Development 127, 2219-2226 (2000)
Printed in Great Britain © The Company of Biologists Limited 2000
DEV0287

TSO1 is a novel protein that modulates cytokinesis and cell expansion in *Arabidopsis*

Bernard A. Hauser[1,§], Jeannie Q. He[1,‡], Sung O. Park[2] and Charles S. Gasser[1,*]

[1]Section of Molecular and Cellular Biology, University of California, Davis, CA 95616, USA
[2]Department of Botany, University of Florida, Gainesville, FL 32611, USA
[§]Present address: Department of Botany, University of Florida, Gainesville, FL 32611, USA
[‡]Present address: Department of Microbiology and Molecular Genetics, Jonsson Comprehensive Cancer Center, University of California, Los Angeles CA 90095, USA
*Author for correspondence (e-mail: csgasser@ucdavis.edu)

Accepted 25 February; published on WWW 18 April 2000

SUMMARY

Previous analyses of *tso1* mutants revealed a loss of control of directional cellular expansion and coordination of growth of adjacent cells, and defects in karyokinesis and cytokinesis. We isolated *TSO1* using a map-based approach, and show that it is a member of a family of at least three genes in *Arabidopsis*. Consistent with the mutant phenotype, *TSO1* transcript was most abundant in flowers, where it accumulated to the highest levels in developing ovules and microspores. The putative TSO1 protein has two cysteine-rich regions that are similar to the CXC domains of a variety of proteins from plants and animals, including a class of kinesins involved in chromosome segregation, and enhancer of zeste-type proteins. Visualization of TSO1-fusion proteins indicated that TSO1 is a nuclear protein. The *tso1* mutant phenotypes and the novelty of the TSO1 sequence suggest the existence of previously unknown participants in regulation of directional processes in eukaryotic cells.

Key words: *Arabidopsis thaliana*, TSO1, Cytokinesis, Flower morphogenesis, Ovule

INTRODUCTION

Plant morphogenesis requires tight coordination among a range of processes. Because plant cells do not migrate, morphogenesis is ultimately dependent on appropriate directional cell expansion and division. Expansion and division must be integrated with positional information such that new structures form at proper locations and in the appropriate orientation. In addition, growth must often be coordinated among large numbers of cells that are not directly in contact with each other to produce bilaterally symmetrical organs such as leaves, sepals, petals, and ovules.

While histology of plant growth and development has been well characterized, less is known about how these processes are regulated at the molecular level. Some of the same components that regulate mitosis in other organisms have similar functions in plants (for review see Doonan and Fobert, 1997). However, cytokinesis occurs via a different mechanism in plants than it does in animals. In a dividing animal cell, constriction of a ring of actin filaments results in the formation of a cleavage furrow and eventual division into two daughter cells. In plants, however, division occurs by deposition of a new cross wall or cell plate separating the two daughter cells (for review see Staehelin and Hepler, 1996). These differences imply the existence of novel proteins involved in cytokinesis in plants.

A number of the proteins involved in plant cytokinesis and directional cell expansion have been identified through mutant screens for plants with altered morphology (Assaad et al., 1996; Liu et al., 1995; Lukowitz et al., 1996). Most such mutants are embryo or seedling lethal because cell division is necessary from the very earliest stages of plant development. The *tso1* mutants, however, exhibit normal vegetative development, but defects in cell division and expansion occur during flower and ovule development (Hauser et al., 1998; Liu et al., 1997). Cells of *tso1-1* flowers showed abnormally high nuclear DNA levels, aberrant spindle apparatuses and incomplete formation of the cell plate (Liu et al., 1997). Gross flower morphology in the weaker *tso1-3* allele was nearly indistinguishable from wild type. In *tso1-3* ovules, however, the embryo sac failed to form and defects in integument development resulted from aberrations in directional expansion of cells and/or coordination of growth among adjacent cells (Hauser et al., 1998). Defects in the assembly, structure, or function of cytoskeletal elements, or processes they direct, might account for all of these phenotypes (Hauser et al., 1998; Liu et al., 1997).

Herein we describe the cloning of *TSO1*. Based on the nature of changes observed in *tso1* mutants, we propose explanations for the differences in phenotypes among the *tso1* alleles. We also describe the relation of the TSO1 protein to other proteins in plants and animals, and discuss possible roles for this protein in cell growth and division.

MATERIALS AND METHODS

Mapping *tso1*

The *tso1-3* mutation is a recessive mutation that was isolated in the Landsberg *erecta* (L*er*) background. To map *tso1*, pollen from *tso1-3* plants was used to pollinate flowers of the Columbia-3 (Co-3) ecotype. F_2 progeny from this cross were analyzed for linkage of *tso1* with genetic markers. Map distances were determined using Mapmaker v. 2.0 (Lander et al., 1987).

To facilitate mapping of *tso1*, we designed a cleaved amplified polymorphic sequence (CAPS) marker (Konieczny and Ausubel, 1993) from the existing m560 RFLP marker (Hauge et al., 1993). Primers that flanked the polymorphic *Eco*RI site (AGTCTCTCC-AGGTCTGTCTTG and GGTATTCCTCCCGTCAAATTG) were used to amplify this DNA region; providing a new CAPS marker.

Cosmid library construction and plant transformation

Bacterial artificial chromosome (BAC) F18G19 (Mozo et al., 1998) was partially digested with *Sau*3AI and fragments were fractionated on a 10-40% sucrose gradient. Fractions containing fragments of 20-25 kb were ligated into pOCA28 (Olszewski et al., 1988) that had been digested with *Bam*HI and treated with calf alkaline phosphatase (Boehringer Mannheim Corp., Indianapolis, IN). This DNA was packaged with GigaPack III and the resultant phage used to infect *E. coli* as recommended by the manufacturer (Stratagene, La Jolla, CA). Cosmid DNAs were mobilized into *Agrobacterium* strain ASE by triparental mating (Figurski and Helinski, 1979). Because *tso1-3* plants are nearly sterile, *tso1-3*/Co-3 heterozygotes were used as recipients in transformations using the method of Clough and Bent (1998).

Sequencing and sequence analysis

A plasmid library was prepared from cosmid cF45 DNA to facilitate sequencing. cF45 was partially digested with *Sau*3AI and fragments averaging 3 kb in length were cloned into the *Bam*HI site in pUC119 (Vieira and Messing, 1987). Randomly chosen plasmids were purified using anion exchange columns (Qiagen Inc., Valencia, CA) and were sequenced on an ABI 377 (PE Applied Biosystems, Foster City, CA). The sequences of mRNAs from mutant lines were determined by (1) amplifying putative genes by RT-PCR (Kawasaki, 1990), (2) DNA purification by PEG precipitation and (3) cycle sequencing of PCR products.

TSO1 constructs

The *TSO1* coding sequence was amplified by RT-PCR (Kawasaki, 1990), with the following primers: GGAATTCTGGAAAATGGACA-AATCCC and AGGATCCCCTGATTTGGGTTGAGAGAAGG. The PCR product was digested with *Bam*HI and *Eco*RI and ligated into pBluescript II KS+ to create pBH6 and the insert was sequenced to verify integrity. The TSO1 coding sequence from pBH6 was isolated as a *Bam*HI/*Eco*RI fragment and inserted into these same sites in pMON999 [a plasmid containing the CaMV 35S promoter with a duplication of the enhancer region (Kay et al., 1987) and the nopaline synthase poly(A) addition site flanking a multi-linker] to form pBH8.

