PERGAMON

Phytochemistry 62 (2003) 513–526

PHYTOCHEMISTRY

www.elsevier.com/locate/phytochem

A family of polyketide synthase genes expressed in ripening *Rubus* fruits

Amrita Kumar[1], Brian E. Ellis*

The Biotechnology Laboratory and Faculty of Agricultural Sciences, University of British Columbia, Bioscience Building, Rm 3508, 6270 University Blvd, Vancouver V6T 1Z4, Canada

Received 26 August 2002; received in revised form 18 October 2002

Dedicated to Meinhart H. Zenk on the occasion of his 70th birthday

Abstract

Quality traits of raspberry fruits such as aroma and color derive in part from the polyketide derivatives, benzalacetone and dihydrochalcone, respectively. The formation of these metabolites during fruit ripening is the result of the activity of polyketide synthases (PKS), benzalcetone synthase and chalcone synthase (CHS), during fruit development. To gain an understanding of the regulation of these multiple PKSs during fruit ripening, we have characterized the repertoire of *Rubus PKS* genes and studied their expression patterns during fruit ripening. Using a PCR-based homology search, a family of ten *PKS* genes (*Ripks1-10*) sharing 82–98% nucleotide sequence identity was identified in the *Rubus idaeus* genome. Low stringency screening of a ripening fruit-specific cDNA library, identified three groups of *PKS* cDNAs. Group 1 and 2 cDNAs were also represented in the PCR amplified products, while group 3 represented a new class of *Rubus* PKS gene. The *Rubus* PKS gene-family thus consists of at least eleven members. The three cDNAs exhibit distinct tissue-specific and developmentally regulated patterns of expression. *RiPKS5* has high constitutive levels of expression in all organs, including developing flowers and fruits, while *RiPKS6* and *RiPKS11* expression is consistent with developmental and tissue-specific regulation in various organs. The recombinant proteins encoded by the three *RiPKS* cDNAs showed a typical CHS-type PKS activity. While phylogenetic analysis placed the three *Rubus* PKSs in one cluster, suggesting a recent duplication event, their distinct expression patterns suggest that their regulation, and thus function(s), has evolved independently of the structural genes themselves.

Keywords: *Rubus idaeus*; Raspberry; Fruit development; Gene families; Benzalacetone; Polyketide synthase; Chalcone synthase; Benzalacetone synthase

1. Introduction

Plant-specific polyketide synthases (PKSs) are dimeric proteins that act directly on Coenzyme A (CoA) thioesters of various carboxylic acids to extend the carbon backbone (reviewed in Schröder, 1997; Dixon, 1999; Jez et al., 2001). Structural diversity within the polyketides is primarily attributed to the "synthase programming" in PKSs that determines the chain length, the choice of starter and extender units, and the subsequent modification and cyclization of the carbon chain (Hopwood and Sherman, 1990; Schröder, 1997). Numerous plant-specific PKSs have been characterized which collectively orchestrate the successive condensation of malonyl CoA with aromatic CoA-thioesters to yield a large family of compounds that include chalcone, stilbene, bibenzyl, acridone, and benzalacetone derivatives. The more complex PKSs such as chalcone synthase (CHS), stilbene synthase (STS), and coumaroyltriacetic acid synthase carry out multiple rounds of condensation with 4-coumaryl CoA. Other plant PKSs, however, catalyze fewer condensation cycles, or use other starter units, to yield structurally distinct polyketides (reviewed in Schröder, 1997; Dixon and Steele, 1999; Dixon, 1999) (Fig. 1). Styrylpyrone synthase (SPS), for example, uses the same precursors as are used by CHS, but carries out only two condensing reactions prior to ring closure

* Corresponding author. Tel.: +1-604-822-3451; fax: +1-604-822-2114.

E-mail address: bee@interchg.ubc.ca (B.E. Ellis), akumar@microbio-emory.edu (A. Kumar).

[1] Present address: Department of Microbiology and Immunology, Emory University School of Medicine, Atlanta 30322, USA.

0031-9422/03/$ - see front matter
PII: S0031-9422(02)00572-1

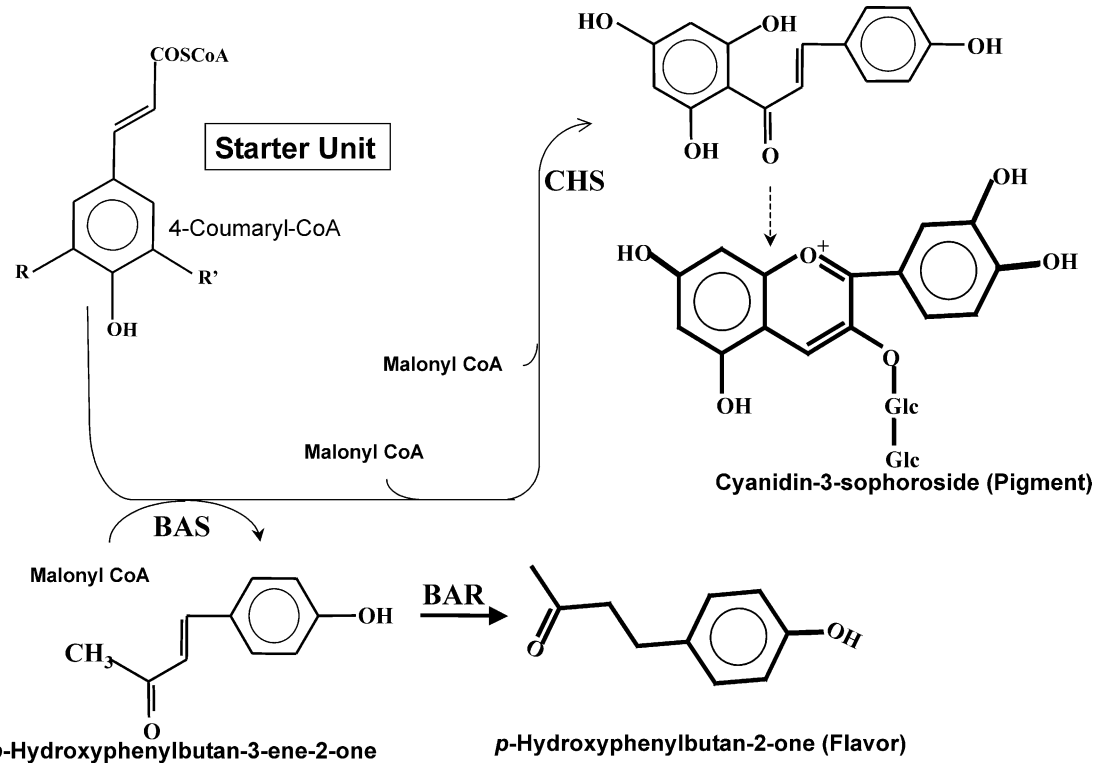

Fig. 1. Origin of the raspberry flavor and color polyketide-derivatives. CHS, chalcone synthase; BAS, benzalacetone synthase.

(Beckert et al., 1997), while pyrone synthase (PS) catalyzes the formation of pyrone derivatives in *Gerbera* by using a starter unit of acetyl CoA and extending the skeleton with two units of malonyl CoA (Eckermann et al., 1998).

Comparative analysis of plant *PKSs* reveals substantial similarity at the nucleotide and protein sequence level. All plant *PKSs* consist of ~400 residue polypeptide chains that share >50% sequence identity. However, functional analysis of some of these sequences using either genetic or biochemical approaches has shown that, despite the considerable sequence similarity, PKSs can exhibit very distinct "synthase programming", and thus, catalytic capabilities. For example, within a family of three *PKS* genes isolated from *Gerbera hybrida*, *gCHS1* and *gCHS3* share 89% amino acid sequence identity and show a typical CHS- type activity, while a third gene (*gPS2*) that shares ~75% amino acid sequence identity to *gCHS1* and *gCHS3*, is a PS-type PKS (Helariutta et al., 1995; Eckermann et al., 1998). Similarly, the two *PKS* genes in *Pinus strobus* share 87% amino acid sequence identity, yet the products of PStrCHS1 performed a typical CHS reaction, while PStrCHS2 was completely inactive with the usual CHS starter substrates and preferred to use a diketide derivative analogous to a CHS-intermediate (Schröder et al., 1998). Thus, it is now generally accepted that the catalytic program of a plant-PKS cannot be predicted solely from sequence information.

While a large repertoire of *PKS* genes sharing the same or divergent catalytic functions has been isolated, the functional basis for this diversity has been analyzed only in a very few species. Large families of *PKS* genes have been cloned from plants such as *Petunia* (Koes et al., 1987), bean (Ryder et al., 1987), pea (Harker et al., 1990; An et al., 1993), alfalfa (Junghans et al., 1993), soybean (Akada et al., 1991; Estabrook and Gopalan-Sengupta, 1991), and *Ipomea* (Durbin et al., 1995), while smaller gene families have been cloned from species such peanut (Schröder et al., 1988), tomato (O'Neill et al., 1990), mustard (Batschauer et al., 1991), *Mathhiola* (Epping et al., 1990), maize (Franken et al., 1991), barley (Rhode et al., 1991), gerbera (Helariutta et al., 1995). Only single *PKS* genes have been reported as yet from *Antirrhinum* (Sommer and Saedler, 1986), *Arabidopsis* (Feinbaum and Ausubel, 1988) and *Petroselinum* (Reimold et al., 1983). A family of *PKS* genes has also been cloned from gymnosperms (Raiber et al., 1995; Schanz et al., 1992; Schröder et al., 1998). The presence of apparently single *CHS* genes in *Arabidopsis, Antirrhinum* and *Peteroselinum* suggests that possession of multiple *PKS* genes is not essential for the survival of the plants.

