the plant journal

SEB SOCIETY FOR EXPERIMENTAL BIOLOGY

The Plant Journal (2019) doi: 10.1111/tpj.14506

Cinnamate-CoA ligase is involved in biosynthesis of benzoate-derived biphenyl phytoalexin in *Malus* × *domestica* 'Golden Delicious' cell cultures

Deepa Teotia[1,†], Mariam Gaid[2,†,‡], Shashank S. Saini[1], Aparna Verma[3], Ragothaman M. Yennamalli[4], Satyajeet P. Khare[5], Kiran Ambatipudi[3], Javid Iqbal Mir[6], Till Beuerle[2], Robert Hänsch[7], Partha Roy[8], Pawan Kumar Agrawal[9], Ludger Beerhues[2] and Debabrata Sircar[1,*]

[1]*Plant Molecular Biology Group, Biotechnology Department, Indian Institute of Technology Roorkee, Roorkee 247667, India,*
[2]*Institute of Pharmaceutical Biology, Technische Universität Braunschweig, Mendelssohnstrasse 1, D-38106 Braunschweig, Germany,*
[3]*Biotechnology Department, Indian Institute of Technology Roorkee, Roorkee 247667, India,*
[4]*VivagenDx Labs, Venkateshwara Nagar, Velachery, Chennai 600042, India,*
[5]*Symbiosis School of Biological Sciences, Symbiosis International, Lavale MH-412115, India,*
[6]*Central Institute of Temperate Horticulture (ICAR-CITH), Srinagar 190 005, Jammu and Kashmir, India,*
[7]*Institute of Plant Biology, Technische Universität Braunschweig, Humboldtstrasse 1, D-38106 Braunschweig, Germany,*
[8]*Molecular Endocrinology Group, Biotechnology Department, Indian Institute of Technology Roorkee, Roorkee 247667, India, and*
[9]*Odisha University of Agriculture and Technology, Bhubaneswar 751003, Odisha, India*

Received 22 June 2019; revised 31 July 2019; accepted 7 August 2019.
*For correspondence (e-mail debsrfbt@iitr.ac.in; dsircar.iitkgp@gmail.com).
†These authors contributed equally to this work.
‡Present address: Department of Bioorganic Chemistry, Leibniz Institute of Plant Biochemistry, Weinberg 3, 06120 Halle (Saale), Germany

SUMMARY

Apple (*Malus* sp.) and other genera belonging to the sub-tribe Malinae of the Rosaceae family produce unique benzoic acid-derived biphenyl phytoalexins. Cell cultures of *Malus domestica* cv. 'Golden Delicious' accumulate two biphenyl phytoalexins, aucuparin and noraucuparin, in response to the addition of a *Venturia inaequalis* elicitor (VIE). In this study, we isolated and expressed a cinnamate-CoA ligase (CNL)-encoding sequence from VIE-treated cell cultures of cv. 'Golden Delicious' (*M. domestica CNL*; *Md*CNL). *Md*CNL catalyses the conversion of cinnamic acid into cinnamoyl-CoA, which is subsequently converted to biphenyls. *Md*CNL failed to accept benzoic acid as a substrate. When scab-resistant (cv. 'Shireen') and moderately scab-susceptible (cv. 'Golden Delicious') apple cultivars were challenged with the *V. inaequalis* scab fungus, an increase in *Md*CNL transcript levels was observed in internodal regions. The increase in *Md*CNL transcript levels could conceivably correlate with the pattern of accumulation of biphenyls. The C-terminal signal in the *Md*CNL protein directed its N-terminal reporter fusion to peroxisomes in *Nicotiana benthamiana* leaves. Thus, this report records the cloning and characterisation of a cinnamoyl-CoA-forming enzyme from apple via a series of *in vivo* and *in vitro* studies. Defining the key step of phytoalexin formation in apple provides a biotechnological tool for engineering elite cultivars with improved resistance.

Keywords: Cinnamate-CoA ligase, phytoalexin biosynthesis, *Venturia inaequalis*, aucuparin, apple, malinae, *Malus* × *domestica*, Golden Delicious, Shireen.

INTRODUCTION

Apple (*Malus* × *domestica* Borkh.; Rosaceae) is an important deciduous fruit crop in temperate areas of world with an annual global production of approximately 8.93 million tons (FAOSTAT, http://www.fao.org/faostat/en). Apple fruits are known for their high nutritional value (Sarkate *et al.*, 2017). Despite its tremendous economic importance, apple production is vulnerable to a number of serious diseases. In apple, the Ascomycete fungus *Venturia inaequalis* is the causative agent of scab disease (Jha *et al.*, 2009), while fire blight disease, caused by the enterobacterium *Erwinia amylovora* (Thomson, 2000), is recognised as the most

destructive disease of apple; both infections trigger severe losses in apple production every year (Carisse and Bernier, 2002). It is noteworthy that most of the top-selling commercial apple cultivars are highly susceptible to scab and fire blight; however, some wild apple cultivars are resistant to these infections. Apple cultivars combat scab and fire-blight infections by producing chemically specialised defence metabolites in the form of biphenyls and dibenzofurans, among multiple other defence strategies (Chizzali and Beerhues, 2012). Biphenyls and dibenzofurans exhibit significant antifungal and antibacterial properties. Recently, it has been shown that when *V. inaequalis* conidia are treated with a mixture of aucuparin, noraucuparin and eriobofuran (5 μM each), a significant reduction (70%) in conidial germination is observed under *in vitro* conditions (Sarkate *et al.*, 2018). Similarly, both aucuparin and noraucuparin show significant inhibitory effects against the fire blight bacterium *E. amylovora* under *in vitro* conditions (Chizzali *et al.*, 2012a). However, their antifungal or antibacterial modes of action remain largely elusive (Chizzali *et al.*, 2012a).

Biosynthesis of the biphenyl skeleton is well studied at both biochemical and molecular levels (Liu *et al.*, 2004; Chizzali *et al.*, 2012b). Biphenyl formation is catalysed by a type III polyketide synthase, biphenyl synthase (BIS; EC 2.3.1.177). Decarboxylative aldol condensation is mediated by BIS, which requires benzoyl-CoA (one molecule) and malonyl-CoA (three molecules) to produce 3,5-dihydroxybiphenyl, the precursor of biphenyl phytoalexins (Saini *et al.*, 2017). The initial substrate for BIS, benzoyl-CoA, is derived from either cinnamoyl-CoA or free benzoic acid, both of which originate from cinnamic acid (Figure 1). Benzoate-CoA ligase (BZL) promotes the thioesterification of benzoic acid with free coenzyme A to form benzoyl-CoA (Figure 1). *In planta*, BZL activity has been biochemically identified in two plants other than *Malus*, namely *Clarkia breweri* and *Hypericum androsaemum* (Abd El-Mawla and Beerhues, 2002; Beuerle and Pichersky, 2002a). Although the core biphenyl biosynthetic pathway is well studied in the Malinae (Chizzali *et al.*, 2012a; Khalil *et al.*, 2015; Sircar *et al.*, 2015; Sarkate *et al.*, 2019), the formation of the precursor, benzoyl-CoA, is not well understood.

Plant benzoic acids arise from either shikimate pathway intermediates or the phenylpropanoid pathway. Examples of shikimate-derived benzoates are 3-hydroxybenzoic acid in *Centaurium erythraea* and *Swertia chirata*, salicylic acid in *Arabidopsis thaliana* and *Nicotiana benthamiana*, gallic acid in *Rhus typhina* and 2,3-dihydroxybenzoic acid in *Catharanthus roseus* (Ishikura *et al.*, 1984; Moreno *et al.*, 1994; Werner *et al.*, 1997; Wang *et al.*, 2001; Catinot *et al.*, 2008). Many plants synthesise isochorismate, which serves as the intermediate for the biosynthesis of *O*-succinylbenzoic acid, salicylic acid and 2,3-dihydroxybenzoic acid (Mustafa *et al.*, 2009; Bartsch *et al.*, 2010).

In phenylpropanoid metabolism, the formation of benzoic acid and its derivatives starts from either *trans*-cinnamic acid (CA) or its derivatives (hydroxy- and methoxy-derivatives). Cinnamic acid originates from L-phenylalanine through the activity of the enzyme phenylalanine ammonia-lyase (PAL). C_2 shortening of the propyl side chain of CA produces benzoic acid via one of the three major pathways. The first route involves a CoA-dependent and β-oxidative mechanism that mirrors fatty acid β-oxidation, which produces CoA thioesters as an intermediate. The CoA-dependent and β-oxidative pathways of benzoic acid biosynthesis have been detected in several plants such as *Cucumis sativus*, *A. thaliana*, *Nicotiana attenuata* and *Petunia hybrida* (Widhalm and Dudareva, 2015). The genes and corresponding enzymes of the β-oxidative and CoA-dependent pathways of benzoic acid biosynthesis have been determined and functionally defined. In this route, CA is first thioesterified into cinnamoyl-CoA via cinnamate-CoA ligase (CNL) (Colquhoun *et al.*, 2012; Gaid *et al.*, 2012; Klempien *et al.*, 2012). Cinnamoyl-CoA hydratase/dehydrogenase (CHD) (Qualley *et al.*, 2012; Bussell *et al.*, 2014) is a bi-functional enzyme that converts cinnamoyl-CoA to 3-oxo-3-phenylpropanoyl-CoA. Finally, benzoyl-CoA is the product of 3-ketoacyl thiolase (KAT)-catalysed conversion of 3-oxo-3-phenylpropanoyl-CoA (Van Moerkercke *et al.*, 2009).

The alternative C_2 side chain shortening route involves a CoA-independent and non-β-oxidative cleavage mechanism. This route for the formation of benzoic acid has been detected in cell cultures of *Sorbus aucuparia* (Gaid *et al.*, 2009) and *Pyrus pyrifolia* (Saini *et al.*, 2017). A key characteristic of this route is the oxidation of the intermediate benzaldehyde by a NAD^+-dependent dehydrogenase. Biosynthesis of methyl benzoate in *Antirrhinum majus* has been shown to start with the catalysis of a mitochondrial benzaldehyde dehydrogenase (Long *et al.*, 2009). A third route of C_2 side chain cleavage that involves a CoA-dependent, non-β-oxidative mechanism has been detected in only a few plant species. In this biosynthetic sequence, cinnamoyl-CoA undergoes C_2 side-chain cleavage to form benzaldehyde. Ibdah and Pichersky (2009) isolated the *CHY* gene encoding 3-hydroxyisobutyryl-CoA hydrolase, which converts cinnamoyl-CoA into benzaldehyde *in vivo*, from *A. thaliana*. A similar chain-shortening mechanism was also observed in *H. androsaemum* cell cultures (Abd El-Mawla and Beerhues, 2002) and in *Vanilla planifolia* pods (Gallage *et al.*, 2014).

In this study, we perform cDNA cloning, functional characterisation and subcellular localisation analysis of CNL from elicitor-treated cell cultures of the apple cv. 'Golden Delicious'. The carbon flux from the general C_6-C_3 phenylpropanoid pathway is directed by CNL towards the formation of the C_6-C_1 core, benzoic acid. Benzoic acid is then converted into biphenyl phytoalexin. The information

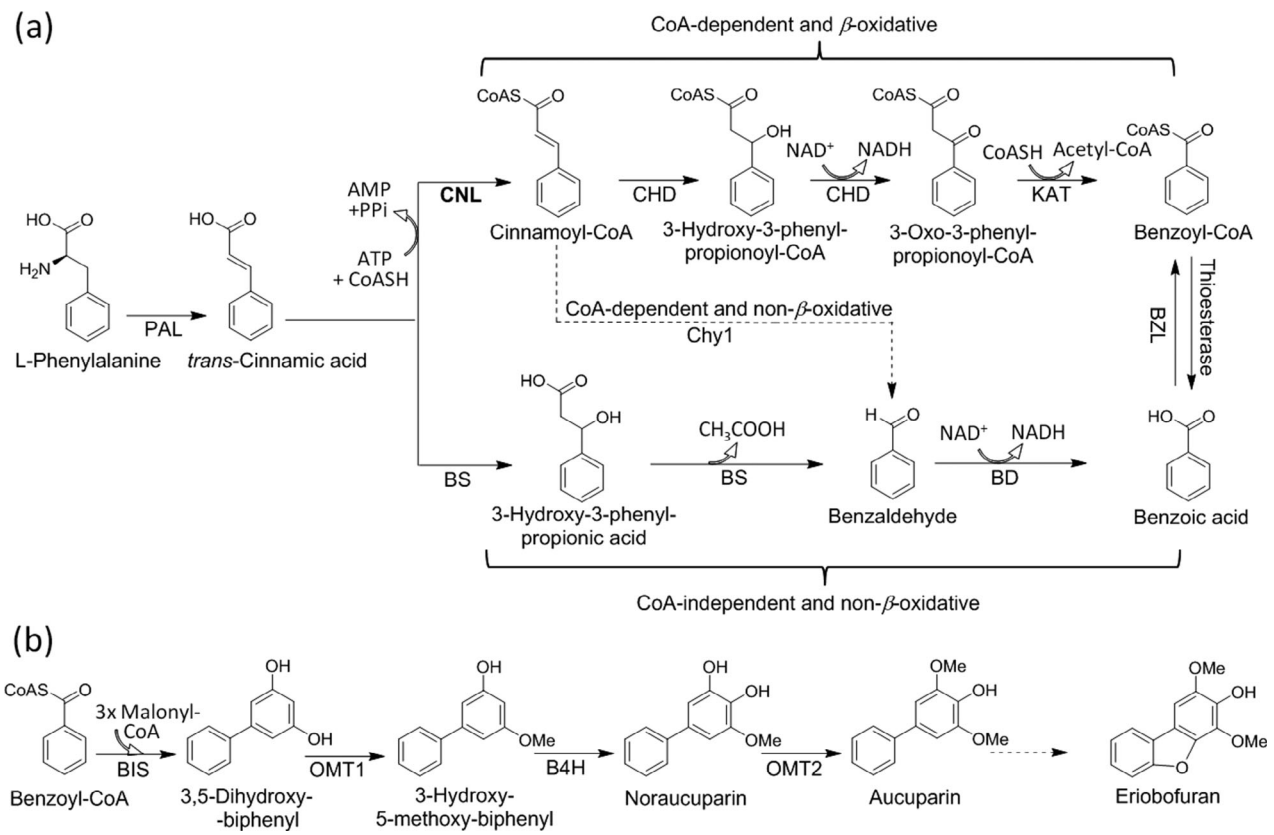

Figure 1. Proposed role of cinnamate-CoA ligase (CNL) in the biosynthesis of biphenyl phytoalexin in Malinae.
(a) Various routes of benzoyl-CoA formation.
(b) Benzoyl-CoA-derived biosynthesis of biphenyl and dibenzofuran in Malinae.
BIS, biphenyl synthase; BS, benzaldehyde synthase; BD, benzaldehyde dehydrogenase; B4H, biphenyl-4-hydroxylase; BZL, benzoate-CoA ligase; CHD, cinnamoyl-CoA hydratase dehydrogenase; KAT, ketoacylthiolase; OMT, biphenyl-*O*-methyltransferase; PAL, phenylalanine ammonia lyase. The dashed arrows indicate putative catalysis for a cinnamoyl-CoA; hydrolase1 (Chy1), which is assumed to convert cinnamoyl-CoA into benzaldehyde under *in vivo* conditions in (a) and a dibenzofuran synthase in (b).

obtained from these *in vitro* and *in vivo* systems suggests promising possibilities for the metabolic engineering of apple plants with improved scab resistance.

