears to belong to a small gene family in *B. oleracea*. Although homologous genomic sequences are present in all *B. oleracea* varieties analyzed, transcripts homologous to *CCE1* have only been detected in the *botrytis* variety, showing a correlation with the formation of a curd. Due to the lack of available information regarding the origin of the cauliflower cultivars used for our analyses, we cannot completely rule out the possibility that the expression of *CCE1* in these cultivars is due to genotype-specific expression in one common ancestral strain. It is certainly probable that some of these cultivars share part of their genetic background. However, it is very unlikely that all these hybrids share a common allele responsible for *CCE1* expression. In fact, Southern blot analyses of DNA from the Snowball and White Magic cultivars using restriction enzyme *Hin*dIII have revealed some polymorphism in the *CCE1* alleles of these two cultivars (results not shown). It seems reasonable to conclude then that expression in these cultivars reflects specific expression in the cauliflower variety.

Both Southern blot analyses and preliminary characterization of genomic clones have indicated that there are three highly homologous genes including the *CCE1* coding sequences in the cauliflower genome. These results are not surprising, since the current hypothesis is that Brassica diploid species, including *B. oleracea*, are secondary polyploids (Prakash and Hinata 1980). In fact, approximately 50% of all cDNA sequences cloned in *B. oleracea* have been mapped to more than one locus (Quiros 1999). Restriction mapping of the *CCE1* genomic clones isolated has shown that the distinct gene copies have almost identical coding sequences (unpublished results). Since the probes used for slot and Northern blot analyses were derived from restriction fragments within the coding region, it is certainly possible that the pattern of expression observed corresponds to the additive expression of more than one gene copy. One or more copies may be expressed specifically or at least preferentially in the *botrytis* variety. However, since the cauliflower is believed to be derived from broccoli (Gómez-Campo and Prakash 1999), the simplest explanation for the observed expression pattern is that a single gene copy was turned on, or its expression increased to a much higher level, concomitantly with the evolution of the *botrytis* variety.

Expression of *CCE1* in the apical meristems of cauliflower takes place during the formation of the curd, and continues, at what appears to be lower levels, after the plants resume the floral transition. The fact that *CCE1* transcripts are present in other cauliflower tissues, and temporally prior to the formation of the curd, indicates that its expression is not simply a consequence of the proliferation or developmental arrest of the meristems in this curd. It is interesting to point out that in Arabidopsis the *TFL1* gene, which has been shown to be involved in determination of inflorescence meristem identity, is also expressed during the vegetative stages of development of the plant (Bradley et al. 1997). The protein encoded by *CCE1* could be directly or indirectly involved in the arrest in the floral transition that takes place in the apical meristems of cauliflower.

Analysis of the amino acid sequence of the putative protein encoded by *CCE1* has indicated that it is a transmembrane protein. The lack of any sorting signal that would target the protein to a specific intracellular compartment suggests that this protein is localized on the plasma membrane. According to the preferred model for membrane topology, most of the amino acid residues, including the N-terminal end would be localized in the extracellular space, whereas the C-terminal end would be in the cytoplasm. The three potential glycosylation sites would be in the exocytoplasmic loop between the second and third transmembrane domains, and therefore the model supports the notion of in vivo glycosylation. Among the possible phosphorylation sites detected in the deduced amino acid sequence, only two would be susceptible to phosphorylation by cellular kinases. Those would be the serine at position 59, in the cytoplasmic loop between the first and second transmembrane domains, susceptible to phosphorylation by a cAMP/cGMP-dependent kinase, and the threonine at position 433, in the C-terminal end, susceptible to phosphorylation by casein kinase II.

