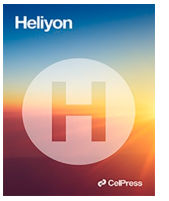
Heliyon 10 (2024) e27698

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Research article

Cloning of three *Alnus sieboldiana* type III polyketide synthases and formation of polyketides in recombinant *Escherichia coli* using cinnamic acid analogs as substrates

Konosuke Takemoto [a,b], Yuichi Mikota [a], Ryota Moriuchi [c], Yuko Yoneda [a], Shingo Kawai [a,*]

[a] *Faculty of Agriculture, Shizuoka University, 836 Ohya, Suruga-ku, Shizuoka, 422-8529, Japan*
[b] *The United Graduate School of Agricultural Science, Gifu University, 1-1, Yanagido, Gifu-shi, 501-1193, Japan*
[c] *Functional Genomics Section, Shizuoka Instrumental Analysis Center, Shizuoka University, 836 Ohya, Suruga-ku, Shizuoka, 422-8529, Japan*

ARTICLE INFO

Keywords:
Betulaceae
Flavonoids
Diarylheptanoids
Type III polyketide synthase
Diketide-CoA synthase
Actinorhizal symbiosis

ABSTRACT

Alnus sieboldiana is an actinorhizal plant that coexists with the nitrogen-fixing actinomycete *Frankia* via nodules. It produces a variety of polyketides, including flavonoids, stilbenoids, and diarylheptanoids. These compounds have beneficial biological activities. Plant polyketides are produced by type III polyketide synthases (PKSIII). In this study, three *A. sieboldiana* PKSIIIs (AsPKSIII1, AsPKSIII2, and AsPKSIII3) predicted from next-generation sequencing analysis of *A. sieboldiana* seedling RNA were amplified and cloned. Phylogenetic tree analysis classified AsPKSIII2 and AsPKSIII3 into the chalcone synthase (CHS) group, whereas AsPKSIII1 was not classified into this group. We attempted to produce polyketides by adding cinnamic acid analogs to the culture medium of *Escherichia coli,* in which the respective PKSIII gene and the acetyl-CoA carboxylase (ACC) and 4-coumarate: CoA ligase (4CL) genes were simultaneously recombined. AsPKSIII1 is an enzyme that condensed only one molecule of malonyl-CoA to cinnamoyl-CoAs. In contrast, AsPKSIII2 and AsPKSIII3 produced chalcones as shown in a phylogenetic tree analysis, but also produced triketide pyrone. The ratio of these products differed between the two enzymes. We determined the gene and amino acid sequences as well as the substrate specificities of the two enzymes involved in flavonoid production and one enzyme potentially involved in diarylheptanoid production in *A. sieboldiana*.

1. Introduction

Alnus sieboldiana is a warm-temperate tree found in the highlands of the Pacific coast of central Japan [1]. It is a pioneer tree that grows in devastated areas after volcanic eruptions and is used for land stabilization, the greening of slopes, and in rough areas [2,3]. The tree is also an actinorhizal plant that symbiotically coexists with the atmospheric nitrogen-fixing actinomycete *Frankia* via rhizobium. These plants absorb ammonia nitrogen from *Frankia* in return for supplying carbon sources to *Frankia* [4]. Based on this property, actinorhizal plants grow well on nitrogen-starved soils; thus, fulfilling a pivotal role in the nitrogen cycle of terrestrial ecosystems. Recently, we demonstrated for the first time that the tree-side signaling compounds involved in symbiosis are cyclic

* Corresponding author.
E-mail address: kawai.shingo@shizuoka.ac.jp (S. Kawai).

https://doi.org/10.1016/j.heliyon.2024.e27698
Received 13 November 2023; Received in revised form 13 February 2024; Accepted 5 March 2024
Available online 9 March 2024

diarylheptanoids [5] and examined the biosynthetic mechanisms of these compounds. We also recently reported an unsaturated, heptane chain-double bond reductase involved in the formation of *A. sieboldiana*-specific diarylheptanoids, which are mostly saturated-type molecules [6].

A. sieboldiana produces a variety of polyketides, such as flavonoids, stilbenoids, and diarylheptanoids [7–10]. *A. sieboldiana* extracts contain polyketide compounds that exhibit beneficial biological activities. Cyclic diarylheptanoids, such as alnusonol, which are tree-side signaling compounds for symbiosis with *Frankia* [5], were isolated from branch wood and exhibit α-glucosidase inhibitory activity [7]. Leaf extracts inhibited cell proliferation against A549 human lung cancer cells and contain flavonol galangin, which inhibits tumor necrosis factor alpha (TNF-α) expression [8]. Therefore, *A. sieboldiana* may not only be used to conserve the forest and mountain environment, but also as a source of woody biomass and useful biological compounds.

These plant polyketides are produced by enzymes belonging to the type III polyketide synthase (PKSIII) superfamily [11]. Chalcone synthase (CHS) is the most recognized PKSIII member, which synthesizes the flavonoid precursor chalcone. CHS condenses the starter substrate *p*-coumaroyl-CoA with three malonyl-CoA molecules to form a tetraketide, followed by a Claisen-type cyclization to form the naringenin chalcone [12]. The biosynthesis of linear-type diarylheptanoids has been reported for curcumin from turmeric rhizome, which occurs through a two-step reaction: condensation of feruloyl-CoA with malonyl-CoA catalyzed by diketide-CoA synthase (DCS) and condensation of diketide-CoA with a second molecule of feruloyl-CoA catalyzed by curcumin synthase (CURS) [13]. Currently, more than 20 PKSIIIs with different functions are recognized, depending on the type of starter and extender substrates as well as on the number of condensations. Because of their various functions, PKSIII also has an important role in the biosynthesis of plant-specific secondary metabolites.

In *A. sieboldiana*, several compounds have been identified that are predicted to contain C_6–C_3 units derived from cinnamoyl-CoA analogs, including *p*-coumaroyl-CoA. In the present study, three PKSIII cDNAs were cloned from *A. sieboldiana* and each gene was simultaneously recombined with the acetyl-CoA carboxylase (ACC) gene and the 4-coumarate: CoA ligase (4CL) gene into *E. coli*. The resulting polyketides were produced by adding cinnamic acid analogs to the recombinant *E. coli* culture medium. Characterization of the three *A. sieboldiana* PKSIII species was done by HPLC and LCMS analysis of the *in vivo* products and phylogenetic tree analysis of the plant PKSIIIs.

