Phytochemistry 70 (2009) 508–516

Contents lists available at ScienceDirect

Phytochemistry

journal homepage: www.elsevier.com/locate/phytochem

Evolution of pyrrolizidine alkaloids in *Phalaenopsis* orchids and other monocotyledons: Identification of deoxyhypusine synthase, homospermidine synthase and related pseudogenes

Niknik Nurhayati [a,1], Daniela Gondé [b,2], Dietrich Ober [b,*]

[a] Institut für Pharmazeutische Biologie, TU Braunschweig, Mendelssohnstr. 1, D-38106 Braunschweig, Germany
[b] Botanisches Institut und Botanischer Garten, Universität Kiel, Olshausenstrasse 40, D-24098 Kiel, Germany

ARTICLE INFO

Article history:
Received 27 November 2008
Received in revised form 29 January 2009
Available online 28 February 2009

Keywords:
Plant secondary metabolism
Pathway evolution
Pyrrolizidine alkaloid
Phalaenopsine
Maize
Zea mays
Homospermidine synthase
Deoxyhypusine synthase
Transcriptionally active pseudogene
Processed pseudogene

ABSTRACT

In order to study the evolution of pathways of plant secondary metabolism, we use the biosynthesis of pyrrolizidine alkaloids (PAs) as a model system. PAs are regarded as part of the plant's constitutive defense against herbivores. Homospermidine synthase (HSS) is the first specific enzyme of PA biosynthesis. The gene encoding HSS has been recruited from the gene encoding deoxyhypusine synthase (DHS) from primary metabolism at least four times independently during angiosperm evolution. One of these recruitments occurred within the monocot lineage. We have used the PA-producing orchid *Phalaenopsis* to identify the cDNAs encoding HSS, DHS and the substrate protein for DHS, i.e., the precursor of the eukaryotic initiation factor 5A. A cDNA identified from maize was unequivocally characterized as DHS. From our study of *Phalaenopsis*, several pseudogenes emerged, of which one was shown to be a "processed pseudogene", and others to be transcribed. Sequence comparison of the HSS- and DHS-encoding sequences from this investigation with those of monocot species taken from the databases suggest that HSS and probably the ability to produce PAs is an old feature within the monocot lineage. This result is discussed with respect to the recent discovery of structural related PAs within grasses.

1. Introduction

Pyrrolizidine alkaloids (PAs) are typical compounds of plant secondary metabolism and are constitutively produced by the plant as a defense against herbivores (Hartmann and Ober, 2000). We use this class of alkaloids as a system to study the way that plants have recruited complex pathways, i.e., the enzymes catalyzing specific reactions and the regulatory factors that ensure the proper integration of the pathway in plant metabolism during evolution. The occurrence of PAs is restricted to the angiosperms, suggesting that, in evolutionary terms, this pathway is relatively "young". PAs are found in many species belonging to the families of the Asteraceae (the tribes Senecioneae and Eupatorieae), Boraginaceae, Heliotropiaceae, and Apocynaceae, in some species of the Orchidaceae, and within the Fabaceae mainly in the genus *Crotalaria* (Hartmann and Witte, 1995). In some further families (Ranunculaceae, Convolvulaceae, Celastraceae), isolated occurrences have been described in only a few species (Hartmann and Witte, 1995; Jenett-Siems et al., 1993).

We have previously presented evidence that the gene coding for the first specific enzyme in PA biosynthesis, viz., homospermidine synthase (HSS, EC 2.5.1.45), was recruited by duplication of the gene coding for deoxyhypusine synthase (DHS, EC 2.5.1.46) (Ober and Hartmann, 1999b). DHS catalyzes the first of two steps activating the eukaryotic initiation factor 5A (eIF5A), which has been shown to be essential for cell proliferation (Chattopadhyay et al., 2008; Nishimura et al., 2005; Park et al., 1993, 1997). The DHS system is characterized by high sequence conservation related to strongly preserved biochemical functions (Gordon et al., 1987; Kyrpides and Woese, 1998; Magdolen et al., 1994). In plants, activated eIF5A might be involved in processes such as seed germination (Moll et al., 2002), senescence, and apoptosis (Wang et al., 2001, 2003, 2005). In addition to a high degree of sequence identity, HSS and DHS share many biochemical properties (Ober and Hartmann, 1999b; Reimann et al., 2004). HSS catalyzes the formation of homospermidine by the transfer of an aminobutyl moiety of spermidine to putrescine. DHS, which also transfers the aminobutyl moiety of spermidine, modifies a specific lysine residue within the eIF5A precursor protein but is also able to catalyze the formation of homospermidine. This "side-activity" of DHS has been

* Corresponding author. Tel.: +49 431 880 4299; fax: +49 431 880 4500.
E-mail addresses: dober@bot.uni-kiel.de (D. Ober).
[1] Present address: Center for Bioindustrial Technology, Agency for Assessment and Application of Technology (BPPT), J1 MH Thamrin No. 8, Jakarta 10340, Indonesia.
[2] Present address: Centre de Recherche en Cancérologie de l'Université LAVAL, Pavillon Hôtel-Dieu de Québec; 9, rue Mc Mahon, Québec (QC), Canada G1R 2J6.

0031-9422/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.phytochem.2009.01.019

proposed as being responsible for small amounts of homospermidine detected in various angiosperm species unable to produce PAs (Ober et al., 2003a, b). Thus, despite their completely different reaction products, HSS and DHS are almost identical with respect to their reaction mechanism with one exception: HSS is unable to bind and modify the eIF5A precursor protein (Ober et al., 2003b). This inability is the distinctive feature between paralogous HSS and DHS (Fig. 1).

A comparison of 23 cDNA sequences coding for HSS and DHS sequences of various angiosperm species has provided evidence that the recruitment of the HSS gene by duplication of the DHS-coding gene occurred at least four times independently during angiosperm evolution (Reimann et al., 2004). Further such recruitments are likely, because HSS sequences have been studied so far only in the Boraginales (one independent origin), Asteraceae (two independent origins within the tribes Eupatorieae and Senecioneae), and the monocots (one independent origin).

Orchidaceae are one of the largest (ca. 30,000 species) and most complex families of flowering plants (Cameron, 2004). According to molecular phylogenies based on the plastid *rbcL* (Cameron et al., 1999; Chase et al., 1994) and *psaB* sequences (Cameron, 2004), members of this family are divided into five subfamilies (Apostasioideae, Cypripedioideae, Epidendroideae, Orchidoideae, Vanilloideae). Within monocots, PAs have only been found within orchids. They have been described in eight genera, viz., *Liparis*, *Malaxis*, *Cysis*, *Phalaenopsis*, *Vanda*, *Vandopsis* (Hartmann and Witte, 1995), *Pleurothallis* (Borba et al., 2001), and *Cremastra* (Ikeda et al., 2005) belonging to five tribes within the subfamily Epidendroideae (Cameron, 2004).

Most of the PAs found in plants are characterized by a 1,2-double bond that, amongst others, is responsible for the liver toxicity of PAs (Fu et al., 2004; Stewart and Steenkamp, 2001). They are usually synthesized, translocated, and stored as *N*-oxides (Hartmann and Witte, 1995). In contrast, the PAs found in the Orchidaceae belong to the phalaenopsine type of PAs as classified by Hartmann and Witte (1995). They are characterized by a 1,2-saturated necine base that is esterified with an aryl, aralkyl, or alkyl necic acid. Furthermore, only 40–60% of the PAs of *Phalaenopsis* occur in their *N*-oxide form (Frölich et al., 2006). This putative structural simplicity, the lack of liver toxicity, and the assumed restricted occurrence of these phalaenopsine-type PAs in only certain genera of the orchids might be interpreted as being indications for a more basal defense system. This interpretation and the observation that HSS, as the first enzyme of PA biosynthesis, has an independent origin within monocots, suggest that the ability of monocots to produce PAs is a young feature, probably restricted to only some orchid genera.