RESULTS

Accumulation of biphenyl phytoalexins in response to elicitor and pathogen treatment

Cell cultures of the apple cultivar 'Golden Delicious' responded to treatment with *Venturia inaequalis* elicitor (VIE) by producing two biphenyl phytoalexins, as shown by HPLC with a diode-array detector (Figure S1 in the online Supporting Information). These compounds were identified as noraucuparin and aucuparin, and the time courses of their induced accumulation were studied (Figure 2). Noraucuparin formation started at 6 h post-elicitation (hpe) and reached a peak at 24 hpe [4.2 $\pm$ 0.2 μg g^{-1} fresh weight (FW)]. Thereafter, the noraucuparin concentration decreased as rapidly as it increased, almost reaching the basal level at 72 hpe (1.1 $\pm$ 0.1 μg g^{-1} FW). In contrast, aucuparin accumulation started at 12 hpe, and its level continued to increase during the entire time course studied. The maximum aucuparin level was recorded at 72 hpe (5.4 $\pm$ 0.1 μg g^{-1} FW). Phytoalexins were only detected inside the cells, with a trace amount being observed in the culture medium (0.1 μg 50 ml^{-1} medium) that was probably due to occasional cell death during the elicitation process. The untreated cell cultures showed no phytoalexin accumulation over the time course studied. In contrast to cv. 'Golden Delicious', the leaves of cv. 'Shireen' did not form a friable callus, hence cell cultures of this cultivar could not be included in the present study. When leaves of the greenhouse-grown plants of cv. 'Shireen' and cv. 'Golden Delicious' were infected with *V. inaequalis* conidia, accumulation of noraucuparin and aucuparin was observed in the internode regions of both cultivars. However, the levels of accumulation of these phytoalexins were higher in the scab-resistant cultivar 'Shireen' compared with the moderately scab-susceptible cultivar 'Golden Delicious' (Figure 3). In contrast, the leaves

failed to show any accumulation of biphenyl phytoalexins. Moreover, when different apple cultivars with varying tolerance towards scab disease were compared for their ability to produce biphenyl phytoalexin after scab infection, a clear correlation was observed between the level of biphenyl accumulation and the previously reported degrees of scab tolerance (Table S1; Krüger, 1991; Kellerhals *et al.*, 1993; Dewdney *et al.*, 2003; Biggs *et al.*, 2010; Khajuria *et al.*, 2014; Shafi *et al.*, 2019). Upon scab infection, greenhouse-grown plants of the highly scab-susceptible cultivar 'Vista Bella' failed to produce any biphenyls. The highly scab-resistant cultivars 'Shireen' and 'Prima' produced approximately 2–2.5-fold higher levels of biphenyls than the moderately scab-susceptible cultivar 'Golden Delicious' (Table S1).

Cloning of a CNL cDNA from cv. 'Golden Delicious'

The genome sequence of 'Golden Delicious' contained genes that shared significant similarities with the CNL sequences from *Hypericum calycinum* (accession no. AFS60176) and *P. hybrida* (accession no. AEO52693). *In silico* analyses of the clean genome sequence from the Genome Database for Rosaceae (GDR; https://www.rosaceae.org/; Table S2) identified one apple unigene (MDP0000576682) that shared about 73% and 67% sequence similarity with *H. calycinum* and *P. hybrida* CNLs, respectively. Therefore, the coding sequences of MDP0000576682 were selected as candidate unigenes to design gene-specific forward and reverse primers for PCR amplification of the corresponding transcript from a cDNA pool isolated from VIE-treated (6 hpe) 'Golden Delicious' cell cultures. The full-length cDNA was again amplified using a proofreading DNA polymerase. The resulting full-length cDNA consisted of a CDS

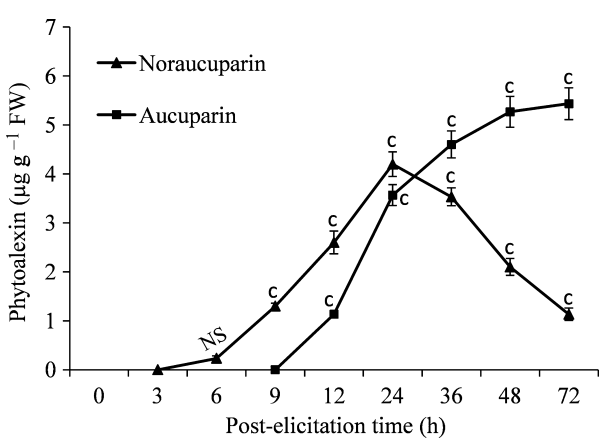

Figure 2. Time course accumulation of biphenyl phytoalexins in cell cultures of *Malus domestica* cv. 'Golden Delicious' after treatment with *Venturia inaequalis* elicitor.
Results are means ± SD values of three independent experiments and statistical significance was calculated using *t*-test. $^{c}P < 0.001$ compared with the untreated cells (0 h), NS, not significant.

of 1725 bp, a 5′ untranslated region (UTR) of 240 bp, a 3′ UTR of 447 bp and an 11 bp polyA tail. The CDS encoded a 574-amino acid-long protein with a mass of 63.67 kDa and an isoelectric point of 7.4. The molecular mass of the recombinant protein was confirmed by matrix-assisted laser desorption/ionisation mass spectrometry (MALDI-MS) analysis (Figure S2). The *Md*CNL amino acid sequence bears a putative AMP-binding motif (box I) and a second box II at a distance of 196 and 395 amino acids from the N-terminus, respectively (Figure 4). A glycine residue at amino acid position 254 (marked by an asterisk in Figure 4) was found to be conserved in CNLs (Schneider *et al.*, 2003). As predicted by a software tool (Rottensteiner *et al.*, 2004), the C-terminal end of the *Md*CNL protein consisted of a peroxisomal targeting signal type 1 (PTS1) comprising the tripeptide SRL. The sequence similarity of putative *Md*CNL with 4-coumarate-CoA ligases (4CLs) was low (26–30%). However, *Md*CNL shared 62.3%, 66.7% and 72.8% homology with CNLs from *A. thaliana* (*At*BZO1; Q9SS01), *P. hybrida* (*Ph*CNL; AEO52693) and *H. calycinum* (*Hc*CNL; AFS60176), respectively. It also presented 72.5% homology with a predicted *o*-succinylbenzoate-CoA ligase 1 from *Medicago truncatula* (OSBZL1; XP_003600627).

Functional characterisation of *Md*CNL

The *Md*CNL coding sequence (CDS) was expressed in *Escherichia coli* as a His_6-tag fusion protein, then purified to homogeneity using Ni-NTA agarose. Upon incubation with cinnamic acid and CoA in the presence of ATP, the recombinant protein formed cinnamoyl-CoA, as shown by spectrophotometric and liquid chromatography (LC)-MS analyses. The temperature and pH optima of the protein were 30°C and 6.5, respectively. Activity of *Md*CNL was linear with time and with the protein concentration up to 10 min and 10 μg, respectively. Cations were a pre-requisite for its enzymatic activity (Figure 5). The activity obtained under supplementation with both divalent Mg^{2+} (2.5 mM) and monovalent K^{+} (100 mM) was highest and was set to 100% (Figure 5). Among tested divalent cations, the least promotion of the catalytic activity was observed with copper Cu^{2+} (2%) followed by Zn^{2+} (14%), Fe^{2+} (18%), Co^{2+} (20%), Ca^{2+} (20%), Mn^{2+} (22%) and Mg^{2+} (26%). Among divalent cations, Mg^{2+} resulted in the best activity. However, K^{+} (100 mM) was the only monovalent cation that increased the activity almost two-fold in the presence of Mg^{2+}, whereas supplementation of Na^{+} alone in the presence of Mg^{2+} had no significant effect compared with supplementation with Mg^{+2} alone as a cationic source. No significant loss of *Md*CNL activity was observed upon storage at −80°C (in 15% glycerol) for 6 months. Approximately 80% of the enzymatic activity was detected after storing the protein for 24 h at 4°C. Based on these optimum conditions, the substrate preference of *Md*CNL was

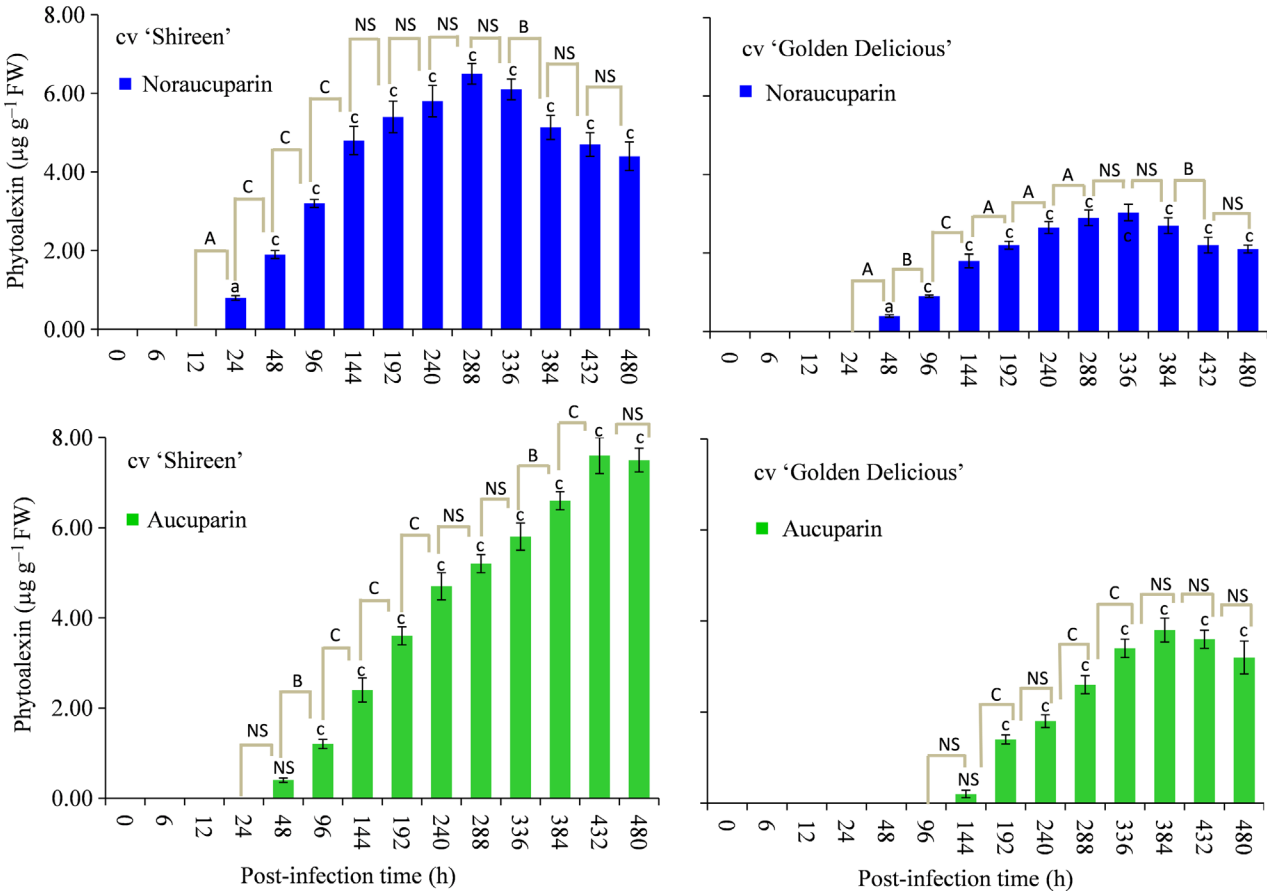

Figure 3. Time course accumulation of biphenyl phytoalexins in internodes of apple plants of cv. 'Shireen' and cv. 'Golden Delicious' infected with *Venturia inaequalis* conidia.
Results are means ± SD of three independent experiments. Statistical significance was calculated using one way ANOVA. Independent *t*-tests were performed to evaluate the differences between two consecutive groups. $^{a}P < 0.05$, $^{b}P < 0.01$, $^{c}P < 0.001$ compared with the untreated shoots (0 h); $^{A}P < 0.05$, $^{B}P < 0.01$, $^{C}P < 0.001$ comparison between two consecutive time points; NS, not significant.

