Plant Molecular Biology (2005) 58:123–139
DOI 10.1007/s11103-005-5219-y

Engrailed-ZmOCL1 fusions cause a transient reduction of kernel size in maize

Abdel-Sabour Khaled[1], Vanessa Vernoud[1], Gwyneth C. Ingram[1,4], Pascual Perez[2], Xavier Sarda[3] and Peter M. Rogowsky[1,*]
[1]*RDP, UMR879 INRA-CNRS-ENSL-UCBL, IFR128 BioSciences Lyon-Gerland, ENS-Lyon, 46 Allée d'Italie, F-69364 Lyon Cedex 07, France (*author for correspondence; e-mail Peter.Rogowsky@ens-lyon.fr);* [2]*Biogemma SA, Laboratoire de Biologie Cellulaire et Moléculaire, Campus Universitaire des Cézeaux, 24 ave des Landais, F-63177 Aubière, France;* [3]*Biogemma SA, Laboratoire de Génomique, 1 rue Pierre Fontaine, F-91058 Evry Cedex, France;* [4]*Institute of Cell and Molecular Biology, Kings Buildings, University of Edinburgh, EH9 3JR Edinburgh, UK*

Received 7 March 2005; accepted in revised form 10 April 2005

Key words: development, *Engrailed*, kernel, maize, *Zea mays*, *ZmOCL* genes

Abstract

ZmOCL1 is the founding member of the *ZmOCL* (*Outer Cell Layer*) family encoding putative transcription factors of the HD-ZIP IV class. It is expressed in the L1 cell layer of the embryo and several other tissues of maize. After determination of the intron/exon structure a *mutator* insertion was isolated in the upstream region. No notable phenotypes and wildtype levels of *ZmOCL1* transcript were observed in homozygous mutant plants. In contrast transgenic plants carrying a fusion of the repressor domain of the *Drosophila Engrailed* gene with the DNA binding and dimerisation domains of ZmOCL1 showed a transient reduction of embryo, endosperm and kernel size that was most obvious around 15 DAP. An inverse relationship was observed between the degree of size reduction and the expression level of the transcript. In reciprocal crosses the size reduction was only observed when the transgenic plants were used as females and no expression of male transmitted transgenes was detected. Smaller kernels resembled younger kernels of wild-type siblings indicating that interference with ZmOCL1 function leads to an overall slow-down of early kernel development. Based on marker gene analysis ZmOCL1 may act *via* a modification of gibberellin levels. Phylogenetic analyses based on the intron/exon structure and sequence similarities of ZmOCL1 and other HD-ZIP IV proteins from maize, rice and *Arabidopsis* helped to identify orthologues and suggested an evolution in the function of individual genes after the divergence of monocots and dicots.

Introduction

Numerous developmental events are under transcriptional control and involve a large variety of target gene-specific transcription factors. Highly conserved DNA binding domains are the basis for their classification into families such as the HD family characterised by the presence of the homeo domain. HD proteins have been found in fungi, animals and plants where they play important roles in diverse developmental pathways (Gehring *et al.*, 1994; Chan *et al.*, 1998). However, the association of a leucine zipper (ZIP) with the HD has only been found in plants raising the question whether HD-ZIP proteins are involved in plant-specific functions (Ruberti *et al.*, 1991; Riechmann *et al.*, 2000). The ZIP domain is essential for the function of these transcription factors because it is

required for dimerisation; contrary to other HD proteins HD-ZIP proteins can only bind to DNA as dimers (Sessa *et al*., 1993). A sub-family of HD-ZIP proteins is characterised by the presence of an additional conserved region, the StAR-related lipid transfer (START) domain (Ponting and Aravind, 1999). In animals this domain has been shown to be involved in lipid/sterol binding and transport. In contrast, in plants it has primarily been found in transcription factors of the HD-ZIP family suggesting a different function more related to signal transduction and possibly involving steroid-type phytohormones. The basic structure of HD-ZIP proteins containing a START domain is reminiscent of nuclear receptors in animals that also reunite a transcription factor (Zinc finger family) with a steroid-binding hormone receptor domain (HR) and are known to control numerous metabolic and developmental pathways (Schrick *et al*., 2004). Both the plant-specificity and the unique association with the START domain raised a particular interest as to the function of HD-ZIP-START proteins.

Soon after their discovery HD-ZIP-START proteins have been divided into the HD-ZIP III and the HD-ZIP IV sub-family based on the presence of a loop in the middle of the ZIP domain of the HD-ZIP IV sub-family, which has also been named HD-ZIP GL2 or simply GL2 family after its founding member, the *Arabidopsis* Glabra2 protein (Di Cristina *et al*., 1996). Interestingly a phylogenetic analysis of the START domain confirms this separation and reveals an important divergence, in particular at the N-terminal part of the START domain (Schrick *et al*., 2004). Even more intriguingly the numerous functional studies show that HD-ZIP III proteins are frequently involved in the establishment of organ polarity, while HD-ZIP IV proteins almost exclusively seem to play a role in the differentiation and/or maintenance of the L1 or epidermal layer. For example, the semidominant *Phabulosa* (*PHB*) or *Phavoluta* (*PHV*) mutations cause an adaxialisation of lateral organs that manifests itself among others by the presence of trichomes and axillary meristems on the entire leaf surface rather than just on the adaxial surface (McConnell *et al*., 2001). Conversely lateral organs of the loss-of-function *revoluta* (*rev*) mutant lack adaxial features such as axillary meristems (Talbert *et al*., 1995). All three genes belong to the HD-ZIP III class and are expressed adaxially in wildtype plants (Bowman *et al*., 2002). Although all semi-dominant mutations of *PHB*, *PHV* and *REV* cluster in the N-terminal part of the START domain the phenotypes are not caused by a functional alteration of the START domain but rather by a change in a miRNA binding site (Emery *et al*., 2003), which is not conserved in HD-ZIP IV genes.

Ten HD-ZIP IV genes have been identified in the *Arabidopsis* genome (Tavares *et al*., 2000). Mutations in *Gl2* lead to trichome abortion and an increased number of root hairs (Rerie *et al*., 1994) and mutations in *Anl2* cause suppression of anthocyanin accumulation in the subepidermal layer (Kubo *et al*., 1999). While single mutations in *AtML1* or *PDF2* do not result in any notable phenotype, double mutants fail to survive after germination and their leaves seem to lack an epidermis (Abe *et al*., 2003). While no functional data are available for other HD-ZIP IV genes, detailed *in situ* or *Gus* fusion expression patterns have been established for *Gl2* (Masucci *et al*., 1996), *AtML1* (Lu *et al*., 1996) and *PDF2* (Abe *et al*., 2003) in *Arabidopsis*, five different *ZmOCL* genes in maize (Ingram *et al*., 2000) *Roc1* in rice (Ito *et al*., 2002) and *PaHB1* in Norway spruce (Ingouff *et al*., 2001). All of these HD-ZIP IV genes are expressed specifically in the outer cell layer of the respective plant organs with the exception of *ZmOCL2*, which is expressed in the subepidermal layer.

The differentiation of an epidermis is one of the three major events in plant embryogenesis (Kaplan and Cooke, 1997). This radial pattern formation occurs after the apico-basal pattern formation and precedes the formation of the shoot apical meristem (SAM) and the root meristem. It has even been postulated that the differentiation of a protoderm and the resulting organisation in outer and inner cell layers is a prerequisite for meristem formation (Bommert and Werr, 2001). The expression of the above HD-ZIP IV genes in the outermost cell layer of the embryo prior to any microscopically visible differentiation of a protoderm makes these HD-ZIP IV transcription factors prime candidates for a role in this important developmental step. Furthermore the distinct albeit sometimes overlapping expression patterns of *ZmOCL* genes in different parts of the embryo suggest the existence of genetically defined territo-

ries within the protoderm (Ingram *et al.*, 2000). To further elucidate the function of individual HD-ZIP IV genes and their role in epidermis differentiation we characterised transgenic maize plants expressing a dominant-negative *ZmOCL1* allele.

Materials and methods

Plant material

The maize inbred line A188 (Gerdes and Tracy, 1993), the maize stock Rscm2 (Heckel *et al.*, 1999), the maize hybrids DH5 × DH7 (Barloy *et al.*, 1989) and Hi-II (Armstrong, 1994) as well as all maize mutants and transgenic plants were grown either in a growth chamber with a 16 h illumination period (100 Wm^{-2}) at 24/19 °C (day/night) and 80% relative humidity or in a greenhouse under the same conditions without regulation of the relative humidity. All plants were propagated by hand pollination.

Genomic library screen

A genomic library of maize hybrid HD5 × HD7 in vector λEMBL3 SP6/T7 (Clontech) was screened with the radioactively labeled insert of plasmid L65/4. This 1973 bp fragment corresponded to a partial cDNA of *ZmOCL1* excluding the homeo domain and the leucine zipper (Figure 1). DNA was isolated from single plaques and subcloned in pBluescript SK + or pBCSK + (Stratagene).

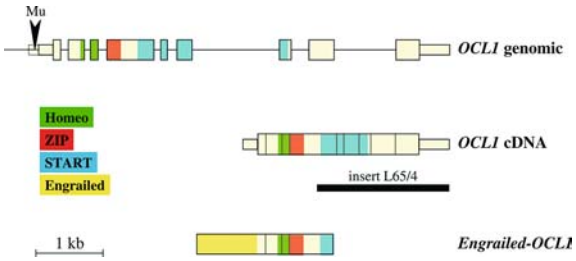

Figure 1. Schematic drawings of the *ZmOCL1* genomic DNA, cDNA and *Engrailed* fusion. Exons are drawn as wide boxes, 5′ and 3′ untranslated regions as narrow boxes and promoter regions and introns as lines. Functional domains are colour coded. The insertion site of the transposon *mutator* upstream of *ZmOCL1* is indicated by an arrow head and the probe used for the screen of the genomic library by a fat horizontal bar.