In some species, the presence of multiple *PKS* genes has been correlated with distinct capacity to code for different but related enzymatic activities and such genes may be regulated in a tissue- and development-specific fashion. In *Antirrhinum* (Sommer and Saedler, 1986)

and *Arabidopsis* (Shirley et al., 1995) expression of the *CHS* gene has been correlated with anthocyanin pigmentation during flower development or with pigmentation of the seed coat, respectively, while in maize, *CHS* gene expression varies during kernel development (Franken et al., 1991). Similarly, in *Gerbera hybrida*, the expression of the two *CHS* genes has been correlated with the accumulation of flavonols, while the pyrone synthase (PS) was expressed throughout the plant (Helariutta et al., 1995). In grape, pine and peanut cell-cultures, it has been demonstrated that the PKS belonging to the STS-class is preferentially induced in response to elicitation as compared to CHS in the same species (Schoeppner and Kindl, 1979; Rolfs et al., 1981; Rolfs and Kindl, 1984). This diversity emphasizes the importance of characterization of all *PKS* gene-family members in a single plant species if we are to fully understand the functional divergence and the contribution to cell function(s) of the PKSs.

In addition, characterization of the full repertoire of *PKS* genes in a genome provides further insight into the evolution of this important family of proteins. Based on the phylogenetic analysis of all PKSs characterized in plants, Tropf et al. (1994) hypothesized that divergent members of the plant PKS superfamily had arisen independently from a CHS-type PKS in several species. However, in a more recent phylogenetic analysis of STS and CHS in *Vitis,* Goodwin et al. (2000) proposed that *Vitis*-STS has a separate origin from *Vitis*-CHSs, and that evolution of PKSs may have followed multiple routes.

In *Rubus*, polyketide derivatives such as anthocyanins are an obvious feature of ripening fruits. The starting materials for the biosynthesis of anthocyanins are the products of enzyme CHS (Fig. 1). The fruits also accumulate another polyketide-derivative, p-hydroxyphenylbut-2-one (pHPB), which is primarily responsible for the characteristic aroma of ripe raspberry fruits (Schinz and Sediel, 1961) (Fig. 1). Biosynthesis of pHPB follows a two-step process in which one unit of *p*-coumaryl CoA is first condensed with one-unit of malonyl CoA to form *p*-hydroxyphenylbut-3-ene-2-one (BA), which is then reduced by a separate oxidoreductase in the presence of NADPH to yield pHPB (Borejsza-Wysocki and Hrazdina, 1996). Formation of BA, catalyzed by benzalacetone synthase (BAS) is reminiscent of the biosynthesis of naringenin chalcone, and partial purification of raspberry BAS protein revealed that it has a M_r, pI, and pH optimum similar to that of CHS (Borejsza-Wysocki and Hrazdina, 1996) (Fig. 1). The authors therefore suggested that BAS, like STS, ACS and other PKS, might be a unique PKS that shares significant similarities to CHS. The recent isolation of BAS from *Rheum palmatum* confirmed that the *Rheum* enzyme is closely related to CHS at the molecular level, but displays a novel catalytic specificity (Abe et al., 2001).

To learn how the capacity to form fruit-associated polyketides is organized in *Rubus*, we have characterized the *PKS* gene-family in this species, and studied the developmental regulation of *PKS* during fruit ripening.

2. Results and discussion

2.1. Amplification and characterization of the Rubus PKS genes

To identify the full complement of *Rubus PKS* gene(s), we initially used a PCR-based approach, employing degenerate PCR primers designed to target evolutionarily conserved regions within a range of PKSs with distinct catalytic capabilities. Despite differences in substrates and reaction mechanisms utilized by plant-specific PKSs, alignment of the different classes of PKS sequences revealed several such conserved regions.

Genomic DNA isolated from young *Rubus* leaf tissue was then used as a template for amplification of *PKS* genes, employing using combinations of degenerate primers PK1–PK2, PK1–PK4, PK3–PK2, and PK3–PK4 (Fig. 2). Amplification with primer pairs PK1–PK2, PK3–PK2, and PK3–PK4 in each case yielded products of expected size. However, amplifications with combination PK1–PK4 yielded either attenuated or no amplification products, and therefore this combination was not considered further. RFLP analysis of about 500 independent clones of the amplicons obtained from each primer pair revealed polymorphism for *Sma*I and *Pst*I. This analysis led to the identification of five classes of amplicon clones from primer pairs PK1–PK2, four classes from primer pairs PK3–PK2, and one class from primer pair PK3–PK4. The *PKS* gene-family in a given plant species can consist of multiple members that share considerable sequence homology, as observed in alfalfa (Junghans et al., 1993), *Gerbera* (Helariutta et al., 1995, 1996), *Ipomea* (Durbin et al., 1995), petunia (Koes et al., 1987), and soybean (Akada and Dube, 1995). It is therefore conceivable that two or more genes could produce identical RFLP patterns with these two restriction enzymes. To test for that possibility, we sequenced multiple clones representing each RFLP class, and compared a stretch of 351 nucleotides (117 aa) within all the sequenced clones, using the CLUSTAL program in PC/GENE. *Rubus idaeus* is vegetatively propagated and, it is likely that *R. idaeus* L. cv. Meeker consists of a homogenous but heterozygous population. To accommodate the contribution to sequence variation of co-amplification of allelic sequences, as well as any errors introduced by *Taq* DNA Polymerase during amplification, we chose an arbitrary variability threshold (10 nucleotides out of 351 nucleotides compared), below which all sequence variability was ascribed to these factors. Using this threshold as our

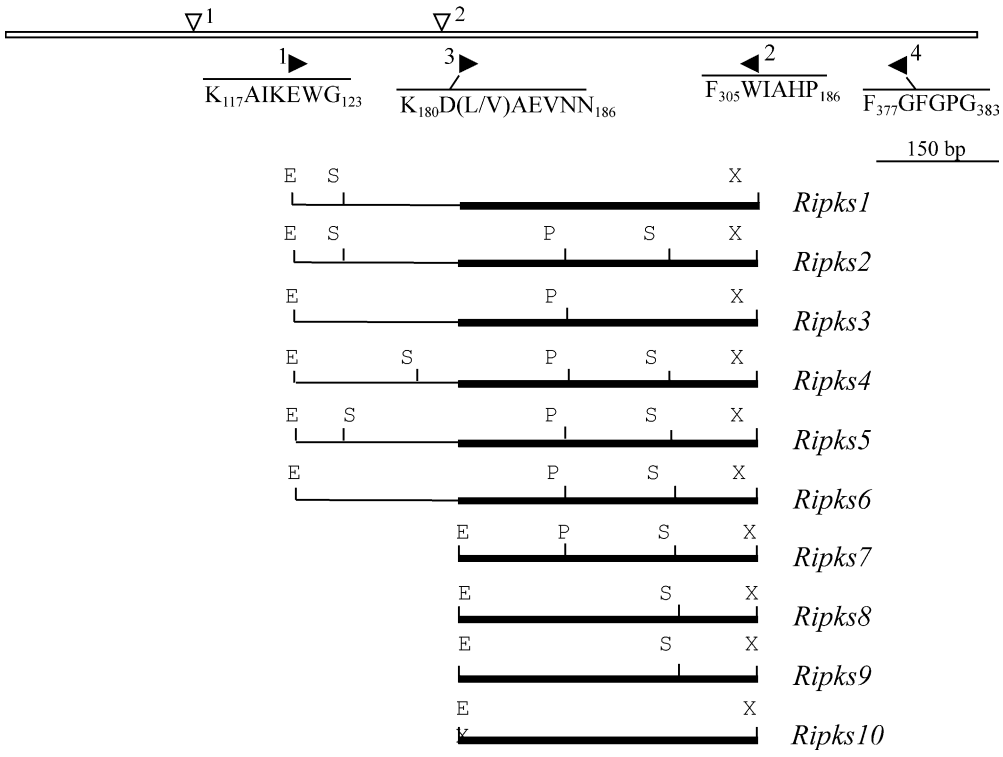

Fig. 2. Positions of the *PKS* gene-specific primers, and of the amplified *Rubus PKS* clones, relative to a generic plant *PKS* gene. Closed arrow heads indicate the positions of the degenerate PCR primers, relative to the coding regions of a generic plant *PKS* gene. The amino acids position corresponding to each primer have been numbered relative to *Arabidopsis* CHS (GenBank, M20308. ▽[1]represents the conserved intron found in all plant-specific *PKS* genes. ▽[2] represents the additional intron characterized in *Antirrhinum* CHS (X03710). The amplified regions of the *Rubus PKS* genes are indicated by solid lines. *Ripks1-Ripks6* were detected by primers PK1-PK2, while *Ripks7-Ripks10* were generated from primer combination PK3-PK4. The dark line indicates the region (351 bp) compared and analyzed in detail. Amplified products had an *Eco*RI (E) and *Xba*I (X) site at their 5′ and 3′-ends respectively. The polymorphic *Pst*I (P) and *Sma*I (S) sites within the amplified regions of the *Rubus pks* genes were used to classify the PCR clones.

standard for distinguishing between genes and alleles, we were able to characterize fragments potentially representing at least ten different *PKS* genes within the *Rubus* genome.

2.2. Rubus PKS sequence classes

Within the characterized regions, the *Ripks1* sequence class had a recognition site for *Sma*1 (60 bp downstream of the primer PK1 binding site), and lacked any recognition site for *Pst*I. *Ripks1* was represented by 11 sequenced clones obtained from two independent PCR reactions (denoted A and B) with primer pairs PK1-PK2. Five out of the eleven clones sequenced were identical over the 351 bp compared and this sequence was therefore designated *Ripks1*. Six clones carried the following changes (designated within brackets): clone A58 (A34T, A62G), clone B27 (T28C, G148A, A222G), clone B240 (G31A), clone B49 (C39T, G85A, A94G), clone B03 (G44A, (C201T) and clone B164 (T13G, A61G, A93G, C196T). Amongst these six clones no base substitutions occurred at the same positions, therefore we conclude that all clones represent a single *PKS* gene with variations attributed to amplifications of allelic sequences or to errors introduced by *Taq* DNA polymerase. However, it is conceivable that clone B164 represents a different locus, since its variation exceeds the error rate of 0.5% estimated by Butland et al. (1998) for *Taq* DNA polymerase in a similar analysis.

