

Molecular cloning of a full-length cDNA and gene from *Coffea arabica* encoding a protein homologous to the yeast translation initiation factor SUI1: expression analysis in plant organs

Charlotte Gaborit, Victoria Caillet, Alain Deshayes and Pierre Marraccini^{1*}

Department of Plant Science, Nestlé Research Centre, 101, avenue Gustave-Eiffel, B.P.9716, Notre Dame D'Oé, 37097 Tours Cedex 2, France; ¹ Present address: Instituto Agronômico do Paraná – IAPAR, CIRAD-CP AMG Laboratorio de Biotecnologia Vegetal, Rodovia Celso Garcia Cid, Km 375, CP 481, 86001-970, Londrina, Paraná, Brasil. * Corresponding author: marraccini@cirad.fr
Received: 10/03/2003. Accepted: 27/03/2003.

A full-length cDNA (*CaSUI1*) was isolated from a *Coffea arabica* cDNA library from beans during maturation. Its putative translational product is highly homologous to the SUI1 protein of *Saccharomyces cerevisiae* which functions in concert with eIF2 and the initiator tRNA-Methionin in directing the ribosome to the proper start site of translation. The corresponding gene from coffee was also cloned and sequenced. Its organization is very similar to what observed for the same gene from rice (*Oryza sativa*) particularly regarding the position and length of its introns. The expression of *CaSUI1* gene was also checked and showed that it was highly expressed in all coffee tissues analyzed, therefore confirming the essential role of the SUI1 protein in cell housekeeping functions.

Key words: *Coffea arabica*, constitutive gene expression, GOS2, SUI1, translational initiation.

Clonagem de uma seqüência completa de cDNA e gene de *Coffea arabica* codificando uma proteína homóloga ao fator de tradução SUI1 de levedura: análise da expressão em órgãos da planta: Uma seqüência completa de cDNA (*CaSUI1*) foi isolada de uma biblioteca de cDNA de frutos de *Coffea arabica* coletados no estágio de maturação. O produto putativo de transcrição é altamente homólogo à proteína SUI1 de *Saccharomyces cerevisiae*, que funciona em combinação com eIF2 e o iniciador tRNA-Metionina no direcionamento do ribossomo para o sítio correto de iniciação da tradução. O gene correspondente de café também foi clonado e sequenciado. Sua organização é similar ao que foi descrito para o mesmo gene em arroz (*Oryza sativa*), particularmente em relação à posição e ao comprimento de seus introns. A expressão do gene *CaSUI1* também foi estudada, observando-se alta expressão em todos os tecidos de café analisados, confirmando, portanto, o papel essencial da proteína de SUI1 em funções básicas de manutenção celular.

Palavras-chave: *Coffea arabica*, expressão gênica constitutiva, GOS2, iniciação de tradução, SUI1.

Initiation of protein synthesis in eukaryotes involves the binding of the initiator methionyl-tRNA and mRNA to the 40S ribosomal subunit, followed by the association to the 60S subunit to form the 80S initiation complex. This process requires the participation of several initiation factors like eIF3 for example (Merrick, 1992). Genetic experiments in yeast (*S. cerevisiae*) identified a mutated form of the *SUI1* gene by its ability to restore the expression of *HIS4* whose AUG initiation codon was

altered (Castilho-Valavicius et al., 1990). Few years later, it was shown that the smallest subunit p16 corresponded to the SUI1 protein of the initiation factor eIF3 therefore confirming its essential role in protein synthesis (Naranda et al., 1996).

Here we report the isolation of a cDNA encoding a SUI1-homologous protein from *Coffea arabica*. It was isolated from a cDNA library of coffee beans at 27 weeks after flowering (WAF) that was initially screened to clone

the 11S-encoding cDNA (Rogers *et al.*, 1999). The full-length cDNA sequence was shown to be 678 nucleotides long and contained an open reading frame of 339 bp encoding a putative protein of 113 amino acids with a theoretical molecular mass of 12675 Da. Database searches revealed that this coffee protein is homologous to the SUI1 protein (Yoon and Donahue, 1992; Naranda *et al.*, 1996). From multiple sequence alignment, we observed that this coffee protein, hereafter named CaSUI1, had high identity (91%) with equivalent proteins from plants like rice and maize (figure 1A) but only 56% identity with SUI1 proteins from mammals (i.e. mouse, human) and yeast. The dendrogram of amino acid sequence relationships also reflected this observation (figure 1B).