Cloning and sequence analysis

If not specified otherwise, all DNA and RNA manipulations were performed according to (Sambrook *et al.*, 1989). For nucleotide sequence analysis, template DNA was isolated with the QIAprep spin plasmid miniprep kit (Qiagen, Courtaboeuf, France), sequenced with the Thermo Sequenase™ II dye terminator cycle sequencing kit (Amersham Biosciences, Saclay, France) and run on an ABI 3100 capillary sequencer (Applied Biosystems, Courtaboeuf, France). Raw sequences were assembled with Sequencher software (GeneCodes, AnnArbor, USA) and clean sequences manipulated with the DNAstar package (Lasergene, Madison, USA). Nucleotide sequences were compared to EMBL and Genebank databases using the BLAST algorithm (Altschul *et al.*, 1997). The genomic sequence reported here was given the EMBL accession no. AJ748731.

Construction of a ZmOCL1-Engrailed *fusion*

The vector pCHRIS containing a *35S* promoter, an Ω leader sequence, the repressor domain of the *Drosophila Engrailed* gene and a *CaMV* terminator (Markel *et al.*, 2002) was kindly provided by Wolfgang Werr. The N-terminal part of ZmOCL1 was amplified with primers A10EN2 (5′ GAGGATCCATGCAGTTCCCGTTCACGACC 3′) and ART6 (5′ TCAATGATGACGAGGCCAGACTCCC 3′) and cloned into the BamHI site of pCHRIS (Figure 1). The entire cassette "Ω leader-*Engrailed-ZmOCL1*-terminator" was then transferred under the control of the rice actin promoter including its intron (Cao *et al.*, 1992).

Phylogenetic analysis

For the comparison of the intron/exon structure of HD-ZIP IV genes the genomic sequences with the following Genbank accession nos. were used: AJ748731 (*ZmOCL1*), AP004656 (*OsRoc1*), AL606669 (*OsRoc2*), AE017122 (*OsRoc3*), AL663003 (*OsRoc4*), AP005885 (*OsRoc5*), AP005864 (*OsRoc6*), AP003896 (*OsRoc7*), AP004758 (*OsRoc8*), AP003274 (*OsRoc9*), AB056455 (*AtPDF2*), AL035527 (*AtML1*), AF178688 (*AtFWA*), AF077335 (*AtANL2*), Z54356 (*AtGL2*), AJ224338 (*AtGL2-1*), AC000098

(*AtGL2-2*), AC005700 (*AtGL2-3*), Z97344 (*AtGL2-4*) and AB0 13394 (*AtGL2-5*).

For phylogenetic trees the following protein accession nos. were used: CAG38614 (ZmOCL1), CAB96422 (ZmOCL2), CAB96423 (ZmOCL3), CAB96424.2 (ZmOCL4), CAB96425 (ZmOCL5), AAR97952 (ZmRLD), BAB85750 (OsRoc1), BAC77155 (OsRoc2), BAC77156 (OsRoc3), BAC 77157 (OsRoc4), BAC77158 (OsRoc5), BAB58961 (AtPDF2), AAB49378 (AtML1), AAG09302 (AtFWA), AAD47139 (AtANL2), CAA91183 (AtGL2), CAB45018 (AtGL2-1), AAB71455 (AtGL2-2), AAC69941 (AtGL2-3), CAB10551 (AtGL2-4), BAB10227 (AtGL2-5), AAG43405 (PaHB1), AAL 83725 (PaHB2), CAD29659 (AtPHB), CAD 29544 (AtPHV) and BAB09842 (AtREV). Sequences BAC77155, BAC77156 and BAC77158 were extended based on the corresponding genomic sequences (see above) and sequences CAB10551 and BAB10227 modified according to Tavares *et al.* (2000). In the case of *OsRoc6*, *OsRoc7*, *OsRoc8* and *OsRoc9* only very short cDNA sequences were present in the databases. Complete protein sequences were obtained with the gene prediction programs GrailExp (http://grail.lsd.ornl.gov/grailexp/), Genscan (http://genes.mit.edu/GENSCAN.html), and Splice Predictor (http://bioinformatics.iastate.edu/cgi-bin/sp.cgi) followed by conceptual translation of the corresponding genomic sequences. Sequences were aligned by the Clustal method (Higgins and Sharp, 1988) and phylogenetic trees were constructed by the neighbour joining (NJ) method with pairwise gap removal and 500 bootstrap replicates using the PhyloWin software (Galtier *et al.*, 1996).

Screen for mutator *insertions*

Segregating families of 16 *emb* mutants of the collection described by Clark and Sheridan (1991) were grown. Genomic DNA was isolated from pools of at least 5+/+ plants and 5+/*emb* plants and screened with primers L93E2 (5′ GAGATGGCGATGAGACCC 3′), L93b7-4 (5′ ACACCGGGGAGGACAGGG3′), A10GH11 (5′ AGTACAGCACTTATGTGGG 3′) or A10PO2 (5′ GATGCGAACGGCATGGTGCGTGCGTGC 3′) in combination with primer Mu9242 (5′ GAGAAGCCAACGCCAWCGCCTCYATTTCGTC 3′). The PCR band obtained in mutant *emb*-8530* with the primer pair L93b7-4/Mu9242 was cloned and sequenced.

Reverse transcription polymerase chain reaction

Total RNA from immature kernels (approximately 500 mg) was extracted as described in Magnard *et al.* (2004) and resuspended in 30 μl RNAse free water. In the case of leaf tissue a disk of approximately 100 mg was ground to powder in a Mixer Mill MM300 (Qiagen, Courtaboeuf, France) and total RNA extracted with 1 ml TRIzol® Reagent according to the instructions of the supplier (Invitrogen, Cergy Pontoise, France). After ethanol precipitation the RNA was resuspended in 30 μl RNAse free water.

In both cases RNA was treated with RNase free DNAse and the DNAse inactivated according to the instructions of the supplier (AMBION, Montrogue, France). Approximately 5 μg of total RNA were reverse transcribed using random hexamers (Amersham Biosciences, Saclay, France) and reverse transcriptase without RNaseH activity (Fermentas) in a final volume of 20 μl. 2.5×10^5 copies of GeneAmplimer pAW109 RNA (Applied Biosystems, Courtaboeuf, France) were added to the reverse transcription reaction. The cDNA was diluted 50 times and 5 μl used for amplification by PCR in a volume of 20 μl.

The gene-specific primers used for RT-PCR experiments were En-A (5′ACAGTCAGCGCCCAGCCCCATTACC 3′) and En-B (5′ CCTCCTCCGCCTCCTCCTCCTCGTT 3′) for the *Engrailed-ZmOCL1* transgene, A10DEL-C1 (5′ TCT GCACGTGAATGGAGTAAG 3′) and A10DEL-C2 (5′ ACTGGAGCATACACGACCATC 3′) for the endogenous copy of *ZmOCL1*, LTP2-N3 (5′CTGCAACTGCCTCAAGAACG 3′) and LTP2-C3 (5′CACACACAAACACAAGGACG A3′) for *Ltp2*, GA20-F (5′ AACATC GGCGAC ACCTTCAT 3′) and GA20-R (5′ CCTTCCCTAAATGGATCATCAGC 3′) for *ZmGA20*, as well as Pin1a-F (5′ CATTCTCAGCACCGCAGTCAT 3′) and Pin1a-R (5′ ACCTCCCACACTTTCCATCCA 3′) for *ZmPin1*. The primers for *ZmGA20* were based on the consensus sequence of the EST sequences with the accession nos. AI920454, AW042221, AY105651, CD973075, CD974784 and CD975344 and those of *ZmPin1* on the consensus sequence of the EST sequences with the accession nos. AW498032, AW499200, AY108500, BI389170, BI430705, BM080181, BM 335134, BM339018, BM378630 and BM378734. The constitutively expressed 18S rRNA gene

(primers 5′ CCATCCCTCCGTAGTTAGCTTCT 3′ and 5′ CCTGTCGGCCAAGGCTATATAC 3′) was used as an internal control of RNA quantity and GeneAmplimer pAW 109 RNA (primers 5′ CATGTCAAATTTCACTGCTTCATC 3′ and 5′ TGACCACCCAGCCATCCTT 3′) as positive control of the RT-PCR efficiency. In order to get semi-quantitative results, the number of cycles of the PCRs were adjusted for each gene to obtain barely visible bands in agarose gels. Aliquots of the PCR were loaded on agarose gels and stained with ethidium bromide.

Quantitative RT-PCR

The cDNA was obtained as described above and quantitative PCR carried out on a GeneAmp® 5700 Sequence Detection System (Applied Biosystems, Courtaboeuf, France). Data were normalised using Ct values for the constitutively expressed *Gapdh* gene and the *Engrailed-ZmOCL1* transgene. Real time PCR efficiencies (E) were calculated for each gene by measuring the Ct on two replicates of a dilution series (2^n with $n = 0$ to 7). The relative expression rates of target gene transcripts were calculated (Pfaffl, 2001) and expressed as a percentage of the value obtained in the non-transformed control plant.