The *Ripks2* gene fragment contained two recognition sites for *Sma*I and a single recognition site for *Pst*I (Fig. 2). It was represented by clones A43 and B157, obtained from the primer pair combination PK1–PK2. These two clones were identical over the 351 nucleotides compared and this sequence was designated *Ripks2*.

Ripks3 had a recognition site for *Pst*I located 308 nucleotides downstream of the primer PK1 binding site. It was represented by two clones, A125 and B113, obtained from the primer pair combination PK1–PK2. The two sequences were identical over the 351 nucleotide region compared and this sequence was therefore designated as the *Ripks3* sequence.

Ripks4 had dual recognition sites for *Sma*I (144 and 462 nucleotides downstream of the primer PK1 binding site) and a single *Pst*I site in common with some of the other *Ripks* genes (Fig. 2). The *Ripks4* sequence was

represented by three clones: A53, B48 and A166, obtained from primer combination PK1–PK2. While two of these clones were identical in sequence (this sequence was designated as *Ripks4*) one clone (A48) had changes at nucleotide positions 49 (C→T), 61 (T→C), 67 (C→T), 116(A→T), 210 (C→A), 219 (T→C). Since this level of change is greater than the estimated level of amplification error, it is likely that clone A48 represents the allelic sequence of *Ripks4*.

Ripks5 was represented by five clones obtained from primer pair PK1-PK2. It had a fingerprint with *Sma*I and *Pst*I similar to that of *Ripks2*, but differed at 14 nucleotide positions from *Ripks2*. Since this level of variation is higher than the estimated amplification error-rate, we designated this sequence as originating from a new gene. However, we cannot exclude the possibility that *Ripks5* represents an allele of *Ripks2*. Four of the five clones were identical in the 351 bp region compared and this sequence was thus designated as the *Ripks5* sequence. One clone (A83) of the same RFLP class had a single nucleotide substitution at position 340 (T→A).

Ripks6 was amplified by all three primer pair combinations. This RFLP class, represented by a single recognition site for *Pst*I and *Sma*I (Fig. 2), was the most prominent class detected in the analysis. It was represented by two sequenced clones from primer pair PK1–PK2, four sequenced clones from primer pair PK3–PK2 and seven clones from the primer pair PK3–PK4. Two clones from primer combination PK1–PK2, three clones from primer pair PK3–PK2 and five clones from primer combination PK3–PK4 were identical over the 351 nucleotides compared; hence this sequence was designated as *Ripks6*. However, in other clones with the same RFLP pattern differences were detected. For example, clone B263 from primer pair PK3–PK2 had changes at nucleotide positions 14 (T→C), 62 (G→T), 75 (G→T) and clones A94 and B22 from primer pair PK3-PK4 had single base substitutions each at nucleotide positions 12 (T→C), 142 (C→T), and 289 (A→G). Since no base substitutions occurred at the same positions in these clones, we conclude that all these clones represent a single *PKS* gene with variations attributable to amplifications of allelic sequences or to the error introduced by *Taq* DNA polymerase.

Ripks7 had single recognition sites for *Pst*I and for *Sma*I (Fig. 2), and was represented by two clones that had identical amino acid sequences over the 351 bp region compared. This sequence was designated *Ripks7*. Even though this class of clones shared common restriction sites for *Pst*I and *Sma*I with *Ripks4*, *5* and *6*, the nucleotide sequence varied at more than 14 nucleotide positions (over 351 nucleotides compared) when compared to *Ripks4*, *5* and *6*. We therefore assigned this to a different gene family member.

Sequence class *Ripks8* did not have a recognition site for *Pst*I (Fig. 2) and was represented by two sequences from two independent PCR reactions with primer pair PK3–PK2. These sequences were identical over the 351 bp compared and this sequence was designated *Ripks8*.

Classes *Ripks9* and *Ripks10* were each represented by multiple clones with a similar RFLP pattern arising from two independent PCR reactions with primer pairs PK3–PK2. We sequenced two representative clone of each class. Each of these sequences varied at least 15 nucleotide positions within the 351 bp region when compared to *Ripks1-6*. Thus, according to our criteria, we designate these sequences as originating from two different genes.

Using a homology-based search method we thus established that there are at least ten diverse *PKS* genes in the *Rubus* genome. The homology between these genes ranges from 86–96% nucleotide sequence identity in the 351 nucleotide region upstream of the PK2 primer binding site, which provides strong correlative evidence that they encode *Rubus* PKSs. Large PKS gene families have also been isolated from other plants. Up to eight genes classified as *CHS* (based on homology) have been reported from legumes such as bean (Ryder et al., 1987), soybean (Estrabrook and Gopalan-Sengupta, 1991), and alfalfa (Junghans et al., 1993).

PKSs other than CHS have also been reported to be encoded by a family of homologous genes. In *Vitis*, at least seven closely linked *STS* genes have been reported (Wiese et al., 1994), while in gymnosperms ten *STS* genes have been isolated and characterized from *P. sylvestris* (Preisig-Muller et al., 1995). Similarly, a family of three transcriptionally-active *PS* genes has been characterized in *Gerbera* (Helariutta et al., 1996). Only *Arabidopsis* and *Antirrhinum* have been proposed to contain single copies of *CHS*-type PKS genes (Sommer and Saedler, 1986; Shirley-Winkel et al., 1995). However, this conclusion may be premature, since a computer-based search of the *Arabidopsis* genome database reveals the presence of several sequences that share homology with the CHS consensus. Weinand et al. (1982) report the presence of several faint bands in a Southern blot analysis of *Antirrhinum* genomic DNA with a *Petroselinum CHS* fragment as a probe, suggesting that *Antirrhinum* might also have additional copies of *PKS*-type genes.

2.3. Isolation and characterization of ripening-related *Rubus PKS* genes

Since we were interested in ripening-related *PKSs*, we proceeded to isolate full-length cDNA sequences corresponding to those *RiPKS* genes that were expressed during fruit-ripening. Screening of a fruit-specific cDNA library with a mixed population of probes consisting of amplified fragments of *Ripks1* and *Ripks7* led to the isolation of 150 positive clones after the initial round of screening, 80 of which were subjected to further analysis. Amplifications of the insert in each positive

plaque with universal primer T3 and a PKS-gene-specific primer (PK4), identified 72 plaques as having potential *PKS* gene inserts >900 bp in length. Restriction fragment length analysis of the amplified fragments of the 72 clones with multiple restriction endonucleases grouped all clones into three discrete classes (Fig. 3). The three *RiPKS* cDNAs share considerable sequence homology amongst themselves and with other *PKSs*. The high homology within the coding regions of the *Rubus PKS* sequences also extended to the 5′ and the 3′-UTRs. However, subtle differences in these regions between the three cDNA clones suggest that each of the clones can be considered as arising from individual genes rather than alleles of the same gene. This observation also indirectly supports the cut-off values used to categorize the PCR amplicons as arising from different genes, as opposed to alleles of the same gene. Confirmatory evidence for these distinctions would require analysis of genetic crosses.

In comparisons with the ten *Ripks* amplicon classes, group I cDNA clones showed 99% nucleotide sequence identity (single nucleotide change) to *Ripks5*, and group II cDNA clones showed 99% nucleotide sequence identity (single nucleotide change) to *Ripks6*. Group III cDNA clones shared 92–96% nucleotide sequence identity with *Ripks1-10*, with the closest being *Ripks9* (96% nucleotide sequence identity). This level of variation most likely represents a new gene rather than an allelic sequence, and we therefore designated group III clones as representing a new class, *RiPKS11*. At the nucleotide level, the open reading frame of *RiPKS5* shares 93% nucleotide sequence identity to *RiPKS6*, and 97% nucleotide sequence identity to *RiPKS11*. The open reading frames of *RiPKS6* and *RiPKS11* share 92% nucleotide sequence identity. The lengths of the 3′-untranslated regions were 256, 308 and 223 nucleotides (respectively) for *RiPKS5, RiPKS6* and *RiPKS11*. The homology between the 3′-untranslated regions varied from 98% between *RiPKS5* and *RiPKS11* to 60% between *RiPKS5* and *RiPKS6*. Differences between the 5′ and 3′-untranslated regions consisted mostly of deletions and insertions rather than significant stretches of nucleotide dissimilarity. In the sequenced regions, the 5′-untranslated regions of *RiPKS5* and *RiPKS6* were identical. This is not an unprecedented observation, since the five alfalfa *CHS* genes (Junghans et al., 1993) also share considerable homology in these 5′-and the 3′-untranslated regions. Similarly, the two soybean *CHS* genes (*chs5* and *chs4*) have identical coding regions and 3′-untranslated regions, while sharing 63% nucleotide identity in the 5′-untranslated region (Akada and Dube, 1995). It is likely that characterization of sequences further upstream would reveal significant differences between the *RiPKS5* and *RiPKS6* genes.

The predicted amino acid sequences of the three cDNAs share 96–98% amino acid sequence identity amongst themselves. The *RiPKS* sequences also shared considerable sequence homology with other *Rubus idaeus PKS* clones identified from *R. idaeus* variety "Royalty" and other Rosaceae PKS sequences; 91–99% amino acid sequence identity with the three other *Rubus PKS* clones (Zheng et al., 2001); 94–95% amino acid sequence identity with strawberry CHS, and 86–90% amino acid sequence identity with apple CHS. The three *RiPKSs* also shared considerable similarity with *Arabidopsis CHS* sequences. The sequencing of the *Arabidopsis* genome has identified several divergent sequences that have been annotated as "*CHS*-like" based on homology. The *RiPKSs* were considerably divergent from these *Arabidopsis CHS*-like sequences and much more closely related to the *bona fide CHS* sequences (78–87% amino acid sequence identity).

