Available online at www.sciencedirect.com
SCIENCE DIRECT
Plant Science 168 (2005) 723–729
PLANT SCIENCE
www.elsevier.com/locate/plantsci

Characterisation of the cacao *somatic embryogenesis receptor-like kinase* (*SERK*) gene expressed during somatic embryogenesis

Marcelo de Oliveira Santos[b,c,*], Eduardo Romano[a], Karla Suemy Clemente Yotoko[c], Maria Laine Penha Tinoco[a,b], Bárbara Barreto Andrade Dias[a], Francisco José Lima Aragão[a]

[a]*Embrapa Recursos Genéticos e Biotecnologia, PqEB Final W5 Norte, 70770-900 Brasília, DF, Brasil*
[b]*Departamento de Biologia Celular, Universidade de Brasília, CEP 70910-900 Brasília, DF, Brasil*
[c]*Departamento de Biologia, ICB, Universidade Federal de Juiz de Fora, 36036-330 Juiz de Fora, MG, Brasil*

Received 17 September 2004; accepted 1 October 2004
Available online 27 October 2004

Abstract

Somatic embryogenesis is a useful tool for plant propagation and consists of a good model for embryo development studies. During somatic embryogenesis, some morphological and biochemical changes occur in response to alterations in gene expression patterns. The *somatic embryogenesis receptor kinase* (*SERK*) gene is an important element related to this process. In this work we have focused on the function of *SERK* in *Theobroma cacao*. Our main questions were whether this gene exists in this plant with another arrangement; if it is functional and, in a positive case, in which tissues it acts. We have found only one copy of *SERK* that is functional and analyzed its expression in a series of tissues. The phylogenetic tree showed that the available sequences of *SERK* are highly conserved and a considerable number of species contain different sequences of *SERK*, confirming it as a gene family with variant repeats.

Keywords: RT-PCR; Molecular phylogeny; Gene family; Gene cloning; Gene expression; Theobroma cacao

1. Introduction

Somatic embryogenesis is a useful tool for plant propagation and consists of an ability of competent cells to change their differentiation pathway in order to become somatic embryos naturally or through tissue cultures [1]. Somatic embryos comprise a good model for embryo development studies since they are easy to manipulate and can be produced in large scale [2]. The proliferation of these somatic embryos from each other is known as repetitive somatic embryogenesis and has been used for genetic transformation in various species [3]. During somatic embryogenesis, biochemical and morphological changes occur throughout the development of induced tissues [4], which is strongly related to alterations in gene expression patterns [2]. Therefore, some genes are differentially expressed during somatic embryogenesis induction, while others are expressed during differentiation from embryo maturation up to full plant development [2].

Among the genes involved with the induction of somatic embryogenesis, the somatic embryogenesis receptor kinase (*SERK*) gene is claimed to have an important role. A single cell of *Daucus carota* that expresses this gene has the ability to form a somatic embryo [5], what was also observed in *Dactilys glomerata* [6]. In *Arabidopsis thaliana*, the expression of *SERK* was observed in transgenic embryogenic *calli* from floral explants and in zygotic embryos [7]. In maize (*Zea mays*), *SERK* is expressed in different tissue types, such as in embryogenic cells Type 2, late embryo stages, ovules and leaves [8]. Concerning the number of copies, *Arabidopsis* contains five [7] and maize three [8], indicating that *SERK* genes constitute a gene family, at least

* Corresponding author. Tel.: +55 3232293206; fax: +55 3232293216.
E-mail addresses: marcelo.santos@ufjf.edu.br (M. de O Santos), aragao@cenargen.embrapa.br (F.J.L. Aragão).

0168-9452/$ – see front matter
doi:10.1016/j.plantsci.2004.10.004

in these species. Only one copy of this gene was found in carrot [5] and *Medicago truncatula*, but it is possible that the latter contains a higher number of copies [9]. Regardless the numbers of genes, all of these organisms present a functional copy of the *SERK* gene that is preferentially expressed during the induction of somatic embryogenesis and probably in zygotic embryogenesis. Phylogenetic studies have shown that the putative functional *SERK* copies of *M. truncatula*, *A. thaliana* and *D. carota* are orthologous sequences [9].

The product of this gene is a protein that carries leucine-rich repeat (LRR) domains exposed at the cell membrane that should be responsible for the interaction with other signal molecules, external signal perceptions and signal transduction during somatic embryogenesis induction [5,7]. This protein belongs to the receptor-like kinase-LRR (RLK) family with a serine-threonine kinase site [5].

The focus of this paper lies on the somatic embryogenesis of cacao (*Theobroma cacao* L.), considered as a recalcitrant species for induction of somatic embryogenesis and seed conservation [10–12]. Regardless the importance of this crop both in the food and pharmaceutical industries, only recently it was demonstrated that secondary somatic embryos of cacao can be obtained from a single cell [13], which finally allowed the development of transgenic plants based on agrobacterium systems [14]. Simultaneously, our group induced the repetitive somatic embryogenesis after continual cultivation of staminodes under 2,4-D and thidiazuron and showed that somatic embryos of cacao can be maintained up to globular stage, making the production of several new secondary embryos possible [15].