Maize transformation

Embryogenic type II calli were initiated by culture of immature embryos (10 JAP) of maize hybrid Hi-II and maintained by regular subculturing for 3 to 6 months (Armstrong, 1994). Plasmid DNA was isolated with the QIAprep spin plasmid miniprep kit (Qiagen, Courtaboeuf, France) and coated onto Tungsten (M10) particles (Klein *et al.*, 1992). Callus pieces of 20 mm^2 were placed in the centre of a Petri dish with a culture medium of high osmotic pressure (0.2 M mannitol, 0.2 M sorbitol) 4 h prior to transformation. They were co-bombarded with the plasmid of interest and plasmid pDM302 carrying a *pat* gene under the control of a rice actin promoter (Cao *et al.*, 1992) using a particle inflow gun (Finer *et al.*, 1992). One day after bombardment the calli were transferred to a medium of normal osmotic pressure containing 2 mg/l Glufosinate-Ammonium (Riedel-de Haën, St. Quentin Fallavier, France) as selective agent. The calli were subcultured under selective pressure every 2 weeks for 3 months and then placed on regeneration medium to regenerate plants (Vain *et al.*, 1993).

Micro-array analysis

The Genoplante maize micro-array contained 26404 cDNAs spotted in duplicate on two microscope slides. This unigene set had been established by clustering the Genoplante maize EST data (http://genoplante-info.infobiogen.fr) with all publicly available EST data. A biological replica of the hybridisation was carried out and each hybridisation included a dye swap of the Cy5 or Cy3 labelled probes. The data were analysed with Spotfire software (Somerville, MA). To be considered as differentially expressed the difference in expression of a gene had to be at least two fold and the P value of the 8 data points had to be below 10^{-3}.

Cytological sections

Maize ovaries were collected at the emergence of silks and caryopses at fixed intervals after controlled pollination. The lateral parts were removed with two longitudinal cuts to allow better penetration of the fixative. The remaining central third of the ovary or caryopsis was immediately fixed in 4% PFA (4% para-formaledhyde, 0.1 M phosphate buffer pH 7.0). After vacuum infiltration for 10 min the fixative was renewed and the samples incubated at 4 °C for 16 h. Standard protocols were used for dehydration and paraffin embedding (Paraplast Plus, Sherwood Medical, St. Louis, Missouri). Microtome sections of 10 μm were placed on ProbeOn™ Plus microscope slides (Fisher Scientific, Pittsburgh, Pennsylvania) and dewaxed with Histo-Clear II (National Diagnostics, Atlanta, Georgia). The periodic acid-Schiff stain was carried out by 10 min oxidation with 1% periodic acid (Sigma, St. Quentin Fallavier, France), 10 min rinsing with tap water and 5 min staining with Schiff reagent (Sigma, St. Quentin Fallavier, France). After dehydration the slides were mounted with ENTELLAN rapid mounting medium (Merck, Darmstadt, Germany) and observed under a Leica MZ16FA fluorescent stereomicroscope. Images were digitised with a Leica DC300F digital camera and analysed with IM1000 imaging software (Leica Microsystems, Rueil-Malmaison, France).

For length measurements the most central section in the embryo sac or embryo was chosen. Eight kernels of each size class were measured per ear and average values calculated. All measurements were carried out on the adaxial side of the kernel linking the most distal points of the pericarp, endosperm or embryo by a straight line. In the case of incomplete suspensors in the measured cytological section the suspensor was virtually extended to intersect the pericarp.

Measurement of ovary and kernel mass

From a single ear 50 ovaries or kernels of each size class were weighed together on a balance before (wet weight) or after drying for 12 h at 80 °C (dry weight). Average values were calculated from 3 ears per stage and type of cross.

Results

Intron/exon structure of ZmOCL1

To prepare the search for insertional mutants in *ZmOCL1*, the complete genomic nucleotide sequence was determined. A genomic lambda library of hybrid DH5 × DH7 (Barloy *et al.*, 1989) was screened with a 1.97 kb partial cDNA of *ZmOCL1* excluding the homeo domain and the leucine zipper (Figure 1). Eight strongly hybridising and 7 weakly hybridising lambda clones were isolated. The genomic sequence of *ZmOCL1* was established from overlapping subclones of two strongly hybridising lambda clones and given the accession no. AJ748731. The comparison with the previously published cDNA sequence (Ingram *et al.*, 1999) (accession no. Y17898) revealed the presence of 8 introns (Figure 1).

The amino acid sequence deduced from the genomic sequence had an N-terminal extension of 19 aa compared to the sequence deduced from the cDNA sequence (accession no. CAB51059). This difference was due to a single base insertion resulting in an in frame ATG located further upstream. Since this insertion was also present in a GSS sequence (Accession no. BZ312574) of the ongoing maize sequencing programme (Martienssen *et al.*, 2004), we believe that the ZmOCL1 protein is actually longer than originally published.

Phylogeny of ZmOCL1

The intron/exon structure of as well as multiple alignments of the HD-ZIP domain had been used previously to establish the phylogenetic relationship of HD-ZIP IV genes or proteins (Tavares *et al.*, 2000; Ingouff *et al.*, 2001). We included *ZmOCL1* in the previous data set and complemented the study with an exhaustive data mining of the rice genome. Starting with the complete or partial cDNA sequences of 9 HD-ZIP IV rice genes deposited in the databases by Matsuoka and co-workers we identified the corresponding genomic sequences and obtained the intron/exon structure and full length protein sequences for all of them using EST based gene prediction programs followed by conceptual translation.

ZmOCL1 contained 7 of the 9 conserved intron positions defined previously (Tavares *et al.*, 2000). It shared the absence of positions 3 and 8 with the *Arabidopsis* genes *AtFWA*, *AtANL2* and *AtGL2-1* and the rice genes *OsRoc4*, *OsRoc5* and *OsRoc6* (Figure 2A). Two other major groups were defined by the absence of position 3 and the presence of both positions 3 and 8. *OsRoc8* was the only gene with position 3 but lacking position 8. The most surprising structure was observed for *OsRoc7*, which lacked 6 of the 9 conserved intron positions. One may speculate that this gene arose from a partially spliced HD-ZIP IV gene, possibly by re-insertion of flanking retro-transposons into the genome.

The construction of phylogenetic trees allowed to clearly identify the orthologues of Zm*OCL1* or other *ZmOCL* genes in rice and, to a lesser extent, in *Arabidopsis*. Contrary to the previous authors we used full length protein sequences and included four HD-ZIP III proteins as an outgroup (Figure 2B). ZmOCL1 was most similar to OsRoc5, ZmOCL2 to OsRoc4, ZmOCL3 to OsRoc6 and ZmOCL4 to OsRoc3. The close similarity of ZmOCL5 with both OsRoc1 and OsRoc2 was further evidence for the existence of an additional, closely related maize gene suggested earlier (Ingram *et al.*, 2000). The presence of OsRoc7 in the clade was not necessarily indicative of the presence of a third OCL5 like protein because OsRoc7 has clearly arisen after the divergence of monocots and dicots and likely after the separation of maize and rice. The relationship between maize and *Arabidopsis* genes was more difficult to

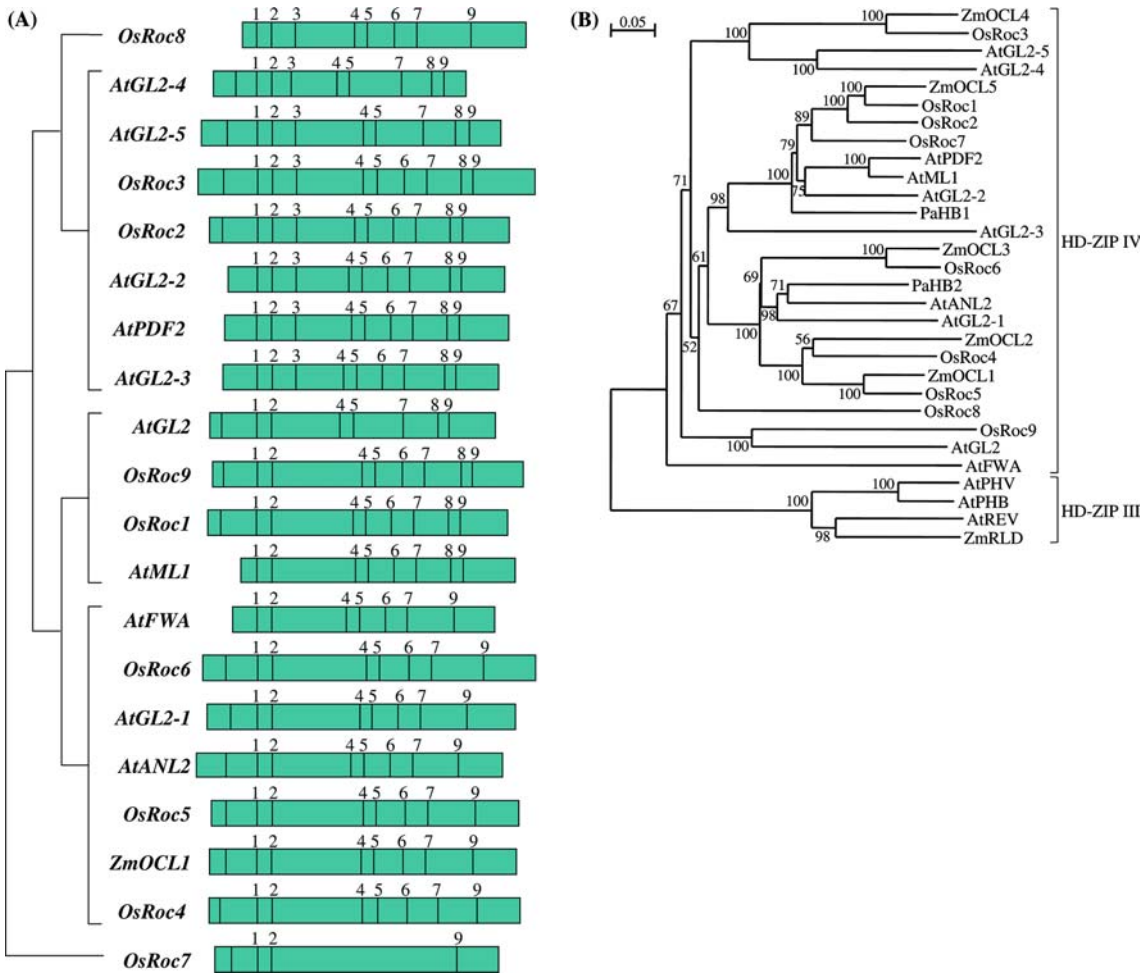

Figure 2. Phylogenetic relationship of *ZmOCL1* and related HD-ZIP IV genes. (A) Intron/exon structure of selected HD-ZIP IV genes. Intron positions are indicated as vertical bars. Conserved intron positions are given the same number in all genes, the numbering starting with a conserved intron in the homeo domain. The accession nos. of the sequences used are listed in Materials and methods. The figure is extension of the work presented by Ingouff *et al*. (2001). (B) Phylogenetic tree of an full lengh protein sequences of selected HD-ZIP IV proteins. The accession nos. of the protein sequences used and the construction of the tree by the NJ method and are described in Material and methods. The numbers are bootstrap values.

