Available online at www.sciencedirect.com
SCIENCE DIRECT®
Phytochemistry 66 (2005) 1346–1357
PHYTOCHEMISTRY
www.elsevier.com/locate/phytochem

Recruitment of alkaloid-specific homospermidine synthase (HSS) from ubiquitous deoxyhypusine synthase: Does *Crotalaria* possess a functional HSS that still has DHS activity?☆

Niknik Nurhayati, Dietrich Ober *

Institut für Pharmazeutische Biologie, Technische Universität Braunschweig, Mendelssohnstrasse 1, D-38106 Braunschweig, Germany

Received 31 January 2005; received in revised form 4 April 2005
Available online 2 June 2005

Abstract

Quinolizidine alkaloids are the most prominent group of alkaloids occurring in legumes, except for many members of the tribe Crotalarieae that accumulate pyrrolizidine alkaloids (PAs). To study the evolution of PA biosynthesis as a typical pathway of plant secondary metabolism in this tribe, we have searched for a cDNA coding for homospermidine synthase (HSS), the enzyme catalyzing the first specific step in this biosynthesis. HSS was shown to have been recruited from deoxyhypusine synthase (DHS) by independent gene duplication in several different angiosperm lineages during evolution. Except for a cDNA sequence coding for the DHS of *Crotalaria retusa*, no data is available concerning the origin of PA biosynthesis within this tribe of the Fabaceae. In addition to several pseudogenes, we have identified one functional DHS in *C. scassellatii* and two in *C. juncea*. Despite *C. juncea* plants under study being devoid of PAs, we have found that the two sequences of *C. juncea* are different with respect to their genomic organization, their tissue-specific expression, and their biochemical activities. Supported by the branching pattern of a maximum likelihood analysis of these sequences, they have been classified as "class 1" and "class 2" DHS. It remains open whether the duplicated DHS belonging to class 2 is involved in the biosynthesis of PAs.

Keywords: Pathway evolution; Pyrrolizidine alkaloids; Homospermidine synthase; Deoxyhypusine synthase; Pseudogenes; Exon skipping; Splicing variants

1. Introduction

The evolution of pathways involved in plant secondary metabolism is far from being understood. We have used the pyrrolizidine alkaloid (PA) system to study the evolution of such a pathway under the assumed selection pressure of herbivory. PAs are a group of plant-defence chemicals (Hartmann, 1999; Hartmann and Ober, 2000) that are constitutively produced by the species of only a few families within the

☆ Sequence data from this article have been deposited with the EMBL/GenBank data libraries under the following accession numbers. *Crotalaria scassellatii* DHS cDNA (DHS-*Cs*, AJ968375); *C. juncea* DHS1 cDNA (DHS1-*Cj*, AJ968376); *C. juncea* DHS2 cDNA (DHS2-Cj, AJ968377); *C. juncea* DHS1 pseudogenes (Y1 *dhs1-Cj*, AJ968380; Y2 *dhs1-Cj*, AJ968379); *C. juncea* DHS2 pseudogene (Y1 *dhs2-Cj*, AJ968378); *C. juncea dhs1* gene (*dhs1-Cj*, AJ968383); *C. juncea dhs2* genes (*dhs2a-Cj*, AJ968384; *dhs2b-Cj*, AJ968385; *dhs2c-Cj*, AJ968386); *C. retusa* DHS pseudogenes (Y1 *dhs-Cr*, AJ968381; Y2 *dhs-Cr*, AJ968382).

Abbreviations: DHS, deoxyhypusine synthase; DOP-PCR, degenerate oligonucleotide-primed polymerase chain reaction; eIF5A, eukaryotic initiation factor 5A; HSS, homospermidine synthase; ORF, open reading frame; PA, pyrrolizidine alkaloid; RT-PCR, reverse transcription-polymerase chain reaction; UTR, untranslated region; GSP, gene-specific primer.

* Corresponding author. Tel.: +49 531 391 5690; fax: +49 531 391 8104.
E-mail address: d.ober@tu-bs.de (D. Ober).

0031-9422/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.phytochem.2005.04.013

angiosperms. Approximately 95% of the PA-containing species studied thus far belong to the four distant families of the Asteraceae (tribes Senecioneae and Eupatorieae), Boraginaceae (many genera), Fabaceae (mainly the genus *Crotalaria*), Orchidaceae (nine genera), and Apocynaceae. Isolated occurrences have also been described for some further families, such as the Convolvulaceae, Ranunculaceae, and Celastraceae (Hartmann and Witte, 1995).

We have recently shown that the first specific enzyme of PA biosynthesis, homospermidine synthase (HSS), was recruited from deoxyhypusine synthase (DHS), an enzyme involved in the post-translational activation of eukaryotic initiation factor 5A (eIF5A) (Ober and Hartmann, 1999a; Reimann et al., 2004). Two post-translational modifications are needed to activate eIF5A, of which the first is catalyzed by DHS, which transfers the aminobutyl moiety of spermidine to a specific protein-bound lysine residue of the eIF5A precursor protein. This lysine residue is modified to the amino acid deoxyhypusine, which is hydroxylated in the second step of the activation procedure to hypusine (Park et al., 1997). The modification of the eIF5A precursor protein by DHS is one of the most specific post-translational modifications known (Krishna and Wold, 1993; Yan et al., 1989). The mechanism is highly conserved within eukaryotes and in Archaea (Bartig et al., 1992; Gordon et al., 1987) not only in terms of function, but also with regard to sequence identity. We have therefore been able to identify DHS-coding and HSS-coding sequences in various plant species by the use of degenerate oligonucleotide-primed polymerase chain reaction (DOP-PCR) (Ober and Hartmann, 1999a,b; Ober et al., 2000; Reimann et al., 2004).

Recently, a total of 23 cDNA sequences coding for DHS and HSS of various angiosperm species, including 17 of species that are known to possess PAs, have been used for phylogenetic analyses. The results indicate at least four independent recruitments of HSS within the evolution of the angiosperms (Reimann et al., 2004). Quinolizidine alkaloids are the most prominent group of alkaloids in legumes, with exception of the tribe Crotalarieae (Wink and Mohamed, 2003), which contain a large number of PA-producing species. With regard to this tribe, only a cDNA-sequence identified as DHS is available from *Crotalaria retusa* (family Fabaceae). To study the evolutionary origin of HSS within the Fabaceae family, we have attempted to identify sequences that are homologous to DHS and that code for the HSS involved in the biosynthesis of PAs in *Crotalaria* species.

2. Results and discussion

2.1. Amplification of sequences homologous to DHS from various Crotalaria species

DHS and HSS cDNA sequences of various angiosperm species show a high degree of sequence identity (Reimann et al., 2004). Therefore, we used degenerate oligonucleotide-primed polymerase chain reaction (DOP-PCR) to search for sequences within *Crotalaria* species with homology to previously known DHS and HSS sequences. The primers P1–P4 were designed according to an alignment of various sequences, as described previously (Ober and Hartmann, 1999a,b). Because root cultures of various *Crotalaria* species are unable to produce PAs (Hartmann et al., 1989; Hartmann, 1994), and because analysis of *C. retusa* roots has not resulted in the identification of an HSS, we focused on the aerial parts of *C. retusa*, *C. anagyroides*, *C. juncea*, and *C. scassellatii* for RNA isolation and DOP-PCR. All identified sequences showed a high

Table 1
Nucleic acid sequence identity (%) among cDNA clones isolated from various species of *Crotalaria* and the length of their overlapping sequences (bp)

Clone	i*Cr*S	i*Ca*S	i*Cs*S	i*Cj*S1	i*Cj*S2
DHS-*Cr*	99.5% 374 bp	95.0% 502 bp	96.5% 658 bp	97.0% 439 bp	90.7% 355 bp
i*Cr*S		95.2% 375 bp	95.7% 375 bp	97.4% 312 bp	90.4% 313 bp
i*Ca*S			99.0% 502 bp	93.9% 439 bp	90.7% 355 bp
i*Cs*S				94.8% 439 bp	91.6% 355 bp
i*Cj*S1					90.4% 355 bp

Sequences generated by DOP-PCR were shortened by 20 nucleotides at both ends prior to comparison to avoid artifacts introduced by the degenerate primers.