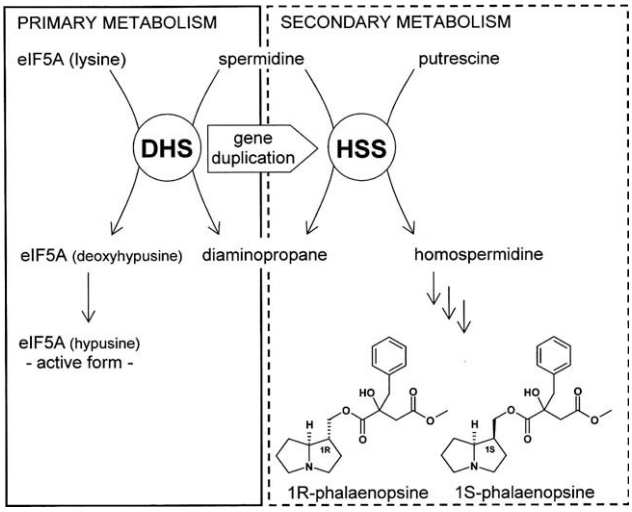

Fig. 1. The role of the paralogous enzymes DHS and HSS in the activation of eIF5A and the biosynthesis of PAs, respectively. Both enzymes catalyze the transfer of an aminobutyl moiety of spermidine, releasing diaminopropane, to modify a specific lysine residue within the eIF5A prescursor protein to deoxyhypusine (in the case of DHS) and to transform putrescine to homospermidine (in the case of HSS). Homospermidine is the first specific intermediate in the biosynthesis of PAs. The two isomers 1R-phalaenopsine and 1S-phalaenopsine are the only PAs found in *Phalaenopsis*.

In the present paper, we describe the identification of the cDNAs encoding DHS in *Phalaenopsis* and maize and of the eIF5A substrate protein of *Phalaenopsis*. The encoded proteins have been heterologously expressed in *Escherichia coli* and used for comparative biochemical characterization with HSS and several HSS-related pseudogenes of *Phalaenopsis*. A phylogenetic analysis of these sequences in addition to sequence data available from the databases suggest that the gene duplication creating the HSS was not a recent event within monocots.

2. Results

2.1. Cloning, expression, and identification of recombinant *Phalaenopsis* HSS

The identification of the open-reading frame (ORF) of the cDNA coding for PA-specific HSS from the total RNA of aerial root tips of an interspecific hybrid of *Phalaenopsis* has been briefly described (Reimann et al., 2004). The ORF codes for 371 amnio acids and is flanked by a 5'-untranslated region (UTR) of 50 bp and a 3'-UTR of 206 bp. To simplify the purification of the recombinant enzyme (HSS-*Ph*), the cDNA in this study was inserted into the pET23a expression vector, which added a hexahistidine tag at the C-terminal end of the protein during expression. The recombinant protein was characterized after metal chelate affinity chromatography. The protein showed only HSS activity (5.04 nkat/mg, Table 1) but no DHS activity, having the characteristics of most identified HSS of plant origin (Ober et al., 2003b).

To rule out the possibility that this sequence codes for a DHS that is just unable to accept the eIF5A precursor protein of *Senecio vernalis* used in the assay as a substrate, we identified, cloned, and expressed an eIF5A protein of *Phalaenopsis* in *E. coli*. This protein, termed eifph1, was accepted as a substrate by the DHS of *S. vernalis* (Ober and Hartmann, 1999b) to almost the same degree as its native substrate, the eIF5A of *S. vernalis* (data not shown). After purification by metal chelate affinity chromatography, eifph1 was used as a substrate for all further DHS assays. The enzyme protein HSS-*Ph* encoded by the identified sequence of *Phalaenopsis* again showed only HSS activity but no DHS activity, supporting the classification of the protein as HSS (Table 1).

2.2. Identification of pseudogenes with homology to the hss gene

The same set of degenerate primers used for the identification of the HSS-coding cDNA of *Phalaenopsis* was used in various combinations to identify the cDNA coding for DHS. For this

Table 1
Specific enzyme activities of affinity-purified recombinant HSS and DHS from *Phalaenopsis* and of DHS of maize. Enzyme assays for DHS were performed with 20 μM eIF5A substrate protein of *Senecio vernalis* (eifsv1), of *Phalaenopsis* (eifph1), and of *Zea mays* (eifzm1), respectively. nd, not detectable, –, not analyzed.

	HSS *P. amabilis*	DHS *P. amabilis*	DHS of *Z. mays*
HSS assay	5.04 nkat/mg	34.6 pkat/mg	198.5 pkat/mg
DHS assay			
With eifsv1	nd	8.9 pkat/mg	–
With eifph1	nd	36.4 pkat/mg	–
With eifzm1	–	–	202.2

purpose, total RNA of various tissues of an interspecific hybrid of *Phalaenopsis* (roots, leaves, stalks, flowers, flower buds) was used for reverse transcription. All fragments amplified showed a high degree of identity to the cDNA coding for the HSS of *Phalaenopsis* (more than 97% at the nucleic acid level). Cloning and sequencing of these homologous sequences resulted in the identification of two pseudogenes (Ψ*hss2-Ph*, Ψ*hss3-Ph*) unequivocally identified by the presence of a premature stop codon corresponding to position 161 of the derived amino acid sequence (see Fig. 2). Another pseudogene with a stop codon at an identical position was amplified from the *Phalaenopsis* hybrid by using the Exact RT-PCR (reverse transcription with polymerase chain reaction) approach (Smith et al., 2001) to avoid co-amplification of contaminating genomic sequences (Ψ*hss4-Ph*), suggesting that this pseudogene sequence is transcribed. Further pseudogenes were identified from *Phalaenopsis amabilis* with the same primer pair (P15/P16) as that for the amplification of the HSS-encoding ORF (Ψ*hss5-Ph*) by a RT-PCR approach and from genomic DNA (Ψ*hss1-Ph*). It is noteworthy that the pseudogene Ψ*hss1-Ph* is devoid of any introns, suggesting that it is a processed pseudogene. The identification of these various pseudogenes with high sequence identity and our inability to identify the cDNA coding for DHS resulted in two hypotheses. (1) Based on the phylogeny of HSS- and DHS-coding cDNAs of various angiosperm species suggesting an independent origin of PA biosynthesis within monocots (Reimann et al., 2004), we might speculate that the ability of *Phalaenopsis* to synthesize PAs is a feature invented "recently" during the evolution of the orchid family. This hypothesis is further supported by the structural simplicity of the phalaenopsine-type PAs found in *Phalaenopsis*. In this case, the two copies encoding HSS and DHS resulting from a recent duplication event of the DHS-coding gene would be expected to share a high degree of sequence identity, suggesting that the identified pseudogenes might also be derived from the DHS-encoding gene. (2) Alternatively, we might speculate that the duplication event of the *dhs* gene is much older than expected. In this case, the two copies would have diverged for a longer period of time, resulting in two sequences with a significant lower degree of identity, suggesting that the identified pseudogenes arose from HSS-coding gene copies. To test the first hypothesis, we performed site-specific mutagenesis of the pseudogenes Ψ*hss2-Ph*, Ψ*hss3-Ph*, and Ψ*hss4-Ph* by replacing the premature stop codon (TGA) with the triplet encoding tryptophan (TGG), which is found at this position within the HSS-encoding cDNA sequence (Fig. 2). Heterologous expression of these proteins in *E. coli* BL21(DE3) resulted in the expression of a soluble protein of approximately 45 kDa in size that showed neither HSS nor DHS activity.