A comparison of the deduced amino acid sequences of *RiPKSs* to strict CHS or STS consensus sequence revealed that the amino acids required in the active site of alfalfa CHS (Cys164, Phe215, His303 and Asn336) (Ferrer et al., 1999) are also conserved both in the three *RiPKS* polypeptides and in the consensus for STS and CHS. Residues Thr197, Ile254, Gly256, Asn336, Ser338, and His303, which define the active site volume for the alfalfa CHS2, were conserved in the *RiPKSs*, indicating that the active site architecture of the *RiPKSs* was

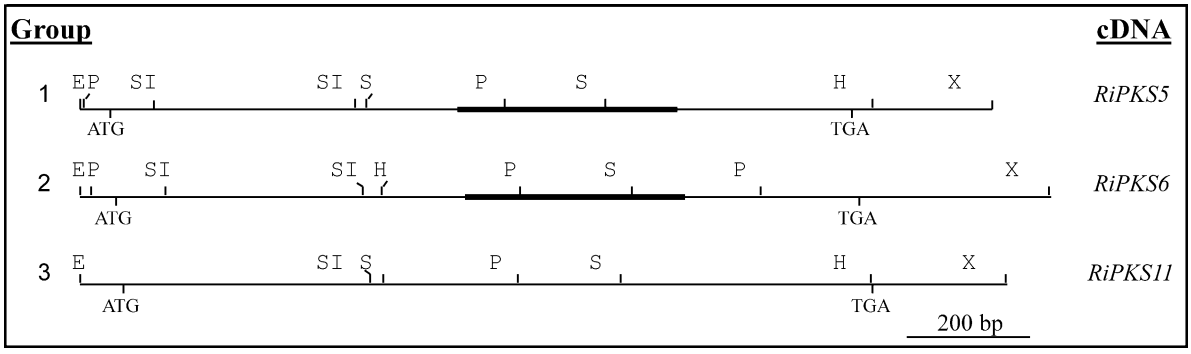

Fig. 3. Representative restriction maps of *Rubus PKS* cDNA clones. The dark line represents the regions of the cDNA also characterized by the PCR-based homology search. The positions of the start and the stop codons are denoted within each cDNA. Restriction enzymes used are: E, *Eco*RI; H, *Hind*III; P, *Pst*I; S, *Sma*I; SI, *Sal*I; X, *Xho*I.

similar to that of the alfalfa CHS2 and thus that *Rubus* enzymes should be capable of accommodating bulkier end products such as resveratrol or chalcones. Among the various *PKSs* that have been identified from plants, *RiPKSs* share higher percent sequence identity with *CHS*-type *PKSs* (82–91% amino acid sequence identity) than with other non-*CHS* type *PKS* (78–88% amino acid sequence identity).

While computer-based homology searches have proven to be fruitful exercises to find and define functional homologues of genes across species, one has to be particularly cautious in using such an approach for predicting the catalytic properties of new PKS. Non-CHS type PKSs are still very similar to CHS-type proteins at the sequence level, and it is clear that only subtle changes in the PKS active site are required to create a novel catalytic functions (Jez et al., 2000). This adaptability may well underlie the remarkable diversification that characterizes PKS-catalyzed metabolism in plants.

2.4. Functional expression of the three RiPKS cDNAs

To establish the functional properties of the *Rubus* PKSs, all three *RiPKS* cDNAs were expressed in *E. coli* as recombinant proteins either containing or lacking an N-terminal His_6-tag. Expression of the recombinant protein was not affected by the presence or the lack of the N-terminal His_6- tag (data not shown). The recombinant proteins had apparent molecular masses of approximately 42,000 Da and were recognized on western blots by polyclonal antibodies raised against parsley CHS (Kreuzaler et al., 1979) (Fig. 4A).

When assayed with *p*-coumaryl CoA and malonyl CoA, naringenin was the major product of the reaction with all three recombinant proteins (Fig. 4B). Negligible benzalacetone synthesis was observed, and control extracts from bacteria transformed with the vector lacking inserts displayed no CHS activity (Fig. 4B). Control assays using recombinant proteins incubated with a single substrate (malonyl-CoA) also did not yield naringenin. These results are consistent with the typical properties of CHS-type PKS. All three recombinant proteins also showed an apparent preference for *p*-coumaryl CoA as the starter unit, relative to other hydroxycinnamyl-CoA esters (data not shown).

2.5. Developmental regulation of the RiCHS cDNA transcripts

A range of *PKS* relatives, including *CHS*, *STS*, and *PS*, have consistently demonstrated differential patterns of expression with respect to specific tissues, and to developmental and environmental cues (Ryder et al., 1987; Bell et al., 1986; O'Neill et al., 1990; Harker et al., 1990; Gong et al., 1997; Koes et al., 1989; Helaruitta et al., 1995). Although there is much information about the expression patterns of individual genes, or the response of multiple gene-family members to environmental cues, little is known about the regulation of multiple family members during a developmental program such as flower and fruit development. Although we have not conducted an analysis with all members of the *RiPKS* family, we did examine the behavior of the three members of the *Rubus CHS* family that are most prominently associated with fruit ripening (Fig. 5).

Two patterns of gene expression were apparent. *RiPKS5* displayed high levels of mRNA expression in all organs of *Rubus*. The expression of the other two cDNAs, (*RiPKS6* and *RiPKS11*), varied significantly in ripening *Rubus* fruits, indicating that these genes are regulated in a ripening-dependent manner. The three *RiPKS* transcripts also differed in their temporal patterns of expression during flower development and fruit maturity. In fruits, *RiPKS11* expression was particularly high in developmental stages that also have higher anthocyanin levels (stages III–V). During flower development, *RiPKS11* was more actively transcribed in fully mature flowers (stage II) than in buds (stage I) or fertilized flowers (stage III). The kinetics of expression of *RiPKS5* and *RiPKS6* in developing flowers generally resembled each other and the accumulation of these transcripts was not obviously correlated with the process of fruit development. Most notably, in the comparison of five developmental stages of fruits that we sampled, *RiPKS5* and *RiPKS6* did not show a temporal relationship to anthocyanin accumulation, in contrast to *RiPKS11*.Our results demonstrate that both fruit-ripening-dependent and independent pathways of gene expression coexist within the *Rubus* CHS gene family, and thus provides a system in which it should be possible in the future to identify cis- and trans-acting factors that contribute to ripening-regulated gene expression.

It is not unusual for one or two family members to account for most of the *CHS* gene expression in a plant. In petunia, with about eight gene family members, one gene (*CHS-A*) accounts for ~90% of the CHS gene expression, while *CHS-J* accounts for ~10%, with *CHS-B* and *CHS-G* being expressed only at very low levels in floral tissues and in UV and light-induced seedlings (Koes et al., 1989). Similarly, in soybean only one of the three *CHS* genes is active in cotyledons (Wingender et al., 1989). In *Ipomea*, out of the thirteen genes characterized from seven species, *I. purpurea CHS-A* and *CHS-C*, and *I. platensis CHS-A* are the only genes actively expressed (Durbin et al., 1995). In contrast, two *CHS* genes characterized from tomatoes are expressed at equivalent levels in cotyledons, hypocotyls and leaves (O'Neill et al., 1990). As determined by the RT-cPCR analysis (Fig. 5B), the absolute levels of RiPKS transcripts in stage III fruits were *RiPKS5* > *RiPKS11* > *RiPKS6* (160: 40: 0.8 pmol cDNA/mg tRNA). Comparison of the relative abundance amongst

