PNAS RESEARCH ARTICLE | BIOCHEMISTRY OPEN ACCESS

Biosynthesis of iridoid sex pheromones in aphids

Tobias G. Köllner[a,1], Anja David[a], Katrin Luck[a], Franziska Beran[b], Grit Kunert[c], Jing-Jiang Zhou[d], Lorenzo Caputi[a], and Sarah E. O'Connor[a,1]

Edited by Richard Dixon, University of North Texas, Denton, TX; received June 30, 2022; accepted September 9, 2022

Iridoid monoterpenes, widely distributed in plants and insects, have many ecological functions. While the biosynthesis of iridoids has been extensively studied in plants, little is known about how insects synthesize these natural products. Here, we elucidated the biosynthesis of the iridoids *cis-trans*-nepetalactol and *cis-trans*-nepetalactone in the pea aphid *Acyrthosiphon pisum* (Harris), where they act as sex pheromones. The exclusive production of iridoids in hind legs of sexual female aphids allowed us to identify iridoid genes by searching for genes specifically expressed in this tissue. Biochemical characterization of candidate enzymes revealed that the iridoid pathway in aphids proceeds through the same sequence of intermediates as described for plants. The six identified aphid enzymes are unrelated to their counterparts in plants, conclusively demonstrating an independent evolution of the entire iridoid pathway in plants and insects. In contrast to the plant pathway, at least three of the aphid iridoid enzymes are likely membrane bound. We demonstrated that a lipid environment facilitates the cyclization of a reactive enol intermediate to the iridoid cyclopentanoid-pyran scaffold in vitro, suggesting that membranes are an essential component of the aphid iridoid pathway. Altogether, our discovery of this complex insect metabolic pathway establishes the genetic and biochemical basis for the formation of iridoid sex pheromones in aphids, and this discovery also serves as a foundation for understanding the convergent evolution of complex metabolic pathways between kingdoms.

iridoids | aphids | pathway | sex pheromone | biosynthesis

Significance

Plants, animals, and microbes produce a plethora of natural products that are important for defense and communication. Most of these compounds show a phylogenetically restricted occurrence, but, in rare instances, the same natural product is biosynthesized by organisms in two different kingdoms. The monoterpene-derived iridoids, for example, have been found in more than 50 plant families but are also observed in several insect orders. The discovery of the aphid iridoid pathway, one of the longest and most chemically complex insect-derived natural product biosynthetic pathways reported to date, highlights the mechanisms underlying the convergent evolution of metabolic enzymes in insects and plants, including the recruitment of different enzyme classes to catalyze the same chemical processes.

Iridoids are a class of atypical bicyclic monoterpenoids that are widely distributed in flowering plants, but, notably, are also found in several insect orders, including Coleoptera, Hymenoptera, and Hemiptera (1). Iridoids therefore present an opportunity to compare and contrast the chemical logic of natural product biosynthesis between plants and insects.

In plants, iridoids largely act as defensive metabolites or biosynthetic intermediates for other natural products (e.g., monoterpenoid indole alkaloids and isoquinoline alkaloids). The pathway leading to the cyclopentanoid-pyran (iridoid) scaffold was first elucidated in the plant Madagascar periwinkle (*Catharanthus roseus*) (2–6) and more recently in the two mint species *Nepeta mussinii* and *Nepeta cataria* (7–9). Iridoid biosynthesis in plants starts with the condensation of the universal terpene precursors isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP) to form geranyl diphosphate (GPP), followed by hydrolysis to geraniol (Fig. 1*A*). Both reactions take place in the plastids and are catalyzed by *trans*-isoprenyl diphosphate synthase (IDS) and geraniol synthase (GES), respectively. Hydroxylation of geraniol by geraniol-8-hydroxylase (G8H) leads to 8-hydroxygeraniol, which is further oxidized in two consecutive reaction steps by 8-hydroxygeraniol oxidase (HGO) to 8-oxogeranial. This dialdehyde is then converted to the iridoid nepetalactol by a two-step reduction–cyclization sequence that involves the formation of a highly reactive 8-oxocitronellyl enol/enolate intermediate. Initially, reduction and cyclization of 8-oxogeranial were thought to be controlled by a single enzyme, iridoid synthase (ISY) (3), though later studies showed that ISY likely catalyzes only the NADPH-dependent reduction of 8-oxogeranial to the 8-oxocitronellyl enol/enolate intermediate (8). This intermediate can nonenzymatically cyclize, or, alternatively, the stereoselective cyclization of this intermediate to nepetalactol is enzymatically mediated by nepetalactol-related short-chain dehydrogenase (NEPS) or by major latex protein-like (MLPL) enzymes (8, 9). In *C. roseus*, nepetalactol is further metabolized to secologanin, which serves as a precursor for the formation of monoterpene indole alkaloids in this plant (10). In *Nepeta*, a NEPS protein oxidizes nepetalactol to nepetalactone (8), with both the alcohol and lactone released as volatiles.

Insects utilize iridoids as both defense compounds and volatile pheromones, but in terms of biosynthesis, comparatively little is understood about insect-derived iridoids. Biosynthetic insights have been obtained from studies on larvae of chrysomelid leaf beetles, which accumulate the iridoid-related monocyclic dialdehydes chrysomelidial

Author affiliations: [a]Department of Natural Product Biosynthesis, Max Planck Institute for Chemical Ecology, 07745 Jena, Germany; [b]Department of Insect Symbiosis, Max Planck Institute for Chemical Ecology, 07745 Jena, Germany; [c]Department of Biochemistry, Max Planck Institute for Chemical Ecology, 07745 Jena, Germany; and [d]State Key Laboratory of Green Pesticide and Agricultural Bioengineering, Guizhou University, Guiyang, 550025 China

Author contributions: T.G.K., J.-J.Z., L.C., and S.E.O. designed research; T.G.K., A.D., K.L., F.B., and G.K. performed research; T.G.K. analyzed data; and T.G.K. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

Copyright © 2022 the Author(s). Published by PNAS. This open access article is distributed under Creative Commons Attribution License 4.0 (CC BY).

See online for related content such as Commentaries.

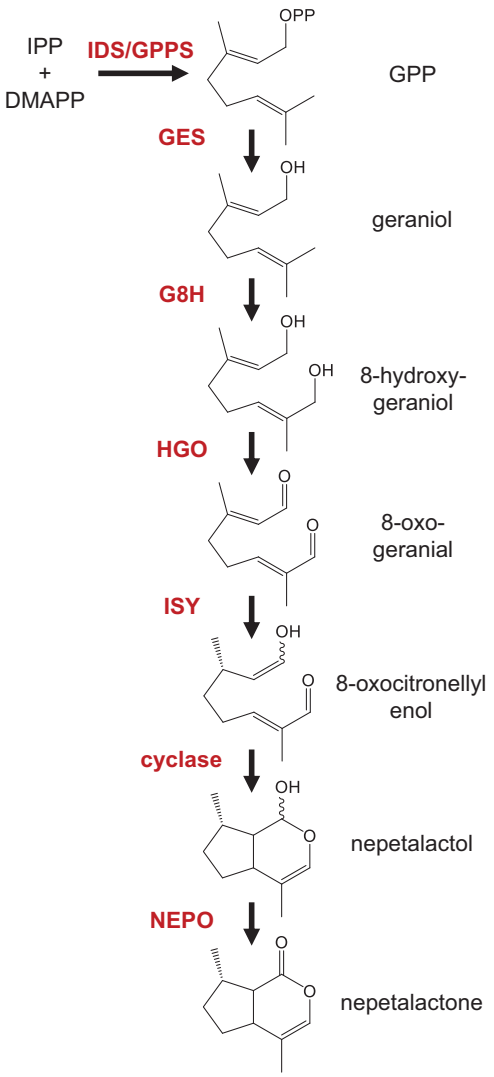
[1]To whom correspondence may be addressed. Email: koellner@ice.mpg.de or oconnor@ice.mpg.de.

This article contains supporting information online at http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2211254119/-/DCSupplemental.

Published October 13, 2022.

Downloaded from https://www.pnas.org by 95.90.176.231 on January 27, 2023 from IP address 95.90.176.231.

A Iridoid pathway in plants and aphids



B Relative expression of iridoid pathway genes in *A. pisum*

gene ID	gene annotation	f-hl	f-fl	af-hl	m-hl
100162815	acetyl-CoA acetyltransferase	1.00	0.01	0.01	0.01
100165154	HMG-CoA synthase	1.00	0.00	0.00	0.00
100165462	HMG-CoA reductase	1.00	0.01	0.01	0.01
100163305	mevalonate kinase	1.00	0.10	0.08	0.07
100574505	phosphomevalonate kinase	1.00	0.02	0.02	0.03
100163413	phosphomevalonate kinase-like	1.00	0.02	0.02	0.03
100158798	diphosphomevalonate decarboxylase	1.00	0.01	0.02	0.02
100166744	isopentenyl diphosphate isomerase	1.00	0.02	0.01	0.02
100144905	**isoprenyl diphosphate synthase (ApIDS)**	1.00	0.05	0.03	0.04
100158803	**dolichyldiphosphatase 1 (ApGES)**	1.00	0.05	0.08	0.07
100162683	inositol polyphosphate 1-phosphatase	1.00	0.05	0.04	0.05
100165972	**P450 (ApG8H)**	1.00	0.00	0.01	0.01
100160284	**P450 reductase (ApRed)**	1.00	0.11	0.20	0.20
100301633	**farnesol dehydrogenase (ApHGO)**	1.00	0.02	0.10	0.12
100162094	retinol dehydrogenase	1.00	0.13	0.01	0.00
103310029	**polyprenol reductase (ApISY)**	1.00	0.00	0.00	0.00
100169582	**GMC oxidase (ApNEPO)**	1.00	0.00	0.00	0.00
100164798	GMC oxidase	1.00	0.00	0.00	0.00
100168586	flavin reductase	1.00	0.12	0.12	0.12