2.3. Cloning, expression, and identification of recombinant *Phalaenopsis* DHS

Based on the second hypothesis (see above), we designed a set of degenerate primers according to an alignment of 23 cDNA sequences coding for HSS and DHS of various angiosperm species (Reimann et al., 2004) allowing us to identify the DHS-encoding cDNA of *Phalaenopsis*. Using the cDNA as a template that was reverse-transcribed from the total RNA of pedicels and protocorms (young seedlings) of *Phalaenopsis*, a full-length sequence has been identified coding for a protein (DHS-*Ph*) of 378 amino acids with UTRs of 41 bp and 211 bp at the 5'- and the 3'-ends, respectively. The recombinant protein showed significant DHS and HSS activity (Table 1).

2.4. Sequence comparison and phylogenetic analysis of monocot HSS and DHS

The cDNA sequences coding for HSS and DHS of *Phalaenopsis* shared identities of 73% and 72% at the nucleotide and the amino acid levels, respectively. This was the lowest degree of identity between these two paralogous sequences that we had ever observed within a single angiosperm species.

We previously presented evidence for at least four independent recruitments of the gene coding for HSS from the DHS-coding gene

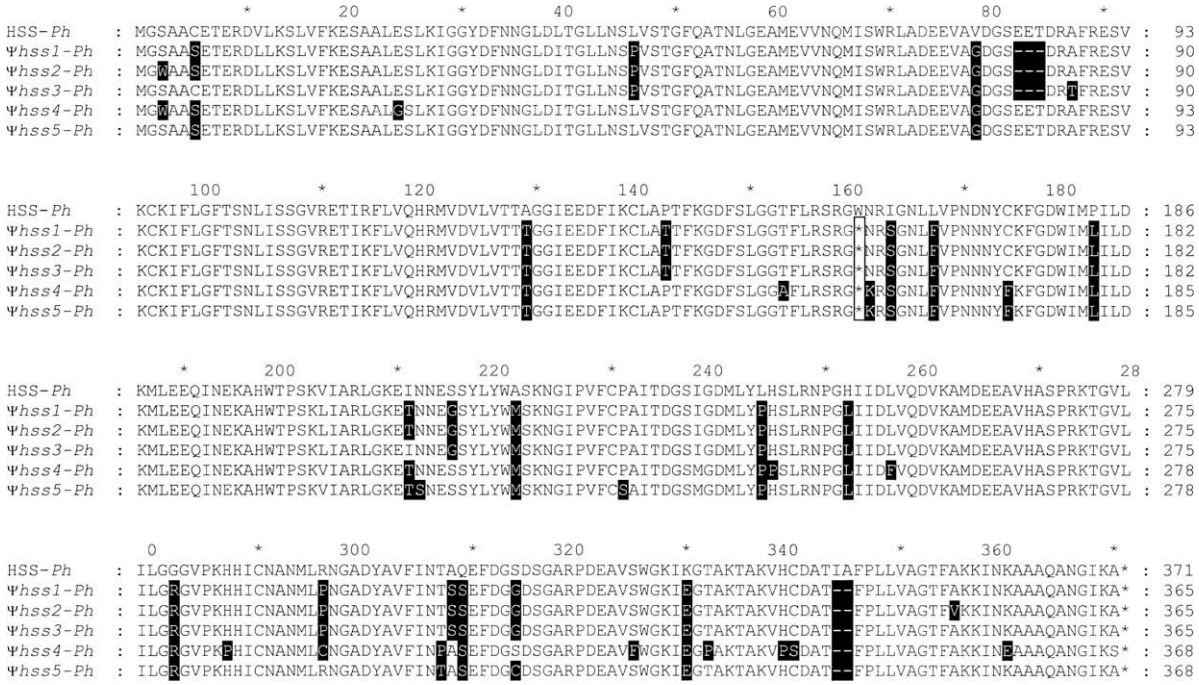

Fig. 2. Amino acid sequences derived from the cDNA sequence encoding functional HSS of *Phalaenopsis* (HSS-*Ph*) and from several related pseudogene sequences (Ψ*hss1-Ph*, Ψ*hss2-Ph*, Ψ*hss3-Ph*, Ψ*hss4-Ph*, Ψ*hss5-Ph*). Sequence Ψ*hss1-Ph* was amplified from genomic DNA, sequences Ψ*hss2-Ph*, Ψ*hss3-Ph*, and Ψ*hss4-Ph* from cDNA of an interspecific hybrid of *Phalaenopsis*, and sequence Ψ*hss5-Ph* from a cDNA preparation of *P. amabilis*. Hyphens indicate sequence gaps in the alignment, asterisks the position of the stop codons. Black boxes highlight those residues within the pseudogene sequences that are different to the sequence of the functional HSS.

by gene duplication during angiosperm evolution, one of these occurring within the monocots (Reimann et al., 2004). At that time, only two monocot sequences coding for HSS or DHS were available, viz., the DHS of banana (*Musa acuminata*) and the HSS of *Phalaenopsis*. The cDNA coding for the putative DHS of *Oryza sativa* has subsequently become available in the database. After the identification of the DHS of *Phalaenopsis*, we were interested in whether the recruitment of HSS was a recent event within the Orchidaceae family of the monocots, or whether the presence of a gene specialized for the biosynthesis of homospermidine as the first step in PA biosynthesis was an old feature among monocots. For these analyses, the limited amount of available cDNA sequences coding for HSS and DHS from monocots was an obstacle. We therefore used the following strategies to increase and improve the available data. First, we decided to identify, express, and characterize one of the putative DHS sequences of the Poaceae, i.e., the DHS of maize, to establish its identity as a DHS. Second, further expressed sequence tag (EST) sequences coding for putative DHS sequences were analyzed and assembled to generate the full ORF. The resulting sequences were combined in a nucleic acid alignment for our phylogenetic analyses.

The cDNA encoding DHS of maize contains an ORF coding for 370 amino acids with 5'-UTRs and 3'-UTRs of 31 bp and 238 bp, respectively. After heterologous expression, the affinity-purified maize DHS showed HSS and DHS activity of 198.5 pkat/mg and 202.2 pkat/mg, respectively, thereby establishing its identity as a DHS.

Analyzing the available EST data, we were able to assemble the whole ORF of *Triticum aesticum* and *Festuca arundinacea*, whereas for *Saccharum officinarum*, *Hordeum vulgare*, and *Sorghum bicolor*, gaps of unidentified sequences had to be filled with 39, 2, and 44