DHS-*Cr*, 1116 bp DHS *C. retusa* (Reimann et al., 2004); i*Cr*S, 415 bp internal cDNA fragment isolated from *C. retusa* shoot; i*Ca*S, 542 bp internal cDNA fragment isolated from *C. anagyroides* shoot; i*Cs*S, 698 bp internal cDNA fragment isolated from *C. scassellatii* shoot; i*Cj*S1, 479 bp internal cDNA fragment isolated from *C. juncea* shoot; i*Cj*S2, 395 bp internal cDNA fragment isolated from *C. juncea* shoot.

degree of sequence identity (94–100%) when they were compared with each other and with the previously identified DHS of *C. retusa* (DHS-*Cr*, Table 1). Only one clone isolated from *C. juncea* shoot tips (termed i*Cj*S2) was an exception, because it showed a lower sequence identity of only 90–92% to all other sequences identified from *Crotalaria* species, including a sequence of the same species (i*Cj*S1).

The high sequence identity between the sequences identified from the various *Crotalaria* species prompted us to try the direct amplification of the complete coding regions of the putative DHS of the different *Crotalaria* species by using the primers designed for the expression of the DHS of *C. retusa* (P9/P10, Reimann et al., 2004). We were able to amplify the open-reading frame (ORF) of the DHS of *C. juncea* and *C. scassellatii* (DHS1-*Cj* and DHS-*Cs*, respectively) and cloned each into the pET3a expression vector. Both sequences showed high identity to the previously known DHS of *C. retusa* of 97.5% and 96.1%, respectively, at the nucleic acid level and of 97.6% and 96.5%, respectively, at the amino acid level. The sequence information of fragment i*Cj*S2 that differed significantly from the other sequences was used to generate gene-specific primers for the amplification of the 3′-end and the 5′-end of the cDNA. The full-length cDNA contained an ORF of 1152 bp, a 5′ untranslated region (UTR) of 34 bp, and a 3′UTR of 492 bp. The coding region of DHS2-*Cj* showed sequence identity of 88.6% and 89.2% to DHS1-*Cj* and DHS-*Cr* at the nucleic acid level, respectively, and 88.7% at the amino acid level. The ORF of DHS2-*Cj* was cloned into the expression vector pET3a-mod by using the gene-specific primers P12/P13, and all newly identified DHS-like sequences were heterologously expressed in *Escherichia coli* for biochemical characterization.

2.2. Biochemical characterization of DHS homologs of *Crotalaria*

Table 2 summarizes the biochemical activities of the newly identified sequences. The assay conditions did not allow quantitative conclusions to be made, because desalted protein extracts of the *E. coli* cells were used without quantification of the level of heterologous protein expression in two independent experiments. As positive controls, the previously identified DHS of *C. retusa* and the HSS of *Phalaenopsis* (Reimann et al., 2004) were used for the DHS and HSS assays, respectively. The data clearly showed that, with exception of the HSS of *Phalaenopsis* (used as reference), all DHS-like sequences identified in *Crotalaria* were able to catalyze the formation of protein-bound deoxyhypusine. Unlike all plant DHS identified to date, the proteins encoded by the DHS-like sequence of *C. scassellatii* (DHS-*Cs*) and of one of the sequences identified from *C. juncea* (DHS1-*Cj*) were unable to catalyze the formation of homospermidine. In contrast, the protein DHS2-*Cj* that was also identified in *C. juncea* and that exhibited conspicuous sequence differences from all other DHS-like sequences of *Crotalaria*, showed high HSS activity and distinct DHS activity.

2.3. Determination of PA concentrations in the Crotalaria species under study

None of the DHS-like sequences from the various *Crotalaria* species had the characteristic features of a HSS showing only HSS but not DHS activity. Therefore we checked the PA content of the *Crotalaria* plants that were used for RNA extraction after they had germinated and been grown under artificial conditions (see Experimental). Plantlets of *C. scassellatii* and of *C. juncea* grew well and survived for more than one year, with *C. juncea* actually reaching the flowering stage. In contrast, *C. retusa* survived for only approximately three months after germination. Table 3 summarizes the results. The seeds examined were of the same batch from which the plantlets were grown. *C. scassellatii* seeds contained 18 mg/g PAs, a value that is well in accordance with previous data (Toppel et al., 1988). PAs were also detected in the cotyledons and the whole plant of *C. retusa*, and in the leaves of a two-year-old *C. scassellatii*. Surprisingly, all analyzed parts of *C. juncea* (cotyledons, leaves, flowers, and roots) were PA-free though this plant is known to contain PAs (Hartmann and Witte, 1995). Because of these results, we repeated our search for an HSS-coding

Table 2
DHS and HSS activities of heterologously expressed cDNAs identified from various *Crotalaria* species

CDNA	HSS activity (%)[a]		DHS activity (cpm)[b]	
	Experiment 1	Experiment 2	Experiment 1	Experiment 2
DHS-*Cr*	6.0		9006	
DHS-*Cs*	0.0		5584	
DHS1-*Cj*	0.0	0.0	4555	2588
DHS2-*Cj*	–	22.7		1381
HSS-*Ph*	49.6	67.9	102	106

Two independent experiments were performed.
[a] HSS activity: % radioactivity incorporated into homospermidine upon feeding of [^{14}C]putrescine (100%) (incubation time: 7.5 min).
[b] DHS activity: radioactivity incorporated into protein-bound deoxyhypusine (incubation time: 60 min). 100–150 cpm represents the background noise for the DHS assay.

Table 3
Alkaloid content (mg/g) in diverse tissues of various *Crotalaria* species

	Seeds	Cotyledons	Whole plant	Leaves	Flowers	Roots
C. retusa (2 months old)	45	33	86	–	–	–
C. juncea (6 months old)	–	n.d.	n.d.	n.d.	n.d.	n.d.
C. scassellatii (2 years old)	18	–	–	1.9	–	–

n.d., Not detectable.
–, Not analyzed.

Table 4
Amino acid identities (%) of DHS-cDNA clones isolated from various *Crotalaria* species in comparison with each other and with other plant DHS

CDNA clones	% Amino acid identity			
	DHS-*Cr*	DHS-*Cs*	DHS1-*Cj*	DHS2-*Cj*
DHS-*Cs*	96.5			
DHS1-*Cj*	97.6	95.4		
DHS2-*Cj*	86.8	87.2	86.8	
DHS-*Sv*	82.5	80.2	80.2	76.7
HSS-*Sv*	74.5	73.4	80.8	74.0
DHS-*Nt*	77.9	79.0	77.0	76.0
DHS-*So*	75.9	77.2	76.1	74.5
HSS-*So*	69.9	70.2	69.3	70.0
DHS-*Ma*	73.7	74.7	73.4	71.3
HSS-*Ph*	71.2	72.0	71.0	69.9

DHS-*Cr*, DHS *Crotalaria retusa* (AJ704838); DHS-*Cs*, DHS *Crotalaria scassellatii*; DHS1-*Cj*S, DHS1 from *Crotalaria juncea*; DHS2-*Cj*S, DHS2 from *Crotalaria juncea*; DHS-*Sv*, DHS *Senecio vernalis* (AJ238622); HSS1-*Sv*, HSS1 *Senecio vernalis* (AJ238623); DHS-*Nt*, DHS *Nicotiana tabacum* (AJ242017); DHS-*So*, DHS *Symphytum officinale* (AJ704852); HSS-*So*, DHS *Symphytum officinale* (AJ704851); DHS-*Ma*, DHS *Musa acuminata* (AF296079); HSS-*Ph*, HSS *Phalaenopsis* spec. (AJ704848).

sequence by DOP-PCR with various tissues of the two *Crotalaria* species that produced PAs under our growth conditions, namely *C. scassellatii* and *C. retusa*. However, we were only able to amplify sequences identical or nearly identical to those previously identified.

