De novo formation of an aggregation pheromone precursor by an isoprenyl diphasphate synthase-related terpene synthase in the harlequin bug

Jason Lancaster, Ashot Khrimian, Sharon Young, Bryan Lehner, Katrin Luck, Anna Wallingford, Saikat Kumar B. Ghosh, Philipp Zerbe, Andrew Muchinski, Paul E. Marek, Michael E. Sparks, James G. Tokuhisa, Claus Tittiger, Tobias G. Köllner, Donald C. Weber, Dawn E. Gundersen-Rindal, Thomas P. Kuhar, and Dorothea Tholl

*Department of Biological Sciences, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061; *Invasive Insect Biocontrol and Behavior Laboratory, Agricultural Research Service, US Department of Agriculture, Beltsville, MD 20705; *Department of Biochemistry and Molecular Biology, University of Nevada, Reno, NV 89557; *Department of Biochemistry, Max Planck Institute for Chemical Ecology, D-07745 Jena, Germany; *Department of Entomology, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061; *Department of Plant Biology, University of California, Davis, CA 95616; and *Department of Biochemistry, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061

De novo formation of an aggregation pheromone precursor by an isoprenyl diphasphate synthase-related terpene synthase in the harlequin bug

Insects use a diverse array of specialized terpene metabolites as pheromones in intraspecific interactions. In contrast to plants and microbes, which employ enzymes called terpene synthases (TPSs) to synthesize terpene metabolites, limited information from few species is available about the enzymatic mechanisms underlying terpene pheromone biosynthesis in insects. Several stink bugs (Hemiptera: Pentatomidae), among them severe agricultural pests, release 15-carbon sesquiterpenes with a bisabolene skeleton as aggregation pheromones. The harlequin bug, Murgantia histrionica, a specialist pest of crucifers, uses two stereoisomers of 10,11-epoxy-1-bisabolen-3-ol as a male-released aggregation pheromone called murgantiol. We show that MhTPS (MhIDS-1), an enzyme unrelated to plant and microbial TPSs but with similarity to trans-isopropyl diphasphate synthases (IDSs) of the core terpene biosynthetic pathway, catalyzes the formation of (15,65,7R)-1,10-bisaboladien-1-ol (sesquipiperitol) as a terpene intermediate in murgantiol biosynthesis. Sesquipiperitol, a so-far-unknown compound in animals, also occurs in plants, indicating convergent evolution in the biosynthesis of this sesquiterpene. RNAi-mediated knockdown of MhTPS mRNA confirmed the role of MhTPS in murgantiol biosynthesis. MhTPS expression is highly specific to tissues lining the cuticle of the abdominal sternites of mature males. Phylogenetic analysis suggests that MhTPS is derived from a trans-IDS progenitor and diverged from bona fide IDSs, including MhIDS-2, which functions as an (E,E)-farnesyl diphasphate (FPF) synthase. Structure-guided mutagenesis revealed several residues critical to MhTPS and MhFPFS activity. The emergence of an IDS-like protein with TPS activity in M. histrionica demonstrates that de novo terpene biosynthesis evolved in the Hemiptera in an adaptation for intraspecific communication.

Hemiptera | Pentatomidae | harlequin bug | aggregation pheromone | terpene synthase

Terpenes play an important role in mediating chemical interactions of microbes, plants, and animals. In particular, volatile terpene compounds function as long- and short-distance semiochemicals in organismal interactions. Insects are well known to release volatile terpenes as interspecific signals in chemical defense or as alarm, aggregation, and sex pheromones in intraspecific communication. Despite these important functions, little is known about the formation of terpene specialized metabolites in insects.

In bacteria, fungi, and plants, volatile terpenes with 10-carbon (monoterpene) and 15-carbon (sesquiterpene) scaffolds are produced from the isoprenyl diphasphates, geranyl diphasphate (GPP) and farnesyl diphasphate (FPF), respectively, by enzymes called terpene synthases (TPSs). As central intermediates in the core terpene metabolism, GPP and trans- or cis-FPF are assembled by trans- or cis-isoprenyl diphasphate synthases (IDSs) in condensations of the 5-carbon unit dimethylallyl diphasphate (DMAPP) with one or two molecules of its isomer isopentenyl diphasphate (IPP). Despite their low sequence similarity, plant- and microbial-type TPS and IDSs share a common functionally active α-helical domain and catalyze reactions initiated by diphasphate ionization and carbocation formation (24–26). These structural and mechanistic similarities have led to the hypothesis of a possible ancient re-cruitment of IDS enzymes from core metabolism to specialized TPS function. Although insects employ a core terpene metabolism that supports juvenile hormone biosynthesis and has been associated with pheromone production (27, 28), to date no plant- or microbial-type TPS genes have been reported from insect genomes. The absence of these genes has led to the general hypothesis that terpenes have been acquired de novo from prokaryotes.