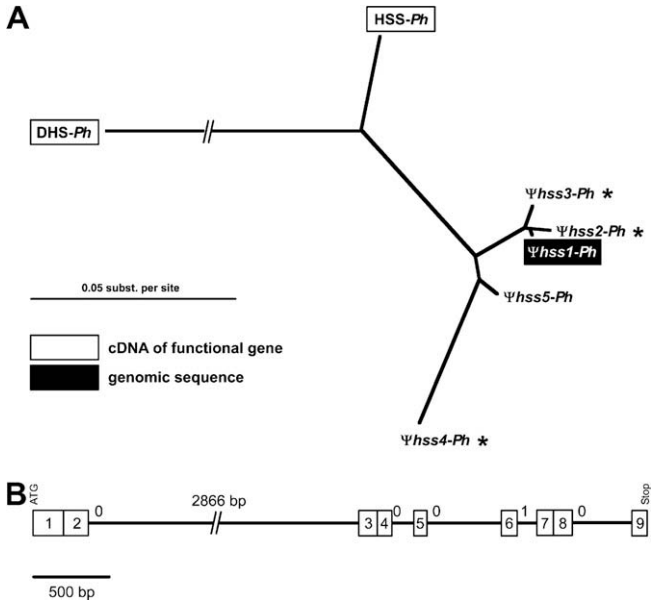

Fig. 3. A: Similarity tree of nucleotide sequences related to HSS encoding cDNA of *Phalaenopsis*. Pseudogene names are labeled by the prefix Ψ. cDNA sequences encoding HSS (HSS-*Ph*) and DHS (DHS-*Ph*) are presented in white boxes, the intronless pseudogene sequence identified from genomic DNA (Ψ*hss1-Ph*) is presented in a black box, and pseudogene sequences labeled with an asterisk were modified by site-specific mutagenesis and expressed in *E. coli* BL21(DE3). The branch length of DHS-encoding cDNA (DHS-*Ph*) was reduced to 10% of its original length. Sequences were aligned using ClustalX (Thompson et al., 1997) and compared with the software DNAML of the PHYLIP program package (Felsenstein, 2001). B: Genomic structure of the *hss* gene of *Phalaenopsis amabilis* containing five introns. Numbered boxes represent the coding stretches of the gene sequence that was separated into nine exons according to (Reimann et al., 2004). Introns are shown by lines with a number indicating the intron phase.

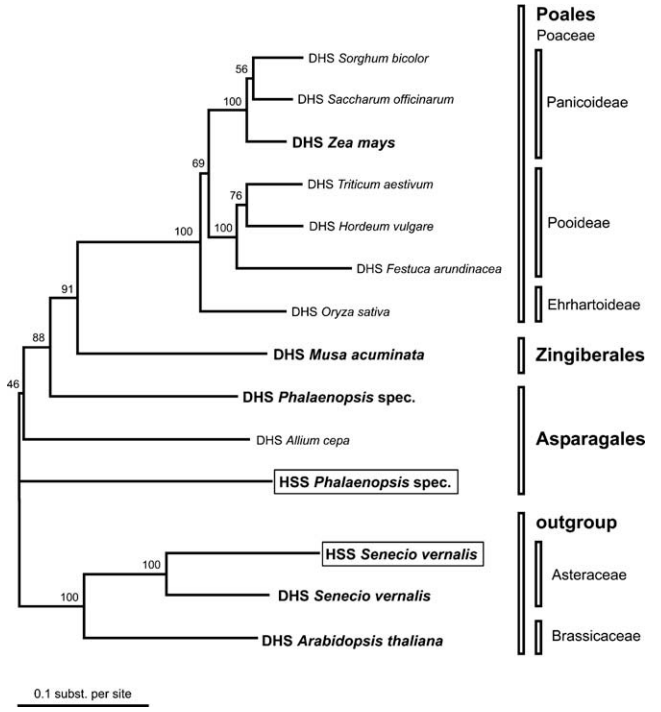

Fig. 4. Rooted neighbor-joining tree of HSS-coding (framed) and DHS-coding cDNA sequences of various monocot species with sequences from *S. vernalis* and *A. thaliana* as the outgroup. The identity of sequences in bold was experimentally established, whereas the others should be regarded as putative DHS-coding cDNA. Bootstrap proportions resulted from 1000 replicates.

replacement characters, respectively. Of monocot species outside the Poaceae, only partial sequences of *Allium cepa*, *Acorus americanus*, and *Asparagus officinalis* were found. As the latter two turned out to be identical, we discarded these sequences and included only that of *A. cepa* into our alignment to compensate for the excess of Poaceae sequences. To avoid misleading results attributable to the missing 3' end of the *Allium* sequence, we used only the central part of the alignment (782 bp in length), to test whether a phylogenetic analysis of these sequences might give a hint about the branching pattern of HSS- and DHS-encoding sequences within the monocots (the alignment is available as Supplementary Fig. S1). A neighbor joining tree of all available monocot sequences with HSS and DHS cDNAs of *S. vernalis* and the cDNA for DHS of *Arabidopsis thaliana* as outgroup is shown in Fig. 4. Despite the limitations concerning the available sequence data, the branching pattern is the same as that from an independent maximum likelihood analysis (data not shown). The splitting of the tree into three subclades within the Poaceae sequences has good bootstrap support and corresponds to three subfamilies, viz., Pooideae, Panicoideae, and Ehrhartoideae, as confirmed by phylogenetic data of the Poaceae family (Duvall and Morton, 1996; Salamin et al., 2002). Moreover, the DHS-coding cDNA sequences of *Phalaenopsis* and of *Allium*, both species belonging to the order Asparagales, cluster together. However, the HSS-coding cDNA of *Phalaenopsis* as one of the most basal branches of all analyzed monocot sequences is well separated from the HSS of *S. vernalis* as one of the outgroup sequences, supporting the independent origin of HSS within the monocots.

3. Discussion

The cDNA sequences encoding paralogous HSS and DHS of *Phalaenopsis*, which are involved in PA biosynthesis and in the post-translational activation of eIF5A, respectively, and the DHS of

maize have been identified in addition to several pseudogenes. Coding cDNAs have been characterized biochemically after the expression of the recombinant proteins in *E. coli* and purification by metal chelate affinity chromatography. Despite the assumed simplicity of PA structures found in orchids, the relative low sequence identity between the cDNAs encoding HSS and DHS of *Phalaenopsis* in comparison to other angiosperm species suggest that the gene coding for HSS was recruited early during monocot evolution and thus, PA biosynthesis is likely to be an old feature within this lineage.

3.1. Strategies for identification of cDNAs encoding HSS and DHS within Phalaenopsis

Using cDNA preparations of aerial root tips of *Phalaenopsis*, we have succeeded in the identification of a sequence with homology to the DHS of other angiosperm species. Heterologous expression of this sequence and biochemical characterization of the resulting protein has confirmed its identity as HSS, because the protein is able to catalyze the formation of homospermidine but is unable to catalyze the modification of the eIF5A substrate protein of *S. vernalis* and of *Phalaenopsis*. A phylogenetic analysis based on 23 cDNA sequences coding for HSS and DHS of various angiosperm species, including that coding for HSS of *Phalaenopsis* and a second monocot sequence coding for DHS of banana, has suggested an independent origin of the PA-specific HSS within the monocot lineage (Reimann et al., 2004). For the identification of the cDNA encoding DHS, we have used young seedlings known as protocorms in orchids, because we have observed an elevated level of DHS expression in young seedlings as shown by Moll et al. (2002) for tobacco. In this tissue, the mRNA concentration of DHS is presumably sufficiently high to allow the successfully identification of the respective cDNA. The heterologously expressed DHS of *Phalaenopsis* shows all the characteristics of a plant DHS as described by Reimann et al. (2004), i.e., the ability to catalyze the formation of homospermidine and to modify the eIF5A precursor protein of *Phalaenopsis* and of *S. vernalis*.