2. Results and discussion

2.1. Sequence analysis of AsPKSIII1, AsPKSIII2, and AsPKSIII3

The RNA-seq library was sequenced using an Illumina NovaSeq 6000. Filtering resulted in a paired-end reads of approximately 516 million, including a total of 77.2 gigabases. It also obtained 486,828 transcripts, which were clustered into 232,492 unigenes. The

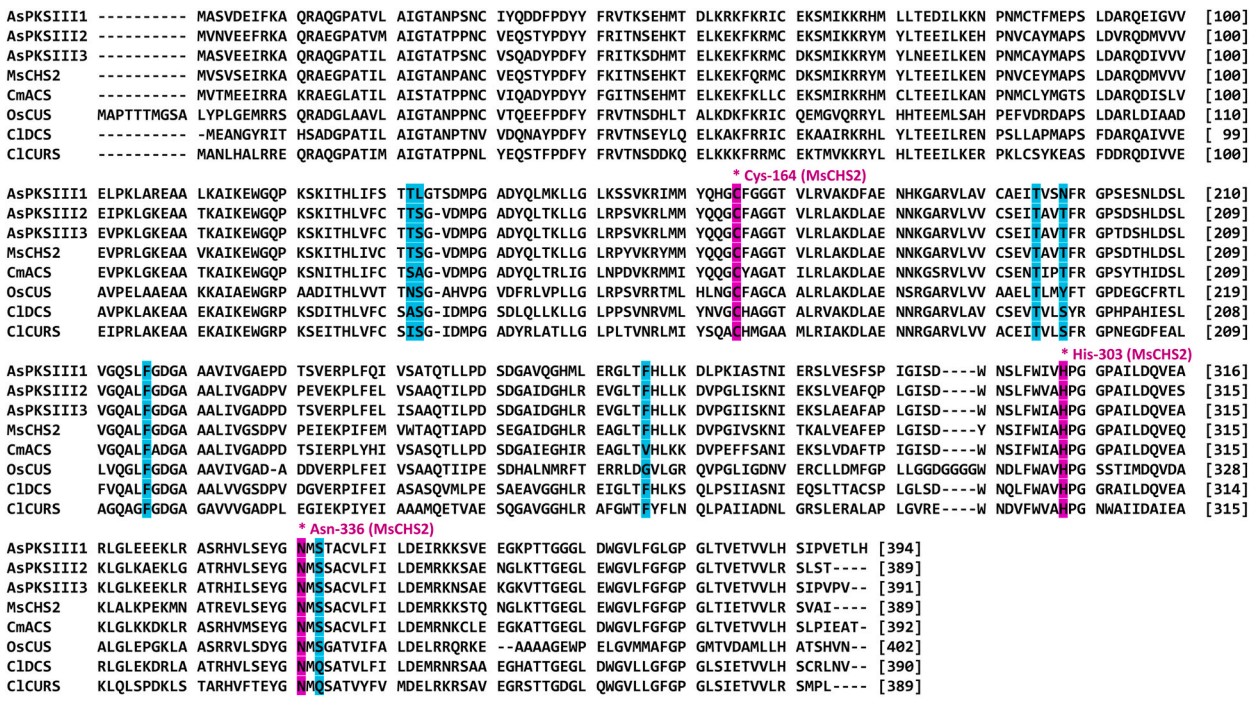


Fig. 1. Comparison of amino acid sequences of AsPKSIII1, AsPKSIII2, and AsPKSIII3 with other plant type III PKSs. The catalytic triad (Cys, His, Asn) and the residues lining the active site cavity in CHS are indicated as magenta and light blue, respectively. The GenBank accession numbers are: MsCHS2, L02902; RpBAS, AF326911; CmACS, AB823699; OsCUS, AK109558; ClDCS, AB535216; ClCURS1, AB495007.

transcripts had mean length of 838.19 bp, medium length of 490 bp, N50 length of 1347 bp, and maximum length of 19,163 bp. Three cDNAs encoding type III PKSs, *AsPKSIII1*, *AsPKSIII2,* and *AsPKSIII3*, were amplified from *A. sieboldiana* seedlings by RT-PCR and cloned. *AsPKSIII1*, *AsPKSIII2,* and *AsPKSIII3* contained protein coding regions (coding sequence, CDS) of 1185 bp, 1170 bp, and 1176 bp encoding 43,222 Da, 42,517 Da, and 42,593 Da proteins of 394, 389, and 391 amino acids, respectively. The CDS of *AsPKSIII1*, *AsPKSIII2,* and *AsPKSIII3* were deposited into the DNA Data Bank of Japan (DDBJ) (accession numbers: LC779619, LC779620, and LC779621, respectively).

AsPKSIII1, AsPKSIII2, and AsPKSIII3 share 87–51% amino acid sequence identity with other functionally characterized plant PKSIIIs. For AsPKSIII1, the amino acid sequence identity is 69% and 65%, respectively, with *Medicago sativa* CHS2 (MsCHS2), which produces naringenin chalcone [12], and *Arachis hypogaea* STS (AhSTS), which produces resveratrol [14], using *p*-coumaroyl-CoA as a substrate and three molecules of malonyl-CoA as extender substrate. Moreover, AsPKSIII1 exhibits 64% and 56% identity to ClDCS and ClCURS1, respectively, which are enzymes involved in the biosynthesis of curcumin [13], and 51% identity to *Oryza sativa* curcuminoid synthase CUS, which produces bisdemethoxycurcumin by itself [15]. It shares 71% identity with *Citrus microcarpa* acridone synthase (CmACS), which produces 1,3-dihydroxy-*N*-methylacridone from the decarboxylative condensation of *N*-methylanthraniloyl-CoA with three molecules malonyl-CoA [16]. The identity of AsPKSIII1 to AsPKSIII2 and AsPKSIII3 is 74% and 78%, respectively.

AsPKSIII2 and AsPKSIII3, respectively, share 87–83% amino acid sequence identity with plant CHSs, 75% and 71% identity with AhSTS, 66% and 68% identity with ClDCS, 61% and 60% identity with ClCURS, 53% and 54% identity with OsCUS, 77% and 78% identity with CmACS. There is 89% identity between AsPKSIII2 and AsPKSIII3.

AsPKSIII2 and AsPKSIII3 have conserved Cys-164, His-303, and Asn-336 residues, which are known catalytic triad of type III PKS. It also conserves Thr-132, Ser-133, Thr-194, Thr-197, Ser-338, and Pro-375, lining the active site cavity of CHS, and the "gatekeepers" Phe-215 and Phe-265 (MsCHS2 numbering) (Fig. 1) [12]. Differences in these residues are thought to regulate the substrate and product specificity of the type III PKS enzyme [11]. AsPKSIII1 has a catalytic triad, Thr-132, Thr-194, Ser-338, Pro-375 residues, and two gatekeeper Phe residues. In contrast, Ser-133 and Thr-197 were changed to Leu and Asn. Furthermore, a phylogenetic tree analysis classified AsPKSIII2 and AsPKSIII3 into the CHS/STS group. In contrast, AsPKSIII1 was classified in the non-CHS group, which produces polyketides that are not chalcone, and placed in a cluster containing ACS and quinolone synthase (QNS) (Fig. 2).