By several reaction of inverse-PCR (not detailed here), we also cloned and sequenced a genomic DNA fragment of 2.7 kbp including the *CaSUI1* whole gene, as well as a

300 bp fragment upstream of the cDNA 5' end. The sequences of the *CaSUI1* cDNA and *CaSUI1* gene were deposited in the GenBank database under the accession numbers AJ519839 and AJ519840, respectively.

Despite the genetic distance separating the *Coffea* genus from *Oryza*, it is interesting to note that the *CaSUI1* gene contains four introns, which are localized exactly at the same position and have approximately the same length of those found in the rice *GOS2* gene (de Pater *et al.*, 1992). The presence of a long intron (1206 bp) in the 5' UTR is particularly interesting because it could be required for the expression of this gene as demonstrated for the leader intron of the rice *Ostub16* β -tubulin gene (Morello *et al.*, 2002). A Southern-blot analysis of *C. arabica* genomic DNA probed with the *CaSUI1* cDNA showed only one hybridizing band indicating that *CaSUI1* is probably unique in the genome, as also observed in rice (data not shown).

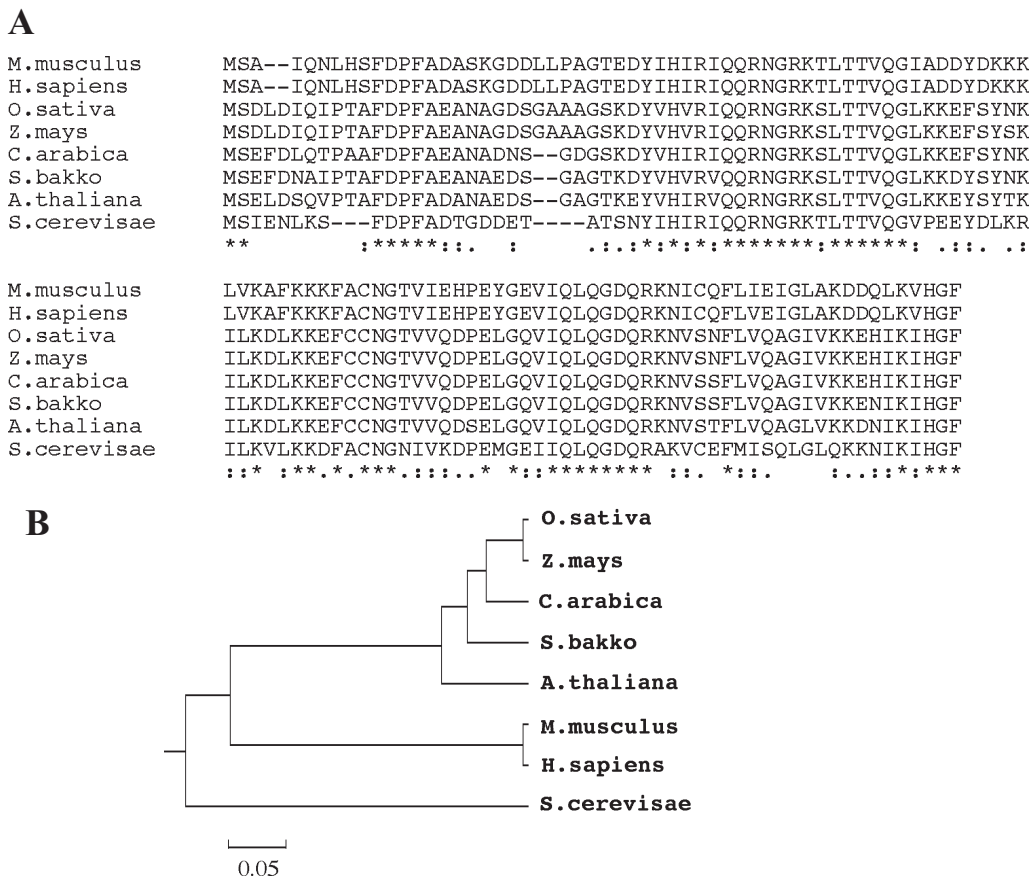


Figure 1. Sequence alignment of SUI1 proteins (A) from *M. musculus* (P48024), *Homo sapiens* (P41567), *O. sativa* (S21636), *Z. mays* (P56330), *C. arabica* (this work), *S. bakko* (O48650), *A. thaliana* (P41568) and *S. cerevisiae* (M77514) performed by ClustalW. Accession numbers of proteins are given in brackets. (B) Dendrogram of amino acid sequence relationship among the proteins aligned in (A). Distance along the horizontal axis separating two sequences is proportional to the divergence between the sequences.