A plasmid, pBH10, encoding 35S::TSO1-green fluorescent protein (GFP; Sheen et al., 1995) fusion protein, was made by ligating the *Bam*HI/*Eco*RI TSO1 fragment from pBH6 and the *Sac*I/*Bam*HI GFP fragment from pGFP1.1.5 (Schumacher et al., 1999) into the *Sac*I/*Eco*RI sites of pMON999. A 35S::GFP construct was made by ligating the *Bam*HI/*Sac*I GFP fragment from pGFP1.1.5 into the *Bgl*II/*Sac*I sites in pMON999, creating pBH9.

The stop codon of the GFP coding region in pGFP1.1.5 was replaced with an *Eco*RI site by PCR amplification using the following primers: AGAATTCTTTGTATAGTTCATCCATG and TGGATCCA-AGGAGATATAACAATG. A 35S::GFP-TSO1 translational fusion vector was made by digesting this PCR product with *Bam*HI and *Eco*RI and inserting this fragment into the *Bgl*II and *Eco*RI sites in pBH8, thereby creating pBH13.

An *Xba*I fragment of cF45 containing the entire *TSO1* gene was inserted into the *Spe*I site of pMLBART (Gleave, 1992) to form pBH12.

The 3′ portions of *SOL1* and *TSO1* cDNAs were amplified by PCR from first strand cDNA template using the following primers: TSO1 5R, CTAGAATGTACTATTACATTACC; TSO1 4F, ACATTTGG-TGGTTCATCAGCC; SOL1 4R, GTAGCAGTATTGCGTTAGTTTT; SOL1 4F, TCCTCAGGGATATACAACATC. The DNA products from these PCR reactions were ligated into the EcoRV site of pLitmus 28 (New England Biolabs, Bethseda, MD), creating pBH5 and pBH3 with 3′ fragments from TSO1 and SOL1, respectively.

Northern blots and in situ hybridizations

Total RNA and poly(A)$^+$ RNA were extracted from plant organs as previously described (Gasser et al., 1989; Jakobsen et al., 1990). Equivalent amounts of RNA were electrophoresed through a 1.2% agarose/7% formaldehyde gel, and blotted to Hybond N$^+$ membrane (Amersham, Boston, MA). Divergent regions of *SOL1*, *SOL2* and *TSO1* were used as probes for the hybridizations so transcripts from paralogous genes would not cross-hybridize. A portion of *SOL2* was amplified by PCR using the following nested primers: SOL1F, GAGTGACACTGTGTTTGTGG; SOL1R, TTTCCTTAGCTAC-ACAACACC; SOL2F, GGTAGTTGAACCAGTTCTG, SOL2R, TCAGTGTTGGGGAGTGAG.

The *SOL2* PCR product, pBH5 and pBH3 were radiolabeled (Feinberg and Vogelstein, 1983) and hybridized to RNA blots. RNA blots were either hybridized in 50% formamide, 50 mM phosphate buffer, pH 7.2, 100 μg/ml sonicated salmon sperm DNA, 2× SSPE (1× SSPE is 150 mM NaCl, 10 mM sodium phosphate, 1 mM EDTA) or QuikHyb with 100 μg/ml sonicated salmon sperm DNA (Stratagene, La Jolla, CA). Blots were washed twice in 2× SSPE, 1% SDS and twice in 0.1× SSPE, 0.1% SDS at 65°C. Blots were exposed to imaging screens and imaged on a Fuji (Tokyo, Japan) BAS1000 phosphorimager.

In situ hybridizations were done as described by Vielle-Calzada et al. (1999). Templates for sense and antisense*TSO1* digoxigenin (dig) probes were made by cleaving pBH5 with *Xba*I or *Eco*RI, respectively. Templates for sense and antisense *SOL1* probes were made by cleaving pBH3 with *Xba*I or *Eco*RI, respectively. Dig-labeled RNAs were transcribed using the Genius Dig RNA Labeling kit and T7 RNA polymerase (Boehringer Mannheim Corp., Indianapolis, IN).

Particle bombardment

Gold particles (1 μm) were coated with plasmid DNA as suggested by the manufacturer (BioRad, Hercules, CA). DNA-coated particles were used to bombard onion epidermal peels using the Biolistic PDS-1000/He System (BioRad) and 1100 psi rupture discs essentially as described by Varagona et al. (1992). Epidermal peels were screened for GFP fluorescence 16-20 hours after bombardment using an Axioskop microscope (Zeiss Inc., Thornwood, NY). GFP fluorescence was examined using 480 nm light excitation and a 510 nm long-pass filter for observation. Phase contrast and fluorescent images were digitized then adjusted and assembled using PhotoShop 4 (Adobe Inc., San Jose, CA).

RESULTS

tso1 mutants

The floral and ovule phenotypes of representative wild-type and *tso1* plants are shown in Fig. 1. The strong *tso1-1* allele produces sepals that are smaller than wild type and that terminate in jagged projections (Fig. 1B; Liu et al., 1997). Petal, stamen and carpel primordia exhibited highly aberrant morphology and were less than one-third the size of wild-type organs. The weaker *tso1-3* mutant had nearly normal flowers,