2.4. Sequence comparison of the DHS of various Crotalaria species with the DHS of other plants

The deduced amino acid sequences of the DHS of *C. retusa*, *C. scassellatii*, and *C. juncea* were compared one with another and with several other previously published plant DHS sequences (Ober and Hartmann, 1999a,b; Reimann et al., 2004) (Table 4). Within the sequences of *Crotalaria*, we found two classes of sequences. The first class comprises DHS-*Cr*, DHS-*Cs*, and DHS1-*Cj*. These sequences share identities of more than 95% on the amino acid level. The second class consists, so far, only of DHS2-*Cj*, which shares only sequence identities of about 87% with the sequences of class one. Both classes show a similar degree of divergence from the DHS and HSS sequences of other angiosperm species. This suggests that the structure of DHS evolved in the genus *Crotalaria* into different forms. Fig. 1 shows a maximum likelihood tree of the four *Crotalaria* sequences by using the DHS-coding sequences of *Arabidopsis thaliana* and *Dianthus caryophyllus* as the outgroup. The branching pattern of the tree suggests that the two sequence classes separated before the speciation events of the *Crotalaria* species included in this study. The amino acid alignment of the four *Crotalaria* DHS (Fig. 2) shows the characteristic feature of DHS2-*Cj*, which possesses a threonine-rich and

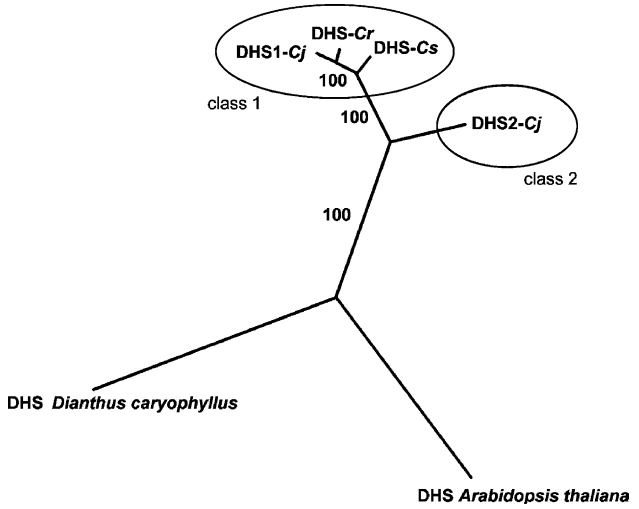

Fig. 1. Unrooted maximum likelihood tree of cDNA sequences coding for DHS from *C. retusa* (DHS-*Cr*), *C. scassellatii* (DHS-*Cs*), *C. juncea* (DHS1-*Cj* and DHS2-*Cj*), and from *D. caryophyllus* and *A. thaliana* (used as outgroup). Bootstrap proportions resulted from 1000 replicates.

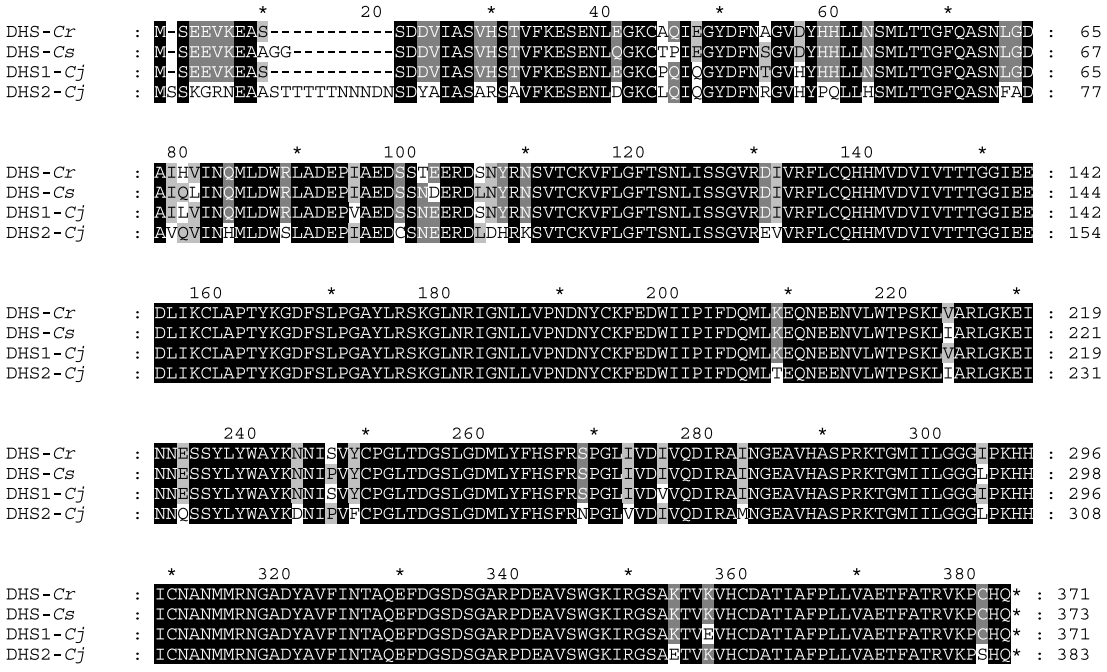

Fig. 2. Alignment of amino acid sequences of DHS from *C. retusa* (DHS-*Cr*), *C. scassellatii* (DHS-*Cs*), and *C. juncea* (DHS1-*Cj* and DHS2-*Cj*).

asparagine-rich region at the N-terminal end of the sequence. To test for a putative N-terminal targeting sequence, we compared the two sequences DHS1-*Cj* and DHS2-*Cj* by using TargetP and PSORT software. Whereas TargetP predicted no N-terminal targeting sequence known for chloroplasts, mitochondria, or the secretory pathway, PSORT suggested a localization of DHS1-*Cj* in the cytoplasm, whereas DHS2-*Cj* was predicted to be localized within the chloroplast stroma. We tested the putative N-terminal signal peptide of DHS2-*Cj* by cloning the ORF of DHS2-*Cj* in frame with the GFP gene of a pBSK-based vector as described by Bauer et al. (2004). After transfer of the construct by a particle gun into the guard cells of tobacco and *C. retusa* leaves, the GFP was localized in the cytosol, suggesting that the sequence did not contain any signals for intracellular targeting (data not shown).

2.5. Tissue-specific expression of DHS1 and DHS2 of C. juncea

In order to study the expression of the genes coding for DHS1-*Cj* and DHS2-*Cj* in various organs of *C. juncea*, we performed reverse transcription-PCR (RT-PCR) experiments with the primer pairs P10/P11 and P12/P13 (Table 6), respectively. Fig. 3(a) and (b) suggests that the genes *dhs1-Cj* and *dhs2-Cj* are expressed in all the investigated organs of *C. juncea*. This observation is in good agreement with previously obtained data concerning the expression of DHS in tobacco plants, *Senecio vernalis* (Moll et al., 2002), and *Eupatorium cannabinum* (Anke et al., 2004). However, the finding of an intronless gene with a high degree of identity to the cDNA coding for DHS2 (termed *dhs2a-Cj*, see below) implies that this genomic sequence is amplified during RT-PCR instead of the cDNA coding for DHS2. For this reason, we repeated the RT-PCR analysis of DHS2 with an exclusive amplification of cDNA ends (EXACT) RT-PCR approach according to Smith et al. (2001) to avoid any coamplification of *dhs2a-Cj* attributable to possible contamination of the template cDNA by genomic DNA. The expression of the gene coding for DHS2-*Cj* was now only detectable at a low level in the roots, leaves, and shoots, but not in flower buds and open flowers (Fig. 3(c)). Because the 3′-UTR is amplified together with the coding region during EXACT RT-PCR, the PCR product in bigger than that in Fig. 3(a) and (b).