The knowledge of the function of the *SERK* gene in somatic embryogenesis in some plant species and the difficulty of the ''in vitro'' culture of cacao suggest queries about the role of *SERK* in these species. Therefore, we have isolated and sequenced the cacao's *SERK* (*TcSERK*) expressed during induction of somatic embryogenesis in cacao and analyzed its expression in different types of tissues. We have also compared the *TcSERK* sequence with sequences of other species available in GenBank and TIGR databases, and performed a phylogenetic hypothesis in order to check the relationships between *TcSERK* and the sequence of this gene in other plants, as well as intended to make inferences about the function(s) of this gene in plants.

2. Materials and methods

2.1. Tissue culture of embryogenic calli and RNA extraction

Floral buds of cacao (genotype Cenargen-2) were surface-sterilized with 2% sodium hypochlorite for 15 min and washed six times in distilled autoclaved water. Staminods were aseptically excised from buds and then cultivated on Plant Callus Growth medium (pH = 5.8), PCG (DKW [16] supplemented with 9 μM 2,4-D, 22.7 nM thidiazuron, 2% sucrose and 0.2% phytagel) [10] for 14 days to obtain the 2-week-old *calli*. In order to induce repetitive somatic embryogenic clusters, cultures were maintained in this medium for 6 weeks more and subcultured at intervals of 14 days [15]. Total RNA was extracted using the RNAeasy kit (Qiagen, Valencia, CA, USA) both from fresh embryogenic 2-week-old *calli* and from repetitive embryogenic clusters (100 mg of each tissue).

2.2. Cloning and sequencing of cacao TcSERK cDNA

Two micrograms of total RNA from each sample were used to produce total cDNA using the Superscript II kit (Invitrogen, Carlsbad, CA, USA). Degenerated primers (forward S3, S4 and S5 and reverse S1, S2) [8] were used to amplify cacao *SERK* gene coding sequence with all possible combinations. Each RT-PCR reaction was carried out in a PTC-100 termocicler (MJ Researcher) in 50 μL solution containing 40 μg of cDNA, 60 mM Tris–SO_4 (pH 8.9), 18 mM $(NH_4)_2SO_4$, 2 mM $MgSO_4$, 250 nM of each dNTP; 200 nM of each primer, 5 U of Platinum® *Taq* DNA Polymerase High Fidelity (Invitrogen, Carlsbad, CA, USA). The mixture was overlaid with mineral oil, denatured at 95 °C (5 min) and subjected to 35 cycles of amplification (95 °C for 1 min, 50 °C for 1 min, 72 °C for 1 min) with a final elongation cycle of 5 min at 72 °C.

The primer pair S1/S5 generated the longest fragment product of the cacao *SERK* gene (1.3 kb), which was cloned into the pGMTeasy vector for PCR products (Promega, Madison, WI, USA). Since this fragment is relatively small, we used the GeneRacer Kit (Invitrogen, Carlsbad, CA, USA) to amplify the remaining coding sequence, using the primers S1 and GeneRacer 5′. The single 1.5 kb fragment was cloned into pCR4 TOPO (Invitrogen, Carlsbad, CA, USA). Twelve clones of each fragment; S1/S5, and S1/GeneRacer 5′; were sequenced by using universal M13 and T7 primers on automatic sequencer (ABI Prism® 3700, Foster City, CA, USA).

2.3. Sequence analysis of DNA and putative protein prediction

The obtained sequences were aligned on the DNAMAN program (Lynnon BioSoft—Version 4.0). Protein predictions were performed using the PSORT (Prediction of Protein Sorting Signals and Localization Sites in Amino Acid Sequences) (http://psort.nibb.ac.jp/) [17]; and PROSITE programs (Database of Protein Families and Domais) (http://www.expasy.org/prosite/) [18].

2.4. TcSERK expression analysis by RT-PCR

Specific primers for *TcSERK*, TCSPF (5′-AAGCGGGAATAGTGTCAC-3′) and TCKR (5′-GCTTCAAACTCCTCATCC-3′) were designed from the obtained sequence to