**Significance**

Many insects release volatile terpenes for chemical communication. However, the biosynthetic origin and evolution of these infochemicals are mostly unknown. We show that the harlequin bug, Murgantia histrionica, a stink bug pest (Hemiptera) of crucifer crops, produces a terpene aggregation pheromone by an enzyme that is unrelated to microbial and plant terpene synthases. M. histrionica terpene synthase activity is highly sex- and tissue-specific and makes a sesquiterpene alcohol, so far unknown in animals, as pheromone precursor. The enzyme evolved from ancestral isoprenyl diphasphate synthases and provides new evidence for de novo biosynthesis of terpenes in hemipteran insects. Knowledge of pheromone biosynthesis in stink bugs may lead to the development of new controls of these pests.

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

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

This open access article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

Data deposition: Sequences reported in this paper have been deposited in the GenBank database (accession nos. MG662378.1 for MhTPS and MG662379.1 for MhFPFS).

To whom correspondence should be addressed. Email: tholl@vt.edu.

This article contains supporting information online at www.pnas.org/lookup/supp/doi/10.1073/pnas.1800008115/-/DCSupplemental.

Published online August 23, 2018.
notion that insects are unable to synthesize terpene specialized metabolites de novo by the use of TPS enzymes and instead largely depend on the sequestration of terpene precursors from their host plants (15). However, studies on the biosynthesis of aggregation pheromones in Coleoptera (beetles) have shown that trans-IDS-like enzymes are able to convert GPP or FPP to terpene pheromones or their respective precursors. Gilg et al. (29, 30) demonstrated that the bark beetle *Phyllophaga striolata*, which synthesizes the cyclic sesquiterpene aggregation pheromone (6R,7S)-bimachala-9,11-diene from (Z,E)-FPP by an IDS-type enzyme (31). From nine different IDS-type transcripts in the *P. striolata* transcriptome, two were found to encode bona fide trans- or cis-IDS enzymes, while four transcripts encode TPSs, suggesting an evolutionary origin of these enzymes from IDS progenitors. Whether insects other than beetles, especially those of earlier evolutionary origin, biosynthesize volatile terpenes de novo is unknown.

Here, we show that within the Hemiptera, stink bugs (Pentatomidae) use IDS-like proteins in terpene biosynthesis. Several pentatomids release sesquiterpene aggregation/sex pheromones with a bisabolene carbon skeleton (32–39). Among these, the harlequin bug *Murganta histrionica*, a crucifer specialist, produces a mixture of (3S,6S,7R,10S) and (3S,6S,7R,10R) stereoisomers of 10,11-epoxy-1-bisabolen-3-ol as a male-released aggregation pheromone dubbed murgantiol (40–42). We demonstrate that an enzyme with homology to IDS proteins (*MhTPS*) converts (E,E)-FPP 2 to sesquipiperitol 3 as the presumed stereospecific alcohol precursor of murgantiol (Fig. 1), while a second trans-IDS protein (*MhFPPS*) catalyzes the formation of the *MhTPS* substrate (E,E)-FPP from IPP and DMAPP. *MhTPS* is transcribed at high levels in males, with a predominant localization in the subcuticular tissue of the abdominal sternites. A significant role of *MhTPS* in pheromone biosynthesis was confirmed by RNAi-mediated knockdown of *MhTPS* mRNA in *M. histrionica* males leading to reduced emission of murgantiol. Phylogenetic comparison of the *M. histrionica* enzymes with other insect IDS proteins suggests that in the Hemiptera proteins with TPS activity evolved from a trans-IDS progenitor. Sequence- and structure-guided mutagenesis identified several residues with critical function in *MhTPS* and *MhFPPS* activity. Together, our study suggests that, in comparison with