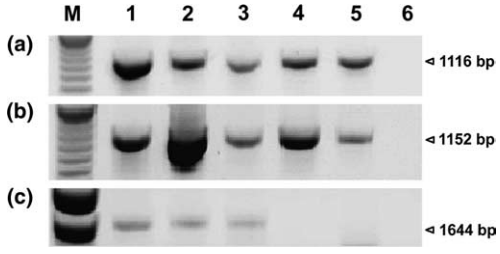

Fig. 3. RT-PCR expression analysis of DHS1 (a) and DHS2 (b) and EXACT RT-PCR analysis of DHS2 (c) with cDNA from various tissues of *C. juncea*. M, 100 bp ladder; 1, roots; 2, leaves; 3, shoot tips; 4, flower buds; 5, flowers; 6, water.

2.6. Analysis of genomic DNA coding for DHS in *C. juncea*

The identification of two different cDNAs in *C. juncea* coding for DHS prompted us to analyze the structure of the corresponding genes. For this purpose, we used the primers previously designed for the amplification of the whole ORF of DHS1-*Cj* (P10/P11) and DHS2-*Cj* (P12/P13) with genomic DNA as the template. In the case of the primer pair P10/P11, a 3432 bp-fragment was amplified comprising the whole *dhs1-Cj* gene consisting of 7 exons that were 99% identical to the cDNA sequence coding for DHS1-*Cj*. The amplification with primer pair P12/P13 resulted in a fragment of

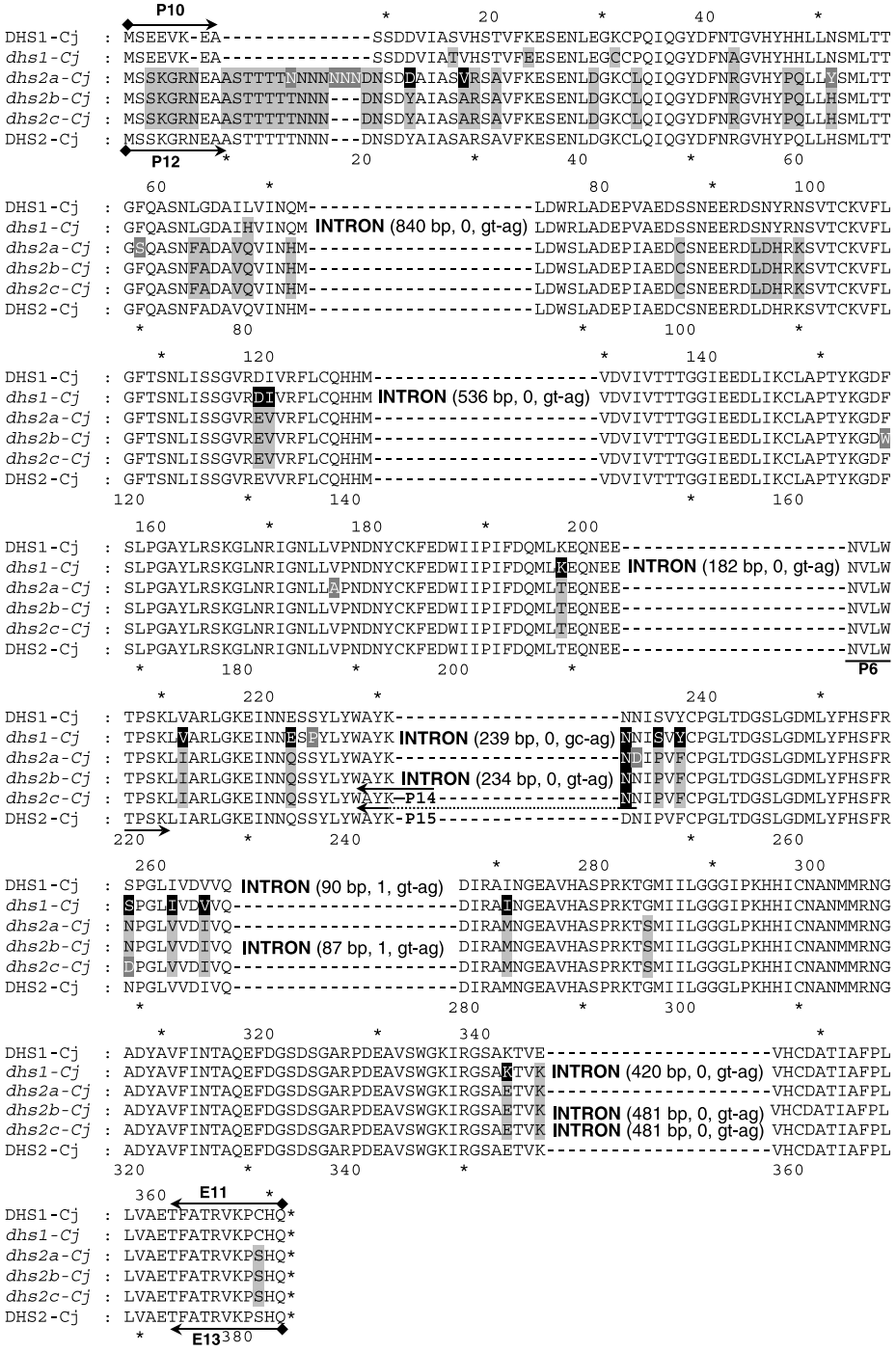

Fig. 4. Deduced amino acid sequences of *C. juncea dhs*1 and *dhs*2 genes in comparison with that of their corresponding cDNA. Amino acids deviating from DHS1-*Cj* and DHS2-*Cj* and from both cDNA sequences are marked as X, X and X, respectively, where X is a general amino acid. Arrows indicate positions of GSPs used for PCR amplification. The presence of introns is indicated together with the length of the intron, their phase, and their 5′/3′ splice-site motive. Numbering refers to the amino acid sequences of DHS1-*Cj* (top) and DHS2-*Cj* (bottom).

1164 bp that showed 99% sequence identity to the cDNA coding for DHS2-*Cj*. This sequence lacked any introns but contained nine additional nucleotides, coding for three additional asparagine residues in the threonine-rich and asparagine-rich region at the *N*-terminus (*dhs2a-Cj*, Fig. 4). Whether this gene is transcribed is unknown, because a cDNA with this characteristic property was never found. We thus continued our search for the genomic sequence encoding DHS2-*Cj* with a combination of the primer P12 with primer P6 that was previously designed for the identification of the 3′-end of the DHS2-*Cj*-cDNA. Amplification resulted in two products of 1.1 and 1.3 kb that both showed 99% sequence identity with the cDNA encoding DHS2-*Cj* and that possessed three and two introns, respectively. To identify the 5′-ends of the genomic DNA sequences, we designed primer P14 by using the exon-intron-boundary of the intron that is exclusively present in the 1.1-kb fragment and primer P15 corresponding to the region of the 1.3-kb fragment lacking this specific intron. Both amplifications resulted in identical fragments of 730 bp without any additional introns and without the insertion of nine nucleotides as observed in *dhs2a-Cj*. The resulting genes *dhs2b-Cj* and *dhs2c-Cj* contained three and one intron, respectively, and showed, in their coding regions, 99% nucleotide sequence identity to the cDNA coding for DHS2-*Cj*. An alignment of the deduced amino acid sequences of the coding regions of these genes and of the DHS-encoding cDNAs is shown in Fig. 4. In all genes that contain introns, their position and phase is conserved, as is their sequence. Hence, the sequence of introns 4, 5, and 6 of *dhs1-Cj* share 77–80% identity to the introns at the corresponding position in *dhs2b-Cj* and *dhs2c-Cj*. The only intron of sequence *dhs2c-Cj* is identical to intron 3 at the corresponding position in sequence *dhs2b-Cj*. All introns identified possess the consensus 5′/3′ splice-site motif GT/AG, except for intron 4 of *dhs1-Cj*, which has the dinucleotide GC as a donor site instead of GT (Fig. 4). The use of GC instead of GT at the donor site is observed in the order of ~1% of the introns in *A. thaliana* (Korning et al., 1996) and was also found in myrosinase genes of the Brassicaceae (Xue and Rask, 1995). With the identification of four different genes of which at least two code for functional DHS, *C. juncea* is the first eukaryotic organism in which more than one gene coding for DHS has been identified. All genome projects of eukaryotic species completed so far have identified the *dhs* gene as a single-copy gene.