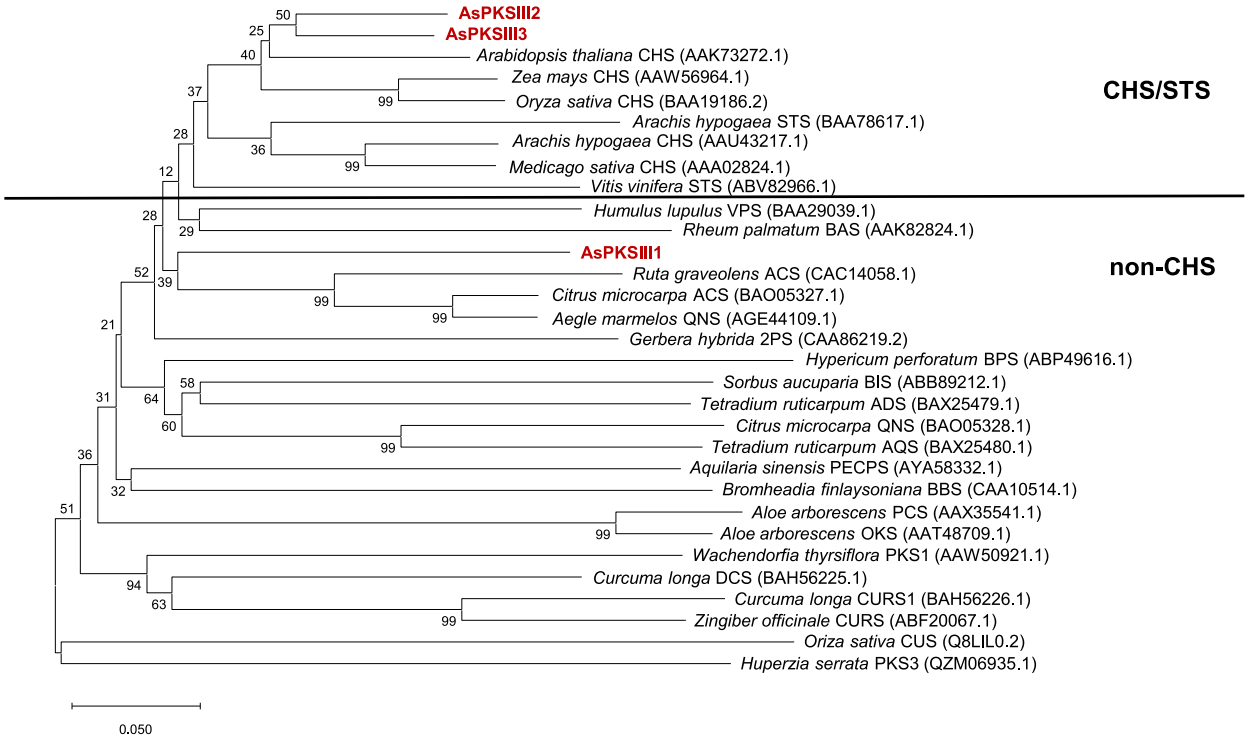


Fig. 2. Phylogenetic tree analysis of plant type III PKSs containing AsPKSIII1, AsPKSIII2, and AsPKSIII3. AsPKSIII1, AsPKSIII2, and AsPKSIII3 are shown in red and bold. The scale represents 0.05 amino acid substitutions per site. CHS, Chalcone synthase; STS, Stilbene synthase; VPS, valerophenone synthase; BAS, benzalacetone synthase; ACS, acridone synthase; QNS, quinolone synthase; 2 PS, 2-pyrone synthase; BPS, benzophenone synthase; BIS, biphenyl synthase; ADS, alkyldiketide-CoA synthase; AQS, alkylquinolone synthase; PECPS, 2-(2-phenylethyl)chromone precursor synthase; BBS; bibenzyl synthase; PCS, pentaketide chromone synthase; OKS, octaketide synthase; DCS, diketide-CoA synthase; CURS, curcumin synthase; CUS, curcuminoid synthase; GenBank registration numbers are shown in parentheses.

2.2. In vivo functional analysis of three PKSIIIs from A. sieboldiana and insights into their roles in plants

Cinnamic acid analogs were added to the culture medium of *E. coli* transformed with a total of three plasmids carrying each *AsPKSIII* gene, 4CL gene and *ACC* gene, respectively, to produce polyketides. 4CL is required for the cinnamoyl-CoA ester synthesis step, and the *Lithospermum erythrohizon* 4CL (*LE4CL-1*) gene [17] of was used. Two subunit genes of *ACC* from *Corynebacterium glutamicum* [18] were used to increase the intracellular pool of malonyl-CoA. HPLC and LCMS analysis revealed the presence of benzalacetone corresponding to each cinnamic acid in the ethyl acetate (EtOAc) extracts of the AsPKSIII1 recombinant *E. coli* culture medium supplemented with *p*-coumaric acid (**1**), ferulic acid (**2**), and caffeic acid (**3**) (Table 1, Figs. S3–S5). Benzalacetones are produced either by benzalacetone synthase (BAS) [19] or by acid treatment hydrolysis and subsequent decarboxylation of diketide-CoA esters produced by DCS [13]. In the combination of LE4CL-1 with AsPKSIII1, 0.1 mM **1** and **2** were completely converted to 4-hydroxybenzalacetone and dehydrozingerone after 40 min (Fig. 3). For **3**, the substrate was completely eliminated after 60 h, but the amount of 3,4-dihydroxybenzalacetone produced after 1 h was minimal (Fig. 3, S5). Although new peaks were observed when dihydro-*p*-coumaric acid (**5**) and cinnamic acid (**6**) were used as substrates, respectively, the products could not be determined (Figs. S6 and S7). During the cloning process, an isozyme with five amino acid mutations (L71H, K79E, N82K, N206H, and Q256E) was obtained simultaneously, but *in vivo* enzymatic reactions revealed no functional differences with AsPKSIII1. By switching the expression vector to pET23a(+) DNA, His-tagged AsPKSIII1 was expressed in transformed *E. coli* and purified by Ni-NTA affinity chromatography (Fig. S8). An *in vitro* condensation reaction of *p*-coumaroyl-CoA with malonyl-CoA by purified AsPKSIII1 produced benzalacetone as in the *in vivo* reaction. Alkaline treatment of the product with NaOH prior to acid treatment resulted in more efficient hydrolysis of *p*-coumaroyldiketide-CoA and increased detection of benzalacetones by HPLC analysis (data not shown). Because there are no reports of benzalacetones being isolated from *A. sieboldiana*, it is likely that AsPKSIII1 is a DCS-like enzyme responsible for the first step of diarylheptanoid biosynthesis in *A. sieboldiana* (Fig. 4). However, LCMS analysis did not capture *p*-coumaroyldiketide-CoA and *p*-coumaroyldiketide acid, so further studies are needed to characterize AsPKSIII1.