### Figure 1

Functional characterization of *MhIDS-1* (*MhTPS*) from *M. histrionica*. Recombinant *MhIDS-1* protein was expressed in *E. coli* and Sf9 cells and partially purified by affinity chromatography. Proteins were incubated with (E,E)-FPP 2 in the presence of Mg²⁺ and products were analyzed by GC-MS. (A) GC-MS chromatograms of enzyme products. Sf9 control cells express a housefly cytochrome P450 reductase. *: nonenzyme product; 3 (15,6S,7R)-sesquipiperitol; Cu, γ-curcumene; far, (2E,6E)-farnesol; sesq, β-sesquiphellandrene; zi, α-zingiberene. (B) Mass spectra of enzymatic products with (15,6S,7R)-sesquipiperitol standard 3. (C) Formation of sesquipiperitol 3 by *M. histrionica* TPS activity (boxed). A putative single or two-step pathway to murgantiol 1 is shown involving isomerization and epoxidation reactions. EV, empty vector; MdCPR, *Musca domestica* cytochrome P450 reductase.
microbial- and plant-type TPSs, insect TPS enzymes evolved more recently from IDS proteins and this event has occurred in multiple insect lineages including hemipteran insects.

Results

Identification and Functional Characterization of IDS-Like Genes in *M. histrionica*. We hypothesized that in the murgantiol biosynthetic pathway of *M. histrionica* an enzyme with terpene synthase activity synthesize a bisabolene-type hydrocarbon or alcohol terpene precursor, which presumably undergoes further modification(s) including oxidation to form the pheromone end products (Fig. 1). To identify genes involved in the formation of the murgantiol pheromone precursor, we queried a transcriptome dataset from different sexes and developmental stages of *M. histrionica* (43) with plant and microbial TPS sequences and *trans*-IDS-type sequences of *I. pini* FPPS (AAAX56631.1) and GPPS/TPS (AAAX56632.1), *Bombbyx mori* (silk moth) FPPS1 (NP_001036889.1) and FPPS2 (NP_001093001.1), and *Drosophila melanogaster* (fruit fly) FPPS (NP_477380). While no genes with sequence similarity to plant or microbial TPS genes could be identified, two *trans*-IDS-like sequences (*MhIDS-1* and *MhIDS-2*) annotated as FPP synthase (FPPS)-like genes were retrieved. According to the transcriptome results by Sparks et al. (43), *MhIDS-1* (GECO1420512.1; MG662378.1) mRNA levels are low in mature females but show a ~3-fold higher accumulation in mature males, while *MhIDS-2* (GECO1414919.1; MG662379.1) transcript levels are equal in both sexes. cDNAs of both genes were amplified from RNA extracted from mature male bugs. *MhIDS-1* encodes a 385-aa protein of 44.30 kDa while the *MhIDS-2* protein contains 405 aa residues and has a size of 46.36 kDa (*SI Appendix, Fig. S1*). *MhIDS-1* shares 22% sequence identity with insect FPPSs from *I. pini, D. melanogaster*, and the black bean aphid *Aphis fabae* (AAX334882.2), whereas comparatively higher sequence identities with these proteins (40–45%) were found for *MhIDS-2*.

