Planta (2008) 227:853–866
DOI 10.1007/s00425-007-0662-4

ORIGINAL ARTICLE

Characterization of *leafy cotyledon1-like* during embryogenesis in *Theobroma cacao* L.

Laurence Alemanno · Martine Devic · Nicolas Niemenak · Christine Sanier · Jocelyne Guilleminot · Mariannick Rio · Jean-Luc Verdeil · Pascal Montoro

Received: 13 March 2007 / Accepted: 19 October 2007 / Published online: 18 December 2007

Abstract *Theobroma cacao* L., an economically important crop for developing countries, can be experimentally propagated by somatic embryogenesis. Because of their potential roles in embryogenesis, a gene candidate strategy was initiated to find gene homologues of the members of the leafy cotyledon family of transcription factors. A homologue of the *leafy cotyledon1-like* gene, that encodes the HAP 3 subunit of the CCAAT box-binding factor, was found in the cocoa genome (*TcL1L*). The translated peptide shared a high amino acid sequence identity with the homologous genes of *Arabidopsis thaliana*, *Phaseolus coccineus* and *Helianthus annuus*. *TcL1L* transcripts mainly accumulated in young and immature zygotic embryos, and, to a lesser extent, in young and immature somatic embryos. In situ hybridization specified the localization of the transcripts as being mainly in embryonic cells of young embryos, the meristematic cells of the shoot and root apex of immature embryos, and in the protoderm and epidermis of young and immature embryos, either zygotic or somatic. Non-embryogenic explants did not show *TcL1L* expression. Ectopic expression of the *TcL1L* gene could partially rescue the *Arabidopsis lec1* mutant phenotype, suggesting a similarity of function in zygotic embryogenesis.

Keywords Embryogenesis · Gene expression · HAP3 · *Leafy cotyledon1-like* · *Theobroma cacao* · Transcription factor

Abbreviations

SERK	*Somatic embryogenesis receptor kinase*
BBM	*Baby boom*
LEC	*Leafy cotyledon*
ABI3	*Abscisic acid insensitive3*
SSEE	Secondary somatic embryogenic explants
VYSE	Very young somatic embryos
YSE	Young somatic embryos
NEE	Non-embryogenic explants
RT	Reverse transcriptase
PCR	Polymerase chain reaction

Introduction

Theobroma cacao L., exclusively cultivated in the intertropical area of the world, is a major source of income for developing countries. According to Eskes (2000), 70% of cultivated trees result from unselected material propagated by seeds. Because of the high heterozygous nature of seedlings in a given field, only 2–3% of the trees in a population of high-yielding families account for 60% of the yield (Irizarry and Rivera 1998). Somatic embryogenesis is one way of increasing yields, homogenizing cocoa production, hence smallholder incomes, by propagating elite material. Since the production of the first cocoa somatic embryo (Esan 1974), it has been possible to produce somatic

L. Alemanno · C. Sanier · M. Rio · J.-L. Verdeil · P. Montoro (✉)
Centre de Coopération Internationale en Recherche Agronomique pour le Développement, CIRAD, UMR DAP, TA A-96/03, Avenue Agropolis, 34398 Montpellier Cedex 5, France
e-mail: pascal.montoro@cirad.fr

M. Devic · J. Guilleminot
Laboratoire Génome et Développement des Plantes, University of Perpignan, UMR 5096, 52 Avenue Paul Alduy, 66860 Perpignan Cedex, France

N. Niemenak
Laboratory of Plant Physiology, Department of Biological Science, Higher Teachers' Training College, University of Yaoundé I, P.O. Box 47, Yaoundé, Cameroon

embryos and plantlets from a large number of genotypes (Maximova et al. 2002). Based on this procedure, field trials were established to evaluate this new planting material (Figueira and Alemanno 2004). The morphogenetic processes of zygotic and somatic embryogenesis have been described (Alemanno et al. 1996, 1997; Maximova et al. 2002). Despite those improvements, the efficiency of the process is not sufficient for a scaling-up step. Moreover, many genotypes remain recalcitrant to somatic embryogenesis (Figueira and Alemanno 2004).

Plant zygotic embryogenesis is classically divided into two distinct phases: the early morphological processes during which embryonic cell types, tissues and different organs, including the meristems, are formed; and the late maturation phase during which the fully developed embryo accumulates storage reserves, and enters the desiccation and dormancy state. Since the beginning of the genomic era, more information has become available about genes involved in the early and late phases of embryogenesis in the *Arabidopsis thaliana* model species. Such genes were discovered either from loss-of-function mutants or from the study of genes differentially expressed during embryogenesis. In the latter strategy, somatic embryogenesis appeared to be a powerful and convenient tool (Boutilier et al. 2002), especially for early phases of zygotic embryogenesis where the embryos are very small, with reduced accessibility. Several groups of genes can be distinguished according their expression during the embryogenesis stages: early embryogenesis, morphogenesis and/or maturation.

Somatic embryogenesis receptor kinase (*SERK*) was identified from in vitro somatic embryos of carrot (Schmidt et al. 1997), *Dactilys glomerata* (Somleva et al. 2000) and *A.thaliana* (Hecht et al. 2001), for which it was shown to play a key role in somatic embryogenesis induction. In *A. thaliana*, ectopic expression of this gene enhanced the embryogenic capacity of already embryogenic culture cells (Feher et al. 2003; Verdeil et al. 2007). It was also expressed during early embryogenesis in *A. thaliana* (Hecht et al. 2001). This gene encodes a protein that is part of the receptor-like kinase-LRR family (Schmidt et al. 1997). The LRR domains likely play a role in signal perception and transduction during somatic embryogenesis induction (Schmidt et al. 1997; Hecht et al. 2001). Competent cells may contain an inactive receptor, activated by the presence of the proper ligand to switch on the embryogenic programme (Hecht et al. 2001). Recent in vivo isolation of the SERK1 signalling complex has provided supportive evidence for a role in the brassinolide signalling pathway (Karlova et al. 2006).

Another class of genes is exclusively expressed during late embryogenesis *abscisic acid insensitive3* (*ABI3*), *ABI4*, *ABI5* (Giraudat et al. 1992; Finkelstein et al. 1998; Finkelstein and Lynch 2000). The role of those genes, identified through loss-of-function mutant studies, involves the preparation of embryos to desiccation and dormancy. Finally, another group of genes, *baby boom* (*BBM*) and *leafy cotyledon* (*LEC)* genes, was found to be expressed in both early and late phases of embryogenesis. *BBM* was isolated by a subtractive hybridization approach with *Brassica napus* microspore-derived embryos, while *leafy cotyledon* genes were isolated from loss-of-function mutants in *A. thaliana*. The protein encoded by the *BBM* gene showed similarity to the transcription factors of the AP2/ERF family (Boutilier et al. 2002). When expressed ectopically in *A. thaliana* and *B.napus*, this gene induced somatic embryo formation. Leafy cotyledon proteins are central regulators of both the early and late phases of embryogenesis. They are defined by mutations at three loci *LEC1*, *LEC2* and *FUSCA3* (*FUS3*). They intervene in several functions: maintaining suspensor cell identity, specification of cotyledon identity, desiccation tolerance, synthesis and accumulation of storage reserves and inhibition of germination (Harada 2001). *LEC1* encodes a protein with sequence similarity to the HAP3 sub-unit of CCAAT binding factors (Lotan et al. 1998). It was expressed in early stages of embryogenesis and accumulated in the embryo proper, the suspensor and the endosperm. Ectopic expression of the *LEC1* gene in *A. thaliana* seedlings was concomitant to other genes usually activated during embryogenesis. Some plants were able to give rise to embryo-like structures on their leaves, indicating that this gene was sufficient to induce somatic embryo formation from vegetative cells (Lotan et al. 1998). *LEC2* also encodes a transcription factor containing a plant specific B3 domain (Stone et al. 2001) and belongs to the same gene family like two other regulators of seed maturation, *ABA insensitive 3* (*ABI3*) and *FUSCA 3* (*FUS3*). B3 domain confers DNA binding activity to these proteins (reviewed in Harada 2001). Ectopic expression of *LEC2* in transgenic *Arabidopsis* induced the formation of somatic embryos and organ-like structures, and induced embryonic characteristics in vegetative tissues (Stone et al. 2001). More recently, Kwong et al. (2003) identified genes encoding *Arabidopsis* HAP3 subunits. One of them, closely related to *LEC1* and named *LEC1-LIKE* (*L1L*), is required for normal embryo development (Kwong et al. 2003). When ectopically expressed, *L1L* can complement *LEC1* functions (Kwong et al. 2003). *LEC1* and *LEC2* are considered to be transcriptional regulators capable of establishing a cellular environment sufficient to initiate embryo development (Feher et al. 2003). They act synergistically together with *ABI3* and *FUS3* to control multiple processes during seed development (Parcy et al. 1997). However, some degree of hierarchy exists in this regulatory network. *LEC1* (Kagaya et al. 2005) and *LEC2* (To et al. 2006) exert their action, at least partially, through *ABI3* and *FUS3*. Furthermore, the recent discovery of the role of *FUS3* on the synthesis of ABA and GA and the influence of its expression by auxin reconciles

the well know requirement for hormones with the transcriptional regulation in seed maturation (Gazzarrini et al. 2004).