2.7. Identification of pseudogenes from C. juncea and C. retusa

During our efforts to identify a cDNA coding for HSS of *C. juncea* and *C. retusa*, some partial cDNA clones were identified with characteristics of pseudogenes (Table 5). In *C. retusa*, two clones covering the 3′-end of the cDNA were identified that both contained deletions (Ψ1*dhs-Cr* and Ψ2*dhs-Cr*). Because the genomic sequence coding for DHS1-*Cr* is not available, we compared both pseudogenes with the genomic sequence of DHS1 from *C. juncea* (*dhs1-Cj*). Whereas Ψ1*dhs-Cr* lacks the first 63 nucleotides of the coding region corresponding to exon 6 in *dhs1-Cj*, Ψ2*dhs-Cr* has lost the whole exon 7 (Fig. 5(a)). Omission of exons has also been reported in *Arabidopsis* mutants (Sieburth et al., 1995) and maize mutants (Lal et al., 1999a,b). This omission of exons, termed exon skipping, is considered to be the result of a mutational alteration of either 5′ or 3′ splice-sites of introns that activate cryptic splice-sites in close proximity to the mutated site (Lal et al., 1999b). The exon skipping in Ψ2*dhs-Cr* may have occurred because of an alteration in the 3′ splice-site of the last intron (corresponding to intron 6 in *dhs1-Cj*) that activates the cryptic splice-site 34 nucleotides downstream of the stop codon in the 3′UTR of Ψ2*dhs-Cr*.

A cDNA exhibiting exon skipping has also been found in *C. juncea* by amplification of the full-length ORF (Ψ2*dhs1-Cj*). Comparison with *dhs1-Cj* has shown that this cDNA lacks 99 nucleotides within exon 6, resulting in a loss of 33 amino acids (Fig. 5(b)). A mutation may have created a new 3′ splice-site resulting in an additional intron, or this cDNA may be a splice variant of the *dhs1-Cj* gene. Because this deletion does not cause a frame shift mutation, we expressed this cDNA in *E. coli* to test it for catalytic activity in the DHS and HSS assays. However, the resulting recombinant protein of 37 kDa had neither DHS nor HSS activity (data not shown). Splice variants are also known from human DHS, which is encoded by a single-copy gene; three transcriptional variants are known of which only one encodes a functional DHS. These variants differ in the kind of internal deletions that are present in exon 7 (Mantuano et al., 1998). Interestingly, exon 7 in the human *dhs* gene corresponds to exon 6 in the *dhs-Cj* gene. The shorter isoforms of DHS might act as modulating factors of DHS activity (Mantuano et al., 1998). Indeed, a regulating role has also been suggested for pseudogenes. So, recently, Hirotsune et al. (2003) have reported

Table 5
Comparative analyses between several notable cDNA sequences cloned from *Crotalaria* and the *dhs1-Cj* genomic DNA sequence from *C. juncea*

Clone	Insertion (nt)	Deletion (nt)	Position[a]
Ψ1*dhs-Cr*		63	Exon 6
Ψ2*dhs-Cr*		116	Exon 7
Ψ1*dhs1-Cj*	89		Intron 4
Ψ1*dhs2-Cj*	55		Intron 1
Ψ2*dhs1-Cj*		99	Exon 6

[a] Relative to the intron/exon structure of the *dhs1-Cj* genomic sequence.