To determine the biochemical function of the detected *IDS*-like genes, cDNAs encoding full-length proteins were cloned in the bacterial protein expression vector pEXP-5, generating an *N*-terminal histidine-tag fusion. When tested for sesquiterpene synthase activity with (E,E)-FPP 2 as a substrate, the partially purified recombinant *MhIDS-1* protein produced a terpene alcohol as its main product (Fig. 1A). Using GC-MS, we identified the alcohol product as sesquipiperitol 3, a sesquiterpene alcohol with a bisabolane skeleton, which is found in different plant species (44–46). The identification of sesquipiperitol was performed by comparisons of mass spectra and retention indices and further verified by chemical correlations (Fig. 1A and B and *SI Appendix, SI Results, SI Materials and Methods*, and Fig. S2). Sesquipiperitol was also the main product in assays upon cleavage of the *N*-terminal histidine tag (Fig. 1A, *SI Appendix, Fig. S3A*). In addition to sesquipiperitol, small and varying amounts of the sesquiterpene olefins, γ-curcumene, zingibereine, and β-sesquiphellandrene were detected (Fig. 1A). Hot sample injection contributed to the formation of the olefin products by thermal dehydration of sesquipiperitol, as could be shown in contrast to cool-on-column injection (Fig. *SI Appendix, Fig. S3B*). Sesquipiperitol was also produced, although at lower levels, from (Z,Z)-FPP, but almost no enzymatic activity was found with (Z,Z)-FPP as the substrate (Fig. *SI Appendix, Fig. S3C*). The recombinant enzyme did further convert GPP to several monoterpene internal structures (Fig. 1A, *SI Appendix, Fig. S3C*). However, when incubated with IPP and DMAPP, no formation of terpene products was observed, indicating that *MhIDS-1* was unable to synthesize prenyl diphosphates for subsequent conversion into terpene products (*SI Appendix, Fig. S3C*). Accordingly, formation of FPP by *MhIDS-1* was not observed. Because of its TPS activity and lack of IDS activity, we designate *MhIDS-1* hereinafter as *MhTPS*. We further tested the activity of *MhTPS* produced in insect *Sf9* cells (Fig. 1A). Recombinant *MhTPS* protein expressed without a histidine tag under these conditions generated the same enzymatic products upon incubation with (E,E)-FPP as those produced by the bacterially expressed enzyme (Fig. 1A). An alignment of the *MhTPS* amino acid sequence with those of *I. pini* and *P. striolata* TPS proteins suggested the presence of a putative N-terminal targeting peptide, although the Rxs motif indicative of a mitochondrial targeting sequence is absent from the *MhTPS* protein (*SI Appendix, Fig. S1*). Truncation of *MhTPS* (M3–R3) resulted in the loss of enzymatic activity.

Kinetic analysis of *MhTPS* with (E,E)-FPP as the substrate revealed an apparent *Km* value of 4.0 ± 0.7 μM and a *Vmax* of 675.3 ± 53.7 pmol/mg. The *kcat* value was 0.03 ± 0.003 s⁻¹ and *kcat/Km* was 7.5 ± 0.5 × 10⁻⁶ s⁻¹·μmol⁻¹. *Km*, *kcat*, and *kcat/Km* values of *MhTPS1* were similar to those of plant sesqui-TPS enzymes such as (E)-β-arylethylene synthase from *Artemisia annua* (47).

Recombinant *MhTPS1* did not produce zingiberenol 4 as a possible precursor of murgantiol. We tested whether changes in cofactor type and concentration or modifications of pH conditions would alter the enzymatic product profile and activity. No change in product specificity was observed when Mg²⁺ was substituted with Co²⁺, although this metal ion has been found to modify product specificity of a GPP/FPP synthase in the leaf beetle *Phaedon brassicae* (PfIDS-1) (48). Activity increased by approximately twofold between pH 0.1 and 10 mM Mg²⁺, while the opposite was the case for Co²⁺ (*SI Appendix, Fig. S3D*). Activity was highest at pH 7 (*SI Appendix, Fig. S3E*) and no change in product outcome was found under lower or higher pH conditions.

In contrast to *MhTPS1*, partially purified recombinant *MhIDS-2* protein did not show any TPS activity when assayed with different isomers of FPP or GPP as substrates. Instead, *MhIDS-2* produced (E,E)-FPP from IPP and DMAPP, indicating that this protein functions as a *trans*-IDS (*SI Appendix, Fig. S4*). The enzyme was unable to synthesize any other isomer of FPP, which suggests that (E,E)-FPP is the main isomeric form produced by *M. histrionica* (*SI Appendix, Fig. S4*). Removal of a putative transit peptide (M1–S30) (*SI Appendix, Fig. S1*) led to a substantially higher production of (E,E)-FPP by the truncated *MhIDS-2* protein (*SI Appendix, Fig. S4*). Because of its FPPS activity, we designate *MhIDS-2* hereinafter as *MhFPPS*.

Absolute Configuration of Sesquipiperitol. To further support the role of sesquipiperitol as a precursor of murgantiol, we determined the stereospecific configuration of sesquipiperitol at C-6 and C-7, which we predicted to be the same as that of murgantiol (*SI Appendix, SI Results, SI Materials and Methods*, and Figs. S2 and S5). Oxidation of enzymatically produced sesquipiperitol to sesquiperipentol 5 concluded a relative 65,7R or 6R,7S configuration (*SI Appendix, Fig. S2*). Further conversion of sesquipiperitol to bisabolane 7 determined the configuration at C-7 to be (R) (*SI Appendix, Fig. S5A*). This result unambiguously confirmed a 65,7R configuration of sesquipiperitol 3, which is identical to the C-6, C-7 configuration of murgantiol 1. Chiral GC analysis and 2D NMR recordings determined an (S) configuration at C-1 (*SI Appendix, SI Results, SI Materials and Methods*, and Fig. S5 B and C).