Because *LEC* genes are essential regulators of the early phases of embryogenesis in *Arabidopsis*, several studies set out to extend those findings to the somatic embryogenesis pathway as well as to other plant species. *LEC1* was isolated in maize (Zhang et al. 2002) and carrot (Yazawa et al. 2004) and *LEC1-L* was isolated from *Phaseolus coccineus* (Kwong et al. 2003) and several other plants reported in the NCBI Gene Database: *Helianthus annuus* lec1L (AJ879074, AJ863113); *Bixa orellana* Lec 1 (AJ489457); *Oryza sativa Lec1* (AY264284). Expression of *LEC1* was demonstrated during maize (Zhang et al. 2002) and carrot (Yazawa et al. 2004) somatic embryogenesis. More recently, Gaj et al. (2005) showed that somatic embryogenesis capacity from *lec* and/or *fus* mutants was highly reduced compared to wild plants, suggesting that those genes are also essential for somatic embryogenesis induction.

Given the recalcitrance of cocoa in somatic embryogenesis and current interest in cocoa genomics (Jones et al. 2002; Bennett 2003; Kuhn et al. 2003), we launched a gene candidate strategy to isolate *leafy cotyledon* gene homologues in *T. cacao*. This study aims at the isolation and, the structural and functional characterization of the *T.cacao* gene homologous to the *Arabidopsis thaliana leafy cotyledon1-like* gene, named *T.cacao leafy cotyledon1-like* (*TcL1L*). We characterized the *TcL1L* messenger RNA (mRNA) in zygotic and somatic embryos at various stages of development and in vegetative organs and tissues. Furthermore, we studied if the *TcL1L* gene was able to restore an *Arabidopsis defective mutan*t.

Materials and methods

Plant materials

Zygotic embryos were harvested for gene isolation from mature and immature pods of hand-pollinated *T.cacao* L. IMC 67 (Upper-Amazon Forastero) × DHS 30 (São Tomé Hybrid) provided by CIAT (Centro de Investigãçaó agronómica e tecnológica, São Tomé-et-Principe). For gene expression analyses, zygotic embryos were isolated from pods of clone Sca-6 provided by Cameroonian and Brazilian Institutes, and then separated into two batches: (1) immature, consisting of a mix of stages IIIz, IVz, and Vz respectively 80, 100 and 110 days after pollination as described by Alemanno et al. (1997) and (2) mature corresponding to stage VIIIz. Somatic embryos were produced according to a two-step protocol consisting of primary somatic embryogenesis (PSE) induced from staminodes of immature flower buds, and secondary somatic embryogenesis (SSE) induced from primary somatic embryos. Two different genotypes, one grown in the greenhouse, UF667, known to be non-embryogenic, and one grown in the field (tree 29) in Côte d'Ivoire, known to be embryogenic were compared (Li et al. 1998). This tree came from progenies of Upper Amazon × Amelonado and Upper Amazon × Trinitario crosses selected by CNRA (Centre National de Recherche Agronomique, Abidjan, Côte d'Ivoire). Explants were sub-cultured successively on three different media: first on PCG (primary callus growth) medium containing 2.4 µM 2,4-D and 22.7 nM TDZ (from day 0 to day 14), second on SCG (secondary callus growth) medium containing 2.4 µM 2,4-D and 1.4 µM BA (from day 15 to day 29) and third on ED (embryo development) medium depleted of plant growth regulators where primary somatic embryos regenerated (from day 30 to day 44). Secondary somatic embryogenesis was induced on SSE medium from fragments of cotyledon from primary somatic embryos of tree 29 only as described by Maximova et al. (2002). Explants were sub-cultured every 2 weeks respectively on SCG and ED media where secondary somatic embryos have been regenerated. For RT-PCR experiments, UF667 non-embryogenic primary calli, and tree 29 embryogenic explants at different stages of secondary somatic embryogenesis, were frozen in liquid N_2 and stored at −80°C for further RNA extraction. Samples from tree 29 were collected at days 0, 7, 14 and 21 corresponding to secondary somatic embryogenic explants (SSEE), at day 28 and day 42 for secondary somatic embryogenic explants bearing very young somatic embryos (SSEE + VYSE), at day 72 both for young somatic embryos (YSE) which are globular and cotyledonary embryos measuring less than 2 mm, and for non-embryogenic explants (NEE). YSE consist of cotyledonary embryos measuring between 2 and 5 mm and cotyledonary embryos measuring from 5 to 10 mm. The same explants were also fixed for in situ hybridization as described later.

Arabidopsis heterozygous *lec1-2* plants, Wassilevskija (WS-2) ecotype, were obtained from the Nottingham Arabidopsis Stock Centre (ID: N3867). The plants were grown in soil or in Petri dishes under constant illumination at 22°C. Transformed seedlings were selected on Murashige and Skoog medium (Murashige and Skoog 1962) supplemented with 100 mg l^{-1} kanamycin and 30 mg l^{-1} hygromycin and transferred to soil after 2 weeks.

Isolation of RNA

Frozen samples were ground in liquid nitrogen and the powder was suspended in an extraction buffer containing 4 M guanidinium isothiocyanate, 0.24 M sodium acetate, 0.03 M N-lauroyl sarcosine sodium salt, 22.5 mM PVP-40.000 and 14 mM β-mercaptoethanol. After initial centrifugation at

10,000*g* for 30 min at 4°C, RNAs were purified onto a cushion of 5.7 M cesium chloride by ultracentrifugation at 82,705*g* for 20 h at 20°C. RNA pellets were then resuspended in sterile water and quantified.

Reverse transcriptase PCR

One microgram of total RNA from each zygotic and somatic embryo preparation was placed with 1 µl oligo$(dT)_{12}$VN (10 µM) and the volume adjusted to 11 µl with sterile water. The mixture was incubated for 5 min at 70°C then chilled on ice. Four microlitres of 5× reaction buffer, 2 µl 10 mM dNTP and 20 units ribonuclease inhibitor were prepared and the volume was adjusted to 19 µl with water, and incubated for 5 min at 37°C. Two hundred units of RevertAid™ M-MuLV Reverse Transcriptase (Fermentas, St Leon-Rot, Germany) were finally added and the samples were incubated for 60 min at 42°C. The reactions were stopped by incubation for 10 min at 70°C. Remaining RNAs were degrade by adding 1 µl EDTA, 0.5 M, pH8 and 1 µl DNAase-free pancreatic RNAase (5 µg/ml) and incubation at 37°C for 30 min. Reverse-transcribed first strand cDNA was precipitated in EtOH and resuspended in 150 µl TE buffer.

TcL1L cDNA isolation and sequencing

Degenerate oligonucleotide primers were designed from the best conserved regions among the *Arabidopsis thaliana* LEC1 proteins (Stone et al. 2001; Lee et al. 2003): D2S1 (5′-CARGARTGYGTNWSNGARTWYATHWS-3′), and D4R1 (5′-CCNARYTTNSWCATNGCCCA-3′).

A PCR reaction was performed with 2 µl of single strand cDNA. The PCR product was cloned in a pGEM-T Easy Vector Systems I (Promega, Madison, WI, USA) and sequenced. TcL1L full length cDNA was obtained by rapid amplification of cDNA ends (RACE; BD Biosciences Clontech, San José, CA, USA). The 5′ end of the TcL1L cDNA was amplified using cDNA from zygotic embryos of clone IMC67×DSH30 at various stages of development with primer (5′-CDS primer A: 5′-$(T)_{25}N_{-1}$N-3′, N = A, C, G or T; N_{-1} = A, G or C) and BD SMART II™ A Oligonucleotide (5′-AAGCAGTGGTATCAACGCAGAGTACGC GGG-3′) provided in the kit. Two 5′-end sequences were obtained with the TcLec140T7-SPb-R (5′-TTGCGCT GCTCACGTTGGCAGCG-3′) /UPM primer pair.

After analysis of the sequences of 3′ and 5′ RACE-PCR products, primers TcLec140FL-S (5′-ACGCGGGGACC CTCTTAAT-3′) and TcLec140FL-R (5′-TTTAGATGGT AATAAGGCT-3′) were designed to amplify a full-length cDNA. The full-length cDNA was purified and cloned into the pGEM-T vector followed by sequencing using both the T7 and SP6 primers.

Sequence analysis

Sequences were analysed using bioinformatics software available on NCBI databases such as Basic Local Alignment Search Tool (BLAST; Altschul et al. 1997).

Southern-blot analysis

Genomic DNA was isolated from *T. cacao* leaves as described by Risterucci et al. (2000). Given that one *Hin*dIII site exists in the cDNA sequence, 2 µg of *T. cacao* DNA (4 million copies of genomic DNA) were digested for 24 h at 37°C with *Hin*dIII. Digested DNA was fractionated by electrophoresis on a 0.8% agarose gel, transferred to a Hybond-N^+ filter (Amersham, Buckinghamshire, UK) and hybridized with a [^{32}P]-labelled 780 bp cDNA *Tc-LEC1-like Hin*dIII fragment. Hybridization was performed at 65°C overnight in 50 mM Tris-HCl buffer (pH 8), 5× SSC (pH 7.4), 10 mM EDTA (pH 8), 10% dextran sulphate, 1% Denhardt reagent, 0.2% SDS and denatured salmon sperm DNA. Membranes were washed twice, 1 and 30 min, respectively, at 65°C in 2× SSC, 0.5% SDS, and then 45 min at 65°C in 1× SSC, 0.5% SDS.