studied using a luciferase-based assay that reflected CoA-ligase activity by measuring the ATP remaining in the assay. Incubation of heat-denatured *Md*CNL without any substrate was considered as a negative control, and the measured luciferase activity was set as 100%, indicating that no ATP was consumed by CoA-ligase (Figure 6a). A wide array of potential substrates for *Md*CNL (cinnamic acid, benzoic acid and their related acids) were tested. The preferred substrate for *Md*CNL was cinnamic acid, followed by 4-coumaric acid. Minor activity was detected in the presence of caffeic, ferulic and sinapic acids. Notably, no enzyme activity was observed in the presence of benzoic acid, identifying this enzyme as CNL and not BZL. The substrate specificity was also checked by a spectrophotometric assay, which confirmed the preference for cinnamic acid as a substrate for *Md*CNL (Figure 6b). Caffeic, ferulic and sinapic acids were unsuitable substrates for *Md*CNL. The product of the *Md*CNL-catalysed reaction with cinnamic acid (i.e. cinnamoyl-CoA) was identified by electrospray ionisation (ESI)-MS in negative ion mode. The molecular ion $[M\text{-}H]^-$ at *m*/*z* 896.2 was an indicative peak for cinnamoyl-CoA formation, which was subjected to further MS fragmentation. The resulting MS/MS spectrum matched the previously reported fragmentation pattern of cinnamoyl-CoA (Beuerle and Pichersky, 2002b) (Figure S3). The kinetic constants of *Md*CNL were determined with cinnamic and 4-coumaric acids (Table 1). Both the affinity and the turnover number of *Md*CNL were significantly higher for cinnamic acid than for 4-coumaric acid. The catalytic efficiency (k_{cat}/K_m) for cinnamic acid was 44 times that for 4-coumaric acid. The K_m values for the co-substrates ATP and CoA were 102.4 and 108.6 μM, respectively.

Induced expression of *MdPAL*, *MdCNL* and *MdBIS3*

In VIE-treated 'Golden Delicious' cell cultures, the *Md*PAL, *Md*CNL and *Md*BIS3 transcript levels were analysed by quantitative real-time PCR (qPCR) over 72 hpe (Figure 7a). Both *MdPAL* and *MdBIS3* are known to be inducible (Chizzali *et al.*, 2012a; Sircar *et al.*, 2015). *Profilin* mRNA was used as an internal standard to verify equal template

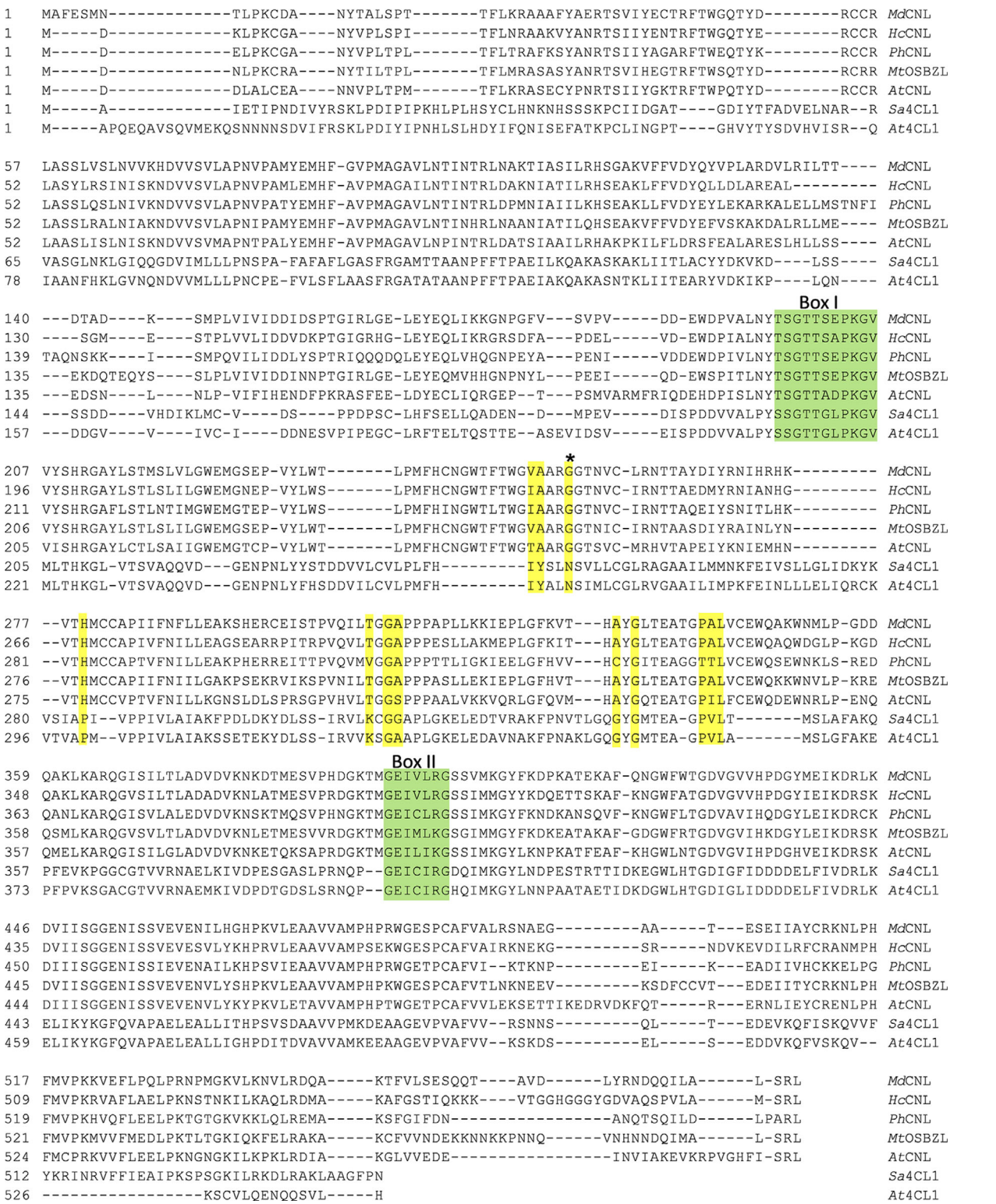

Figure 4. Amino acid sequence alignments.
Amino acid sequence alignments of cinnamate-CoA ligase (CNL) of *Malus domestica* cv. 'Golden delicious' (*Md*CNL; MG334585), *Hypericum calycinum* CNL (*Hc*CNL AFS60176), *Petunia hybrida* CNL (*Ph*CNL AEO52693), predicted *o*-succinylbenzoate-CoA ligase 1 (OSBZL1) from *Medicago truncatula* (*Mt*OSBZL; XP_003600627), *Arabidopsis thaliana* CNL (*At*CNL, *At*BZO1; Q9SS01) and 4-coumarate-CoA ligase 1 (4CL1) from *A. thaliana* (*At*4CL1; NP_001077697) and *Sorbus aucuparia* (*Sa*4CL1; ADF30254). The conserved motifs (box I and box II) and 12 amino acids predicted to function as a 4CL substrate specificity code are highlighted in green and yellow, respectively. The asterisk indicates a glycine residue conserved in all CNLs. C-terminal sequences (SRL, ARL) represent peroxisomal targeting signals. Gaps introduced to maximise the alignment are indicated by dashes.

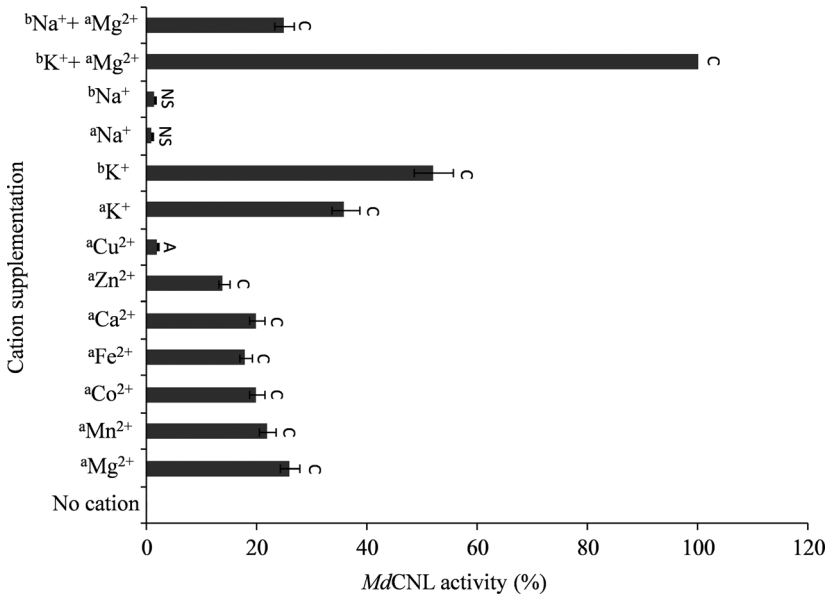

Figure 5. Effect of different cations on *Malus domestica* cinnamate-CoA ligase (*Md*CNL) activity.
Assays contained 2.5 and 100 mM concentrations of the divalent and monovalent cations, respectively. Maximum activity was recorded for assays containing 2.5 mM Mg^{2+} and 100 mM K^{+} and it was set as 100%. a = 2.5 mM, b = 100 mM concentration. Results are means ± SD (n = 3). Statistical significance was calculated using one-way ANOVA. $^{C}P < 0.001$ compared with assay supplemented with no cations. NS, not significant.

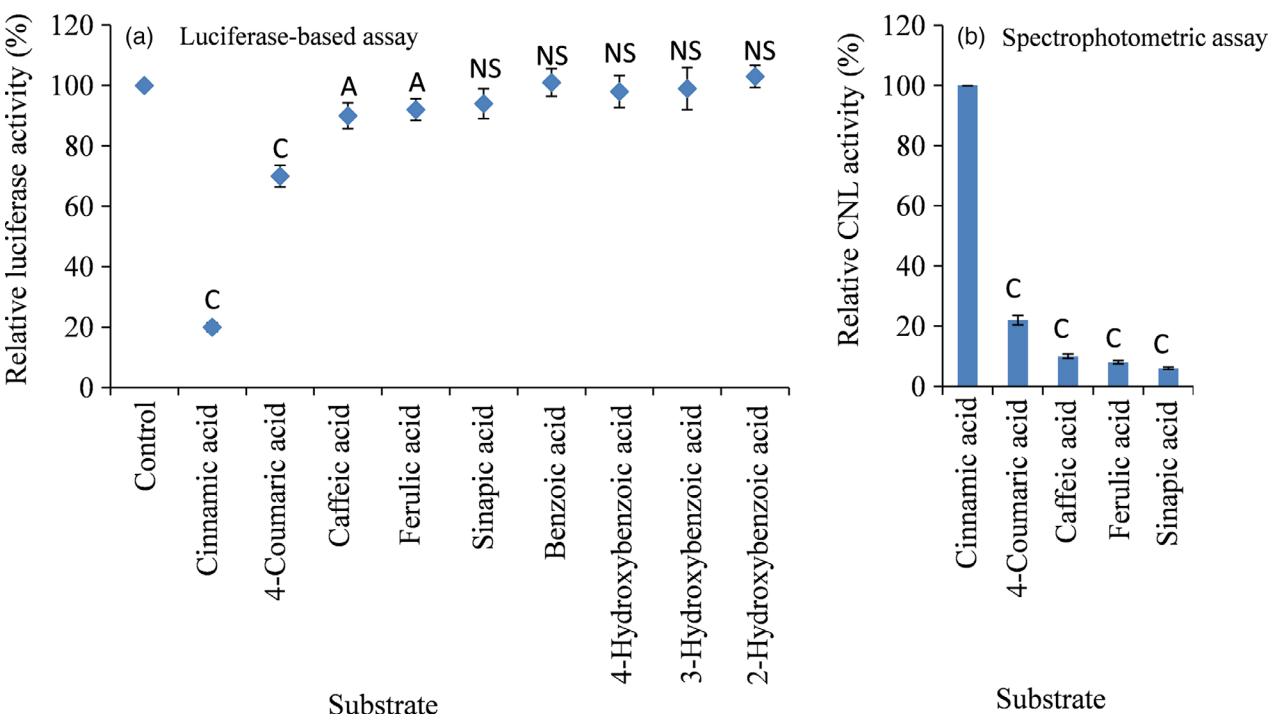

Figure 6. Substrate specificity of recombinant *Malus domestica* cinnamate-CoA ligase (*Md*CNL).
(a) Luciferase-based assay. The left-over ATP in the CNL assay was measured as luciferase-dependent bioluminescence, which inversely correlated with CNL activity. Reaction with heat-denatured *Md*CNL was used as negative control to normalise bioluminescence. Results are mean ± SD values of three independent experiments and statistical significance was calculated using the t-test. $^{A}P < 0.05$, $^{C}P < 0.001$ compared with negative control; NS, not significant.
(b) Spectrophotometric-based assay, in which CNL activity was calculated using the previously reported extinction coefficients of the respective CoA esters, considering activity with cinnamic acid (cinnamoyl-CoA formation) as 100%. $^{C}P < 0.001$ compared with activity recorded with cinnamic acid.

amounts. Transcripts of all the studied genes were detectable at 0.5 hpe. The mRNA levels of *Md*PAL and *Md*CNL peaked at 6 h and decreased thereafter. However, a faster decrease was observed for *Md*CNL. The expression pattern of *Md*BIS3 was different. Despite the similar timing of initial detection (0.5 h), it peaked 6 h later (after 12 h). Thereafter, the expression level rapidly decreased and reached the basal level at 72 h. Following the expression of *Md*PAL,

Table 1 Kinetic properties of *Malus domestica* cv. 'Golden Delicious' cinnamate-CoA ligase. Values are means ± SD (n = 3)

Substrates	K_m (μM)	V_{max} (nkat mg^{-1} protein)	K_{cat} (sec^{-1})	K_{cat}/K_m (μM^{-1} sec^{-1})
Cinnamic acid	19.6 ± 0.88	13.8 ± 1.05	0.88 ± 0.06	0.044 ± 0.001
4-Coumaric acid	126.4 ± 3.34	2.7 ± 0.37	0.17 ± 0.02	0.001 ± 0.0003

For ATP and CoA, the K_m values were 102.4 ± 4.6 and 108.6 ± 5.2 μM, respectively.