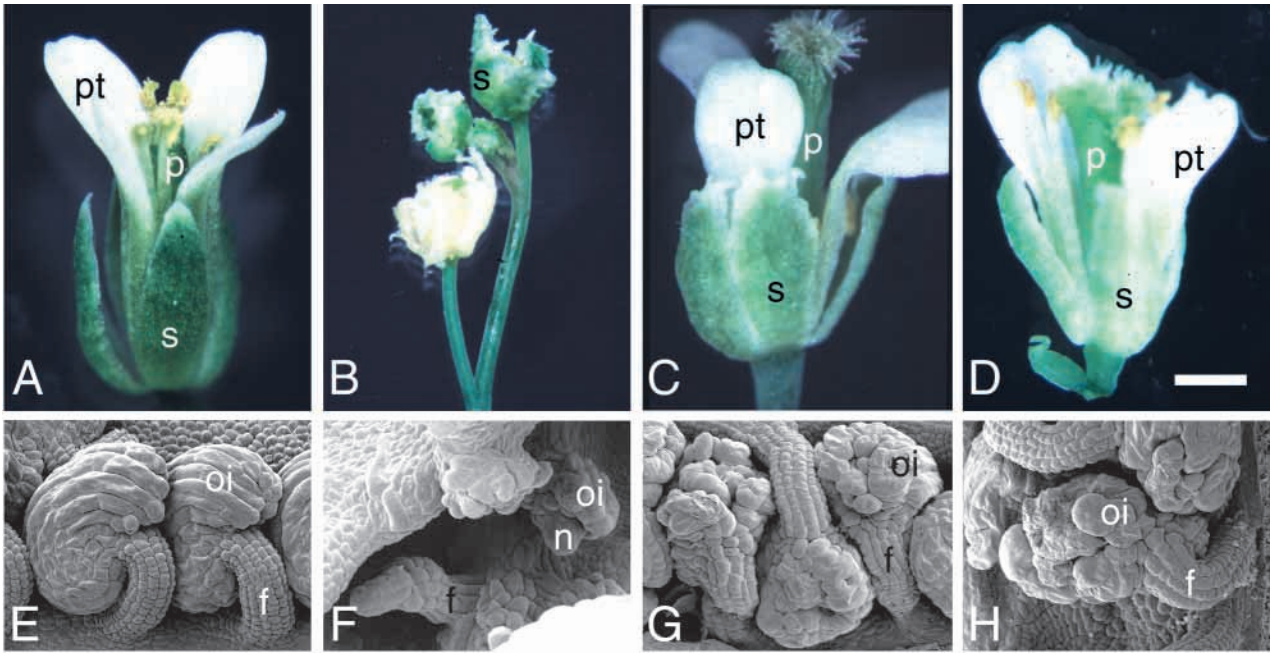

Fig. 1. Wild-type and *tso1* floral structures. (A) A wild-type flower. p, pistil; pt, petal; s, sepal. (B) *tso1-1* flowers are reduced in size and have misshapen floral organs. (C) *tso1-3* flowers have sepals with ragged fringes, but other floral organs appear normal. A stamen was removed to expose the pistil. (D) Flower of a homozygous *tso1-1* plant containing two copies of the cF45 transgene. Flower is similar to wild type, except for a slight bifurcation of the stigma. (E) Wild-type ovules at anthesis. f, funiculus; oi, outer integument. (F) In *tso1-1* ovules, the integuments consist of a small number of irregularly-shaped cells. n, nucellus. (G) *tso1-3* ovules at anthesis. Integuments consist of irregularly shaped cells that are not organized into files. (H) Ovules of *tso1-1* plants containing the cF45 transgene. Ovules appear similar to those of *tso1-3*. Bar corresponds to 2 mm in A-D, 50 μm in E, G and H, and 25 μm in F.

except for small defects in the edges of the sepals (Fig. 1C; Hauser et al., 1998).

tso1-3 ovules superficially resembled those of wild type, but had integuments consisting of misshapen cells which did not show the regular organization observed in wild type (Fig. 1G; Hauser et al., 1998). Examination of the rarely formed ovules in a *tso1-1* flower showed a more extreme manifestation of this phenotype where integuments were reduced to rudimentary structures of aberrantly shaped cells (Fig. 1F). The common features among the *tso1* alleles were anomalies in floral cell growth and division, which led to aberrant organ morphology.

Mapping *tso1*

TSO1 was previously mapped to the middle of chromosome 3 (Hauser et al., 1998; Liu et al., 1997). Further mapping revealed that *tso1* was flanked by markers m560 (Hauge et al., 1993) and DMC (Klimyuk and Jones, 1997) (Fig. 2A). Sixty-two plants with recombinations between m560 and DMC were identified from 962 plants in a mapping population, and were used to refine the position of *tso1* using additional markers (Fig. 2A). Hybridizations were then used to identify a bacterial artificial chromosome, F18G19, containing markers flanking *TSO1* on both sides (Fig. 2A).

Complementation of *tso1*

A cosmid library was constructed from F18G19 DNA and an overlapping set of cosmid DNAs was assembled through fingerprinting and hybridizations (Fig. 2B). Because homozygous mutants exhibit reduced fertility, individual cosmids were introduced into *tso1-3* heterozygotes from the *tso1-3* L*er*/Col-3 mapping population. The genotype at the *TSO1* locus could then be evaluated in transgenic progeny using the flanking PAP140 and DMC CAPS markers, which differentiate between these two ecotypes.

A single cosmid, cF45, was identified that complemented *tso1-3* (Fig. 2B). The same cosmid only partially rescued *tso1-1* mutants, producing plants with flowers similar to wild type

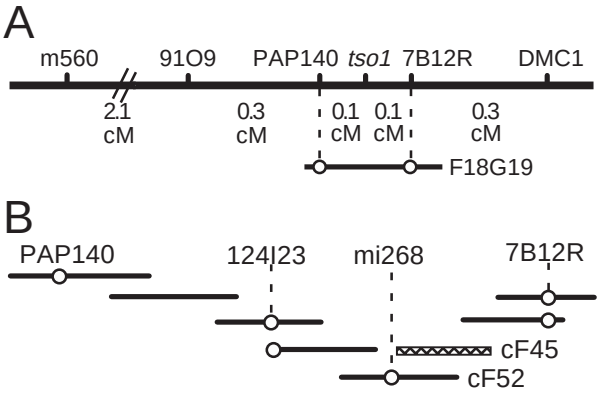

Fig. 2. Mapping of *tso1*. (A) Genetic distances between the *tso1* locus and nearby mapping markers (above line) are displayed in centiMorgans (cM). One BAC, F18G19, hybridized with the 7B12R and PAP140 markers, which flank *tso1*. (B) Cosmid contig spanning BAC F18G19 DNA. Markers used in mapping *TSO1* are shown at the top, 124I23, and mi268 cosegregated with *tso1*. Each of these cosmids was individually transformed into *tso1* plants and only cosmid cF45 (hatched) complemented this mutation.

(Fig. 1D), but with defective ovules (Fig. 1H), which were similar to those of *tso1-3* (Fig. 1G). Complementation was verified through segregation analysis of progeny plants, where the presence of the T-DNA always correlated with partial or complete rescue of *tso1-1* or *tso1-3* mutants, respectively (data not shown).

Identifying *TSO1* in the complementing cosmid

Randomly generated subclones from cF45 were sequenced and the sequences were assembled into DNA contigs. Analysis of the sequence identified four putative genes in cF45, and the subsequent availability of the complete sequence for this region (MWI23, accession no. AB022223) confirmed this analysis. The cDNA for each gene was amplified from wild-type and mutant lines by RT-PCR, and was sequenced to identify potential mutations. In one of the putative genes we identified a nonsense mutation in the *tso1-3* and *tso1-4* alleles, and missense mutation in the *tso1-1* allele (Fig. 3A). No mutations were found in the other three genes on this cosmid.

To demonstrate that we had identified *TSO1*, a plasmid (pBH12) containing the wild-type copy of the gene was transformed into *tso1*/Co-3 plants. In the resulting transgenic plants, the transgene completely complemented all of the homozygous *tso1-3* mutants and partially complemented all of the homozygous *tso1-1* mutants (data not shown). The partially complemented *tso1-1* flowers exhibited a phenotype indistinguishable from that shown in Fig. 1D,H for complementation with cF45 (data not shown).

A second gene (*SOL1*) contained within cF45 was significantly similar in sequence to *TSO1* (see below). A cosmid (cF52) overlapping with cF45 (Fig. 2B) contains all of *SOL1*, but does not include *TSO1*. Introduction of this cosmid into *tso1* plants produced no discernable alteration in the mutant phenotype (data not shown); demonstrating that extra copies of this gene could not substitute for the defective *TSO1* gene.