Nevertheless, our efforts to identify the DHS-coding cDNA have been hampered by the co-amplification of processed pseudogenes with a high degree of identity to the identified HSS. Pseudogenes are defined as sequences that are related to certain functional sequences but that have lost their functionality because of defects within the sequences, such as deletions, insertions, or mutations (Vanin, 1985). Vanin (1985) distinguishes two kind of pseudogenes, i.e., "duplicated pseudogenes" that are characterized by the presence of introns, in contrast to "processed pseudogenes" that are devoid of these non-coding stretches of DNA. Although duplicated pseudogenes are assumed to originate by gene duplication with succeeding pseudogenization, processed pseudogenes originate by retroposition, a process that involves reverse transcription of a mRNA sequence and the integration of the resulting cDNA into the genome. As these pseudogenes originate from mRNA, they are characterized by the lack of introns, by the lack of a 5'-promotor, and by the presence of a polyA-stretch. Furthermore, they are flanked by target site duplications (TSD) of 7 to 17 bp in length that are introduced by retroposition (Balakirev and Ayala, 2003; Maestre et al., 1995; Vanin, 1985; Zhang et al., 2005). Of these characteristics, "intronless" is the most persistent trait of pseudogenes within the genome, as poly-(A) tails and TSDs tend to decrease in length with time (Gilbert et al., 2002; Symer et al., 2002; Zhang et al., 2005).

For the functional gene encoding HSS of *Phalaenopsis*, we have identified five introns (Fig. 3B). The finding that the pseudogene amplified from genomic DNA is devoid of any introns (Ψ*hss1-Ph*) confirms its identity as a "processed pseudogene". Expression of processed pseudogenes should only occur if integration into the genome results in the control of the pseudogene by a promoter of an intact unrelated gene. In our study, four of the identified pseudogenes were amplified from cDNA preparations. Sequences Ψ*hss2-Ph* and Ψ*hss3-Ph* were amplified from preparations that might have been contaminated by genomic DNA. These sequences show a high degree of identity to the genomic pseudogene sequence Ψ*hss1-Ph* (Figs. 2 and 3A). For the identification of the sequence Ψ*hss4-Ph*, the EXACT-PCR approach (Smith et al., 2001) was used to exclude co-amplification of contaminating genomic DNA, suggesting that this pseudogene is transcribed. Comparing the derived amino acid sequence with that of the other pseudogenes, this sequence shows significant differences (e.g., the lack of the deletion of three amino acids at position 79–81). One might speculate whether the Ψ*hss4-Ph* gene is not a "processed pseudogene" but rather a "duplicated pseudogene". However so far, we have been unable to identify, within the genome, a pseudogene sequence containing introns. All identified pseudogene sequences share several of the amino acids that are exchanged in comparison with the amino acid sequence of the functional HSS, e.g., the position of the premature stop codon (Figs. 2 and 3A). Site-specific mutagenesis experiments have shown that not only the nonsense mutation resulting in the premature stop codon, but also other mutations are responsible for pseudogenization and consequential for the loss of enzymatic activity of the encoded protein.

Regarding plant pseudogenes, Drouin and Dover (1987) have suggested that processed pseudogenes in plants are genetically stable in cases when the gene with which they share a common ancestor is expressed in vegetative and reproductive meristems whose cells may ultimately differentiate into gametic mother cells (i.e., apical meristem). This hypothesis is the result of the observation that, within mammals, processed pseudogenes are expressed mainly in cells of the germ line where they accumulate through successive generations, whereas somatic cells seem to be devoid of processed pseudogenes. The presence of several HSS-related pseudogenes is consistent with this hypothesis, since HSS is highly expressed in the apical meristem of aerial roots and in young flower buds (Anke et al., 2008) suggesting a high stability for HSS pseudogenes. Remarkably, in this context, we have identified only pseudogenes of the gene encoding HSS, but none of the gene encoding DHS (Fig. 3). Whether pseudogenes have a function, such as the regulation of gene expression or of mRNA stability of its homologous coding gene as suggested by Korneev et al. (1999) and Hirotsune et al. (2003), respectively, is a matter of debate (Gray et al., 2006; Zheng and Gerstein, 2007).

3.2. Gene recruitment for PA-specific HSS

We have recently presented evidence that the gene coding for PA-specific HSS was recruited by duplication of the DHS-coding gene at least four times independently during angiosperm evolution. One of these gene duplications has been suggested to have occurred within the monocots (Reimann et al., 2004). With respect to the assumed simplicity of PA structures found within orchids, the identification of the cDNAs coding for HSS and DHS of *Phalaenopsis* allows us to speculate whether the recruitment of the gene coding for HSS as the first step in PA biosynthesis is an old or a recent feature among monocots. The neighbor-joining tree in Fig. 4 suggests that the gene duplication resulting in the establishment of an independent gene coding for HSS occurred early in the monocot lineage. The branching pattern of the tree is nicely supported by modern phylogenies of the monocots separating the Poaceae of the Poales order from the sequences identified from species belonging to the Zingiberals and Asparagales. These data suggest that the gene duplication of the DHS-coding gene generating the HSS-coding gene copy occurred early in the monocot lineage. An early origin is supported by the low degree of identity between

the HSS- and DHS-coding cDNAs of only 73% and 72% at the nucleotide and amino acid levels, respectively. A similar low degree of identity between these two paralogous sequences of single species has been observed in members of the Boraginales order (Reimann et al., 2004). However, the monocot lineage is, in contrast to dicots, monophyletic (Soltis and Soltis, 2000) and was among the first lineages of angiosperms to diversify (Chase, 2004). The age of the extant monocot lineages is estimated as between 127 and 176 million years (Bremer, 2002; Chase, 2004; Janssen and Bremer, 2004; Wikström et al., 2001), whereas the age of the order of the Boraginales is estimated as about 77–81 million years (Wikström et al., 2001). Because of the weak bootstrap support at the base of the tree (Fig. 4), conclusions concerning the sequence of gene duplication and speciation events in the early monocot evolution are inappropriate. Further analyses have to show whether the gene duplication resulting in HSS and DHS occurred early within a certain lineage, e.g. within the Asparalaes or even within the Orchidaceae. Recently, Koulman et al. (2008) have identified a pyrrolizidine alkaloid conjugate as a major component of perennial ryegrass (*Lolium perenne*, Poaceae, Pooideae). The necine base of these grass derived alkaloids possess the same absolute configuration as the PAs of *Phalaenopsis*, suggesting that even the common ancestor of orchids and grasses may have possessed the ability to produce PAs. HSS-coding cDNA of *L. perenne* and more DHS-homologuous sequences are needed from species belonging to the various families of the monocots for any conclusions to be made. These species should also be tested for the occurrence of PAs. This might aid the determination of whether the specific HSS-coding gene was a property of the common ancestor of all monocot species and was later lost in most of the extant lineages or whether the gene duplication resulting in the HSS-coding gene occurred early in the evolution of the lineage leading to the Orchidaceae family. Of importance, these sequences have to be experimentally expressed and functionally characterized to confirm their identity as HSS or DHS, a distinction that is impossible merely from sequence data (Ober and Hartmann, 1999b; Reimann et al., 2004), and to exclude related pseudogenes from the phylogeny. We have shown in this and in earlier studies (Nurhayati and Ober, 2005) that an array of HSS-related pseudogenes seem to be "by products" of gene duplication activity. This characterization is particularly required in those cases in which DHS is not represented by a single-copy gene in a specific genome. If only one copy of a gene with homology to DHS is identified within a genome, as is the case for all genomes analyzed to date from eukaryotes including rice (http://mips.gsf.de/proj/plant/jsf/rice/index.jsp), then this gene presumably encodes a functional DHS, as DHS has been shown to be an essential enzyme within eukaryotes (Park et al., 1997, 1998). In the case of *Zea mays*, we have established here the identity of the DHS homolog as being a functional DHS; this is, so far, the first functionally characterized DHS of the agriculturally important Poaceae family.