Sex- and Tissue-Specific Expression of *MhTPS*. We compared transcript abundance of *MhTPS* and *MhFPPS* between different sexes, developmental stages, and tissues by RT-PCR and qRT-PCR analyses. In agreement with mRNA levels determined by transcriptome analysis (43), *MhTPS* transcript levels were significantly higher (~37-fold) in mature males than in mature and immature females (Fig. 2A and *SI Appendix, Fig. S6A*). *MhTPS* transcript abundance in mature males was also significantly higher than that of nymphs and immature males, which do not emit murgantiol (Fig. 2A). As expected from transcriptome data (43), *MhFPPS* transcript levels were comparatively higher than those of *MhTPS*.
in mature females (SI Appendix, Fig. S6A). Tissue-specific transcripts of MhTPS localized to the tissue lining the cuticle of the abdominal sternites of mature males, while 200- to 2,000-fold lower transcript levels of MhTPS were observed in the thorax, fat body, and midgut (Fig. 2B and SI Appendix, Fig. S6B).

**Sesquipiperol Synthase Activity in M. histrionica Protein Extracts.** To further confirm that the enzymatic activity for the formation of sesquipiperol is present in protein lysates of M. histrionica, we extracted protein from the cuticle-specific tissue of mature male bugs and incubated the lysate with (E,E)-FPP. GC-MS analysis of hexane extracts obtained from the aqueous phase of the assay demonstrated the specific formation of sesquipiperol in the presence of (E,E)-FPP (Fig. 3A). By contrast, protein extracts from cuticle-associated tissue of mature females did not produce sesquipiperol, confirming the male-specific biosynthesis of this compound (Fig. 3A). Sesquipiperol was not synthesized in extracts of the male head and thorax, fat body, or gut tissue when incubated with (E,E)-FPP, which further supports the tissue specificity of this enzymatic reaction (Fig. 3B). We did not find further conversion of sesquipiperol to the putative downstream intermediate zingiberenol or to murgantiol or other products, suggesting that possible downstream enzymatic activities were not supported or potentially inhibited under our selected extraction and assay conditions. Alternatively, compound conversion might have been hampered by the loss of tissue integrity and compartmentalization.

**Verification of in Vivo MhTPS Function in Murgantiol Biosynthesis.** To verify the in vivo role of MhTPS in murgantiol biosynthesis, we injected males 3–5 d posteclosion with MhTPS-derived sequence-specific dsRNA. Transcript abundance was significantly reduced (~25- to 40-fold) 13–15 d postinjection compared with males injected with lacZ or Ringer’s solution, or non-injected males (P < 0.0001) (Fig. 4A). Pheromone analysis of males at the same time postinjection showed significantly lower emission (approximately sevenfold) of murgantiol in MhTPS dsRNA males than control males (Fig. 4B), confirming a substantial role of MhTPS in murgantiol biosynthesis.

**Phylogenetic Analysis of M. histrionica TPS and FPPS and Determination of Residues with Function in TPS and IDS Catalytic Activity.** A phylogenetic analysis based on maximum likelihood was performed to assess the evolutionary relationship of M. histrionica TPS and FPPS with other insect IDS and TPS enzymes (Fig. 5). The dataset included trans-IDS proteins from Coleoptera, Lepidoptera, and Hemiptera with known GPPS/FPPS or FPPS activities, the GPPS/TPS from I. pini, and the recently characterized IDS and TPS enzymes from P. striolata. To compare the relationship of the M. histrionica enzymes with similar enzymes in the Pentatomidae, we retrieved IDS-like sequences from the brown marmorated stink bug *Halyomorpha halys* based on publicly available transcriptome datasets of this insect (49). For *H. halys* two IDS-like sequences were identified, of which the IDS-1 sequence clusters with the Pentatomidae, while the second sequence is more closely related to *H. halys* IDS-2 protein and positioned in a larger clade of insect proteins with bona fide trans-FPPS or GPPS/FPPS activity from Coleoptera, Lepidoptera, and Hemiptera (Fig. 5). A broader phylogenetic analysis including insect GPPS/FPPS and plant trans-IDS proteins supports an evolutionary origin of the pentatomid TPSs together with the coleopteran TPSs from a trans-IDS progenitor that gave rise to proteins with GPPS/FPPS or TPS activities (SI Appendix, Fig. S7).