analyze the expression of this gene. Therefore, we searched for *TcSERK* in eight different tissues, staminodes, embryogenic *calli* after 2 weeks of culture, repetitive embryogenic clusters, mature somatic embryos, mature zygotic embryos, roots, leaves and petals. In order to accomplish that, the total RNA was extracted separately from each tissue (samples of 200 mg) with the RNAeasy kit (Qiagen, Valencia, CA, USA). The remaining genomic DNA was eliminated by DNase digestion of the RNA samples. Total RNA was used to produce cDNA using the reverse transcriptase Superscript II (Invitrogen, Carlsbad, CA, USA), according to the protocol suggested by the manufacturer—except for the use of random decamers (Ambion Inc. Ustin, TX, USA). PCR reactions were performed with the specific primers described above to amplify a fragment of 635 bp of *TcSERK*. These reactions were carried out as described above (cloning and sequencing topic), with 24 cycles of amplification. As internal control, the primers RNATc381 (5′-AACGGCTACCACATCCAAGG-3′) and TcRNA820C (5′-TCATTACTCCGATCCCGAAG-3′) were used to amplify 393 bp within the *18S* rRNA gene from *T. cacao* (GenBank Accession No. AF207040). Each reaction was repeated three times and submitted to electrophoresis in 1.0% agarose gel.

2.5. Phylogenetics

In order to check the relationship between the *SERK* in cacao and this gene in other plant species, we aligned it with the *SERK* sequences available at GenBank (www.ncbi.nih.nlm.gov) and TIGR plant databases (www.tigr.org). These sequences were found by searching for sequences annotated as "*SERK*" or "*SERK* homologues". In addition, sequences not identified as *SERK* were retrieved by BLAST searches [19] at these databases, performing a total of 35 sequences out of 16 different species of plants. The alignment was initially performed using CLUSTAL W (http://www.ebi.ac.uk/clustalw). Subsequently, it was eye inspected and corrected manually using the BioEdit 5.0.9 program [20].

After alignment, the MEGA 2.1 [21] program was used to analyze the sequences, which, in turn, showed a high proportion of transversions in relation to the transitions, indicating that some corrections, in principle Kimura-2 parameters [22], must be implemented in order to construct the phylogenetic tree with a Neighbour-Joining (NJ) distance method [23]. Nevertheless, the *p*-distance was 0.23 on average, which is considered a large distance, and indicates that a more complex model must be used. Thus, we used the Model Test 3.06 program [24] to establish the model of DNA evolution that would best fit our data. The selected model was the general time reversible (GTR) one with a proportion of invariant sites and a correction for gamma distribution. This complex model requires a maximum likelihood (ML) approach [25,26]. We thus chose to run our phylogenetic analysis using the MrBayes 3.0 [27] program and to perform a Bayesian inference of phylogeny using a variant Markov Chain Monte Carlo (MCMC). This method was chosen because the same models of DNA substitution used in ML approaches can be used in a Bayesian analysis [28], with an advantage of less time consumption and the possibility to perform a large quantity of replications, giving an estimate of the statistical support of the branches. We performed the MCMC replications for 1,000,000 cycles replications and built and drew the consensus over these replications with the CONSENCE and DRAWTRE programs of the Phylip Package v. 3.6 [29].

So as to ensure that our results were consistent [30], we have also constructed the tree with NJ using Kimura-2 parameters model. The robustness of each internal branch of this tree was estimated using the nonparametric bootstrap test (PB) [31,32], and the confidence probability test (CP) [33].

3. Results

In the present work, we considered sectors of *calli* that turn into compact bright yellow masses of cells (dark arrow in Fig. 1) as repetitive embryos clusters. Those sectors were cultivated in induction media for 8 weeks. Most sectors of these *calli* became a mass of friable aqueous cells, which were highly oxidized and lost their ability to turn into repetitive embryos (white arrow in Fig. 1).

3.1. Tissue culture, cloning and sequence analysis of cacao SERK gene

After 2 weeks of culture, 100% of staminods formed embryogenic *calli*, but only 5% of them produced repetitive embryos after 8 weeks of culture. RT-PCRs showed that both embryogenic *calli* (after 2 weeks of culture) and repetitive somatic embryos contain the same sequence of *SERK*, since all combinations with the primers S4 and S5 (S1–S4, S2–S4, S1–S5, S2–S5 and S5–GeneRacer 5′) produced identical aligned sequences for both tissues (the combinations with the primer S3 did not work). The longest sequence (S5–GeneRacer 5′) was compared with other sequences of *SERK* available in GenBank and showed a high similarity (≈70%) with each and every one of them (data not shown).

The putative protein prediction analysis of this sequence showed that it contains all domains found in SERK protein of other species such as the signal peptide (SP), with a possible cleavage site in position 29, followed by a leucine ziper (LZ) domain (between 36–57 positions) similar to *AtSERK1* [7]; five LRR domains with glycosilation sites, followed by a single SPP pro-rich domain which is the hallmark of SERK protein [7]; and a transmembrane Ala-Rich domain, followed by a Kinase domain (Fig. 2). The Kinase domain contains an ATP-binding site, a ser/thre domain and a protein kinase C site.