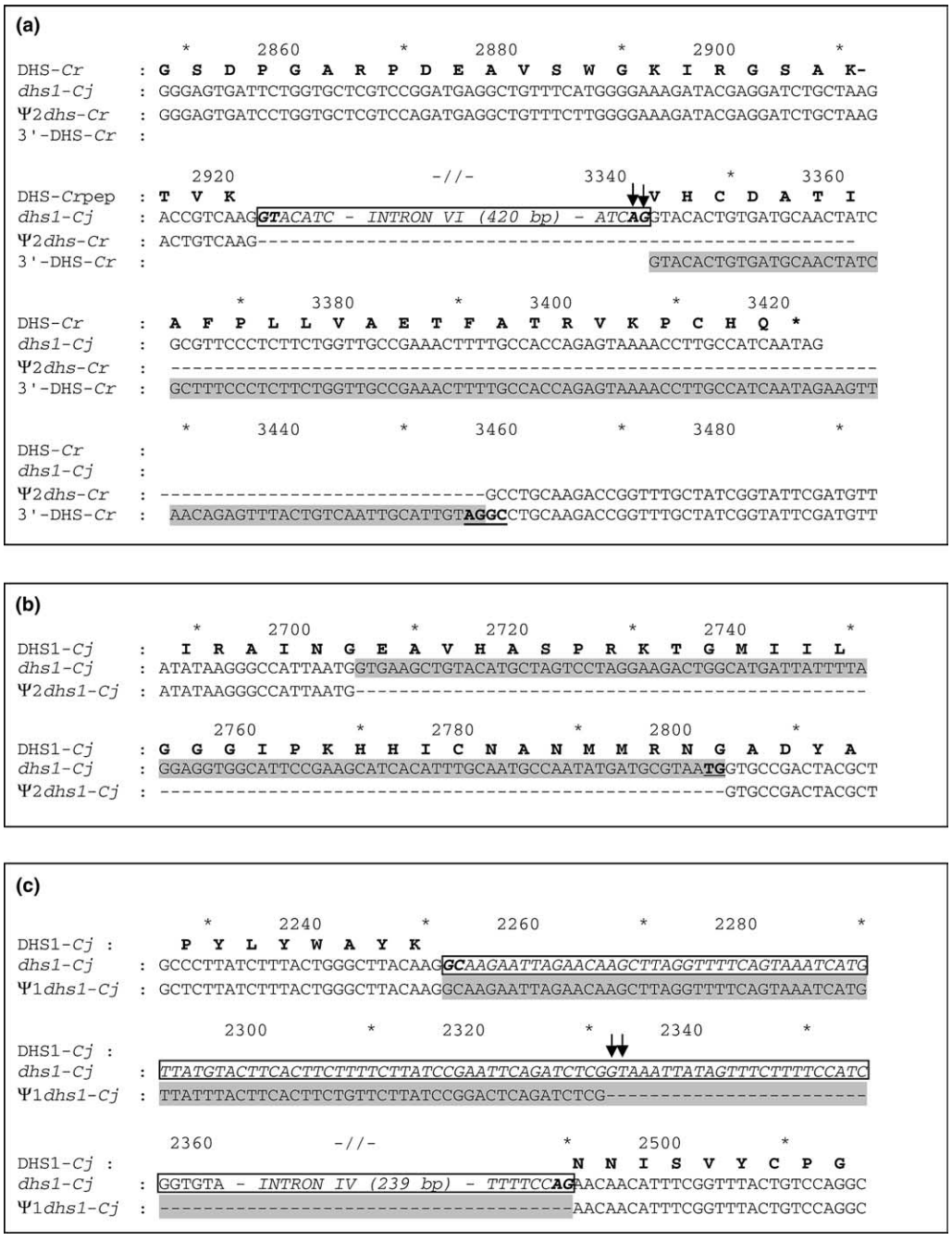

Fig. 5. Sequence analysis of *Crotalaria* cDNA clones with characteristics of pseudogenes. (a) Exon skipping resulted in a loss of 116 nucleotides (gray box) within Ψ2*dhs-Cr* encompassing the complete last exon and the beginning of the 3′-untranslated region. The sequence is aligned with the amino acid sequence of the DHS of *C. retusa* (DHS-*Cr*), with the genomic sequence *dhs1-Cj* from *C. juncea*, and with the cDNA coding for DHS-*Cr*, including the 3′UTR. The inactive 3′-splice-site of the last intron (open box) is indicated by two arrows, and the activated cryptic 3′ splice-site 34 nucleotides downstream of the stop codon is underlined. (b) In sequence Ψ2*dhs1-Cj*, exon skipping resulted in the loss of 99 nucleotides of the coding region (gray box). The two nucleotides within the genomic sequence *dhs1-Cj* that may have mutated to a new 3′ splice-site in the gene coding for Ψ2*dhs1-Cj* are underlined. (c) Sequence Ψ1*dhs1-Cj* with an intron-like insertion of 89 nucleotides (gray box) aligned with the genomic sequence *dhs1-Cj* and the predicted amino acid sequence encoded by the exons of *dhs1-Cj*. The new activated 5′ splice-site within the intron 4 of *dhs1-Cj* (open box) is labeled by two arrows. The numbering above the alignments refers to the genomic sequence *dhs1-Cj*.

the role of an expressed pseudogene in regulating the mRNA stability of its homologous coding region.

The alteration of splice-sites may also be responsible for the occurrence of two further cDNAs identified in *C. juncea* as containing additional sequences (Ψ1*dhs1-Cj* and Ψ1*dhs2-Cj*). The sequence of Ψ1*dhs1-Cj* is 100% identical to the cDNA coding for DHS1-*Cj*, with the exception of 63 additional nucleotides. The position

and the phase of this intron-like sequence corresponds to that of intron 4 in *dhs1-Cj*. Obviously, this insertion is derived directly from intron 4, because it shares 95% sequence identity with the first 80 bp of the intron, including the unusual 5′ splice-site consensus sequence GC (Fig. 5(c)). This cDNA is probably a splice variant of a gene closely related to *dhs1-Cj* in which a cryptic 5′ splice-site (GT) 63 nucleotides downstream of the original GC-containing splice-site is activated within the intron. As a result, 63 additional nucleotides of intron 4 are included into the cDNA (Fig. 5(c)). The cDNA Ψ1*dhs2-Cj* is 99.3% identical to the cDNA coding for DHS2-*Cj* but contains 55 additional nucleotides in the same position and with the same phase as intron 1 in the *dhs1-Cj* gene, an intron not found within the *dhs2-Cj* gene. Because the minimum length of functional introns that have to form a lariat during splicing is about 64 nucleotides (Goodall and Filipowicz, 1990), this intron-like sequence may be too short to be efficiently spliced.

Thus, *C. juncea* is the first eukaryotic organism in which more than one gene coding for DHS has been identified. All genome projects of eukaryotic species completed to date have identified the *dhs* gene as a single-copy gene. Several pseudogenes with homology to the DHS-coding genes have also been identified, indicating that duplication and diversification is an ongoing process within the genus *Crotalaria*.

3. Conclusion

In order to identify the key-enzyme HSS in PA biosynthesis in *Crotalaria*, we tried to find a cDNA coding for a protein that has homology to DHS but that has lost its ability to catalyze the DHS-specific modification of the eIF5A precursor protein without losing its ability to catalyze the formation of homospermidine. This is a characteristic property of all HSSs identified to date (Reimann et al., 2004). Analyzing various species of the genus *Crotalaria*, we were able to identify several cDNA sequences coding for functional DHS, the corresponding genomic sequences, and several cDNA variants with characteristic properties of pseudogenes. The finding that no sequence could be unequivocally identified as HSS leads to the following two hypotheses:

1. DHS2-*Cj*, identified in *C. juncea*, may be involved in PA biosynthesis as an HSS. In this case, the HSS has retained its original DHS activity. This idea is supported by several observations. First, DHS2-*Cj* shows high HSS activity and low DHS activity in comparison with all other identified DHS of *Crotalaria* (Table 2). Second, DHS2-*Cj* is not expressed in all plant tissues, but only in the roots, leaves, and shoots (Fig. 3). In contrast, all plant DHS studied so far are expressed ubiquitously in all tested tissues. This suggests that after duplication of the DHS-coding gene early in the evolution of the genus *Crotalaria*, one copy retained its original function (sequences of class 1, such as DHS1-*Cj*), whereas the other copy (class 2, namely DHS2-*Cj*) was recruited for a new function, viz., the biosynthesis of PAs. Since this recent recruitment (indicated by the high degree of identity between the introns of *dhs1-Cj* and *dhs2-Cj*), the coding region has only been slightly modified, whereas mutations within the regulatory elements have resulted in a modified expression pattern. This model is in accordance with results obtained from genome analyses that indicate that differential expression is often the primary event allowing duplicated genes to be maintained in the genome (Hogeweg, 2002; Pickett and Meeks, 1995; Wendel, 2000).
2. The analyzed *C. juncea* is unable to produce PAs. Although *C. juncea* is known to produce PAs (Hartmann and Witte, 1995), the plant investigated in this study was devoid of alkaloids (Table 3). The ability to produce PAs is probably dependent on the developmental stage of the plant as has been described for the biosynthesis of monoterpenoid indole alkaloids (De Luca and St-Pierre, 2000; Meijer et al., 1993). Because no PAs are detected until the flowering stage, one may speculate whether PA biosynthesis is restricted to plants producing seeds. *Crotalaria* seeds are known to contain large amounts of PAs (Chang and Hartmann, 1998; Toppel et al., 1988). However, inappropriate growth conditions or the absence of nitrogen-fixing symbiotic rhizobia may also be the reason for the lack of PAs and, consequently, for an unsuccessful approach to identify the PA-specific HSS. Two groups of rhizobia were isolated from *Crotalaria*, namely *Bradyrhizobium* spp. (Doignon-Bourcier et al., 2000) and *Methylobacterium* (Sy et al., 2001a,b). It is well known that nodulation induces gene expression as shown for specific genes in pea roots (Gloudemans et al., 1989) or for chalcone synthase in *Vicia* (Recourt et al., 1992). DHS2-*Cj* shows only weak expression in the PA-free *C. juncea* plant under the given growth conditions (Fig. 3). Thus, it remains speculative whether DHS2-*Cj* provides *C. juncea* plants with homospermidine and might be upregulated under conditions that allow the plant to produce PAs.

4. Experimental

4.1. Plant material, RNA isolation and cDNA synthesis

Seeds of *C. retusa* L., *C. juncea* L., *C. scassellatii* Chiov., and *C. anagyroides* Kunth. were collected in Brazil, Zimbabwe, Kenya, and Indonesia, respectively. They were germinated on moistened cotton sheets

Table 6
List of primers used for cloning of *Crotalaria*-DHS

1.1.1.1. Oligonucleotide sequence	
P1	5′-GA**R** GA**R** GA**Y** **Y**T**N** AT**H** AA**R** TG**Y** **Y**T-3′
P2	5′-CC**R** TC**R** **W**A**Y** TC**Y** TGN GCN GT**R** TT-3′
P3	5′-AA**R** TT**Y** GA**R** GA**Y** TGG AT**H** AT**H** CC-3′
P4	5′-GC**Y** TC**R** TCN GG**W** CK**N** G**N**R CC-3′
P5	5′-GTC GAC TCG AGA ATT CTT TTT TTT TTT TTT TTT-3′
P6	5′-GAA TGT CTT ATG GAC ACC ATC TAA GTT AA-3′
P7	5′-TGC AAA TGT GAT GCT TT-3′
P8	5′-CAG ACC AGG GTT GCG GAA GGA ATG A-3′
P9	5′-GAA AGT ACA GCA TGT CCC CCA GTG A-3′
P10	5′-ATA TAT CAT ATG AGT GAA GAA GTA AAG GAA GC-3′
P11	5′-AT GGA TCC TAT TGA TGG CAG GTT TTA CTC-3′
P12	5′-TAT CTC GAG ATG AGT AGT AAG GGA AGA AAT GAA-3′
P13	5′-TAA GGA TCC CTA TTG ATG GCT AGG TTT CAC TCT T-3′
P14	5′-GTT CTC ATT CTT ACC TTG TAT GCC CAA TA-3′
P15	5′-CTG GAA TGT TGT CCG TGT ATG CCC AAT A-3′
P16	5′-CTA ATA CGA CTC ACT ATA GGG CGT CGA CTC GAG AAT-3′
P17	5′-CTA ATA CGA CTC ACT ATA GGG C-3′
P18	5′-CTA TAG GGC GTC GAC TCG AGA AT-3′