Northern-blot analysis revealed that a single hybridisation band was clearly detected at approximately 700 bp (figure 2A), therefore indicating that the cloned cDNA of 678 bp corresponded probably to a full-length cDNA. The expression was also investigated by RT-PCR using primers internal to the coding sequence, leading to a *CaSUI1*-specific amplification product of 470 bp (figure 2C). From these experiments, we showed that the *CaSUI1* gene is highly expressed in all tissues of *C. arabica* L. cv. Caturra tested, as in leaves at various developmental stage, flowers buds, roots, seeds from mature fruits at 35 WAF and germinated seeds at twenty days after imbibition (DAI). Recent screening of coffee ESTs from the on going Brazilian Coffee Genome Project also revealed the presence of sequences homologous to *CaSUI1* sequence in cDNA libraries from all tissues (L.G.E. Vieira, personal communication), hence confirming the housekeeping function of the SUI1 protein (Fields and Adams, 1994). By comparing the intensities of the bands revealed after probing (figure 2A) with total RNA loaded on the same gel (figure 2B), it seems that the *CaSUI1* gene is highly expressed in mature beans (figure 2A, lane 7), at the time when the expression of α -galactosidase and storage-protein (11S) encoding genes is not detected (Marraccini et al., 1999, 2001). We suppose that this large accumulation of *CaSUI1* mRNA in mature coffee beans is a prerequisite for the intense transcription and protein synthesis required for the following germination of coffee seeds (Giorgini, 1988). We also observed that the sequences of *C. arabica* also recognized similar sequences in "robusta", i.e. somatic embryos from *C. canephora* 126 (figures 2A and 2C, lane 6). This could be explained by the fact that *C. arabica* was originated by hybridisation between *C. eugenioides* and *C. canephora* (Lashermes et al., 1999), but also by a high conservation of the SUI1-coding sequence during the evolution due to its essential role in the initiation of translation. The high similarities of plant SUI1 proteins (figure 1A) led us to systematically use the *GOS2* primers in RT-PCR tests (see legend in figure 2C) to evaluate the integrity of RNA samples extracted from various tissues of different coffee species during the construction of cDNA libraries (data not shown).

The inspection of the coffee genomic DNA fragment located upstream the 5' end of the *CaSUI1* cDNA revealed the presence of a TATA box and several TGACG motifs, also found in rice and known to function as a binding site of various plant nuclear factors (de Pater et al., 1992).

Experiments are on going to see if this sequence could function as a constitutive promoter in transformed tobacco plants.