On the other hand, triketide pyrones and flavanones were detected in EtOAc extracts of AsPKSIII2 and AsPKSIII3 recombinant *E. coli* cultures supplemented with **1**, **3**, **5**, and **6** (Table 1, Fig. S3, S5–S7). Because acid treatment is known to cause nonstereospecific ring-C closure of chalcones, we assumed that the enzymatic product was a chalcone, but extracted as a flavanone. In fact, the product showing an *m/z* value of 271 detected in negative ion mode when **1** was used as a substrate, matched the retention time of naringenin in the HPLC chromatogram (Fig. S3). After 60 min, substrates were completely consumed in the *in vivo* reaction of AsPKSIII2 recombinant *E. coli* when **1**, **2**, **5**, and **6** were added, while a large amount of unreacted residual substrate was detected in AsPKSIII3 recombinant *E. coli* (Fig. 3). Whether this was the result of enzyme activity or differences in protein expression in *E. coli* cells cannot be determined. However, because little **6** were consumed even after 60 h, AsPKSIII3 appeared to have very low activity against **6** (Fig. S7). There are a number of flavonoids derived from cinnamic acid isolated from *A. sieboldiana* [10]. This strongly suggests that AsPKSIII2, rather than AsPKSIII3, is involved in the production of these flavonoids. Interestingly, neither AsPKSIII2 nor AsPKSIII3 converted **2** to

Table 1
Products detected by HPLC and LCMS analysis of the *A. sieboldiana* PKSIII recombinant *E. coli in vivo* reactions 60 h after the addition of cinnamic acid analogs.

starting substrates	enzyme name	main products	byproducts
p-coumaric acid (**1**)	AsPKSIII1	4-hydroxybenzalacetone[a]	–
	AsPKSIII2	naringenin chalcone[a,b]	(*E*)-4-hydroxy-6-(4-hydroxystyryl)-2*H*-pyran-2-one
	AsPKSIII3	naringenin chalcone[a,b]	(*E*)-4-hydroxy-6-(4-hydroxystyryl)-2*H*-pyran-2-one
ferulic acid (**2**)	AsPKSIII1	dehydrozingerone	–
	AsPKSIII2	(*E*)-4-hydroxy-6-(4-hydroxy-3-methoxystyryl)-2*H*-pyran-2-one	–
	AsPKSIII3	(*E*)-4-hydroxy-6-(4-hydroxy-3-methoxystyryl)-2*H*-pyran-2-one	–
caffeic acid (**3**)	AsPKSIII1	3,4-dihydroxybenzalacetone	–
	AsPKSIII2	2′,3,4,4′,6′-pentahydroxychalcone	(*E*)-6-(3,4-dihydroxystyryl)-4-hydroxy-2*H*-pyran-2-one
	AsPKSIII3	(*E*)-6-(3,4-dihydroxystyryl)-4-hydroxy-2*H*-pyran-2-one	2′,3,4,4′,6′-pentahydroxychalcone
sinapinic acid (**4**)	AsPKSIII1	–	–
	AsPKSIII2	–	–
	AsPKSIII3	–	–
dihydro-*p*-coumaric acid (**5**)	AsPKSIII1	–	–
	AsPKSIII2	6-(4-hydroxyphenethyl)-4-hydroxy-2*H*-pyran-2-one	phloretin
	AsPKSIII3	6-(4-hydroxyphenethyl)-4-hydroxy-2*H*-pyran-2-one	phloretin
cinnamic acid (**6**)	AsPKSIII1	distinct product peek (MS spectrum not detected)	–
	AsPKSIII2	pinocembrin chalcone	(*E*)-4-hydroxy-6-styryl-2*H*-pyran-2-one
	AsPKSIII3	(*E*)-4-hydroxy-6-styryl-2*H*-pyran-2-one	pinocembrin chalcone

The determination of whether a product was a main or secondary product was based on the area value of the product peak in the HPLC chromatogram using the absorbance at 280 nm.

[a] Products that were compared to the standard are marked with an asterisk.

[b] Naringenin chalcone was detected as flavanone (naringenin) after the extraction operation and compared to the standard.