A

```
SOL2 MDTPQKSIT.QIGTPISKSRFEDSPVFNYINSLSPIRPVRSIPNPNQFSSLNFTSPPSVF
TSO1 MDKSQKNPTSQIGTSTPKSKFEDSPVFNYISNLSPIESVKSISTAQTFSSLSFTSPPPVF
SOL1 MDTPEKSET.QIGTPVSKLKVEDSPVFSYICNLSPIKTIKPIPITCPLSSLNYASPPSVF
                                                        100
SOL2 TSPHLTSSHKESRFFKTHNSSSSDPTNSVESQEDESTSHEEVPAEGED.TKGLNIDDCMR
TSO1 TSPHVI.SHRESRFFRCHN..SVDRSKHLESLDGSAVKGEVVVPLVEDLNKEASLED..E
SOL1 TSPHAV.SHKESRF..............RSQKDVSASKEV.................GE

SOL2 EEASVETNLD.......DSVA.....SPCGGNTTDLSLVPYA.PTRGEDG.SCEDNGMEL
TSO1 EETSVETSSELPQILKFDSQTSEHSDSPC...TEDVVIEASSDPPRGDNGSSSEDVTMGL
SOL1 EEALV..GSEPEQSYKNDCNTPRVLN........DVK.........DNGC.....GKDL
                                  200
SOL2 QKMHDNVQGKTETPDWESLIADASELLIFDSPDASEAFRCFMMQRASNSEARFRNGVEKQ
TSO1 QNMLVVREG.NDTPGCGRLISDATELLVFRSPNDSEAFRC.LVDKISSSERRFCAGV.KS
SOL1 QVMMDNVKKKSDTPDWETLIAATTE.LIYGSPRESEAFSC.LLKKTSNSEARLRGSIT.A

SOL2 TMQHDSNKE.PESANAIPYEVNSGV.ISQAVSLLHR.GIRRRCLDFEMPGNK.......Q
TSO1 TKRPDINKDIPANGSSNENQPLAVLPTNESVFNLHRGGMRRRCLDFEMPGKRKKDIVDDQ
SOL1 TSVAVTNTDVVNNESESV..........DALSILHR.GVRRRCLDFEVKGN.........
               300
SOL2 TSSENNTAACESSSRCVVPSIGLHLNAILMSSKDCKTNVTQDYSCSANIQVGLQRSISTL
TSO1 QSVCDNNVAGESSSSCVVPGIGLHLNAVAMSAKDSNISVIHGYSISGEIQKSFSGSTTPI
SOL1 ....NQQTLGESSSSCVVPSIGLHLNTIAMSSKDK..NVANEYSFSGNIKVGVQSSLTPV
                                                          400
SOL2 ....QDSLDQTEN.EIREDADQDVPVEPALQELNLSSPKKKS.................
TSO1 ..QSQDTVQETSD.QAENEPVEEVPKALVFPELNLGSLKKKMRKSEQAGEGE.SCKRCNC
SOL1 LHSQHDIVRENESGKDSGQIIEVVPKSLASVDLTPISPKKKRRKSEQSGEGDSSCKRCNC
                                                  R→Stop(tso1-3,4)
SOL2 ........YCECFAAGVYCIEPCSCIDCFNKPIHEDVVLATRKQIESRNPLAFAPKVIRN
TSO1 KKSKCLKLYCECFAAGVYCIEPCSCIDCFNKPIHEETVLATRKQIESRNPLAFAPKVIRN
SOL1 KKSKCLKLYCECFAAGFYCIEPCSCINCFNKPIHKDVVLATRKQIESRNPLAFAPKVIRN
                                      C→Y(tso1-1)
SOL2 SDSVQETGDDASKTPASARHKRGCNCKKSNCLKKYCECYQGGVGCSINCRCEGCKNAFGR
TSO1 ADSIMEASDDASKTPASARHKRGCNCKKSNCMKKYCECYQGGVGCSMNCRCEGCTNVFGR
SOL1 SDSIIEVGEDASKTPASARHKRGCNCKKSNCLKKYCECYQGGVGCSINCRCEGCKNAFGR

SOL2 KDGS.SIDMEAEQEEENETSEKSRTAKSQQNTEVL..MRKDMSSALPTTPTPIYRPELVQ
TSO1 KDGSLLVIMESKLEENQETYEK.RIAKIQHNVEVSKEVEQNPSSDQPSTPLPPYRHLVVH
SOL1 KDGSLF.....EQDEENETSGTPGTKKTQQNVELFKP......AAPPSTPIP.FRQPLAQ
                  600
SOL2 LPFSSSKNRMPPPQSLLGGGSSSGIFNSQYLRKPDISLSQSRIEKS.FETVAVDGAEQMP
TSO1 QPF.LSKNRLPPTQFFLGTGSSS.......FRKPNSDLAQSQNEKKPLETVTEDKTEIMP
SOL1 LPISSNNRLLPPQSHFHHGAIGSSSSGIYNIRKPDMSL....LSHSRIETITED.IDDMS

SOL2 EILIHSPIPNIKSVSPNGKRVSPPHMESSSSGSILGRRNGGRKLILQSIPSFPSLTPQH
TSO1 EILLNSPIANIKAISPNSKRVSPPQPGSSESGSILRRRGNGRKLILRSIPAFPSLNPNQ
SOL1 ENLIHSPIT...TLSPNSKRVSLSHLDSPES.TPWRRNGEGRNLI.RSFPTFPSLTPHH
```

B

```
1-TSO1        CNCKKSKCLKLYCECFAAGVYCIEPCSCIDCFN
1-GmPC        CNCKKSKCLKLYCDCFAAGTYCTDPCACQGCLN
1-CeJC8.6a    CNCTKSQCLKLYCDCFANGEFC.RDCNCKDCHN
1-AtF19B15.3 CNCKHSRCLKLYCECFASGTYC.DGCNCVNCFN
1-HsTes              LAGYCDCFASGDFCNNCNCNNCCNN
2-TSO1        CNCKKSNCMKKYCECYQGGVGCSMNCRCEGCTN
2-GmPC        CNCKRSMCLKKYCECYQANVGCSSGCRCEGCKN
2-CeJC8.6a    CHCKKSGCLKNYCECYEAKVPCTDRCKCKGCQN
2-AtF19B15.3 CHCKKSGCLKKYCECFQANILCSENCKCLDCKN
2-HsTes       CNCRRSGCLKNYCECYEAQIMCSSICKCIGCKN
  AtCLF       CHCAKSQCRSRQCPCFAADREC.DPDVCRNCWV
  DmKLP3A     CKCRTK.CTTKRCGCLSGNNACSETCVCKNCRN
  Cons        C-C--S-CLK-YC-C------C---C-C--C-N
```

C

```
 TSO1        RNPLAFAPKV
 GmPC        RNPIAFAPKI
 CeJC8.6a    RNPNAFKPKI
 AtF19B15.3 RNPFAFRPKI
 HsTes       RNPEAFQPKI
 Cons        RNP-AF-PKI
```

Fig. 3. Sequence alignment of TSO1 with homologous proteins. (A) Sequence similarity was observed throughout the TSO1 (L*er*), SOL1 and SOL2 (Co-3) proteins. CXC domains are shaded in gray. The RGD sequence (bold) found in TSO1 was not conserved in SOL1 or SOL2. Amino acids altered in *tso1-1*, *tso1-3* and *tso1-4* are highlighted and in bold, with the defect shown above the sequences. Numbering is according to the TSO1 sequence. (B) Alignment of representative CXC domains. The first and second repeats are numbered 1-, and 2-, respectively, while proteins with a single sequence are not numbered. The selected sequences were TSO1, a cysteine-rich 'polycomb-like' protein from soybean (GmPC, AJ010165), a protein of unknown function from *Arabidopsis thaliana* (At19B15.3, CAB43905), a protein of unknown function from *C. elegans* (CeJC8.6a, Z82274), a human tesmin (HsTes, NM_004923), the CURLY LEAF protein from *Arabidopsis thaliana* (AtCLF, Y10580) and *D. melanogaster* KLP3A (DmKLP3A, L19117). Consensus (Cons) indicates residues identical in ≥70% of the illustrated sequences. (C) A short region between the cysteine-rich repeats of TSO1 was conserved in several of the other proteins. Protein designations and consensus are as in B. The *TSO1* and *SOL1* cDNA sequences have been deposited in GenBank under accession numbers AF204059, and AF205142, respectively.