The underlined bases belong to restriction sites used for cloning. The bold-typed bases are IUB code for mixed bases: B = C,G,T; D = A,G,T; H = A,C,T; K = G,T; M = A,C; N = A,C,G,T; R = A,G; S = G,C; V = G,A,C; W = A,T; Y = C,T.

according to Chang and Hartmann (1998) or germinated under sterile conditions on MS medium (Hartmann et al., 1989). Total RNA was isolated from freshly harvested plant tissues or frozen samples (−80 °C) by using the RNeasy Plant Mini Kit (Qiagen) according to the manufacturer's protocol. An aliquot containing 1 μg total RNA was reverse-transcribed with oligo-dT primer P5 (Table 6) by using Superscript II reverse transcriptase (Invitrogen).

4.2. Degenerate oligonucleotide primed-PCR

Several combinations of degenerate primers (P1–P4, see Table 6, each 1 μM) and an oligo-dT primer (P5, see Table 6, 0.4 μM) were used with 1 μl cDNA for amplification in a reaction volume of 25 μl containing 0.025 U *Taq*-DNA polymerase (Invitrogen). A "touch down" temperature program with decreasing annealing temperatures was used as described previously (Ober and Hartmann, 1999b). After electrophoretic purification, the PCR products were subcloned by means of the pCR2.1 TOPO TA Cloning Kit (Invitrogen) according to the manufacturer's instructions.

4.3. Amplification of 3′-end and 5′-end of cDNA

A gene-specific primer (GSP) was designed according to the fragment i*Cj*S2 identified by DOP-PCR to identify the 3′-end of cDNA (P6, Table 6). It was amplified in a 25-μl reaction mixture containing 1 μl oligo-dT (P5) primed cDNA with P6 and P5 (each 0.4 μM; Table 6). To identify the 5′-end cDNA fragment, GSPs were designed according to the 3′-end of the cDNA (P7–P9, Table 6). After reverse transcription of the cDNA with primer P7, its 5′-end was tailed with dCTP. The dC-tailed cDNA was then amplified by using the rapid amplification of 5′ cDNA ends system (Invitrogen) with the GSPs P8 and P9 (each 0.4 μM). The resulting PCR fragments were electrophoretically purified and subcloned by using the pCR2.1 TOPO-TA Cloning Kit (Invitrogen).

4.4. Amplification, expression, and biochemical identification of recombinant DHS

In order to amplify the full-length ORF of cDNAs homologous to DHS, the same pair of primers was used for *C. juncea* and *C. scassellatii* as for the expression of DHS-*Cr* (P10/P11, Reimann et al., 2004). For the amplification of DHS2-*Cj* a pair of GSPs (P12/P13) was designed containing *Xho*I and *Bam*HI restriction sites, respectively, that allowed cloning into the pET3a-mod expression vector (Reimann et al., 2004). Amplification was performed in 33 cycles with an annealing temperature of 62 °C in a reaction volume of 50 μl containing the primers (4 μM each), *Pf*x-DNA Polymerase (Invitrogen), and 2 μl oligo-dT primed cDNA. After ligation into the linearized expression vectors, transformation and expression of the recombinant protein were achieved as described (Ober and Hartmann, 1999b). The activities of HSS and DHS were measured in desalted protein extracts with eifsv1 as the protein substrate for the DHS assay as described by Ober and Hartmann (1999b).

4.5. Exclusive amplification of cDNA ends RT-PCR

Exclusive amplification of cDNA ends (EXACT) RT-PCR was performed according to a method described by Smith et al. (2001) with slight modifications. cDNA synthesized with primer P4 was used as the template for two rounds of PCR with a combination of a gene-specific primer (P12) and primers P16–P18 as reverse primers. The first PCR was conducted with *Taq*-DNA polymerase (Invitrogen) in a 25-μl reaction volume containing P16 (0.4 μM), P17 (0.02 μM) and P12 (0.04 μM), and the second PCR in 25 μl containing 1 μl diluted PCR product of the first PCR, P12 (0.4 μM), and P18 (0.04 μM).

4.6. Isolation and PCR of genomic DNA

Genomic DNA was isolated by using the DNeasy Plant Mini Kit (Qiagen) according to the manufacturer's instructions. For PCR, 300 ng genomic DNA was used

as the template in a 25-μl reaction mixture with *Pfu*-DNA Polymerase and various combinations of GSPs (P9–P15). Amplificates were subcloned into pCR2.1 TOPO XL vector (Invitrogen).

4.7. Computer-assisted sequence analyses

DNA sequences were analyzed by the Wisconsin Sequence Analysis Package (version 10, Genetics Computer Group, Madison, Wis., USA). Predictions of targeting sequences and of subcellular localization were performed by using TargetP (v1.01; http://www.cbs.dtu.dk/services/TargetP/) and PSORT (http://www.psort.org/). The maximum likelihood tree and the bootstrap proportions were calculated as described previously (Reimann et al., 2004).

4.8. Extraction and quantification of PAs

Extraction and analysis of PAs from various *Crotalaria* tissues was performed according to Witte et al. (1993).

Acknowledgments

The authors thank T. Hartmann for helpful discussions, R. Hänsch for his advice regarding the in vivo analysis of the putative signal peptide, and T. Beuerle and S. Burkhard for their assistance in PA analyses. This work was supported by the Deutsche Forschungsgemeinschaft and by the Deutscher Akademischer Austauschdienst (grant to N.N.).

References

Anke, S., Niemüller, D., Moll, S., Hänsch, R., Ober, D., 2004. Polyphyletic origin of pyrrolizidine alkaloids within the Asteraceae. Evidence from differential tissue expression of homospermidine synthase. Plant Physiol. 136, 4037–4047.

Bartig, D., Lemkemeier, K., Frank, J., Lottspeich, F., Klink, F., 1992. The archaebacterial hypusine-containing protein. Structural features suggest common ancestry with eukaryotic translation initiation factor 5A. Eur. J. Biochem. 204, 751–758.

Bauer, M., Dietrich, C., Nowak, K., Sierralta, W.D., Papenbrock, J., 2004. Intracellular localization of *Arabidopsis* sulfurtransferases. Plant Physiol. 135, 916–926.

Chang, A., Hartmann, T., 1998. Solubilization and characterization of a senecionine *N*-oxygenase from *Crotalaria scassellatii* seedlings. Phytochemistry 49, 1859–1866.

De Luca, V., St-Pierre, B., 2000. The cell and developmental biology of alkaloid biosynthesis. Trends Plant Sci. 5, 168–173.

Doignon-Bourcier, F., Willems, A., Coopman, R., Laguerre, G., Gillis, M., de Lajudie, P., 2000. Genotypic characterization of *Bradyrhizobium* strains nodulating small Senegalese legumes by 16S–23S rRNA intergenic gene spacers and amplified fragment length polymorphism fingerprint analyses. Appl. Environ. Microbiol. 66, 3987–3997.