The molecular origin of HSS clearly shows that the assumed "simple structure" of the phalaenopsine-type PAs typical for Orchidaceae is not an argument for the recent origin of PA biosynthesis in the Orchidaceae. Of course, the finding that the HSS might be an ancient enzyme within monocots does not exclude the possibility that other enzymes involved in the biosynthesis of PAs in orchids are younger than the HSS. However, the particular structural features of the phalaenopsine-type PAs, such as the missing 1,2-double bond within the ring structure that is jointly responsible for the toxicity of many PAs or the observation that phalaenopsine PAs are frequently found in their tertiary form in plants, might also be the consequence of environmental factors that are different from those that shaped the well-characterized PA-defense system within *Senecio* species (for reviews see Hartmann, 1999; Hartmann and Ober, 2000; Ober, 2003). This idea finds further support following two observations: first, 1,2-saturated PAs are abundant in commercial cultivars of grasses, suggesting that these compounds have a positive effect on plant performance selected unwittingly by plant breeders (Koulman et al., 2008), and second, these PAs preferentially accumulate mainly in young organs, peripheral stem tissue, and the flower parts of orchids (Anke et al., 2008; Frölich et al., 2006), i.e., in those tissues with the highest probability of herbivore attack and with the highest value for reproductive fitness.

4. Experimental

4.1. Plant material

P. amabilis plants were purchased from Hennis Orchideen (Hildesheim, Germany). An interspecific hybrid of *Phalaenopsis equestris* x (*P. aphrodite* Rchb. F. x *P. mannii* Rchb. F.) was obtained as mature plants from Wichmann Orchideen (Celle, Germany). Plants were grown and maintained at 23 °C and approximately 50% humidity in a greenhouse. Protocorms of *Phalaenopsis* hybrids were obtained from Kaiser Orchideen (Schonungen, Germany), frozen in liquid nitrogen, and stored at −80 °C until used.

4.2. RNA isolation and cDNA synthesis

Total RNA was isolated from plant tissues (freshly harvested or frozen at −80 °C) by using the RNeasy Plant Mini Kit (Qiagen, Hilden, Germany). Total RNA (1 μg) was reverse-transcribed with an oligo$(dT)_{17}$ primer (0.1 μM, P5) and Superscript II reverse transcriptase (Invitrogen) in a total volume of 25 μl. All primer sequences are given in Supplementary Table S1.

4.3. Polymerase chain reaction with degenerate primers

Degenerate primers were designed according to an alignment of the amino acid sequences of DHSs and HSSs from various plants. *Phalaenopsis* cDNA (1 μl) was amplified in a standard reaction with *Taq*-DNA polymerase in a volume of 25 μl. A "touch down" temperature program (Don et al., 1991) with decreasing annealing temperatures from 60 °C to 45 °C (0.5 °C per cycle, 45 °C constant for 10 further cycles) was used. The primer combination P1/P2 and P3/P4 (1 μM, each) resulted in the amplification of 479-bp- and 656-bp fragments from the cDNA derived from aerial root tips and peduncles, respectively. After electrophoretic purification, the PCR products were subcloned by using the pCR2.1 TOPO TA Cloning Kit (Invitrogen) according to the manufacturer's protocol and sequenced.

4.4. Amplification of 3'- and 5'-cDNA ends

To identify the 3'- and 5'-ends of sequences, gene-specific primers were designed (P8 to P10 for HSS; P11 to P14 for DHS). Amplification of the 3'-cDNA fragments in a standard reaction containing 1 μl oligo$(dT)_{17}$-primed cDNA with primer pairs P8/P5 and P11/P5 (0.4 μM each) and *Taq*-DNA polymerase resulted in 535-bp and 872-bp fragments for HSS and DHS, respectively. To identify 5'-cDNA ends, total RNA was reverse-transcribed with primers P9 and P12 for HSS and DHS, respectively. The resulting cDNA was tailed at the 5'-end with dCTP before being used as a template for amplification with *Taq* DNA polymerase according to the Rapid Amplification of 5' cDNA Ends System (Invitrogen) with the primer pair P10/P6 in the case of HSS and with the primer pairs P13/P6 and P14/P7 in a nested PCR in the case of DHS (primers: 0.4 μM each). The resulting PCR fragments of 700 bp and 814 bp in length were electrophoretically purified and subcloned by using the pCR2.1 TOPO TA Cloning Kit (Invitrogen).

4.5. Amplification, expression, purification, and biochemical identification of Phalaenopsis recombinant HSS and DHS

A pair of gene-specific primers was synthesized for each cDNA (P15/P16 and P17/P18 for HSS and DHS, respectively) to amplify the full-length ORF. Primer pair P15/P16 contained *Nde*I and *Bam*HI restriction sites for the cloning of the cDNA into the pET3a expression vector; primer pair P17/P18 contained *Nde*I and *Xho*I restriction sites for cloning into the pET28a expression vector, which added a hexahistidine tag to the N-terminus of the protein during expression. Amplifications with *Pfx*-DNA polymerase (Invitrogen) and cloning for expression of the recombinant proteins in *E. coli* BL21(DE3) were performed as described by Ober and Hartmann (1999a) using a cDNA preparation of *P. amabilis*. The heterologously expressed HSS was purified to apparent homogeneity by the following procedure: Cells of a 100-ml *E. coli* BL21(DE3) culture expressing HSS of *Phalaenopsis* were harvested and sonicated, and the supernatant applied to a 2.5 × 5.0 cm DEAE-Fractogel column (Merck, Darmstadt) in buffer A (50 mM potassium phosphate pH 8.7, 2 mM dithioerythritol, 0.1 mM EDTA). Under these conditions, HSS did not bind to the matrix. The flow-through fraction was applied directly onto a Phenyl Sepharose CL-4B column (Amersham Biosciences) equilibrated with buffer B (5 mM potassium phosphate pH 8.7, 2 mM dithioerythritol, 0.1 mM EDTA, 1 M sodium chloride). Elution of HSS was achieved with buffer B without sodium chloride. Peak fractions were loaded onto a MonoQ column (Amersham Biosciences) equilibrated with buffer A and eluted with a gradient from 0 to 1.0 M sodium chloride. Active fractions were pooled and concentrated. Later in our project, HSS cDNA was cloned into the pET23a vector (primer pair P15/P19) for expression with a hexahistidine tag at the C-terminus of the recombinant protein to allow metal chelate affinity purification with Ni-NTA Agarose (Qiagen, Hilden, Germany) of both enzymes. Elution fractions were rebuffered to assay buffer (0.1 M glycine–NaOH buffer, pH 9.5, 0.1 mM dithiothreitol, 0.1 mM EDTA) by using PD10 columns (Amersham Biosciences) and assayed for HSS and DHS activities as described (Ober and Hartmann, 1999a).

4.6. Genomic DNA isolation and amplification

Small-scale genomic DNA isolation of *P. amabilis* (up tp 100 mg plant material) was conducted by using the DNeasy Plant Mini Kit (Qiagen) according to the manufacturer's instructions. Large scale genomic DNA isolation (up to 1 g plant material) was carried out with the CTAB method developed by Doyle and Doyle (1990). For PCR-based amplification of genomic DNA, ca. 300 ng DNA were set up in a 25 μl reaction mixture with *Pfu*-DNA Polymerase (Promega) and gene-specific primers. DNA fragments were subcloned into the pCR2.1 TOPO XL vector (Invitrogen).

4.7. Exclusive amplification of cDNA ends (EXACT) RT-PCR

Three 5' unique reverse primers (P36–P38) were synthesized to perform PCR amplification with oligo-dT primed cDNA in a two-step PCR as described by Smith et al. (2001). The 1st PCR was conducted in a 25-μl reaction mixture containing P35 (0.4 μM), P36 (0.02 μM) and P37 (0.04 μM), and *Taq*-DNA polymerase (Invitrogen). The 2nd PCR was performed in a 25-μl reaction mixture containing 1 μl diluted 1st PCR product with P35 (0.4 μM) and P38 (0.02 μM), and *Taq*-DNA polymerase.

