


Article

# Comparative Transcriptomics Analysis for Gene Mining and Identification of a Cinnamyl Alcohol Dehydrogenase Involved in Methyleugenol Biosynthesis from *Asarum sieboldii* Miq.

Jinjie Liu <sup>1</sup>, Chong Xu <sup>2</sup>, Honglei Zhang <sup>3</sup>, Fawang Liu <sup>1</sup>, Dongming Ma <sup>2,\*</sup> and Zhong Liu <sup>1,\*</sup> 

<sup>1</sup> School of Pharmacy, Shanghai Jiao Tong University, Shanghai 200240, China; LHJW161903@sjtu.edu.cn (J.L.); fawang90@126.com (F.L.)

<sup>2</sup> Research Center of Chinese Herbal Resource Science and Engineering, Guangzhou University of Chinese Medicine, Guangzhou 510006, China; 15013053420@163.com

<sup>3</sup> Jiusan administration of Heilongjiang farms & land reclamation, Harbin 161441, China; ZHL5555@163.com

\* Correspondence: madm@gzucm.edu.cn (D.M.); liuzhong@sjtu.edu.cn (Z.L.)

Received: 20 November 2018; Accepted: 1 December 2018; Published: 3 December 2018



**Abstract:** *Asarum sieboldii* Miq., one of the three original plants of TCM ASARI RADIX ET RHIZOMA, is a perennial herb distributed in central and eastern China, the Korean Peninsula, and Japan. Methyleugenol has been considered as the most important constituent of *Asarum* volatile oil, meanwhile asarinin is also employed as the quality control standard of ASARI RADIX ET RHIZOMA in Chinese Pharmacopeia. They both have shown wide range of biological activities. However, little was known about genes involved in biosynthesis pathways of either methyleugenol or asarinin in *Asarum* plants. In the present study, we performed de novo transcriptome analysis of plant tissues (e.g., roots, rhizomes, and leaves) at different developmental stages. The sequence assembly resulted in 311,597 transcripts from these plant materials, among which 925 transcripts participated in ‘secondary metabolism’ with particularly up to 20.22% of them falling into phenylpropanoid biosynthesis pathway. The corresponding enzymes belong to seven families potentially encoding phenylalanine ammonia-lyase (PAL), trans-cinnamate 4-monooxygenase (C4H), *p*-coumarate 3-hydroxylase (C3H), caffeoyl-CoA *O*-methyltransferase (CCoAOMT), cinnamoyl-CoA reductase (CCR), cinnamyl alcohol dehydrogenase (CAD), and eugenol synthase (EGS). Moreover, 5 unigenes of DIR (dirigent protein) and 11 unigenes of CYP719A (719A subfamily of cytochrome P450 oxygenases) were speculated to be involved in asarinin pathway. Of the 15 candidate CADs, four unigenes that possessed high FPKM (fragments per transcript kilobase per million fragments mapped) value in roots were cloned and characterized. Only the recombinant AsCAD5 protein efficiently converted *p*-coumaryl, coniferyl, and sinapyl aldehydes to their corresponding alcohols, which are key intermediates employed not only in biosynthesis of lignin but also in that of methyleugenol and asarinin. qRT-PCR revealed that AsCAD5 had a high expression level in roots at three developmental stages. Our study will provide insight into the potential application of molecular breeding and metabolic engineering for improving the quality of TCM ASARI RADIX ET RHIZOMA.

**Keywords:** *Asarum sieboldii* Miq.; ASARI RADIX ET RHIZOMA; transcriptome; cinnamyl alcohol dehydrogenase; methyleugenol biosynthesis pathway

## 1. Introduction

The genus *Asarum*, containing about 100 species which have conventionally been used as herbal medicines, are distributed in East Asia (65 species), North America (15 species), and Europe

(1 species) [1,2]. The Chinese Pharmacopeia has all along enlisted ASARI RADIX ET RHIZOMA, which is stipulated to be dried roots and rhizomes of three different taxa, that is, *Asarum heterotropoides* Fr. Schmidt var. *mandshuricum* (Maxim.) Kitag. (called Bei Xixin in Chinese), *A. sieboldii* Miq. var. *seoulense* Nakai (called Hancheng Xixin in Chinese), or *A. sieboldii* Miq. (called Hua Xixin in Chinese) [3]. *A. heterotropoides* var. *mandshuricum* and *A. sieboldii* var. *seoulense* are both distributed in mountainous areas around Changbaishan within Northeast China. For the similar habitat and overlapped distribution, *A. heterotropoides* var. *mandshuricum* and *A. sieboldii* var. *seoulense* usually tend to be mixed together so that they are collectively called Liao Xixin in Chinese. In contrast, *A. sieboldii* grows in central and eastern China with the Qinling–Dabashan mountainous areas as the center of its distribution, clearly separating it from the two former varieties on the basis of occupation [1].

However, the reports available until now demonstrated that all three taxa have very similar chemical characteristics and biological activities. Both lignans (especially for asarinin and sesamin) and volatile oil (especially for methyleugenol and kakuol) have been regarded as the major active chemicals of ASARI RADIX ET RHIZOMA [4–6]. In traditional Chinese medicine, ASARI RADIX ET RHIZOMA has been widely therapeutically used in numbers of formula for a long history, mainly functioning in treatment of toothache, cough, chronic bronchitis, headache, etc. due to its anti-inflammatory, analgesic, and memory enhancing effects [7–12].

Asarinin, often found in vascular plants like *Sesamum* spp, is the epimer of Sesamin, a furofuran class lignan. In *Sesamum indicum* cytochrome P450 (SiP450), CYP81Q1, has been reported to catalyze sesamin biosynthesis by forming methylenedioxy bridges [13,14]. Pinoresinol could be formed from two coniferyl alcohol radical substrates by dirigent protein (DIR) [15–19]. In *A. sieboldii*, asarinin and sesamin hold stronger antitumor, antibacterial, and antiallergy effects [20–23]. However, asarinin biosynthesis pathway in *A. sieboldii* remains unknown.

Methyleugenol, possessing antitussive [24] and antinociceptive effect [25], is one of the most important constituents of *Asarum* volatile oil. Some studies even revealed the possibility of methyleugenol as a leading compound for anorexia treatment [26]. It is worth highlighting that the biosynthesis of methyleugenol shares the common precursor, coniferyl alcohol, with G-lignin. Due to the importance of lignin in the economy, as well as of methyleugenol in food and cosmetic industry, the elucidation of this metabolic pathway has drawn more attention. Very recently, coniferyl alcohol acyl transferase (CAAT) has been cloned and characterized in apple fruit [27]. Furthermore, eugenol synthase genes (EGS) have also been reported in some species such as Rose, Ocimum, and *Gymnadenia* [28–30]. The genes coding O-methyltransferases (EOMT) involved in methyleugenol have been isolated and functionally characterized in loquat (*Eriobotrya japonica*), *Ocimum tenuiflorum*, *Clarkia breweri*, sweet basil (*Ocimum basilicum*), and ripe apple fruit [31–35]. Cinnamyl alcohol dehydrogenase (CAD; EC 1.1.1.195), a member of alcohol dehydrogenase, functions in converting hydroxycinnamyl aldehydes to corresponding alcohols, the various direct precursors of lignin [36], among which the coniferyl alcohol also acts as the precursor in methyleugenol biosynthesis. Therefore, the production of coniferyl alcohol and its metabolic flow plays an essential role in the biosynthesis of methyleugenol. Many studies have performed in its structure properties and activity assay in vitro. For example, the nine putative cinnamyl alcohol dehydrogenase genes were reclassified in *Arabidopsis*. They were divided into three types: two (AtCAD4, AtCAD5) had the highest activity and regarded as bona fide CADs; four (AtCAD2, 3, as well as AtCAD7 and 8) had low activity and highest homology to sinapyl alcohol dehydrogenase; the rest (AtCAD6, AtCAD9) were considered functionally redundant in CAD metabolic networks [37]. In recent years, CADs have been characterized in many plants such as poplars, *Oryza sativa*, *Rauvolfia*, etc. [38–40]. Although the entire biosynthesis pathway of methyleugenol has been completely elucidated in sweet basil [32], methyleugenol biosynthesis pathway in *A. sieboldii* remains poorly understood. Thus, it is significant to isolate and characterize AsCADs in *A. sieboldii*, which is involved in coniferyl alcohol production.

In this study, transcriptomic sequencing and analysis from five tissues at different developmental stages from *A. sieboldii* were performed to uncover candidate genes involved in methyleugenol

and asarinin biosynthesis pathway, with more attention to the CADs. Of the 15 candidate CADs, 4 unigenes with the higher FPKM (fragments per transcript kilobase per million fragments mapped) in root were cloned and confirmed by qRT-PCR. Of the four recombinant proteins, only AsCAD5 displayed the capability for converting aldehydes (*p*-coumaryl, coniferyl, and sinapyl aldehydes) to their corresponding alcohols *in vitro*. Enzyme kinetics of three aldehydes was conducted to understand the suitable substrates. All of the results provide a good scenario for pathway elucidation, as well as CADs functions possibly involved in methyleugenol and asarinin pathway. Our study will help in improving the quality of *A. sieboldii* through breeding and metabolic regulations.