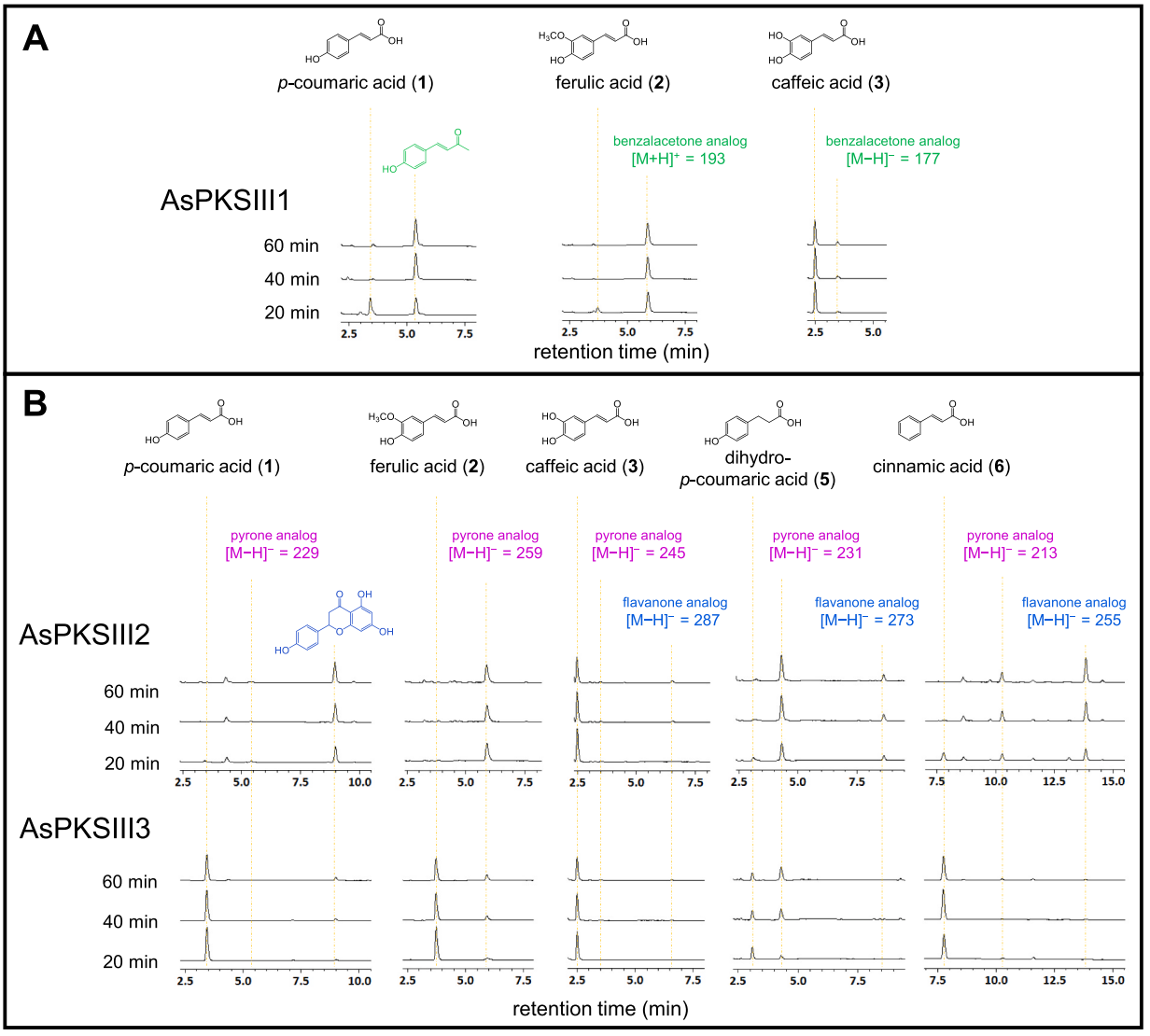


Fig. 3. HPLC chromatograms of AsPKSIII1, AsPKSIII2, and AsPKSIII3 recombinant *E. coli* culture medium extracts at 20, 40, and 60 min following the addition of cinnamic acid analogs. A: HPLC chromatograms of AsPKSIII1 reaction products. The chemical structure or *m/z* values of the product corresponding to benzalacetones are shown in green. B: HPLC chromatograms of AsPKSIII2 and AsPKSIII3 reaction products. The chemical structures or *m/z* values of the products corresponding to triketide pyrone and flavanone are shown in purple and blue, respectively.

chalcones, and produced only triketide pyrones (Fig. 3, S4). The ratio of triketide pyrone to chalcone formation differed depending on the type of cinnamic acid analogs used as a substrate. The pattern in which more triketide pyrones than chalcones were detected was when **5** was added to the AsPKSIII2 recombinant *E. coli* medium and when **3**, **5**, or **6** was added to the AsPKSIII3 recombinant *E. coli* (Table 1, Figs. S3–S7). Therefore, AsPKSIII3 was considered less active during the third malonyl-CoA condensation.

None of the AsPKSIIIs converted sinapinic acid (**4**) to the corresponding product. Because LE4CL-1 was previously reported to catalyze CoA esterification of **4** [18], we determined that all three AsPKSIII could not recognize sinapoyl-CoA as a substrate. Also, the consumption rate of **3** was significantly lower for all three AsPKSIIIs. Thus, the low activity of the three AsPKSIIIs toward caffeoyl-CoA was suggested. However, the detailed preference of each PKSIII for cinnamoyl-CoA analogs in this study is unclear because of the possible low activity of LE4CL-1 for **3**.

As mentioned above, two types of PKSIIIs (DCS and CURS) are involved in the synthesis of the diarylheptanoid curcumin. We constructed vectors in which the *AsPKSIII1* and *AsPKSIII1iso* were inserted into pETDuet-1 DNA, whereas *AsPKSIII2* and *AsPKSIII3* were inserted into pACYCDuet-1 DNA, respectively. However, no diarylheptanoids were detected in the *in vivo* reaction using *E. coli* transformed with four types of PKSIII genes. Therefore, none of the enzymes obtained in the present study exhibited CURS-like activity to catalyze the second-step reaction of the diarylheptanoids. In *A. sieboldiana*, linear-type diarylheptanoids were isolated from male flowers and cyclic-type diarylheptanoids from branch wood [7,9]. We propose that for *A. sieboldiana* diarylheptanoids, the C_6–C_7–C_6

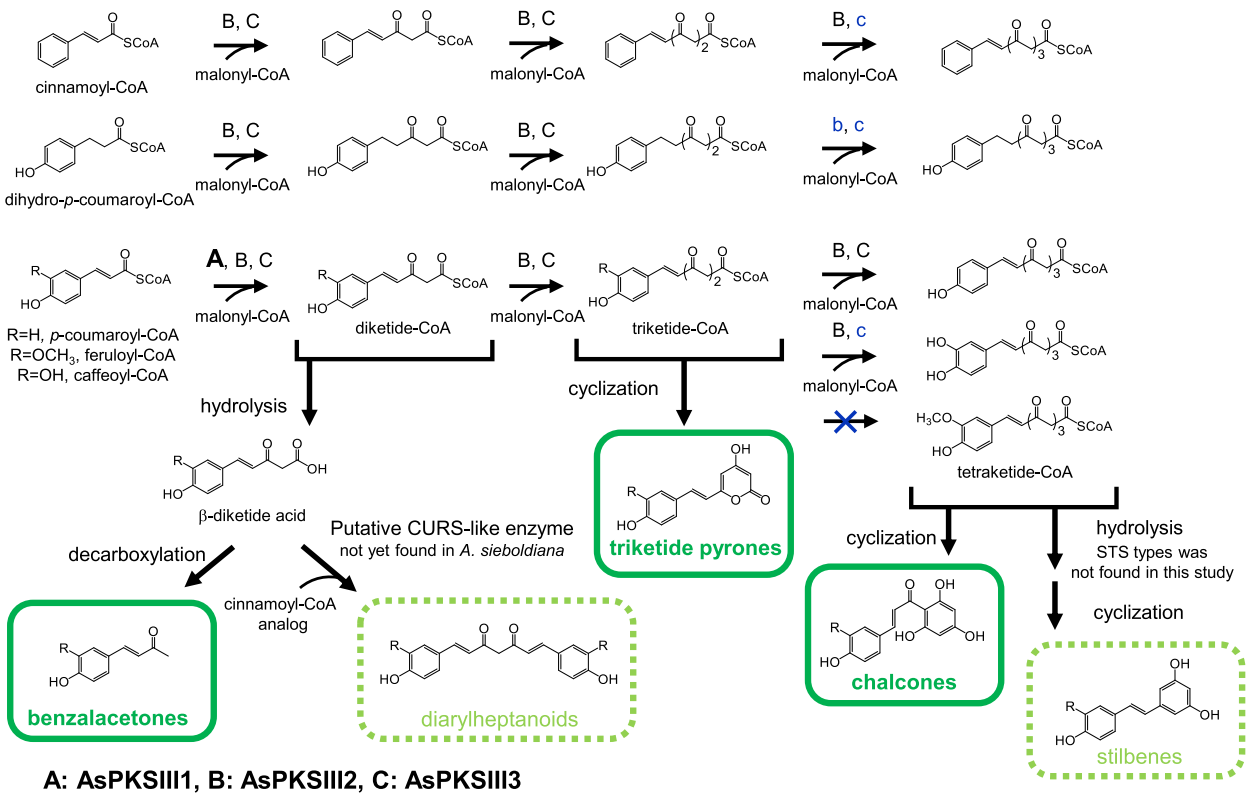


Fig. 4. Plant polyketide biosynthetic pathway and condensation reactions for three *A. sieboldiana* PKSIIIs. A, B, and C indicate the reaction pathways catalyzed by AsPKSIII1, AsPKSIII2, and AsPKSIII3, respectively. In the case of low detection levels of chalcones compared to triketide pyrones in the *in vivo* assay, the alphabet is shown in blue lowercase letters for the third malonyl-CoA condensation reaction pathway by CHS.

scaffolds are formed while retaining the double bond, and the heptane chain is subsequently saturated by double bond reductases [6]. The failure of AsPKSIII1 to recognize dihydro-*p*-coumaroyl-CoA as a substrate also supports our previous suggestion; however, CURS enzymes that utilize dihydrocinnamoyl-CoAs as substrates have been identified in *Zingiberaceae* plants [20,21]. Therefore, the acquisition and characterization of enzymes that form C_6–C_7–C_6 scaffolds, such as CURS-like, is essential to elucidate the diarylheptanoid biosynthesis pathway in *A. sieboldiana*. Also, stilbene was isolated from male flowers, but AsPKSIIIs cloned from seedlings did not produce stilbenes. AsSTS produces a small amount of chalcone [14], so further study of the assay conditions may reveal stilbene-producing potential for the three AsPKSIIIs. In the future, we expect to identify new PKSIIIs using the genetic information from these tissues.