Comparison of the deduced amino acid sequence with sequences in the databases failed to reveal extensive homology to any known protein. Therefore, it is not possible to unmistakably assign a function to the putative protein encoded by *CCE1*. However, the deduced amino sequence includes motifs characteristic of the rhodopsin-like subfamily (type A subfamily) of GPCRs (Attwood and Findlay 1993). GPCRs comprise a wide variety of 7-TM receptors in yeast and animals, including hormone, neurotransmitter, cAMP, pheromone and light receptors. There is increasing biochemical evidence that

in plants G-protein signal transduction pathways are also involved in responses to a variety of stimuli such as hormones, light, and pathogens (Ma 1994, Plakidou-Dymock et al. 1998). However very little is known at the molecular level of GPCRs and G-proteins in plants. Only one gene encoding a GPCR has been cloned in a plant species to date: the *GCR1* gene in Arabidopsis (Plakidou-Dymock et al. 1998). The similarity in amino acid sequences between the putative protein encoded by *CCE1* and the rhodopsin-like GPCR superfamily may reflect some functional analogy. The extent of overall sequence homology between CCE1 and rhodopsin-like GPCRs is very limited (CCE1 only shows 15.4%, 12.2% and 11.3% similarity with chicken pineal opsin, a *C. elegans* putative GPCR and rat cholecystokinin type A receptor, respectively). However, as pointed out earlier, the sequence similarity between members of the rhodopsin-like GPCR superfamily is limited as well (for example, there is only 20.8% similarity between the rat and chicken receptors, 15.8% between the rat and the *C. elegans* receptors, and 12.8% between the chicken and *C. elegans* receptors). CCE1 could be a novel type of GPCR, containing only three of the 7-TM elements that constitute the characteristic fingerprint.

Recently an increasing number of plant developmental signals have been shown to be perceived on the cell surface. The putative large extracellular domain of CCE1 could provide the structural framework for effective binding of an extracellular ligand. Since many of the identified cell surface receptors in plants contain characteristic leucine-rich repeats (LRRs) in their extracellular domains (Torii et al. 1996, Clark et al. 1997, Li and Chory 1997), we analyzed the position of the leucines residues in the extracellular loop of CCE1. However, this analysis revealed that in this protein the spacing of leucines does not match the one present in the LRRs, and indeed does not seem to follow any repetitive pattern.

Finally, it is possible that the proteomic tools utilized failed to predict accurately all the transmembrane domains within CCE1. In this respect, it is intriguing that three additional regions containing motifs present in the 7-TM fingerprint of GPCRs have been localized within the CCE1 sequence (Fig. 1B).

The results suggest that *CCE1* encodes a novel plant receptor. Future experiments will be critical to determine if indeed the protein is localized on the cell membrane, and whether it functions in cauliflower as the receptor of a developmental signal regulating cell division and/or differentiation in the SAM.

Acknowledgements

We would like to thank Sakata Seed America, Inc. for providing cauliflower (cvs. White Magic, Candid Charm and Incline) and broccoli (cv. Marathon) seeds. We thank Dr. E. Lifschitz for making available to us the cDNA library. We would also like to thank Dr. Takashi Ueda for critical reading of the manuscript. This research was supported by Salem-Teikyo University Graduate Research funds and by a Research Initiation Grant to M.C.-A. from the NASA-West Virginia Space Grant Consortium.

References

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Anang, Z., Miller, W. and Lipman, D.J. (1997) Gapped-Blast and Psi-Blast; a new generation of protein database search programs. *Nucl. Acids Res.* 25: 3389–3402.

Anthony, R.G., James, P.E. and Jordan, B.R. (1996) Cauliflower (*Brassica oleracea* var. *botrytis* L.) curd development: the expression of meristem identity genes. *J. Exp. Bot.* 47: 181–188.

Attwood, T.K. and Findlay, J.B.C. (1993) Design of a discriminating fingerprint for G-protein-coupled receptors. *Protein Eng.* 6: 167–176.

Attwood, T.K., Flower, D.R., Lewis, A.P., Mabey, J.E., Morgan, S.R., Scordis, P., Selley, J. and Wright, W. (1999) PRINTS prepares for the new millennium. *Nucl. Acids Res.* 27: 220–225.