establish. In the clade of the three maize proteins ZmOCL1, ZmOCL2 and ZmOCL3 only two *Arabidopsis* proteins were present, namely AtANL2 and AtGL2-1. In trees based only on HD-ZIP or START domains ZmOCL1 was most similar to AtGL2-1 and ZmOCL2 to AtANL2 (data not shown). In the tree based on full length sequences ZmOCL4 fell in the same clade as AtGL2-4 and AtGL2-5 and ZmOCL5 clustered with AtML1, AtPDF2 and AtGL2-2.

In summary the intron/exon structure and the phylogenetic trees gave rather similar results. The most similar genes to *ZmOCL1* were *OsRoc5* in rice and *AtGL2-1* or *AtANL2* in *Arabidopsis*. The existence of more than the 5 published HD-ZIP genes in the maize genome was very likely.

Isolation and characterisation of a mutator *insertion upstream of* ZmOCL1

During a systematic screen of a small collection of kernel mutants for *mutator* insertions in *ZmOCL1*, an insertion was detected 261 bp upstream of the start codon of mutant *emb*-8530* (Clark and Sheridan, 1991). DNA gel blot analysis allowed us to estimate the size of the insertion at

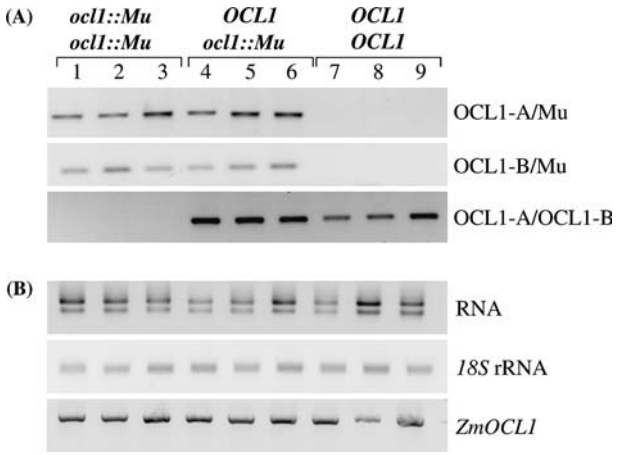

Figure 3. Characterisation of a *mutator* insertion upstream of *ZmOCL1*. The structure of the *ZmOCL1* locus in the maize genome (A) and the expression of *ZmOCL1* in leaves (B) were analysed in three plants homozyogous for the *mutator* insertion upstream of *ZmOCL1* (lanes 1 to 3), three heterozyogous plants (lanes 4 to 6) and three wildtype plants (lanes 7 to 9). (A) A PCR was carried out on genomic DNA with the primer pairs indicated on the right. (B) The expression of *ZmOCL1* and the constitutively expressed *18S* rRNA gene were visualised by RT-PCR analysis of total leaf RNA.

approximately 3 kb (data not shown). A co-segregation analysis of 30 individuals excluded a genetic linkage between the insertion and the *emb* phenotype. Nevertheless the mutation was back-crossed over three generations to look for more subtle phenotypes. After self-pollination plants homozygous for the insertion were identified by PCR with primers flanking the insertion (Figure 3A). However, the homozygous *ocl1::Mu*/*ocl1::Mu* plants showed normal *ZmOCL1* transcript levels in leaves (Figure 3B) and no cytological aberration in the embryo (data not shown).

It was surprising that an insertion close to the putative transcription and translation start sites had no influence on the transcript level. The insertion site was located in a 52 bp region that was conserved at 83% between *ZmOCL1* and its rice orthologue *OsRoc5* but separated from the putative START codon by a non-conserved stretch. In rice this conserved region seemed to be part of a rather extensive 5′UTR of 361 bp as indicated by the perfect alignment of an EST (accession no. CK074793) with the genomic sequence. Transposing these data to maize implied that the true 5′UTR was longer than inferred by the longest cDNA clone (dashed box in Figure 1) and that the transcription start site was situated upstream of the *mutator* insertion. It was tempting to speculate that part of the resulting chimeric 5′UTR of more than 3 kb was spliced, either at natural splice sites present in maize but absent in rice or at novel splice sites created by the conjunction of *mutator* and *OCL1* sequences. An alternative explanation for the *ZmOCL1* transcript in the *ocl1::Mu*/*ocl1::Mu* plants was transcription from a promoter in the inverted repeat of *mutator* as in the case of the *hcf106* mutant (Barkan and Martienssen, 1991).

Production of transgenic maize plants harbouring a ZmOCL1-Engrailed *fusion*

The functional analysis of transcription factors allows the use of specific tools in addition to the general approaches for gene inhibition. One of them is the translational fusion of the repressor domain of the *Drosophila Engrailed* gene to the transcription factor of interest (John *et al*., 1995). In animals as in plants the fusion proteins generally have dominant-negative functions and mimic the phenotype of knockouts (Markel *et al*., 2002). The N-terminal half of ZmOCL1 containing the homeo domain, leucine zipper and a small part of the START domain was fused to the *Engrailed* repressor domain and placed under the control of the constitutive rice actin promoter (Figure 1). Transgenic maize plants were produced *via* biolistics, pollinated by a standard line and T1 seed obtained for 21 independent transformation events.

Transient reduction of kernel size in plants harbouring a ZmOCL1-Engrailed *fusion*

For the phenotypic characterisation at least two hemizygous T1 plants per transformation event were selected based on their resistance to the herbicide Basta. The expression of the *Engrailed-ZmOCL1* fusion was evaluated on leaf material by semi-quantitative RT-PCR with *Engrailed*-specific primers. Only 5 transformation events showed an expression of the *Engrailed-ZmOCL1* fusion that was clearly detectable at 32 PCR cycles. Quantitative PCR was carried out on these 5 events (H302, H305, H310, H313 and H321) and the control event H308 that was Basta resistant but had not shown any tangible *Engrailed-ZmOCL1* expression (Figure 4A). Compared to the background level of the control event H308, transgene expression was increased in events H313 and H302 by a factor of 7.4 and 5.9, respectively.

After self-pollination kernels were harvested at 12 DAP and the size of the kernel, the endosperm and the embryo was measured in cytological sections. Two size classes (Figure 5) with a segregation ratio close to 1:1 could be clearly distinguished in the case of events H313 and H302. Only the values of the smaller class are represented in Figure 4A. For the other three transformation events no visible distinction of size classes was possible and average values were calculated among randomly chosen kernels. A clear inverse relationship existed between the expression level of the *Engrailed-OCL1* transgene and the size of the kernel, the endosperm and the embryo. The reduction in size was more important in the case of H313 than in H302 or the other three transgenics and was more pronounced in the case of the embryo (21% of normal size for H313) than the endosperm (37%) and the kernel (63%). The kernel, endosperm and embryo sizes of the control plant H308 were comparable to those habitually observed for the standard line A188.

A closer examination of the cytological sections showed that the embryos in the small kernels of events H313 and H302 were delayed in their development in comparison to the wildtype control (Figure 4B). The embryos of event H313 displayed transition stage morphology normally found at 7 DAP. They exhibited radial rather than bilateral symmetry and lacked coleoptile or scutellum. The embryos of event H302 resembled coleoptilar embryos typically found at 10 DAP. Taken together these results indicated that strong expression of the *Engrailed-OCL1* transgene might have caused a developmental delay that manifested itself by a reduced kernel size at 12 DAP.

Need for maternal expression of the Engrailed-ZmOCL1 *transgene*

The reduction in size concerned less than half of the kernels obtained after self-pollination of hemizygous plants. This result was unexpected for a dominant transgene that should affect $\frac{3}{4}$ of the kernels. To clarify the genetics reciprocal crosses were carried out between hemizygous T1 plants of event H313 and the standard line A188. The cross T1 × A188 using the transgenic as female resulted in two size classes at 15 DAP (Figure 5A), whereas only a single homogenous class was observed in the reciprocal cross A188 × T1 (Figure 5B). Measurements of the kernel weight at 9 and 15 DAP allowed to quantify this observation (Figure 5C) and proved that the differences between small and big kernels of the same ear as well as between small kernels and the kernels of the reciprocal cross were statistically significant at 15 DAP. The further development of the small kernels was variable. Marking of small kernels with a permanent marker pen and observations at 25 DAP (Figure 5D) and maturity (Figure 5E) showed that the difference between small and big kernels became less and less noticeable, although a number of small kernels could still be clearly distinguished at maturity. Finally measurements of the plant height at the 7 leaf stage or at maturity excluded a pleiotrophic effect of the transgene (data not shown).