## 2. Results

### 2.1. De Novo Assembly and Sequence Annotation

A morphological feature of the intact plant and the representative morphology of root, leaf (including leaf blade and petiole) and rhizome were shown in Figure 1. A total of 89,141,416; 80,100,512; 86,269,700; 88,819,410; and 92,256,930 clean reads were obtained, respectively from different tissues at different developmental stages: root in May (Sample 1), rhizome in May (Sample 2), rhizome in March (Sample 3), rhizome in July (Sample 4) and leaf in May (Sample 5). All the clean reads of five transcriptome sequencing databases were integrated and assembled into 311,597 unigenes. Taking together, 128,934 unigenes (41.38%) were annotated via public databases including NCBI non-redundant protein (Nr), Swiss-Prot protein, Cluster of Orthologous Groups (COG), and the Kyoto Encyclopedia of Genes and Genomes (KEGG). Among them, 9741 unigenes showed significant matches to all four databases. Unigenes that were annotated as unique in various public databases are as follows: 127,167 unigenes in the Nr database, 76,685 unigenes in the Swiss-Prot database, 22,386 unigenes in the COG database, and 35,326 unigenes in the KEGG database (Table 1). Furthermore, about 58.62% of unigenes (182,663) did not show any matches to known genes. They may be considered novel transcripts and specific genes from *A. sieboldii*.

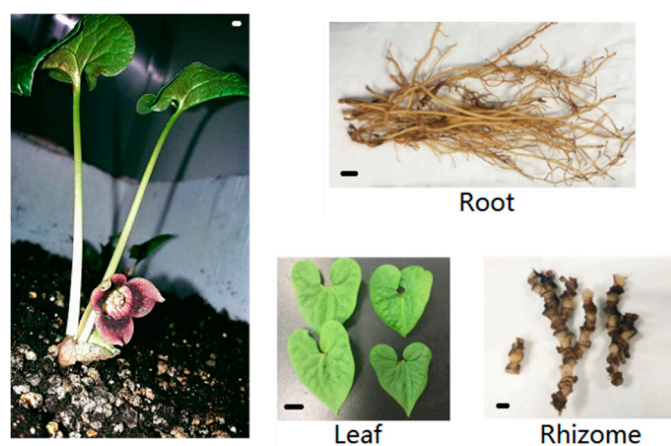


Figure 1. Representative materials diagrams. Scale bars = 1 cm.

Table 1. Summary of number of unigenes annotated in *Asarum* as compared to public database

Public Database	Number of Unigenes Annotated in <i>Asarum sieboldii</i>
Nr	127,167
Swiss-Prot	76,685
COG	22,386
KEGG	35,326

Note: Total number of annotated unigenes: 128,934; Rate of total annotation: 41.38%.

## 2.2. Gene Ontology Classification, COG Classification, and Metabolic Pathway Assignment by KEGG

GO-Term Finder was frequently used to identify Gene Ontology (GO) terms. In total, 157,713 unigenes possessed assignment to one or more gene ontology categories, while 46,782 unigenes were from the cellular component, 61,524 from the biological process, and 49,407 from the molecular function. All unigenes were predicted and classified against the COG database. As a result, 22,386 unigenes were annotated and grouped into 25 COG classifications. Among them, the cluster for metabolic process was 16,611 (10.53%). Some unigenes were assigned to multiple COG classifications; totally 69,969 COG functional annotations were obtained (Table S1, Figures S1 and S2).

The KEGG annotation results showed 35,326 unigenes (11.63%) were obtained and assigned to 369 KEGG pathways, with leading roles in six categories including metabolism, cellular processes, genetic information processing, environmental information processing, organismal systems, and human diseases (Table S2, Figure 2). Moreover, metabolism occupied highest number of transcripts from *A. sieboldii*, accounting for 26.59% (22,194). Altogether 925 transcripts were involved in biosynthesis of various secondary metabolites. The cluster for ‘phenylpropanoid biosynthesis [PATH: ko00940]’ and ‘terpenoid backbone biosynthesis [PATH: ko00900]’ represented the largest group (20.22%), which was indicated in Figure 3. This information will provide useful guidance for further research on specific pathways and related genes in *A. sieboldii*.

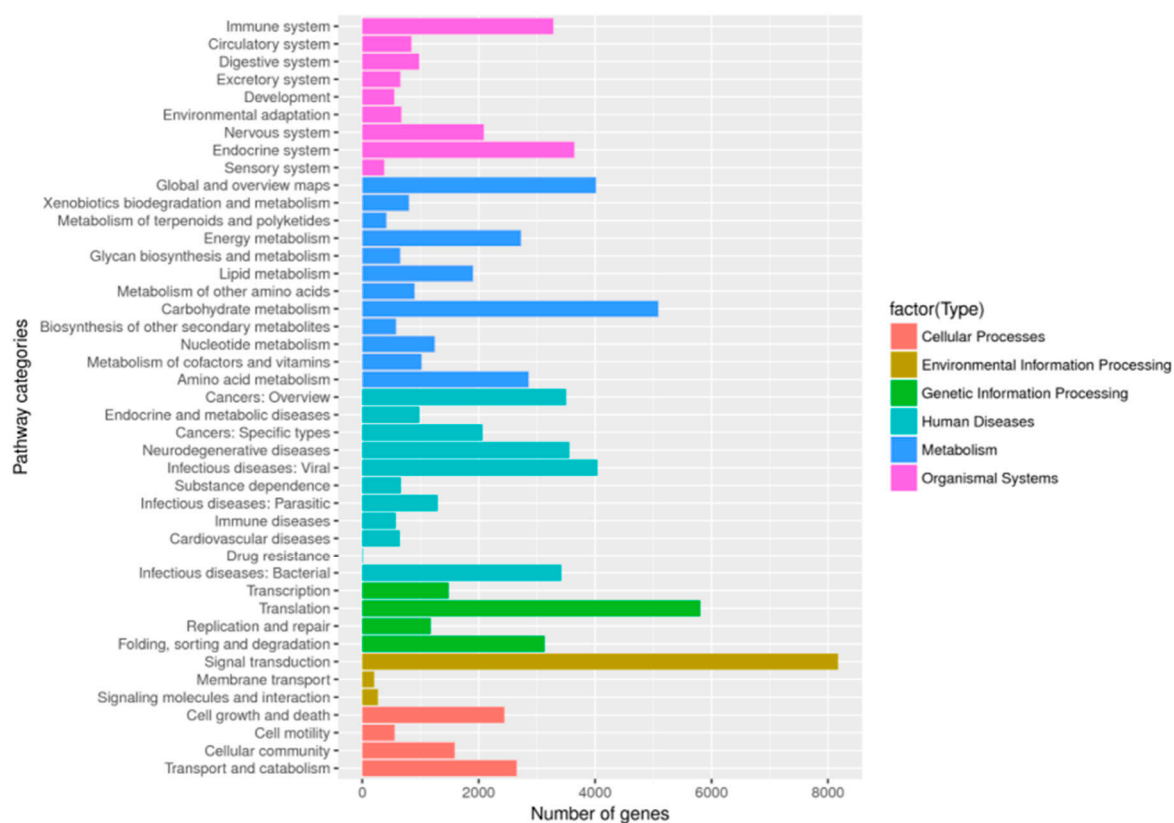
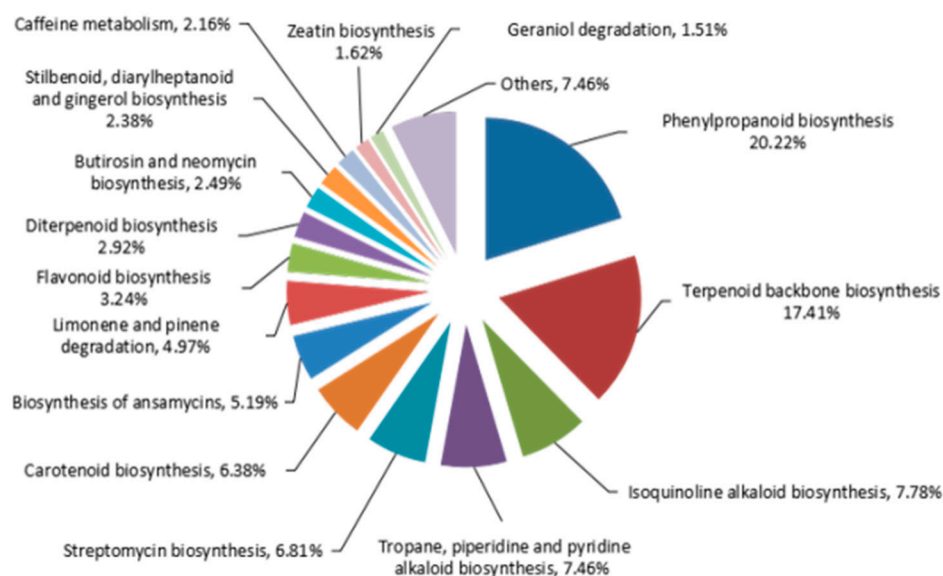


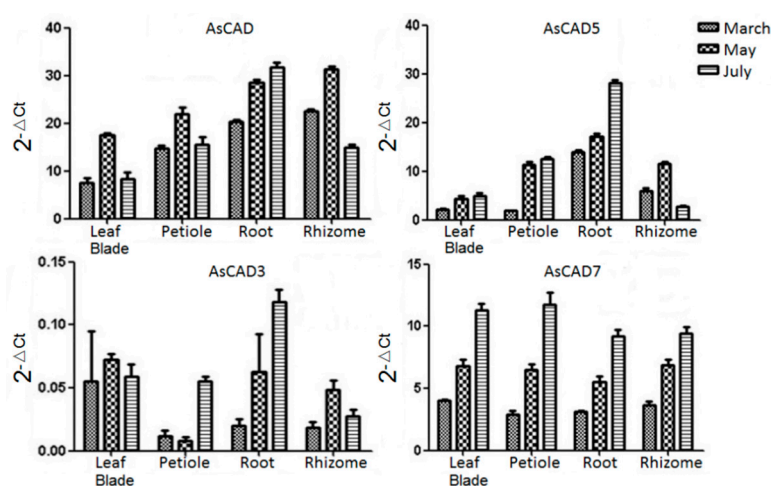
Figure 2. KEGG annotation and classification of unigenes.