In this study, three PKSIII genes were cloned from *A. sieboldiana*: two CHS-like enzymes and a DCS-like enzyme that condenses cinnamic acid analog CoA esters with a molecule of malonyl-CoA, and the *in vivo* enzyme products were examined. Further characterization of these PKSIIIs will be possible through *in vitro* functional analysis of these enzymes and examination of the products and phenotypes of recombinant or genome-edited plants that target the DNA sequences. Recently, we found that alnusonol, a cyclic-type diarylheptanoid isolated from *A. sieborudiana*, increases the number of nodules formed in *Frankia*-inoculated plants [5]. In *Casuarina glauca*, however, a decrease in the number of nodules was reported in CHS-like gene knockdown plants, suggesting that flavonoids are involved in actinorhizal symbioses [22]. In any case, it is very likely that low molecule polyketides, such as flavonoids and diarylheptanoids produced by PKSIII, are involved in actinorhizal symbiosis. The characterization of the genes and proteins of *A. sieboldiana* PKSIIIs in the present study may provide insight for the synthesis of beneficial plant polyketides and for the elucidation of the symbiotic mechanism between actinorhizal plants and *Frankia*.

3. Conclusions

We amplified and cloned three types of PKSIII (AsPKSIII1, AsPKSIII2, and AsPKSIII3) from *A. sieboldiana*. AsPKSIII1, AsPKSIII2, and AsPKSIII3 contain 1185 bp, 1170 bp, and 1176 bp CDSs encoding proteins of 43,222 Da, 42,517 Da, and 42,593 Da, respectively, sharing 87–51% amino acid sequence identity with other functionally characterized plant PKSIIIs.

In vivo reactions were performed in which *p*-coumaric acid, ferulic acid, caffeic acid, sinapinic acid, dihydro-*p*-coumaric acid, or cinnamic acid were added to recombinant *E. coli* cultures in which the 4CL and *ACC* genes were simultaneously transformed with the respective *AsPKSIII* genes. As a result, AsPKSIII1 synthesized diketide-CoAs corresponding to *p*-coumaric acid, ferulic acid, and caffeic

acid, suggesting that it is involved in the first step reaction in diarylheptanoid biosynthesis. On the other hand, AsPKSIII2 and AsPKSIII3 synthesized triketido pyrones and chalcones corresponding to *p*-coumaric acid, caffeic acid, dihydro-*p*-coumaric acid and cinnamic acid. The very low activity of AsPKSIII3 toward cinnamoyl-CoA strongly suggested that AsPKSIII2 is responsible for the biosynthesis of cinnamic acid-derived flavonoids in *A. sieboldiana*. The ratio of triketide pyrones to chalcones produced in the AsPKSIII2 and AsPKSIII3 assays differed depending on the substrate.

No products were produced from sinapinic acid in either AsPKSIII assay. Since neither diarylheptanoids nor stilbenoids were detected in this study, we considered it necessary to reexamine the enzymatic reaction conditions and obtain new AsPKSIII.

4. Materials and methods

4.1. Chemicals and materials

A. sieboldiana seeds from the Forestry and Forest Products Research Institute were used. Benlate wettable powder was purchased from Sumitomo Chemical Garden Products. Ampicillin sodium, kanamycin sulfate, EtOAc, isopropyl-1-thio-β-D-galactopyranoside (IPTG), and methanol (MeOH) were purchased from FUJIFILM Wako Pure Chemical. Streptomycin sulfate, *p*-coumaric acid (**1**), ferulic acid (**2**), caffeic acid (**3**), sinapinic acid (**4**), dihydro-*p*-coumaric acid (**5**), and naringenin were purchased from the Tokyo Chemical Industry. *trans*-Cinnamic acid (**6**) was purchased from Sigma-Aldrich. PrimeSTAR® Max DNA Polymerase, *Escherichia coli* JM109, restriction enzymes (*Nde*I, *Bam*HI, *Bgl*II) and T-Vector pMD19 were purchased from Takara Bio. A-overhang mixture used was included in the Mighty TA-cloning Reagent Set for PrimeSTAR® from Takara Bio. pET23a(+) DNA and pETDuet-1 DNA were purchased from Novagen. The restriction enzyme *Nco*I and *E. coli* BL21(DE3) were purchased from NIPPON GENE. HPLC-grade acetonitrile (CH_3CN) was purchased from Kanto Chemical. LCMS-grade CH_3CN was purchased from Supelco.

4.2. Plant materials

A. sieboldiana seeds were soaked in running water for a few days. The seeds surface was sterilized by stirring for 1 h in a benomyl solution adjusted to 0.5% (w/v) benlate wettable powder. Excess water was removed from the seeds using Kimwipes and they were transferred aseptically onto 0.9% agar plates prepared with 1/4 strength Hoagland's solution [5,23], 1/4 strength nitrogen-free Hoagland's solution, or pure water. The seeded agar plates were then placed in a growth chamber and germinated. Growth conditions in the chamber were daytime at 25 °C, 14 h, 115.5 μmol m^{-2} s^{-1} and nighttime at 20 °C, 10 h, 0 μmol m^{-2} s^{-1}. After germination, plants with root lengths of approximately 2 cm were used for RNA extraction.

4.3. Total RNA extraction, cDNA library preparation, and sequencing

RNA samples were prepared from *A. sieboldiana* seedlings by powdering them in liquid nitrogen using a mortar and pestle. Total RNA extraction was done using a previously reported method with slight modification [24]. Total RNA was DNase-treated using RNase-Free DNase Set (Qiagen) and purified using the RNeasy Plant Mini Kit (Qiagen). The quantity and quality of total RNA for next-generation sequencing was determined using an Agilent 2100 Bioanalyzer (Agilent Technologies), a Varioskan Lux microplate reader (Thermo Fisher Scientific), and a Qubit RNA HS Assay Kit (Thermo Fisher Scientific). The cDNA library was constructed with 1 μg of total RNA using the TruSeq Stranded mRNA Sample Preparation Kit (Illumina) and its instructions. The indexed cDNA libraries were pooled and sequenced using the next-generation sequencer NovaSeq 6000 platform (Illumina) to generate 2 × 151-bp paired-end sequence reads. Since the amount of RNA was insufficient in a single seedling, about 10 plants were combined as one RNA sample and applied for RNA-seq (n = 3 for each growth condition).