While it was easy to select big or small kernels from T1 × A188 ears for measurements or RNA isolation at 15 DAP, it was not possible to determine precisely the ratio of small and big kernels. This was mainly due to a wide range of intermediate sizes on a given ear but also to a certain variability between ears. To obtain a more precise segregation ratio 20 randomly chosen mature kernels of two T1 × A188 crosses (H313 sister plants) were germinated. Genotyping of the 40 seedlings by PCR showed that 11 and 7 of them carried the *Engrailed-ZmOCL1* transgene, respectively. The overall ratio of 45% transgenic kernels was close to the 1:1 segregation expected for a dominant transgene.

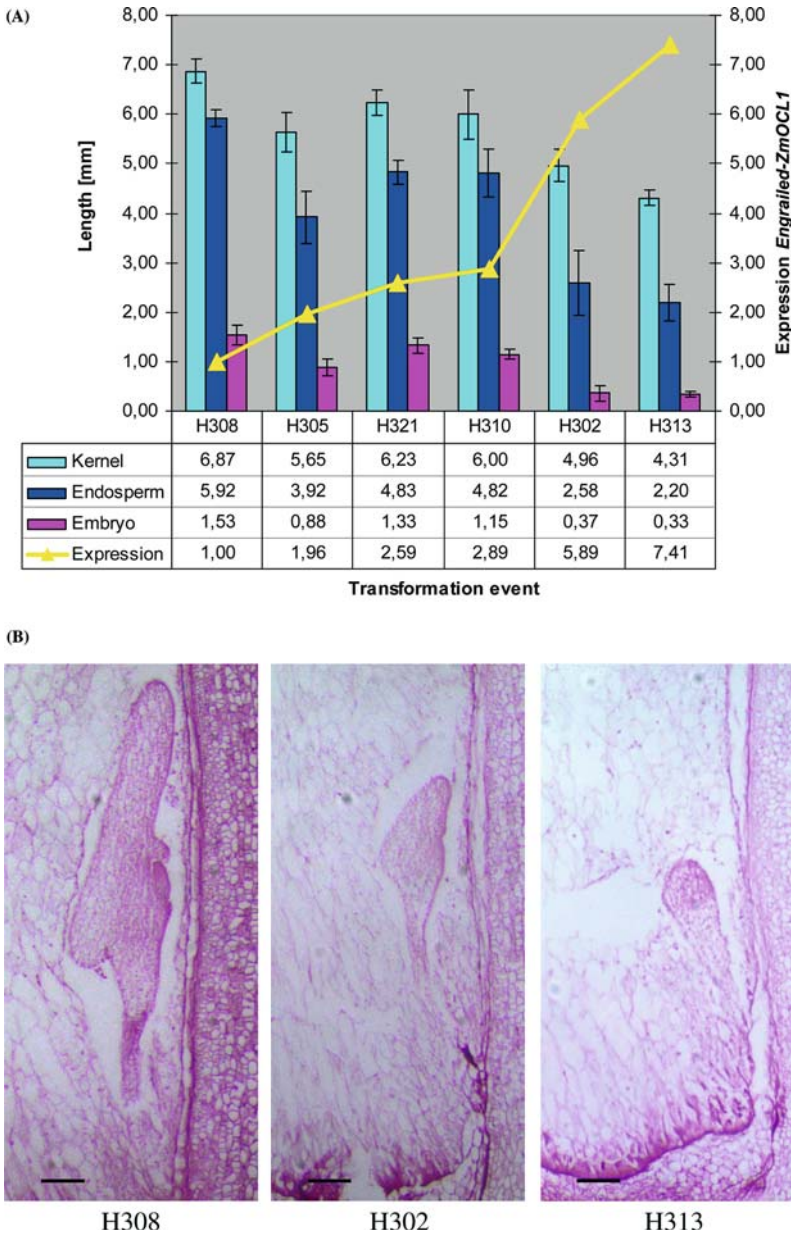

Figure 4. Inverse correlation between transgene expression and kernel size. (A) Total leaf RNA of hemizygous T1 plants corresponding to 5 different transformation events expressing the *Engrailed-ZmOCL1* transgene and the non-expressing control plants H308 was used to quantify the expression level of the *Engrailed-ZmOCL1* transgene by quantitative RT-PCR with *Engrailed*-specific primers. After pollination by the standard line A188 the resulting T2 kernels were harvested at 12 DAP, embedded in wax and sectioned to allow length measurements of the kernel, endosperm and embryo. Average values were calculated from 8 different kernels. Error bars indicate the standard deviation. (B) Schiff stain of cytological sections at 12 DAP.

To irrevocably link small kernel size to transgene expression semi-quantitative RT-PCR experiments were carried out. In fact, the heterogenous kernel size at 12 or 15 DAP had been shown to depend on the presence of the *Engrailed-ZmOCL1* transgene but it remained to be shown that the small but not the large kernels carried and expressed the transgene in embryo and endosperm. From T1 × A188 kernels the maternal pericarp tissue was removed

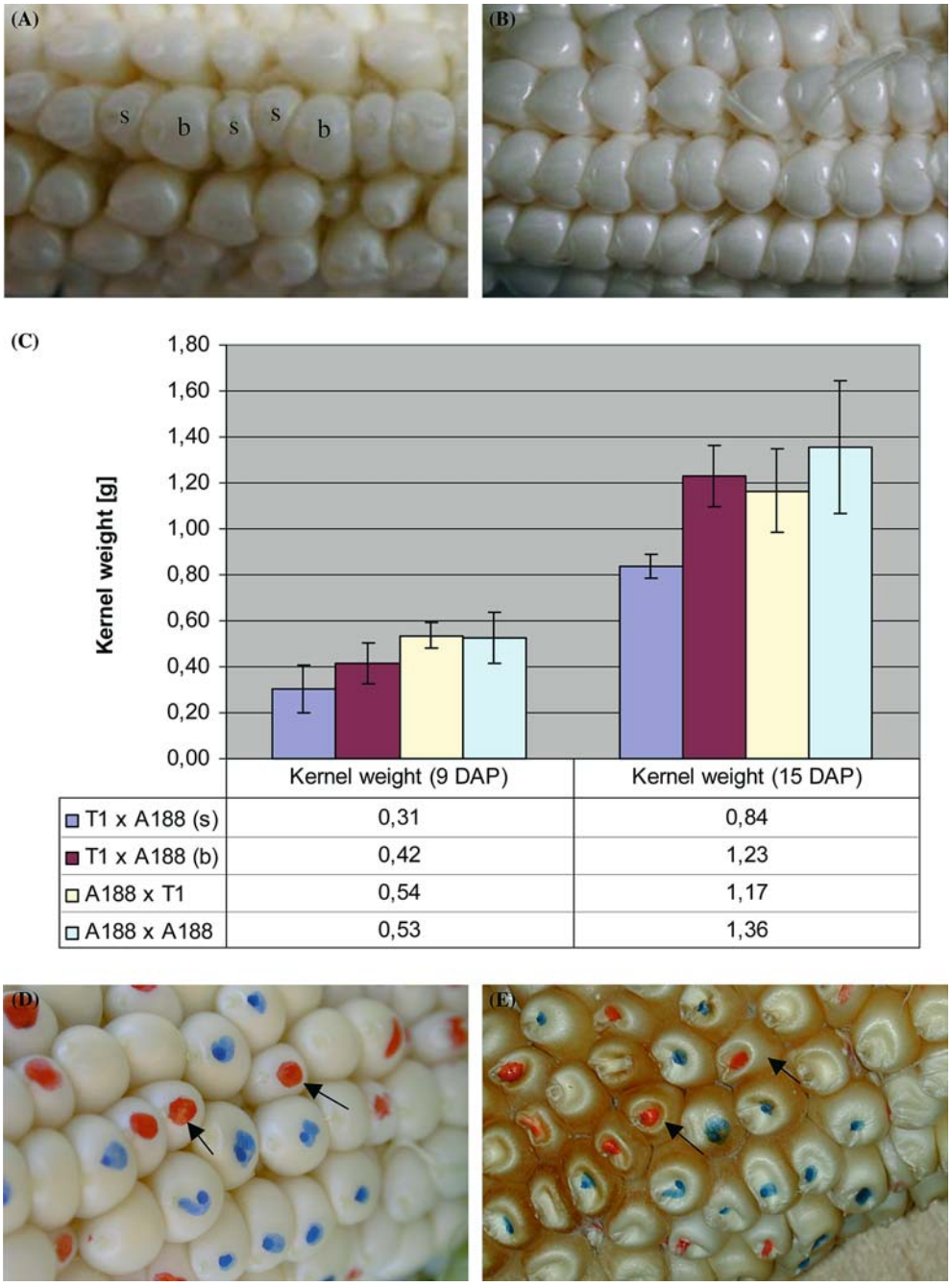

Figure 5. Differences in kernel morphology after reciprocal crosses. Ears obtained by reciprocal crosses of hemizygous T1 plants with the standard line A188 were observed at 9 or 15 DAP. (A) Photograph of an ear resulting from a cross T1 × A188 at 15 DAP. s, small; b, big. (B) Photograph of an ear resulting from a cross A188 × T1 at 15 DAP. (C) Quantification of the kernel wet weight at 9 DAP or 15 DAP. Error bars indicate the standard deviation (D) Ear from a T1 × A188 cross at 25 DAP with small kernels marked in red and big kernels marked in blue. (E) The same ear as in (D) at maturity. Arrows indicate kernels that were small at 25 DAP but normal size at maturity.

and RNA isolated from the entity ''endosperm and embryo''. In order to obtain a reliable RNA quantification, two pools of 10 ''embryo and endosperm'' from small kernels and two pools of 10 ''embryo and endosperm'' from big kernels were obtained. Only the pools of small ''embryo

and endosperm'' expressed the *Engrailed-ZmOCL1* transgene (Figure 6B).