Barton, M.K. and Poethig, S. (1993) Formation of the shoot apical meristem in *Arabidopsis thaliana*: an analysis of development in the wild-type and in the *SHOOT MERISTEMLESS* mutant. *Development* 119: 823–831.

Bowman, J.L., Alvarez, J., Weigel, D., Meyerowitz, E.M. and Smyth, D.R. (1993) Control of flower development in *Arabidopsis thaliana* by *APETALA1* and interacting genes. *Development* 119: 721–743.

Bradley, D., Ratcliffe, O., Vincent, C., Carpenter, R. and Coen, E. (1997) Inflorescence commitment and architecture in *Arabidopsis*. *Science* 275: 80–83.

Carr, S.M. and Irish, V.F. (1997) Floral homeotic gene expression defines developmental arrest stages in *Brassica oleracea* L. vars. *botrytis* and *italica*. *Planta* 201: 179–188.

Chen, L., Cheng, J.C., Castle, L. and Sung, Z.R. (1997) *EMF* genes regulate *Arabidopsis* inflorescence development. *Plant Cell* 9: 2011–2024.

Clark, S.E., Running, M.P. and Meyerowitz, E.M. (1995) *CLAVATA3* is a specific regulator of shoot and floral meristem development affecting the same processes as *CLAVATA1*. *Development* 121: 2057–2067.

Clark, S.E., Williams, R.W. and Meyerowitz, E.M. (1997) The *CLAVATA1* gene encodes a putative receptor kinase that controls shoot and floral meristem size in *Arabidopsis*. *Cell* 89: 575–585.

Das, O.P., Cruz-Alvarez, M., Chaudhuri, S. and Messing, J. (1990) Molecular methods for genetic analysis of maize. *Methods Mol. Cell. Biol.* 1: 213–222.

Denecke, J., De Rycke, R. and Botterman, J. (1992) Plant and mammalian sorting signals for protein retention in the endoplasmic reticulum contain a conserved epitope. *EMBO J.* 11: 2345–2355.

Feramisco, J.R., Glass, D.B. and Krebs, E.G. (1980) Optimal spatial requirements for the location of basic residues in peptide substrates for the cyclic AMP-dependent protein kinase. *J. Biol. Chem.* 255: 4240–4245.

Fletcher, J.C., Brand, U., Running, M.P., Simon, R. and Meyerowitz, E.M. (1999) Signaling of cell fate decisions by *CLAVATA3* in *Arabidopsis* shoot meristems. *Science* 283: 1911–1914.

Franco-Zorrilla, J.M., Fernández-Calvín, B., Madueño, F., Cruz-Alvarez, M., Salinas, J. and Martínez-Zapater, J.M. (1999) Identification of genes specifically expressed in cauliflower reproductive meristems. Molecular characterization of BoREM1. *Plant Mol. Biol.* 39: 427–436.

Frohman, M.A., Dush, M.K. and Martin, G.R. (1988) Rapid production of full-length cDNAs from rare transcripts. *Proc. Natl. Acad. Sci. USA* 85: 8998–9002.

Gómez-Campo, C. and Prakash, S. (1999) Origin and domestication. *In* Biology of Brassica Coenospecies. Edited by Gómez-Campo, C. pp. 33–58. Elsevier Publ., Amsterdam.

Henikoff, S. and Henikoff, J.G. (1994) Protein family classification based on searching a database of blocks. *Genomics* 19: 97–107.

Hofmann, K., Bucher, P., Falquet, L. and Bairoch, A. (1999) The PROSITE database, its status in 1999. *Nucl. Acids Res.* 27: 215–219.

Hofmann, K. and Stoffel, W. (1993) TMbase-a database of membrane spanning protein segments. *Biol. Chem. Hoppe-Seyler* 374: 166.

Kempin, S.A., Savidge, B. and Yanofsky, M.F. (1995) Molecular basis of the cauliflower phenotype in *Arabidopsis*. *Science* 267: 522–525.