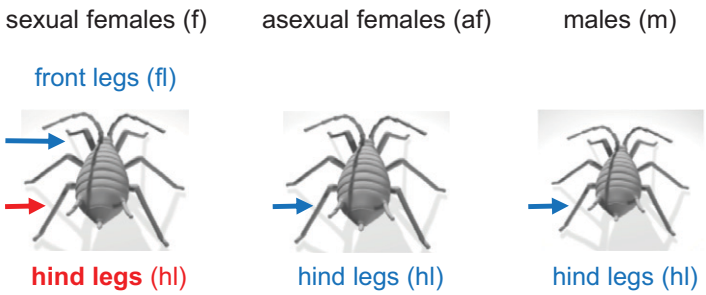


Fig. 1. The formation of iridoids in plants and aphids. (*A*) Labeling studies suggest that the biosynthesis of iridoids in the pea aphid *A. pisum* mimics the biosynthetic pathway in iridoid-producing plants. IPP, isopentenyl diphosphate; DMAPP, dimethylallyl diphosphate; GPP, geranyl diphosphate; IDS, isoprenyl diphosphate synthase; GES, geranyl diphosphate synthase; G8H, geraniol 8-hydroxylase; HGO, 8-hydroxygeraniol oxidoreductase; ISY, iridoid synthase; NEPO, nepetalactol oxidase. (*B*) Relative expression of mevalonate and putative nepetalactone pathway genes in hind legs and front legs of different sexual stages of *A. pisum*. Relative expression data are based on RPKM values obtained by RNAseq. f-hl, hind legs of sexual females; f-fl, front legs of sexual females; af-hl, hind legs of asexual females; m-hl, hind legs of males.

and plagiodial (11). Feeding experiments with isotopically labeled precursors and the discovery of some of the enzymes involved in chrysomelidial formation demonstrated that leaf beetles produce these compounds by a series of chemical reactions similar to those that occur in plants (12–15). Although the enzymatic basis for this pathway has not been completely established, the fact that the known enzymes are unrelated to their counterparts in plants suggests independent evolution of the pathway occurred (14).

Cis-trans-nepetalactol and *cis-trans*-nepetalactone are the major iridoids produced by catnip (*N. mussinii*) and catmint (*N. cataria*) (16). These molecules are responsible for the euphoric effect these plants have on cats, but their ecological function is unclear, though they may play roles in mediating interactions with insects (17). Interestingly, *cis-trans*-nepetalactol and *cis-trans*-nepetalactone occur also in aphids, which produce these compounds as volatile sex pheromones (18, 19). The pea aphid *Acyrthosiphon pisum*, for example, has been reported to biosynthesize (1*R*,4*aS*,7*S*,7*aR*)-*cis-trans*-nepetalactol and (4*aS*,7*S*,7*aR*)-*cis-trans*-nepetalactone in glandular structures on the hind legs of sexual female aphids, from where they are released to attract male conspecifics (18, 20). Recent studies with isotopically labeled iridoid precursors suggest that the iridoid pathway in aphids follows the reaction sequence described for plants (21). However, the underlying enzymatic machinery of this pathway is completely unknown.

Here, we report the elucidation of the entire iridoid pathway in the pea aphid *A. pisum*. By searching for genes expressed exclusively in hind legs of sexual female aphids, the site of iridoid production, we could rapidly identify all six biosynthetic genes/enzymes responsible for the conversion of IPP and DMAPP to *cis-trans*-nepetalactone. The discovery of the insect nepetalactone pathway in its entirety now allows a comparison of the chemical solutions that have evolved for nepetalactone

Downloaded from https://www.pnas.org by 95.90.176.231 on January 27, 2023 from IP address 95.90.176.231.

biosynthesis in plants and animals. Although the chemical steps from GPP to nepetalactone are the same in both *Nepeta* and pea aphids, the enzymes of these pathways have clearly evolved independently.

Results

Transcriptome-Enabled Discovery of Iridoid Genes in *A. pisum*. Iridoids are produced exclusively in the hind legs of sexual aphid females (20), which allowed us to search for iridoid genes by comparing transcriptomic data from legs of sexual females, asexual females, and males. Since most aphid species have a complex life cycle with multiple asexual generations over the summer and only one sexual female/male generation in fall (22), we subjected a colony of pea aphids to day length/temperature conditions that mimic the fall season to generate sexual females. After verifying that *cis-trans*-nepetalactol and *cis-trans*-nepetalactone were in fact produced in the aphids (*SI Appendix*, Fig. S1), we then collected hind and front legs of sexual females, hind legs of asexual females, and hind legs of males. RNA was extracted and subjected to Illumina sequencing. Out of the 20,918 gene models of the *A. pisum* v3 genome, only 96 appeared to be specifically expressed in hind legs of sexual females (*SI Appendix*, Tables S1 and S2). Notably, among these transcripts were 8 genes encoding the entire mevalonate pathway starting from acetyl-CoA to IPP and DMAPP (Fig. 1*B*). Previously reported isotopic labeling studies indicated that the iridoid pathways in pea aphids and plants share at least some of the same biosynthetic intermediates (21). We therefore assumed that the biochemical transformations in pea aphids were the same as those established in *Nepeta*. Then we compiled a list of candidates from the 96 genes specifically expressed in sexual female hind legs that also encoded enzymes that could, in principle, carry out these predicted reactions. One putative *IDS* gene, two putative phosphatase genes, one putative cytochrome P450 gene, one putative P450 reductase gene, and six putative oxidase/reductase genes were selected for further characterization (Fig. 1*B*).

ApIDS Catalyzes the Metal Ion Cofactor–Dependent Formation of GPP. The formation of GPP in the horseradish leaf beetle *Phaedon cochleariae* is catalyzed by a bifunctional IDS (PcIDS) that shows a metal ion cofactor–dependent product specificity, producing primarily GPP with cobalt or manganese, or farnesyl diphosphate (FPP) with magnesium (12). A homolog of PcIDS (ApIDS, gene ID, 100144905) was specifically expressed in sexual female hind legs. Phylogenetic analysis revealed that PcIDS and ApIDS clustered together in a clade of beetle and aphid GPP/FPP synthases (Fig. 2*A*). In vitro characterization of recombinant ApIDS (*SI Appendix*, Tables S3 and S4) showed IDS activity when incubated with IPP and DMAPP. Using magnesium as a cofactor, ApIDS produced similar amounts of GPP, FPP, and geranylgeranyl diphosphate (GGPP), while cobalt and, to a lesser extent, manganese shifted the product specificity to GPP as the major product (Fig. 2*B* and *SI Appendix*, Fig. S2). This suggests that ApIDS, like PcIDS, uses cobalt or manganese as a metal ion cofactor to produce GPP for iridoid formation in vivo.

The Phosphatase ApGES Hydrolyzes GPP to Geraniol. In plants, the formation of the mono- and sesquiterpene alcohols geraniol and farnesol from GPP and FPP is catalyzed by terpene synthases (TPSs) (23). However, in insects, farnesol, which is an intermediate in juvenile hormone biosynthesis, is produced from FPP by a phosphatase belonging to the haloalkanoic acid dehalogenase (HAD) super family (24, 25). Our candidate search in the pea aphid transcriptome did not reveal any TPS- or HAD-like proteins that were selectively expressed in sexual female hind legs. Instead, we identified two putative phosphatases annotated as dolichyldiphosphatase (gene ID, 100158803) and inositol polyphosphate-1-phosphatase (gene ID, 100162683) (Fig. 1*B* and *SI Appendix*, Tables S3 and S5). The putative inositol polyphosphate-1-phosphatase, expressed in *Escherichia coli* as a soluble protein, showed no GPP hydrolysis activity. In contrast, *Saccharomyces cerevisiae* microsomes harboring this recombinant membrane-bound putative dolichyldiphosphatase protein hydrolyzed GPP to geraniol (Fig. 2*C* and *SI Appendix*, Fig. S3). Thus, we named this dolichyldiphosphatase homolog ApGES.

The P450 ApG8H and the P450 Reductase ApRed Act Together to Produce 8-Hydroxygeraniol. Both plants and the leaf beetle *P. cochleariae* utilize P450 enzymes to catalyze the hydroxylation of geraniol to 8-hydroxygeraniol (2, 9, 15). Only a single P450 in the pea aphid transcriptome displayed selective expression in sexual female hind legs, and this enzyme grouped together with PcG8H from *P. cochleariae* in a phylogenetic analysis (Fig. 1*B* and *SI Appendix*, Fig. S4), though these two proteins share only 35% amino acid sequence identity. The aphid gene was named *ApG8H* and the complete open reading frame (ORF) was expressed in *S. cerevisiae* either alone or together with a putative P450 reductase gene (*ApRed*; gene ID, 100162683), which had a similar expression pattern to *ApG8H*. In the presence of the cosubstrate NADPH and ApRed, ApG8H converted geraniol to 8-hydroxygeraniol (Fig. 2*D*). A heterologously expressed P450 from maize (BX2) (26) that was used as a negative control, and ApG8H expressed without ApRed, showed no activity. Notably, ApG8H exhibited a relatively broad substrate specificity, hydroxylating citronellol, nerol, linalool, and neral, though not the monoterpene hydrocarbons limonene and myrcene (*SI Appendix*, Fig. S5), suggesting that an oxygen atom at C1 is critical for binding or catalysis.