**Fig. 2.** Relative transcript abundance of MhTPS in M. histrionica determined by qRT-PCR. (A) MhTPS transcript abundance at different developmental stages and in different sexes. Young is 3-d adult; mature is 15-d adult. (n = 3, ±SD). (B) MhTPS transcript abundance in different tissues of adult males. Ab. Cuticle, tissue lining the abdominal cuticle; Ab. ST, Abdominal soft tissue minus midgut. (n = 3, ±SD). Gene expression was normalized against 18S RNA. Transcript abundance is shown relative to that in nymphs (A) or the male head tissue (B) (set to 1). Significance was determined using one-way ANOVA and means grouped by Tukey’s honestly significant difference test.

**Fig. 3.** Terpene synthase activity in crude protein extracts of M. histrionica by sex (A) and tissue (B). (A) Tissue of the abdominal cuticle of mature male or female M. histrionica was homogenized in assay buffer and assayed with 100 μM (E,E)-FPP. (B) Different tissues from mature male M. histrionica were homogenized in assay buffer and assayed with 50 μM (E,E)-FPP. Volatile compounds were extracted with an equal volume of hexane and analyzed with GC-MS. 3, sesquipiperol; far, (E,E)farnesol.
Sequence comparisons between insect IDS and IDS-derived TPS enzymes show that both types of proteins contain a conserved first and second aspartate-rich motif (FARM and SARM, respectively) (SI Appendix, Fig. S1). These motifs are required for the coordinated binding of Mg\(^{2+}\) ions with the allylic substrate to initiate catalysis through carbocation formation and rearrangement (24–26). While residue substitutions occur in the SARM of TPSs from I. pini and P. striolata, no such substitutions are present in the TPSs of stink bugs (SI Appendix, Fig. S1). Notably, in comparison with bona fide trans-IDS proteins, the TPS proteins substitute aromatic with nonaromatic amino acid residues at positions 4 and 5 upstream of the FARM (SI Appendix, Fig. S1, box M1). To determine whether these substitutions affect the orientation of (E,E)-FPP at the active site of the enzyme, we positioned (E,E)-FPP in the active site of crystallized G. gossypii FPPS and an M. histrionica TPS homology model. This comparative positioning via ligand docking indicated that the amino acid changes upstream of the FARM (GgFPPS F112, F113 – MtTPS M118, S119) likely cause a different orientation of the prenyl side chain of the FPP substrate (Fig. 6A). Accordingly, M118F and S119F substitutions through site-directed mutagenesis resulted in the loss of TPS activity of the recombinant MtTPS protein (Fig. 6B). By contrast, the reciprocal substitutions F95M and F96S in a truncated MtFPPS protein did not abolish ID activity but converted the enzyme into a 20-carbon geranylgeranyl diphosphate (GGPP) synthase (SI Appendix, Fig. S8B). In comparison with (E,E)-FPP, docking of DMAPP in both MtTPS and GgFPPS models did not result in different positions of this allylic diphosphate (SI Appendix, Fig. S9A). However, DMAPP did not have substantial inhibitory effects on the MtTPS-catalyzed reaction of (E,E)-FPP to sesquiperipetal, indicating a limited affinity of the enzyme for this allylic diphosphate (SI Appendix, Fig. S9B).

To further define determinants of TPS or IDS catalytic specificity, we identified amino acids that are exclusively conserved in IDS proteins of related insect species. Amino acids conserved exclusively in TPSs (SI Appendix, Fig. S8) were used to predict active sites of TPS enzymes with known or putative TPS activity are highlighted by the gray box.