Kishimoto, A., Nishiyama, K., Nakanishi, H., Uratsuji, Y., Nomura, H., Takeyama, Y. and Nishizuka, Y. (1985) Studies on the phosphorylation of myelin basic protein kinase C and adenosine 3′:5-monophosphate-dependent protein kinase. *J. Biol. Chem.* 260: 12492–12499.

Kyte, J. and Doolittle, R.F. (1982) A simple method for displaying the hydrophobic character of a protein. *J. Mol. Biol.* 157: 105–132.

Li, J. and Chory, J. (1997) A putative leucine-rich repeat receptor kinase involved in brassinosteroid signal transduction. *Cell* 90: 929–938.

Long, J.A., Moan, E.I., Medford, J.I. and Barton, M.K. (1996) A member of the *KNOTTED* class of homeodomain proteins encoded by the *STM* gene of *Arabidopsis*. *Nature* 379: 66–69.

Lütcke, H.A., Chow, K.C., Mickel, F.S., Moss, K.A., Kern, H.F. and Scheele, G.A. (1987) Selection of AUG initiation codons differs in plants and animals. *EMBO J.* 6: 43–48.

Ma, H. (1994) GTP-binding proteins in plants: new members of an old family. *Plant Mol. Biol.* 26: 1611–1636.

Mandel, M.A., Gustafson-Brown, B.S., Savidge, B. and Yanofsky, M.F. (1992) Molecular characterization of the floral homeotic gene *APETALA1*. *Nature* 360: 273–277.

Marshall, R.D. (1972) Glycoproteins. *Annu. Rev. Biochem.* 41: 673–702.

McGeoch, D.J. (1985) On the predictive recognition of signal peptide sequences. *Virus Res.* 3: 271–286.

Nakai, K. and Kanehisa, M. (1992) A knowledge base for predicting protein localization sites in eukaryotic cells. *Genomics* 14: 897–911.

Pinna, L.A. (1990) Casein kinase 2: an 'eminence grise' in cellular regulation. *Biochim. Biophys. Acta* 1054: 267–284.

Plakidou-Dymock, S., Dymock, D. and Hooley, R. (1998) A higher plant seven-transmembrane receptor that influences sensitivity to cytokinins. *Curr. Biol.* 8: 315–324.

Prakash, S. and Hinata, K. (1980) Taxonomy, cytogenetics and origin of crop brassicas, a review. *Opera Bot.* 55: 1–57.

Quiros, C.F. (1999) Genome structure and mapping. *In* Biology of Brassica Coenospecies. Edited by Gómez-Campo, C. pp. 217–245. Elsevier Publ., Amsterdam.

Ratcliffe, O.J., Bradley, D.J. and Coen, E.S. (1999) Separation of shoot and floral identity in *Arabidopsis*. *Development* 126: 1109–1120.

Sadik, S. (1962) Morphology of the curd of cauliflower. *Amer. J. Bot.* 49: 290–297.

Sanderfoot, A.A. and Raikhel, N.V. (1999) The specificity of vesicle trafficking: coat proteins and SNAREs. *Plant Cell* 11: 629–641.

Torii, K.U., Mitsukawa, N., Oosumi, T., Matsuura, Y., Yokoyama, R., Whittier, R.F. and Komeda, Y. (1996) The *Arabidopsis ERECTA* gene encodes a putative receptor protein kinase with extracellular leucine-rich repeats. *Plant Cell* 8: 735–746.

von Heijne, G. (1986) A new method for predicting signal sequence cleavage sites. *Nucl. Acids Res.* 14: 4683–4690.

Wang, Z. and Brown, D.D. (1991) A gene expression screen. *Proc. Natl. Acad. Sci. USA* 88: 11505–11509.

Weigel, D., Alvarez, J., Smyth, D.R., Yanofsky, M.F. and Meyerowitz, E.M. (1992) *LEAFY* controls floral meristem identity in *Arabidopsis*. *Cell* 69: 843–859.

(Received August 8, 2000; Accepted January 29, 2001)