Sequence analysis of the putative TSO1 protein

The *TSO1* coding region, isolated by RT PCR, includes a 2.4 kb open reading frame (ORF) encoding a putative protein of 695 amino acids (Fig. 3A). Two additional start codons, each immediately followed by a stop codon, precede the start codon for this long ORF (data not shown). Short upstream open reading frames have previously been shown to attenuate expression of downstream ORFs (Damiani and Wessler, 1993).

The TSO1 protein contained an arginine-glycine-aspartate (RGD) sequence and a potential nuclear localization sequence (NLS, amino acids 268-285, RRRCLDFEMPGKRKKDI; Fig. 3A). RGD sequences can be sufficient for recognition and binding by some integrins (Giancotti and Ruoslahti, 1999). TSO1 has two copies of a cysteine-rich region, with each region containing three copies of the sequence cysteine-X-cysteine (Fig. 3B). Regions containing several such motifs have been referred to as 'CXC domains' (Hobert et al., 1996).

Database searches identified two *Arabidopsis* paralogs (***TSO1*-L**ike, *SOL*) and one soybean homolog (accession no. AJ010165) of TSO1 which showed significant sequence conservation throughout their length, especially in the CXC domains (Fig. 3A,B; data not shown). The function of these proteins is not known. While *SOL2* (dl3425c, Z97337) resides on chromosome 4, *SOL1* is adjacent to *TSO1* and is included in cF45. Notably, the RGD sequence of TSO1 was not conserved in any of the other three sequences (Fig. 3A; data not shown).

In addition to these three homologous sequences, a number of putative proteins were identified which were similar to TSO1 only in the two CXC domains and a conserved ten amino acid sequence in the region between the CXC domains (Fig. 3B,C). This arrangement of conserved regions was found in five other *Arabidopsis* sequences (T2G17.9, AC00681; F19B15.30, AL078470; T9J14.20, AC009465; F18A17, AC006081; MSL1, AB012247), one *Caenorhabditis elegans* sequence (CeJC8.6a, Z82274.1), one *Drosophila melanogaster* sequence (AC007452), and three animal tesmins (testes-specific proteins of unknown function; Sugihara et al., 1999; accession nos. NM_004923, U67176, U77383). The first CXC domain in the tesmins, and in some of the other proteins (e.g. T2G17.9, AC00681), was shorter by eight to ten amino acids at its N-terminal end than the corresponding CXC domain of TSO1 (Fig. 3B; data not shown).

The CXC repeats of TSO1 also aligned with the single CXC domains of other classes of proteins (Fig. 3B). These included a class of animal kinesins including *Drosophilia* KLP3A (associated with the spindle apparatus during meiosis and fertilization; Williams et al., 1997, 1995) and the *enhancer of zeste* [E(z)]-type proteins, such as *Arabidopsis* CURLY LEAF (CLF; Goodrich et al., 1997), where the TSO1 CXC domains aligned with the C-terminal region of a larger CXC domain.

The mutation in *tso1-3* and *tso1-4* changes the first amino acid of the conserved motif shown in Fig. 3C to a stop codon. This leads to elimination of both this motif and the second CXC domain (Fig. 3). The mutation in *tso1-1* causes a change of a conserved cysteine in the second CXC domain to tyrosine (Fig. 3A). Surprisingly, the amino acid substitution in *tso1-1* produces a more severe phenotype than seen for the nonsense mutation in *tso1-3* and *-4*.

mRNA accumulation patterns

RNA blots were hybridized with probes specific for *TSO1* and each of the *Arabidopsis TSO1* paralogs (Fig. 4). A 2.8 kb *TSO1* transcript was observed in all organs examined, with the highest levels detected in flowers (Fig. 4B). *SOL1* mRNA (2.2 kb) was also most abundant in flowers, but in addition was found at significant levels in leaves (Fig. 4C). *SOL2* mRNA (2.4 kb) was detected in all aerial organs (Fig. 4D). The relative signal intensities from RNA blots indicated that the *TSO1* and *SOL1* transcripts were present at similar levels in flowers, while the *SOL2* mRNA was less abundant.

In situ hybridizations of *TSO1* and *SOL1* transcripts were performed on floral meristems, flowers and embryos (Fig. 5). Sense strand probes produced little or no signal compared to anti-sense probes (data not shown), indicating that even weak signals with the anti-sense *TSO1* probe represented mRNA detection rather than non-specific background. In wild-type plants, *TSO1* mRNA was detected in the central region of floral meristems (Fig. 5A), and throughout flowers, with the highest level of mRNA in developing microspores and ovules (Fig. 5B-E). Following fertilization, the *TSO1* transcript was detected throughout globular, heart and torpedo stage embryos (Fig. 5H; data not shown). *SOL1* showed a similar pattern of mRNA accumulation to *TSO1* in developing microspores and ovules, but showed a relatively higher level of transcript accumulation in petals (Fig. 5I,J).

TSO1 mRNA was present at a higher level in *tso1-1* than in wild-type flowers (Fig. 5F,G). The *TSO1* mRNA was also more evenly distributed throughout the flowers, staining all floral structures including the sepals (Fig. 5G; data not shown).

Subcellular localization of TSO1 fusion proteins

To determine the subcellular localization of TSO1, translational fusions of TSO1 with green fluorescent protein (GFP; Sheen et al., 1995) were produced by transient expression of chimeric gene constructs in onion epidermal cells. Both TSO1-GFP, and GFP-TSO1 were confined exclusively to the nucleus, in clear contrast to GFP alone (Fig. 6).

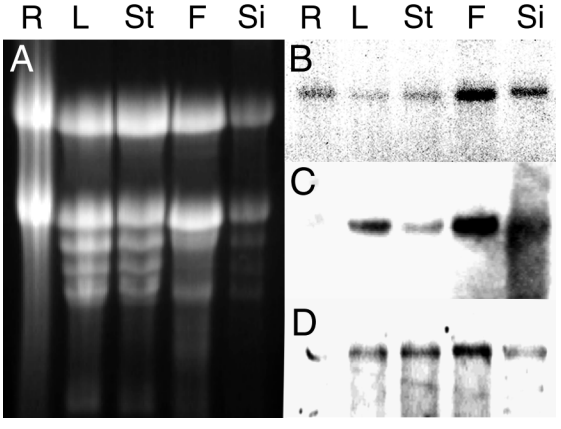

Fig. 4. RNA blot hybridizations with probes specific for *TSO1*, *SOL1* and *SOL2*. (A) Ethidium bromide staining of total RNA from roots (R), rosette leaves (L), stems (St), flowers (F) and siliques (Si). (B,C,D) Hybridizations of a blot of the gel shown in A with *TSO1*, *SOL1* and *SOL2* probes, respectively.