To elucidate the absence of a kernel phenotype in crosses where the transgenic plant had been used as male parent, RNA was isolated from 18 random 15 DAP kernels of two A188 × T1 crosses. In none of them expression of the *Engrailed-ZmOCL1* transgene was detected by semi-quantitative RT-PCR, while the *18S* rRNA gene control was equally expressed in all samples (9 lanes shown in Figure 6A). The most likely explanation of these data was a systematic paternal silencing of the transgene.

Expression of marker genes

To further characterise the transgenic plants, marker gene expression was studied by semi-quantitative RT-PCR in the kernel samples of reciprocal crosses as well as in ovaries of transgenic and control plants. No influence of the transgene was observed on the expression of the endogenous *ZmOCL1* gene or the L1 layer-specific *Ltp2* gene (Figure 6). Similarly expression of the *Pin1* gene involved in polar auxin transport was not significantly altered between transgenic and non-transgenic kernel or ovary samples (Figure 6). Higher expression in whole kernels (Figure 6A) as compared to dissected endosperm and embryo without the maternal pericarp (Figure 6B) was likely indicative of substantial *Pin1* expression in the pericarp. Finally a decrease in the expression of *GA20* needed for the biosynthesis of gibberellins was observed in ''endosperm and embryo'' pools of small kernels of the T1 × A188 crosses (Figure 6).

Target genes of ZmOCL1

In an attempt to identify potential direct or indirect target genes of the transcription factor ZmOCL1 we hybridised a genome wide cDNA micro-array with RNA isolated from ovaries of transgenic *Engrailed-ZmOCL1* plants and non-transformed siblings (unpublished preliminary data). Semi-quantitative RT-PCR experiments on ovaries of three transgenic plants and three control plants confirmed the up-regulation of CF006057 and CF007764 (Figure 7). The sequences of these two EST clones did not show significant homology to any known protein sequences or domains. They were likely indirect target genes of ZmOCL1 because direct target genes were expected to be repressed rather than activated by the chimeric Engrailed-OCL1 protein.

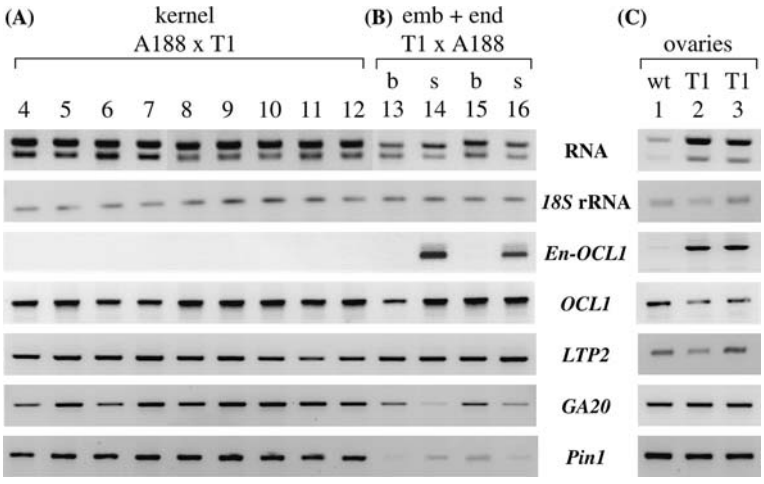

Figure 6. Marker gene expression. The expression of the *Engrailed-ZmOCL1* transgene (*En-OCL1*), the endogenous *ZmOCL1* gene (*OCL1*) and several marker genes was evaluated by semi-quantitative RT-PCR with gene-specific primers. The constitutively expressed *18S* rRNA gene was used as an internal control of RNA quantity. (A) Expression in individual 15 DAP kernels of a A188 × T1 cross. (B) Expression in pools of 10 endosperms containing the embryo obtained by removal of the maternal pericarp tissue from 15 DAP big (b) or small (s) kernels of T1 × A188 crosses. (C) Expression in pools of 10 immature ovaries from wild-type (wt) or two different T1 plants (T1).

Discussion

Here we present evidence that ZmOCL1, a member of the plant-specific HD-ZIP IV class of transcription factors, has a crucial function during early kernel development in maize. Transgenic plants expressing a fusion of the repressor domain of the *Drosophila* Engrailed ptotin with the DNA binding and dimerisation domains of ZmOCL1 exhibit a transient reduction of kernel size. The transient nature of this trait suggests functional redundancy with other genes, possibly other *ZmOCL* genes.

ZmOCL1 *is required for normal kernel development*

An implication of *ZmOCL1* in protoderm differentiation of maize embryos had been suggested by its expression pattern, which is restricted to the outer-most cell layer of the embryo proper and precedes the characteristic cytological changes observed in this cell layer at the transition stage (Ingram *et al.*, 1999). To substantiate this hypothesis and to determine the precise function of *ZmOCL1* in embryo development we took two complementary approaches, namely the search for insertional mutants and maize transformation with a construct impairing *ZmOCL1* function. With the first approach we were able to isolate a maize line homozygous for a *mutator* insertion 261 bp upstream of the start codon but did not detect any morphological changes in the epidermis of the embryo or any other notable phenotype. A similar lack of phenotypic changes has been reported for *Arabidopsis* plants carrying insertions in the *AtML1* or *Pdf2* genes that belong to the same HD-ZIP IV family as *ZmOCL1* (Abe *et al.*, 2003). Severe phenotypic changes of the double mutant demonstrated that functional redundancy between *AtML1* and *Pdf2* is the cause for the absence of phenotypic changes in single mutants. While we cannot exclude functional redundancy of *ZmOCL1* with another gene, the presence of normal *ZmOCL1* transcript levels indicates that the insertion may rather not interfere at all with *ZmOCL1* transcription or function, possibly because it is located in a yet unknown leading intron and spliced without negative effects on promoter activity. Alternatively, the *ZmOCL1* transcript could be due to transcription initiating from an outward facing promoter within the long terminal inverted repeat of *mutator* as shown for example for *Mu1* in the *hcf106* mutant (Barkan and Martienssen, 1991).

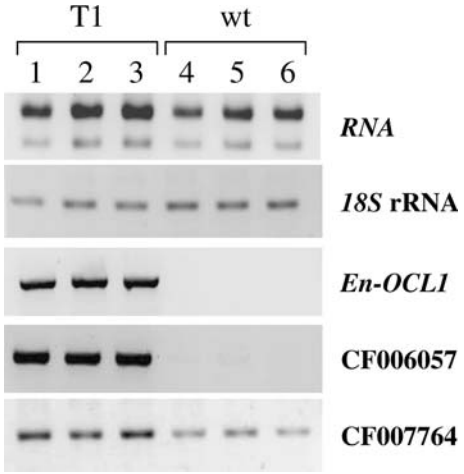

Figure 7. Putative target gene expression. The expression of the candidate target genes with the accession nos. CF006057 and CF007764 in immature ovaries of three transgenic T1 plants (lanes 1 to 3) or wildtype siblings (lanes 4 to 6) was evaluated by semi-quantitative RT-PCR. The *Engrailed-ZmOCL1* transgene (*En-OCL1*) and the constitutively expressed *18S* rRNA gene were used as controls.

With the second approach involving transgenic plants we observed a transient reduction in kernel size due to a developmental delay that was most obvious around 15 DAP. Several lines of evidence support the idea that the phenotype was caused by the *ZmOCL1* transgene. Firstly, several independent transformation events showed an effect on kernel size ruling out a role of the region disrupted by the insertion of the transgene. Secondly, an inverse quantitative relationship exists between the expression level of the transgene and the extent of size reduction. Finally, an effect of the sole *Engrailed* domain on maize kernel development can be ruled out because transgenic maize plants harbouring fusions of the Engrailed domain to the DNA binding domain of HsfB (Gagliardi *et al.*, 1995) under the control of the actin promoter did not show any comparable phenotypes (data not shown). Concerning the exact mode of action of the transgene several alternative explanations exist. The most attractive one is a repression rather than activation of ZmOCL1 target genes mimicking a loss-of-function phenotype. Two lines of evidence suggests that wildtype ZmOCL1 acts

indeed as an activator of transcription. Firstly in loss of function mutants of the related HD-ZIP IV genes *AtML1* and *PDF2* the transcription of the direct target gene *PDF1* is abolished (Abe *et al*., 2003). Secondly, fusions of ZmOCL1 without its homeo domain to the DNA binding domain of GAL4 activate UAS driven reporter genes in yeast (our unpublished results). Alternatively, a loss-of-function phenotype may also be obtained by squelching, i.e. a depletion of the pool of proteins normally interacting with ZmOCL1 (Markel *et al*., 2002). The chimeric protein contains the leucine zipper of ZmOCL1, which is likely involved in protein/protein interactions, and the highly expressed chimeric protein could either interfere with the homo- or hetero-dimerisation of the native ZmOCL1 protein or compete with it for the binding of interacting proteins. Theoretically it is also possible that the chimeric protein interacts with proteins that are not naturally interacting with ZmOCL1. Gene silencing or over-expression did not contribute substantially to the phenotypes observed in four case studies of translational *Engrailed* fusions (Markel *et al*., 2002). In the present case we can rule out gene silencing because the expression of the endogenous ZmOCL1 gene was not diminished in transgenic ovaries or endosperm with embryo (Figure 6). On the other hand we cannot rule out that the observed phenotype is caused by the ectopic expression of the ZmOCL1 domains present in the chimeric protein. In summary our data show that an interference with ZmOCL1 function leads to a transient delay in maize kernel development.