**Figure 3.** KEGG classification based on secondary metabolism categories. Others containing: tetracycline biosynthesis; novobiocin biosynthesis; brassinosteroid biosynthesis; sesquiterpenoid and triterpenoid biosynthesis; glucosinolate biosynthesis; biosynthesis of siderophore group nonribosomal peptides; betalain biosynthesis; monoterpene biosynthesis; polyketide sugar unit biosynthesis; indole alkaloid biosynthesis; flavone and flavonol biosynthesis.

### 2.3. Gene Expression Analysis of AsCADs by qRT-PCR

For the purpose of clarifying the expression of AsCAD genes in different tissues, the relative expression level of AsCADs was evaluated at three stages, with 18S rRNA as reference gene. Before qRT-PCR experiment, we speculated petiole possibly contained lignin to some extent. Thus, leaf blade and petiole have been separated as qRT-PCR templates. The results showed AsCAD5 in roots had higher expression than in other tissues (e.g., leaf blades, petioles, and rhizomes) at each stage. AsCAD showed a similar expression pattern between roots and rhizomes, AsCAD7 showed a similar expression pattern between different tissues. For AsCAD3, the expression level remained low either in roots or in other tissues when compared with AsCAD, AsCAD5, and AsCAD7 (Figure 4).

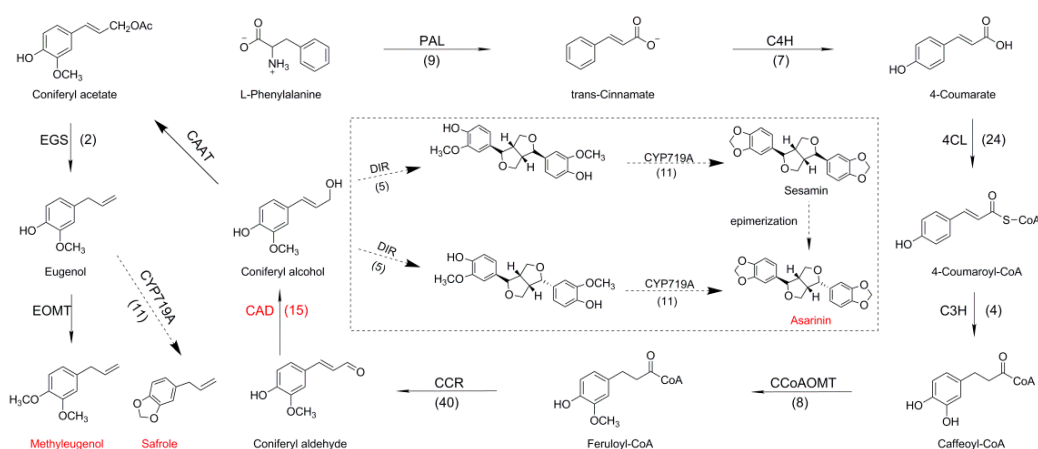


**Figure 4.** Analysis of AsCADs relative expression in leaf blades, rhizomes, roots, and petioles at three developmental stages (collected in March, May, and July) by real-time quantitative PCR. 18S rRNA was used as reference gene.



#### 2.4. Identifying Genes Involved in Biosynthesis of Phenylpropanoids Pathway in *A. sieboldii*

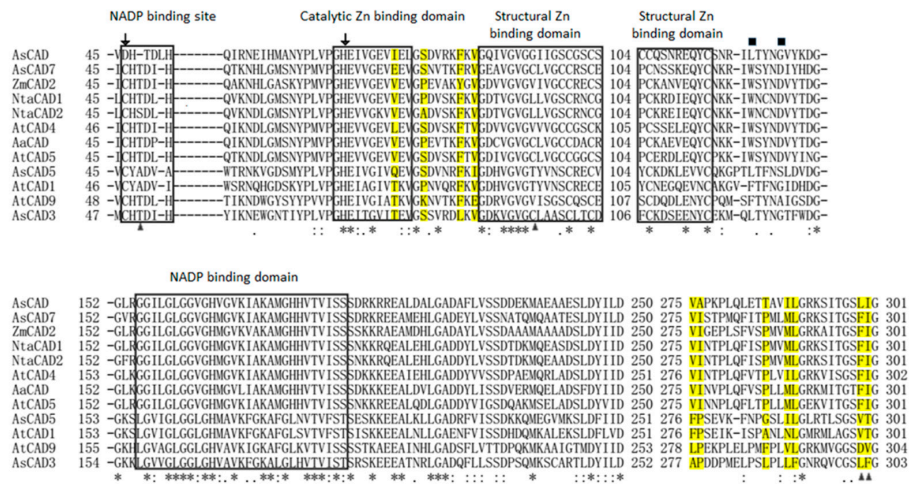
According to the reported methyleugenol biosynthesis pathway, the corresponding enzymes belonged to eight families potentially encoding phenylalanine ammonia-lyase (PAL) (9 unigenes), trans-cinnamate 4-monooxygenase (C4H) (7 unigenes), 4-coumarate-CoA ligase (4CL) (24 unigenes), *p*-coumarate 3-hydroxylase (C3H) (4 unigenes), caffeoyl-CoA *O*-methyltransferase (CCoAOMT) (8 unigenes), cinnamoyl-CoA reductase (CCR) (40 unigenes), cinnamyl alcohol dehydrogenase (CAD) (15 unigenes), and eugenol synthase (EGS) (2 unigenes) (Figure 5). One family potentially encoding dirigent proteins (DIR) (five unigenes) involved in asarinin pathway has also been identified. Plant cytochrome P450 (P450), CYP719A subfamily, can catalyze methylenedioxy bridge-forming reactions in isoquinoline alkaloid-producing plant species [41]. Eleven unigenes of CYP719A were found in *A. sieboldii* and a heatmap based on transcription levels of 512 CYP450s unigenes in the five samples are shown in Figure S3. It is worth noting that 8 of 11 CYP719A had high expression in roots.



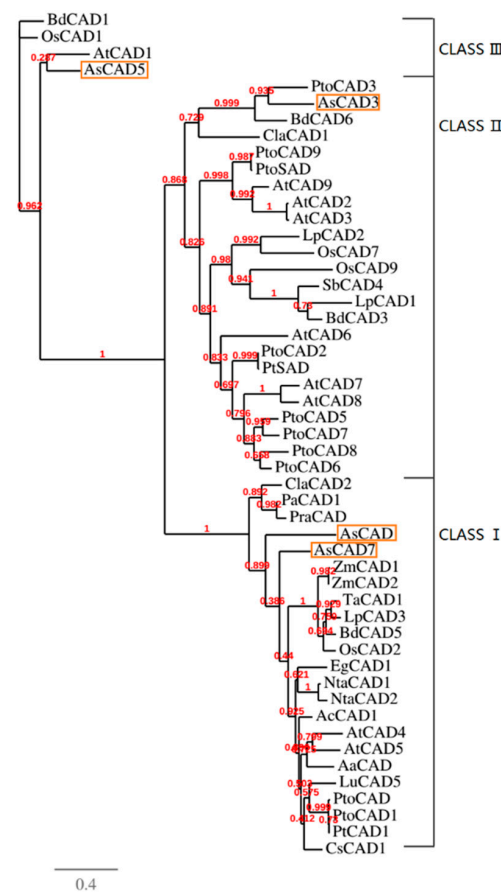
**Figure 5.** A proposed pathway for methyleugenol and asarinin biosynthesis in *A. sieboldii*. Abbreviations: PAL, phenylalanine ammonia-lyase; C4H, trans-cinnamate 4-monooxygenase; 4CL, 4-coumarate-CoA ligase; C3H, *p*-coumarate 3-hydroxylase; CCoAOMT, caffeoyl-CoA *O*-methyltransferase; CCR, cinnamoyl-CoA reductase; CAD, cinnamyl alcohol dehydrogenase; CAAT, cinnamyl alcohol acyl transferase; EGS, eugenol synthase; EOMT, eugenol *O*-methyltransferase; DIR, dirigent proteins; CYP719As belong to cytochrome P450 family. Bracketed numbers represent numbers of transcripts respectively.

#### 2.5. Cloning and Phylogenetic Analysis of AsCADs

Of the 15 candidate CADs, four unigenes AsCAD, AsCAD3, AsCAD5, and AsCAD7 (Genbank accession number: KT454711, KY446442, KY446444, and KY446446) with the higher FPKM were cloned. As shown in Figure 6, the alignment of AsCADs protein sequences with the representatives of the CAD family in *Arabidopsis thaliana*, *Artemisia annua*, *Nicotiana tabacum*, and *Zea mays* uncovered that all of the candidates for AsCADs had a highly conserved catalytic zinc binding domain, and AsCAD and AsCAD7 were highly identical with AtCAD5 at all key residues [42]. In phylogeny analysis AsCADs were gathered into three major groups concerning the CAD protein family (Figure 7). Several classification treatments of the CAD family have been proposed based on substrate specificity, expression pattern and gene structure [43–47]. According to them, AsCAD and AsCAD7 nested in Class I, also called ‘bona fide’. This group was highly conserved in terms of substrate specificity, which was represented by AtCAD4 and AtCAD5, enzymes associated with primary lignin synthesis in *Arabidopsis*. AsCAD3 belonged to Class II, whereas AsCAD5 fell into Class III for its relatively high similarity with AtCAD1.