Analysis of *TcL1L* expression by RT-PCR analysis

The semi-quantitative RT-PCR analyses were carried out as described by Zegzouti et al. (1999). The *TcLec1L* cDNA was amplified using primers TcLec1L-RT-S: (5′-CGAGCC AGAAAGAGATGGAA-3′) and TcLec1L-RT-R: (5′-TGA TGACCCTGCATTGGATG-3′), generating a 600 bp fragment. Primers were designed from a *T. cacao* actin gene sequence (C. Lanaud, CIRAD, pers. communication): TcActin-KZ-S: (5′-GCGTCGTCGTTTTCTGCTTC-3′) and TcActin-KZ-R: (5′-CAGGAGCAACACGGAGT TCA-3′). A 380 bp *TcActin* cDNA was amplified concomitantly with *TcLec1L* cDNA and used as the internal standard. To obtain the same actin intensity signal, *TcActin* primers were added to the PCR reaction 3 or 5 cycles after the beginning of *TcLec1L* clone amplification. Samples were subjected to 30 or 32 cycles. In addition, PCR amplification was performed with total RNA extracts without reverse transcription as a negative control.

Localization of *TcL1L* expression by mRNA in situ hybridization

Probe synthesis

Specific primers were designed for PCR amplification of the specific region of 18S ribosomal RNA (RibHis-P-Up: 5′-CCGACCCTGATCTTCTGTGAAGGG-3′, RibHis-P-Down: 5′-CCAAGTCAGACGAACGATTTGCACG-3′

for 18S ribosomal RNA probe) and *Lec1-like* cDNAs (TcLec1-lHis-P-Up: 5′-GATGGTAATAAGGCTAAG-3′, TcLec1-lHis-P-Down: 5′-GAACGTGGTTCGATCAGA-3′ for Lec1-L probe). The same primers including RNA Polymerase T7 promoter sequence 5′-GCGAAATTAATACG ACTCACTATAGGGAGA-3′ were also designed: namely RibT7His-P-Up: 5′-CCGACCCTGATCTTCTGTGAAG GGGCGAAATTAATACGACTCACTATAGGGAGA-3′, RibT7His-P-Down: 5′-GCGAAATTAATACGACTCAC TATAGGGAGACCAAGTCAGACGAACGATTTGCAC G-3′; TcLec1-lT7His-P-Up: 5′-GCGAAATTAATACGA CTCACTATAGGGAGAGATGGTAATAAGGCTAAG-3′, TcLec1-lT7His-P-Down: 5′-GCGAAATTAATACGACT CACTATAGGGAGAGAACGTGGTTCGATCAGA-3′. Finally, one primer corresponding to the T7 end was also designed: E-T7His: 5′-GCGAAATTAATACGACTCAC-3′. Sense and reverse probes were synthesized in two steps: an initial PCR was performed with one primer P-Up or one primer T7-P-Down, and one primer P-Down and one primer T7-P-Up. PCR was carried out with 2 ng of DNA solution, which was supplemented with 1× reaction buffer, 1.5 mM $MgCl_2$, 10 mM each of dATP, dCTP, dGTP and dTTP, 0.4 µM each of 5′ and 3′ oligonucleotide primers and 5 U Taq polymerase, and then adjusted to a final volume of 50 µl. Samples were subjected to 1 cycle of 3 min at 95°C, 35 cycles (95°C, 30 s; 55°C, 30 s; 72°C, 1 min) and 1 cycle of 7 min at 72°C. Amplification products were fractionated by 1% agarose gel electrophoresis. A second PCR was performed using "100 times" dilutions of the first PCR products, E-T7 primer and primer P-Up or P-Down, to obtain the desired probe: sense or reverse. PCR was carried out with 1 µl of DNA dilution, which was supplemented with 1× reaction buffer, 1.5 mM $MgCl_2$, 10 mM each of dATP, dCTP, dGTP and dTTP, 0.4 µM each of 5′ and 3′ oligonucleotide primers and 5 U of Taq polymerase, and then adjusted to final volume of 50 µl. Samples were subjected to 1 cycle of 3 min at 95°C, 35 cycles (95°C, 30 s; 56°C, 30 s; 72°C, 1 min) and 1 cycle of 7 min at 72°C. Amplification products were fractionated by 1% agarose gel electrophoresis. After purification in ethanol, the fragments were used to generate sense and reverse digoxigenin-labelled transcripts by in vitro transcription (T7 MAXIScript Kit, Ambion, Austin, TX, USA). Final purification of the probes was done using ethanol precipitation. The final lengths of the different probes were: ribosome probes (150 bp); lec1-like probes (400 bp).

Zygotic and somatic embryogenesis explants were fixed in paraformaldehyde 4%, potassium phosphate buffer (10 mM, pH 7.2) overnight at 4°C, dehydrated in ethanol series and butanol, and embedded in Paraplast X-TRA (Labonord, Templemars, France). Cross sections (8 µm) were placed on silanized slides (Dako, Trappes, France). Nucleic probes for *TcLEC1*were specifically designed to reach a length >300 bp. Probes for ribosomal 18S were used as controls. Sense and antisense probes were labelled with UTP-digoxigenin during transcription. Labelling efficiency was estimated by dot blotting on membrane with a phosphatase alkaline conjugate antibody raised against digoxigenin (Roche Diagnostics, Meylan, France). Slides were dewaxed, rehydrated, and then incubated for 40 min in proteinase K (2 µg ml^{-1}) at 37°C. A probe concentration of 2 µg/ml was used in an overnight hybridization of tissue sections at 45°C in 10% dextran sulfate, 1× Denhart solution, 50% formamide, 1 µg $µl^{-1}$ tRNA, and 2× SSC buffer. Post hybridization washing at graded stringency (2× SSC at 50°C and 1× SSC at 55°C) was followed by detection. Slides were first pretreated with a blocking solution (Roche Diagnostics) to prevent nonspecific binding of antibodies then incubated with antidigoxigenin mouse antibody (Roche Diagnostics). Two developing system were used a Vector Blue Development and a Alexa488 Fluorescent Development. Both, immunochemical and fluorescent detection of transcripts were used as complementary methods. For the Vector blue development, slides were incubated with antibody [1:500 dilution of anti-DIG combined with alkaline phosphatase (Roche, Basel, Switzerland) for 1 h at 30°C] and washed (PBS for three times 10 min at 30°C, colour development buffer 100 mM Tris-HCl pH 8.2 for twice 5 min at 30°C). Hybridization signals were detected with Vector® Blue Alkaline Phosphatase Substrate Kit III (Vector Laboratories, Burlingame, CA, USA). Slides were observed with a DMRXA microscope (Leica, Wetzlar, Germany) under white light. For the Alexa488 Fluorescent Development, slides were incubated with a mouse anti-DIG antibody and then with an anti-DIG FluoProbes 488 antibody IgG (rom Interchim) at room temperature. Sections were mounted in Mowiol. Fluorescence was detected using a confocal laser-scanning microscope (argon 488-nm laser; Zeiss 510 Meta). Specificity of labelling was checked by spectral analysis.

Transformation of Arabidopsis *lec1* mutant

The full-length *TcL1L* cDNA inserted into the pGEM-T vector was digested by *Sal*I and *Pvu*II and introduced into the pENTR2B gateway (Invitrogen, Carlsbad, CA, USA) vector at the *Sal*I-*EcoR*V sites by ligation. The TcL1Lp ENTR2B plasmid was recombined into the pH7WG2D vector allowing expression of *TcL1L* cDNA under the control of the 35SCaMV constitutive promoter (Karimi et al. 2002). The plasmid was transferred into *Agrobacterium tumefaciens* strain C58C1. Heterozygous *lec1* plants were transformed by the floral dip protocol of Clough and Bent (1998). Transformants were selected on kanamycin for the presence of the initial T-DNA that tags the *lec1* mutation and on hygromycin for the 35SCaMV:TcL1L construct. The genotype of the

transformants was determined by PCR. Use of the lec1-2F: (5′-GGTAGCTGATAGTGAGAGAAGGG-3′) and lec1-2F: (5′-AGGTCATGGACTGGGCCGTAAAC-3′) primers that amplify a fragment of 858 bp for the wild type copy of *LEC1*, discriminated between *lec1/LEC1* or *lec1/lec1* plants. The primers TcLec140B1RT-S: (5′-CGAGCCAG AAAGAGATGGAA-3′) and TcLec140B1T-R (5′-TG ATGACCCTGCATTGGATG-3′) were used to verify that the cocoa gene was present. Siliques of the primary transformants were opened under the binocular microscope to score the seeds presenting a characteristic *lec1* phenotype. To analyse the embryo phenotype in the *lec1-2* mutant line and the primary *lec1-2* transformants, siliques at different stages of maturity were opened under the binocular microscope in order to remove the immature seeds. The seeds were cleared in Hoyer's solution (2.5 g gum arabic, 100 g chloralhydrate, 5 ml glycerol in 30 ml of H_2O) for 1 h to overnight and observed using Nomarski optics on a Zeiss Axioskop2 microscope (Carl Zeiss, Jena, Germany). The T2 seeds of individual transformants were sown on kanamycin only to verify the presence of the *lec1-2* mutated allele and on hygromycin only to estimate the number of copies of the TcL1L T-DNA.