Epidermal cell fate and kernel development

In *Arabidopsis* the knock-outs of *AtML1* or *Pdf2* cause no visible phenotype (Abe *et al*., 2003) and knock-outs in other HD-ZIP IV genes have rather subtle phenotypes that concern anthocyanin accumulation in *anl2* (Kubo *et al*., 1999) or root hair and trichome formation in *gl2* (Masucci *et al*., 1996). Only the *atml1*/*pdf2* double mutant is seedling lethal and prohibits the formation of a morphologically normal epidermis as expected for a key gene in epidermis differentiation. Transgenic *Engrailed-ZmOCL1* plants have normal leaf surfaces, normal plant height and their endosperm and embryo show normal expression of the L1 marker gene *Ltp2* at 15 DAP (Figure 6) and normal cytology of epidermal cells at 30 DAP (data not shown). No information is available on anthocyanin accumulation because maize transformation was carried out in a colourless genetic background (Armstrong, 1994). The transient reduction in embryo, endosperm and kernel size is most severe in the embryo, possibly indicating that it is a hiatus or slow-down of the embryo development that delays the development of the entire caryopsis. One of several possible explanations of this delay would be a strong but not absolute need for ZmOCL1 during protoderm differentiation and a less important role during the maintenance of epidermal cell fate. Functional redundancy with one or several genes that have a delayed expression pattern would also account for our observations. According to the hypothesis of Bommert and Werr (2001) the absence of a differentiated outer cell layer would compromise meristem formation and interrupt embryo development. Independent recent findings support the idea that the functioning of meristems is dependent on an outer cell layer to which polar auxin transport is limited (Reinhardt *et al*., 2003). Finally the arrest and re-start of embryo development has been observed in apomictic plants of the closely related species *Tripsacum* and a cross-talk between embryo and endosperm development has been suggested (Grimanelli *et al*., 2003).

Potential targets of ZmOCL1

To gain some insight into the mode of action of the *Engrailed-ZmOCL1* the expression of widely used marker genes was evaluated. The *Ltp2* gene (Sossountzov *et al*., 1991) is a classical epidermis marker in maize (Bommert and Werr, 2001) and in other species (Kalla *et al*., 1994; Vroemen *et al*., 1996) and therefore a potential target gene of ZmOCL1. GA20 oxidase is one of enzymes needed for the synthesis of gibberellins (GA) and it has been shown that repression of the corresponding gene is an important component in the action of KNOX type homeo domain proteins in the shoot apical meristem (Sakamoto *et al*., 2001). Pin1 is an auxin efflux carrier necessary for polar auxin transport (Galweiler *et al*., 1998) and the double mutants *pin1 stm* or *pin1 cuc* fail to establish bilateral symmetry in the embryo (Aida *et al*.,

2002). Out of the three marker genes only *GA20* expression showed a correlation with the presence or absence of the *Engrailed-ZmOCL1* transgene. Its expression was decreased in "endosperm and embryo" of small kernels indicating a decrease in GA synthesis and possibly GA steady state levels. Since reduced GA levels cause dwarfism on the whole plant level (Olszewski *et al.*, 2002) it is tempting to speculate that the small kernel phenotype is caused by a reduction of GA levels in the kernel.

In an more open minded approach to identify target genes of ZmOCL1 the transcriptome of ovaries of transgenic and non-transgenic plants was compared by the use of a genome wide microarray (unpublished preliminary data). In theory this approach detects expression changes in both genes directly regulated by ZmOCL1 *via* binding of the ZmOCL1 protein to their promoter region and genes indirectly regulated by ZmOCL1 *via* diverse pathways. Differential expression was confirmed for two of the genes identified but in neither case the sequence analysis of the EST contig revealed any significant homology to proteins or protein domains of known function.

Phylogenetic position of ZmOCL1

The genomic sequence allowed us to include *ZmOCL1* in phylogenetic studies based on the intron/exon structure and sequence similarities of HD-ZIP IV genes (Ingouff *et al.*, 2001). To have a more complete view of the family in cereals we also deduced the complete protein sequence of four additional rice genes from genomic sequences and determined the intron/exon structure of all 9 rice genes. Our enlarged data set indicates that contrary to previous conclusions the groups defined by common intron/exon structures or by sequence similarities do not always coincide. While the group around ZmOCL1 characterised by the absence of intron positions 3 and 8 (Figure 2A) falls with the exception of AtFWA into a single clade in a tree based on full length protein sequences (Figure 2B), the members of the other two major groups defined by the presence of intron position 8 and the presence or absence of intron position 3 are scattered and interspersed over several clades. A prime example are genes *AtML1* and *AtPDF2* that differ in intron position 3 and yet code for the highly similar and functionally redundant proteins (Abe *et al.*, 2003). We conclude from these data that not all changes in the intron/exon structure are ancestral events in evolution and that such changes occurred after the divergence of monocots and dicots and even after the separation of maize and rice. A striking argument in favour of this hypothesis is *OsRoc7* that lacks 6 of the 9 conserved intron positions, while none of the 10 HD-ZIP genes present in the completely sequenced *Arabidopsis* genome has less than 7 of the 9 conserved intron positions.

The construction of phylogenetic trees based on similarities of full length protein sequences was hampered by high sequence divergence in the stretches between conserved domains as well as by substantial insertions or deletions in individual protein sequences. On the other hand trees based on partial sequences corresponding to the highly conserved HD-ZIP or START domains were not discriminatory enough and resulted in low bootstrap values and minor differences between trees. Based on the full length sequences the following maize/rice orthologues were identified: ZmOCL1/OsRoc5, ZmOCL2/OsRoc4, ZmOCL3/OsRoc6 and ZmOCL4/OsRoc3. ZmOCL5 was close to both OsRoc1 and OsRoc2 providing further evidence for the existence of a closely related *ZmOCL5-like* gene in the maize genome that had already been suspected earlier (Ingram *et al.*, 2000). Discarding OsRoc7, which likely is the result a recent duplication event, the rice proteins OsRoc8 and OsRoc9 remained without counterpart in maize suggesting the existence of additional *ZmOCL* genes in the maize genome.

The orthologue of ZmOCL1 in *Arabidopsis* could not be unambiguously identified. The clade of ZmOCL1 in the phylogenetic tree contains three maize genes, three rice genes but only two *Arabidopsis* genes. The sequence diversity of the three maize genes, the available mapping information and the clearly distinct expression patterns make it unlikely that any of the three maize genes are the result of the well documented recent duplication events in the maize genome (Helentjaris *et al.*, 1988; Gaut, 2001). It follows that the function of HD-ZIP IV genes has certainly evolved since the divergence of monocots and dicots and that specific functional analyses in maize and rice are necessary to fully understand the function of this gene family.

Acknowledgements

We thank Wolfgang Werr for the Engrailed vector, Denise Garcia and co-workers for maize transformation as well as Françoise Le Boulaire and co-workers for the micro-array analysis. Pascal Condamine and Véronique Boltz provided excellent technical assistance in the lab. Alexis Lacroix, Armand Guillermin, Claudia Bardoux and Hervé Leyral are acknowledged for their implication in plant culture and media preparation. This work was financed in part by the programs NO-1999-013 and ZmS3P1 of the French genomics project Génoplante. ASK was supported by a PhD fellowship of the Egyptian Government and GI by a postdoctoral bourse from INRA.

References

Abe, M., Katsumata, H., Komeda, Y. and Takahashi, T. 2003. Regulation of shoot epidermal cell differentiation by a pair of homeodomain proteins in Arabidopsis. Development 130: 635–643.

Aida, M., Vernoux, T., Furutani, M., Traas, J. and Tasaka, M. 2002. Roles of PIN-FORMED1 and MONOPTEROS in pattern formation of the apical region of the Arabidopsis embryo. Development 129: 3965–3974.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25: 3389–3402.

Armstrong, C.L. 1994. Regeneration of plants from somatic cell cultures: applications for *in vitro* genetic manipulation. In: M. Freeling and V. Walbot (Eds.), The Maize Handbook, Springer, New York, pp. 663–671.

Barkan, A. and Martienssen, R.A. 1991. Inactivation of maize transposon Mu suppresses a mutant phenotype by activating an outward-reading promoter near the end of Mu1. Proc. Natl. Acad. Sci. USA 88: 3502–3506.

Barloy, D., Denis, L. and Beckert, M. 1989. Comparison of the aptitude for anther culture in some androgenetic doubled haploid maize lines. Maydica 34: 303–308.

Bommert, P. and Werr, W. 2001. Gene expression patterns in the maize caryopsis: clues to decisions in embryo and endosperm development. Gene 271: 131–142.

Bowman, J.L., Eshed, Y. and Baum, S.F. 2002. Establishment of polarity in angiosperm lateral organs. Trends Genet. 18: 134–141.

Cao, J., Duan, X.L., McElroy, D. and Wu, R. 1992. Regeneration of herbicide resistant transgenic rice plants following microprojectile-mediated transformation of suspension culture cells. Plant Cell Rep. 11: 586–591.

Chan, R.L., Gago, G.M., Palena, C.M. and Gonzalez, D.H. 1998. Homeoboxes in plant development. Biochim. Biophys. Acta 1442: 1–19.

Clark, J.K. and Sheridan, W.F. 1991. Isolation and characterisation of 51 *embryo-specific* mutations of maize. Plant Cell 3: 935–951.

Di Cristina, M., Sessa, G., Dolan, L., Linstead, P., Baima, S., Ruberti, I. and Morelli, G. 1996. The Arabidopsis Athb-10 (GLABRA2) is an HD-Zip protein required for regulation of root hair development. Plant J. 10: 393–402.

Emery, J.F., Floyd, S.K., Alvarez, J., Eshed, Y., Hawker, N.P., Izhaki, A., Baum, S.F. and Bowman, J.L. 2003. Radial patterning of Arabidopsis shoots by class III HD-ZIP and KANADI genes. Curr. Biol. 13: 1768–1774.