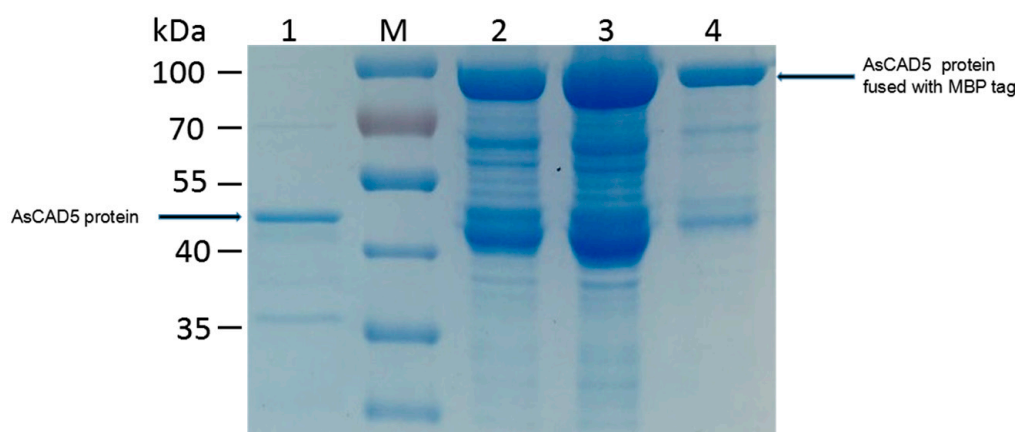
**Figure 6.** Alignment of putative AsCAD proteins with eight known CAD proteins from other plants. *Arabidopsis thaliana* (AtCAD1, AtCAD4, AtCAD5, AtCAD9), *Nicotiana tabacum* (NtaCAD1, NtaCAD2), *Artemisia annua* (AaCAD), and *Zea mays* (ZmCAD2) (alignment—Clustal Omega software 1.2.2). Arrow: residues common for Zn catalytic activity and substrate specificity; triangle: residues responsible for substrate specificity; square: key residues for substrate binding; asterisk: a single, fully conserved residue; colon: conservation between groups of strongly similar properties; period: conservation between groups of weakly similar properties; yellow mark: key residues for primary lignin synthesis.



**Figure 7.** A phylogenetic relationship of AsCADs with CADs from other plant species. The phylogeny analysis was done on a set of 51 CAD proteins, analyzed using PhyML 3.0 (software based on the maximum-likelihood principle) and visualized using TreeDyn software.

## 2.6. Cloning and Heterologous Expression of AsCADs

To elucidate biochemical functions of AsCADs, we attempted to clone the candidates (AsCAD, AsCAD3, AsCAD5, and AsCAD7) from wild type *A. sieboldii*. The cDNAs were successfully cloned with open reading frame (ORF) of 1071, 1077, 1068, and 1068 bp respectively. Firstly, the ORFs of them were constructed to His-tagged vector (pDEST 17). Unfortunately, all of recombinant AsCAD proteins exhibited a low-level expression. Thus, MBP-tagged vector (pMAL-C5X) was used for recombinant AsCAD proteins production. They were successfully expressed as soluble protein in Rosetta (DE3) cells and purified with Dextrin Beads 6FF (MBP Affinity chromatography). Before enzyme activities were tested, purified proteins were cleaved by Factor Xa protease to remove MBP tag. All of the four CAD proteins exhibited similar molecular masses of ca. 40 kDa on SDS-PAGE (Figure 8).



**Figure 8.** Expression and purification of AsCAD5 recombinant proteins in *E. coli* Rosetta (DE3) harboring pMAL-C5X-AsCAD. M, marker; 1, purified proteins after removal of MBP tag; 2, soluble proteins purified in MBP binding buffer with maltose elution; 3, total proteins induced by IPTG; 4, purified proteins dialyzed in desalting columns.

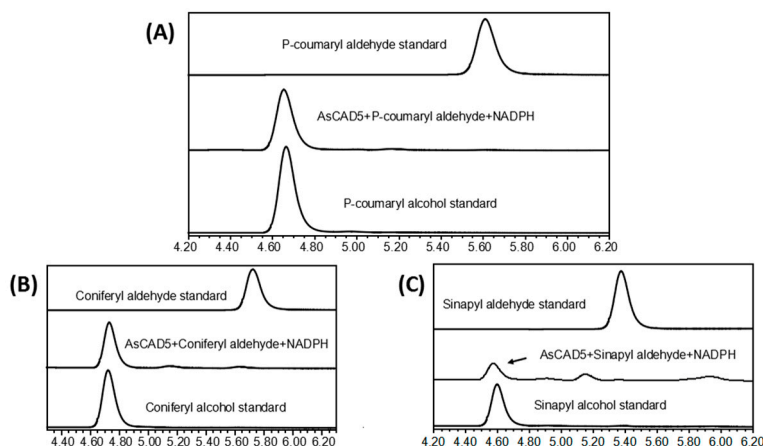
## 2.7. CAD Activity and Kinetic Parameters of the Recombinant AsCADs

Three substrates, coniferyl aldehyde, sinapyl aldehyde, and *p*-coumaryl aldehyde, were used to evaluate the enzymatic activity. The enzyme assays revealed that only the recombinant protein coded by AsCAD5 had efficient catalytic activity in converting the substrates to corresponding alcohol products (Figure 9A–C). In addition, kinetic analysis was carried out to characterize  $K_m$ ,  $V_{max}$ ,  $K_{cat}$ , and  $K_{cat}/K_m$  values for the three substrates. The microplate reader was used to measure three products' concentrations considering that they had maximum absorbance values at 340 nm. The resulting data showed that the  $K_m$  value of coniferyl aldehyde was lower than those of sinapyl aldehyde and *p*-coumaryl aldehyde, which supported a catalytic preference for coniferyl aldehyde. The  $K_{cat}/K_m$  values for sinapyl aldehyde and coniferyl aldehyde were similar and higher than for *p*-coumaryl aldehyde (Table 2).

**Table 2.** Kinetic parameters of the purified recombinant AsCAD5 with three substrates

Enzyme Kinetic Parameters	Coniferyl Aldehyde	Sinapyl Aldehyde	<i>P</i> -coumaryl Aldehyde
$K_m$ ( $\mu\text{M}$ )	$27.32 \pm 6.186$	$35.02 \pm 8.043$	$45.52 \pm 8.888$
$V_{max}$ ( $\text{nmol s}^{-1} \text{mg protein}^{-1}$ )	$342 \pm 26.29$	$448.5 \pm 50.54$	$383.8 \pm 31.63$
$K_{cat}$ ( $\text{s}^{-1}$ )	221.7	290.8	248.8
$K_{cat}/K_m$ ( $\mu\text{M}^{-1} \text{s}^{-1}$ )	8.115	8.304	5.466





**Figure 9.** Characterization of enzymatic activity with HPLC analysis and kinetic features: (A) profiles of HPLC show conversion of *p*-coumaryl aldehyde to *p*-coumaryl alcohol; (B) HPLC profiles show conversion of coniferyl aldehyde to coniferyl alcohol; (C) profiles of HPLC show conversion of sinapyl aldehyde to sinapyl alcohol.

### 3. Discussion

Asarinin, a tetrahydrofurofurano lignan, is the quality evaluation standard of *A. sieboldii* in Chinese Pharmacopoeia; and sesamin is the epimer of asarinin. Both have been regarded as the main active ingredients from water-extract of *A. sieboldii* roots. Until now, sesamin biosynthesis pathway has been elucidated in *Sesamum indicum*, in which a cytochrome P450, SiCYP81Q1, has been reported to catalyze sesamin biosynthesis by forming two methylenedioxy bridges [13,14]. Although CYP81 in *A. sieboldii* was not found, 11 unigenes of CYP719A subfamily have been found in our transcriptomics data (Figure 5, Figure S3). It has been reported that CYP719A23 from *Podophyllum hexandrum* and CYP719A24 from *Podophyllum peltatum* were capable of converting (–)-matairesinol into (–)-pluviatolide by catalyzing methylenedioxy bridge formation in the biosynthesis of podophyllotoxin, an aryltetralin lignan extensively used for treating cancer [48]. Likewise, the CYP719A subfamily has also been reported in alkaloid biosynthesis in catalyzing the formation of a methylenedioxy bridge. In cases of alkaloid biosynthesis—CYP719A1, CYP719A2, CYP719A3, CYP719A5, CYP719A9, CYP719A13, and CYP719A14—have been reported to be involved in methylenedioxy bridge-formation [41,49,50]. Thus, a proposed pathway was speculated in Figure 3, where CYP719s might participate in the synthesis of sesamin or asarinin by forming two methylenedioxy bridges. Indeed, five DIR unigenes were found in *A. sieboldii*, suggesting two molecules of coniferyl alcohols could be used for formation of pinosresinol. Taking together, the biosynthesis pathways of asarinin in *A. sieboldii* were deduced in the light of transcriptome sequencing data. One possible pathway of asarinin is that two molecules of conifer alcohols were catalyzed by DIRs and CYPs to form sesamin, then transformed to asarinin by epimerization [51]. The second possible pathway of asarinin biosynthesis is that two molecules of conifer alcohols were catalyzed by DIRs to form pinosresinol epimer, then directly transformed to asarinin by CYP719As (Figure 5).