The Short-Chain Reductase ApHGO Catalyzes the $NADP^+$-Dependent Oxidation of 8-Hydroxygeraniol to the Iridoid Precursor 8-Oxogeranial. While the oxidation of 8-hydroxygeraniol to 8-oxogeranial in plants is catalyzed by alcohol dehydrogenases (5, 9), *P. cochleariae* beetles use a flavin-dependent glucose-methanol-cholin (GMC) oxidase to catalyze this reaction (13). Thus, we tested two aphid short-chain alcohol dehydrogenase (SDR) candidates, one annotated as farnesol dehydrogenase (gene ID, 100301633) and the other as retinol dehydrogenase (gene ID, 100162094), as well as two GMC candidates (gene IDs, 100169582 and 100164798) that were selectively expressed in sexual female hind legs (Fig. 1*B* and *SI Appendix*, Fig. S6). The complete ORFs were expressed in *E. coli* and purified proteins were assayed with 8-hydroxygeraniol in the presence of $NADP^+$. Enzyme activity could only be observed for the putative farnesol dehydrogenase (named ApHGO), which catalyzed this two-step oxidation (Fig. 2*E* and *SI Appendix*, Fig. S7). Further characterization showed that ApHGO preferred $NADP^+$ over NAD^+ as cosubstrate and exhibited a broader substrate specificity, also oxidizing geraniol and nerol, but not β-citronellol (*SI Appendix*, Fig. S8). The putative retinol dehydrogenase 100162094 and the GMC oxidase 100169582, although not able to accept 8-hydroxygeraniol as substrate, converted geraniol to geranial (*SI Appendix*, Fig. S7).

The Iridoid Synthase ApISY Is a Membrane Protein Catalyzing the Reduction of 8-Oxogeranial. The crucial formation of the cyclopentanoid-pyran scaffold occurs with the reductive cyclization

Downloaded from https://www.pnas.org by 95.90.176.231 on January 27, 2023 from IP address 95.90.176.231.

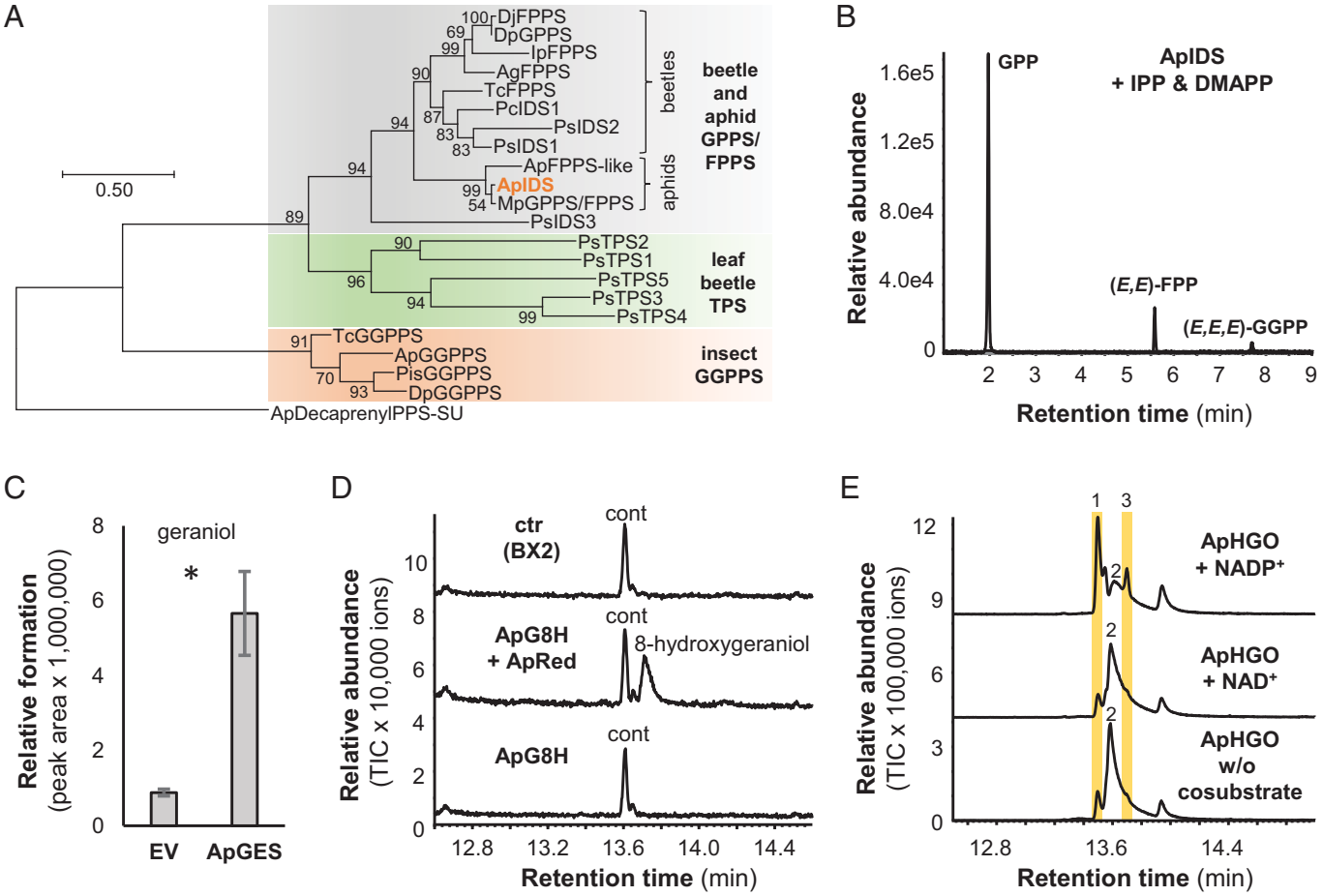


Fig. 2. Biochemical characterization of ApIDS, ApGES, ApG8H, and ApHGO. (*A*) Cladogram analysis of characterized IDSs and TPSs from insects. The tree was inferred by using the maximum likelihood method based on the Jones-Thornton-Taylor (JTT) model. Bootstrap values (n = 1,000) are shown next to each node. The rooted tree is drawn to scale, with branch length measured in the number of amino acid substitutions per site. Putative decaprenyldiphosphate synthase subunit 2 from *A. pisum* was used as an outgroup. (*B*) Characterization of ApIDS. ApIDS lacking the N-terminal signal peptide was expressed as an N-terminal His-tag fusion protein in *E. coli*, purified, and incubated with the substrates IPP and DMAPP in the presence of 1 mM $CoCl_2$. Enzyme products GPP, (*E,E*)-FPP, and (*E,E,E*)-GGPP were analyzed using liquid chromatography–tandem mass spectrometry. (*C*) Characterization of ApGES. The phosphatase ApGES was expressed in *S. cerevisiae* and microsomes harboring the recombinant protein were incubated with GPP. Geraniol was extracted with hexane and analyzed using GC-MS. Mean peak area ± SE are shown (n = 3 independent expression constructs). The amount of geraniol formed was significantly higher in ApGES samples than in the negative controls (empty vector [EV]) (t = 5.219; df = 2.026; P = 0.034). (*D*) Characterization of ApG8H. *S. cerevisiae* microsomes containing either ApG8H, ApG8H in combination with the P450 reductase ApRed, or the maize P450 BX2 as negative control were assayed with geraniol as substrate and NADPH as cosubstrate. Reaction products were analyzed using GC-MS. cont, di-tert-butylphenol (contamination). (*E*) Characterization of ApHGO. ApHGO was expressed as N-terminal His-tag fusion protein in *E. coli*, purified, and incubated with 8-hydroxygeraniol either in the absence or presence of NAD(P). Enzyme products were extracted from the assays and analyzed using GC-MS. 1,8-oxogeraniol (partially oxidized product); 2,8-hydroxygeraniol (starting material); 3,8-oxogeranial (fully oxidized product). Note: The tailing of the peaks is due to the polar nature of the aldehydes and the alcohol.

of 8-oxogeranial. In plants, this step is initiated by iridoid synthase, an SDR that belongs to the progesterone 5β-reductase super family (3). The enzyme that insects use to catalyze this reduction is unknown. In vitro assays of a putative retinol dehydrogenase (gene ID, 100162094) and two GMC oxidases (gene IDs, 100169582 and 100164798) from our list of candidate genes showed no activity with the substrate 8-oxogeranial and the cosubstrate NADPH (*SI Appendix*, Fig. S9). However, a putative oxidoreductase annotated as membrane-bound polyprenol reductase (gene ID, 103310029) (*SI Appendix*, Table S5) was also selectively expressed in sexual female hind legs. Yeast microsomes containing this recombinant protein converted the substrate 8-oxogeranial to *cis-trans*-nepetalactol as the major product, along with minor amounts of *cis-trans*-iridodial and five other unidentified compounds (Fig. 3*A* and *SI Appendix*, Figs. S9 and S10). Thus, the tested protein was named ApISY. As with plant iridoid synthases, without NADPH, ApISY showed no activity (Fig. 3*A*).

In *Nepeta*, and likely also in other plants, iridoid synthase works in concert with cyclases that mediate the stereoselective cyclization of the initial reduction product of ISY, 8-oxocitronellyl enol/enolate, into different nepetalactol stereoisomers (8). When any plant ISY is assayed in vitro without a cyclase, the product profile is strongly dependent on the assay conditions. In high buffer concentrations or at low pH values, spontaneous tautomerization of the 8-oxocitronellyl enol/enolate intermediate to 8-oxocitronellal is favored, while low buffer conditions or higher pH values lead to the spontaneous cyclization to *cis-trans*-nepetalactol as the predominant product. Moderate buffer concentrations or pH values lead to a mixture of monocyclic dialdehydes (8). Using ISY from the plant *C. roseus* (CrISY) as a point of comparison (Fig. 4 and *SI Appendix*, Fig. S11), we were surprised to observe that assays with yeast microsomes containing ApISY showed specificity for formation of *cis-trans*-nepetalactol over a broad buffer concentration range. Only at a buffer concentration of 0.5 M, yield of *cis-trans*-nepetalactol was affected in the ApISY reaction (Fig. 4). Since the plant ISY is a soluble protein, while ApISY is an integral membrane protein with seven predicted transmembrane domains (*SI Appendix*, Table S5), we assayed CrISY in the presence of yeast

Downloaded from https://www.pnas.org by 95.90.176.231 on January 27, 2023 from IP address 95.90.176.231.