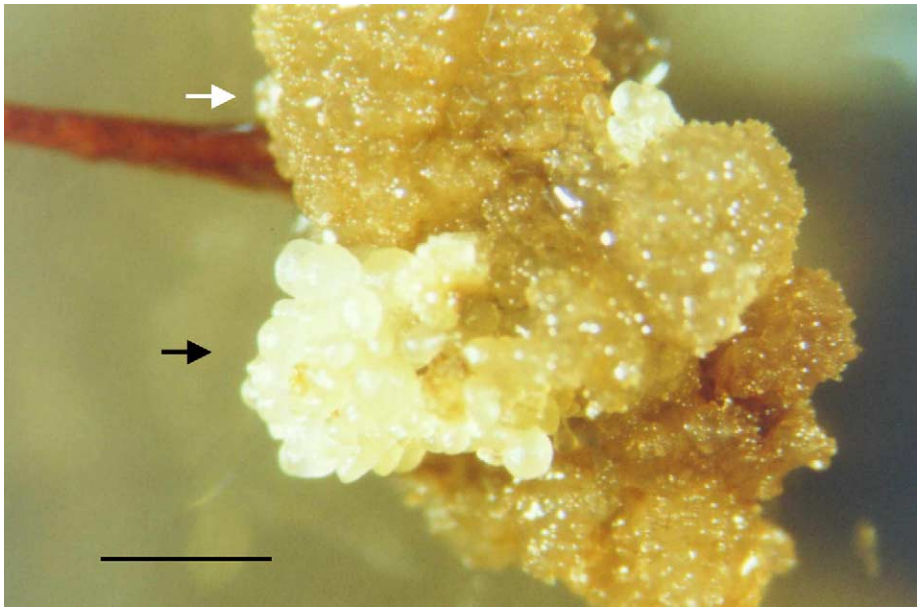

Fig. 1. Induced embryogenic cacao *calli* from staminode after 8 weeks of culture on PCG medium (black arrow—repetitive somatic embryos; white arrow—oxidised sector of *calli* with high phenol content) (bar = 0.5 cm).

3.2. Expression analysis of cacao SERK gene

Once established the *TcSERK* gene sequence, specific primers were designed to detect the *TcSERK* expression in different tissues (Fig. 3). For instance, *TcSERK* was highly expressed at initial induced embryogenic *calli* and at repetitive embryogenic sectors, while a weak signal was found in leaves and somatic and zygotic mature embryos and no signals were found in roots, petals and staminods.

3.3. Phylogenetic analysis

The majority rule consensus tree taken from every 100th tree sampled during generations 1–1,000,000 is shown in Fig. 4. The average ln likelihood score was −22128.4 and the maximum likelihood score in the chains was −22096.8. For each replication, the program calculated the proportion of invariable sites that averaged 0.275010 and the alpha parameter of gamma distribution, which was 1.06 on average.

The rootless phylogenetic tree obtained in the phylogenetic analysis of *SERK* showed three main branches. The first includes the *Pinnus* species, the only gymnosperm species included in our tree. The second main branch encompasses all the sequences of monocots, and this branch appeared in 9,984 replications. The third main branch comprises all the sequences of dicotyledonous, which appeared in 9,838 replications.

The branch of the monocots was split into three conspicuous branches. From the monocot sequences, only three maize sequences were described as functional [8]. The copies 2 and 3 derive apparently from a recent duplication event, while the copy 1 resulted from a more ancient duplication, which, following this topology, occurred before the raising speciation of maize. In addition, there is a fourth copy in GenBank (*Zm* AY104736), that may have resulted from another event of duplication. Similarly to these events concerning the fourth copy, there is no information about the functionally of the remaining sequences. Nevertheless, the tree also shows that other species, such as wheat (*Triticum aestivum*) and rice (*Oriza sativa*) contain copies of *SERK* derived from recent and old duplication events as well, which was evidenced by the division of sequences of these species into different branches. This pattern suggests that some copies were inherited from ancient monocots while others occurred more recently.

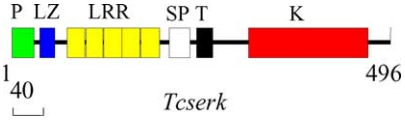

Fig. 2. Schematic representation of *T. cacao* SERK putative protein domains, according to *pfam* search domain program: P—signal peptide domain; LZ—leucine zipper; LRR—leucine rich repeat; SP—proline rich region; T—transmembrane domain; K—kinase domain.

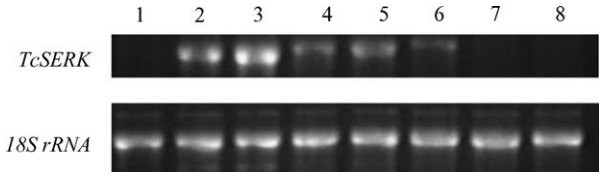

Fig. 3. RT-PCR of *TcSERK* fragment gene from cacao tissues (635 bp long). Upper lanes: (1) staminodes; (2) 2 weeks old *calli*; (3) repetitive embryogenic cluster after 8 weeks of culture; (4) somatic mature embryo; (5) zygotic mature embryo; (6) leaves; (7) petals; (8) roots. Lower lanes = -cacao 18S rRNA internal control.