Methyleugenol, eugenol, safrole, and other eugenol derivatives, the natural constituents of many plant essential oils, exist not only in *A. sieboldii* (Aristolochiaceae) but also in plants of other families such as Piperaceae, Labiatae, Myristicaceae, and Lauraceae (trees) [52–62]. Safrole and eugenol are known carcinogens in animals and methyleugenol is a suspected carcinogen [63]. For this harmful effects or uncertain safety of these mentioned chemicals, the utilization and content control of them in food and everyday articles have obtained more and more attention. A decoction procedure, similar to that traditionally used for Chinese herbal preparations, is able to effectively reduce the amount of safrole and methyleugenol in *Asiasari radix* [64]. In our study, two unigenes potentially encoding eugenol synthase (EGS) have been found in the transcriptomes of five tissues.

Furthermore, safrole (C<sub>10</sub>H<sub>10</sub>O<sub>2</sub>) could be speculated to be formed from eugenol (C<sub>10</sub>H<sub>12</sub>O<sub>2</sub>) through the formation of the methylenedioxy bridge by CYP719As. As a result, a proposed pathway of safrole biosynthesis was shown in Figure 5. If the EGS, EOMT, and CYP719A functions would be elucidated, reduction of eugenol, safrole, and methyleugenol could be achieved by knockdown or knockout of the corresponding genes.

From pathway shown in Figure 5 we can see that asarinin, sesamin, methyleugenol, eugenol, and safrole share a common precursor coniferyl alcohol, which is generated from coniferyl aldehyde under the catalysis of CAD proteins. The amino acid sequence alignment showed that two key residues for substrate binding (square), two residues for substrate specificity (triangle), and seven key residues for primary lignin synthesis in Figure 6 (yellow background) and Table 3 (circle) were totally different between AsCAD5, AtCAD5, and AaCAD, among which the latter two CADs have been demonstrated for bona fide cinnamyl alcohol dehydrogenase [37,65]. The amino acids discrepancy of AsCAD5 with AtCAD5 and AaCAD might be one of the reasons of clustering, where AsCAD5 is close to AtCAD1, and belongs to Class III but distant from Class I (e.g., AtCAD4 and AtCAD5) (Figure 7). Although AsCAD5 amino acids and clustering differs from AtCAD5, AsCAD5 exhibited the bona fide cinnamyl alcohol dehydrogenase activity similarly with AtCAD5. Furthermore, AsCAD5 displayed its substrate preference in the order coniferaldehyde > sinapaldehyde > *p*-coumaraldehyde, with a similar catalytic efficiency of coniferaldehyde and sinapaldehyde, but higher than that of *p*-coumaraldehyde, indicating AsCAD5 preferred coniferaldehyde as substrate to form cinnamyl alcohol. qRT-PCR showed AsCAD5 remained highly expressed in active compounds-rich roots in an ascending trend during the engaged developmental stages, namely, pre-anthesis, anthesis, and post-anthesis; but in another main medicinal tissue, rhizome, the expression level of AsCAD5 appeared highest at anthesis and then sharply decreased at post-anthesis. Considering the contents of either volatile oil or methyleugenol and asarinin reached the highest level at anthesis, it was suggested that AsCAD5 probably played an important role in biosynthesis of methyleugenol and asarinin.

**Table 3.** Key amino acids diversities among AsCAD5 (*Asarum sieboldii*), AaCAD (*Artemisia annua*), and AtCAD5 (*Arabidopsis thaliana*)

	▲ 48	● 75	● 86	▲ 94	■ 118	■ 122	● 275	● 276	● 285	● 299	● 300
AaCAD	T	V	V	C	W	D	V	I	P	F	I
AtCAD5	T	V	V	C	W	D	V	I	P	F	I
AsCAD5	A	Q	I	T	119T	123S	276F	277P	G	V	T

Triangle: residues responsible for substrate specificity; Square: key residues for substrate binding; Circle: key residues for primary lignin synthesis.

In conclusion, our transcriptomics data from roots, rhizomes, and leaves revealed candidate genes in biosynthesis of methyleugenol, asarinin, and safrole. Moreover, a cinnamyl alcohol dehydrogenase AsCAD5 was demonstrated to be responsible for the formation of coniferyl alcohol, the common precursor of methyleugenol and G-ligin.

## 4. Materials and Methods

### 4.1. Plant Materials

Three-year-old seedlings of *A. sieboldii* were collected from its cultivation base in Ningqiang County, Shaanxi Province, China. cDNA libraries were prepared from different tissues at different developmental stages: root in May (Sample 1), rhizome in May (Sample 2), rhizome in March (Sample 3), rhizome in July (Sample 4), and leaf (leaf blade and petiole, used as a whole or separately) in May (Sample 5). All the fresh samples were frozen at −80 °C and used to extract total RNA.

#### 4.2. Illumina Sequencing for Transcriptome Analysis

Total RNA of each sample was extracted according to manufacturer's protocol of RNeasy® Plant Mini Kit (Tiangen, Beijing, China), then quantified and qualified by Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA, USA), NanoDrop (Thermo Fisher Scientific Inc., Waltham, MA, USA) and 1% agarose gel. 1 µg of total RNA with RIN value above 7 was used for sequencing on the Illumina HiSeq™ 4000 sequencing platform. All raw transcriptome data were deposited in the GenBank Short Read Archive. The accession numbers were SRR6838612, SRR6838611, SRR6838610, SRR6838609, and SRR6838613 for *Asarum-1* to *Asarum-5*, respectively.

#### 4.3. De Novo Assembly and Functional Annotation Analysis

In order to remove technical sequences, pass filter data of Fasta format were processed by Cutadapt version 1.9.1 (Dortmund, Germany) to be high quality clean data. With paired-end reads, contigs can be identified from the same transcript and the distance between these contigs can be estimated. Then they were assembled by Trinity [66] to get the unigenes sequence file. The generated unigenes were annotated based on the Nr, Swiss-Prot protein, COG, and KEGG public databases. GO-Term Finder was used to annotate a list of enriched genes with a significant p-value less than 0.05. To elucidate the active biochemical pathways in *A. sieboldii*, unigenes were compared against the KEGG using BLASTx with an e-value  $< 1 \times 10^{-10}$  and the corresponding pathways were established.

#### 4.4. Quantitative Real-Time PCR Validation

To investigate expression profiles of AsCADs in different tissues at three developmental stages (collected in March, May and July, representing pre-anthesis, anthesis and post-anthesis accordingly), total RNA was extracted from leaf blades, rhizomes, roots, and petioles and the cDNA was reversely transcribed from 1 µg of total RNA using a FastQuantRT Kit (with gDNase) purchased from Tiangen. Quantitative real-time PCR was performed using 0.5 µL of cDNA product and each primer of 10 µL in a 20 µL reaction volume. The primers for qRT-PCR were designed based on the cDNA sequence (Table S3). 18S rRNA was used as reference gene.

#### 4.5. Identifying Genes Related to Biosynthesis of Methyleugenol and Asarinin in *A. sieboldii*

KEGG pathway annotation is helpful for predicting potential genes and their functions at a whole transcriptome level. By searching the gene name or the EC number, we screened the key enzymes coding sequences involved in the pathway and determined FPKM values for the transcripts. *A. sieboldii* CAD sequences were subjected to homology search using BLASTp against NR database. Heatmap of CYP450s was plotted using the OmicShare tools, a free online platform for data analysis ([www.omicshare.com/tools](http://www.omicshare.com/tools)).

#### 4.6. Multiple Sequence Alignments and Phylogenetic Analysis of AsCADs

Deduced protein sequences of AsCADs and functional CADs identified from other plant species were retrieved from the NCBI (<https://blast.ncbi.nlm.nih.gov>) database. These amino acid sequences were engaged in multiple alignment through Clustal Omega (<http://www.ebi.ac.uk/Tools/msa/clustalo>), following a phylogeny analysis. 51 CAD proteins were taken into the phylogeny analysis (Table S4), which was treated accordingly with the MUSCLE, PhyML (<http://phylogeny.lirmm.fr/>), and TreeDyn 198.3 (via [www.phylogeny.fr](http://www.phylogeny.fr)) softwares (Information Génomique et Structurale, Marseille, France) [67,68].

#### 4.7. Cloning of AsCADs

PCR was performed using PrimeSTAR Max DNA Polymerase (Takara, China), accompanying with a touch-down method: initial denaturation at 94 °C for 5 min, following 10 cycles of denaturation (94 °C, 1 min), annealing (1 min) with the temperature decreasing from 65 to 55 °C during these

10 cycles, and then extension at 72 °C (1 min); subsequently, 25 cycles (1 min each) of denaturation (94 °C), annealing (55 °C) and extension (72 °C); and finally extension at 72 °C for 7 min. The specific primers for AsCAD genes were designed according to the screened sequences and provided in Table S3. The amplification products were analyzed and the amplified fragments of interest in each case were further picked up and purified with the QIAquick gel extraction kit (Qiagen, Hilden, Germany). Then the purified fragments were ligated into pLB cloning vector (Tiangen) and transformed into *E. coli* DH5 $\alpha$  cells. The sequence alignments for each gene were performed using DNAMAN (Lynnon Biosoft Company, San Ramon, CA, USA).