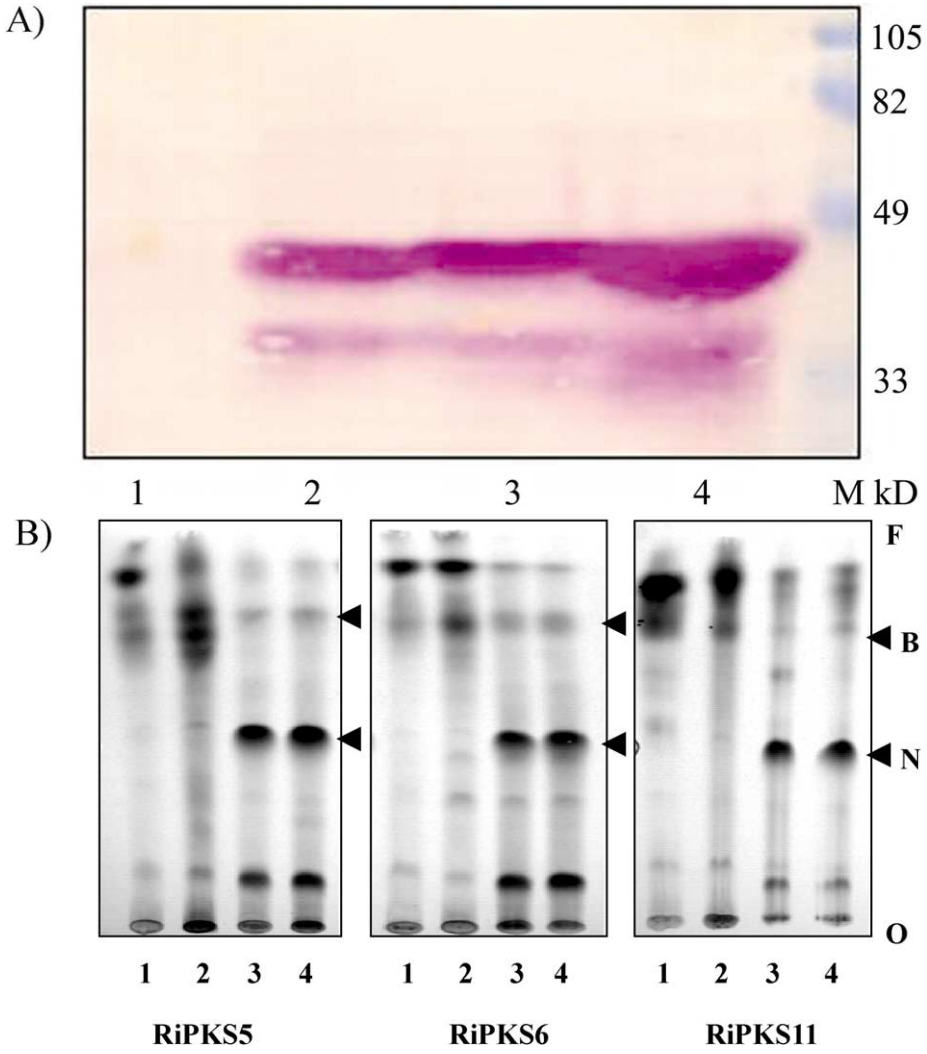

Fig. 4. Immunoblot analysis and enzymatic activity of the recombinant RiPKS protein. (A) Immunoblot analysis of RiPKS expressed proteins. Soluble proteins were isolated from IPTG-induced *E. coli* harboring plasmid pQE30 (lane 1), pQE30-PKS5 (lane 2), pQE30-PKS6 (lane 3), or pQE30-PKS11 (lane 4). M, Molecular mass standard in kD. The proteins were blotted onto a PVDF membrane and developed with antiserum specific to parsley CHS. (B) TLC analysis of products of the assay with the three *Rubus* PKS recombinant proteins. Ethyl acetate-soluble assay products were separated by TLC and radioactive products labeled from [2-^{14}C] malonyl-CoA visualized by autoradiography. The substrates used were *p*-coumaryl-CoA as the starter unit and malonyl-CoA as the extender moiety. Lane 1, vector-only *E.coli* control with both CoA esters; lane 2, RiPKS5, 6 or 11 recombinant proteins incubated with malonyl-CoA; lane 3 and lane 4, *Rubus* recombinant proteins incubated with both CoA esters. The relative mobilities of the radioactive products were compared against authentic samples of the expected end-products. O, origin of the chromatogram; F, solvent front of the chromatogram; N, position of authentic naringenin; B, position of authentic benzalacetone.

the three *RiPKS* transcripts over all the tissues revealed that *RiPKS5* mRNA is more abundant than either *RiPKS6* or *RiPKS11* (Table 1). In fact, *RiPKS6* was expressed only at very low levels in all the tissues investigated. During the screening of the cDNA library representing the stage III of fruits, we also observed that *RiPKS5* was the most abundant cDNA, which is consistent with the absolute levels of specific *RiPKS* transcripts determined by RT-cPCR.

It is important to recognize that, while we have conclusively demonstrated differential expression among three of the eleven *Rubus PKS* genes, it remains to be established whether the other *Rubus PKS* genes are functional, and if so, where and when they are expressed.

2.6. Phylogenetic origin of Rubus PKSs

Phylogenetic analysis of the *RiPKS* genes (*RiCHS5*, *RiCHS6* and *RiCHS11*) and other plant PKS sequences with defined catalytic activity yielded a tree in which the three *RiPKS* sequences clustered together, indicating that they have probably evolved by species-specific gene duplication event(s) from a single ancestor (data not shown). An earlier phylogenetic study concluded that non-CHS-like PKSs had evolved from *CHS* sequences several times independently during the course of plant evolution (Tropf et al., 1994). Later analyses from Durbin et al. (1995) and Helariutta et al. (1996) support the working hypothesis of Tropf et al. (1994), and

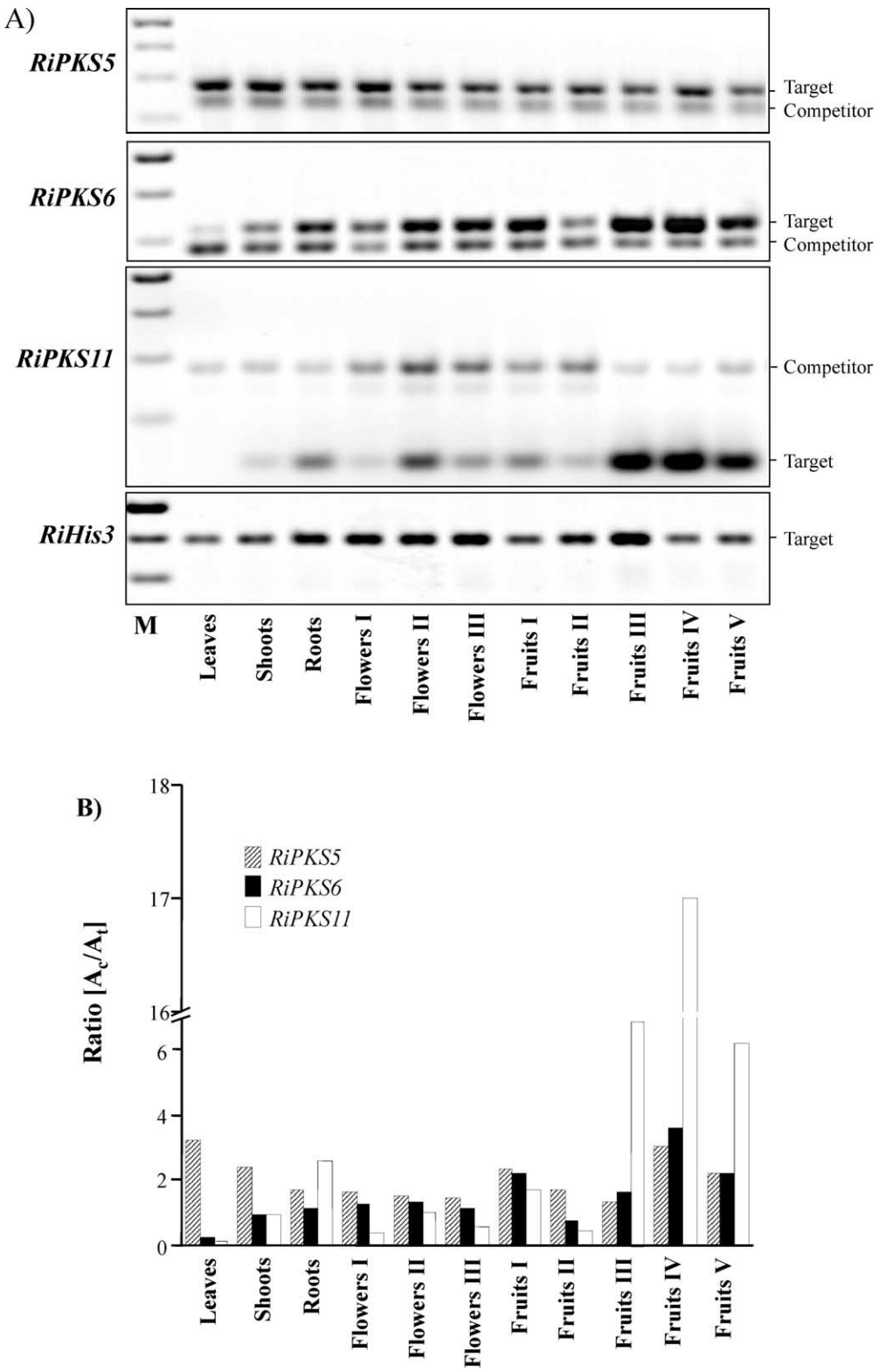

Fig. 5. Semi-quantitative RT-cPCR analysis of the accumulation of specific *RiPKS* transcripts in different organs of *Rubus*. (A) RT-cPCR was performed using 100 ng of total RNA isolated from young leaves, shoots, roots and different developmental stages of flowers and fruits. Following 32-repeated cycles of amplification, the products (50% of each amplification mix) were resolved in a 3% TAE-agarose/EtBr gel. (B) The relative amounts of target and competitor amplification product were calculated based on the intensity of the products, and the ratio of the two products has been graphed. Similar results were obtained in two independent experiments. The expression level of a given *RiPKS* genes can be compared between tissues, but expression amongst the three *RiPKS* genes cannot be compared within a tissue in this graph. The intensity of the bands was normalized to the average intensity of *RiHIS3* products as a control for starting RNA equivalence.

Table 1
Absolute levels of specific *RiPKS* transcripts in different organs

	RiPKS transcript levels (pmol/mg of tRNA)		
	RiPKS 5	*RiPKS 6*	*RiPKS1 1*
Leaves	390	0.1	0.4
Shoots	288	0.5	5
Roots	206	0.5	15
Flowers I	190	0.6	2
Flowers II	182	0.6	6
Flowers III	169	0.6	3
Fruits I	279	1	10
Fruits II	210	0.4	3
Fruits III	159	0.8	40
Fruits IV	361	2	101
Fruits V	266	1	36

The absolute levels of the three *RiPKS* transcripts were determined in the developmental stage III of fruits (Fig. 5A). The levels of transcripts of each cDNA in other organs were then determined based on the relative ratio of expression of each cDNA in those tissues compared to that seen in stage III fruits (Fig. 5B).

suggest that a similar mechanism underlies the evolution of catalytically divergent PS in *Gerbera* (Helariutta et al., 1996), and the family of *CHS*-like sequences in *Ipomea* (Durbin et al., 1995). Our phylogenetic analysis is consistent with this picture. Gene duplication followed by differentiation can result in the production of closely-related proteins with novel functions. The *Rubus CHS* genes seem have to undergone such differentiation, at least with their regulatory regions, since the expression patterns of the three genes do not correlate with their close phylogenetic clustering.