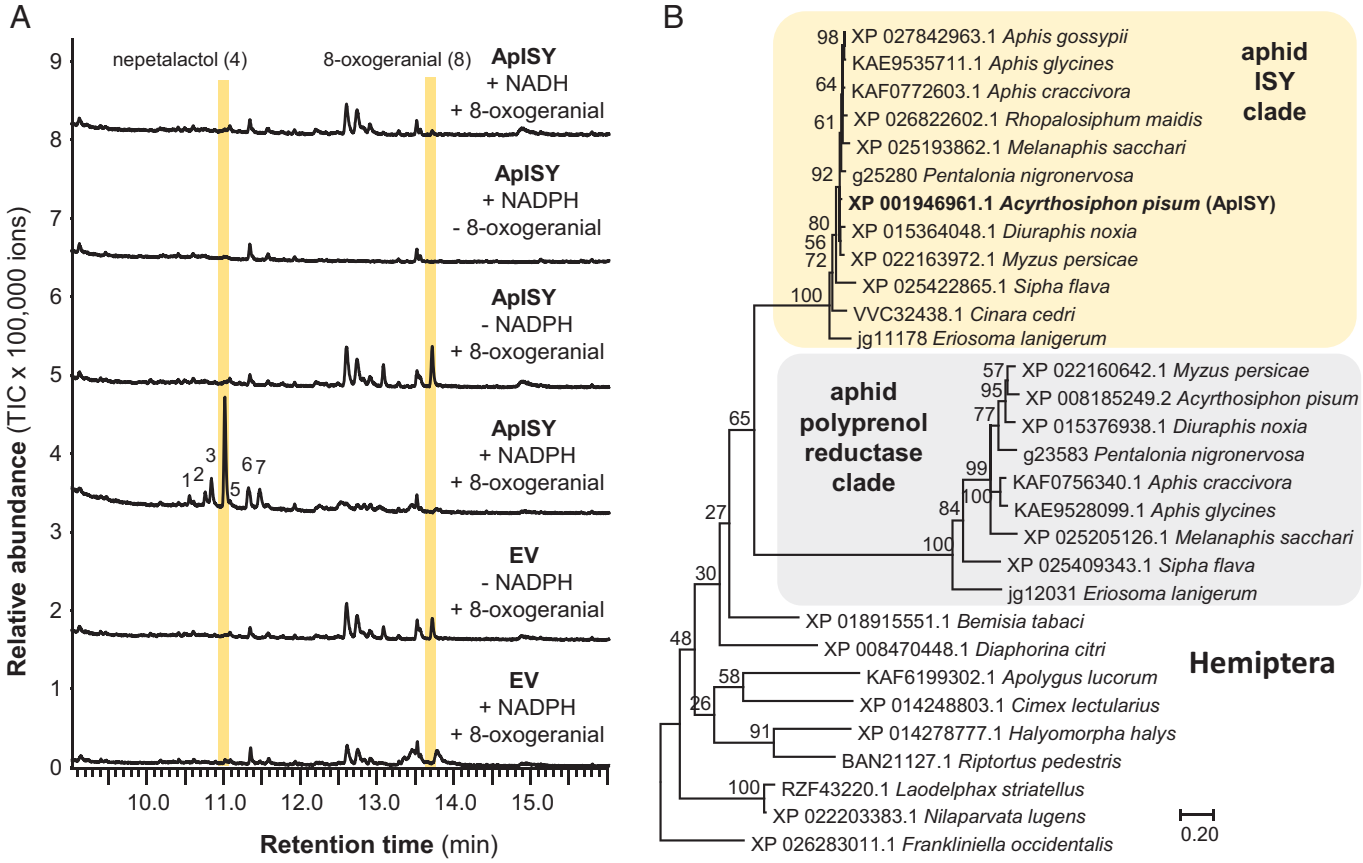


Fig. 3. ApISY is a membrane-bound reductase that likely evolved from a polyprenol reductase ancestor. (*A*) Yeast (*S. cerevisiae*) microsomes made from a strain harboring the EV or expressing ApISY were assayed with 8-oxogeranial either in the presence or absence of NAD(P)H. Reaction products were extracted with ethyl acetate and analyzed using GC-MS. (*B*) Rooted tree of polyprenol reductases and iridoid synthases from Hemiptera. The tree was inferred by using the maximum likelihood method based on the Le_Gascuel_2008 model. Bootstrap values (n = 1,000) are shown next to the branches. A discrete gamma-distribution was used to model evolutionary rate differences among sites. The tree is drawn to scale, with branch lengths measured in the number of amino acid substitutions per site. All positions with less than 95% site coverage were eliminated. A putative polyprenyol reductase from the Western flower thrips, *Frankliniella occidentalis*, was used to root the tree. NCBI accession numbers for all sequences are given in the tree.

microsomes to determine whether the membranes had an impact on product profile. Notably, *cis-trans*-nepetalactol was the predominant product of CrISY throughout the buffer concentration range after addition of microsomes (Fig. 4). This suggests that the lipid environment of the membranes likely facilitates spontaneous cyclization to the iridoid scaffold by preventing contact of the reactive intermediate with general acid catalysts such as buffer components in the assay.

ApISY Likely Evolved from a Polyprenol Reductase Ancestor. Polyprenol reductases, ubiquitous in eukaryotes, catalyze reduction of the α-isoprene unit of polyprenols to form dolichols, the precursors for dolichol-linked monosaccharides that are required for protein *N*-glycosylation (27–29). Together with steroid 5α-reductases and very-long-chain enoyl-CoA reductases, polyprenol reductases belong to the steroid 5α-reductase (SRD5A) family (Pfam, PF02544). A BLAST analysis with ApISY as query revealed two polyprenol reductase-like genes in most of the available aphid genomes. In a phylogenetic analysis, these genes formed two distinct and aphid-specific clades among the polyprenol reductases of eukaryotes (Fig. 3*B* and *SI Appendix*, Fig. S12).

The Flavin-Dependent GMC Oxidase ApNEPO Converts *Cis-Trans*-Nepetalactol into *Cis-Trans*-Nepetalactone. The SDR NEPS1 from *Nepeta* catalyzes the oxidation of *cis-trans*-nepetalactol to *cis-trans*-nepetalactone (8). To identify the enzyme that catalyzes this reaction in pea aphids, the putative retinol dehydrogenase (gene ID, 100162094) and the two FAD-dependent GMC oxidases (gene IDs, 100169582 and 100164798) from our candidate gene list (Fig. 1*B*) were assayed with the *cis-trans*-nepetalactol substrate and $NADP^+$. While the SDR 100162094 and the GMC oxidase 100164798 were not active, GMC oxidase 100169582 converted *cis-trans*-nepetalactol into *cis-trans*-nepetalactone (Fig. 5 and *SI Appendix*, Fig. S13). A phylogenetic analysis revealed that this enzyme, designated *A. pisum* nepetalactol oxidase (ApNEPO), belonged to the ε-clade of GMC oxidoreductases (*SI Appendix*, Fig. S14) and was not related to the leaf beetle enzyme PcHGO, which clustered into a beetle-specific GMC clade (*SI Appendix*, Fig. S14). Sequence prediction suggested that ApNEPO contains a signal peptide targeting the protein into the lumen of the endoplasmatic reticulum (ER) (*SI Appendix*, Table S5).

Discussion

Here, we elucidated the entire iridoid pathway in the pea aphid *A. pisum*. Previously reported feeding studies in pea aphids (21), as well as the spatial localization of nepetalactone biosynthesis (20), guided the identification of the six biosynthetic genes (Fig. 1*A*). The characterization of these enzymes indicates that the plant and aphid nepetalactone biosynthetic pathways are composed of the same chemical transformations. However, each of the respective enzymes clearly evolved independently in plants and aphids. Our data show that in some cases plants and aphids recruited enzymes from different protein families to catalyze the

Downloaded from https://www.pnas.org by 95.90.176.231 on January 27, 2023 from IP address 95.90.176.231.