*Md*CNL and *Md*BIS3, accumulation of noraucuparin and aucuparin was observed (Figure 2), supporting their involvement in biosynthesis of biphenyl phytoalexin.

When greenhouse-grown apple plants were infected with *V. inaequalis* conidiospores, the leaves of cv. 'Shireen' failed to show any visible scab symptoms after 16 days of infection; however, 'Golden Delicious' leaves showed minor scab symptoms after 16 days (Figure S4). Moreover, pinpoint pits, the typical sign of a hypersensitive response (HR), were observed in 'Shireen' leaves. These pinpoint pits are associated with the scab resistance properties of 'Shireen', as observed previously in the scab-resistant apple cultivar 'Florina' (Sarkate *et al.*, 2019). The expression of *Md*CNL in the *V. inaequalis*-infected stems of the two cultivars was measured in relation to that observed in mock-inoculated stems over a time-course of 16 days (384 h). Compared with 'Golden Delicious', a faster and stronger increase in the *Md*CNL transcript level was observed in 'Shireen' (Figure 7b). After the onset of scab fungus infection, *Md*CNL expression in 'Shireen' peaked at day 6, whereas the transcript level in 'Golden Delicious' did not reach a maximum until day 8. Furthermore, compared with mock-inoculated shoots, the *Md*CNL expression

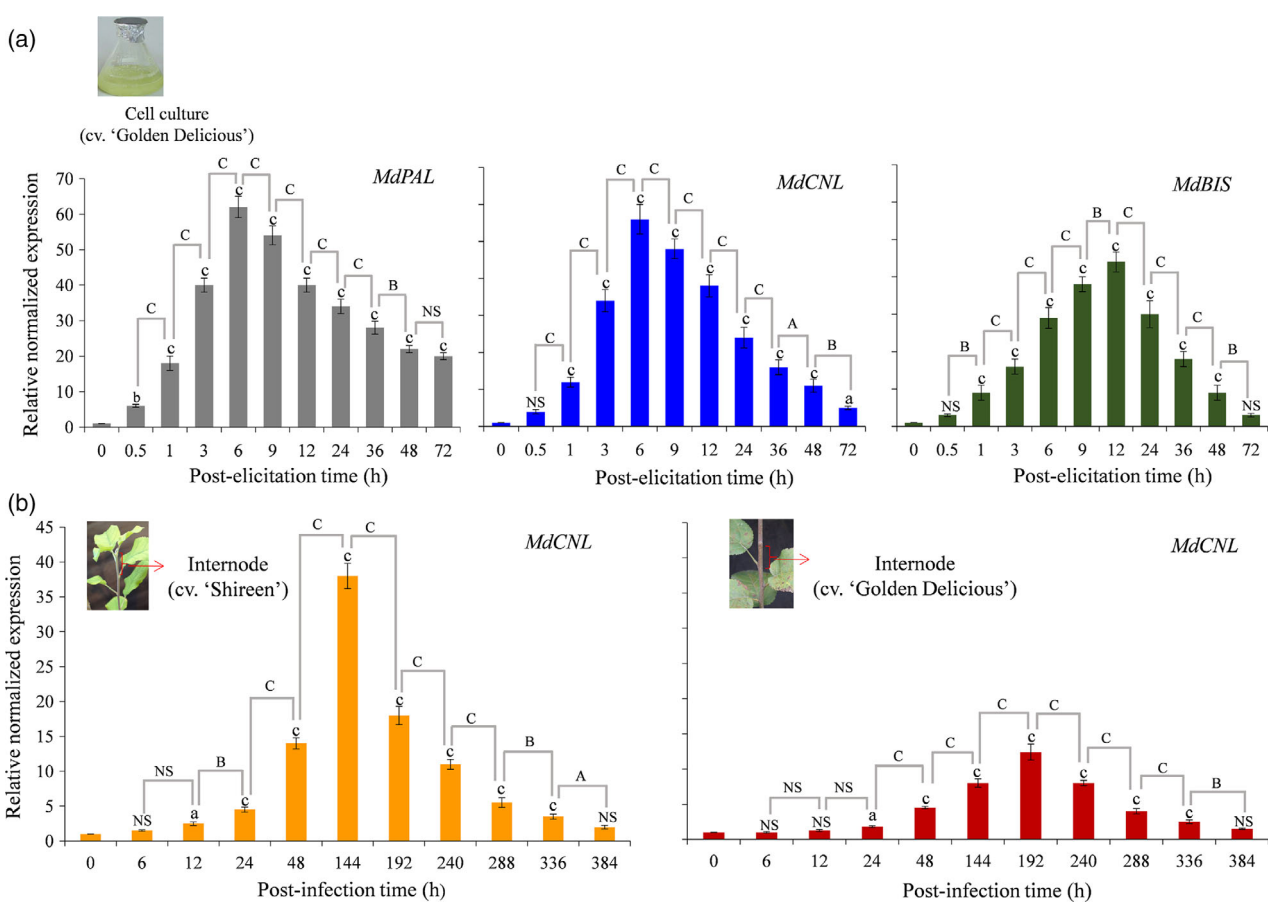

Figure 7. Changes in gene expression levels in treated apple cell cultures and plants.
(a) Changes in the expression of *MdPAL* (grey), *MdCNL* (blue) and *MdBIS3* (green) in *Venturia inaequalis* elicitor-treated cell cultures of cv. 'Golden Delicious'.
(b) Changes in the expression level of *MdCNL* in scab-infected shoots of cv. 'Shireen' (orange) and cv. 'Golden Delicious' (red). The relative transcript levels were determined by quantitative real-time PCR. Results are mean ± SD (n = 3). Statistical significance was calculated using one-way ANOVA. Independent t-tests were performed to evaluate the differences between two consecutive groups. $^{a}P < 0.05$, $^{b}P < 0.01$, $^{c}P < 0.001$ compared with the expression at 0 h; $^{A}P < 0.05$, $^{B}P < 0.01$, $^{C}P < 0.001$ comparison of expression level between two consecutive time points; NS, not significant.

level was 38-fold higher in 'Shireen' but only 12.5-fold higher in 'Golden Delicious'.

Phylogenetic analyses of *Md*CNL

To generate a phylogenetic tree, benzoate- and cinnamate-specific-CoA ligases and 4-coumarate-CoA ligase (4CL) enzymes from angiosperms were taken into consideration (Figure 8). The amino acid sequence accession numbers are listed in Table S4. The resulting phylogenetic tree consisted of two main evolutionarily distinct clusters of CoA ligases. The first cluster comprised the 4CL sequences, whereas the second cluster comprised benzenoid-related-CoA ligases including *Md*CNL. The clade that consisted of *Md*CNL and other functionally characterised CNLs grouped together with a clade that comprises the benzenoid-related ligases from *Clarkia* and *Medicago*.

Subcellular localisation of *Md*CNL

Bioinformatic processing of the *Md*CNL sequence showed the presence of a putative C-terminal PTS1 signal represented by the SRL amino acids (Figure 4). To determine the functionality of this PTS1 signal, Gateway™ constructs encoding either N- or C-termini–YFP fusion proteins (YFP-*Md*CNL or *Md*CNL-YFP, respectively) of *Md*CNL were generated and subsequently transformed into *N. benthamiana* leaf epidermal cells for transient expression. A peroxisomal marker protein, cyan fluorescent protein (CFP) (Nowak *et al.*, 2004), was used for co-localisation studies. Expressions of all the constructs were driven by 35S promoter originating from the cauliflower mosaic virus. For the N-terminal YFP-*Md*CNL construct, characteristic punctate fluorescence was observed, indicating peroxisomal localisation (Figure 9a,c). In contrast, when the C-terminal *Md*CNL-YFP construct was used, yellow fluorescence was dispersed in the cytoplasm, indicating masking of the PTS1 signal peptide of *Md*CNL (Figure 9b). The co-localisation of the marker construct CFP-PTS1 was also directed toward the peroxisomes (Figure 9d). When merged, the yellow dots of YFP exactly overlapped with the cyan dots of CFP, thereby demonstrating the co-localisation (Figure 9f).

Molecular modelling of *Md*CNL

No well-characterised three-dimensional (3D) structure of the CNL from *M. domestica* is available in the PDB database. Therefore, a comparative or homology modelling approach was adopted using Robetta to construct a high-quality 3D structure; this generated five models, as shown in Figure 10 (superposed figure). These models developed using the X-ray-solved structure of acetyl CoA synthase from *Salmonella enterica*, which has a query coverage of 90% and sequence identity of 27%. The acetyl CoA-ligase (PDB ID: 1pg4) from *Salmonella* sp. was used as the template to develop the model of the query sequence. The confidence score of the resulting models was 0.66, suggesting a medium-quality model (Song *et al.*, 2013).

To validate the predicted structure and obtain an equilibrated model, a molecular dynamics simulation of the Theseus superimposed structure was performed for 25 ns. The simulation trajectory indicated convergence, as was evident from the root mean square deviation (RMSD) analysis of the involved trajectory (Figure S5). In particular, until 5 ns, the structure showed minor fluctuations, but thereafter, it showed minimal RMS fluctuations (between 0.5 and 0.6 nm). These observations indicated that after 5 ns, the structure reached equilibrium, and the obtained trajectory was that of a stable conformation. In addition, the residues showing most of the fluctuations in the trajectory were located at the C-terminus (Figure S6), which is predicted to be in the active site. During the simulation, the protein did not denature and remained stable throughout most of the simulation process, as was evident from the radius obtained through gyration analysis (Figure S7). It fluctuated only between 2.5 and 2.6 nm. Therefore, molecular dynamics simulation studies demonstrated that this homology-modelled structure derived from Robetta is stable in terms of energy. This energetically stable model was found to be suitable for molecular docking with cinnamic acids and other preferred substrates of *Md*CNL for comparative interaction analysis.

The docking of six molecules was performed on the 25 ns-equilibrated structure of *Md*CNL using AutoDock4.2 with settings for detailed or exhaustive docking in a blind approach. Here, as the active site was not specified, the molecule identified locations/regions that were favourable for binding. We examined the clustering of the low-energy conformations because any molecule with a high affinity would present a large cluster of similar conformations in a specific spatial location of the protein. Using this method, we observed differences in the molecule's specificity. The docking analysis suggested that cinnamic acid binds mainly to the largest cavity of the protein, where 109 conformations out of 200 presented the identical location (Figure 11a). In contrast, 4-coumaric acid showed weak affinity to the same cavity (76 out of 200 conformations presented the same location; Figure 11b). Similarly, ferulic and sinapic acids showed fewer conformations (38 out of 200 and 42 out of 200 conformations) in the same spatial location as cinnamic acid. However, the lowest-energy-conformation clusters of benzoic acid and caffeic acid were bound to another location of the protein, indicating that they exhibit much lower specificity than the others.

DISCUSSION

In plants, benzoic acid and its derivatives serve as biosynthetic building blocks for the synthesis of an array of metabolites with established roles in plant–pathogen interactions. One interesting example is provided by the

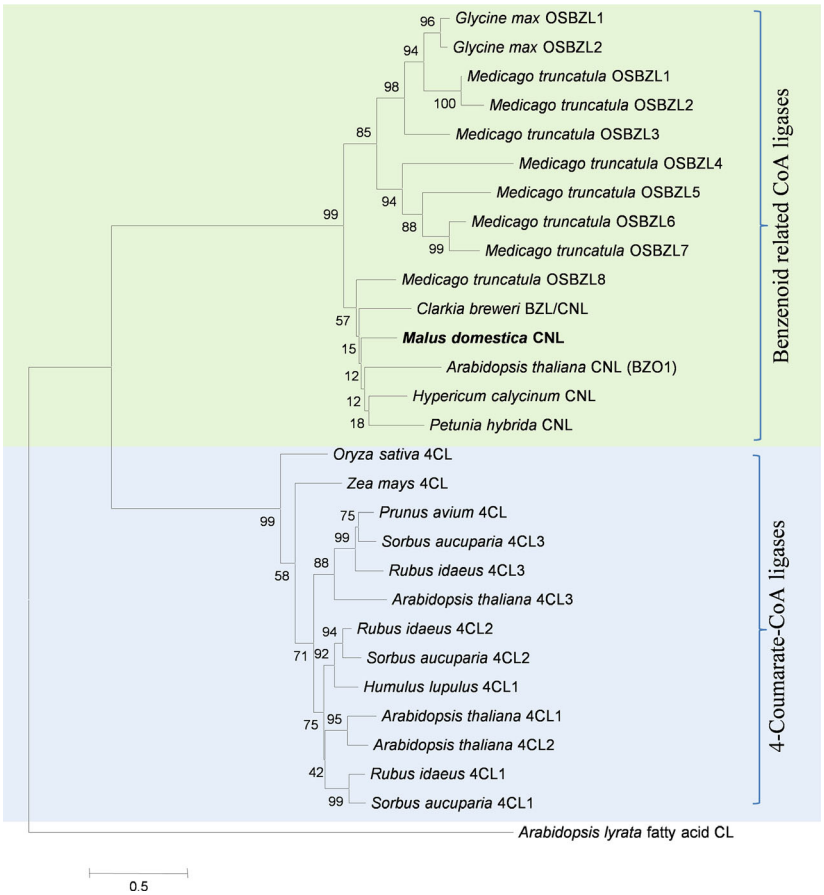

Figure 8. Phylogenetic tree demonstrating the evolutionary relationships between *Malus domestica* cv. 'Golden Delicious' cinnamate-CoA ligase (*Md*CNL) and other benzoate- and cinnamate-related-CoA ligases.
The neighbour-joining tree was constructed using amino acid sequences through MEGA 5.05. The bootstrap values are indicated at the branch points. The scale bar represents 0.5 amino acid substitutions per site. *Medicago truncatula o*-succinylbenzoate-CoA ligase (OSBZL) sequences were numbered according to the order of appearance in the tree. 4CL, 4-coumarate-CoA ligase.