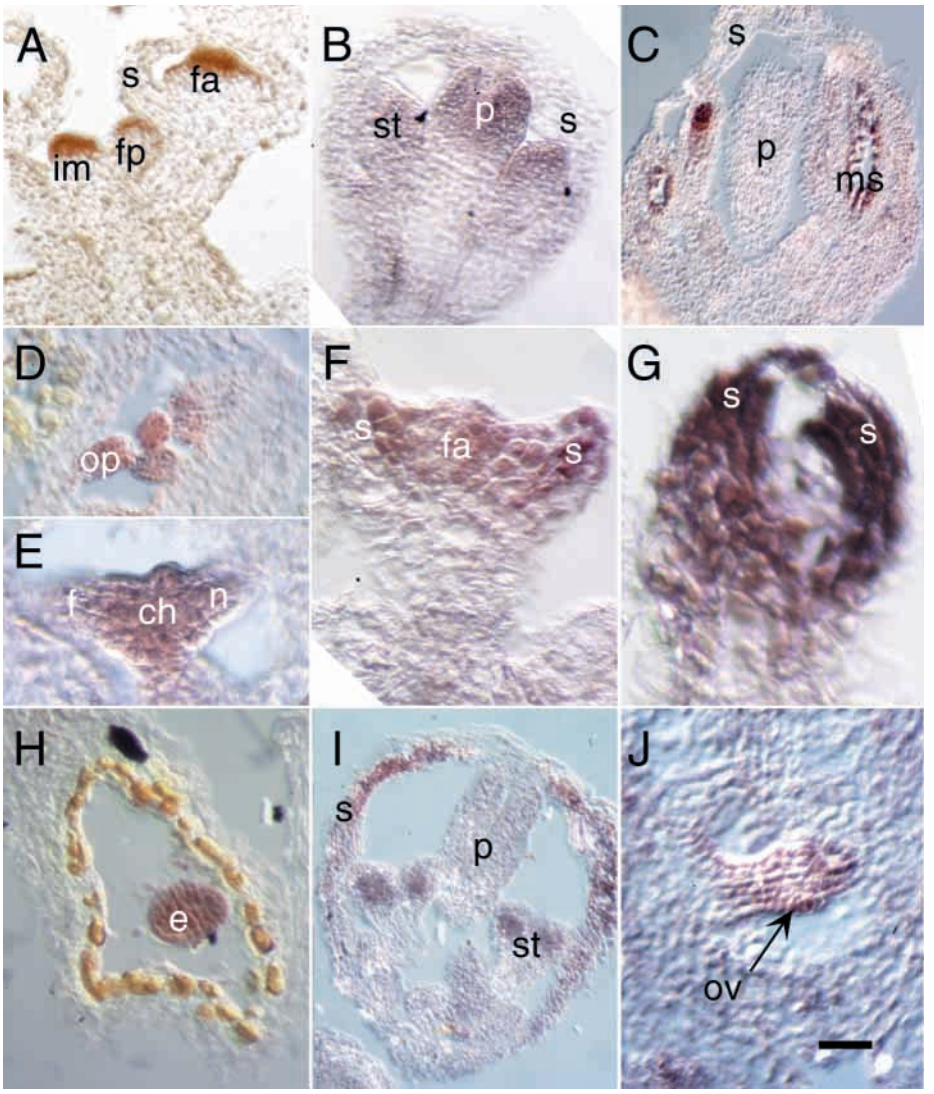

Fig. 5. Patterns of *TSO1* and *SOL1* mRNA accumulation examined by in situ hybridization. (A) *TSO1* mRNA was detected in the inflorescence meristem (im), stage 2 floral primordia (fp) and floral apex (fa). s, sepals. (B) In stage 7 flowers, *TSO1* mRNA was in stamen (st) and pistil (p) primordia. (C) In stage 11 flowers *TSO1* transcript was abundant in microspores (ms) and (D) ovule primordia (op). (E) In developing ovules, *TSO1* mRNA was found throughout the funiculus (f), chalaza (ch) and nucellus (n). (F) In *tso1-1* flowers at stage 3, *TSO1* transcript was found throughout the floral apex and developing sepals. (G) *TSO1* mRNA levels were somewhat variable in the highly reduced *tso1-1* flowers, but were almost always higher than in wild type. (H) Following fertilization, *TSO1* mRNA was detected throughout a globular stage embryo (e). (I) *SOL1* transcripts accumulated in sepals and microspore mother cells in the stamens of stage 8 flowers. (J) *SOL1* message was also found throughout immature ovules (ov). Bar, 20 μm in A, D, E and J; 5 μm in F; 50 μm in B and H; 30 μm in G and 100 μm in C and I. Stages according to Smyth et al. (1990).

DISCUSSION

Map-based methods were used to identify the *TSO1* gene, and identification was confirmed by complementation of *tso1* mutants through transformation. All three *tso1* isolates had mutations which would lead to production of defective proteins. The identical mutations observed in the *tso1-3* and *tso1-4* isolates suggests that the previously described differences in phenotype between these lines (Hauser et al., 1998) were due to second site mutations or other differences in their genetic backgrounds.

Constructs that fully complemented *tso1-3* only partially complemented the *tso1-1* mutation (Fig. 1; data not shown). The missense mutation in *tso1-1* should lead to production of a full-length protein with a single amino acid substitution of tyrosine for a conserved cysteine in the second CXC domain (Fig. 3). We hypothesize that the aberrant TSO1-1 protein might have the potential for deleterious effects on cell division and expansion. Since *tso1-1* appears to be fully recessive, such effects must be sensitive to the ratio between wild-type and mutant protein levels. One model for this would be that TSO1-1 competes with wild-type TSO1 protein in some molecular interaction and that with the one-to-one ratio in heterozygous plants the wild-type activity would predominate. In homozygous mutants harboring a wild-type transgene, however, the higher relative level of mutant protein would result in an altered phenotype. Despite their weaker phenotypes, *tso1-3* and *tso1-4* may be null alleles due to the premature stop codon, which leads to elimination of the conserved central sequence and the second CXC domain. Other recessive mutants have been reported where a null allele exhibits a less severe phenotype than an allele expressing an altered protein (e.g. *superman*; Sakai et al., 1995).

The expression of *TSO1* in all parts of flowers is consistent with the phenotype of *tso1-1* mutants, in which all floral organs are affected. The pattern was altered in *tso1-1* mutants where we observed increased levels and an altered pattern of accumulation of *TSO1* mRNA (Fig. 5F,G; data not shown). This could indicate that expression of *TSO1* is self-regulating. Normal developmental progress, which requires TSO1, might lead to repression of this gene later in development. Absence of normal progress might lead to increased *TSO1* expression to attempt to correct this defect. The feedback mechanism controlling *TSO1* expression could be imprecise, resulting in the variable *TSO1* mRNA levels observed in *tso1-1* flowers.