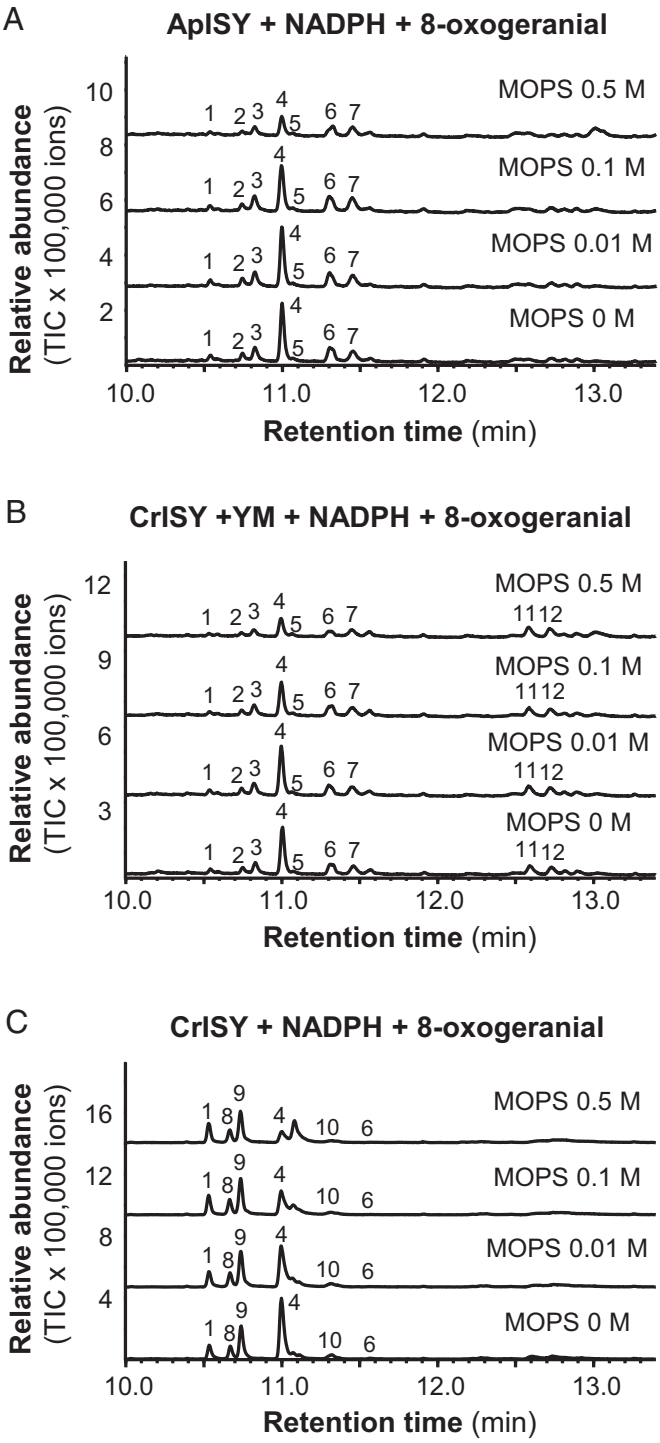


Fig. 4. Buffer concentration and the lipid environment influence iridoid synthase activity. (*A*) Yeast (*S. cerevisiae*) microsomes containing ApISY were assayed with 8-oxogeranial in the presence of NADPH under different buffer concentrations. (*B*) CrISY was expressed in *E. coli*, purified, and assayed with 8-oxogeranial and NADPH under different buffer concentrations in the background of yeast microsomes (YM). (*C*) Purified CrISY was incubated with 8-oxogeranial and NADPH under different buffer concentrations. Reaction products were extracted with ethyl acetate and analyzed using gas chromatography–mass spectrometry. 1, *cis-trans*-iridodial; 2, unidentified; 3, unidentified; 4, *cis-trans*-nepetalactol; 5, unidentified; 6, unidentified; 7, unidentified; 8, *trans-trans*-iridodial; 9, *cis-trans*-iridodial; 10, tetrahydro-8-oxogeranial; 11, unidentified; and 12, unidentified.

same reactions (*SI Appendix*, Fig. S15). The formation of geraniol in *Catharanthus* and *Nepeta*, for example, is mediated by terpene synthases (5, 9), while the pea aphid uses a phosphatase to produce geraniol by direct hydrolysis of the phosphodiester bond of GPP (Fig. 2*C*). Moreover, the reductive cyclization of 8-oxogeranial to *cis-trans*-nepetalactol and the subsequent oxidation of this alcohol to *cis-trans*-nepetalactone in aphids involves the action of a polyprenol reductase-like protein and a flavin-dependent oxidase from the GMC family, respectively, while plants recruited members of the SDR family to catalyze both reactions (*SI Appendix*, Fig. S15).

Only one other natural product pathway, the three-step biosynthesis of the cyanogenic glycoside linamarin, has been fully elucidated in both plants and insects. Recent work has demonstrated that linamarin biosynthesis consists of two cytochrome P450s and one glucosyl transferase in both plants and insects, and that these pathways arose independently (30, 31). Additionally, terpene synthases, the key enzymes in terpene formation, have been identified in both plants and insects, and these enzymes are also the result of independent evolution in the different kingdoms (32–34).

Iridoids and iridoid-related compounds are widespread among insects and have been observed in different insect orders, including Coleoptera, Hymenoptera, and Hemiptera (1). The discovery of the aphid nepetalactone pathway provides an opportunity to determine whether iridoids evolved convergently in two divergent species of insects. The biosynthesis of the iridoid-related dialdehyde chrysomelidial has been partially elucidated in the leaf beetle *P. cochleariae* (11). Although chrysomelidial lacks the cyclopentanoid-pyran scaffold that defines the iridoids, its formation shares many of the same reactions as the iridoid core pathway in aphids (*SI Appendix*, Fig. S15). GPP is produced in the leaf beetle by PcIDS, which is obviously phylogenetically related to ApIDS (12) (Fig. 2*A*). Geraniol, which is produced in beetles by an as yet undiscovered enzyme, acts as substrate for PcG8H, a P450 that hydroxylates this alcohol to 8-hydroxygeraniol (15). Although PcG8H and ApG8H both belong to clan 3 of insect P450s (*SI Appendix*, Fig. S4), their low sequence identity suggests independent origins. Independent evolution of enzyme activities is even more obvious for the last two steps of the chrysomelidial pathway. The oxidation

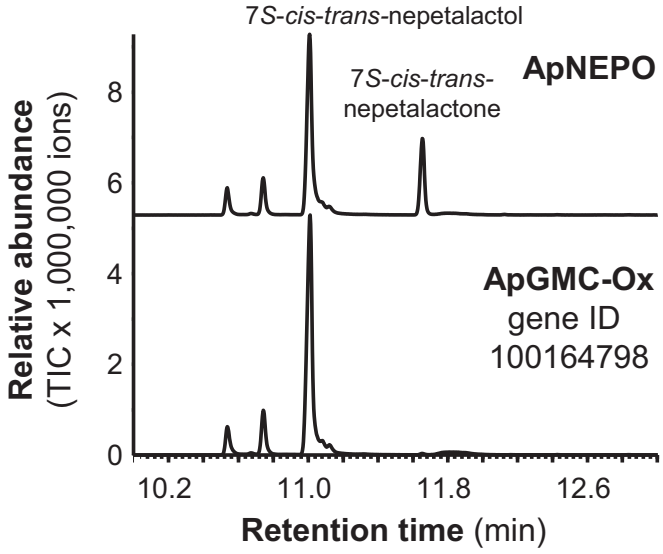


Fig. 5. Biochemical characterization of ApNEPO. ApNEPO and another putative GMC-oxidase (ApGMC-Ox) highly expressed in hind legs of *A. pisum* sexual females were expressed as N-terminal His-tag fusion proteins in *E. coli*, purified, and incubated with 7*S-cis-trans*-nepetalactol in the presence of NADP. Enzyme products were extracted with ethyl acetate and analyzed using gas chromatography–mass spectrometry.

Downloaded from https://www.pnas.org by 95.90.176.231 on January 27, 2023 from IP address 95.90.176.231.

of 8-hydroxygeraniol in *P. cochleariae* is catalyzed by a GMC oxidase (13), in contrast to aphids, which recruited an SDR for the same reaction, and the final nonreductive cyclization of the formed 8-oxogeranial to chrysomelidial is presumably mediated by a PcTo-like juvenile hormone-binding protein, although conclusive evidence of this enzyme activity is still lacking (35). Overall, the elucidation of the iridoid pathway in aphids presented here shows that although the reaction sequence is conserved, iridoid formation evolved independently not only in different kingdoms, but also in different insect orders through convergent evolution (*SI Appendix*, Fig. S15).

From a chemical perspective, the first committed step in the iridoid pathway, the cyclization of 8-oxogeranial by ISY, is of mechanistic interest. All known plant iridoid synthases belong to the SDR protein family and have been described to catalyze the reduction of 8-oxogeranial to a highly reactive 8-oxocitronellyl enol/enolate intermediate, which is then cyclized by NEPS or MLPL proteins to different stereoisomers of nepetalactol (8, 9). In the absence of a cyclase, 8-oxocitronellyl enol/enolate can react spontaneously to form various compounds depending on the assay conditions. Higher pH values or low buffer concentrations lead to cyclization to *cis-trans*-nepetalatol, while acidic conditions or high buffer concentrations favor spontaneous formation of 8-oxocitronellal or other dialdehydes (8). In contrast to plant ISY, a soluble protein likely located in the cytosol, the iridoid synthase in the pea aphid is an integral membrane protein predicted to possess seven membrane domains (*SI Appendix*, Table S5). Interestingly, yeast microsomes harboring ApISY produced mainly *cis-trans*-nepetalactol independent of the buffer concentration (Fig. 4*A*). Moreover, when CrISY from the plant *C. roseus* was tested in the presence of microsomes prepared from a control yeast strain, the same trend, the production of mainly *cis-trans*-nepetalactol at all buffer concentrations tested, was observed (Fig. 4*B*). This indicates that the lipid environment of membranes favors the spontaneous cyclization of the 8-oxocitronellyl enol/enolate to *cis-trans*-nepetalactol, presumably by preventing contact of the reactive intermediate with protons or other acidic compounds. Since the lipid composition of membranes can vary significantly depending on cell type and developmental stage, a detailed lipid analysis may help to better understand this process. Although iridoid formation in pea aphids does not appear to require the action of a cyclase, we cannot rule out that ApISY itself or other, as yet unidentified, aphid proteins fulfill this function in vivo. Most aphid species produce *cis-trans*-nepetalactol (18, 36). The damson-hop aphid *Phorodon humili*, however, produces the *cis-cis* isomer (37, 38). This species must have either an iridoid synthase with different stereoselectivity, catalyzing both the reduction and cyclization of 8-oxogeranial to the final nepetalactol stereosiomer, or a partner cyclase that cyclizes the potential 8-oxocitronellyl enol/enolate intermediate to *cis-cis*-nepetalactol. Elucidating the biosynthetic pathway for *cis-cis*-nepetalactol in *P. humili* would provide additional insight into the complex chemistry underlying the formation of the iridoid backbone in animals.

Sequence comparisons revealed that ApISY is related to polyprenol reductases, a class of integral membrane proteins involved in *N*-glycosylation of secreted and membrane-bound proteins (27–29) (*SI Appendix*, Fig. S12). Interestingly, most of the aphid species sequenced to date possess two putative polyprenol reductase copies that form two distinct and aphid-specific clades in a phylogenetic tree of the polyprenol reductases of Hemiptera (Fig. 3*B* and *SI Appendix*, Fig. S12). Thus, it is likely that the ApISY-containing clade represents aphid iridoid synthases, while the other clade contains true polyprenol reductases, which is a metabolic enzyme that is essential for survival. The close relationship of these two clades suggests that iridoid synthase activity arose by gene duplication and subsequent neofunctionalization of a polyprenol reductase gene early in aphid evolution or in an ancestor of the aphids. Furthermore, the striking sequence similarities among proteins within the two clades (Fig. 3*B*) indicates a high degree of purifying selection to preserve their respective enzymatic functions. We cannot predict the evolutionary origin of other iridoid pathway genes because the function of their closest homologs in aphids is still unknown. For example, although ApGES was annotated as a dolichyldiphosphatase, an enzyme acting together with polyprenol reductase in the process of *N*-glycosylation of secreted and membrane-bound proteins (39, 40), there is no experimental evidence for dolichyldiphosphatase activity of ApGES, and BLAST analysis did not reveal any other *ApGES*-related gene that could be dedicated to dolichyldiphosphatase function in the *A. pisum* genome.