biphenyl phytoalexins of *Malus* species. The biosynthesis of benzoic acid in Malinae is not yet completely understood. In the phenylpropanoid pathway, cinnamic acid acts as a central metabolite which diverts the carbon flux into various sub-pathways. In most plant species, cinnamic acid is first converted into 4-coumaric acid by the enzyme cinnamate-4-hydroxylase. Subsequently, 4-coumaric acid is thioesterified by 4CL to form 4-coumaroyl-CoA, which serves as a precursor for several plant secondary metabolites, flavonoids, anthocyanins and phenolics (Li *et al.*, 2015). In addition to the 4CL route, cinnamoyl-CoA is the thioesterified product of CNL when incubated with cinnamic acid and CoA, directing the carbon flux from the phenylpropanoid pathway towards the biosynthesis of benzoate-derived metabolites (Figure 1). The thioesterification of cinnamic acid to form cinnamoyl-CoA is catalysed by *Md*CNL. *Md*CNL possesses a conserved AMP-binding domain, which is a characteristic feature of all members of the adenylate-forming enzyme family, such as 4CLs, fatty acid-CoA ligases and luciferases (Fulda *et al.*, 1994). The AMP-binding domain (box I in Figure 4), which is present in *Md*CNL, is also a characteristic feature in CNLs from *P. hybrida* (*Ph*CNL), *H. calycinum* (*Hc*CNL) and *A. thaliana* (*At*CNL/*At*BZO1) as well as OSBZL from *M. truncatula*. Furthermore, a second conserved domain (box II in Figure 4) is present in all these proteins.

There are previous reports of the cloning and functional characterisation of CNL cDNAs from *H. calycinum* cell cultures (Gaid *et al.*, 2012) and *P. hybrida* petals (Klempien *et al.*, 2012). The affinity of *Md*CNL for cinnamic acid (K_m = 19.6 μM) is markedly higher than that of *Ph*CNL (K_m = 285.7 μM) but slightly lower than that of *Hc*CNL (K_m = 11.14 μM). Similarly, the turnover number of *Md*CNL (k_{cat} = 0.88 sec^{-1}) is lower than that of *Hc*CNL (k_{cat} = 1.73 sec^{-1}) but higher than that of *Ph*CNL (k_{cat} = 0.472 sec^{-1}). For all three studied CNLs, 4-coumaric acid was the second preferred substrate, and benzoic acid in particular was not accepted as a substrate. The affinity of *Md*CNL for CoA (K_m = 108.6 μM) was slightly lower than that of *Hc*CNL (K_m = 95.6 μM) but markedly higher than that of *Ph*CNL (K_m = 775.2 μM). However, the affinity of *Md*CNL towards ATP (K_m = 102.4 μM) was comparable to that of *Hc*CNL (K_m = 104.7 μM), whereas the ATP K_m value of *Ph*CNL was not reported (Klempien *et al.*, 2012). Notably, cations were a pre-requisite for the activities of *Md*CNL, *Hc*CNL and *Ph*CNL, where the combination of Mg^{2+} and K^+ provided the best result. A similar need for cations was previously observed in *Hc*CNL (Gaid *et al.*, 2012). The

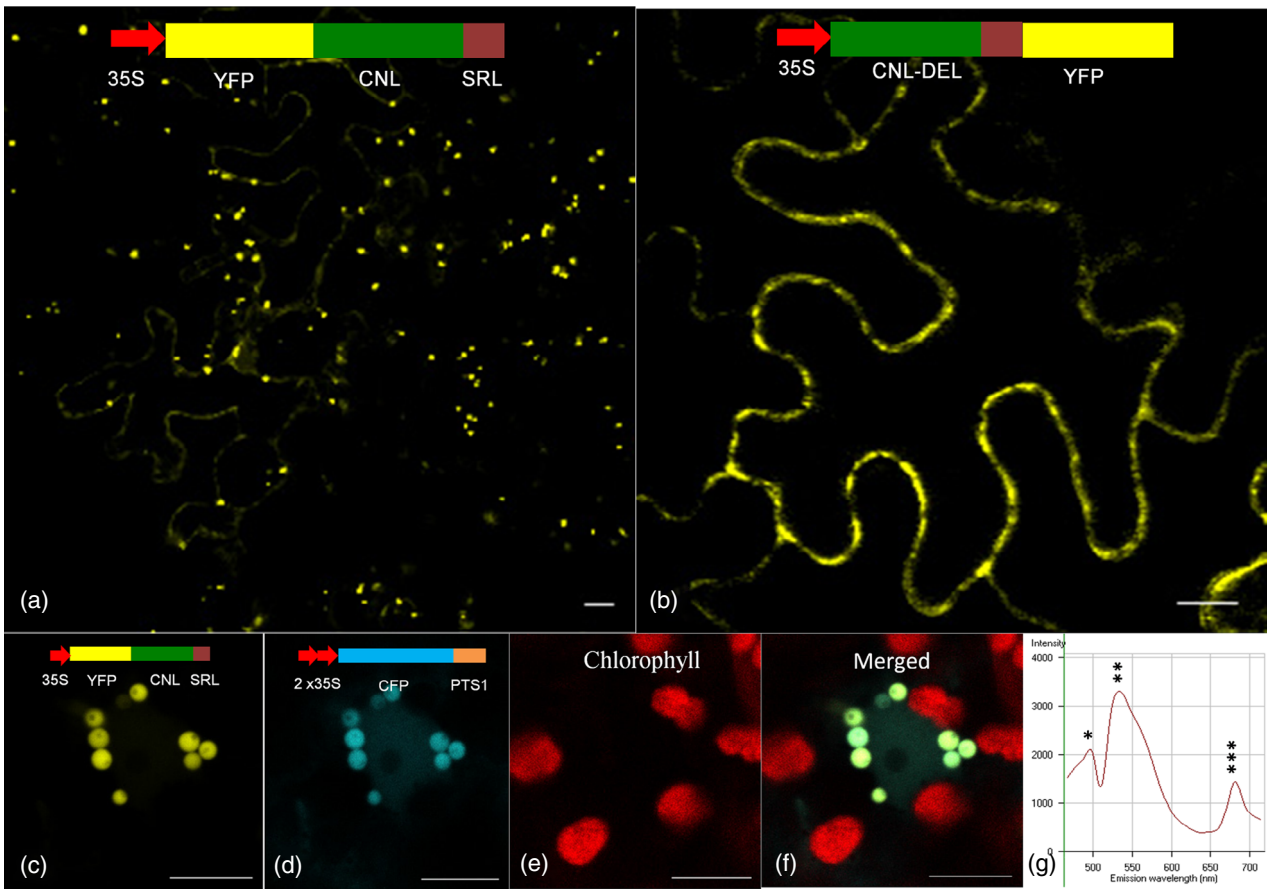

Figure 9. Localisation studies of *Malus domestica* cv. 'Golden Delicious' cinnamate-CoA ligase (*Md*CNL) in leaf epidermis cells *of Nicotiana benthamiana*. Transient expression of YFP-CNL (a) and CNL-YFP (b) constructs resulted in peroxisomal and cytosolic localisation of CNL fusion proteins, respectively. Co-transformation with YFP-CNL (c) and the marker construct CFP-PTS1 (d), autofluorescence of chloroplasts (e) and their merged images (f).
(g) Emission fingerprints of cyan fluorescent protein (CFP) (*), YFP (**) and chlorophyll (***). 35S, promoter from *Cauliflower mosaic* virus; DEL, deleted stop codon; SRL, type 1 peroxisomal targeting signal (PTS1). Bars = 10 μm.

highest *Md*CNL activity was observed with 100 mM potassium. The previously reported *Hc*CNL also showed the highest activity at a 100 mM potassium concentration. However, in the case of *Ph*CNL, 50% of the maximum activity was observed when 10.5 mM potassium supplementation was performed (Klempien *et al.*, 2012).

The preferred substrate for *Md*CNL, *Hc*CNL and *Ph*CNL was cinnamic acid. However, 4CLs showed little or no activity toward cinnamic acid; rather, they preferred 4-coumaric acid as a substrate (Schneider *et al.*, 2003). The *Ri*4CL2 isoform isolated from *Rubus idaeus* (Kumar and Ellis, 2003) preferred cinnamic acid (153%) as a substrate over 4-coumaric acid (100%). It has been reported that the 12 amino acids surrounding the substrate-binding pocket of 4CL actually function as a signature motif for 4CL substrate specificity. These authors reported an Arabidopsis *At*4CL2 triple mutant, which presented approximately 30-fold higher affinity towards cinnamic acid as a substrate. The triple mutation N256A/M293P/K320L is known to increase the hydrophobicity of the substrate-binding pocket. Interestingly, *Md*CNL exhibits hydrophobic glycine residues at the corresponding position of N256 in *At*4CL2 (Figure 4). In contrast, cinnamate-preferring *Ri*4CL2 bears an asparagine residue at the comparable position. The other two amino acid positions that were altered in the *At*4CL triple mutant are occupied by hydrophilic histidine and threonine residues in all the functionally characterised CNLs. Furthermore, phylogenetic analyses showed that the benzenoid-related-CoA ligases and 4CLs form two evolutionarily distant clusters, indicating their origins from different ancestral genes (Figure 8).

In cell cultures of *S. aucuparia*, C_2-chain cleavage of cinnamic acid is known to proceed via dehydrogenase activity in the CoA-dependent, non-β-oxidative route, giving rise to free benzoic acid (Gaid *et al.*, 2009). It is likely that CNL-formed cinnamoyl-CoA is first converted into benzaldehyde, which is then transformed into benzoic acid by the reported dehydrogenase activity. A *Chy1* gene reported in *A. thaliana* (Ibdah and Pichersky, 2009) encodes 3-hydroxyisobutyryl-CoA hydrolase, which is assumed to convert

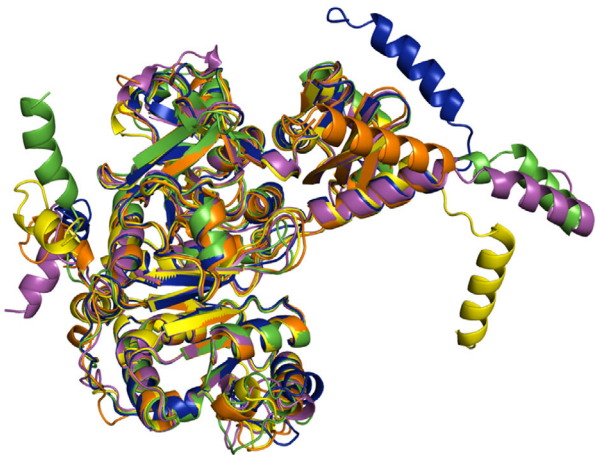

Figure 10. Superposition of *Malus domestica* cv. 'Golden Delicious' cinnamate-CoA ligase models.
The five models obtained from Robetta are superposed using the CE algorithm, where the root mean square deviation among the five models is about 0.6 Å. The main differences among the models are in the N- and C-termini of the protein.

cinnamoyl-CoA into benzaldehyde under *in vivo* conditions. Benzaldehyde dehydrogenase (BD) activity has also been detected in cell cultures of *P. pyrifolia* (Saini *et al.*, 2017). The resulting benzoic acid appears to be converted to benzoyl-CoA in a reaction catalysed by an as yet unidentified plant BZL. Benzoyl-CoA then enters the biphenyl phytoalexin biosynthetic pathway by combining with three molecules of malonyl-CoA, catalysed by biphenyl synthase. Thus, in 'Golden Delicious' cell cultures, the sequence of biphenyl biosynthesis appears to involve CNL, BD and BZL activities. The increases in the gene expression levels of *PAL, CNL* and *BIS3* preceded the accumulation of noraucuparin and aucuparin, which function as phytoalexins in Malinae (Figures 2 and 7a). Moreover, the coordinated changes in *Md*PAL, *Md*CNL and *Md*BIS3 transcript levels (Figure 7a) argue for the involvement of these enzymes in the formation of biphenyl phytoalexins through intermediate formation of benzoic acid and its derivatives. The accumulation of biphenyls as a defence mechanism in apple was repeatedly reported in *in vitro* cultures (Sircar *et al.*, 2015; Sarkate *et al.*, 2019) and in greenhouse-grown plants (Chizzali *et al.*, 2012a,b; Sarkate *et al.*, 2018). Similarly, in our study, the level of biphenyls in the scab-resistant 'Shireen' cultivar was almost 2.5-fold higher than that in cv. 'Golden Delicious' (Figure 3). Similar biphenyl accumulation patterns were observed when greenhouse-grown apple plants were infected with the scab fungus (Table S1). This is in agreement with the earlier reports, where scab-