The identification and initial characterization of the eleven *Rubus PKS* genes presents us with an unprecedented opportunity to fully understand the fundamental function(s) and diversity of the *PKS* gene family members within a single species. We have also successful developed a set of primers that can be used to characterize the *PKS* gene-family from other plant species. Characterization of the catalytic properties of all the RiPKS isoforms would generate important insights into their potential roles, while creation of transgenic plants that are blocked in specific PKS-catalyzed reactions would permit an unambiguous association of specific genes with the functions of their encoded proteins.

3. Experimental

3.1. PCR primers and amplification reactions

The following plant-specific PKS sequences were aligned to generate an amino acid consensus: chalcone synthases [*Antirrhinum* (X03710), *Arabidopsis* (M20308), *Gerbera* (Z308096, Z380980), *Petroselinum* (P16107), *Pueraria* (D63855), *Pinus* (X60754), *Sinapis* (X16437), *Zea* (X60204, X60205)]; stilbene synthases [*Arachis* (X62300, L00952), *Pinus* (Z46915, Z46914, S50350), *Vitis* (X76892, P28343)]; bibenzyl synthases [*Phalaenopsis* (X79903, X79904)] and acridone synthase [*Ruta* (Z34088)] and a *Gerbera* PKS (Z38097) whose function was unknown at the time this study was started.

PK1 [5′-CGGAATTCAA(A/G)GCIAT(TCA)AA(AG) GA(GA)TGGG] and PK3 [5′-cgGAATTCAA(G/A) GA(T/C)CT IGCIGA(A/G)AA(T/C)AA] are the sense primers based on amino acids KAIKEWG and KD(L/V) AEVNN, respectively. PK2 [5′-ctctagaIGG(A/G)TG IGC(A/G/T)ATCCA(A/G)AA] and PK4 [5′-GC TCTAGA CCIGGICC(A/G)AAICC(A/G)AA] are the fully- degenerate antisense PCR primers based on amino acids FWIAHP and FGFGPG, respectively. The target locations of the primers within a typical plant *PKS* gene are illustrated in Fig. 2. To facilitate subcloning of the PCR-amplified products, *Eco*RI recognition sites (underlined) were incorporated in the sense primers PK1 and PK3, and *Xba*I recognition sites (underlined) were incorporated in the antisense primers PK2 and PK4.

Rubus idaeus genomic DNA was isolated from young leaves of the variety 'Meeker' using the method described by Doyle and Doyle (1990). Amplification reactions contained 100 ng genomic DNA, 1× Appligene buffer (providing a final concentration of 1.5 mM $MgCl_2$), 0.5 μM each dNTP, 2 μM each primer and 2.5 U *Taq* DNA polymerase in a final volume of 50 μl. The reaction mixture was incubated at 95 °C for 10 min and then subjected to 35 cycles of amplification (95 °C for 50 s, 55 °C for 50 s, 72 °C for 1 min), and completed by a final 10 min extension at 72 °C in a Techne PHC-3 thermal cycler (Mandel Scientific). PCR reactions were analyzed by 1% TAE-agarose gel electrophoresis.

To verify that the amplification products did not contain any chimeric artifacts due to recombination among related gene family members (Saiki et al., 1988), products of two independent PCR reactions with each primer combination were processed separately. To examine the size and the specificity of the amplicon inserts, multiple transformants from each primer combination were directly amplified in a 50 μl reaction volume containing 2 μM each vector-specific primers M13R and M13F, 50 μM each dNTP, 0.5 U *Taq* DNA polymerase (Appligene), and 1x PCR buffer mix (Appligene). Amplified products (4 μl) were digested with the restriction enzymes *Pst*I and *Sma*I (GIBCO-BRL) for 15 h at 37 °C to ensure complete digestion. The digested products were electrophoresed through a 1.5% TAE-agarose gel, and clones were grouped according to the resulting restriction fragment patterns.

3.2. Construction and screening of a Rubus cDNA library

Total RNA was isolated from *R. ideaus* cv. Meeker raspberry fruits (stage III) using the RNeasy Maxi Kit (Qiagen) following the manufacturer's protocol. Poly(A)$^+$ RNA was isolated from 1.5 mg total RNA using Dynabeads Oligo (dT_{25}) (DYNAL) following the manufacturer's instructions. A cDNA library was constructed from 5 μg poly(A)$^+$ RNA using Uni-ZAP XR Library Construction Kit (Stratagene). The cDNA library, consisting of $\sim 10^7$ independent clones, was amplified once.

Approximately 5×10^5 plaques of the amplified cDNA library were blotted in duplicate on Hybond N$^+$ nylon membrane (Amersham). The membranes were screened with a mixed population of *Ripks5* and *Ripks6*, radiolabeled to a high specific activity with [α-^{32}P]dATP using a Random Primer Labeling kit (Life Technologies). After three rounds of screening, the inserts of positive plaques were amplified with vector-specific primer T3 and *PKS* gene-specific primer PK4. Crude amplified products (5 μl) were treated for 15 h at 37 °C with 10 U of each enzyme (*Bam*HI, *Eco*RI, *Eco*RV, *Hin*dIII, *Kpn*I, *Pst*I, *Sma*I, *Sal*I, and *Xho*I) to ensure complete digestion. The digested products were electrophorezed through a 1.5% TAE-agarose gel and clones were grouped according to their restriction enzyme fingerprints. One representative plaque of each class, harboring the largest size insert, was rescued as a pBluescript II SK(−) phagemid using the ExAssist helper phage.

3.3. Sequencing and sequence analysis

Plasmid DNA from selected clones was isolated for sequencing following a mini-alkaline lysis/PEG precipitation procedure (Ausubel et al., 1995). Both strands of the inserts were sequenced with the M13 universal primer and/or synthetic oligonucleotide primers as needed to extend the sequence. Sequencing reactions were carried out at the Nucleic Acid Protein Service Unit (NAPS, UBC) in an Applied Biosystems ABI 373 DNA sequencer using Applied Biosystem AmpliTaq Dye-Deoxy Terminator Cycle Sequencing.

All sequences were edited and analyzed using the PC/GENE Software (Intelligenetics). Database searches for sequence homology and comparisons were performed using various web-based analytical tools compiled at the web-site http://www.sdsc.edu/ResTools/.

3.4. Recombinant protein expression and extraction

To generate the expression constructs, the open reading frame of each *RiPKS* cDNA was amplified in a PCR reaction with a gene-specific primer and the vector-specific universal primer T7. Gene-specific primers anchored at the 5′-end of the cDNA introduced a unique *Sph*I site upstream of the start codon. Vector-specific primer T7 anchored at the 3′-end of the cDNA introduce a unique *Xho*I site downstream of the stop codon. The nucleotide sequences of the gene-specific primers were 5′-ACATGCATGCATGGTGACCGTC GATGAA-3′ for *RiPKS5*, 5′-ACATGCATGCATGGT GACCGTCGATGAA-3′ for *RiPKS6*, and 5′-ACATG CATGCATGGTGACCGTCGATGAA-3′ for *RiPKS11*. The resulting PCR-amplified products (about 1.2 kb each) were ligated into *Sph*I-*Sal*I digested pQE30 or pQE50 to yield plasmids pQE30-RiPKS5, pQE30-RiPKS6, pQE30-RiPKS11, pQE50-RiPKS5, pQE50-RiPKS6, and pQE50-RiPKS11. pQE50-based expression plasmids express the recombinant proteins without the N-terminal His_6-tag, while recombinant proteins from the pQE30-based vector have an in-frame N-terminal His_6- tag.

To produce samples for enzymatic assays, plasmids containing the cDNA of interest were transformed into *E. coli* strain RM82, which contained an additional plasmid, pBUS25, carrying the *argU* (*dnaY*) gene (Brinkmann et al., 1989). Expression vectors without any insert, transformed into RM82 [pBUS250] cells, served as the negative control. Recombinant protein was obtained by inducing 50 ml cultures in the logarithmic-phase (grown at 37 °C) with 1.0 mM IPTG for 4 h. The culture was then centrifuged and the bacterial pellet resuspended in 5 ml of 50 mM HEPES buffer (pH 7.5). Cells were disrupted in a French press at 1100 psi, and the extract clarified by centrifugation at $10{,}000\times g$ for 2 min. The resulting supernatant containing the soluble protein was immediately used for further analysis.

3.5. Enzymatic synthesis of hydroxycinnamyl-CoenzymeA derivatives

Cinnamyl CoA and other hydroxycinnamyl CoA derivatives were enzymatically synthesized using poplar 4CL6 recombinant protein expressed in baculovirus-infected insect cells (Cukovic, 2001). Hydroxycinnamic acid derivatives (0.2 mM) were incubated with recombinant protein (200 μl) in a final volume of 1 ml. Co-factors were added at the same concentration as used for a typical 4CL assay as described by Knobloch and Hahlbrock (1975). The reaction was allowed to proceed for 2 h and was monitored at regular intervals at wavelengths appropriate for various CoA derivatives (Stöckigt and Zenk, 1975). At the end of the incubation, the mixture was loaded on a pre-equilibrated reverse-phase C_{18}-column (Waters, bed volume 1ml). Prior to applying the reaction mixture, the column was washed with three column volumes of methanol and equilibrated with three column volumes of 0.2 M MOPS, pH 7.5. After sample application, the sample was washed with 0.2 M MOPS, pH 7.5 until the effluent had no

absorbance. The aromatic CoA derivatives were then eluted with 30 ml water and immediately lyophilized. The freeze-dried powder was dissolved in 1 ml water and stored at −80 °C until further use. The authenticity and concentration of the synthesized product were determined by a UV-spectral scan of the product and absorbance measurements at the appropriate wavelength for each CoA derivative (Stöckigt and Zenk, 1975).