Sequence conservation in the two paralogous genes in *Arabidopsis* and a soybean gene of unknown function was seen throughout the length of TSO1 (Fig. 3A,B; data not shown). This implies that functional residues are found throughout the protein, and that functions of these proteins might be conserved

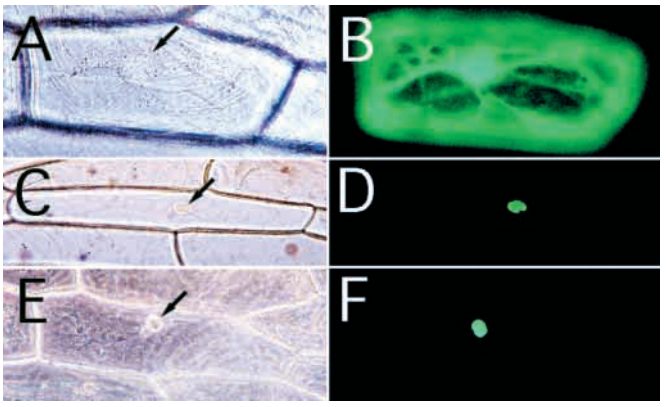

Fig. 6. Localization of GFP fusions with TSO1 in onion epidermal cells. (A,C,E) Phase contrast optics; (B,D,F) fluorescence images using filters for detection of GFP fluorescence. (A,B) Cell transformed with a 35S:GFP plasmid. Because GFP protein is smaller than the nuclear exclusion limit, it diffuses across the nuclear membrane and is distributed throughout the cytosol and nucleus (Scott et al., 1999). (C,D) Cell transformed with 35S:TSO1-GFP translational fusion. The TSO1-GFP fusion protein was confined to the nucleus. (E,F) Cell transformed with 35S:GFP-TSO1 translational fusion. The GFP-TSO1 translational fusion was confined to the nucleus. Nuclei are indicated by arrows in A, C and E.

in a variety of plants. The high degree of similarity between *TSO1* and the paralogous genes raises the question of functional redundancy. Despite the similarity between *SOL1* and *TSO1* sequences and expression patterns (Figs 3-5), additional wild-type copies of *SOL1* were not sufficient to complement *tso1* mutations (data not shown). The simplest explanation for these data is that these two genes do not encode interchangeable protein products.

The TSO1 protein sequence includes an RGD sequence, which is the recognition sequence for integrins (for review see Giancotti and Ruoslahti, 1999). While RGD-mediated signaling does occur in plants (Canut et al., 1998; Schindler et al., 1989), the nuclear localization of TSO1 would preclude the normal extracellular interaction of RGD sequences with integrins. In addition, the absence of the RGD motif in the two *Arabidopsis* TSO1 paralogs does not support an important role for this motif in the TSO1 protein family.

The presence and spacing of cysteine residues are conserved between the CXC domains of TSO1 and the CXC domain in C-terminal region of the *Drosophila* kinesin-like protein KLP3A (Fig. 3B). The C-terminal regions of kinesins have been hypothesized to function as cargo domains, so that the association of KLP3A with cytoskeletal spindles during meiosis and fertilization (Williams et al., 1997, 1995) may indicate interactions between the CXC domain and chromatin. TSO1 CXC domains also align with part of the CXC domains of E(z) proteins (Fig. 3B). Mutations in E(z) CXC domains disrupt chromosome binding of these proteins, indicating likely interaction between this motif and chromatin (Carrington and Jones, 1996). Thus, in two cases, CXC domains may participate in interactions with protein or DNA components of chromatin.

Other proteins from both plants and animals not only have conserved cysteine residues in the CXC domains (Fig. 3B), but also have two such domains flanking a conserved amino acid sequence (Fig. 3C). This conservation across kingdoms implies that this entire region may act as a single functional unit. Unfortunately, these other sequences derived primarily from genomic surveys, and no information on the function of these proteins is available. While the significant differences in sequence make it unlikely that TSO1 and these other related proteins functionally duplicate E(z)- or KLP3A-type proteins, the CXC domains could still participate in chromatin or DNA interactions.

We have previously hypothesized that TSO1 could be a regulator of cytoskeletal dynamics, and that this could explain all of the phenotypic effects of *tso1* mutations. The nuclear localization of TSO1 appears to preclude a direct interaction between TSO1 and the cytoskeleton. However, if TSO1 can bind directly to DNA, then it could function as a transcriptional regulator which could modulate expression of genes encoding cytoskeletal elements or factors which regulate cytoskeletal activity. TSO1 could thus indirectly regulate cytoskeletal function.

Alternatively, TSO1 could bind chromatin and participate in alignment or separation of chromosomes during cell division. Failure of these processes could explain the aberrant nuclei and DNA content of *tso1* mutant cells (Liu et al., 1997). How closely linked this process is to the subsequent events of cell plate formation, cell division, and cell expansion is not clear. If these processes are linked, then aberrant chromosome segregation could lead to defects in these later processes, explaining these effects of *tso1* mutations. However, in *tso1-3*, the effects on cell growth were prominent, while defects in chromosome separation or nuclear structures were not detected. For this model to be correct, subtle disruptions in chromosome/nuclear events would have to have significant ramifications in cell division/expansion.

If TSO1 does not interact with chromatin, then it could participate in other aspects of cytoskeletal interactions with chromosomes, or in regulation of structure or dynamics of the nuclear matrix or nuclear envelope. Defects in these processes in *tso1* mutants could also disrupt cell expansion and division. The novelty of the TSO1 sequence suggests the existence of previously unknown modes of regulation of fundamental directional processes in cells. The conservation of regions of TSO1 sequence in animals indicates that such functions may be common to all higher eukaryotes. Different models for TSO1 action would be suggested by the nature of proteins with which TSO1 interacts, and the cloning of TSO1 now provides the tools necessary for identification of such interacting proteins. The association of this member of a new class of CXC domain proteins with a mutant phenotype provides the initial key to determination of the cellular functions of the entire family.

We thank Laura Wilhelmi, Daphne Pruess, Linda Ehler and Zhongchi Liu for sharing information prior to publication, Bo Liu for assistance with microscopy, Maureen Clancy for assistance with particle bombardment, Neil Olszewski for the pOCA28 vector, Joanne Chory and Danielle Friedrichsen for the pGFP1.1.5 construct, the NSF Arabidopsis Biological Resources Center at Ohio State for mapping stocks, and Jean Broadhvest and other members of the Gasser lab for helpful discussions. This work was supported by the National Science Foundation (IBN-9808395) and an NSF Plant Cell Biology Training Grant Fellowship to J. Q. H.

REFERENCES

Assaad, F. F., Mayer, U., Wanner, G. and Juergens, G. (1996). The *KEULE* gene is involved in cytokinesis in Arabidopsis. *Mol. Gen. Genet.* **253**, 267-277.

Canut, H., Carrasco, A., Galaud, J., Cassan, C., Bouyssou, H., Vita, N., Ferrara, P. and Pont-Lezica, R. (1998). High affinity RGD-binding sites at the plasma membrane of *Arabidopsis thaliana* links the cell wall. *Plant J.* **16**, 63-71.

Carrington, E. and Jones, R. (1996). The *Drosophila Enhancer of zeste* gene encodes a chromosomal protein: examination of wild-type and mutant protein distribution. *Development* **122**, 4073-4083.

Clough, S. J. and Bent, A. F. (1998). Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* **16**, 735-743.

Damiani, R. D. and Wessler, S. R. (1993). An upstream open reading frame represses expression of LC, a member of the R-B family of maize transcriptional activators. *Proc. Natl. Acad. Sci. USA* **90**, 8244-8248.