Results

Isolation and sequence analysis of a homologue of *At leafy cotyledon1-like* full length cDNA from *T. cacao* L.

A full length cDNA sequence homologous to the *A. thaliana leafy cotyledon1-like* gene was isolated in *T. cacao* by a PCR-based approach (GenBank accession number DN237955). We named it *T. cacao leafy cotyledon1-like* (*TcL1L*). This partial cDNA sequence of 491 bp was isolated from first-strand cDNA derived from a mixture of *T. cacao* zygotic embryos at various stages. Amplification products of 5′ RACE (579 bp) and 3′ RACE (484 bp) were then only obtained from first-strand cDNA synthesized from immature zygotic embryos only. By sequence analysis of RACE products, specific primers at the 5′ and 3′ ends were designed and led to the isolation of a 917 bp full length cDNA sequence. The full length cDNA contained a 639 bp open reading frame (ORF), 108 nucleotides of a 5′-untranslated region (UTR), and 170 of 3′-UTR. The ORF encoded a putative peptide of 213 amino acids accounting for a theoretical pI of 8.23 and a calculated molecular mass of 24.01 kDa. The translated *TcL1L* nucleotide sequence shared a high degree of identity, 76–78%, with *Phaseolus coccineus*, *Helianthus annuus* and *A. thaliana* LEC1L (GenBank accession numbers AAN 01148, CAI 48078, NP_199578, respectively). *Tc*L1L cDNA encodes a putative transcription factor homologue, the CCAAT box-binding factor (CBF, also known as NF-Y) HAP3 (heme-activated protein) subunit. The HAP3 subunit consists of three regions, an amino-terminal A domain, a central B domain, and a carboxyl-terminal C domain (Lotan et al. 1998). HAP3 subunits are well recognized by their central B domain, a = 95 amino acid region of the protein that is conserved across eukaryotic organisms (Fig. 1). The translated amino acid sequence of *Tc*L1L shares all the characteristic features of the HAP3 subunit with high identity (>94%) and high similarity (>98%) in the B domain of all known L1L. Sixteen residues, indicated in red in Fig. 1, were conserved between *Tc*L1L and the LEC1-type HAP3 subunits, including *At*L1L. The B domain possesses amino acid residues required for interaction of HAP3 with other CBF subunits and for the DNA-binding activity of the CBF complex, thus, sharing sequence identity with the histone like-transcription factor domain H2B, a structurally conserved region consisting of four α helices (α1, α2, α3 and αC) connected by two short loops (Fig. 1, L1, L2 and L3). In this domain, Lee et al. (2003) showed that Asp (D) residue at position 55 in *Arabidopsis* plays an important role in the LEC function. Southern hybridization of *Hin*dIII digested cocoa genomic DNA with *TcL1L* cDNA revealed the presence of a single gene in the *T. cacao* genome (Fig. 2). Two additional slight bands may reflect partial hybridization with genes sharing homology with *TcL1L*, such as *LEC1* or *LEC2*.

Tc leafy cotyledon1-like is expressed in early zygotic embryogenesis and at different stages of somatic embryogenesis

TcL1L was strongly expressed in both immature zygotic and immature somatic embryos (Fig. 3). When RT-PCR cycles were increased from 30–32, very low expression was also observed in roots and cotyledons which lost capacity to regenerate secondary somatic embryos (data not shown). With regard to the somatic embryogenesis process, *TcL1L* expression was recorded in fragments of cotyledon from immature somatic embryos (SSEE d0), in the same fragments bearing secondary embryos (SSEE + VYSE d28-42) and in both immature (YSE d72) and developed regenerated embryos (SE d72). Interestingly, at the beginning of secondary embryogenesis, when cotyledon fragments were dissected, the initial low *TcL1L* expression decreased within 7 days (SSEE d7, SSEE d14, SSEE d21). Expression in fragments of cotyledon bearing embryos was probably related to new embryo development since non-embryogenic cotyledon fragments (NEE d42-56) did not show any expression, as did non-embryogenic primary calli (NEPC). To conclude, *TcL1L* was preferentially and highly expressed in somatic and zygotic embryo tissues, even

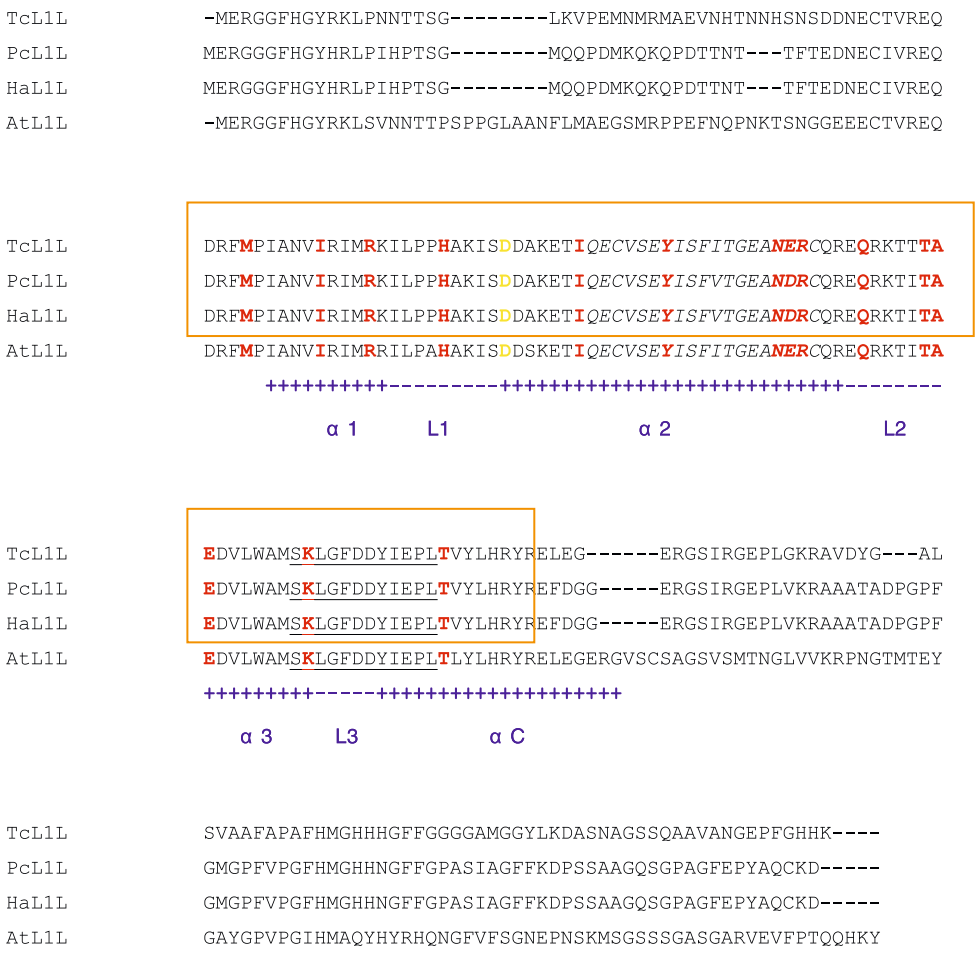

Fig. 1 Comparison of the deduced *Theobroma cacao* (*Tc*) leafy cotyledon1-like protein (*Tc*L1L) and L1L from *Phaseolus coccineus* (*Pc*L1L), *Helianthus annuus* (*Ha*L1L), and *Arabidopsis thaliana* (*At*L1L). The B domain of the HAP3 subunit of the CCAAT binding factor is framed in orange. The histone fold motif (HBF) is in *italics*. The *underlining* indicates the consensus sequence which interacts with the TATA-binding protein. Residues unique to LEC1-type HAP3 subfamilies proteins are highlighted in red. The positions of α-helices and loops in the histone fold motif are indicated as *blue crosses* and *dashed lines*, respectively. The Asp (D) residue which is critically required for the LEC function is highlighted in *green* (Lee et al. 2003)

though its expression was higher during zygotic embryogenesis.