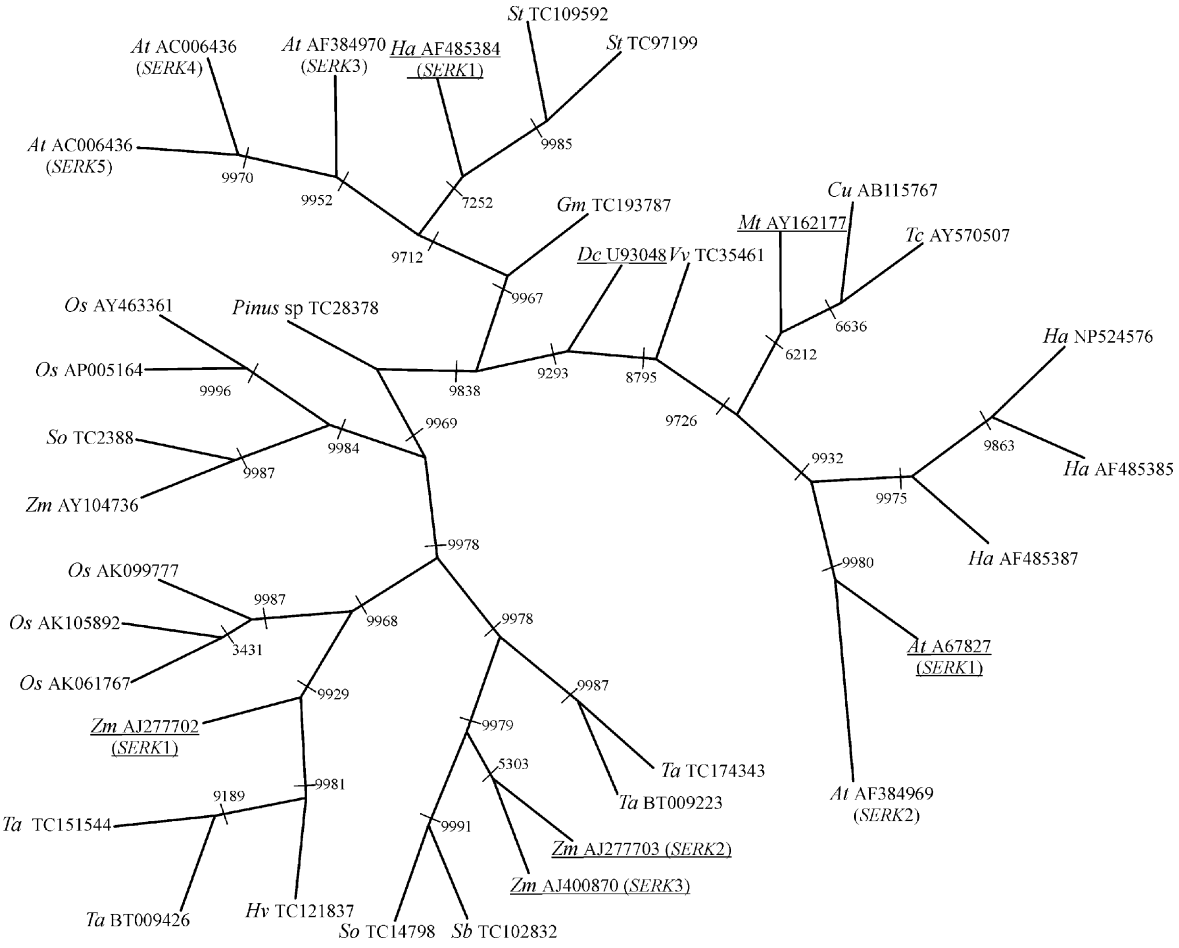

Fig. 4. Majority rule consensus tree taken from every 100th tree sampled during generations 1–1,000,000 of Bayesian analysis. The tree includes all sequences of *SERK* available in GenBank and TIGR databases. At = *Arabidopsis thaliana*; Dc = *Daucus carota*; Zm = *Zea mays*; Mt = *Medicago truncatula*; Tc = *Theobroma cacao*; Os = *Oryza sativa*; Ha = *Helianthus annuus*; St = *Solanum tuberosum*; Gm = *Glycine max*; Vv = *Vitis vinifera*; Ta = *Triticum aestivum*; Cu = *Citrus unshiu*; So = *Saccharum officinarum*; Hv = *Hordeum vulgare*; Sb = *Sorghum bicolor*. Besides the abbreviation of the species, the accession number of each sequence is represented. The underlined sequences are the ones confirmed as functional in other studies.