A defining feature of the iridoid biosynthetic pathway in aphids is that many of the enzymes appear to be membrane anchored. In addition to the iridoid synthase ApISY, the phosphatase ApGES, the cytochrome P450 ApG8H, and its reductase ApRed were predicted to be integral membrane proteins (*SI Appendix*, Table S5). Moreover, the prediction of an ER signal peptide in ApNEPO suggests that this protein is localized in the lumen of the ER (*SI Appendix*, Table S5). This leads us to speculate that these enzymes may form a metabolon, most likely, based on signal sequence prediction, on the membrane of the ER (Fig. 6). Given the predicted mitochondrial signal peptide of ApIDS, the pathway likely starts in the mitochondria with the formation of GPP, which is then transported to the ER. Notably, the list of candidate iridoid genes exclusively expressed in hind legs of sexual female aphids contained seven genes annotated as transporters for the inner and outer membrane of mitochondria that could be involved in GPP transport (*SI Appendix*, Table S1). The hydrolysis of GPP and the formation of the iridoid backbone could then be catalyzed by the putative metabolon in the ER membrane, which may

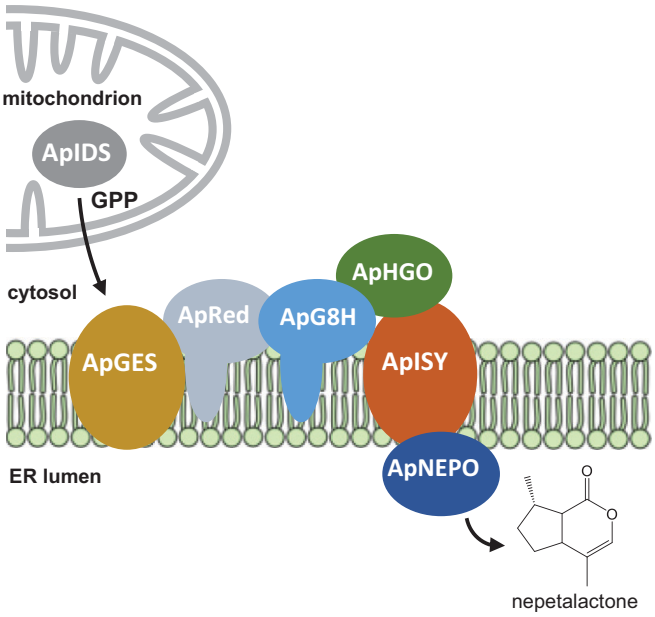


Fig. 6. The iridoid pathway in aphids might be organized in a membrane-associated metabolon. While ApHGO is predicted to be localized as soluble protein in the cytosol and ApNEPO is likely localized in the lumen of the ER, ApGES and ApISY are transmembrane proteins.

Downloaded from https://www.pnas.org by 95.90.176.231 on January 27, 2023 from IP address 95.90.176.231.

provide efficient substrate channeling, preventing the release of highly reactive pathway intermediates such as 8-oxogeranial. The final oxidation of the comparatively stable nepetalactol product to nepetalactone could then occur in the lumen of the ER, and formation of nepetalactone-containing vesicles and their transport to the cell membrane might represent a possible mechanism for the active release of these volatile iridoids.

Overall, chemical logic, along with the discreet spatial localization of the site of biosynthesis, facilitated the discovery of the six-step pathway for nepetalactone biosynthesis in animals. This provides a foundation for understanding how complex natural products have evolved in two kingdoms of life. The insect pathway also provides insights into the relatively understudied field of insect natural product biosynthesis.

Materials and Methods

Cultivation of *A. pisum* and Generation of Sexual Female Aphids. Asexual females of pea aphid (*A. pisum* [Harris]) clone JML06 were reared on 4-wk-old broad bean (*V. faba*) cv. "The Sutton" plants under long-day conditions (16/8 h light/dark, 22 °C, 60% humidity). To avoid escape of aphids, plants were covered with air-permeable cellophane bags (18 × 38.5 cm, Griesinger Verpackungsmittel). To generate sexual female and male aphids, asexual L3 aphid larvae were transferred to short-day conditions mimicking fall season (12/12 h light/dark, 14 °C, 60% humidity). Two generations later, sexual females and males were produced, and adult aphids (6 to 10 d old) were used for experiments. The emission of iridoids by sexual female aphids was tested by placing a solid phase microextraction (SPME) fiber for 3 h into the headspace of *V. faba* plants with the aphids. The SPME fiber was then loaded into the injector of a gas chromatograph coupled with a mass spectrometer (GC-MS) as described below.

Transcriptome Sequencing and Gene Identification. For RNA extraction, 20 hind legs of sexual females, 20 front legs of sexual females, 20 hind legs of male aphids, and 20 hind legs of asexual aphids were collected and directly placed in 450 μL lysis buffer containing guanidinium thiocyanate (innuPREP RNA Mini Kit, IST Innuscreen). Material was shredded by shaking with metal beads using a Tissue Lyzer II (Qiagen) for 2 × 4 min (frequency 50/s). Total RNA was extracted with the innuPREP RNA Mini Kit according to the manufacturer's instructions, eluted in 30 μL RNase-free water, and sent to Novogene for RNA-seq library construction (polyA enrichment) and sequencing (NovaSeq PE150, paired reads, 6 gigabytes of raw data per sample). Trimming of the obtained sequencing reads and mapping to the pea aphid genome (version 3) were performed with the program CLC Genomics Workbench (Qiagen Bioinformatics) (mapping parameter: length fraction, 0.8; similarity fraction, 0.9; maximum number of hits, 25). In order to identify pea aphid genes involved in iridoid formation, we performed Pearson correlation based on the hypothesis that iridoid genes are exclusively expressed in hind legs of sexual female aphids. Genes with a Pearson correlation coefficient ≥0.99, a Reads per kilobase million (RPKM) value ≥10 in hind legs of sexual female aphids, and a fold change ≥5 (hind legs of sexual female aphids versus other samples) were considered as candidates (*SI Appendix*, Table S1).

Prediction of Signal Peptides and Transmembrane Domains. Prediction of signal peptides (*SI Appendix*, Table S3) and transmembrane domains (*SI Appendix*, Table S5) was performed using TargetP-2.0 (https://services.healthtech.dtu.dk/service.php?TargetP-2.0) and DeepTMHMM (https://dtu.biolib.com/DeepTMHMM/), respectively.

Gene Synthesis and Cloning. The complete ORFs of *ApHGO*, *ApNEPO*, 100162094, 100164798, 100162683, and 100168586, as well as the N-terminal truncated ORF of *ApIDS* lacking the predicted signal peptide were synthesized after codon optimization for heterologous expression in *E. coli* by Twist Bioscience and inserted as *Bam*HI/*Hin*dIII fragments into the vector pET-28a(+) that allows expression as N-terminal His-tag fusion protein in *E. coli*. *ApGES* was codon optimized for *S. cerevisiae* and synthesized by Twist Bioscience. The complete ORFs of *ApG8H*, *ApRed*, and *ApISY* were amplified from cDNA obtained from hind legs of sexual female pea aphids using the primers listed in *SI Appendix*, Table S6. *ApG8H* and *ApRed* were cloned as sticky-end fragments into the same pESC-Leu-2d vector using the two different cloning sites (41). *ApISY* and *ApGES* were separately cloned as sticky-end fragments into pESC-Leu-2d. cDNA was synthesized from total RNA (1 μg) treated with DNaseI (Thermo Fisher Scientific) using SuperScript III reverse transcriptase and oligo (dT)20 primers (Invitrogen) according to the manufacturer's instructions. All synthesized or amplified sequences are given in *SI Appendix*, Table S4.

Heterologous Expression of Candidate Genes in *E. coli*. Expression constructs were transferred to *E. coli* strain BL21 (DE3) (Invitrogen). Liquid cultures were grown in lysogeny broth at 37 °C and 220 rpm until an OD_{600} of 0.7, induced with a final concentration of 0.5 mM isopropyl beta-D-1-thiogalactopyranoside, and subsequently incubated at 18 °C and 220 rpm for 16 h. The cells were harvested by centrifugation at 3,200 × *g* for 10 min, resuspended in refrigerated extraction buffer (50 mM Tris·HCl pH 8, 500 mM NaCl, 20 mM imidazole, 5% [vol/vol] glycerol, 50 mM glycine, ethylenediaminetetraacetic acid (EDTA)-free protease inhibitor (1 tablet/50 mL buffer, freshly added), and lysozyme (10 mg/50 mL buffer, freshly added) and disrupted by sonication for 2 min (2 s on, 3 s off) on ice (Bandelin UW 2070). Cell debris were removed by centrifugation (35,000 × *g* at 4 °C for 20 min) and the N-terminal His-tagged proteins were purified from the supernatant using NiNTA agarose (Qiagen) according to the manufacturer's instructions. The buffer of the eluted protein samples was exchanged for assay buffer (for details see paragraph enzyme assays) 100 mM 3-(N-morpholino)propanesulfonic acid (MOPS) pH 7.5, 10% (vol/vol) glycerol by using Amicon 10K Concentrator columns (Merck Millipore). Sodium dodecyl sulfate-polyacrylamid gel electrophoresis and spectrophotometric analysis was used to check purity and approximate quantity of proteins.