**Discussion**

**M. histrionica TPS Functions as a Sesquiterpene Alcohol Synthase.** We found that the TPS activity associated with the recombinant MtTPS enzyme and with crude lysates of male M. histrionica stereospecifically convert (E,E)-FPP to the (13,65,7R) isomer of sesquipiperitol. The synthesis of sesquipiperitol most likely proceeds by formation of a bisabolyl cation followed by a hydride shift and subsequent quenching of the carbocation by water (SI Appendix, Fig. S10). The 65,7R configuration of sesquipiperitol corroborates its function as an intermediate in the formation of sesquiterpene alcohol synthesis.
the murgantiol stereoisomers. An RNAi-based approach further confirmed the role of the MhTPS-catalyzed reaction in murgantiol biosynthesis. The 10,11-epoxy-1-bisabolene-3-ol stereoisomers also constitute pheromone components of the brown marmorated stink bug *H. halys* (32), suggesting that the IDS-1 gene of *H. halys* has a function similar to that of *MhTPS*. Moreover, related TPS activities presumably operate in other pentatomids to make bisabolene-type sex pheromones or their corresponding precursors, such as isomers of zingiberenol in the rice stalk stink bug, *Tibraca limbavitentris*, and the rice stink bug *Oebalus picipes* (34, 37); the sesquiterpenes alpha-cucumene, zingiberene, and beta-sesquiphellandrene in the red-shouldered stink bug *Thyanta pallidovirens* (38); and (Z)-alpha-bisabolene epoxides in the Southern green stink bug *Nezara viridula* (35, 50). Until our finding of sesquipiperitol as an enzymatic product in *M. histrionica*, this terpenoid and its biosynthesis had not been reported in animals. The occurrence of sesquipiperitol as a natural product in different plant species (44–46) suggests convergent biosynthetic evolution of this compound by both plant- and insect-derived TPS proteins. Since several terpenes are used as identical semiochemicals by both plants and insects, it can be assumed that similar processes of convergent evolution have occurred in the synthesis of these compounds.

**Terpene Formation in *M. histrionica* Is Highly Sex- and Tissue-Specific.**

According to the sex-specific release of murgantiol, we found that *MhTPS* lost activity when nonaromatic amino acids at positions 4 and 5 upstream of the FAR motif were substituted with aromatic amino acids of equivalent position in insect IDSs. This effect was likely caused by a modified position of the FPP substrate. Molecular docking of (E,E)-FPP in the active-site cavity of an *MhTPS* homology model showed an extension of the FPP prenyl side chain into the cavity (Fig. 6A), which appears to be critical for facilitating a subsequent cyclization and formation of the terpene product. For IDS enzymes, substitutions of aromatic and nonaromatic amino acids upstream of the FARM determine the chain length of the enzymatic product: nonaromatic residues facilitate the synthesis of products with extended prenyl chains in long-chain trans-IDS enzymes (≥C20) (54). Indeed, substitution of the two Phe residues with Met and Ser upstream of the FARM in *MhTPP* resulted in a change from FPPS to GGPPS activity (*SI Appendix, Fig. S8B*). This change in IDS product specificity correlates with the presence of nonaromatic residues in other insect GGPP synthases (55).

The basic residues of this motif interact with the diphosphate moiety of IPP (56). Therefore, the loss of *MhTPS* activity was also observed when residues Ser235→DAW were substituted with KKxR, a highly conserved motif in insect IDS proteins. The charged side chain of the introduced Arg residue facing toward the hydrocarbon tail of (E,E)-FPP in this *MhTPS* variant may prevent further cyclization (*SI Appendix, Fig. S9C*). The KKxR motif in insect IDSs is equivalent to the KRLR motif in the FPPS of *Escherichia coli*. The basic residues of this motif interact with the diphosphate moiety of IPP (56). Therefore, the loss of *MhTPS* activity by reversely substituting the KKxR motif with the SDAW sequence is most likely due to improper binding of the IPP substrate.

Reciprocal substitutions of the highly conserved Lys and Cys residues downstream of the FARM of TPS or IDS proteins, respectively, also caused a loss of activity in both mutants. Although, according to our model, these residues are positioned on a loop without immediate proximity to the docked substrate (*SI Appendix, Fig. S9D*), they may be involved in substrate binding and/or in the closure of the active site upon substrate binding, as has been shown, for example, for loop residues on the β-domain of TPS enzymes (57).

In summary, we were able to identify residues with critical function in TPS and IDS activity. Further structural analysis, ideally via protein crystallization and combinatorial mutations paired with the identification of epistatic residue networks (58), will be necessary to fully identify the residues controlling the transition from insect IDS to TPS enzyme activity. This evolutionary transition likely required a combination of residue substitutions to change substrate affinities and specificities for

**IDS-Type Terpene Synthases Evolved in the Hemiptera.**

Our finding of a functional TPS in *M. histrionica* suggests that adaptations of terpene-specialized metabolism for intraspecific communication have occurred in multiple lineages throughout insect evolution. To date, IDS-type TPSs with similar functions have only been characterized in the coleopterans *I. pini* and *P. striolata* (29–31). Phylogenetic analysis indicates that pentatomid and coleopteran TPSs are derived from a trans-IDS progenitor that diverged early from true trans-IDS enzymes. The separate clustering of pentatomid TPSs from *P. striolata* and *I. pini* TPSs (Fig. 5 and *SI Appendix, Fig. S7*) suggests that proteins with TPS activity might have emerged independently in these and other insect taxa. However, phylogenetic comparisons with additional TPSs from other insect lineages will be necessary to further support this hypothesis.