The branch of dicots contains four functional sequences of *SERK* from *Arabidopsis* [7], *Daucus* [5], sunflower [34] and *Medicago* [9]. Unfortunately, only three species (*Arabidopsis*, sunflower and potato) have more than a copy of *SERK*. In these species, as found for monocots, some duplications are recent (such as *At SERK*s 1 and 2 that grouped with three copies of sunflower; and *At SERK*s 3, 4 and 5, which grouped with another copy of sunflower and the two copies of potato) and others are older, and may have occurred before the events of speciation of these species.

4. Discussion

The main purpose of this work was to verify the structure and functionality of the *SERK* gene in *Theobroma cacao*, and, in a positive case, search for the cacao tissues that actually express this gene. In addition, we intended to make inferences about the function of this gene in plants. To implement that, we searched for the *TcSERK* expression at induced *calli* after 2 weeks of culture and at repetitive somatic embryogenic clusters, and found, through a cDNA analysis, that this gene is expressed in both embryogenic stages (Fig. 1).

The coding sequences obtained from these two types of *callus* were cloned and sequenced, and their nucleotide sequences were identical, providing evidence that the same copy is transcribed during both stages. Unfortunately, we could not determine the exact number of copies of this gene by Southern Blot, which must be investigated in a further study, using different procedures, such as chromosome localization by cytological techniques.

The functionality of *TcSERK* could be inferred by the fact that all important domains found in putative functional copies of SERK protein were detected. Indeed, the Kinase domain contains a ser/thre site and a binding site to the protein kinase C, which are characteristic sites in RLK signal transduction proteins [35]. In addition, the persistence of induction conditions to obtain the somatic embryogenic clusters maintains both the transcription of *TcSERK* in high

levels and the repetitive status, which can be maintained for at least 3 months (data not shown). Therefore, it is reasonable to infer that the repetitive globular stage can be maintained, at least in part, by the expression of high amounts of *SERK*, which was also found in *Arabidopsis thaliana* [7].

Regarding the embryogenic tissues where *TcSERK* is expressed, we have found that it is highly expressed at initial induced embryogenic *calli* and at repetitive embryogenic sectors, what was also found in carrot [5], *Dactylis glomerata* [6] and *Arabidopsis thaliana* [7]. Interestingly, the *TcSERK* expression was found in somatic and zygotic mature embryos, suggesting that this gene plays a role during the development of cacao embryos, while it is expressed only up to globular stage in carrot [5] and *D. glomerata* [6] and up to heart stage in *A. thaliana* [7] and *Hieracium* [36]. In maize, the second copy of *SERK* (*ZMSERK*2) is expressed in further embryogenic stages as well as in *Medicago truncatula*, where the *MtSERK* is expressed in different stages of embryos.

Concerning the *TcSERK* expression in other tissues, a weak signal was found in leaves, though we could not detect the expression of *TcSERK* in roots, petals and staminods. In *Arabidopsis* and *Hieracium*, transcripts of *SERK* were found in seedlings and adult vascular tissues and in apical meristems [7,36]. In *Dactylis*, the expression of *SERK* was found both in apical and radicular meristems [6]; in *Medicago*, its expression was found during rhizogenesis [9]; whereas in maize, *SERK* transcripts were found in leaves and ovules [8]. All these data provide some evidence that, although the expression of *SERK* is extremely related to the induction of embryogenesis, this gene may play a broader and unknown role in the metabolism of plants.

It is important to notice that in some species, such as carrot and *Arabidopsis*, both the somatic embryos and a high *SERK* expression are induced by the addition of auxin in the culture medium, whereas in *M. truncatula*, somatic embryogenesis and the transcription of *SERK* are induced under the addition of both auxin and cytokinin, which was also found in cacao [10,37], suggesting an up-regulation over *MtSERK1* [9], and consequently over *TcSERK* expression.

The topology shown in Fig. 4 provides evidence that most sequences of *SERK* represented in this tree are functional in their respective organisms. This conclusion follows the fact that the known functional sequences (underlined in Fig. 4) are widely scattered along all the branches of the tree, and not concentrated in a single or few branches. Indeed, each important branch of our tree contains at least one known functional sequence. The *T. cacao SERK* sequence was placed in a branch containing the *M. trucatula SERK*, a putative functional sequence [9].

Interestingly, some species represented in this tree, such as maize, rice, wheat, sugarcane, sunflower, potato and *Arabidopsis*, contain more than a copy of *SERK*, and that copies of the same plant were frequently placed in different branches of the tree, indicating that some duplications occurred before the division among some of these species. On the other hand, some duplications of the same plant appear in the same branch, showing that these duplications occurred after the events of speciation of the species represented in this tree. In any way, all copies of *SERK* present in each organism surely belong to a gene family with variant repeats (>50% of similarity among copies in the same species). Variant repeats may differ in function or regulation to some extent, so that they contribute to the fine-turning of physiological functions of an organism [38]. Unfortunately, it is hard to state whether some of these sequences correspond to pseudo-genes, but if that is the case for some sequences, the duplications must be very recent, since none of the sequences included in this work present frame shifts or premature stop codons, which are inevitable as evolution progresses [39,40].