#### 4.8. Heterogeneous Expression of Recombinant AsCADs

After confirmation of the cDNA sequence for each CAD homologue, the ORF for each gene was individually cloned in pMAL-C5X between restriction sites of *Nde*I and *Eco*RI using In-Fusion Cloning method. The primers were listed in Table S3. The positive constructs were then transformed into the competent Rosetta (DE3) cells for expression. A single colony of each transformed *E. coli* cell line was first incubated at 37 °C in LB medium containing carbenicillin (25 mg/L) and chloramphenicol (17 mg/L). When OD<sub>600</sub> reached 0.4–0.6, 0.1 mM isopropyl- $\beta$ -D-1-thiogalactopyranoside (IPTG) was added into the medium at 16 °C to induce the protein. The cells were allowed to grow for 20 h and finally harvested by centrifugation at 11,500 $\times$  *g* for 15 min, with the pellets frozen and stored at –80 °C. Frozen cell pellets for each CAD homologue were individually resuspended in binding buffer (20 mL, pH 7.4) containing Tris-HCl (20 mM), NaCl (200 mM), EDTA (1 mM), and DTT (1 mM) and then lysed by sonication for 10 min on ice (3 s pulse at 30W). Subsequently the cell debris for each CAD homologue was removed by centrifugation at 6000 $\times$  *g* for 20 min at 4 °C.

All purifications were individually performed at room temperature by using a Dextrin beads 6FF (Smart Lifesciences, Changzhou, China) column (25 mL), previously equilibrated in binding buffer. Samples of each of the enzyme preparations were then applied to the affinity column, in which the target proteins would be held by the tags. The unbound proteins would be removed through washing with binding buffer. Target proteins were individually eluted with elution buffer (20 mM maltose in binding buffer), resulting in 1-mL fraction collection of individual sample. For each CAD homologue, aliquots (15  $\mu$ L) of every other fraction were subjected to SDS-PAGE with gradient (5% acrylamide) gels; proteins were visualized by eStain L1 Protein Staining System (Genscript, Piscataway, NJ, USA). Fractions containing each MBP-tagged CAD (molecular mass ca. 80 kDa) were combined and dialyzed in PD-10 Desalting Columns (GE Healthcare, Uppsala, Sweden). Considering the molecular mass of MBP tag was similar with AsCADs, proteins after removal of MBP tag were used for enzyme assay by Factor Xa protease (New England Biolabs, Hertfordshire, UK).

#### 4.9. CAD Activity Assay and Determination of Kinetic Parameters

Three lignin substrates—coniferyl aldehyde, sinapyl aldehyde, and *p*-coumaryl aldehyde (Sigma-Aldrich, St. Louis, MO, USA)—were used to examine the catalytic activity of the recombinant AsCADs. The reactions were initiated by adding 2  $\mu$ g of the purified protein to a 300  $\mu$ L reaction mixture containing 0.1 mM substrate, 50 mM Tris-HCl (pH 7.5), 2 mM DTT, and 0.5 mM NADPH as cofactor. The reactions were allowed to proceed for 30 min at 30 °C, stopped by adding 15  $\mu$ L glacial acetic acid. After mixed with 300  $\mu$ L methanol and then centrifuged at 6000 $\times$  *g* for 10 min, the resulting supernatant was performed high performance liquid chromatography (HPLC) analysis with Waters 2695 equipped with sharpsil-U C18 reverse column (250  $\times$  4.6 mm, 5  $\mu$ m). Methanol-water (60:40, *v/v*) was employed as the mobile system, at the flow rate of 0.8 mL/min. The detection wavelengths were 260 nm and 340 nm, and the injected volume was 5  $\mu$ L. The AsCADs were further characterized the kinetic parameters with spectrophotometric method, employing each reaction in total 200  $\mu$ L of mixture in 96-well microplates and possessions of enzymatic reaction data through Epoch Microplate reader (BioTek instruments Inc., Winooski, VT, USA). For each substrate a serial of concentrations was subjected to test, in which all other components of the reaction mixture were kept the same as

described above. Extinction coefficients for both aldehyde and NADPH were utilized to calculate the relative contribution of each to the 340-nm signal [69]. The resulting data were analyzed using GraphPad Prism 7 software (San Diego, CA, USA) to determine the  $K_m$  and  $V_{max}$  values.

**Supplementary Materials:** Supplementary materials are available online, Figure S1: Gene ontology classification of assembled unigenes; Figure S2: COG function classification of *Asarum sieboldii* Miq.; Figure S3: A heatmap of all CYP450s from five samples: root-May, rhizome-May, rhizome-March, rhizome-July and leaf-May; Table S1: Gene ontology classification; Table S2: Mapping of *Asarum sieboldii* Miq. unique sequences to KEGG biochemical pathways; Table S3: Primers used in this study; Table S4: The accession number(s) of 51 CAD proteins for the phylogeny analysis.

**Author Contributions:** Z.L. and D.M. conceived and designed the research; J.L., C.X., H.Z., and F.L. performed the experiments and analyzed the data; Z.L., D.M., and J.L. wrote the manuscript. D.M. wrote the discussion. All authors read and approved the manuscript.

**Funding:** This work was supported by National Natural Science Foundation of China (NSFC) projects 31570325 and 81373962, key platform construction project of Department of Education of Guangdong Province (grant no. 2014KTSP016) and High-Level University of Guangdong Province (grant no. A1-AFD018171Z0302).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Lawrence, M.K. Phylogenetic relationships in *Asarum* (Aristolochiaceae) based on morphology and ITS sequences. *Am. J. Bot.* **1998**, *85*, 1454–1467.
2. Michl, J.; Bello, O.; Kite, G.C.; Simmonds, M.S.J.; Heinrich, M. Medicinally used *Asarum* species: High-resolution LC-MS analysis of aristolochic acid analogs and in vitro toxicity screening in HK-2 cells. *Front. Pharmacol.* **2017**, *8*, 215. [[CrossRef](#)] [[PubMed](#)]
3. Chinese Pharmacopoeia Commission. *Chinese Pharmacopoeia Part I*, 2015 ed.; Chemical Industry Press: Beijing, China, 2015; pp. 230–231.
4. Kopyt'ko, Y.F.; Shchurevich, N.N.; Sokol'skaya, T.A.; Markaryan, A.A.; Dargaeva, T.D. Uses, chemical composition, and standardization of plant raw material and medicinal substances from plants of the genus *Asarum* L. *Pharm. Chem. J.* **2013**, *47*, 157–168. [[CrossRef](#)]
5. Li, C.; Xu, F.; Cao, C.; Shang, M.-Y.; Zhang, C.-Y.; Yu, J.; Liu, G.-X.; Wang, X.; Cai, S.-Q. Comparative analysis of two species of *Asari Radix et Rhizoma* by electronic nose, headspace GC-MS and chemometrics. *J. Pharm. Biomed. Anal.* **2013**, *85*, 231–238. [[CrossRef](#)] [[PubMed](#)]
6. Wu, H.; Li, J.; Zhang, F.; Li, L.; Liu, Z.; He, Z. Essential oil components from *Asarum sieboldii* Miquel are toxic to the house dust mite *Dermatophagoides farinae*. *Parasitol. Res.* **2012**, *111*, 1895–1899. [[CrossRef](#)]
7. Han, Y.; Kim, S.J. Memory enhancing actions of *Asiasari radix* extracts via activation of insulin receptor and extracellular signal regulated kinase (ERK) I/II in rat hippocampus. *Brain Res.* **2003**, *974*, 193–201. [[CrossRef](#)]
8. Jang, J.Y.; Lee, J.H.; Shin, H.K.; Choi, Y.H.; Lee, J.D.; Choi, B.T. Partially purified *Asiasari radix* inhibits melanogenesis through extracellular signal-regulated kinase signaling in B16F10 cells. *Int. J. Mol. Med.* **2010**, *25*, 287–292. [[CrossRef](#)]
9. Jeong, S.H.; Lee, J.E.; Jin, S.H.; Ko, Y.; Park, J.B. Effect of *Asiasari radix* on osteoblastic differentiation of stem cells derived from gingiva. *J. Tradit. Chin. Med.* **2016**, *36*, 756–759. [[CrossRef](#)]
10. Kim, H.M.; Moon, Y.S. *Asiasari radix* inhibits immunoglobulin E production on experimental models in vitro and in vivo. *Immunopharmacol. Immunotoxicol.* **1999**, *21*, 469–481. [[CrossRef](#)]
11. Lee, J.A.; Lee, M.Y.; Seo, C.S.; Ha, H.; Lee, H.; Kim, J.H. *Asiasari sieboldii* suppresses inflammatory mediators through the induction of hemeoxygenase-1 expression in RAW264.7 cells. *Immunopharmacol. Immunotoxicol.* **2012**, *34*, 15–20. [[CrossRef](#)]
12. Oh, S.M.; Kim, J.; Lee, J.; Yi, J.M.; Oh, D.S.; Bang, O.S.; Kim, N.S. Anticancer potential of an ethanol extract of *Asiasari radix* against HCT-116 human colon cancer cells in vitro. *Oncol. Lett.* **2013**, *5*, 305–310. [[CrossRef](#)] [[PubMed](#)]
13. Ono, E.; Nakai, M.; Fukui, Y.; Tomimori, N.; Fukuchi-Mizutani, M.; Saito, M.; Satake, H.; Tanaka, T.; Katsuta, M.; Umezawa, T.; et al. Formation of two methylenedioxy bridges by a *Sesamum* CYP81Q protein yielding a furofuran lignan, (+)-sesamin. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 10116–10121. [[CrossRef](#)] [[PubMed](#)]