Gloudemans, T., Bhuvaneswari, T.V., Moerman, M., Van Brussel, T., Van Kammen, A., Bisseling, T., 1989. Involvement of *Rhizobium leguminosarum* nodulation genes in gene expression in pea root hairs. Plant Mol. Biol. 12, 157–168.

Goodall, G.J., Filipowicz, W., 1990. The minimum functional length of pre-mRNA introns in monocots and dicots. Plant Mol. Biol. 14, 727–733.

Gordon, E.D., Mora, R., Meredith, S.C., Lee, C., Lindquist, S.L., 1987. Eukaryotic initiation factor 4D, the hypusine-containing protein, is conserved among eukaryotes. J. Biol. Chem. 262, 16585–16589.

Hartmann, T., Ehmke, A., Eilert, U., von Borstel, K., Theuring, C., 1989. Sites of synthesis, translocation and accumulation of pyrrolizidine alkaloid *N*-oxides in *Senecio vulgaris* L. Planta 177, 98–107.

Hartmann, T., 1994. *Senecio* spp.: biochemistry of the formation of pyrrolizidine alkaloids in root cultures. In: Bajaj, Y.P.S. (Ed.), Biotechnology in Agriculture and Forestry, Vol. 26. Springer, Berlin, Heidelberg, New York, pp. 339–355.

Hartmann, T., Witte, L., 1995. Pyrrolizidine alkaloids: chemical, biological and chemoecological aspects. In: Pelletier, S.W. (Ed.), Alkaloids: Chemical and Biological Perspectives, Vol. 9. Pergamon Press, Oxford, pp. 155–233.

Hartmann, T., 1999. Chemical ecology of pyrrolizidine alkaloids. Planta 207, 483–495.

Hartmann, T., Ober, D., 2000. Biosynthesis and metabolism of pyrrolizidine alkaloids in plants and specialized insect herbivores. In: Leeper, F.J., Vederas, J.C. (Eds.), Topics in Current Chemistry, Vol. 209. Springer, Berlin, Heidelberg, New York, pp. 207–244.

Hirotsune, S., Yoshida, N., Chen, A., Garrett, L., Sugiyama, F., Takahashi, S., Yagami, K.-I., Wynshaw-Boris, A., Yoshiki, A., 2003. An expressed pseudogene regulates the messenger-RNA stability of its homologous coding gene. Nature 423, 91–96.

Hogeweg, P., 2002. Multilevel evolution: the fate of duplicated genes. Z. Phys. Chem. 216, 77–90.

Korning, P.G., Hebsgaard, S.M., Rouze, P., Brunak, S., 1996. Cleaning the GenBank *Arabidopsis thaliana* data set. Nucleic Acids Res. 24, 316–320.

Krishna, R.G., Wold, F., 1993. Post-translational modification of proteins. Adv. Enzymol. Relat. Areas Mol. Biol. 67, 265–298.

Lal, S., Choi, J.H., Hannah, L.C., 1999a. The AG dinucleotide terminating introns is important but not always required for pre-mRNA splicing in the maize endosperm. Plant Physiol. 120, 65–72.

Lal, S., Choi, J.H., Shaw, J.R., Hannah, L.C., 1999b. A splice site mutant of maize activates cryptic splice sites, elicits intron inclusion and exon exclusion, and permits branch point elucidation. Plant Physiol. 121, 411–418.

Mantuano, E., Trettel, F., Olsen, A.S., Lennon, G., Frontali, M., Jodice, C., 1998. Localization and genomic structure of human deoxyhypusine synthase gene on chromosome 19p13.2–distal 19p13.1. Gene 215, 153–157.

Meijer, A.H., Verpoorte, R., Hoge, J.H.C., 1993. Regulation of enzymes and genes involved in terpenoid indole alkaloid biosynthesis in *Catharanthus roseus*. J. Plant Res. 3, 145–164.

Moll, S., Anke, S., Kahmann, U., Hänsch, R., Hartmann, T., Ober, D., 2002. Cell-specific expression of homospermidine synthase, the entry enzyme of the pyrrolizidine alkaloid pathway in *Senecio vernalis*, in comparison with its ancestor, deoxyhypusine synthase. Plant Physiol. 130, 47–57.

Ober, D., Hartmann, T., 1999a. Homospermidine synthase, the first pathway-specific enzyme of pyrrolizidine alkaloid biosynthesis, evolved from deoxyhypusine synthase. Proc. Natl. Acad. Sci. USA 96, 14777–14782.

Ober, D., Hartmann, T., 1999b. Deoxyhypusine synthase from tobacco: cDNA isolation, characterization, and bacterial expression of an enzyme with extended substrate specificity. J. Biol. Chem. 274, 32040–32047.

Ober, D., Harms, R., Hartmann, T., 2000. Cloning and expression of homospermidine synthase from *Senecio vulgaris*: a revision. Phytochemistry 55, 305–309.

Park, M.H., Lee, Y.B., Joe, Y.A., 1997. Hypusine is essential for eukaryotic cell proliferation. Biol. Signals 6, 115–123.

Pickett, F.B., Meeks, W.D.R., 1995. Seeing double: appreciating genetic redundancy. Plant Cell 7, 1347–1356.

Recourt, K., van Tunen, A.J., Mur, L.A., van Brussel, A.A., Lugtenberg, B.J., Kijne, J.W., 1992. Activation of flavonoid biosynthesis in roots of *Vicia sativa* subsp. *nigra* plants by inoculation with *Rhizobium leguminosarum* biovar *viciae*. Plant Mol. Biol. 19, 411–420.

Reimann, A., Nurhayati, N., Backenköhler, A., Ober, D., 2004. Repeated evolution of the pyrrolizidine alkaloid-mediated defense system in separate angiosperm lineages. Plant Cell 16, 2772–2784.

Sieburth, L.E., Running, M.P., Meyerowitz, E.M., 1995. Genetic separation of third and fourth whorl functions of *AGAMOUS*. Plant Cell 7, 1249–1258.

Smith, R.D., Ogden, C.W., Penny, M.A., 2001. Exclusive amplification of cDNA template (EXACT) RT-PCR to avoid amplifying contaminating genomic pseudogenes. BioTechniques 31, 776–782.

Sy, A., Giraud, E., Jourand, P., Garcia, N., Willems, A., de Lajudie, P., Prin, Y., Neyra, M., Gillis, M., Boivin-Masson, C., Dreyfus, B., 2001a. Methylotrophic *Methylobacterium* bacteria nodulate and fix nitrogen in symbiosis with legumes. J. Bacteriol. 183, 214–220.

Sy, A., Giraud, E., Samba, R., de Lajudie, P., Gillis, M., Dreyfus, B., 2001b. Nodulation of certain legumes of the genus *Crotalaria* by the new species *Methylobacterium*. Can. J. Microbiol. 47, 503–508.

Toppel, G., Witte, L., Hartmann, T., 1988. *N*-oxidation and degradation of pyrrolizidine alkaloids during germination of *Crotalaria scassellatii*. Phytochemistry 27, 3757–3760.

Wendel, J.F., 2000. Genome evolution in polyploids. Plant Mol. Biol. 42, 225–249.

Wink, M., Mohamed, G.I.A., 2003. Evolution of chemical defense traits in the Leguminosae: mapping of distribution patterns of secondary metabolites on a molecular phylogeny inferred from nucleotide sequences of the *rbc*L gene. Biochem. Syst. Ecol. 31, 897–917.

Witte, L., Rubiolo, P., Bicchi, C., Hartmann, T., 1993. Comparative analysis of pyrrolizidine alkaloids from natural sources by gas chromatography–mass spectrometry. Phytochemistry 32, 187–196.

Xue, J., Rask, L., 1995. The unusual 5′ splicing border GC is used in myrosinase genes of the Brassicaceae. Plant Mol. Biol. 29, 167–171.

Yan, S.C., Grinnell, B.W., Wold, F., 1989. Post-translational modifications of proteins: some problems left to solve. Trends Biochem. Sci. 14, 264–268.