It is important to reinforce that the topology consistently separated monocots and dicots sequences. Additionally, the only sequence obtained from a gymnosperm was placed as a third branch of the rootless tree (and it is, in fact, the root of a rooted tree—data not shown). This separation, that in a way follows the known phylogeny and taxonomy of these plants, indicates that this gene has a conserved function in plants and that it plays a key role for their survival, which explains the restrictions for nucleotide (and amino acids) substitutions.

Finally, the results found in this article strongly suggest that the *TcSERK* is a functional gene that plays an important role during cacao embryo development and the maintenance of repetitive status. Moreover, it has been found that this gene is expressed in tissues not involved in embryogenesis in cacao, which has also been found in other species [6–9,36]. This suggests a broader function for this gene, which must be investigated in further studies. Our phylogenetic analysis confirmed that *SERK* belongs to a gene family in some of the studied species, which leads to the conclusion that more studies must be carried out in order to determine if this is a general pattern. After all, the high conservation of the *SERK* sequences implies restrictions to aminoacid substitutions in this gene and strongly suggests that it is extremely relevant to the survival of the plants.

References

[1] E.G. Willians, G. Maheswaran, Somatic embryogenesis: factors influencing coordinated behaviour as an embryogenic group, Ann. Bot. 57 (1986) 443–462.
[2] J.L. Zimmerman, Somatic embryogenesis: a model for early development in higher plants, Plant Cell 5 (1993) 1141–1423.
[3] S.A. Merkle, W.A. Parrott, E.G. Williams, Applications of somatic embryogenesis and embryo cloning, in: S.S. Bhojwani (Ed.), Plant Tissue Culture: Applications and Limitations, Elsevier, Amsterdam, 1990, pp. 67–101.
[4] W.R. Sharp, M.R. Sondhal, L.S. Caldas, S.B. Maraffa, The physiology of in vitro assexual embryogenesis, Horticult. Rev. 2 (1980) 47–54.