14. Pathak, N.; Bhaduri, A.; Bhat, K.V.; Rai, A.K. Tracking sesamin synthase gene expression through seed maturity in wild and cultivated sesame species—A domestication footprint. *BMC Plant Biol.* **2015**, *17*, 1039–1046. [[CrossRef](#)] [[PubMed](#)]
15. Davin, L.B.; Bedgar, D.L.; Katayama, T.; Lewis, N.G. On the stereoselective synthesis of (+)-pinoresinol in *Forsythia suspensa* from its achiral precursor, coniferyl alcohol. *Phytochemistry* **1992**, *31*, 3869–3874. [[CrossRef](#)] [[PubMed](#)]
16. Halls, S.C.; Davin, L.B.; Kramer, D.M.; Lewis, N.G. Kinetic study of coniferyl alcohol radical binding to the (+)-pinoresinol forming dirigent protein. *Biochemistry* **2004**, *43*, 2587–2595. [[CrossRef](#)] [[PubMed](#)]
17. Kazenwadel, C.; Klebensberger, J.; Richter, S.; Pfannstiel, J.; Gerken, U.; Pickel, B.; Schaller, A.; Hauer, B. Optimized expression of the dirigent protein AtDIR6 in *Pichia pastoris* and impact of glycosylation on protein structure and function. *Appl. Microbiol. Biotechnol.* **2013**, *97*, 7215–7227. [[CrossRef](#)] [[PubMed](#)]
18. Li, N.; Zhao, M.; Liu, T.; Dong, L.; Cheng, Q.; Wu, J.; Wang, L.; Chen, X.; Zhang, C.; Lu, W.; et al. A novel Soybean dirigent gene GmDIR22 contributes to promotion of lignan biosynthesis and enhances resistance to *Phytophthora sojae*. *Front. Plant Sci.* **2017**, *8*, 1185. [[CrossRef](#)] [[PubMed](#)]
19. Pickel, B.; Pfannstiel, J.; Steudle, A.; Lehmann, A.; Gerken, U.; Pleiss, J.; Schaller, A. A model of dirigent proteins derived from structural and functional similarities with allene oxide cyclase and lipocalins. *FEBS J.* **2012**, *279*, 1980–1993. [[CrossRef](#)]
20. Jin, J.S.; Hattori, M. A new mammalian lignan precursor, asarinin. *Food Chem.* **2011**, *124*, 895–899. [[CrossRef](#)]
21. Kushiro, M.; Masaoka, T.; Hageshita, S.; Takahashi, Y.; Ide, T.; Sugano, M. Comparative effect of sesamin and episesamin on the activity and gene expression of enzymes in fatty acid oxidation and synthesis in rat liver. *J. Nutr. Biochem.* **2002**, *13*, 289–295. [[CrossRef](#)]
22. Gu, J.; Zhang, L.; Wang, Z.; Chen, Y.; Zhang, G.; Zhang, D.; Wang, X.; Bai, X.; Li, X.; Lili, Z. The effect of asarinin on toll-like pathway in rats after cardiac allograft implantation. *Transplant. Proc.* **2015**, *47*, 545–548. [[CrossRef](#)]
23. Jeong, M.; Kim, H.M.; Lee, J.S.; Choi, J.H.; Jang, D.S. (–)-Asarinin from the roots of *Asarum sieboldii* induces apoptotic cell death via caspase activation in human ovarian cancer cells. *Molecules* **2018**, *23*, 1849. [[CrossRef](#)]
24. Kosuge, T.; Yokota, M.; Nagasawa, M.; Nukaya, H.; Gotoh, Y. Studies on antitussive principles of *Asiasari Radix*. *Chem. Pharm. Bull.* **1978**, *26*, 2284–2285. [[CrossRef](#)]
25. Yano, S.; Suzuki, Y.; Yuzurihara, M.; Kase, Y.; Takeda, S.; Watanabe, S.; Aburada, M.; Miyamoto, K. Antinociceptive effect of methyleugenol on formalin-induced hyperalgesia in mice. *Eur. J. Pharmacol.* **2006**, *553*, 99–103. [[CrossRef](#)] [[PubMed](#)]
26. Zhu, T.; Yan, Y.; Yan, Y.; Deng, S.; Liu, Y.M.; Fan, H.R.; Ma, B.; Meng, B.; Mei, B.; Li, W.G.; Li, F. Methyleugenol counteracts anorexigenic signals in association with GABAergic inhibition in the central amygdala. *Neuropharmacology* **2018**, *141*, 331–342. [[CrossRef](#)] [[PubMed](#)]
27. Yauk, Y.K.; Souleyre, E.J.F.; Matich, A.J.; Chen, X.; Wang, M.Y.; Plunkett, B.; Dare, A.P.; Espley, R.V.; Tomes, S.; Chagné, D.; et al. Alcohol acyl transferase 1 links two distinct volatile pathways that produce esters and phenylpropenes in apple fruit. *Plant J.* **2017**, *91*, 292–305. [[CrossRef](#)] [[PubMed](#)]
28. Anand, A.; Jayaramaiah, R.H.; Beedkar, S.D.; Singh, P.A.; Joshi, R.S.; Mulani, F.A.; Dholakia, B.B.; Punekar, S.A.; Gade, W.N.; Thulasiram, H.V.; Giri, A.P. Comparative functional characterization of eugenol synthase from four different *Ocimum* species: Implications on eugenol accumulation. *Biochim. Biophys. Acta* **2016**, *1864*, 1539–1547. [[CrossRef](#)]
29. Gupta, A.K.; Schauvinhold, I.; Pichersky, E.; Schiestl, F.P. Eugenol synthase genes in floral scent variation in *Gymnadenia* species. *Funct. Integr. Genom.* **2014**, *14*, 779–788. [[CrossRef](#)]
30. Yan, H.; Baudino, S.; Caissard, J.C.; Florence, N.; Zhang, H.; Tang, K.; Li, S.; Lu, S. Functional characterization of the eugenol synthase gene (RcEGS1) in rose. *Plant Physiol. Biochem.* **2018**, *129*, 21–26. [[CrossRef](#)]
31. Koeduka, T.; Kajiyama, M.; Furuta, T.; Suzuki, H.; Tsuge, T.; Matsui, K. Characterization of an O-methyltransferase specific to guaiacol-type benzenoids from the flowers of loquat (*Eriobotrya japonica*). *J. Biosci. Bioeng.* **2016**, *122*, 679–684. [[CrossRef](#)]
32. Koeduka, T.; Fridman, E.; Gang, D.R.; Vassao, D.G.; Jackson, B.L.; Kish, C.M.; Orlova, I.; Spassova, S.M.; Lewis, N.G.; Noel, J.P.; et al. Eugenol and isoeugenol, characteristic aromatic constituents of spices, are biosynthesized via reduction of a coniferyl alcohol ester. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 10128–10133. [[CrossRef](#)] [[PubMed](#)]