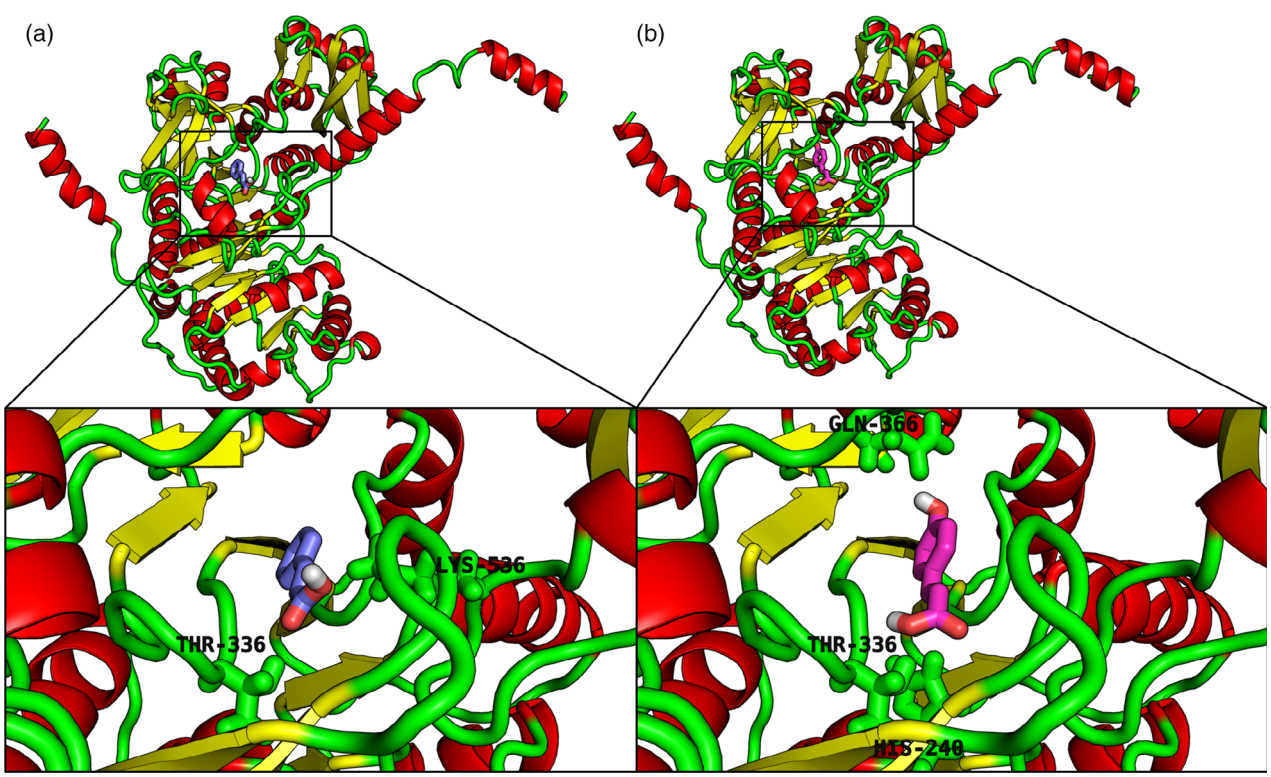

Figure 11. Molecular docking results of cinnamic acid and 4-coumaric acid. The equilibrated structure of *Malus domestica* cv. 'Golden Delicious' cinnamate-CoA ligase is shown in cartoon representation with α-helices coloured red, β-strands coloured yellow and loop regions coloured green.
(a) The top-ranked conformation of cinnamic acid, in stick representation, is bound to the protein making polar contacts with Thr336 and Lys536.
(b) The top-ranked conformation of 4-coumaric acid, in stick representation, is a bound to the protein making polar contacts with Gln366, Thr336 and His240.

susceptible cultivars infected with *V. inaequalis* failed to accumulate phytoalexins (Kellerhals *et al.*, 1993; Dewdney *et al.*, 2003; Sarkate *et al.*, 2018). On the contrary, the scab-resistant cultivars 'Prima' and 'Shireen' are reported to accumulate high levels of phytoalexins upon scab infection (Krüger, 1991; Khajuria *et al.*, 2014; Shafi *et al.*, 2019). Deeper insight was gained upon conducting expression studies of *CNL* in the scab-resistant cv. 'Shireen' and moderately scab-susceptible cv. 'Golden Delicious'. The expression level of the *CNL* gene in the scab-resistant cultivar was almost two- to three-fold higher than that in cv. 'Golden Delicious' between 48 and 192 h post-infection (hpi) (Figure 7b), which was followed by the differential accumulation of biphenyl phytoalexins in each cultivar.

At the biochemical level, only three plant CoA ligases reported thus far exhibit affinity towards benzoic acids: 3-hydroxybenzoate-CoA ligase from *C. erythraea* (Barillas and Beerhues, 2000), BZL from *Clarkia breweri* (Beuerle and Pichersky, 2002a) and OSBZL from *Galium mollugo* (Sieweke and Leistner, 1992).

Similar to *Hc*CNL, *Md*CNL was also localised to the peroxisomes; interestingly, peroxisomes are the site of β-oxidation in plants. Masking of the SRL tripeptide at the C-terminus of *Md*CNL resulted in cytoplasmic localisation. A similar observation was made previously for *Hc*CNL (Gaid *et al.*, 2012) and *Ph*CNL (Klempien *et al.*, 2012). *Hc*CNL carries a C-terminal SRL tripeptide, while *Ph*CNL carries a C-terminal ARL tripeptide as a peroxisomal targeting signal. The CNL protein from *A. thaliana* (Lee *et al.*, 2012) also contains a PTS1 (SRL) signal at the C-terminal end. Interestingly, the enzyme 3-ketoacyl thiolase (*Ph*KAT), which catalyses benzoyl-CoA formation from 3-oxo-3-phenylpropionyl-CoA in the β-oxidative pathway, is also located in the peroxisomes (Van Moerkercke *et al.*, 2009). *Ph*KAT utilises 3-oxo-3-phenylpropionyl-CoA as a substrate, which is derived from cinnamoyl-CoA in a reaction catalysed by cinnamic acid-CoA hydratase/dehydrogenase (*Ph*CHD) (Qualley *et al.*, 2012).

Molecular modelling of *Md*CNL showed that all five structures belong to the α+β class of protein structures, with an overhanging C-terminus. When the five models were compared, it was found that the structural differences between the five models were minimal (the average RMSD between all five models was 0.6 Å). The docking analysis suggested that cinnamic acid was the most preferred substrate for the enzyme. Specifically, cinnamic acid presented a binding energy of −6.19 kcal mol^{-1} and 4-coumaric acid presented a binding energy of −6.14 kcal mol^{-1}, indicating negligible differences in terms of the predicted binding energy. While this finding appears to contradict the above results in terms of the number of conformations, it is well aligned with the experimental observation that cinnamic acid and 4-coumaric acid are preferred relative to the other potential substrates (i.e. caffeic acid, ferulic acid and sinapic acid). Similarly, the docking results for benzoic acid were in agreement with experimental results, where the lowest-energy conformation is not docked at the cavity of *Md*CNL.

One of the strategies adopted for disease management in apple is the genetic production of resistant cisgenic lines. The use of a virus-based system that utilises a plant episomal expression vector (Li *et al.*, 2012) was successful in increasing the resistance of the susceptible apple cultivar 'Gala Galaxy' via the insertion of a scab resistance gene (Vanblaere *et al.*, 2014; Cusin *et al.*, 2017). Similarly, homologous expression of apple CNL in susceptible cultivars is expected to produce resistant traits.

Biosynthesis of biphenyl phytoalexin in Malinae involves the intermediate formation of benzoyl-CoA; however, the biosynthesis of benzoyl-CoA remains poorly investigated in Malinae. Overall, this paper reports *Md*CNL-catalysed cinnamoyl-CoA formation as a required step for biphenyl phytoalexin biosynthesis in apple. This finding provides a basis for future cisgenic engineering of susceptible apple cultivars to fortify their scab resistance potential. Nevertheless, further studies are required to confirm the gain of resistant traits.

EXPERIMENTAL PROCEDURES

Plant material and chemicals

The apple (*Malus* × *domestica* Borkh.) cultivars 'Shireen', 'Prima', 'Golden Delicious' and 'Vista Bella' were collected from the Central Institute of Temperate Horticulture (CITH), Srinagar, India. Apple plants were grown in a greenhouse as described previously (Sarkate *et al.*, 2017). Unless otherwise noted, standards were purchased from Sigma-Aldrich Chemical Co. Ltd (https://www.sigmaaldrich.com/). Pure aucuparin and noraucuparin were chemically synthesised following a previously reported protocol (Saini *et al.*, 2017). Growth regulators and plant growth media were purchased from HiMedia (http://www.himedialabs.com/).

Treatment of cell suspension and greenhouse-grown plants

Primary callus cultures of 'Golden Delicious' were derived from young leaves. Cell cultures were subsequently developed from those calli (Sarkate *et al.*, 2017). The VIE was applied to 8-day-old cell cultures following the published protocol (Sarkate *et al.*, 2018). After VIE treatment, cell cultures were extracted at particular post-elicitation time points (0–72 hpe) to estimate the quantity of biphenyl phytoalexins and transcript expression levels. Cells without VIE were treated with an equal volume of sterile distilled water.

To check scab-induced phytoalexin accumulation in intact plants, greenhouse-grown apple plants were infected with the scab fungus as described previously (Sarkate *et al.*, 2018). For scab infection, two scab-resistant (cv. 'Shireen' and 'Prima'), one moderately scab susceptible (cv. 'Golden Delicious') and one highly scab susceptible (cv. 'Vista Bella'), apple cultivars were selected. Infected shoots (internode region between two successive infected leaves) of cv. 'Shireen' and cv. 'Golden Delicious' at various stages were stored for the RNA isolation and phytoalexin analyses (0–480 hpi). Infected plants were routinely photographed until 20 days post-infection to capture disease symptoms. Infected

shoots of cv. 'Prima' and 'Vista Bella' were only used for analysing phytoalexin accumulation at defined time points (Table S1) to correlate phytoalexin accumulation with scab resistance.

Extraction and investigation of biphenyl formation

Elicitor-treated cell cultures were collected at 0, 3, 6, 9, 12, 24, 36, 48 and 72 hpe and kept at −80°C until further analyses. Frozen cell cultures (about 2 g) and infected shoots were used for the estimation of phytoalexin formation, as described previously (Saini *et al.*, 2017). Biphenyl contents were analysed by HPLC as described previously (Teotia *et al.*, 2016).

Isolation and cloning of an apple cDNA encoding a CoA-ligase

The genome sequence (v.0.1 cDNA) of apple (Velasco *et al.*, 2010) was screened for putative CNL sequences involved in the conversion of cinnamic acid to cinnamoyl-CoA via the TBLASTN server of the GDR (https://www.rosaceae.org/), using available published plant cinnamate-specific CoA-ligase sequences as queries (Table S2). Based on bioinformatic analyses and sequence identity, the apple unigene MDP0000576682 was selected as a promising candidate for the design of 5′ and 3′ primers for the amplification of the CDS of a putative *Md*CNL cDNA prepared from the VIE-treated 'Golden Delicious' cell culture (Table S2). Total RNA was extracted, and cDNA was synthesised following the protocols of Sarkate *et al.* (2018). The *MdCNL* coding sequence was proofread from the apple cell culture cDNA pool (6 hpe) using Phusion High-Fidelity DNA polymerase (NEB, https://www.neb.com/). The forward and reverse primers (1 and 2 with *Nhe*I and *Kpn*I restriction sites, respectively; Table S3) were used to proofread *Md*CNL. The PCR program was set as described previously by Sarkate *et al.* (2018) except that the annealing temperature was set to 65°C. The amplified PCR product (1725 bp) was purified with the GenElute™ PCR product clean-up kit (Sigma-Aldrich). The 3′ UTR was PCR amplified using gene-specific forward primer 3 (targeting a site 88 bp upstream of the stop codon), following the 3′ rapid amplification of cDNA ends (RACE) protocol of the SMART RACE cDNA Amplification Kit (Clontech, https://www.takarabio.com/) (Table S3). Genomic DNA was extracted using the DNeasy Mini Kit (Qiagen, https://www.qiagen.com/) and employed to obtain the 5′ UTR sequence using gene-specific forward and reverse primers 4 and 5 (targeting sites 240 and 131 bp upstream and downstream of the start codon of MDP0000576682, respectively) (Table S3). In the case of greenhouse-grown apple plants, a reverse-transcribed RNA prepared from 48 h *V. inaequalis*-inoculated internodal segments of the shoots (Sarkate *et al.*, 2018) was used as a template to amplify the *Md*CNL CDS using the aforementioned PCR program.

Heterologous expression of *Md*CNL in *Escherichia coli*

The *Nhe*I/*Kpn*I-digested PCR product of the *Md*CNL open reading frame (ORF) was ligated into the *Nhe*I/*Kpn*I-linearised pRSET B expression vector (Invitrogen, part of ThermoFisher, http://www.thermofisher.com). After sequence verification of the ORF from both strands of the *Md*CNL-pRSETB construct, the coding sequence of N-terminal His_6-tagged *Md*CNL was expressed in *E. coli* BL21 (DE3)-RIL cells (developed by Stratagene and distributed by Fisher Scientific, https://www.fishersci.com). The induction of expression of the recombinant *Md*CNL protein (64 kDa) and its affinity purification in Ni-NTA agarose (Qiagen) were performed according to the manufacturer's description. Sodium dodecyl sulphate-PAGE was routinely performed to check the purification efficiency.