[5] E.D.L. Schmidt, F. Guzzo, M.A.J. Toonen, S.C. de Vries, A leucine-rich repeat containing receptor-like kinase marks somatic plant cells competent to form embryos, Development 124 (1997) 2049–2062.
[6] M.N. Somleva, E.D.L. Schmidt, S.C. de Vries, Embryogenic cells in *Dactylis glomerata* L. (Poaceae) explants identified by cell tracking and by SERK expression, Plant Cell Rep. 19 (2000) 718–726.
[7] V. Hecht, J.P. Vielle-Calzada, M.V. Hartog, D.L. Schmidt, K. Boutilier, U. Grossnicklaus, S.C. de Vries, The *Arabdopsis* SOMATIC EMBRYOGENESIS RECEPTOR KINASE 1 gene is expressed in developing ovules and embryos and enhances embryogenic competence in culture, Plant Physiol. 127 (2001) 803–816.
[8] S. Baudino, S. Hansen, R. Brettschneider, V.F.G. Hecht, T. Dresselhaus, H. Lörz, C. Dumas, P.M. Rogowsky, Molecular characterization of two novel maize LRR receptor-like kinase, which belong to the *SERK* gene family, Planta 213 (2001) 1–10.
[9] K.E. Nolan, R.R. Irwanto, R.J. Rose, Auxin up-regulates *MtSERK1* expression in both *Medicago truncatula* root-forming and embryogenic cultures, Plant Physiol. 133 (2003) 218–230.
[10] Z. Li, A. Traore, S. Maximova, M.J. Guiltinan, Somatic embryogenesis and plant regeneration from floral explants of cacao (*Theobroma cacao* L.) using thidiazuron, In Vitro Cell Dev. Biol-Plant 34 (1998) 293–299.
[11] J.Y. Fang, A. Wetten, P. Hadley, Cryopreservation of cocoa (*Theobroma cacao* L.) somatic embryos for long-term germplasm storage, Plant Sci. 166 (2004) 669–675.
[12] L. Alemanno, T. Ramos, A. Gargadenec, C. Andary, N. Ferriere, Localization and identification of phenolic compounds in *Theobroma cacao* L. somatic embryogenesis, Ann. Bot. 92 (2003) 613–623.
[13] S.N. Maximova, L. Alemanno, A. Young, N. Ferriere, A. Traore, M.J. Guiltinan, Efficiency, genotypic variability, and cellular origin of primary and secondary somatic embryogenesis of *Theobroma cacao* L., In Vitro Cell Dev. Biol-Plant 38 (2002) 252–259.
[14] S. Maximova, C. Miller, G. Antunez de Mayolo, S. Pishak, A. Young, M.J. Guiltinan, Stable transformation of *Theobroma cacao* L. and influence of matrix attachment regions on GFP expression, Plant Cell Rep. 21 (2003) 872–883.
[15] M.O. Santos, E.V.S.A. Barros, M.L.P. Tinoco, A.C.M. Brasileiro, F.J.L. Aragão, Repetitive somatic embryogenesis in cacao and optimisation of gene expression by particle bombardment, J. Plant Biotech. 4 (2002) 71–76.
[16] J.A. Driver, A.H. Kuniyuki, In vitro propagation of Paradox walnut rootstock, Hortscience 19 (1984) 507–509.
[17] K. Nakai, M. Kanehisa, A knowledge base for predicting protein localization sites in eukaryotic cells, Genomics 14 (1992) 897–911.
[18] K. Hofmann, P. Bucher, L. Falquet, A. Bairoch, The PROSITE database: its status in 1999, Nucl. Acids Res. 27 (1999) 215–219.
[19] S.F. Altschul, W. Gish, W. Miller, E.W. Myers, D.J. Lipman, Basic local alignment search tool, J. Mol. Biol. 215 (1990) 403–410.
[20] T.A. Hall, BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT, Nucl. Acids Symp. Ser. 41 (1999) 95–98.
[21] S. Kumar, K. Tamura, I.B. Jakobsen, M. Nei, MEGA2: molecular evolutionary genetics analysis software, Bioinformatics 17 (2001) 1244–1245.
[22] M. Kimura, A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences, J. Mol. Evol. 16 (1980) 111–120.
[23] N. Saitou, M. Nei, The neighbor-joining method: a new method for reconstructing phylogenetic trees, Mol. Biol. Evol. 4 (1987) 406–425.
[24] D. Posada, K.A. Crandall, Modeltest: testing the model of DNA substitution, Bioinformatics 14 (1998) 817–818.
[25] L.L. Cavalli-Sforza, A.W.F. Edwards, Phylogenetic analysis: models and estimation procedures, Am. J. Hum. Genet. 19 (1967) 122–257.
[26] J. Felsenstein, Evolutionary trees from DNA sequences: a maximum likelihood approach, J. Mol. Evol. 17 (1981) 368–376.
[27] F. Ronquist, J.P. Huelsenbeck, MrBayes 3: Bayesian phylogenetic inference under mixed models, Bioinformatics 19 (2003) 1572–1574.
[28] J.P. Huelsenbeck, F. Ronquist, MrBayes: Bayesian inference of phylogeny, Bioinformatics 17 (2001) 754–755.
[29] J. Felsenstein, PHYLIP (Phylogeny Inference Package) version 3.6 (distributed by the author), Department of Genome Sciences, University of Washington, Seattle, 2004.
[30] M.J. Sanderson, H.B. Shaffer, Troubleshooting molecular phylogenetic analyses, Ann. Rev. Ecol. Syst. 33 (2002) 49–72.
[31] J. Felsenstein, Confidence limits on phylogenies: and approach using the bootstrap, Evolution 39 (1985) 783–791.
[32] B. Efron, E. Harlloran, S. Holmes, Bootstrap confidence levels for phylogenetic trees, Proc. Natl. Acad. Sci. U.S.A. 93 (1996) 13429–13434.
[33] A. Rzhetsky, M. Nei, A simple method for estimating and testing minimum-evolution trees, Mol. Biol. Evol. 9 (1992) 945–967.
[34] C. Thomas, D. Meyer, C. Himber, A. Steinmetz, Spatial expression of a sunflower *SERK* gene during induction of somatic embryogenesis and shoot organogenesis, Plant Physiol. Biochem. 42 (2004) 35–42.
[35] J.M. Stone, J.C. Walker, Plant protein kinase families and signal transduction, Plant Physiol. 108 (1995) 451–457.
[36] M.R. Tucker, A.C.G. Araujo, N.A. Paech, V. Hecht, Ed. D.L. Schmidt, J.B. Rossell, S.C. de Vries, A.M.G. Koltunow, Sexual and apomictic reproduction in *Hieracium* subgenus pilosella are closely interrelated developmental pathways, Plant Cell. 15 (2003) 1524–1537.
[37] C.L. Tan, D.B. Furtek, Development of an in vitro regeneration system for *Theobroma cacao*, from mature tissues, Plant Sci. 164 (2003) 407–412.
[38] W.H. Li, Molecular Evolution, Sinauer Associates, Sunderland, MA, USA, 1997, 487 pp.
[39] E.F. Vanin, Pseudogenes: characteristics and evolution, Annu. Rev. Genet. 19 (1985) 253–272.
[40] P.M. Harrison, M. Gerstein, Studying genomes through the aeons: protein families, pseudogenes and proteome evolution, J. Mol. Biol. 318 (2002) 1155–1174.