33. Renu, I.K.; Haque, I.; Kumar, M.; Poddar, R.; Bandopadhyay, R.; Rai, A.; Mukhopadhyay, K. Characterization and functional analysis of eugenol *O*-methyltransferase gene reveal metabolite shifts, chemotype specific differential expression and developmental regulation in *Ocimum tenuiflorum* L. *Mol. Biol. Rep.* **2014**, *41*, 1857–1870. [[CrossRef](#)] [[PubMed](#)]
34. Wang, J.; Pichersky, E. Characterization of S-adenosyl-L-methionine: (Iso)eugenol *O*-methyltransferase involved in floral scent production in *Clarkia breweri*. *Arch. Biochem. Biophys.* **1998**, *349*, 153–160. [[CrossRef](#)] [[PubMed](#)]
35. Yauk, Y.K.; Chagne, D.; Tomes, S.; Matich, A.J.; Wang, M.Y.; Chen, X.; Maddumage, R.; Hunt, M.B.; Rowan, D.D.; Atkinson, R.G. The *O*-methyltransferase gene MdoOMT1 is required for biosynthesis of methylated phenylpropenes in ripe apple fruit. *Plant J.* **2015**, *82*, 937–950. [[CrossRef](#)] [[PubMed](#)]
36. Boudet, A.M.; Lapiere, C.; Grima-Pettenati, J. Biochemistry and molecular biology of lignification. *New Phytol.* **1995**, *129*, 203–236. [[CrossRef](#)]
37. Kim, S.J.; Kim, M.R.; Bedgar, D.L.; Moinuddin, S.G.; Cardenas, C.L.; Davin, L.B.; Kang, C.; Lewis, N.G. Functional reclassification of the putative cinnamyl alcohol dehydrogenase multigene family in *Arabidopsis*. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 1455–1460. [[CrossRef](#)] [[PubMed](#)]
38. Dorairaj, D.; Ismail, M.R. Distribution of silicified microstructures, regulation of cinnamyl alcohol dehydrogenase and lodging resistance in silicon and paclobutrazol mediated *Oryza sativa*. *Front. Physiol.* **2017**, *8*, 491. [[CrossRef](#)]
39. Geissler, M.; Burghard, M.; Volk, J.; Staniek, A.; Warzecha, H. A novel cinnamyl alcohol dehydrogenase (CAD)-like reductase contributes to the structural diversity of monoterpenoid indole alkaloids in *Rauvolfia*. *Planta* **2016**, *243*, 813–824. [[CrossRef](#)]
40. Ozparpucu, M.; Gierlinger, N.; Burgert, I.; Van Acker, R.; Vanholme, R.; Boerjan, W.; Pilate, G.; Déjardin, A.; Rüggeberg, M. The effect of altered lignin composition on mechanical properties of cinnamyl alcohol dehydrogenase (CAD) deficient poplars. *Planta* **2018**, *247*, 887–897. [[CrossRef](#)]
41. Ikezawa, N.; Iwasa, K.; Sato, F. CYP719A subfamily of cytochrome P450 oxygenases and isoquinoline alkaloid biosynthesis in *Eschscholzia californica*. *Plant Cell Rep.* **2009**, *28*, 123–133. [[CrossRef](#)]
42. Saathoff, A.J.; Hargrove, M.S.; Haas, E.J.; Tobias, C.M.; Twigg, P.; Sattler, S.; Sarath, G. Switchgrass PviCAD1: Understanding residues important for substrate preferences and activity. *Appl. Biochem. Biotechnol.* **2012**, *168*, 1086–1100. [[CrossRef](#)]
43. Barakat, A.; Bagniewska-Zadworna, A.; Choi, A.; Plakkat, U.; DiLoreto, D.S.; Yellanki, P.; Carlson, J.E. The cinnamyl alcohol dehydrogenase gene family in *Populus*: Phylogeny, organization, and expression. *BMC Plant Biol.* **2009**, *9*, 26:1–26:15. [[CrossRef](#)]
44. Guo, D.M.; Ran, J.H.; Wang, X.Q. Evolution of the cinnamyl/sinapyl alcohol dehydrogenase (CAD/SAD) gene family: The emergence of real lignin is associated with the origin of bona fide CAD. *J. Mol. Evol.* **2010**, *71*, 202–218. [[CrossRef](#)]
45. Jin, Y.; Zhang, C.; Liu, W.; Qi, H.; Chen, H.; Cao, S. The cinnamyl alcohol dehydrogenase gene family in melon (*Cucumis melo* L.): Bioinformatic analysis and expression patterns. *PLoS ONE* **2014**, *9*, 101730. [[CrossRef](#)]
46. Rong, W.; Luo, M.; Shan, T.; Wei, X.; Du, L.; Xu, H.; Zhang, Z. A wheat cinnamyl alcohol dehydrogenase TaCAD12 contributes to host resistance to the sharp eyespot disease. *Front. Plant Sci.* **2016**, *7*, 1723. [[CrossRef](#)]
47. Saballos, A.; Ejeta, G.; Sanchez, E.; Kang, C.; Vermerris, W. A genomewide analysis of the cinnamyl alcohol dehydrogenase family in sorghum [*Sorghum bicolor* (L.) Moench] identifies *SbCAD2* as the *Brown midrib6* gene. *Genetics* **2009**, *181*, 783–795. [[CrossRef](#)]
48. Marques, J.V.; Kim, K.W.; Lee, C.; Costa, M.A.; May, G.D.; Crow, J.A.; Davin, L.B.; Lewis, N.G. Next generation sequencing in predicting gene function in podophyllotoxin biosynthesis. *J. Biol. Chem.* **2013**, *288*, 466–479. [[CrossRef](#)]
49. Ikezawa, N.; Tanaka, M.; Nagayoshi, M.; Shinkyō, R.; Sakaki, T.; Inouye, K.; Sato, F. Molecular cloning and characterization of CYP719, a methylenedioxy bridge-forming enzyme that belongs to a novel P450 family, from cultured *Coptis japonica* cells. *J. Biol. Chem.* **2003**, *278*, 38557–38565. [[CrossRef](#)]
50. Diaz Chavez, M.L.; Rolf, M.; Gesell, A.; Kutchan, T.M. Characterization of two methylenedioxy bridge-forming cytochrome P450-dependent enzymes of alkaloid formation in the Mexican prickly poppy *Argemone mexicana*. *Arch. Biochem. Biophys.* **2011**, *507*, 186–193. [[CrossRef](#)]

51. Li, C.Y.; Chow, T.J.; Wu, T.S. The epimerization of sesamin and asarinin. *J. Nat. Prod.* **2005**, *68*, 1622–1624. [[CrossRef](#)]
52. Bos, R.; Hendriks, H.; van Os, F.H. The composition of the essential oil in the leaves of *Coleus aromaticus* Benth and their importance as a component of the species antiaphthosae Ph. Ned. Ed. V. *Pharm. Weekbl. Sci.* **1983**, *5*, 129–130. [[CrossRef](#)]
53. Chokechajaroenporn, O.; Bunyapraphatsara, N.; Kongchuensin, S. Mosquito repellent activities of *Ocimum* volatile oils. *Phytomedicine* **1994**, *1*, 135–139. [[CrossRef](#)]
54. Da Silva, J.K.; Andrade, E.H.; Guimaraes, E.F.; Maia, J.G. Essential oil composition, antioxidant capacity and antifungal activity of *Piper divaricatum*. *Nat. Prod. Commun.* **2010**, *5*, 477–480.
55. De Vincenzi, M.; Silano, M.; Stacchini, P.; Scazzocchio, B. Constituents of aromatic plants: I. Methyleugenol. *Fitoterapia* **2000**, *71*, 216–221. [[CrossRef](#)] [[PubMed](#)]
56. Hashimoto, K.; Yanagisawa, T.; Okui, Y.; Ikeya, Y.; Maruno, M.; Fujita, T. Studies on anti-allergic components in the roots of *Asiasarum sieboldi*. *Planta Med.* **1994**, *60*, 124–127. [[CrossRef](#)]
57. Kamdem, D.P.; Gage, D.A. Chemical composition of essential oil from the root bark of *Sassafras albidum*. *Planta Med.* **1995**, *61*, 574–575. [[CrossRef](#)]
58. Liu, X.; Zhao, M.; Luo, W.; Yang, B.; Jiang, Y. Identification of volatile components in *Phyllanthus emblica* L. and their antimicrobial activity. *J. Med. Food* **2009**, *12*, 423–428. [[CrossRef](#)]
59. Medina, A.L.; Lucero, M.E.; Holguin, F.O.; Estell, R.E.; Posakony, J.J.; Simon, J.; O’Connell, M.A. Composition and antimicrobial activity of *Anemopsis californica* leaf oil. *J. Agric. Food Chem.* **2005**, *53*, 8694–8698. [[CrossRef](#)]
60. Miele, M.; Ledda, B.; Falugi, C.; Mazzei, M. Methyleugenol and eugenol variation in *Ocimum basilicum* cv. Genovese gigante grown in greenhouse and in vitro. *Boll. Soc. Ital. Biol. Sper.* **2001**, *77*, 43–50.
61. Pino Benitez, N.; Melendez Leon, E.M.; Stashenko, E.E. Eugenol and methyl eugenol chemotypes of essential oil of species *Ocimum gratissimum* L. and *Ocimum campechianum* Mill. from Colombia. *J. Chromatogr. Sci.* **2009**, *47*, 800–803. [[CrossRef](#)] [[PubMed](#)]
62. Sayyah, M.; Valizadeh, J.; Kamalinejad, M. Anticonvulsant activity of the leaf essential oil of *Laurus nobilis* against pentylenetetrazole- and maximal electroshock-induced seizures. *Phytomedicine* **2002**, *9*, 212–216. [[CrossRef](#)]
63. Schiestl, R.H.; Chan, W.S.; Gietz, R.D.; Mehta, R.D.; Hastings, P.J. Safrole, eugenol and methyleugenol induce intrachromosomal recombination in yeast. *Mutat. Res.* **1989**, *224*, 427–436. [[CrossRef](#)]
64. Chen, C.; Spriano, D.; Lehmann, T.; Meier, B. Reduction of safrole and methyleugenol in Asari radix et rhizoma by decoction. *Forsch. Komplementmed.* **2009**, *16*, 162–166. [[CrossRef](#)]
65. Li, X.; Ma, D.; Chen, J.; Pu, G.; Ji, Y.; Lei, C.; Du, Z.; Liu, B.; Ye, H.; Wang, H. Biochemical characterization and identification of a cinnamyl alcohol dehydrogenase from *Artemisia annua*. *Plant Sci.* **2012**, *193–194*, 85–95. [[CrossRef](#)]
66. Grabherr, M.G.; Haas, B.J.; Yassour, M.; Levin, J.Z.; Thompson, D.A.; Amit, I.; Adiconis, X.; Fan, L.; Raychowdhury, R.; Zeng, Q.; et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **2011**, *29*, 644–652. [[CrossRef](#)]
67. Guindon, S.; Lethiec, F.; Duroux, P.; Gascuel, O. PhyML Online—A web server for fast maximum likelihood-based phylogenetic inference. *Nucleic Acids Res.* **2005**, *33*. [[CrossRef](#)]
68. Guindon, S.; Dufayard, J.F.; Lefort, V.; Anisimova, M.; Hordijk, W.; Gascuel, O. New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst. Biol.* **2010**, *59*, 307–321. [[CrossRef](#)]
69. Saathoff, A.J.; Sarath, G.; Chow, E.K.; Dien, B.S.; Tobias, C.M. Downregulation of cinnamyl-alcohol dehydrogenase in switchgrass by RNA silencing results in enhanced glucose release after cellulase treatment. *PLoS ONE* **2011**, *6*, 16416. [[CrossRef](#)]

**Sample Availability:** Not available.



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).