Tc leafy cotyledon1-like expression is mainly localized in embryonic and meristematic cells of young zygotic and somatic embryos

To determine the localization of *Tc*L1L RNAs in zygotic and somatic embryos at different stages, we designed in situ hybridization experiments (Figs. 4, 5). For each experiment, different controls were used (data not shown): one positive (reverse ribosome probe to verify the conservation of the RNAs) and three negative controls (sense ribosome probe; no probe to check the antibody link specificity; sense L1L was used to check the specificity of the reverse L1L link). Sense probes did not bind with the sections, showing the specificity of the hybridization reactions (Fig. 4b, d). Equivalent expression results were found either by using Vector Blue or fluorescent detection of the RNAs. During zygotic embryogenesis, *L1L* RNA was detected in all the tissues and cells of the globular embryos (Stage Iz; Fig. 4a, c) and young cotyledonary stages (Stage IIz; Fig. 4e), and especially in the young epidermis, the protoderm. At stage IIIz, during the embryo growth phase where cell divisions and elongations took place, the gene was expressed everywhere but more in the root and shoot apex, in meristematic cells (Fig. 4f). At stage V, when the embryo had reached its final length, but was still immature (no reserves inside the cells) expression drastically decreased and was restricted to the meristematic cells of the shoot (Fig. 4g, h) and root apex (Fig. 4i, j). This expression pattern corresponded with the results from RT-PCR analyses. During secondary somatic embryogenesis, at day 0, very slight expression was detected in cotyledonary explants excepted in the epidermis (Fig. 5a). Expression then disappeared in the explant, while oxidation increased up to the end of the experiment (Fig. 5b). *L1L* transcripts were detected again when young somatic embryos arose from the explant. Here again, the signal was strongly observed in the protoderm of young somatic embryos. As was found for zygotic embryos, expression was located in all the cells of young somatic embryos (Fig. 5c, d, e, f). Compared to the expression observed in zygotic embryos at equivalent stages, the intensity was lower in somatic embryos. These results confirmed what was previously assumed from RT-PCR analyses. The variation in expression during somatic

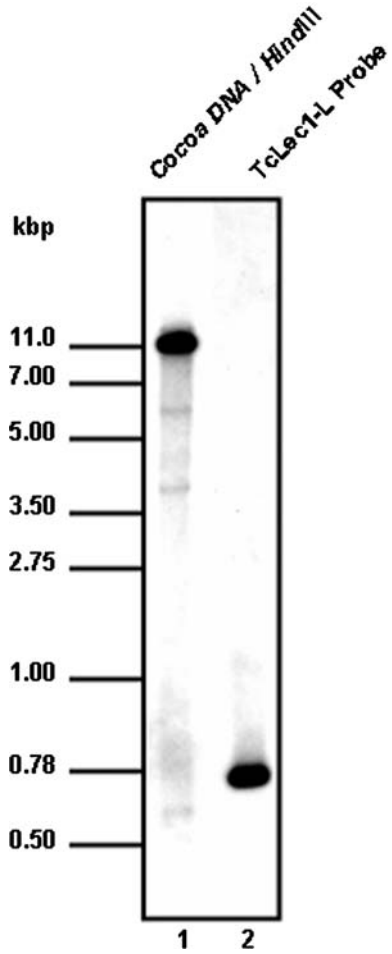

Fig. 2 Southern-blot analysis of *TcL1L* was carried out on cocoa genomic DNA (*lane 1*) and probe fragment as positive control (*lane 2*). Genomic DNA was digested with *Hin*dIII and fractionated by gel electrophoresis. Fragments were hybridized with a 770-bp radiolabelled *Eco*RI–*Hin*dIII fragment from *TcL1L* cDNA

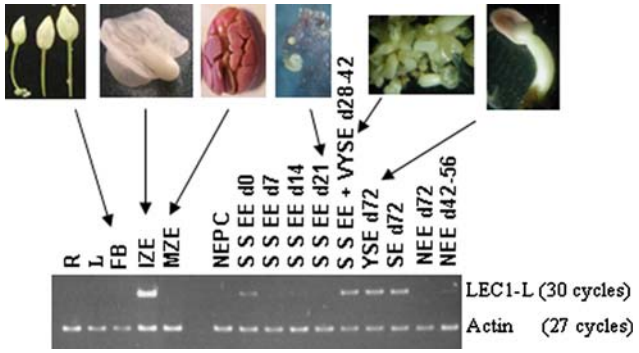

Fig. 3 Analysis of the expression of *TcL1L* gene by semi-quantitative RT-PCR in various organs: roots (*R*); leaves (*L*); flower buds (*FB*); immature zygotic embryos (*IZE*); mature zygotic embryos (*MZE*); non embryogenic primary callus (*NEPC*); secondary somatic embryogenic explants (*SSEE*) day 0, day 7, day 14, day 21; secondary somatic embryogenic explants (*SSEE*) bearing very young somatic embryos (*VYSE*) day 28–42; young somatic embryos (*YSE*) day 72; non embryogenic explants (*NEE*) day 72 and day 42–56. *TcL1L*: 30 amplification cycles *Tc-Actin*: 27 amplification cycles

embryo growth was equivalent to what was observed for zygotic embryos. In fully grown somatic embryos, *L1L* expression was restricted to the shoot (Fig. 5g) and root meristems (Fig. 5h). In non-embryogenic primary callus obtained from staminodes, expression was detected is some rare zones of internal meristematic tissue (Fig. 5i) even though no detection was recorded by RT-PCR analyses. Expression was so low that it could hardly be detected with fluorescent probes (Fig. 5j).

Over-expression of *Tc leafy cotyledon1-like* gene in *A. thaliana* can rescue some aspects of Arabidopsis *lec1* defects

To assess the functionality of the protein encoded by *TcL1L*, we used the mutant collection of *Arabidopsis*. In the absence of a well-characterized *l1l* mutant, we tested the ability of *TcL1L* to suppress the *lec1* mutation. Kwong et al. (2003) had demonstrated that ectopically expressed *Arabidopsis L1L* could restore some LEC1 functions such as desiccation intolerance. However, postembryonic expression of *L1L* conferred embryonic characteristics to seedlings, which were mostly unable to develop further. Heterozygous *lec1-2* plants were obtained by *Agrobacterium*-mediated transformation harbouring the 35SCaMV::*TcL1L* construct. Thirty-two primary transformants were analysed after antibiotic selection and compared to untransformed *lec1/LEC1* plants. The *lec1* mutant line was maintained as heterozygous since homozygous *lec1* embryos are intolerant of desiccation and therefore, unviable as dry seeds. The siliques of *lec1/LEC1* plants contained 25% *lec1* seeds, which could be easily distinguished from phenotypically wild type seeds. Mutant embryos often displayed upright cotyledons with anlages of trichomes, a reduced hypocotyle and large rounded cotyledons (Fig. 6a) and were often viviparous (Fig. 6d). The genotype of the *lec1/LEC1* transformed plants was assessed by PCR. All of the plants were heterozygous for the *lec1* mutation and contained the cocoa *L1L* transgene. In most primary transformants, the ratio of seeds exhibiting the *lec1* phenotype differed significantly from untransformed *lec1/LEC1* plants. In 11 plants, all the seeds displayed a wild phenotype or presented a very similar wild-type seed phenotype. In 3 lines, the siliques contained only 1/16 of *lec1* seeds, a result consistent with complementation by a single locus of the *35SCaMV:TcL1L* transgene. Collectively, these data demonstrated that the *TcL1L* gene could complement the morphological defects of the *lec1* embryos. In ten lines, the number of seeds exhibiting a *lec1* phenotype or some aspect of the *lec1* phenotype was much higher than in untransformed *lec1/LEC1* plants. The percentage of *lec1* seed phenotype ranged between 40 and 100% depending on