Heterologous Expression of ApGES, ApG8H, and ApISY in *S. cerevisiae*. For heterologous expression in yeast, constructs were transformed into the *S. cerevisiae* strain INVSc1 (Thermo Fisher) using the S.c. EasyComp Transformation Kit (Invitrogen) according to the manufacturer's instructions. Subsequently, 30 mL Sc-Leu minimal medium (6.7 g/L yeast nitrogen base without amino acids, but with ammonium sulfate; 100 mg/L of each L-adenine, L-arginine, L-cysteine, L-lysine, L-threonine, L-tryptophan, and uracil; 50 mg/L of each L-aspartic acid, L-histidine, L-isoleucine, L-methionine, L-phenylalanine, L-proline, L-serine, L-tyrosine, L-valine; 20 g/L d-glucose) was inoculated with single yeast colonies and grown overnight at 28 °C and 180 rpm. For main cultures, 100 mL yeast peptone glucose agar (YPGA) (Glc) full medium (10 g/L yeast extract, 20 g/L bactopeptone, 74 mg/L adenine hemisulfate, 20 g/L d-glucose) was inoculated with one unit OD_{600} of the overnight cultures and incubated under the same conditions for 30 to 35 h. After centrifugation (5,000 × *g*, 16 °C, 5 min), the expression was induced by resuspension of the cells in 100 mL YPGA (Gal) medium (see above, but including 20 g/L galactose instead of D-glucose) and grown for another 15 to 18 h at 25 °C and 160 rpm. The cells were harvested by centrifugation (7,500 × *g*, 10 min, 4 °C), resuspended in 30 mL TEK buffer (50 mM Tris·HCl pH 7.5, 1 mM EDTA, 100 mM KCl) and centrifuged again. Then, the cells were carefully resuspended in 2 mL TES buffer (50 mM Tris·HCl pH 7.5, 1 mM EDTA, 600 mM sorbitol; freshly added: 10 g/L bovine serum fraction V protein and 1.5 mM β-mercaptoethanol) and disrupted by shaking five times for 1 min with glass beads (0.45 to 0.50 mm diameter, Sigma-Aldrich). The crude extracts were recovered by washing the glass beads four times with 5 mL TES. The combined washes were centrifuged (7,500 × *g*, 10 min, 4 °C), and the supernatant containing the microsomes was transferred into an ultracentrifuge tube. After ultracentrifugation (100,000 × *g*, 90 min, 4 °C), the supernatant was carefully removed and the microsomal pellet was gently washed with 2.5 mL TES buffer, then with 2.5 mL TEG buffer (50 mM Tris·HCl pH 7.5, 1 mM EDTA, 30% glycerol). The microsomal fractions were homogenized in 2 mL TEG buffer using a glass homogenizer (Potter-Elvehjem, Carl Roth) and aliquots were stored at −20 °C until further use.

Enzyme Assays. IDS assays were carried out using 3 μg of purified ApIDS protein desalted in IDS assay buffer (25 mM MOPSO, pH 7.2, 10% [vol/vol] glycerol), 1 mM metal ion cofactor ($MgCl_2$, $MnCl_2$, or $CoCl_2$), and the substrates IPP and DMAPP (each 50 μM) in 100 μL assay buffer at 30 °C for 1 h. Product formation was monitored by liquid chromatography–tandem mass spectrometry (LC-MS/MS) as described below.

Downloaded from https://www.pnas.org by 95.90.176.231 on January 27, 2023 from IP address 95.90.176.231.

GES activity was determined in assays (total volume, 100 μL) containing 20 μL yeast microsomes harboring ApGES, 25 mM Tris·HCl (pH 7.5) and 50 μg/μL GPP. The assays were overlayed with 100 μL hexane and incubated for 20 min at 22 °C. Enzyme products were extracted by vortexing for 1 min, and 1 μL of the hexane phase was injected into the GC-MS (see below).

For measuring G8H activity, 10 μL microsomes harboring either ApG8H alone or in combination with the P450 reductase ApRed were incubated in 25 mM sodium phosphate buffer (pH 7.0) with 25 mM substrate (geraniol, geranial, citronellol, or citronellal, respectively) and 1 mM NADPH in a total volume of 100 μL for 2 h at 30 °C. Assays were then overlayed with 100 μL ethyl acetate and vortexed for 1 min. G8H products were analyzed by injecting 1 μL of the ethyl acetate phase into the GC-MS. The indole hydroxylase BX2 that is involved in benzoxazinoid formation in maize (26) was used as a negative control. Screening ApG8H activity with other substrates including citral A+B, nerol, linalool, limonene, and myrcene was performed by adding 10 μL of the substrate (0.5 mM dissolved in methanol) to 500 μL of living yeast cells induced with galactose-containing medium (see above). Cells were further incubated for 24 h at 28 °C and 200 rpm, and afterward extracted with 200 μL ethyl acetate. An aliquot (1 μL) of the organic phase was injected into GC-MS for enzyme product analysis.

HGO activity was analyzed using assays containing 40 μg purified protein, 1 mM NAD^+ or $NADP^+$, respectively, and 0.5 mM substrate (8-hydroxygeraniol, geraniol, nerol, or β-citronellol) in a total volume of 50 μL MOPS buffer (0.1 M). Assays were overlayed with 200 μL ethyl acetate, incubated for 2 h at 30 °C, and enzyme products were extracted by vortexing the assay for 1 min. One microliter of the organic phase was injected into GC-MS for enzyme product analysis.

ISY activity was measured in assays (total volume, 50 μL) containing 20 μL microsomes, 50 mM MOPS pH 7.5, 1 mM NADPH, and 0.5 mM 8-oxogeraniol. Assays were incubated for 2 h at 30 °C, overlayed with 100 μL ethyl acetate, and products were extracted by vortexing the assays for 1 min. One microliter of recombinant and purified CrISY from *C. roseus* (8) was tested under the same conditions as described above either in the presence or absence of 20 μL yeast microsomes as negative control.

NEPO activity was determined as described above for HGO with 3 μg of purified protein and 0.5 mM *7S-cis-trans*-nepetalactol as substrate.

Gas Chromatography–Mass Spectrometry Analysis. Qualitative analysis of volatile sex pheromones released from sexual female pea aphids was conducted using an Agilent 6890 Series gas chromatograph coupled to an Agilent 5973 quadrupole mass selective detector (Agilent Technologies; injector temperature, 220 °C; interface temp, 250 °C; quadrupole temp, 150 °C; source temp, 230 °C; electron energy, 70 eV). The constituents of the volatile bouquet were separated using a ZB5 column (Phenomenex; 30 m × 0.25 mm × 0.25 μm) and He as carrier gas (2 mL/min). The SPME sample was injected without split at an initial oven temperature of 70 °C. The temperature was held for 2 min and then increased to 220 °C with a gradient of 7 °C min^{-1}, and then further increased to 300 °C with a gradient of 60 °C min^{-1} and a hold of 2 min. Enzyme products extracted in hexane or ethyl acetate were analyzed using the same GC-MS system with a carrier gas flow of 1.5 mL min^{-1}, splitless injection (1 μL sample), and a temperature program from 60 °C (2-min hold) at 10 °C min^{-1} to 220 °C, and a further increase to 300 °C with a gradient of 100 °C min^{-1} and a hold of 2 min. Compounds were identified by comparison of retention times and mass spectra to those of authentic standards or by comparison with reference spectra in the Wiley and National Institute of Standards and Technology libraries.

Liquid Chromatography–Tandem Mass Spectrometry Analysis of IDS Products. ApIDS products GPP, FPP, and GGPP were analyzed as recently described in Lackus et al. (42) using an Agilent 1200 HPLC system (Agilent Technologies) coupled to an API 6500 triple-quadrupole mass spectrometer (Applied Biosystems). For separation, a ZORBAX Extended C-18 column (1.8 μm, 50 mm × 4.6 mm; Agilent Technologies) was used. The mobile phase consisted of 5 mM ammonium bicarbonate in water as solvent A and acetonitrile as solvent B, with the flow rate set at 0.8 mL/min and the column temperature kept at 20 °C. Separation was achieved by using a gradient starting at 0% B (vol/vol), increasing to 10% B in 2 min, 64% B in 12 min, and 100% B in 2 min (1-min hold), followed by a change to 0% B in 1 min (5-min hold) before the next injection. The injection volume for samples and standards was 1 μL. The mass spectrometer was used in the negative electrospray ionization (EI) mode. Multiple-reaction monitoring (MRM) was used to monitor analyte parent ion-to-product ion formation: *m/z* 312.9/79 for GPP, *m/z* 380.9/79 for FPP, and *m/z* 449/79 for GGPP.

Phylogenetic Analysis. Amino acid alignments were constructed using the MUSCLE algorithm (gap open, −2.9; gap extend, 0; hydrophobicity multiplier, 1.2; clustering method, unweighted pair group B method (UPGMB)) implemented in MEGA7 (43). Tree reconstruction was done with MEGA7 using a maximum likelihood algorithm (model/method, given in the respective figure legends; substitutions type, amino acids; rates among sites, uniform rates; gaps/missing data treatment, partial deletion; site coverage cutoff, 80%). Bootstrap resampling analyses with 1,000 replicates were performed to evaluate the tree's topologies.

Statistical Analysis. Differences in phosphatase activities between the empty vector control and ApGES were compared with the Welch two-sample *t* test. Data were analyzed with R 4.2.0 (https://www.R-project.org/).

Data, Materials, and Software Availability. Raw reads from the transcriptome sequencing were deposited in the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) under the BioProject accession PRJNA866370 (44). Amplified gene sequences were deposited in NCBI GenBank with the accessions ON862918 (*ApG8H*) (45), ON862919 (*ApRed*) (46), and ON862920 (*ApISY*) (47). All other study data are included in the article and/or *SI Appendix*.

ACKNOWLEDGMENTS. We thank Michael Reichelt and Maritta Kunert for help with LC-MS/MS and GC-MS analysis, respectively; Nestor Jose Hernandez-Lozada for providing purified recombinant CrISY protein; Kimberly Falk for drawing the aphid cartoons shown in Fig. 1; Julie Jaquiery for helpful advice regarding the generation of sexual female aphids; and Elke Goschala, Claudia Pakebusch, and Andreas Weber from the greenhouse team of the Max Planck Institute for Chemical Ecology for rearing *V. faba* plants. We gratefully acknowledge Fen Li for preliminary experiments. This work was supported by the Max Planck Society. Additional support was provided by the European Research Council (788301).