4.8. Site-specific mutagenesis

To introduce specific mutations into the pseudogene sequences, a protocol was used according to the instructions of the QuickChange Site-Directed Mutagenesis Kit (Stratagene). To replace the premature stop codon (TGA) in the pseudogenes by the triplet encoding tryptophane (TGG), primer pair P31/P33 was used to mutate Ψ*hss2-Ph*/Ψ*hss3-Ph* and primer pair P32/P34 for Ψ*hss4-Ph*. After control sequencing, the mutated pseudogenes were expressed in *E. coli* BL21(DE3). Desalted crude cell extracts were used to assay the recombinant proteins for HSS and DHS activity.

4.9. Amplification, expression, and biochemical identification of maize recombinant DHS

We used EST data (accession No. CD440049 and CF243878) to assemble most of the cDNA coding for the putative DHS of *Zea mays* lacking approximately 60 amino acids at the N-terminus. Using primers P24–P26, we performed a 5'-RACE with cDNA that had been reverse-transcribed from seedling total RNA (*Z. mays* inbred line B73) resulting in a 762-bp fragment. Based on this sequence data, we designed a pair of specific primers (P27/P28) for the amplification of the full-length ORF. Primer P27 contained a *Nde*I restriction site, and primer P28 encoded a hexahistidine tag in front of a *Bam*HI restriction site that allowed cloning into the linearized pET3a expression vector. Expression in *E. coli* BL21 (DE3), metal chelate affinity purification, and enzyme assays were performed as described above. For the assays, the affinity-purified native eIF5A precursor protein of maize was employed. For this purpose, the ORF of the maize eIF5A cDNA was amplified by using primer pair P29/P30 designed according to the previously published sequence (Dresselhaus et al., 1999) and was cloned into the vector pET23b for expression and metal chelate affinity purification.

4.10. Identification and expression of Phalaenopsis eIF5A precursor protein

To identify the cDNA coding for the eIF5A precursor protein of the *Phalaenopsis* hybrid two degenerate primers (P20, P21) were designed based on an alignment of plant eIF5A precursor proteins and resulted in the amplification of a 353-bp fragment from cDNA of *Phalaenopsis* leaves. To perform 3'-RACE PCR, a gene-specific primer (P22) was designed with a *Nde*I restriction site containing the start ATG of the ORF. Combination of this primer with the oligo$(dT)_{17}$ primer (P5, 0.4 μM each) in a standard PCR resulted in a 708-bp fragment including the ORF and the 3'-UTR. Using a second gene-specific primer with a *Xho*I restriction site instead of the stop codon (P23) in combination with P22 the whole ORF of 493 bp was amplified and cloned into the *Nde*I/*Xho*I-linearized pET23a expression vector (Novagen) for expression in *E. coli* BL21(DE3) with a C-terminal hexahistidine motive.

4.11. Computer-assisted sequence analyses

Sequence analyses were performed by using the GCG Wisconsin software package. Nucleotide sequences were aligned by using ClustalX (Thompson et al., 1997). Only the central part of the alignment of 782 bp in length (the alignment is available in supplementary Fig. S1) was used for phylogenetic analyses with the following software from the PHYLIP program package (Felsenstein, 2001): DNADIST with the two-parameter model of Kimura (Kimura, 1980) in combination with NEIGHBOR and DNAML (Saitou and Nei, 1987). For all algorithms, the input order of the sequences was randomized, and global rearrangements were allowed. Bootstrap values were estimated from 1000 replicates with the programs SEQBOOT and CONSENSE of the PHYLIP package.

4.12. Accession numbers

Sequence data from this article have been deposited in EMBL/GenBank: *P. amabilis* HSS genomic DNA (**FM201299**), *P. amabilis*

DHS cDNA (**AM411616**), *Phalaenopsis* spec. eIF5A precursor cDNA (**AM411618**), and *Zea mays* DHS cDNA (**AM411617**). Accession numbers for EST data taken from the databases are given in Supplementary Fig. S1.

Acknowledgements

We are grateful to Thomas Hartmann for helpful discussions of the data, to Orchideen-Kaiser (Schonungen, Germany) for providing us with the protocorms of *Phalaenopsis*, and to Monika Frey (Technical University Munich) and Jörg Degenhardt (University of Halle) for providing us with the maize seeds. We thank A. Backenköhler and D. Glindemann for their excellent technical assistance. This work was supported by the grant OB162/6 of the Deutsche Forschungsgemeinschaft Priority Program 1152 'Evolution of Metabolic Diversity' (grant to D.O.) and the German Academic Exchange Service (DAAD, scholarship to N.N.).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.phytochem.2009.01.019.

References

Anke, S., Gonde, D., Kaltenegger, E., Hansch, R., Theuring, C., Ober, D., 2008. Pyrrolizidine alkaloid biosynthesis in phalaenopsis orchids: developmental expression of alkaloid-specific homospermidine synthase in root tips and young flower buds. Plant Physiol. 148, 751–760.

Balakirev, E.S., Ayala, F.J., 2003. Pseudogenes: are they "junk" or functional DNA? Annu. Rev. Genet. 37, 123–151.

Borba, E.L., Trigo, J.R., Semir, J., 2001. Variation of diastereoisomeric pyrrolizidine alkaloids in *Pleurothallis* (Orchidaceae). Biochem. Syst. Ecol. 29, 45–52.

Bremer, K., 2002. Gondwanan evolution of the grass alliance of families (Poales). Evolution 56, 1374–1387.

Cameron, K.M., 2004. Utility of plastid psaB gene sequences for investigating intrafamilial relationships within Orchidaceae. Mol. Phylogenet. Evol. 31, 1157–1180.

Cameron, K.M., Chase, M.W., Whitten, W.M.M., Kores, P.J., Jarrell, D.C., Albert, V.A., Yukawa, T., Hills, H.G., Goldman, D.H., 1999. A phylogenetic analysis of the Orchidaceae: evidence from *rbcL* nucleotide sequences. Am. J. Bot. 86, 208–224.

Chase, M.W., 2004. Monocot relationships: an overview. Am. J. Bot. 91, 1645–1655.

Chase, M.W., Cameron, K.M., Hills, H.G., Jarrell, D.C., 1994. DNA sequences and phylogenetics of the Orchidaceae and other lilioid monocots. In: Pridgeon, A. (Ed.), Proceedings of the Fourteenth World Orchid Conference. Her Majesty's Stationary Office, Galsgow, pp. 61–73.

Chattopadhyay, M.K., Park, M.H., Tabor, H., 2008. Hypusine modification for growth is the major function of spermidine in *Saccharomyces cerevisiae* polyamine auxotrophs grown in limiting spermidine. Proc. Natl. Acad. Sci. USA 105, 6554–6559.

Don, R.H., Cox, P.T., Wainwright, B.J., Baker, K., Mattick, J.S., 1991. 'Touchdown' PCR to circumvent spurious priming during gene amplification. Nucleic Acids Res. 19, 4008.

Doyle, J.J., Doyle, J.L., 1990. Isolation of plant DNA from fresh tissue. Focus 12, 13–15.

Dresselhaus, T., Cordts, S., Loerz, H., 1999. A transcript encoding translation initiation factor eIF-5A is stored in unfertilized egg cells of maize. Plant Mol. Biol. 39, 1063–1071.

Drouin, G., Dover, G.A., 1987. A plant processed pseudogene. Nature 328, 557–558.