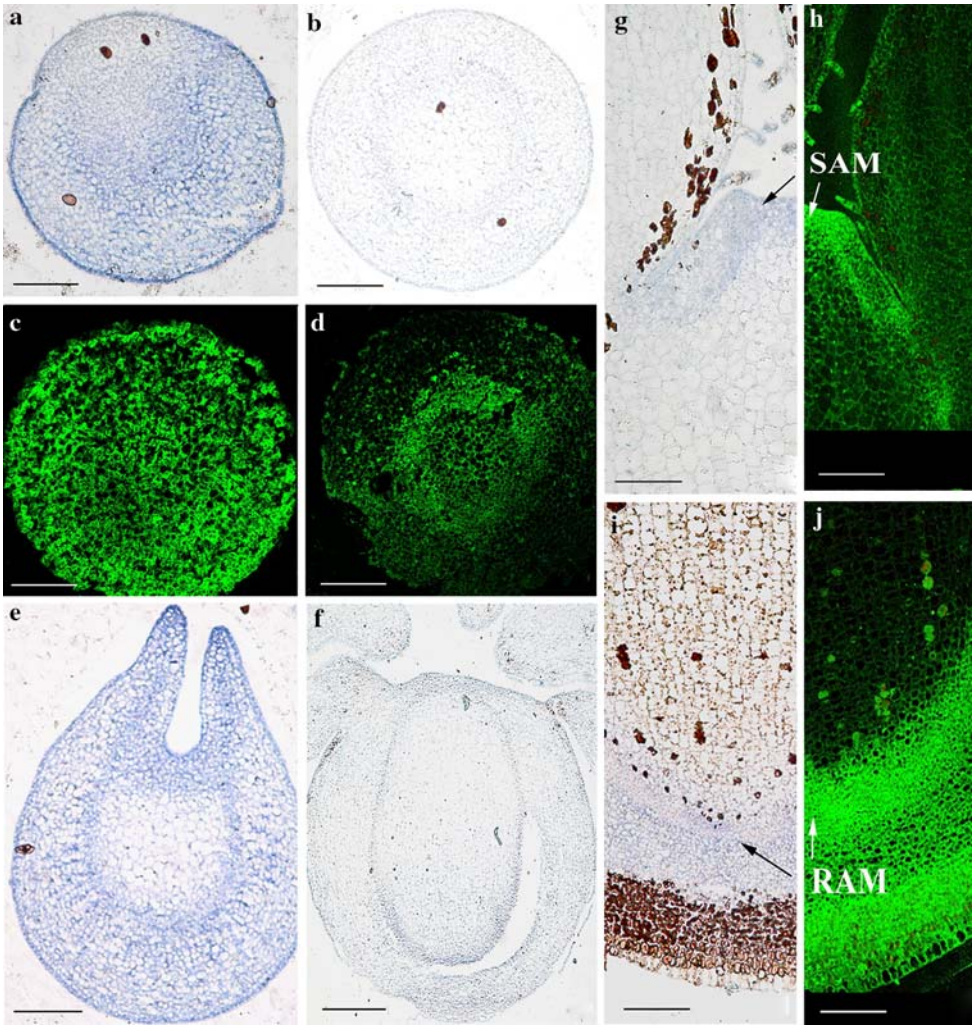

Fig. 4 Localization of *TcL1L* transcripts by in situ hybridization during zygotic embryo development. The sections were probed using either a Vector blue (**a**, **e**, **f**, **g**, **i**) or fluorescent RNA antisense probes (**c**, **h**, **j**). Sense-labelled *TcL1L* RNA (**b**, **d**) was used as a negative control. **a**, **c** Globular zygotic embryo (Iz: 3–5 days after pollination) showing a high specific signal with the antisense *TcL1L* probe. On the slide representing the signal obtained with the fluorescent probe (**c**), the signal appears inside the cell cytoplasm. **e** Cotyledonary zygotic embryo (IIz: 28 days after pollination) showing a specific signal in all its cells with the antisense *TcL1L* probe. **f** Cotyledonary zygotic embryo (IIIz: 80 days after pollination) showing a low specific signal in all its cells with the antisense TcL1L probe. **g**, **h**, **i**, **j** Cotyledonary zygotic embryo (Vz: 110 days after pollination) showing a specific signal restricted to some cells of the shoot apical meristem (**g**, **h**) and to some cells of the root meristem (**i**, **j**) with the antisense TcL1L probe. No significant signal was detected with the *TcL1L* sense probe, results shown for Iz Globular zygotic embryos (**b**, **d**). SAM shoot apical meristem; RAM root apical meristem. *Scale bars* 250 μm, except 500 μm (**f**)

the line. In those plants, 40% of the seed defects were mild, such as twisted seeds (Fig. 6c) or a partially developed root apex with normal cotyledons (Fig. 6b), while those defects only amounted to 11% of the seed phenotype in non-transformed *lec1/LEC1* plants. This dominant effect of the ectopic expression of a *L1L* transgene had already been observed with the *Arabidopsis L1L* gene on germination (Kwong et al. 2003). Finally, eight plants had the same ratio of *lec1* to wild type seeds as untransformed *lec1/LEC1* plants, suggesting that the transgene was inactive in those plants. The T2 seeds of each transformant were analysed on germination. All the lines produced T2 seeds that were, for most of them, able to germinate on hygromycin (selection for the *TcL1L* transgene) and seemingly develop normally. Several minor defects were observed on the cotyledons. The cotyledons were either misshapen or supernumerary (data not shown). A closer examination of the T2 seedlings showed that the leaves were slightly abnormally shaped and exhibited cotyledonary traits. An example is presented in Fig. 6h–k. The first leaf (Fig. 6i) had a shape and thickness similar to a cotyledon (Fig. 6j) and produced a smaller number of trichomes. The second leaf (Fig. 6k) had a relatively normal morphology though contained less chlorophyll. In some instances, we also noticed that a few seeds did not fully germinate and exhibited morphological defects (Fig. 6g–i) similar to those observed by (Kwong

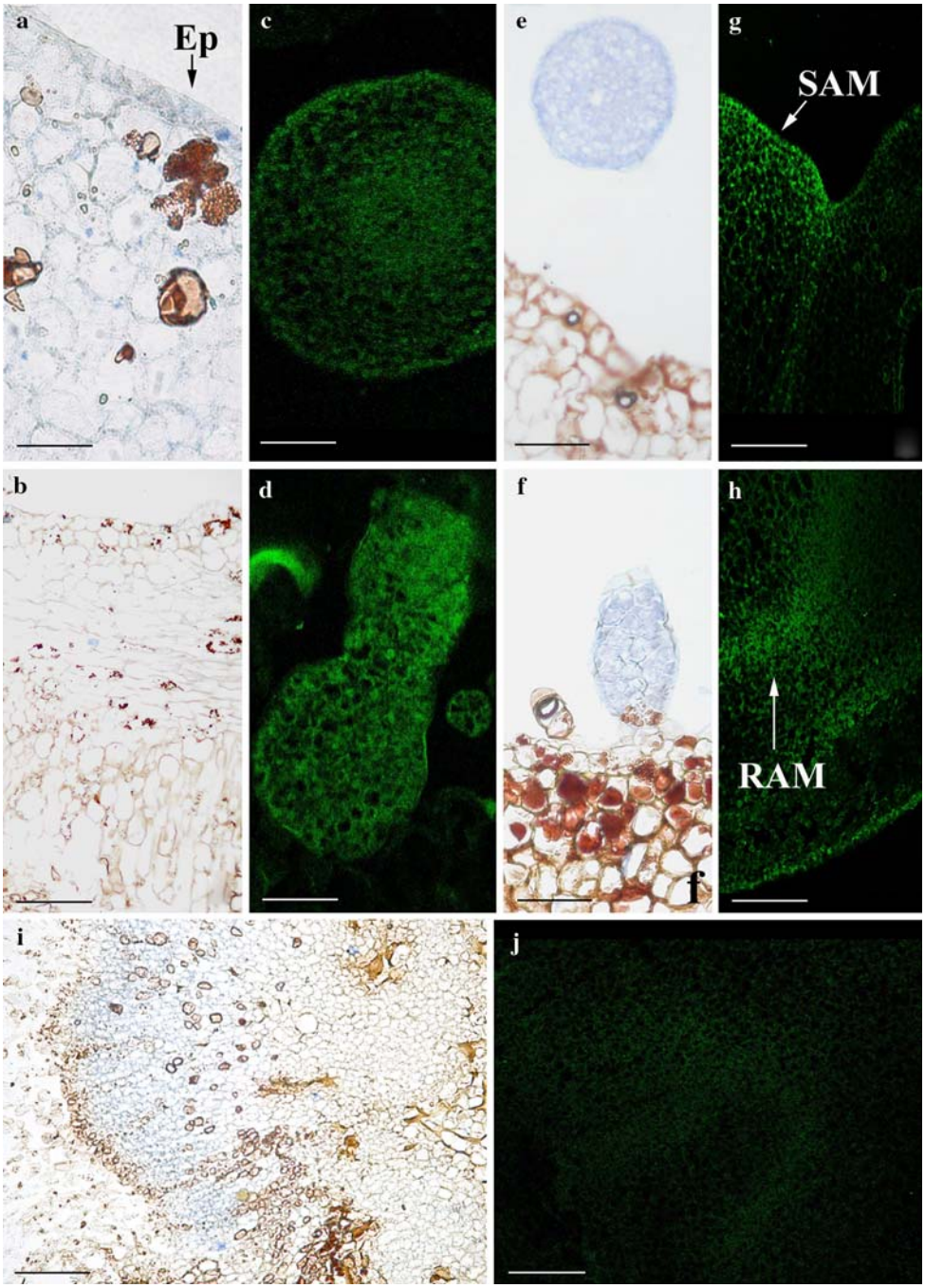

Fig. 5 Localization of *TcL1L* transcripts by in situ hybridization during secondary somatic embryogenesis from cotyledon sections. The slides were probed using either a Vector blue [(**a**, **b**, **e**, **f**, **i**) or fluorescent RNA antisense probes (**c**, **d**, **g**, **h**, **j**)]. **a** A low signal was detected in the cells of a cotyledon section of a somatic embryo, except in the epidermis, with the antisense *TcL1L* probe. **b** After 7 days of culture, the expression of *TcL1L* completely disappeared. **c**, **d**, **e**, **f** *TcL1L* expression was observed again in very young globular somatic embryos. **g**, **h** In older fully grown secondary somatic embryos, the expression of *TcL1L* was, as in zygotic embryos, restricted to the shoot (**g**) and root apex (**h**). **i**, **j** Very low expression of *TcL1L* observed in meristematic zones of non embryogenic UF667 callus. Ep epidermis; SAM shoot apical meristem; RAM root apical meristem. *Scale bars* 30 μm (**a**), 250 μm (**b**, **g**, **h**, **i**, **j**), and 60 μm (**c**, **d**, **e**, **f**)

et al. 2003) with *L1L* or with *LEC1* (Lotan et al. 1998). Comparable results were obtained on a non-selective medium, confirming that the inability to germinate was not due to the presence of the antibiotic. The root pole did not develop (Fig. 6e) or its growth was limited (Fig. 6f). The cotyledons did not expand (Fig. 6e–g) and the new leaves were reduced and abnormally shaped (Fig. 6f, g).