*Md*CNL activity assay

*Md*CNL activity was tested by the spectrophotometric method, where activity was calculated by measuring the rate of increase in the absorbance at 311 nm due to the formation of cinnamoyl-CoA, as described previously (Gaid *et al.*, 2012). The reaction was initiated at 25°C by the addition of 0.2 mM CoA, and cinnamoyl-CoA formation was monitored by recording the change in absorbance at 311 nm up to 5 min. Then CNL activity was calculated using the extinction coefficient of cinnamoyl-CoA (ε = 22 cm^{-1} mM^{-1}; Gaid *et al.*, 2012). In separate assays, a range of phenolic acids (4-coumaric, caffeic, ferulic and sinapic acids) with related structures were tested as potential substrates for *Md*CNL, keeping the other assay conditions unchanged. The substrate preference was calculated using previously published ε values of the corresponding CoA esters (Gaid *et al.*, 2011). For the determination of pH optima, values from 5.5 to 11.5 were tested. Other incubation parameters were also evaluated, including incubation temperatures and times and protein concentrations ranging from 20 to 60°C, 1 to 20 min and 1 to 40 μg protein, respectively. Enzyme stability was tested by placing the protein at 4 or −80°C for 24 h. Stability was further tested by storing recombinant protein at −80°C for 6 months with the addition of glycerol (20%, v/v).

Analyses of enzymatic product by ESI-MS

See Method S1.

Substrate specificity test by luciferase assay

Due to limitations of the ability of a spectrophotometric assay to screen the pattern of substrate preference among a wide range of potential CNL substrates we carried out a luciferase-based assay, exactly as described in Gaid *et al.* (2012). The data obtained were expressed using the remaining ATP concentration, which is relative to the luciferase activity but inversely correlates with the ATP used in the CoA-ligase assay (Gaid *et al.*, 2012). Quantification of ATP was normalised by a negative control reaction containing heat-denatured *Md*CNL and without any substrate. This reaction failed to utilise any ATP and thereby its ATP content (luciferase activity) was considered as 100%, which corresponded to zero CoA-ligase activity.

Analyses of the kinetic parameters of *Md*CNL

The kinetic constants were determined using different concentrations of cinnamic and 4-coumaric acids (1–1000 μM) at fixed concentrations of ATP (2.5 mM), CoA (0.2 mM) and $MgCl_2$ (2.5 mM). To determine the K_m value for ATP, its concentration was varied from 1 to 1000 μM keeping a fixed concentration of cinnamic acid (0.4 mM), CoA (0.2 mM) and $MgCl_2$ (2.5 mM). Similarly, the K_m value for CoA was calculated by varying the concentration from 1 to 1000 μM keeping the concentrations of cinnamic acid (0.4 mM), ATP (2.5 mM) and $MgCl_2$ (2.5 mM) fixed. Three biological repeats were used to estimate the kinetic parameters by the Hanes plot method using the Hyper 32 program (https://hyper32.software.informer.com/).

Phylogeny analyses

Functional plant CoA-ligase sequences related to CNL reactions were used to reconstruct a phylogenetic tree. The sequence accession numbers are listed in Table S4. The neighbour-joining algorithm was followed to create phylogenetic tree using MEGA 5.05 software (https://www.megasoftware.net/) with 1000 bootstrap

supports (Tamura *et al.*, 2011). Pair-wise deletion was used to eliminate gaps and missing data.

Quantitative real-time PCR

Tissue samples (100 mg) extracted from VIE-treated cell cultures of 'Golden Delicious' and shoots of cv. 'Shireen' and cv. 'Golden Delicious' infected with scab fungus were used to isolate the total RNA. Total RNA (2 μg) was reverse-transcribed at 42°C for 60 min to synthesise the cDNA. Quantitative real-time PCR was performed in a QuantStudio 3.0 system (Thermo Fisher Scientific, https://www.thermofisher.com/) using PowerUp™ SYBR™ Green Master Mix following the protocols described by Sarkate *et al.* (2018). Differential expression levels of *MdPAL*, *MdCNL* and *biphenyl synthase* isoform 3 (*MdBIS3*) genes were tested in the VIE-treated 'Golden Delicious' cell cultures at defined post-elicitation time points (Method S2; Pfaffl, 2001).

Identification of the molecular weight of *Md*CNL by MALDI-MS

The purified recombinant protein was eluted and deposited onto the target plate using sinapic acid (Sigma-Aldrich) matrix prepared in 70% acetonitrile (MS grade Sigma-Aldrich) and 0.1% TFA (Loba Chemie, https://www.lobachemie.com/). Consequently, the sample was analysed using a Bruker ultrafleXtreme MALDI-TOF/TOF mass spectrometer (Bruker Daltonics, https://www.bruker.com/) coupled with a Smartbeam-II laser fitted in linear positive mode. Spectrum acquisition was done with FlexControl 3.4 and data analysis was performed using and FlexAnalysis 3.4 software (Bruker Daltonics).

Subcellular localisation

For subcellular localisation studies of *Md*CNL, two separate fusion constructs bearing N- and C-terminal YFP fusions were cloned using Gateway cloning system (Invitrogen). Fusion constructs were transiently expressed in *N. benthamiana* leaves. To generate N-terminal YFP fusion of *Md*CNL, PCR amplification of the *Md*CNL ORF was carried out using a forward primer *att*B1 (primer 14) and a reverse primer *att*B2a (primer 15) bearing a stop codon. Similarly, C-terminal YFP fusion of *Md*CNL was generated by PCR amplification of the *Md*CNL ORF with forward primer *att*B1 (primer 14) and reverse primer *att*B2b (primer 16) without a stop codon (primer details are listed in Table S3). N- and C-terminal fusions of *Md*CNL with YFP were subjected to subcellular localisation studies in *N. benthamiana* leaves using confocal microscopy (Method S3).

Homology modelling

The structure of *Md*CNL was predicted by the Robetta server (Kim *et al.*, 2004) using an *ab initio* method which implements the Rosetta algorithm. The *Md*CNL protein sequence was provided as an input to the Robetta server for computational structure prediction. Five models predicted by Robetta were subjected to superimposition using Theseus (Theobald and Steindel, 2012) to identify the median structure using the maximum likelihood method. Molecular dynamics simulation of the structure obtained from Theseus was performed in GROMACS v.5.1.2 (Berendsen *et al.*, 1995; Pronk *et al.*, 2013) for 25 ns to validate and obtain an equilibrated structure of *Md*CNL (Method S4).

Statistical analyses

Statistical analyses of all data were carried out using GraphPad Prism (v.4.0). The data obtained from three or more independent experiments are presented as the mean values ± standard deviations, and statistical significance was determined using the one-way analysis of variance (ANOVA) tool of GraphPad Prism. Independent *t*-tests were performed to evaluate the differences between two consecutive groups ($P < 0.05$ was considered statistically significant).

ACCESSION NUMBER

*Md*CNL sequence information is available in the data libraries (GenBank/EMBL) under the accession number MG334585.

ACKNOWLEGEMENTS

This work was supported by a research grant (NASF/ABP 6023/2017-18 to DS) from the National Agricultural Science Fund (ICAR-NASF), India. DT acknowledges the Council of Scientific and Industrial Research (CSIR) for the award of a junior research fellowship (CSIR-JRF). SSS is grateful to the Indian Institute of Technology Roorkee for providing a MHRD scholarship.

AUTHOR CONTRIBUTIONS

DT and MG performed cloning and functional characterisation of CNL. SSS performed *in vivo* studies and statistical analyses. AV and KA performed MALDI-TOF studies. RMY and SPK performed molecular modelling and docking studies and both contributed equally. JIM provided apple samples and performed scab infection studies. TB performed ESI-MS analyses. RH and PR helped in the analyses of microscopy images. DS and MG wrote the manuscript with support from DT, SS, KA, PR, RH, JIM and SPK. LB and PKA provided critical feedback on data analyses. DS conceived the original idea, secured funding for this research, designed experiments and supervised the entire work. DS and MG critically revised the manuscript.

CONFLICTS OF INTEREST

Authors declare that there are no conflicts of interest.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Table S1. Accumulation of the biphenyl phytoalexins (noraucuparin and aucuparin) in the internodes of *V. inaequalis* conidia-infected different apple cultivars agrees with the reported levels of scab tolerance. Biphenyl phytoalexins were measured 10 days after *V. inaequalis* conidia infection. Results are means ± SD ($n = 3$).

Table S2. Homology between apple unigenes present in GDR and aromatic-CoA ligase sequences.

Table S3. Primers used.

Table S4. Accession numbers of amino acid sequences used for phylogenetic reconstruction.

Figure S1. HPLC analysis showing the accumulation of the biphenyl phytoalexins noraucuparin (NA) and aucuparin (A) in VIE-treated cell cultures of apple cv. 'Golden Delicious'. (a) Reference compounds (R1: noraucuparin; R2: aucuparin); (b) Profile of non-elicited cells after 24 h; (c) Phytoalexin profile of VIE-treated cells after 24 h. The chromatograms were monitored at 269 nm.

Figure S2. Molecular mass determination of the *Md*CNL by MALDI-TOF. The average molecular mass of the recombinant *Md*CNL protein was observed at 64.3 kDa (asterisk). The inset shows a sodium dodecyl sulphate (SDS) polyacrylamide gel for the purification of recombinant *Md*CNL (lane 1: Crude protein from un-induced bacterial cultures; lane 2: crude protein prepared from IPTG-induced cultures; lane 3: Ni-NTA affinity-purified CNL; lane M: molecular marker). Eluted CNL protein is indicated by asterisk.

Figure S3. ESI-MS/MS analysis of CNL-catalyzed cinnamoyl-CoA formation, represented in the mass spectrum by the molecular ion $[M\text{-}H]^-$ at m/z 896.2. MS/MS analysis showed characteristic fragments of phosphoadenosine-containing compounds as indicated by boxes at m/z 408.1 and 426.2 (Beuerle and Pichersky, 2002b).

Figure S4. Progression of scab symptoms in *V. inaequalis*-infected leaves of apple cv. 'Shireen' (a) and cv. 'Golden Delicious' (b) after 0, 6 and 16 days of inoculation with conidiospores.

Figure S5. Root mean square deviation (RMSD) analysis of *Md*CNL. The RMSD of *Md*CNL in the trajectory was analyzed for the entire protein, where the protein became stabilized and equilibrated after ~5ns.

Figure S6. Root mean square fluctuation analysis of *Md*CNL. The root mean square fluctuation of *Md*CNL in the trajectory was analyzed indicating that the C-terminus of the protein is more flexible compared to the rest of the protein.

Figure S7. Radius of gyration analysis of *Md*CNL showing that the protein does not undergo unfolding as the Rg is fluctuating between 2.5-2.6 nm.

Method S1. Analyses of the enzymatic product by electrospray ionization-mass spectrometry (ESI-MS).

Method S2. Quantitative real-time PCR.

Method S3. *Agrobacterium*-mediated transient expression and subcellular localisation.

Method S4. Homology modelling of *Malus domestica* cinnamate-CoA ligase.

REFERENCES

Abd El-Mawla, A.M. and Beerhues, L. (2002) Benzoic acid biosynthesis in cell cultures of *Hypericum androsaemum*. *Planta*, **214**, 727–733.

Barillas, W. and Beerhues, L. (2000) 3-Hydroxybenzoate:coenzyme A ligase from cell cultures of Centaurium erythraea: Isolation and characterization. *Biol. Chem.* **381**, 155–160.

Bartsch, M., Bednarek, P., Vivancos, P.D., Schneider, B., Von Roepenack-Lahaye, E., Foyer, C.H., Kombrink, E., Scheel, D. and Parker, J.E. (2010) Accumulation of isochorismate-derived 2,3-dihydroxybenzoic 3-O-β-D-xyloside in Arabidopsis resistance to pathogens and ageing of leaves. *J. Biol. Chem.* **285**, 25654–25665.

Berendsen, H.J.C., Van der Spoel, D. and Van Drunen, R. (1995) GROMACS: a message-passing parallel molecular dynamics implementation. *Comput. Phys. Commun.* **91**, 43–56.

Beuerle, T. and Pichersky, E. (2002a) Purification and characterization of benzoate:coenzyme A ligase from *Clarkia breweri*. *Arch. Biochem. Biophys.* **400**, 258–264.

Beuerle, T. and Pichersky, E. (2002b) Enzymatic synthesis and purification of aromatic coenzyme A esters. *Anal. Biochem.* **302**, 305–312.

Biggs, A.R., Sundin, G.W., Rosenberger, D.A., Yoder, K.S. and Sutton, T.B. (2010) Relative susceptibility of selected apple cultivars to apple scab caused by *Venturia inaequalis*. *Plant Health Prog.* **11**, 20.

Bussell, J.D., Reichelt, M., Wiszniewski, A.A.G., Gershenzon, J. and Smith, S.M. (2014) Peroxisomal ATP-binding cassette transporter COMATOSE and the multifunctional protein ABNORMAL INFLORESCENCE MERISTEM are required for the production of benzoylated metabolites in Arabidopsis seeds. *Plant Physiol.* **164**, 48–54.

Carisse, O. and Bernier, J. (2002) Effect of environmental factors on growth, pycnidial production and spore germination of Microsphaeropsis isolates with biocontrol potential against apple scab. *Mycol. Res.* **106**, 1455–1462.

Catinot, J., Buchala, A., Abou-Mansour, E. and Métraux, J.P. (2008) Salicylic acid production in response to biotic and abiotic stress depends on isochorismate in *Nicotiana benthamiana*. *FEBS Lett.* **582**, 473–478.

Chizzali, C. and Beerhues, L. (2012) Phytoalexins of the Pyrinae: biphenyls and dibenzofurans. *Beilstein J. Org. Chem.* **8**, 613–620.