Finer, J.J., Vain, P., Jones, M.W. and Mcmullen, M.D. 1992. Development of the Particle Inflow Gun for DNA Delivery to Plant Cells. Plant Cell Rep. 11: 323–328.

Gagliardi, D., Breton, C., Chaboud, A., Vergne, P. and Dumas, C. 1995. Expression of heat shock factor and heat shock protein 70 genes during maize pollen development. Plant Mol. Biol. 29: 841–856.

Galtier, N., Gouy, M. and Gautier, C. 1996. SeaView and PhyloWin, tow graphic tools for sequence alignment and molecular phylogeny. Comput. Appl. Biosci. 12: 543–548.

Galweiler, L., Guan, C., Muller, A., Wisman, E., Mendgen, K., Yephremov, A. and Palme, K. 1998. Regulation of polar auxin transport by AtPIN1 in Arabidopsis vascular tissue. Science 282: 2226–2230.

Gaut, B.S. 2001. Patterns of chromosomal duplication in maize and their implications for comparative maps of the grasses. Genome Res. 11: 55–66.

Gehring, W.J., Affolter, M. and Burglin, T. 1994. Homeodomain proteins. Annu. Rev. Biochem. 63: 487–526.

Gerdes, J.T. and Tracy, W.F. 1993. Pedigree diversity within the Lancaster surecrop heterotic group of maize. Crop Sci. 33: 334–337.

Grimanelli, D., Garcia, M., Kaszas, E., Perotti, E. and Leblanc, O. 2003. Heterochronic expression of sexual reproductive programs during apomictic development in Tripsacum. Genetics 165: 1521–1531.

Heckel, T., Werner, K., Sheridan, W.F., Dumas, C. and Rogowsky, P.M. 1999. Novel phenotypes and developmental arrest in early *embryo specific* (*emb*) mutations of maize. Planta 210: 1–8.

Helentjaris, T., Weber, D. and Wright, S. 1988. Identification of the Genomic Locations of Duplicate Nucleotide-Sequences in Maize by Analysis of Restriction Fragment Length Polymorphisms. Genetics 118: 353–363.

Higgins, D.G. and Sharp, P.M. 1988. CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. Gene 73: 237–244.

Ingouff, M., Farbos, I., Lagercrantz, U. and von Arnold, S. 2001. PaHB1 is an evolutionary conserved HD-GL2 homeobox gene expressed in the protoderm during Norway spruce embryo development. Genesis 30: 220–230.

Ingram, G.C., Boisnard-Lorig, C., Deguerry, F., Dumas, C. and Rogowsky, P.M. 2000. Expression patterns of genes encoding HD-ZipIV homeodomain proteins define specific domains in the maize embryos and meristems. Plant J. 22: 401–414.

Ingram, G.C., Magnard, J.L., Vergne, P., Dumas, C. and Rogowsky, P.M. 1999. ZmOCL1, an HDGL2 family homeobox gene, is expressed in the outer cell layer throughout maize development. Plant Mol. Biol. 40: 343–354.

Ito, M., Sentoku, N., Nishimura, A., Hong, S.K., Sato, Y. and Matsuoka, M. 2002. Position dependent expression of GL2-type homeobox gene, Roc1: significance for protoderm

differentiation and radial pattern formation in early rice embryogenesis. Plant J. 29: 497–507.
John, A., Smith, S.T. and Jaynes, J.B. 1995. Inserting the Ftz homeodomain into engrailed creates a dominant transcriptional repressor that specifically turns off Ftz target genes in vivo. Development 121: 1801–1813.
Kalla, R., Shimamoto, K., Potter, R., Nielsen, P.S., Linnestad, C. and Olsen, O.A. 1994. The promoter of the barley aleurone-specific gene encoding a putative 7 kDa lipid transfer protein confers aleurone cell-specific expression in transgenic rice. Plant J. 6: 849–860.
Kaplan, D.R. and Cooke, T.J. 1997. Fundamental Concepts in the Embryogenesis of Dicotyledons: A Morphological Interpretation of Embryo Mutants. Plant Cell 9: 1903–1919.
Klein, T.M., Arentzen, R., Lewis, P.A. and Fitzpatrick-McElligott, S. 1992. Transformation of microbes, plants and animals by particle bombardment. Biotechnology 10: 286–291.
Kubo, H., Peeters, A.J., Aarts, M.G., Pereira, A. and Koornneef, M 1999. ANTHOCYANINLESS2, a homeobox gene affecting anthocyanin distribution and root development in Arabidopsis. Plant Cell 11: 1217–1226.
Lu, P., Porat, R., Nadeau, J.A. and O'Neill, S.D. 1996. Identification of a meristem L1 layer-specific gene in Arabidopsis that is expressed during embryonic pattern formation and defines a new class of homeobox genes. Plant Cell 8: 2155–2168.
Magnard, J.L., Heckel, T., Massonneau, A., Wisniewski, J.P., Cordelier, S., Lassagne, H., Perez, P., Dumas, C. and Rogowsky, P.M. 2004. Morphogenesis of maize embryos requires ZmPRPL35–1 encoding a plastid ribosomal protein. Plant Physiol. 134: 649–663.
Markel, H., Chandler, J. and Werr, W. 2002. Translational fusions with the engrailed repressor domain efficiently convert plant transcription factors into dominant-negative functions. Nucleic Acids Res. 30: 4709–4719.
Martienssen, R.A., Rabinowicz, P.D., O'Shaughnessy, A. and McCombie, W.R. 2004. Sequencing the maize genome. Curr. Opin. Plant Biol. 7: 102–107.
Masucci, J.D., Rerie, W.G., Foreman, D.R., Zhang, M., Galway, M.E., Marks, M.D. and Schiefelbein, J.W. 1996. The homeobox gene GLABRA2 is required for position-dependent cell differentiation in the root epidermis of Arabidopsis thaliana. Development 122: 1253–1260.
McConnell, J.R., Emery, J., Eshed, Y., Bao, N., Bowman, J. and Barton, M.K. 2001. Role of PHABULOSA and PHAVOLUTA in determining radial patterning in shoots. Nature 411: 709–713.
Olszewski, N., Sun, T.P. and Gubler, F. 2002. Gibberellin signaling: biosynthesis, catabolism, and response pathways. Plant Cell 14: S61–S80.
Pfaffl, M.W. 2001. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res. 29: e45.
Ponting, C.P. and Aravind, L. 1999. START: a lipid binding domain in StAR, HD-ZIP and signalling proteins. Trends Biochem. Sci. 24: 130–132.
Reinhardt, D., Pesce, E.R., Stieger, P., Mandel, T., Baltensperger, K., Bennett, M., Traas, J., Friml, J. and Kuhlemeier, C. 2003. Regulation of phyllotaxis by polar auxin transport. Nature 426: 255–260.
Rerie, W.G., Feldmann, K.A. and Marks, M.D. 1994. The GLABRA2 gene encodes a homeo domain protein required for normal trichome development in Arabidopsis. Genes Dev. 8: 1388–1399.
Riechmann, J.L., Heard, J., Martin, G., Reuber, L., Jiang, C., Keddie, J., Adam, L., Pineda, O., Ratcliffe, O.J., Samaha, R.R., Creelman, R., Pilgrim, M., Broun, P., Zhang, J.Z., Ghandehari, D., Sherman, B.K. and Yu, G. 2000. Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes. Science 290: 2105–2110.
Ruberti, I., Sessa, G., Lucchetti, S. and Morelli, G. 1991. A novel class of plant proteins containing a homeodomain with a closely linked leucine zipper motif. Embo J 10: 1787–1791.
Sakamoto, T., Kamiya, N., Ueguchi-Tanaka, M., Iwahori, S. and Matsuoka, M. 2001. KNOX homeodomain protein directly suppresses the expression of a gibberellin biosynthetic gene in the tobacco shoot apical meristem. Genes Dev. 15: 581–590.
Sambrook, J., Fritsch, E.F. and Maniatis, T. 1989. Molecular cloning: a laboratory manual 2nd edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor NY.
Schrick, K., Nguyen, D., Karlowski, W.M. and Mayer, K.F. 2004. START lipid/sterol-binding domains are amplified in plants and are predominantly associated with homeodomain transcription factors. Genome Biol. 5: R41.
Sessa, G., Morelli, G. and Ruberti, I. 1993. The Athb-1 and -2 HD-Zip domains homodimerize forming complexes of different DNA binding specificities. Embo J. 12: 3507–3517.
Sossountzov, L., Ruiz-Avila, L., Vignols, F., Jolliot, A., Arondel, V., Tchang, F., Grosbois, M., Guerbette, F., Miginiac, E., Delseny, M., Puigdomenech, P. and Kader, J.-C. 1991. Spatial and temporal expression of a maize lipid transfer protein gene. Plant Cell 3: 923–933.
Talbert, P.B., Adler, H.T., Parks, D.W. and Comai, L 1995. The REVOLUTA gene is necessary for apical meristem development and for limiting cell divisions in the leaves and stems of Arabidopsis thaliana. Development 121: 2723–2735.
Tavares, R., Aubourg, S., Lecharny, A. and Kreis, M 2000. Organization and structural evolution of four multigene families in Arabidopsis thaliana: AtLCAD, AtLGT, AtMYST and AtHD-GL2. Plant Mol. Biol. 42: 703–717.
Vain, P., Keen, N., Murillo, J., Rathus, C., Nemes, C. and Finer, J.J 1993. Development of the particle inflow gun. Plant Cell Tissue Organ Cult. 33: 237–246.
Vroemen, C.W., Langeveld, S., Mayer, U., Ripper, G., Jurgens, G., Van Kammen, A. and De Vries, S.C 1996. Pattern Formation in the Arabidopsis Embryo Revealed by Position-Specific Lipid Transfer Protein Gene Expression. Plant Cell 8: 783–791.