Duvall, M.R., Morton, B.R., 1996. Molecular phylogenetics of Poaceae: an expanded analysis of *rbcL* sequence data. Mol. Phylogenet. Evol. 5, 352–358.

Felsenstein, J., 2001. PHYLIP, Phylogeny Inference Package, Version 3.6(alpha2), Univ. of Washington, Seattle.

Frölich, C., Hartmann, T., Ober, D., 2006. Tissue distribution and biosynthesis of 1,2-saturated pyrrolizidine alkaloids in Phalaenopsis hybrides (Orchidaceae). Phytochemistry 67, 1493–1502.

Fu, P.P., Xia, Q., Lin, G., Chou, M.W., 2004. Pyrrolizidine alkaloids – genotoxicity, metabolism enzymes, metabolic activation, and mechanisms. Drug Metab. Rev. 36, 1–55.

Gilbert, N., Lutz-Prigge, S., Moran, J.V., 2002. Genomic deletions created upon LINE-1 retrotransposition. Cell 110, 315–325.

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.

Gray, T.A., Wilson, A., Fortin, P.J., Nicholls, R.D., 2006. The putatively functional Mkrn1–p1 pseudogene is neither expressed nor imprinted, nor does it regulate its source gene in trans. Proc. Natl. Acad. Sci. USA 103, 12039–12044.

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, pp. 207–244.

Hartmann, T., Witte, L., 1995. Chemistry, biology and chemoecology of the pyrrolizidine alkaloids. In: Pelletier, S.W. (Ed.), Alkaloids: Chemical and Biological Perspectives, vol. 9. Pergamon Press, Oxford, pp. 155–233.

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.

Ikeda, Y., Nonaka, H., Furumai, T., Igarashi, Y., 2005. Cremastrine, a pyrrolizidine alkaloid from *Cremastra appendiculata*. J. Nat. Prod. 68, 572–573.

Janssen, T., Bremer, K., 2004. The age of major monocot groups inferred from 800+*rbcL* sequences. Bot. J. Linn. Soc. 146, 385–398.

Jenett-Siems, K., Kaloga, M., Eich, E., 1993. Ipangulines, the first pyrrolizidine alkaloids from the Convolvulaceae. Phytochemistry 34, 437–440.

Kimura, M., 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. 16, 111–120.

Korneev, S.A., Park, J.-H., O'Shea, M., 1999. Neuronal expression of neural nitric oxide synthase (nNOS) protein is suppressed by an antisense RNA transcribed from an NOS pseudogene. J. Neurosci. 19, 7711–7720.

Koulman, A., Seeliger, C., Edwards, P.J.B., Fraser, K., Simpson, W., Johnson, L., Cao, M., Rasmussen, S., Lane, G.A., 2008. *E/Z*-Thesinine-*O*-4'-α-rhamnoside, pyrrolizidine conjugates produced by grasses (Poaceae). Phytochemistry 69, 1927–1932.

Kyrpides, N.C., Woese, C.R., 1998. Universally conserved translation initiation factors. Proc. Natl. Acad. Sci. USA 95, 224–228.

Maestre, J., Tchenio, T., Dhellin, O., Heidmann, T., 1995. MRNA retroposition in human cells: processed pseudogene formation. EMBO J. 14, 6333–6338.

Magdolen, V., Klier, H., Woehl, T., Klink, F., Hirt, H., Hauber, J., Lottspeich, F., 1994. The function of the hypusine-containing proteins of yeast and other eukaryotes is well conserved. Mol. Gen. Genet. 244, 646–652.

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.

Nishimura, K., Murozumi, K., Shirahata, A., Park, M.H., Kashiwagi, K., Igarashi, K., 2005. Independent roles of eIF5A and polyamines in cell proliferation. Biochem. J. 385, 779–785.

Nurhayati, N., Ober, D., 2005. Recruitment of alkaloid-specific homospermidine synthase (HSS) from ubiquitous deoxyhypusine synthase: does *Crotalaria* possess a functional HSS that still has DHS activity? Phytochemistry 66, 1346–1357.

Ober, D., 2003. Chemical ecology of alkaloids exemplified with the pyrrolizidines. In: Romeo, J.T. (Ed.), Integrative Phytochemistry: From Ethnobotany to Molecular Ecology, vol. 37. Pergamon, Amsterdam, pp. 203–230.

Ober, D., Gibas, L., Witte, L., Hartmann, T., 2003a. Evidence for general occurrence of homospermidine in plants and its supposed origin as by-product of deoxyhypusine synthase. Phytochemistry 62, 339–344.

Ober, D., Harms, R., Witte, L., Hartmann, T., 2003b. Molecular evolution by change of function: alkaloid-specific homospermidine synthase retained all properties of deoxyhypusine synthase except binding the eIF5A precursor protein. J. Biol. Chem. 278, 12805–12812.

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

Park, M.H., Joe, Y.A., Kang, K.R., 1998. Deoxyhypusine synthase activity is essential for cell viability in the yeast *Saccharomyces cerevisiae*. J. Biol. Chem. 273, 1677–1683.

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

Park, M.H., Wolff, E.C., Folk, J.E., 1993. Hypusine: its post-translational formation in eukaryotic initiation factor 5A and its potential role in cellular regulation. Biofactors 4, 95–104.

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.

Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4, 406–425.

Salamin, N., Hodkinson, T.R., Savolainen, V., 2002. Building supertrees: an empirical assessment using the grass family (Poaceae). Syst. Biol. 51, 136–150.

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.

Soltis, E.D., Soltis, P.S., 2000. Contributions of plant molecular systematics to studies of molecular evolution. Plant Mol. Biol. 42, 45–75.

Stewart, M.J., Steenkamp, V., 2001. Pyrrolizidine poisoning: a neglected area in human toxicology. Ther. Drug Monit. 23, 698–708.

Symer, D.E., Connelly, C., Szak, S.T., Caputo, E.M., Cost, G.J., Parmigiani, G., Boeke, J.D., 2002. Human l1 retrotransposition is associated with genetic instability *in vivo*. Cell 110, 327–338.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The CLUSTAL_X windows interface. Flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25, 4876–4882.

Vanin, E.F., 1985. Processed pseudogenes: characteristics and evolution. Annu. Rev. Genet. 19, 253–272.

Wang, T.-W., Lu, L., Zhang, C.-G., Taylor, C., Thompson, J.E., 2003. Pleiotropic effects of suppressing deoxyhypusine synthase expression in *Arabidopsis thaliana*. Plant Mol. Biol. 52, 1223–1235.

Wang, T.-W., Zhang, C.-G., Wu, W., Nowack, L.M., Madey, E., Thompson, J.E., 2005. Antisense suppression of deoxyhypusine synthase in tomato delays fruit softening and alters growth and development. Plant Physiol. 138, 1372–1382.

Wang, T.W., Lu, L., Wang, D., Thompson, J.E., 2001. Isolation and characterization of senescence-induced cDNAs encoding deoxyhypusine synthase and eucaryotic translation initiation factor 5A from tomato. J. Biol. Chem. 276, 17541–17549.

Wikström, N., Savolainen, V., Chase, M.W., 2001. Evolution of the angiosperms: calibrating the family tree. Proc. R. Soc. Lond., B, Biol. Sci. 268, 2211–2220.

Zhang, Y., Wu, Y., Liu, Y., Han, B., 2005. Computational identification of 69 retroposons in *Arabidopsis*. Plant Physiol. 138, 935–948.

Zheng, D., Gerstein, M.B., 2007. The ambiguous boundary between genes and pseudogenes: the dead rise up, or do they? Trends Genet. 23, 219–224.