1. F. Beran, T. G. Köllner, J. Gershenzon, D. Tholl, Chemical convergence between plants and insects: Biosynthetic origins and functions of common secondary metabolites. *New Phytol.* **223**, 52–67 (2019).
2. G. Collu *et al.*, Geraniol 10-hydroxylase, a cytochrome P450 enzyme involved in terpenoid indole alkaloid biosynthesis. *FEBS Lett.* **508**, 215–220 (2001).
3. F. Geu-Flores *et al.*, An alternative route to cyclic terpenes by reductive cyclization in iridoid biosynthesis. *Nature* **492**, 138–142 (2012).
4. A. J. Simkin *et al.*, Characterization of the plastidial geraniol synthase from Madagascar periwinkle which initiates the monoterpenoid branch of the alkaloid pathway in internal phloem associated parenchyma. *Phytochemistry* **85**, 36–43 (2013).
5. K. Miettinen *et al.*, The seco-iridoid pathway from *Catharanthus roseus*. *Nat. Commun.* **5**, 3606 (2014).
6. V. Salim, B. Wiens, S. Masada-Atsumi, F. Yu, V. De Luca, 7-deoxyloganetic acid synthase catalyzes a key 3 step oxidation to form 7-deoxyloganetic acid in *Catharanthus roseus* iridoid biosynthesis. *Phytochemistry* **101**, 23–31 (2014).
7. N. H. Sherden *et al.*, Identification of iridoid synthases from Nepeta species: Iridoid cyclization does not determine nepetalactone stereochemistry. *Phytochemistry* **145**, 48–56 (2018).
8. B. R. Lichman *et al.*, Uncoupled activation and cyclization in catmint reductive terpenoid biosynthesis. *Nat. Chem. Biol.* **15**, 71–79 (2019).
9. B. R. Lichman *et al.*, The evolutionary origins of the cat attractant nepetalactone in catnip. *Sci. Adv.* **6**, eaba0721 (2020).
10. S. E. O'Connor, J. J. Maresh, Chemistry and biology of monoterpene indole alkaloid biosynthesis. *Nat. Prod. Rep.* **23**, 532–547 (2006).
11. A. Burse *et al.*, Always being well prepared for defense: The production of deterrents by juvenile *Chrysomelina beetles* (Chrysomelidae). *Phytochemistry* **70**, 1899–1909 (2009).
12. S. Frick *et al.*, Metal ions control product specificity of isoprenyl diphosphate synthases in the insect terpenoid pathway. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 4194–4199 (2013).
13. P. Rahfeld *et al.*, Independently recruited oxidases from the glucose-methanol-choline oxidoreductase family enabled chemical defences in leaf beetle larvae (subtribe *Chrysomelina*) to evolve. *Proc. Biol. Sci.* **281**, 20140842 (2014).
14. A. Burse, W. Boland, Deciphering the route to cyclic monoterpenes in Chrysomelina leaf beetles: Source of new biocatalysts for industrial application? *Z. Naturforsch. C J. Biosci.* **72**, 417–427 (2017).
15. N. Fu *et al.*, A cytochrome P450 from the mustard leaf beetles hydroxylates geraniol, a key step in iridoid biosynthesis. *Insect Biochem. Mol. Biol.* **113**, 103212 (2019).
16. C. Formisano, D. Rigano, F. Senatore, Chemical constituents and biological activities of *Nepeta* species. *Chem. Biodivers.* **8**, 1783–1818 (2011).
17. T. Eisner, Catnip–Its Raison d'Être. *Science* **146**, 1318–1320 (1964).

Downloaded from https://www.pnas.org by 95.90.176.231 on January 27, 2023 from IP address 95.90.176.231.

18. G. W. Dawson *et al.*, Aphid semiochemicals–A review, and recent advances on the sex pheromone. *J. Chem. Ecol.* **16**, 3019–3030 (1990).
19. J. A. Pickett, R. K. Allemannb, M. A. Birketta, The semiochemistry of aphids. *Nat. Prod. Rep.* **30**, 1277–1283 (2013).
20. K. Murano, K. Ogawa, T. Kaji, T. Miura, Pheromone gland development and monoterpenoid synthesis specific to oviparous females in the pea aphid. *Zoological Lett.* **4**, 9 (2018).
21. S. J. Partridge *et al.*, Iridoid sex pheromone biosynthesis in aphids mimics iridoid-producing plants. *Chemistry* **27**, 7231–7234 (2021).
22. J. C. Simon, C. Rispe, P. Sunnucks, Ecology and evolution of sex in aphids. *Trends Ecol. Evol.* **17**, 34–39 (2002).
23. J. Degenhardt, T. G. Köllner, J. Gershenzon, Monoterpene and sesquiterpene synthases and the origin of terpene skeletal diversity in plants. *Phytochemistry* **70**, 1621–1637 (2009).
24. L. Cao, P. Zhang, D. F. Grant, An insect farnesyl phosphatase homologous to the N-terminal domain of soluble epoxide hydrolase. *Biochem. Biophys. Res. Commun.* **380**, 188–192 (2009).
25. P. Nyati *et al.*, Farnesyl phosphatase, a Corpora allata enzyme involved in juvenile hormone biosynthesis in *Aedes aegypti*. *PLoS One* **8**, e71967 (2013).
26. M. Frey *et al.*, Analysis of a chemical plant defense mechanism in grasses. *Science* **277**, 696–699 (1997).
27. V. Cantagrel *et al.*, SRD5A3 is required for converting polyprenol to dolichol and is mutated in a congenital glycosylation disorder. *Cell* **142**, 203–217 (2010).
28. A. Jozwiak *et al.*, POLYPRENOL REDUCTASE2 deficiency is lethal in *Arabidopsis* due to male sterility. *Plant Cell* **27**, 3336–3353 (2015).
29. M. B. Jones, J. N. Rosenberg, M. J. Betenbaugh, S. S. Krag, Structure and synthesis of polyisoprenoids used in N-glycosylation across the three domains of life. *Biochim. Biophys. Acta* **1790**, 485–494 (2009).
30. A. M. Takos *et al.*, Genomic clustering of cyanogenic glucoside biosynthetic genes aids their identification in *Lotus japonicus* and suggests the repeated evolution of this chemical defence pathway. *Plant J.* **68**, 273–286 (2011).
31. N. B. Jensen *et al.*, Convergent evolution in biosynthesis of cyanogenic defence compounds in plants and insects. *Nat. Commun.* **2**, 273 (2011).
32. F. Beran *et al.*, Novel family of terpene synthases evolved from trans-isoprenyl diphosphate synthases in a flea beetle. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 2922–2927 (2016).
33. J. Lancaster *et al.*, De novo formation of an aggregation pheromone precursor by an isoprenyl diphosphate synthase-related terpene synthase in the harlequin bug. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E8634–E8641 (2018).
34. J. Lancaster *et al.*, An IDS-type sesquiterpene synthase produces the pheromone precursor (Z)-α-Bisabolene in *Nezara viridula*. *J. Chem. Ecol.* **45**, 187–197 (2019).
35. R. R. Bodemann *et al.*, Precise RNAi-mediated silencing of metabolically active proteins in the defence secretions of juvenile leaf beetles. *Proc. Biol. Sci.* **279**, 4126–4134 (2012).
36. M. A. Birkett, J. A. Pickett, Aphid sex pheromones: From discovery to commercial production. *Phytochemistry* **62**, 651–656 (2003).
37. C. A. M. Campbell *et al.*, Responses of the aphids *Phorodon humuli* and *Rhopalosiphum padi* to sex pheromone stereochemistry in the field. *J. Chem. Ecol.* **29**, 2225–2234 (2003).
38. C. A. M. Campbell *et al.*, Sex attractant pheromone of damson-hop aphid *Phorodon humuli* (Homoptera, aphididae). *J. Chem. Ecol.* **16**, 3455–3465 (1990).
39. F. Fernandez *et al.*, The CWH8 gene encodes a dolichyl pyrophosphate phosphatase with a luminally oriented active site in the endoplasmic reticulum of *Saccharomyces cerevisiae*. *J. Biol. Chem.* **276**, 41455–41464 (2001).
40. J. S. Rush, S. K. Cho, S. Jiang, S. L. Hofmann, C. J. Waechter, Identification and characterization of a cDNA encoding a dolichyl pyrophosphate phosphatase located in the endoplasmic reticulum of mammalian cells. *J. Biol. Chem.* **277**, 45226–45234 (2002).
41. D. K. Ro *et al.*, Induction of multiple pleiotropic drug resistance genes in yeast engineered to produce an increased level of anti-malarial drug precursor, artemisinic acid. *BMC Biotechnol.* **8**, 83 (2008).
42. N. D. Lackus *et al.*, Identification and characterization of *trans*-isopentenyl diphosphate synthases involved in herbivory-induced volatile terpene formation in *Populus trichocarpa*. *Molecules* **24**, 2408 (2019).
43. S. Kumar, G. Stecher, K. Tamura, MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* **33**, 1870–1874 (2016).
44. T. G. Köllner *et al.*, Biosynthesis of iridoid sex pheromones in aphids. NCBI BioProject. https://www.ncbi.nlm.nih.gov/sra/?term=PRJNA866370. Deposited 5 August 2022.
45. T. G. Köllner *et al.*, Acyrthosiphon pisum clone JML06. GenBank. https://www.ncbi.nlm.nih.gov/nuccore?term=ON862918. Deposited 27 June 2022.
46. T. G. Köllner *et al.*, Acyrthosiphon pisum clone JML06. GenBank. https://www.ncbi.nlm.nih.gov/nuccore?term=ON862919. Deposited 27 June 2022.
47. T. G. Köllner *et al.*, Acyrthosiphon pisum clone JML06. GenBank. https://www.ncbi.nlm.nih.gov/nuccore?term=ON862920. Deposited 27 June 2022.

Downloaded from https://www.pnas.org by 95.90.176.231 on January 27, 2023 from IP address 95.90.176.231.