4.4. Search and cloning of the putative A. sieboldiana type III polyketide synthase gene

The raw reads generated were constructed into transcripts and protein coding sequences (CDS) within the transcript sequences were predicted as described previously [6]. A BLAST [25] search was performed (E-value < 1e−6) to estimate the *A. sieboldiana* PKSIII gene. Multiple alignment analysis of plant PKSIII sequences was conducted using ClustalW in Molecular Evolutionary Genetic Analysis (MEGA). Phylogenetic trees were generated using the neighbor-joining method.

For cloning, 1.5 μg of total RNA was reverse-transcribed using the AccuScript High Fidelity First Strand cDNA Synthesis Kit (Agilent) and the cDNA containing the full-length CDS of AsPKSIII1, AsPKSIII2, and AsPKSIII3 was amplified by PCR. The reactions were done using PrimeSTAR® Max DNA Polymerase and primers designed based on transcript sequences obtained by next-generation sequencing (Table S1). After adding dA to the 3′ end of the PCR fragment using an A-overhang mixture, it was ligated with the pMD19-T-vector using Ligation Mighty Mix and introduced into *E. coli* JM109 to confirm the sequence.

4.5. Construction of expression vectors and recombinant E. coli

The CDS region of each PKSIII DNA was amplified from pMD19-T-PKSIIIs vectors by PCR using primes containing restriction sites (Table S1). The resulting PCR fragments and pETDuet-1 vector were digested with the corresponding restriction enzymes, ligated using Ligation Mighty Mix, and introduced into *E. coli* JM109 to confirm DNA insertion and sequence. As a result, pETDuet-AsPKSIII1, pETDuet-AsPKSIII2, and pETDuet-AsPKSIII3 were constructed (Fig. S1). Each pETDuet-AsPKSIII, the pCDF-LE4CL-1 carrying the 4-

coumarate: CoA ligase gene from *Lithospermum erythrohizon* [17] for the cinnamoyl-CoA ester synthesis step, and the pRSF-ACC carrying the two subunit genes for acetyl-CoA carboxylase from *Corynebacterium glutamicum* [18], were simultaneously introduced into *E. coli* BL21(DE3) (Fig. S2). Finally, cinnamic acids were added to the culture medium of recombinant *E. coli* to produce polyketides.

4.6. In vivo production of polyketides in E. coli by the multiplasmid methods

In vivo reactions were performed with some modifications of previous methods [17]. Briefly, recombinant *E. coli* BL21(DE3) cells were precultured in 5 ml LB medium overnight at 37 °C. Subsequently, 600 μl of the pre-culture solution was transferred to 30 ml of LB medium and incubated at 37 °C. When the OD_{600} reached 0.6, IPTG (1 mM) was added and incubated at 26 °C for 5 h. The recombinant *E. coli* were then collected by centrifugation and 25 g/l wet weight of *E. coli* cells were added to 1 ml of M9 medium containing 0.1 mM substrate (cinnamic acid analogs; **1**, **2**, **3**, **4**, **5**, or **6**), 25 g/l calcium carbonate, 40 g/l glucose, antibiotics (50 mg/l streptomycin, 50 mg/l kanamycin, and 100 mg/l ampicillin), and 1 mM IPTG and incubated at 26 °C for 60 h. The reaction was quenched by adding 3 drops of 6 M HCl using a Pasteur pipette, the product was extracted three times with 800 μl of EtOAc, and the organic layer was evaporated to dryness. The residue was dissolved in 1.0 ml of methanol and further diluted 6-fold to produce a sample for HPLC and LCMS analysis. HPLC analysis of the extracts was performed 20, 40, and 60 min after the start of the reaction to compare the amount of each substrate converted over a short period of time. The reactions were performed in a total volume of 500 μl. After ethyl acetate extraction and evaporation of the solvent, the residue was dissolved in 1.0 ml of methanol and used for HPLC analysis. Technical replicates of *in vivo* experiments were performed at least twice.

4.7. HPLC and LCMS analytical conditions

Analysis of the products was performed using a liquid chromatograph (Shimadzu Prominence) equipped with a mass spectrometer (LCMS-2020). TSKgel-ODS-100V 3 μm reversed-phase column (2.0 mm × 150 mm, TOSOH) was used to separate the enzymatic reaction products. Mobile phase A (CH_3CN)/mobile phase B (water containing 0.1% formic acid) was used as elution solvent. For the eluent gradient, the flow was from 30% A to 100% A for 30 min, then maintained at 100% A for 10 min, then from 100% A to 30% A for 10 min, then maintained at 100% A for 10 min. The column temperature was 40 °C and the flow rate was constant at 0.25 ml/min. The injection volume was generally 2 μl, but was changed to 1 μl when applied to the LCMS.

Data availability statement

Data will be made available on request.

CRediT authorship contribution statement

Konosuke Takemoto: Writing – original draft, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization. **Yuichi Mikota:** Project administration, Methodology, Investigation. **Ryota Moriuchi:** Software, Investigation. **Yuko Yoneda:** Writing – review & editing, Project administration. **Shingo Kawai:** Writing – review & editing, Supervision, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This study was supported by JSPS KAKENHI Grant Numbers-16H04951 and 21H02254, and a PHOENIX Grant-in-Aid (K. Takemoto, 2021). We are deeply grateful to Dr. Hisashi Kawasaki at Tokyo University for granting us permission to use pRSF-ACC and Dr. Kazufumi Yazaki at Kyoto University for granting us permission to use the *LE4CL-1* cDNA. Also, we are deeply grateful to Dr. Yohei Katsuyama at Tokyo University for his valuable advice for the experiments and helping us with the procedure for the distribution of pRSF-ACC and pCDF-LE4CL-1.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e27698.