3.6. Enzymatic assay for plant PKS

PKS activity was determined by following the formation of radiolabeled end-product during incubation of the enzyme with radiolabeled [2-^{14}C] malonyl Coenzyme A (Amersham). Malonyl CoA (1 nmol) and *p*-coumaryl CoA (1 nmol) were incubated with 85 µl protein extract at 30 °C for 30 min in a final assay volume of 100 µl. The reaction was terminated by addition of glacial acetic acid (20 µl) and the assay mixture spiked with authentic samples of benzalacetone (20 nmol) and naringenin (80 nmol), the expected end-products. The reaction was extracted with 500 µl ethyl acetate, and the organic phase evaporated to dryness and re-dissolved in 50 µl ethanol.

The entire product was applied to the origin of a plastic-backed polyamide TLC plate (Macherey–Nagel) and developed using methanol:acetic acid:water (45:1:1) as the mobile phase. Products were visualized by autoradiography and samples having the same R_f as that of the authentic standard were scraped from the plate, eluted in methanol, and analyzed by liquid scintillation counting. Enzymatic activity is expressed as nkat/mg protein, where katal defines the formation of one mole of end-product in 1 s.

3.7. RT-cPCR

To distinguish between the transcripts of the full-length *RiPKS* genes, primers were designed to target unique regions within the 5′- and 3′-untranslated regions of the corresponding cDNAs. *RiPKS5* was amplified with forward primers 55 [5′-GAGACGGTTGTGCTT CACAGTGTG] and a gene-specific reverse primer 53 [5′-ACAATATG AAATGGAAACTGATA], to yield a 294 bp product. *RiPKS6* was amplified with sense primer 65 [5′-GAGACCGTTGTGCTTCACAGTGTGG] and a gene-specific antisense primer 63 [5′-GCAAAG CAATCAGAACTTTTTATC], yielding a product of 330 bp. *RiPKS11* was selectively amplified with gene-specific sense primer 115 [5′-GATCACTGCAACACC CCAAAC] and antisense primer 113 [5′-TACAG TTGGGAGGAGTTGCC] to amplify a product of 141 bp. Gene-specific primer pairs for amplification of the *Rubus Histone H3* (*RiHISH3*; GenBank,AF304365) consisted of H15 [5′-ATGGCGCGGACGAAGGA] and H13 [5′-GCCTACGCCGCCCGCTCAACCTA].

To generate PCR competitors, a 330 bp fragment of a spruce *coniferin-β-glucosidase* gene (M. Gray-Mitsumune and B.E. Ellis, unpublished data) was first amplified with primers C15 (5′-CCCCTAACAGGAA TTCTGCG-3′) and C13 (5′-ACCATCGCAGATTGAA GGAC-3′). The PCR product was then re-amplified with composite primers containing both *coniferin-β-glucosidase* and *RiPKS5*/*RiPKS6, RiPKS11*gene-specific sequences, and finally with *RiPKS5*/*RiPKS6*/*RiPKS11*-gene-specific primers only. This procedure generated two 370 bp non-homologous DNA fragments (competitors) containing at their ends appropriate templates for *RiPKS5*/*RiPKS6*/*RIPKS11* primers.

For amplification of the cDNA, first-strand cDNA reaction (1 µl) was amplified in a total volume of 20 µl containing a known amount of gene-specific competitor, 200 nM each PCR primer, 200 µM each dNTP, and 2.5 U of *Taq* DNA polymerase in 1×PCR buffer (Qiagen) and 1×Q solution (Qiagen). Thermal cycling was carried out at 94 °C for 5 min followed by 25 cycles (for *RiHISH3*) or 32 cycles (for *RiPKSs)* of 94 °C for 20 s, 59 °C for 50 s, and 72 °C for 50 s and a final extension of 5 min at 72 °C in a T-Gradient 96 thermal cycler (Biometra). Under the PCR conditions used, these primers only amplified their cognate plasmid cDNAs (data not shown). Amplified PCR products (10 µl) were separated in a 4% TAE-agarose gel and stained with 5 µg/ml ethidium bromide. The staining intensity was digitally quantified using the Scion Image (Scion Corporations). Relative intensities of the target and competitor bands were expressed as arbitrary units (A.U.) after correction for band size differences between the competitor and target. This analysis was repeated twice with consistent results and representative data from one analysis have been presented.

To determine the absolute levels of the two transcripts across different tissues, the absolute levels were first determined in fruits (stage III); since it had been established by cDNA screening that both genes are expressed in this tissue. Constant aliquots of cDNA (2 µl) were used as PCR template with *RiPKS5*/*RIPKS6*/*RiPKS11*gene-specific primer pairs in the presence of a series of dilutions of each competitor ranging from 100 to 3×10^{-3} attomoles. The targets from *RiPPKS5*/*RiPKS6*/*RiPKS11* mRNA and the control competitor (370 bp) fragments are simultaneously amplified by the polymerase in proportions that reflect the relative abundance of each target species. As decreasing amounts of competitor are added, the amount of *RiPKSs* cDNA being amplified increases. Hence, the ratio of the two bands (target:competitor) observed after agarose gel electrophoresis changes across the dilution series. After densitometric scanning of the stained gel, the logarithm of the molar ratio of the target to competitor was plotted as a function of the logarithm of the amount of competitor added, to establish

the point at which the amount of target is equal to the amount of competitor.

Once the absolute amounts of *RiPKS5*/*RIPKS6*/*RIPKS11*genes in developing fruits (stage III) were known, it was possible to calculate the amounts of each transcript in other tissues by comparing the ratio (target:competitor) obtained for each tissue to the ratio (target:competitor) established for expression in fruits stage III (Fig. 4B). This analysis was repeated twice with similar results, and representative values from one experiment are shown.

3.8. Phylogenetic analysis

Various PKS protein sequences were downloaded from the GenBank database and aligned using Clustal W (Altschul et al., 1990). The alignment was manually optimized and used for finding the most parsimonious tree, employing the heuristic search option algorithm in the program PAUP 4.02b (Smithsonian Institution). For statistical analysis, 1000 bootstrap replications (Felsenstein, 1995) were analyzed.

Acknowledgements

This research was supported by the Natural Science and Engineering Research Council of Canada. A.K. was the recipient of a University Fellowship from the University of British Columbia. The authors are greatly indebted to Hugh Daubney and Chaim Kempler (Agriculture and Agri-Food Canada, Summerland, BC) for their generous supply of plant materials, and to Dr. Carl Douglas for providing the anti-CHS antibody.

References

Abe, I., Takahashi, Y., Morita, H., Noguchi, H., 2001. Benzalacetone synthase: a novel polyketide synthase that plays a crucial role in biosynthesis of phenylbutanones in *Rheum palmatum*. European Journal of Biochemistry 268, 3354–3359.

An, C., Ichinose, Y., Yamada, T., Tanaka, Y., Shiraishi, T., Oku, H., 1993. Organization of the genes encoding chalcone synthase in *Pisum sativum*. Plant Molecular Biology 21, 789–803.

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., Struhl, K. (Eds.), 1995. Current Protocols in Molecular Biology. John Wiley, New York.

Akada, S., Kung, S.D., Dube, S.K., 1991. The nucleotide sequence of gene 1 of the soybean chalcone synthase multigene family. Plant Molecular Biology 16, 751–752.

Akada, S., Dube, S.K., 1995. Organization of soybean chalcone synthase clusters and characterization of a new member of the family. Plant Molecular Biology 29, 189–199.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. Journal of Molecular Biology 215, 403–410.

Batschauer, A., Ehlmann, B., Schäfer, E., 1991. Cloning and characterization of chalcone synthase gene from mustard and its light-dependent expression. Plant Molecular Biology 16, 175–185.

Beckert, C., Horn, C., Schnitzler, J.P., Lehning, A., Heller, W., Veit, M., 1997. Styrylpyrone biosynthesis in *Equisetum arvense* L. Phytochemistry 44, 275–283.

Bell, J.N., Ryder, T.B., Wingate, V.P.M., Bailey, J.A., Lamb, C.J., 1986. Differential accumulation of plant defense gene transcripts in a compatible and an incompatible plant-pathogen interaction. Molecular and Cellular Biology 6, 1615–1623.

Brinkmann, U., Mattes, R.E., Buckel, P., 1989. High-level expression of recombinant genes in *Escherichia coli* dependent on the availability of the *dnaY* gene product. Gene 85, 109–114.

Borejsza-Wysocki, W., Hrazdina, G., 1996. Aromatic polyketide synthases: purification, characterization and antibody development to beanzalacetone synthase from raspberry fruits. Plant Physiology 110, 791–799.

Butland, S.L., Chow, M.L., Ellis, B.E., 1998. A diverse family of phenylalanine ammonia-lyase genes in pine tree and cell cultures. Plant Molecular Biology 37, 15–24.

Cukovic, D., Ehlting, J., VanZiffle, J.A., Douglas, C.J., 2001. Structure and evolution of 4-coumarate:coenzyme A ligase (*4CL*) gene families. Biological Chemistry 382, 645–654.

Dixon, R.A., 1999. Plant natural products: the molecular genetic basis of biosynthetic diversity. Current Opinion in Biotechnology 10, 192–197.

Dixon, R.A., Steele, C.L., 1999. Flavonoids and isoflavonoids—a gold mine for metabolic engineering. Trends in Plant Science 4, 1360–1385.

Durbin, M.L., Learn Jr., G.H., Huttley, G.A., Clegg, M.T., 1995. Evolution of the chalcone synthase gene family in the genus *Ipomoea*. Proceedings of the National Academy of Science of the United States of America 92, 3338–3342.

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

Eckermann, S., Schroder, G., Schmidt, J., Strack, D., Edrada, R.A., Helariutta, Y., Elomaa, P., Kotilainen, M., Kilpelainen, I., Proksch, P., Teeri, T.H., Schröder, J., 1998. New pathway to polyketides in plants. Nature 396, 387–390.

Epping, B., Kittel, M., Ruhnau, B., Hemleben, V., 1990. Isolation and sequence analysis of a chalcone synthase cDNA of *Matthiola incana* RBr (Brassicaceae). Plant Molecular Biology 14, 1061–1063.