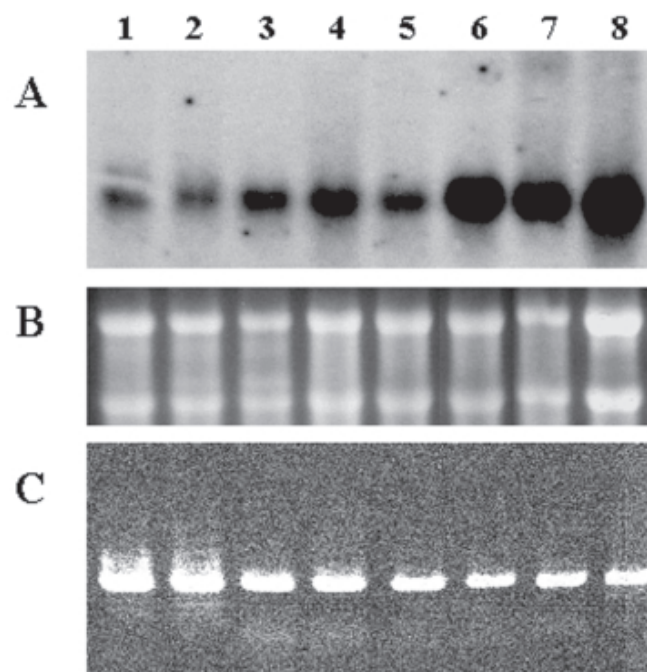


Figure 2. Analysis of *CaSUI1* gene expression. For the Northern blot (A), the probe used corresponded to the *CaSUI1* full-length cDNA which was ^{32}P labelled. Total RNA (10 ng) stained by ethidium bromide demonstrates homogeneity of well loading (B). For RT-PCR (C), total RNAs were incubated (RT step: 48°C, 45 min) with the primers GOS26 5' TACCCGACCCGAACCCCAATTTTC 3' and GOS27 5' ACACCAGATGAATGCACACTGTCTC 3', and further subjected to a PCR reaction for 35 cycles (94°C-90 sec, 55°C-90 sec, 72°C-2 min). PCR products were after separated on 1,2% agarose gel stained with ethidium bromide. Lanes: 1, young leaves (< 1 cm); 2, medium leaves (length 2-10 cm); 3, old leaves (> 10 cm); 4, flowers buds; 5, roots; 6, somatic embryos from *C. canephora*; 7, seeds from mature fruits (35 WAF); 8, germinated seeds (20 DAI). RNA samples were extracted from *C. arabica* L. cv. Caturra 2308 as described previously (Rogers et al., 1999), with the exception of somatic embryos RNA which were extracted from *C. canephora* 126 (kindly given by Dr T. Leroy).

Acknowledgments: We would like to thank Drs L.G.E. Vieira, W.J. Rogers and V. Pétiard for discussions and critical reading of the manuscript.

REFERENCES

- Castilho-Valavicius B, Yoon H, Donahue TF (1990) Genetic characterization of the *Saccharomyces cerevisiae* translational initiation suppressors *sui1*, *sui2* and *SUI3* and their effects on *HIS4* expression. *Genetics* 124:483-495.
- Fields C, Adams MD (1994) Expressed sequence tags identify a human isolog of the *SUII* translation initiation. *Biochem. Biophys. Res. Commun.* 198:288-291.
- Giorgini JF (1988) Ribonucleic acid characterization and synthesis in germinating coffee seed endosperm. *Braz. J. Med. Biol. Res.* 21:811-824.
- Lashermes P, Combes MC, Robert J, Trouslot P, D'Hont A, Anthony F, Charrier A (1999) Molecular characterisation and origin of the *Coffea arabica* L. genome. *Mol. Gen. Genet.* 261:259-266.
- Marraccini P, Deshayes A, Pétiard P, Rogers WJ (1999) Molecular cloning of the complete 11S seed storage protein gene of *Coffea arabica* and promoter analysis in transgenic tobacco plants. *Plant Physiol. Biochem.* 37:273-282.
- Marraccini P, Allard C, André M-L, Courjault C, Gaborit C, Lacoste N, Meunier A, Michaux S, Petit V, Priyono P, Rogers WJ, Deshayes A (2001) Update on coffee biochemical compounds, protein and gene expression during bean maturation and in other tissues. In: Proceedings of the 19th International Scientific Colloquium on Coffee, Trieste, Italy, Abstract B214 CD-rom ISBN 2-90012-18-9.
- Merrick WC (1992) Mechanisms and regulation of eukaryotic protein synthesis. *Microbiol. Rev.* 56:291-315.
- Morello L, Bardini M, Sala F, Breviaro D (2002) A long leader intron of the *Ostub16* rice β -tubulin gene is required for high-level gene expression and can autonomously promote transcription both *in vivo* and *in vitro*. *Plant J.* 29:33-44.
- Naranda T, MacMillan SE, Donahue TF, Hershey JWB (1996) SUI1/p16 is required for the activity of eukaryotic translation initiation factor 3 in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 16:2307-2313.
- de Pater S, van der Marck F, Rueb S, Katagiri F, Chua N-H, Schilperoort RA, Hensgens LAM (1992) The promoter of the rice gene *GOS2* is active in various different monocot tissues and binds rice nuclear factor ASF-1. *Plant J.* 2:837-844.
- Rogers WJ, Bézard G, Deshayes A, Pétiard V, Marraccini P (1999) Biochemical and molecular characterization and expression of the 11S-type storage protein from *Coffea arabica* endosperm. *Plant Physiol. Biochem.* 37:261-272.
- Yoon H, Donahue TF (1992) The *sui1* suppressor locus in *Saccharomyces cerevisiae* encodes a translation factor that functions during tRNA_i^{Met} recognition of the start codon. *Mol. Cell. Biol.* 12:248-260.