References

[1] S. Kitamura, G. Murata, Colored Illustrations of Woody Plants of Japan, vol. II, Hoikusha, Osaka, 1979, pp. 284–290 (in Japanese).
[2] T. Yamanaka, H. Okabe, Distribution of *Frankia*, ectomycorrhizal fungi, and bacteria in soil after the eruption of Miyake-Jima (Izu Islands, Japan) in 2000, J. For. Res. 11 (2006) 21–26, https://doi.org/10.1007/s10310-005-0179-2.
[3] M. Komukai, K. Fukunaga, R. Tachibana, Relationship between seed dispersal and invasive plants on a revegetation slope dominated by *Alnus sieboldiana* Matsum, JJpnSocRevegetTech 38 (2012) 21–26, https://doi.org/10.7211/jjsrt.38.21.
[4] J.O. Dawson, Ecology of actinorhizal plants, in: K. Pawlowski, W.E. Newton (Eds.), Nitrogen-Fixing Actinorhizal Symbioses, Springer Nature, Switzerland, 2008, pp. 199–234.
[5] A. Tsurugi-Sakurada, T. Kaneko, K. Takemoto, Y. Yoneda, T. Yamanaka, S. Kawai, Cyclic diarylheptanoids as potential signal compounds during actinorhizal symbiosis between *Alnus sieboldiana* and *Frankia*, Fitoterapia 162 (2022) 105284, https://doi.org/10.1016/j.fitote.2022.105284.
[6] K. Takemoto, A. Tsurugi-Sakurada, R. Moriuchi, Y. Yoneda, S. Kawai, Cloning and characterization of NADPH-dependent double-bond reductases from *Alnus sieboldiana* that recognize linear diarylheptanoids as substrates, Phytochemistry 215 (2023) 113850, https://doi.org/10.1016/j.phytochem.2023.113850.
[7] K. Chiba, H. Ichizawa, S. Kawai, T. Nishida, α-Glucosidase inhibition activity by cyclic diarylheptanoids from *Alnus sieboldiana*, J. Wood Chem. Technol. 33 (2013) 44–51, https://doi.org/10.1080/02773813.2012.723778.
[8] A. Ludwiczuk, A. Saha, T. Kuzuhara, Y. Asakawa, Bioactivity guided isolation of anticancer constituents from leaves of *Alnus sieboldiana* (Betulaceae), Phytomedicine 18 (2011) 491–498, https://doi.org/10.1016/j.phymed.2010.10.005.
[9] T. Hashimoto, M. Tori, Y. Asakawa, Five new diarylheptanoids from the male flowers of *Alnus sieboldiana*, Chem. Pharm. Bull. 34 (1986) 1846–1849, https://doi.org/10.1248/cpb.34.1846.
[10] Y. Asakawa, Chemical constituents of *Alnus sieboldiana* (Betulaceae) II. The isolation and structure of flavonoids and stilbenes, Bull. Chem. Soc. Jpn. 44 (1971) 2761–2766, https://doi.org/10.1246/bcsj.44.2761.
[11] H. Morita, C.P. Wong, I. Abe, How structural subtleties lead to molecular diversity for the type III polyketide synthases, J. Biol. Chem. 294 (2019) 15121–15136, https://doi.org/10.1074/jbc.REV119.006129.
[12] J.L. Ferrer, J.M. Jez, M.E. Bowman, R.A. Dixon, J.P. Noel, Structure of chalcone synthase and the molecular basis of plant polyketide biosynthesis, Nat. Struct. Biol. 6 (1999) 775–784, https://doi.org/10.1038/11553.
[13] Y. Katsuyama, T. Kita, N. Funa, S. Horinouchi, Curcuminoid biosynthesis by two type III polyketide synthases in the herb *Curcuma longa*, J. Biol. Chem. 284 (2009) 11160–11170, https://doi.org/10.1074/jbc.M900070200.
[14] T. Yamaguchi, F. Kurosaki, D.Y. Suh, U. Sankawa, M. Nishioka, T. Akiyama, M. Shibuya, Y. Ebizuka, Cross-reaction of chalcone synthase and stilbene synthase overexpressed in *Escherichia coli*, FEBS Lett. 460 (1999) 457–461, https://doi.org/10.1016/S0014-5793(99)01403-9.
[15] Y. Katsuyama, M. Matsuzawa, N. Funa, S. Horinouchi, *In vitro* synthesis of curcuminoids by type III polyketide synthase from *Oryza sativa*, J. Biol. Chem. 282 (2007) 37702–37709, https://doi.org/10.1074/jbc.M707569200.
[16] T. Mori, Y. Shimokawa, T. Matsui, K. Kinjo, R. Kato, H. Noguchi, S. Sugio, H. Morita, I. Abe, Cloning and structure-function analyses of quinolone- and acridone-producing novel type III polyketide synthases from *Citrus microcarpa*, J. Biol. Chem. 288 (2013) 28845–28858, https://doi.org/10.1074/jbc.M113.493155.
[17] Y. Katsuyama, N. Funa, I. Miyahisa, S. Horinouchi, Synthesis of unnatural flavonoids and stilbenes by exploiting the plant biosynthetic pathway in *Escherichia coli*, Chem. Biol. 14 (2007) 613–621, https://doi.org/10.1016/j.chembiol.2007.05.004.
[18] I. Miyahisa, M. Kaneko, N. Funa, H. Kawasaki, H. Kojima, Y. Ohnishi, S. Horinouchi, Efficient production of (2*S*)-flavanones by *Escherichia coli* containing an artificial biosynthetic gene cluster, Appl. Microbiol. Biotechnol. 68 (2005) 498–504, https://doi.org/10.1007/s00253-005-1916-3.
[19] I. Abe, Y. Takahashi, H. Morita, H. Noguchi, Benzalacetone synthase. A novel polyketide synthase that plays a crucial role in the biosynthesis of phenylbutanones in *Rheum palmatum*, Eur. J. Biochem. 268 (2001) 3354–3359, https://doi.org/10.1046/j.1432-1327.2001.02255.x.
[20] L. Zhang, B. Gao, X. Wang, Z. Zhang, X. Liu, J. Wang, T. Mo, Y. Liu, S. Shi, P. Tu, Identification of a new curcumin synthase from ginger and construction of a curcuminoid-producing unnatural fusion protein diketide-CoA synthase::curcumin synthase, RSC Adv. 6 (2016) 12519–12524, https://doi.org/10.1039/C5RA23401H.
[21] R. Chen, T. Hu, M. Wang, Y. Hu, S. Chen, Q. Wei, X. Yin, T. Xie, Functional characterization of key polyketide synthases by integrated metabolome and transcriptome analysis on curcuminoid biosynthesis in *Curcuma wenyujin*, Synth. Syst. Biotechnol. 7 (2022) 849–861, https://doi.org/10.1016/j.synbio.2022.04.006.
[22] K. Abdel-Lateif, V. Vaissayre, H. Gherbi, C. Verries, E. Meudec, F. Perrine-Walker, V. Cheynier, S. Svistoonoff, C. Franche, D. Bogusz, V. Hocher, Silencing of the chalcone synthase gene in *Casuarina glauca* highlights the important role of flavonoids during nodulation, New Phytol. 199 (2013) 1012–1021, https://doi.org/10.1111/nph.12326.
[23] D.I. Arnon, D.R. Hoagland, Crop production in artificial solutions and soils with special reference to factors influencing yield and absorption of inorganic nutrients, Soil Sci. 50 (1940) 463–471.
[24] Y. Suzuki, T. Mae, A. Makino, RNA extraction from various recalcitrant plant tissues with a cethyltrimethylammonium bromide-containing buffer followed by an acid guanidium thiocyanate-phenol-chloroform treatment, Biosci. Biotechnol. Biochem. 72 (2008) 1951–1953, https://doi.org/10.1271/bbb.80084.
[25] S.F. Altschul, W. Gish, W. Miller, E.W. Myers, D.J. Lipman, Basic local alignment search tool, J. Mol. Biol. 215 (1990) 403–410, https://doi.org/10.1016/S0022-2836(05)80360-2.