Estabroook, E., Gopalan-Sengupta, C., 1991. Differential expression of phenylalanine ammonia lyase and chalcone synthase during soybean nodule development. The Plant Cell 3, 299–308.

Flesenstein, J., 1995. Confidence limits on phylogenies: an approach using bootstrap. Evolution 39, 783–791.

Ferrer, J.-L., Jez, J.M., Bowman, M.E., Dixon, R.A., Noel, J.P., 1999. Structure of chalcone synthase and the molecular basis of plant polyketide biosynthesis. Nature Structure Biology 6, 775–784.

Feinbaum, R.L., Ausubel, F.M., 1988. Transcriptional regulation of *Arabidopsis thaliana* chalcone synthase gene. Molecular and Cellular Biology 8, 1985–1992.

Franken, P., Niesbach-Klösgen, U., Weydemann, U., Maréchal-Drouard, L., Saedler, H., Wienand, U., 1991. The duplicated chalcone synthase genes *C2* and *Whp* (white pollen) of *Zea mays* are independently regulated; evidence for translational control of *Whp* expression by the anthocyanin intensifying gene. EMBO Journal 10, 2065–2612.

Gong, Z., Yamazaki, M., Sigiyama, M., Tanaka, Y., Saito, K., 1997. Cloning and molecular analysis of structural genes involved in anthocyanin biosynthesis and expressed in a forma-specific manner in *Perilla frutescens*. Plant Molecular Biology 35, 915–917.

Goodwin, P.H., Hsiang, T., Erickson, L., 2000. A comparison of stilbene and chalcone synthases including a new stilbene synthase gene from *Vitis riparia* cv. Gloire de Montpellier. Plant Science 151, 1–8.

Harker, C., Ellis, N.T.H., Coen, E.S., 1990. Identification and genetic regulation of the chalcone synthase multigene family in pea. The Plant Cell 2, 185–194.

Helariutta, Y., Elomaa, P., Kotilainen, M., Griesbach, R.J., Schröder, J., Teeri, T.H., 1995. Chalcone synthase-like genes are active during corolla development are differentially expressed and encode enzymes with differential catalytic properties in *Gerbera hybrida* (Asteraceae). Plant Molecular Biology 28, 47–60.

Helariutta, Y., Kotilainen, M., Elomaa, P., Kalkkinen, N., Bremer, K., Teeri, T.H., Albert, V.A., 1996. Duplication and functional divergence in the chalcone synthase gene family of Asteraceae: evolution with substrate change and catalytic simplification. Proceedings of the National Academy of Sciences of the United States of America 93, 9033–9038.

Hopwood, D.A., Sherman, D.H., 1990. Molecular genetics of polyketides and its comparisons to fatty acid biosynthesis. Annual Review of Genetics 24, 37–66.

Jez, J.M., Ferrer, J.L., Bowman, M.E., Dixon, R.A., Noel, J.P., 2000. Dissection of malonyl-CoA decarboxylation from polyketide formation in the reaction mechanism of a plant polyketide synthase. Biochemistry 39, 890–902.

Jez, J.M., Ferrer, J.L., Bowman, M.E., Austin, M.B., Schröder, J., Dixon, R.A., Noel, J.P., 2001. Structure and mechanism of chalcone synthase-like polyketide synthases. Journal of Industrial Microbiology and Biotechnology 27, 393–398.

Junghans, H., Dalkin, K., Dixon, R.A., 1993. Stress response in alfalfa (*Medicago sativa* L.) 15. Characterization and expression patterns of members of a subset of the chalcone synthase multigene family. Plant Molecular Biology 22, 239–253.

Knobloch, K.-H., Hahlbrock, K., 1975. Isoenzymes of *p*-coumarate:CoA ligase from cell suspension cultures of *Glycine max*. European Journal of Biochemistry 52, 311–320.

Koes, R.E., Spelt, C.E., Mol, J.N.M., Gerats, A.G.M., 1987. The chalcone synthase multigene family of *Petunia hybrida* (V30): sequence homology, chromosomal localization and evolutionary aspects. Plant Molecular Biology 10, 375–385.

Koes, R.E., Spelt, C.E., Mol, J.N.M., 1989. The chalcone synthase multigene family of *Petunia hybrida* (V30): differential light-regulated expression during flower development and UV light induction. Plant Molecular Biology 12, 213–225.

Kreuzaler, F., Ragg, H., Heller, W., Tesch, R., Witt, I., Hammer, D., Hahlbrock, K., 1979. Flavanone synthase from *Petroselinum hortense*. Molecular weight, subunit composition, size of messenger RNA, and absence of pantheinyl residue. European Journal of Biochemistry 99, 89–96.

O'Neill, S.D., Tong, Y., Spörlein, B., Forkmann, G., Yoder, J.I., 1990. Molecular genetic analysis of chalcone synthase in *Lycopersicon esculentum* and an anthocyanin-deficient mutant. Molecular and General Genetics 224, 279–288.

Preisig-Müller, R., Gnau, P., Kindl, H., 1995. The inducible 9,10-dihydrophenanthrene pathway: characterization and expression of bibenzyl synthase and *S*-adenosylhomocysteine hydrolase. Archives of Biochemistry and Biophysics 317, 201–207.

Raiber, S., Schröder, G., Schröder, J., 1995. Molecular characterization of two stilbene synthases from eastern white pine (*Pinus strobus*): a single Arg/His difference determines the activity and the pH dependence of the enzymes. FEBS Letters 361, 299–302.

Reimold, U., Kroeger, M., Kreuzaler, F., Hahlbrock, K., 1983. Coding and 3′ non-coding nucleotide sequence of chalcone synthase mRNA and assignment of amino acid sequence of the enzyme. EMBO Journal 2, 1801–1805.

Rhode, W., Dorr, S., Salamini, F., Becker, D., 1991. Structure of a chalcone synthase gene from *Hordeum vulgare*. Plant Molecular Biology 16, 1103–1106.

Rolfs, C.H., Fritzemeier, K.H., Kindl, H., 1981. Cultures cells of *Arachis hypogaea* susceptible for induction of stilbene synthase (resveratrol-forming). Plant Cell Reports 1, 83–85.

Rolfs, C.H., Kindl, H., 1984. Stilbene synthase and chalcone synthase. Plant Physiology 75, 489–492.

Ryder, T.B., Hedrick, S.A., Bell, J.N., Liang, X., Clouse, S.D., Lamb, C.J., 1987. Organization and differential activation of a gene family encoding the plant defense enzyme chalcone synthase in *Phaseolus vulagris*. Molecular and General Genetics 210, 219–233.

Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B., Erlich, H.A., 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239, 487–491.

Schanz, S., Schröder, G., Schröder, J., 1992. Stilbene synthase from Scots pine (*Pinus sylvestris*). FEBS Letters 313, 71–74.

Schinz, H., Seidel, C.F., 1961. Nachtrag zu der Arbeit Nr. 194. Helvetica Chimica Acta 44, 278.

Schoeppner, A., Kindl, H., 1979. Stilbene synthase (pinosylvine synthase) and its induction by ultraviolet light. FEBS Letters 108, 349–352.

Schröder, G., Brown, W.S., Schröder, J., 1988. Molecular analysis of resveratrol synthase. CDNA, genomic clones and relationship with chalcone synthase. European Journal of Biochemistry 172, 161–169.

Schröder, J., 1997. A family of plant-specific polyketide synthases: facts and predictions. Trends in Plant Science 2, 373–378.

Schröder, J., Raiber, S., Berger, T., Schmidt, A., Schmidt, J., Soares-Sello, A.M., Bradshiri, E., Strack, D., Simpson, T.J., Veit, M., Schröder, S., 1998. Plant polyketide synthases: a chalcone synthase-type enzyme which performs a condensation reaction with the methylmalonyl-CoA in the biosynthesis of *C*-methylated chalcones. Biochemistry 37, 8417–8425.

Shirley-Winkel, B., Kubasek, W., Storz, G., Bruggemann, W., Koorneef, M., Ausubel, F., Goodman, H.M., 1995. Analysis of *Arabidopsis* mutants deficient in flavonoid biosynthesis. The Plant Jounral 8, 659–671.

Sommer, H., Saedler, H., 1986. Structure of the chalcone synthase gene of *Antirrhinum majus*. Molecular and General Genetics 202, 429–434.

Stöckigt, J., Zenk, M.H., 1975. Chemical synthesis and properties of hydroxycinnamoyl-coenzyme A derivatives. Z. Naturforsch. 30c, 352–358.

Tropf, S., Lanz, T., Rensing, S.A., Schröder, J., Schröder, G., 1994. Evidence that stilbene synthases have developed from chalcone synthases several times in the course of evolution. Journal of Molecular Evolution 38, 610–618.

Wiese, W., Vorman, E., Krause, H., Kindl, H., 1994. Structural organization and differential expression of three stilbene synthase genes located on a 13-kb grapevine fragment. Plant Molecular Biology 26, 667–677.

Weinand, U., Sommer, H., Schwarz, Z.S., Shepherd, N., Saedler, H., Kreuzler, F., Ragg, H., Fautz, E., Hahlbrock, K., Harrison, B., Peterson, P.A., 1982. A general method to identify plant structural genes among genomic clones using transposable elements induced mutations. Molecular and General Genetics 187, 195–201.

Wingender, R., Rohrig, H., Horicke, C., Wing, D., Schell, J., 1989. Differential regulation of soybean chalcone synthase genes in plant defense, symbiosis and upon environmental stimuli. Molecular and General Genetics 315–322.

Zheng, D., Schröder, G., Schröder, J., Hrazdina, G., 2001. Molecular and biochemical characterization of three aromatic polyketide synthase genes from *Rubus idaeus*. Plant Molecular Biology 46, 1–15.