Taken together, our results demonstrated that *TcL1L* shares some of the *AtLEC1* and *AtL1L* functions. Expression of *TcL1L* during embryogenesis was able to rescue the morphological defects of *lec1* seeds. However, *TcL1L* was not as efficient as the *Arabidopsis LEC1* and *L1L* genes in conferring embryonic characteristics to seedlings. As for *AtL1L*, ectopic expression of *TcL1L* did not result in the production somatic embryos on seedlings.

Discussion

During embryogenesis, the basic patterns of plant structures are established and the functional embryo domains are determined. To understand the molecular mechanism

Fig. 6 Phenotypes of seeds and seedlings of Arabidopsis *lec1* plants expressing the *TcL1L* gene. **a–d** images of cleared seeds. **a, d**, Typical images of upright *lec1* embryos with anlages of trichomes on the cotyledons (**a**) and viviparous root pole (**d**). **b**, **c** Seeds of *lec1* plants containing the 35SCaMV:*TcL1L* transgene showing an embryo with fairly normal morphology (**b**) or folding defects (**c**). **e–h**, Four-day-old seedlings containing the 35SCaMV:*TcL1L* construct and exhibiting various morphological defects. **e**, **f** Seedlings arrested soon after germination. **g** Close-up view of seedling in **f** showing the lack of cotyledon expansion (C), a cotyledon-like organ (Cl) at the shoot apex and the dual embryonic and vegetative morphologies of the hypocotyle (the *white arrow* indicates the limit between the two regions). **h** Minor defects visible on the leaves of most 35SCaMV:*TcL1L* seedlings. **i–k** Close-up view of the cotyledon (**i**), the first (**j**) and second leaf (**k**). Trichomes are marked by an *asterisk* in **a** and **b**. *Bar* 50 μm

underlying embryogenesis, isolation of the relevant genes is essential. Apart from its agricultural and economic importance, the study of cocoa embryogenesis can result in advances in our understanding of cell differentiation, growth and development in a recalcitrant tropical plant. Molecular characterization of *Arabidopsis* mutants has highlighted several genes that regulate embryogenesis (review in Feher et al. 2003). Application of *Arabidopsis* molecular knowledge to crop species is becoming a challenge.

In this paper, we report on the characterization of *TcL1L*, a homologue of *AtL1L*, that encodes a 213 amino acid polypeptide with sequence similarity to the HAP3 subunit of the CCAAT binding transcription factor, also called CBF or NF-Y (Lotan et al. 1998). The *TcL1L* conceptual translation product showed sequence similarity with the putative L1L protein isolated from *Phaseolus coccineus* (Kwong et al. 2003), *Helianthus annuus* (Fambrini et al. 2006) with 68% amino acid identity; and from *Arabidopsis* with 60% amino acid identity. The central B domain of *Tc*L1L (amino acid 49–138) displayed substantial sequence identity (>94%) with the central B domain in the L1L and HAP3 subunits of the CCAAT-binding transcription factors in other organisms. The sequence similarity was mainly found in the B domain. *Tc*L1L did not show large stretches of sequence similarity with other LEC proteins outside of the B domain binding region. The B domain is an ≈95-amino acid residue region found to be required for the interaction of yeast and mammalian HAP3 subunits with the other CBF subunits, HAP2 and HAP5, and for the DNA binding activity of the CBF complex (Xing et al. 1993; Sinha et al. 1996; Li et al. 1998).The B domain of HAP3 subunits shares sequence identity with the histone fold motif of histone H2B, a structurally conserved region consisting of three α helices (α_1, α_2, and α_3) connected by two short loops (L1 and L2; Sinha et al. 1996; Bolognese et al. 2000). The histone fold motif mediates protein–protein and DNA binding interactions (Zemzoumi et al. 1999). These results strongly

suggest that *Tc*L1L is a member of the CBF class of eukaryotic transcriptional activators. *Tc*L1L belongs to the group containing the LEC1-type AtHAP3 subunits, *Arabidopsis* LEC1 and L1L. The B domains of the LEC1-type AtHAP3 subunits have 16 conserved amino acids that are not conserved in the non-LEC1-type AtHAP3 subunits (Kwong et al. 2003). Some or all of these residues are conserved among most HAP3 subunits from other plants that contain the LEC1-type B domain and are primarily expressed in embryos or seeds. These residues were well conserved in *Tc*L1L. Moreover, the Asp (D) residue at position 55 in *Arabidopsis LEC* gene, which plays an important role in *LEC* function, was also recognized in *Tc*L1L. It is assumed that Asp-55 in *Arabidopsis LEC* B domain mediates specific interactions with other CBF subunits or with other transcription factors (Sinha et al. 1996), and thus regulates genes required for embryo development (Lee et al. 2003).

Due to its potential role in embryogenesis, we analysed *TcL1L* gene expression in both zygotic and somatic processes. High *TcL1L* mRNA levels were detected in young and developing zygotic and somatic embryos. The pattern of expression in zygotic embryos was closer to that observed in sunflower (Fambrini et al. 2006), rather than that observed in *A. thaliana* (Kwong et al. 2003). Its expression was restricted to young and immature zygotic embryos, and no expression was detected in mature embryos. This was different from what was observed in *Arabidopsis* where *L1L* was also expressed during maturation and in mature zygotic embryos (Kwong et al. 2003). A common feature between the *L1L* expression of *A. thaliana, P. coccineus, H. annuus* and *T. cacao* is its enhancement in the epidermal layer of embryos, especially at the globular stage and throughout embryogenesis. The observation that *TcL1L* expression was detected before inside/outside patterning was initiated and was then strongly maintained in the protoderm confirmed the role of LEC genes in coordinating primary events leading to embryonic competence (Harada 2001; Kwong et al. 2003) and their involvement in establishing a cellular environment that promotes embryo development (Lotan et al. 1998). Another important common feature between *AtL1L, PcL1L, HaL1L* and *Tc1L* is the expression at very low levels, in vegetative organs. This point differs from the *LEC1* gene that is only expressed in embryo tissues (Kwong et al. 2003). We showed *L1L* expression throughout the different steps of the cocoa somatic embryogenesis process. This expression was very similar to what was observed during zygotic embryogenesis: *TcL1L* was mainly expressed in young and immature somatic embryos, at the same location as that observed in zygotic embryos. However, even if true quantification of expression was not achieved, it is clear that it was lower in somatic embryos than in zygotic embryos. Our observations confirm the results of Fambrini et al. (2006), who found *HaL1L* expression in *Helianthus* somatic embryos and in ectopic embryos of a somaclonal variant. They also showed expression in cells at the origin of ectopic embryos, the parenchymal cells around the vascular bundles of epiphyllous leaves, suggesting that *HaL1L* could be involved in switching somatic cell fate towards embryogenic competence.

To assess the functionality of the cocoa *L1L* gene, we used the mutant facilities of the *Arabidopsis* model plant. *Lec1-2* mutant lines ectopically expressing the *TcL1L* gene were produced. Since uncertainties about the ability of *TcL1L* to fully complement the *Arabidopsis* gene existed, heterozygous *lec1/LEC1* plants were transformed in order to circumvent the intolerance of desiccation in *lec1* seeds. Kwong et al. (2003) showed by transforming homozygous *lec1* plants, that expression of the *Arabidopsis LEC1* or *L1L* genes was able to confer tolerance of desiccation to *lec1* seeds but not expression of two other genes encoding less conserved HAP3 subunits. We demonstrated that expression of *TcL1L* was able to complement most of the morphological defects of *lec1* embryos. Moreover, post-embryonic expression of *TcL1L* in mutant or wild type backgrounds was responsible for the development of leaves with embryonic traits, although the phenotypes on germination were less severe than those observed by Kwong et al. (2003). To conclude, our results strongly support a role for *TcL1L* in cocoa embryogenesis similar to that of *L1L* in *Arabidopsis*.

TcL1L is expressed very early during cocoa zygotic and somatic embryogenesis, and especially into the protoderm, and epidermis. The cells of this special cell layer are precisely the place of origin of secondary somatic embryos. Because of that, and despite the fact that its over-expression in a heterologous system such as *Arabidopsis* does not induce supernumerary ectopic embryos, *TcL1L* is a potential gene to improve cocoa somatic embryogenesis. It would be very interesting and informative to over-express *TcL1L* in *T. cacao*.

Acknowledgments The authors would like to thank Peter Biggins for checking the English and Fabienne Lapeyre-Montes for technical help. Plant Gateway vectors were provided by the Functional Genomics Division of the Department of Plant Systems Biology (VIB-Ghent University).