Chizzali, C., Khalil, M.N.A., Beuerle, T., Schuehly, W., Richter, K., Flachowsky, H., Peil, A., Hanke, M.V., Liu, B. and Beerhues, L. (2012a) Formation of biphenyl and dibenzofuran phytoalexins in the transition zones of fire blight-infected stems of *Malus domestica* cv. 'Holsteiner Cox' and *Pyrus communis* cv. 'Conference'. *Phytochemistry*, **77**, 179–185.

Chizzali, C., Gaid, M.M., Belkheir, A.K., Hansch, R., Richter, K., Flachowsky, H., Peil, A., Hanke, M.-V., Liu, B. and Beerhues, L. (2012b) Differential expression of biphenyl synthase gene family membersin fire-blight-infected apple 'Holsteiner Cox'. *Plant Physiol.* **158**, 864–875.

Colquhoun, T.A., Marciniak, D.M., Wedde, A.E., Kim, J.Y., Schwieterman, M.L., Levin, L.A., Van Moerkercke, A., Schuurink, R.C. and Clark, D.G. (2012) A peroxisomally localized acyl-activating enzyme is required for volatile benzenoid formation in a *Petunia*×*hybrida* cv. 'Mitchell Diploid' flower. *J. Exp. Bot.* **63**, 4821–4833.

Cusin, R., Revers, L.F. and Maraschin, F.S. (2017) New biotechnological tools to accelerate scab-resistance trait transfer to apple. *Genet. Mol. Biol.* **40**(1 suppl 1), 305–311.

Dewdney, M., Charest, J., Paulitz, T. and Carisse, O. (2003) Multivariate analysis of apple cultivar susceptibility to *Venturia inaequalis* under greenhouse conditions. *Can. J. Plant Pathol.* **25**, 387–400.

Fulda, M., Heinz, E. and Wolter, F.P. (1994) The fadD gene of *Escherichia coli* K12 is located close to rnd at 39.6 min of the chromosomal map and is a new member of the AMP-binding protein family. *Mol. Gen. Genet.* **242**, 241–249.

Gaid, M.M., Sircar, D., Beuerle, T., Mitra, A. and Beerhues, L. (2009) Benzaldehyde dehydrogenase from chitosan-treated *Sorbus aucuparia* cell cultures. *J. Plant Physiol.* **166**, 1343–1349.

Gaid, M.M., Scharnhop, H., Ramadan, H., Beuerle, T. and Beerhues, L. (2011) 4-Coumarate: CoA ligase family members from elicitor-treated *Sorbus aucuparia* cell cultures. *J. Plant Physiol.* **168**, 944–951.

Gaid, M.M., Sircar, D., Muller, A., Beuerle, T., Liu, B., Ernst, L., Hansch, R. and Beerhues, L. (2012) Cinnamate:CoA ligase initiates the biosynthesis of a benzoate-derived xanthone phytoalexin in *Hypericum calycinum* cell cultures. *Plant Physiol.* **160**, 1267–1280.

Gallage, N.J., Hansen, E.H., Kannangara, R., Olsen, C.E., Motawia, M.S., Jørgensen, K., Holme, I., Hebelstrup, K., Grisoni, M. and Møller, B.L. (2014) A single enzyme catalyses vanillin formation from ferulic acid in *Vanilla planifolia*. *Nat. Commun.* **5**, 4037.

Ibdah, M. and Pichersky, E. (2009) *Arabidopsis* Chy1 null mutants are deficient in benzoic acid-containing glucosinolates in the seeds. *Plant Biol.* **11**, 574–581.

Ishikura, N., Hayashida, S. and Tazaki, K. (1984) Biosynthesis of gallic and ellagic acids with ^{14}C-labeled compounds in *Acer* and *Rhus* leaves. *Bot. Mag.* **97**, 355–367.

Jha, G., Thakur, K. and Thakur, P. (2009) The *venturia* apple pathosystem: pathogenicity mechanisms and plant defense responses. *J. Biomed. Biotechnol.* **2009**, 1–10.

Kellerhals, M., Fouillet, A. and Lespinasse, Y. (1993) Effect of the scab inoculum and the susceptible parent on resistance to apple scab (*Venturia inaequalis*) in the progenies of crosses to the scab resistant cv'Florina'. *Agrono. EDP Sci.* **13**, 631–636.

Khajuria, Y.P., Kaul, S., Wafai, B.A. and Dhar, M.K. (2014) Screening of apple germplasm of Kashmir Himalayas for scab resistance genes. *Indian J. Biotechnol.* **13**, 448–454.

Khalil, M.N.A., Brandt, W., Beuerle, T., Reckwell, D., Groeneveld, J., Hänsch, R., Gaid, M.M., Liu, B. and Beerhues, L. (2015) O-Methyltransferases involved in biphenyl and dibenzofuran biosynthesis. *Plant J.* **83**, 263–276.

Kim, D.E., Chivian, D. and Baker, D. (2004) Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Res.* **32**, 526–531.

Klempien, A., Kaminaga, Y., Qualley, A., Nagegowda, D.A., Widhalm, J.R., Orlova, I., Shasany, A.K., Taguchi, G., Kish, C.M. and Cooper, B.R. (2012) Contribution of CoA ligases to benzenoid biosynthesis in *Petunia* flowers. *Plant Cell*, **24**, 2015–2030.

Krüger, J. (1991) Schorfbefall von Nachkommen ausKreuzungen mit der schorfresistenten Apfelsorte 'Prima'. *J. Plant Dis. Protect.* **98**, 73–76.

Kumar, A. and Ellis, B.E. (2003) 4-coumarate:CoA ligase gene family in *Rubusidaeus*: cDNA structures, evolution, and expression. *Plant Mol. Biol.* **51**, 327–340.

Lee, S., Kaminaga, Y., Cooper, B., Pichersky, E., Dudareva, N. and Chapple, C. (2012) Benzoylation and sinapoylation of glucosinolate R-groups in Arabidopsis. *Plant J.* **72**, 411–422.

Li, Y.Y., Mao, K., Zhao, C., Zhao, X.Y., Zhang, H.L., Shu, H.R. and Hao, Y.J. (2012) MdCOP1 Ubiquitin E3 Ligases interact with MdMYB1 to regulate light-induced anthocyanin biosynthesis and red fruit coloration in apple. *Plant Physiol.* **160**, 1011–1022.

Li, Y., Kim, J.I., Pysh, L. and Chapple, C. (2015) Four isoforms of *Arabidopsis* 4-Coumarate: CoA ligase have overlapping yet distinct roles in phenylpropanoid metabolism. *Plant Physiol.* **169**, 2409–2421.

Liu, B., Beuerle, T., Klundt, T. and Beerhues, L. (2004) Biphenyl synthase from yeast-extract-treated cell cultures of *Sorbus aucuparia*. *Planta*, **218**, 492–496.

Long, M.C., Nagegowda, D.A., Kaminaga, Y., Ho, K.K., Kish, C.M., Schnepp, J., Sherman, D., Weiner, H., Rhodes, D. and Dudareva, N. (2009) Involvement of snapdragon benzaldehyde dehydrogenase in benzoic acid biosynthesis. *Plant J.* **59**, 256–265.

Moreno, P.R.H., Van der Heijden, R. and Verpoorte, R. (1994) Elicitor-mediated induction of isochorismate synthase and accumulation of 2,3-dihydroxybenzoic acid in *Catharanthus roseus* cell suspension and shoot cultures. *Plant Cell, Rep.* **14**, 188–191.

Mustafa, N.R., Kim, H.K., Choi, Y.H., Erkelens, C., Lefeber, A.W.M., Spijksma, G., Van der Heijden, R. and Verpoorte, R. (2009) Biosynthesis of salicylic acid in fungus elicited *Catharanthus roseus* cells. *Phytochemistry*, **70**, 532–539.

Nowak, K., Luniak, N., Meyer, S., Schulze, J., Mendel, R.R. and Hänsch, R. (2004) Fluorescent proteins in poplar: a useful tool to study promoter function and protein localization. *Plant Biol.* **6**, 65–73.

Pfaffl, M.W. (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* **29**, e45.

Pronk, S., Páll, S., Schulz, R., Larsson, P., Bjelkmar, P., Apostolov, R., Shirts, M.R., Smith, J.C., Kasson, P.M. and Van der Spoel, D. (2013) GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics*, **29**, 845–854.

Qualley, A.V., Widhalm, J.R., Adebesin, F., Kish, C.M. and Dudareva, N. (2012) Completion of the core -oxidative pathway of benzoic acid biosynthesis in plants. *Proc. Natl Acad. Sci. USA*, **109**, 16383–16388.

Rottensteiner, H., Kramer, A., Lorenzen, S., Stein, K., Landgraf, C., Volkmer-Engert, R. and Erdmann, R. (2004) Peroxisomal membrane proteins contain common Pex19p-binding sites that are an integral part of their targeting signals. *Mol. Biol. Cell*, **15**, 3406–3417.

Saini, S.S., Teotia, D., Gaid, M., Thakur, A., Beerhues, L. and Sircar, D. (2017) Benzaldehyde dehydrogenase-driven phytoalexin biosynthesis in elicitor-treated *Pyrus pyrifolia* cell cultures. *J. Plant Physiol.* **215**, 154–162.

Sarkate, A., Banerjee, S., Mir, J.I., Roy, P. and Sircar, D. (2017) Antioxidant and cytotoxic activity of bioactive phenolic metabolites isolated from the yeast-extract treated cell culture of apple. *Plant Cell Tissue Organ Cult.* **130**, 641–649.

Sarkate, A., Saini, S.S., Teotia, D., Gaid, M., Mir, J.I., Roy, P., Agrawal, P.K. and Sircar, D. (2018) Comparative metabolomics of scab-resistance and susceptible apple cell cultures in response to scab fungus elicitor treatment. *Sci. Rep.* **8**, 17844.

Sarkate, A., Saini, S.S., Gaid, M., Teotia, D., Mir, J.I., Agrawal, P.K., Beerhues, L. and Sircar, D. (2019) Molecular cloning and functional analysis of a biphenyl phytoalexin-specific O-methyltransferase from apple cell suspension cultures. *Planta*, **249**, 677–691.

Schneider, K., Hövel, K., Witzel, K., Hamberger, B., Schomburg, D., Kombrink, E. and Stuible, H. P. (2003) The substrate specificity-determining amino acid code of 4-coumarate:CoA ligase. *Proc. Natl Acad. Sci. USA*, **100**, 8601–8606.

Shafi, S.M., Sheikh, M.A., Nabi, S.U., Mir, M.A., Ahmad, N., Mir, J.I., Raja, W.H., Rasool, R. and Masoodi, K.Z. (2019) An overview of apple scab, its cause and management strategies. *EC Microbiol.* **15**, 283–287.

Sieweke, H.-J. and Leistner, E. (1992) O-succinylbenzoate: coenzyme a ligase from anthraquinone producing cell suspension cultures of *Galium mollugo*. *Phytochemistry*, **31**, 2329–2335.

Sircar, D., Gaid, M., Chizzali, C., Reckwell, D., Kaufholdt, D., Beuerle, T., Broggini, G.A.L., Flachowsky, H., Liu, B. and Hänsch, R. (2015) Biphenyl 4-hydroxylases involved in aucuparin biosynthesis in Rowan and Apple are CYP736A proteins. *Plant Physiol.* **168**, 428–442.

Song, Y., Di Maio, F., Wang, R.Y.-R., Kim, D., Miles, C., Brunette, T., Thompson, J. and Baker, D. (2013) High-resolution comparative modeling with Rosetta CM. *Structure*, **21**, 1735–1742.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* **28**, 2731–2739.

Teotia, D., Saini, S.S., Gaid, M., Beuerle, T., Beerhues, L. and Sircar, D. (2016) Development and validation of a new HPLC method for the determination of biphenyl and dibenzofuran phytoalexins in rosaceae. *J. Chromatogr. Sci.* **54**, 918–922.

Theobald, D.L. and Steindel, P.A. (2012) Optimal simultaneous superpositioning of multiple structures with missing data. *Bioinformatics*, **28**, 1972–1979.

Thomson, S. (2000) Epidemiology of fire blight. In *Fire blight: the disease and its causative agent, Erwinia amylovora*. (Vanneste, J. L., ed). Wallingford: CAB International, pp. 9–36.

Van Moerkercke, A., Schauvinhold, I., Pichersky, E., Haring, M.A. and Schuurink, R.C. (2009) A plant thiolase involved in benzoic acid biosynthesis and volatile benzenoid production. *Plant J.* **60**, 292–302.

Vanblaere, T., Flachowsky, H., Gessler, C. and Broggini, G.A.L. (2014) Molecular characterization of cisgenic lines of apple 'Gala' carrying the Rvi6 scab resistance gene. *Plant Biotechnol. J.* **12**, 2–9.

Velasco, R., Zharkikh, A., Affourtit, J., Dhingra, A., Cestaro, A., Kalyanaraman, A., Fontana, P., Bhatnagar, S.K., Troggio, M. and Pruss, D. (2010) The genome of the domesticated apple (*Malus x domestica* Borkh.). *Nat. Genet.* **42**, 833–839.

Wang, C.Z., Maier, U.H., Eisenreich, W., Adam, P., Obersteiner, I., Keil, M., Bacher, A. and Zenk, M.H. (2001) Unexpected biosynthetic precursors of amarogentin - a retrobiosynthetic ^{13}C NMR study. *Eur. J. Org. Chem.* **8**, 1459–1465.

Werner, I., Bacher, A. and Eisenreich, W. (1997) Retrobiosynthetic NMR studies with ^{13}C-labeled glucose. Formation of gallic acid in plants and fungi. *J. Biol. Chem.* **272**, 25474–25482.

Widhalm, J.R. and Dudareva, N. (2015) A familiar ring to it: biosynthesis of plant benzoic acids. *Mol. Plant*, **8**, 83–97.