Doonan, J. and Fobert, P. (1997). Conserved and novel regulators of the plant cell cycle. *Curr. Opin. Cell Biol.* **9**, 824-830.

Feinberg, A. P. and Vogelstein, B. (1983). A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. *Analyt. Biochem.* **132**, 6-13.

Figurski, D. H. and Helinski, D. R. (1979). Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in *trans*. *Proc. Natl. Acad. Sci. USA* **76**, 1648-1652.

Gasser, C. S., Budelier, K. A., Smith, A. G., Shah, D. M. and Fraley, R. T. (1989). Isolation of tissue-specific cDNAs from tomato pistils. *Plant Cell* **1**, 15-24.

Giancotti, F. and Ruoslahti, E. (1999). Integrin signaling. *Science* **285**, 1028-1032.

Gleave, A. P. (1992). A versatile binary vector system with a T-DNA organisational structure conducive to efficient integration of cloned DNA into the plant genome. *Plant Mol. Biol.* **20**, 1203-1207.

Goodrich, J., Puangsomlee, P., Martin, M., Long, D., Meyerowitz, E. and Coupland, G. (1997). A polycomb-group gene regulates homeotic gene expression in Arabidopsis. *Nature* **386**, 44-51.

Hauge, B. M., Hanley, S. M., Cartinhour, S., Cherry, J. M., Goodman, H. M., Koornneef, M., Stam, P., Chang, C., Kempin, S., Medrano, L., et al. (1993). An integrated genetic RFLP map of the *Arabidopsis thaliana* genome. *Plant J.* **3**, 745-754.

Hauser, B. A., Villanueva, J. M. and Gasser, C. S. (1998). Arabidopsis TSO1 regulates directional processes in cells during floral organogenesis. *Genetics* **150**, 411-423.

Hobert, O., Jallal, B. and Ullrich, A. (1996). Interaction of Vav with ENX-1, a putative transcriptional regulator of homeobox gene expression. *Mol. Cell. Biol.* **16**, 3066-3073.

Jakobsen, K., Breivold, E. and Hornes, E. (1990). Purification of messenger RNA directly from crude plant tissues in 15 minutes using magnetic oligo-dT microspheres. *Nucl. Acids Res.* **18**, 3669.

Kawasaki, E. S. (1990). Amplification of RNA. In *PCR Protocols*, (ed. M. A. Innis, D. H. Gelfand, J. J. Sninsky and T. J. White), pp. 21-27. San Diego, CA: Academic Press.

Kay, R., Chan, A., Daly, M. and McPherson, J. (1987). Duplication of CaMV 35S promoter sequences creates a strong enhancer for plant genes. *Science* **236**, 1299-1302.

Klimyuk, V. and Jones, J. (1997). *AtDMC1*, the *Arabidopsis* homologue of the yeast *DMC1* gene: characterization, transposon-induced allelic variation and meiosis-associated expression. *Plant J.* **11**, 1-14.

Konieczny, A. and Ausubel, F. M. (1993). A procedure for mapping *Arabidopsis* mutations using co-dominant ecotype-specific PCR-based markers. *Plant J.* **4**, 403-410.

Lander, E. R., Green, P., Abrahamson, J., Barlow, A., Daly, M., Lincoln, S. E. and Newburg, L. (1987). MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics* **1**, 174-181.

Liu, C.-M., Johnson, S. and Wang, T. L. (1995). *cyd*, a mutant of pea that alters embryo morphology is defective in cytokinesis. *Dev. Genet.* **16**, 321-331.

Liu, Z., Running, M. P. and Meyerowitz, E. M. (1997). *TSO1* functions in cell division during *Arabidopsis* flower development. *Development* **124**, 665-672.

Lukowitz, W., Mayer, U. and Jürgens, G. (1996). Cytokinesis in the Arabidopsis embryo involves the syntaxin-related KNOLLE gene product. *Cell* **84**, 61-71.

Mozo, T., Fischer, S., Shizuya, H. and Altmann, T. (1998). Construction and characterization of the IGF Arabidopsis BAC library. *Mol. Gen. Genet.* **258**, 562-570.

Olszewski, N. E., Martin, F. B. and Ausubel, F. M. (1988). Specialized binary vector for plant transformation: expression of the *Arabidopsis thaliana AHAS* gene in *Nicotiana tabacum*. *Nucl. Acids Res.* **16**, 10765-10782.

Sakai, H., Medrano, L. J. and Meyerowitz, E. M. (1995). Role of *SUPERMAN* in maintaining *Arabidopsis* floral whorl boundaries. *Nature* **378**, 199-203.

Schindler, M., Meiners, S. and Cheresh, D. (1989). RGD-dependent linkage between plant cell wall and plasma membrane: consequences of growth. *J. Cell Biol.* **108**, 1955-1965.

Schumacher, K., Vafeados, D., McCarthy, M., Sze, H., Wilkins, T. and Chory, J. (1999). The Arabidopsis *det3* mutant reveals a central role for the vacuolar H+-ATPase in plant growth and development. *Genes Dev.* **13**, 3259-3270.

Scott, A., Wyatt, S., Tsou, P.-L., Robertson, D. and Allen, N. S. (1999). Model system for plant cell biology: GFP imaging in living onion epidermal cells. *Biotechniques* **26**, 1125-1132.

Sheen, J., Hwang, S. B., Niwa, Y., Kobayashi, H. and Galbraith, D. W. (1995). Green-fluorescent protein as a new vital marker in plant cells. *Plant J.* **8**, 777-784.

Smyth, D. R., Bowman, J. L. and Meyerowitz, E. M. (1990). Early flower development in *Arabidopsis*. *Plant Cell* **2**, 755-767.

Staehelin, L. and Hepler, P. (1996). Cytokinesis in higher plants. *Cell* **84**, 821-824.

Sugihara, T., Wadhwa, R., Kaul, S. C. and Mitsui, Y. (1999). A novel testis-specific metallothionein-like protein, tesmin, is an early marker of male germ cell differentiation. *Genomics* **57**, 130-136.

Varagona, M. J., Schmidt, R. J. and Raikhel, N. V. (1992). Nuclear localization signal(s) required for nuclear targeting of the maize regulatory protein Opaque-2. *Plant Cell* **4**, 1213-1227.

Vieira, J. and Messing, J. (1987). Production of single stranded plasmid DNA. *Methods Enzymol.* **153**, 3-11.

Vielle-Calzada, J.-P., Thomas, J., Spillane, C. and Grossniklaus, U. (1999). Maintenance of genomic imprinting at the Arabidopsis *MEDEA* locus requires zygotic DDM1 activity. *Genes Dev.* **13**, 2971-2982.

Williams, B., Dernburg, A., Puro, J., Nokkala, S. and Goldberg, M. (1997). The Drosophila kinesin-like protein KLP3A is required for proper behavior of male and female pronuclei at fertilization. *Development* **124**, 2365-2376.

Williams, B. C., Riedy, M. F., Williams, E. V., Gatti, M. and Goldberg, M. L. (1995). The *Drosophilia* kinesin-like protein KLP3A is a midbody component required for central spindle assembly and initiation of cytokinesis. *J. Cell Biol.* **129**, 709-723.