References

Alemanno L, Berthouly M, Michaux-Ferrière N (1996) Histology of somatic embryogenesis from floral tissues in *Theobroma cacao* L. Plant Cell Tissue Organ Cult 46:187–194

Alemanno L, Berthouly M, Michaux-Ferrière N (1997) A comparison between *Theobroma cacao* L. zygotic and somatic embryogenesis from floral explants. In Vitro Cell Dev Biol Plant 33:163–172

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402

Bennett AB (2003) Out of the Amazon: *Theobroma cacao* enters the genomic era. Trends Plant Sci 8:561–563

Bolognese F, Imbriano C, Caretti G, Mantovani R (2000) Cloning and characterization of the histone-fold proteins YBL1 and YCL1. Nucleic Acids Res 28:3830–3838

Boutilier K, Offringa R, Sharma VK, Kieft H, Ouellet T, Zhang L, Hattori J, Liu C-M, van Lammeren AM, Miki BLA, Custers JBM, van Lookeren Campagne MM (2002) Ectopic expression of *BABY BOOM* triggers a conversion from vegetative to embryonic growth. Plant Cell 14:1737–1749

Clough SJ, Bent AF (1998) Floral dip: a simplified method for Agrobacterium mediated transformation of *Arabidopsis thaliana*. Plant J 16:735–743

Esan EB (1974) Development of adventive embryos from immature embryos of *Theobroma cacao* L. culture in vitro 10th annual conference of Agricultural Society Nigeria 43

Eskes B (2000) Introductary notes. International workshop on new technologies and cocoa breeding. Kota Kinabalu, Malaysia, pp 8–11

Fambrini M, Durante C, Cionini G, Geri C, Giorgetti L, Michelotti V, Salvini M, Pugliesi C (2006) Characterization of *LEAFY COTYLEDON1-LIKE* gene in *Helianthus annuus* and its relationship with zygotic and somatic embryogenesis. Dev Genes Evol 216:253–264

Feher A, Pasternak T, Dudits D (2003) Transition of somatic plant cells to an embryogenic state. Plant Cell Tissue Organ Cult 74:201–228

Figueira A, Alemanno L (2004) *Theobroma cacao*. In: Litz RE (ed) Biotechnology of fruit and nut crops 29:639–669

Finkelstein RR, Lynch TJ (2000) The *Arabidopsis Abscisic Acid Response* gene *ABI5* encodes a basic leucine zipper transcription factor. Plant Cell 12:599–610

Finkelstein RR, Wang ML, Lynch TJ, Rao S, Goodman HM (1998) The *Arabidopsis Abscisic Acid Response* locus *ABI4* encodes an APETALA 2 domain protein. Plant Cell 10:1043–1054

Gaj MD, Zhang S, Harada JJ, Lemaux PG (2005) *Leafy cotyledon* genes are essential for induction of somatic embryogenesis of *Arabidopsis*. Planta 22:977–988

Gazzarrini S, Tsuchiya Y, Lumba S, Okamoto M, McCourt P (2004) The transcription factor FUSCA3 controls developmental timing in *Arabidopsis* through the hormones gibberellin and abscisic acid. Dev Cell 7:373–385

Giraudat J, Hauge BM, Valon C, Smalle J, Parcy F, Goodman HM (1992) Isolation of the *Arabidopsis ABI3* gene by positional cloning. Plant Cell 4:1251–1261

Harada JJ (2001) Role of *Arabidopsis LEAFY COTYLEDON* genes in seed development. J Plant Physiol 158:405–409

Hecht V, Vielle-Calzada J-P, Hartog MV, Schmidt DL, Boutilier K, Grossniklaus U, de Vries SC (2001) The *Arabidopsis Somatic Embryogenesis Receptor Kinase 1* gene is expressed in developing ovules and embryos and enhances embryogenic competence in culture. Plant Physiol 127:803–816

Irizarry H, Rivera E (1998) Early yield of five cacao families at three locations in Puerto Rico. J Agric 82:163–171

Jones PG, Allaway D, Gilmour MD, Harris C, Rankin D, Retzel ER, Jones CA (2002) Gene discovery and microarray analysis of cacao (*Theobroma cacao* L.) varieties. Planta 216:255–264

Karimi M, Inzé D, Depicker A (2002) Gateway vectors for *Agrobacterium*-mediated plant transformation. Trends Plant Sci 7:193–195

Karlova R, Boeren S, Russinova E, Aker J, Vervoort J, de Vries S (2006) The Arabidopsis SOMATIC EMBRYOGENESIS RECEPTOR-LIKE KINASE1 protein complex includes BRASSINOSTEROID-INSENSITIVE1. Plant Cell 18:626–638

Kagaya Y, Toyoshima R, Okuda R, Usui H, Yamamoto A, Hattori T (2005) *LEAFY COTYLEDON1* controls *Seed Storage Protein* genes through its regulation of *FUSCA3* and *ABSCISIC ACID INSENSITIVE3*. Plant Cell Physiol 46:399–406

Kuhn DN, Heath M, Wisser RJ, Meerow A, Brown JS, Lopes U, Schnell RJ (2003) Resistance gene homologues in *Theobroma cacao* as useful genetic markers. Theor Appl Genet 107:191–202

Kwong RW, Bui AQ, Lee H, Kwong LW, Fischer RL, Goldberg RB, Harada JJ (2003) LEAFY COTYLEDON1-LIKE defines a class of regulators essential for embryo development. Plant Cell 15:5–18

Lee H, Fischer RL, Goldberg RB, Harada JJ (2003) Arabidopsis LEAFY COTYLEDON1 represents a functionally specialized subunit of the CCAAT binding transcription factor. Proc Natl Acad Sci USA 100:2152–2156

Li Z, Traore A, Maximova S, Guiltinan M (1998) Somatic embryogenesis and plant regeneration from floral explants of cacao (*Theobroma cacao* L.) using thidiazuron. In Vitro Cell Dev Biol Plant 34:293–299

Lotan T, Ohto M, Yee KM, West MAL, Lo R, Kwong RW, Yamagishi K, Fisher RL, Golberg RB, Harada JJ (1998) *Arabidopsis* LEAFY COTYLEDON1 is sufficient to induce embryo development in vegetatives cells. Cell 93:1195–1205

Maximova NS, Alemanno L, Young A, Ferrière N, Traore A, Guiltinan MJ (2002) Efficiency, genotypic variability and cellular origin of primary and secondary somatic embryogenesis of *Theobroma cacao* L. In Vitro Cell Dev Biol Plant 38:252–259

Murashige T, Skoog F (1962) A revised medium for rapid growth and bioassays with tobacco tissue cultures. Physiol Plant 15:473–497

Parcy F, Valon C, Kohara A, Misera S, Giraudat J (1997) The *ABSCISIC ACID-INSENSITIVE3*, *FUSCA3*, and *LEAFY COTYLEDON1* loci act in concert to control multiple aspects of *Arabidopsis* seed development. Plant Cell 9:1265–1277

Risterucci AM, Grivet L, N'Goran JKA, Pieretti I, Flament MH, Lanaud C (2000) A high density linkage map of *Theobroma cacao* L. Theor Appl Genet 101:948–955

Schmidt E, Guzzo F, Toonen M, de Vries S (1997) A leucine-rich repeat containing receptor-like kinase marks somatic plant cells competent to form embryos. Development 124:2049–2062

Sinha S, Kim IS, Sohn KY, de Crombrugghe B, Maity SN (1996) Three classes of mutations in the A subunit of the CAAT-binding factor CBF delineate functional domains involved in the three-step assembly of the CBF-DNA complex. Mol Cell Biol 16:328–337

Somleva MN, Schmidt ED, de Vries SC (2000) Embryogenic cells in *Dactylis glomerata* L. (Poaceae) explants identified by cell-tracking and by *SERK* expression. Plant Cell Rep 19:718–726

Stone SL, Kwong LW, Yee KM, Pelletier J, Lepiniec L, Fischer RL, Goldberg RB, Harada JJ (2001) *LEAFY COTYLEDON2* encodes a B3 domain transcription factor that induces embryo development. Proc Natl Acad Sci USA 98:11806–11811

To A, Valon C, Savino G, Guilleminot J, Devic M, Giraudat J, Parcy F (2006) A network of local and redundant gene regulation governs *Arabidopsis* seed maturation. Plant Cell 18:1642–1651

Verdeil J-L, Alemanno L, Niemenak N, Tranbarger TJ (2007) Pluripotent versus totipotent plant stem cells: dependence versus autonomy? Trends Plant Sci 12:243–252

Xing Y, Fikes JD, Guarente L (1993) Mutations in yeast *HAP2/HAP3* define a hybrid CCAAT box binding domain. EMBO J 12:4647–4655

Yazawa K, Takahata K, Kamada H (2004) Isolation of the gene encoding carrot leafy cotyledon1 and expression analysis during somatic and zygotic embryogenesis. Plant Physiol Biochem 42:215–223

Zegzouti H, Jones B, Frasse P, Marty C, Maitre B, Latche A, Pech J-C, Bouzayen M (1999) Ethylene-regulated gene expression in

tomato fruit: characterization of novel ethylene-responsive and ripening-related genes isolated by differential display. Plant J 18:589–600

Zemzoumi K, Frontini M, Bellorini M, Mantovani R (1999) NF-Y histone fold? 1 helices help impart CCAAT specificity. J Mol Biol 286:327–337

Zhang S, Wong L, Meng L, Lemaux PG (2002) Similarity of expression patterns of *Knotted1* and *ZmLEC1* during somatic and zygotic embryogenesis in maize (*Zea mays* L.). Planta 215:191–194