In our approach to identify residues that discriminate between TPS or IDS catalytic activity, we found that *MhTPS* lost activity when nonaromatic amino acids at positions 4 and 5 upstream of the FAR motif were substituted with aromatic amino acids of equivalent position in insect IDSs. This effect was likely caused by a modified position of the FPP substrate. Molecular docking of (E,E)-FPP in the active-site cavity of an *MhTPS* homology model showed an extension of the FPP prenyl side chain into the cavity (Fig. 6A), which appears to be critical for facilitating a subsequent cyclization and formation of the terpene product. For IDS enzymes, substitutions of aromatic and nonaromatic amino acids upstream of the FARM determine the chain length of the enzymatic product: nonaromatic residues facilitate the synthesis of products with extended prenyl chains in long-chain trans-IDS enzymes (≥C20) (54). Indeed, substitution of the two Phe residues with Met and Ser upstream of the FARM in *MhTPP* resulted in a change from FPPS to GGPPS activity (*SI Appendix, Fig. S8B*). This change in IDS product specificity correlates with the presence of nonaromatic residues in other insect GGPP synthases (55).

Loss of *MhTPS* activity was also observed when residues Ser235→DAW were substituted with KKxR, a highly conserved motif in insect IDS proteins. The charged side chain of the introduced Arg residue facing toward the hydrocarbon tail of (E,E)-FPP in this *MhTPS* variant may prevent further cyclization (*SI Appendix, Fig. S9C*). The KKxR motif in insect IDSs is equivalent to the KRLR motif in the FPPS of *Escherichia coli*. The basic residues of this motif interact with the diphosphate moiety of IPP (56). Therefore, the loss of *MhTPS* activity by reversely substituting the KKxR motif with the SDAW sequence is most likely due to improper binding of the IPP substrate.

Reciprocal substitutions of the highly conserved Lys and Cys residues downstream of the FARM of TPS or IDS proteins, respectively, also caused a loss of activity in both mutants. Although, according to our model, these residues are positioned on a loop without immediate proximity to the docked substrate (*SI Appendix, Fig. S9D*), they may be involved in substrate binding and/or in the closure of the active site upon substrate binding, as has been shown, for example, for loop residues on the β-domain of TPS enzymes (57).

In summary, we were able to identify residues with critical function in TPS and IDS activity. Further structural analysis, ideally via protein crystallization and combinatorial mutations paired with the identification of epistatic residue networks (58), will be necessary to fully identify the residues controlling the transition from insect IDS to TPS enzyme activity. This evolutionary transition likely required a combination of residue substitutions to change substrate affinities and specificities for
DMAPP and IPP and gain a cyclization function following the binding of (E,E)-FPP as a single allylic substrate.

Materials and Methods
Two putative trans-IDS–like genes (MhIDS-1, MhIDS-2) identified in the transcriptome of M. histrionica were amplified as full-length cDNAs from RNA extracted from adult males and inserted in the prokaryotic expression vector pEXPS-N/TOPO, generating an N-terminal histidine tag. A truncated version of MhIDS-2 lacking a putative N-terminal transit peptide was cloned in the same vector, and a truncated version of MhIDS-1 was synthesized and cloned in the pET11b expression vector. For expression in insect cells, the full-length MhIDS-1 cDNA was cloned without an N-terminal His-tag into the BaculoDirect vector. Recombinant partially purified MhIDS-1 and MhIDS-2 proteins expressed in E. coli and lysates of Sf9 cells expressing MhIDS-1 were tested for TPS and IDS activities and the reaction products were analyzed by GC-MS and liquid chromatography tandem MS, respectively.

Identification and determination of the stereospecific configuration of the main enzymatic product of MhIDS-1 (MhTPS), sesquipiperitol, were performed by synthesis and chemical transformations in combination with chiral GC and NMR analysis. Kinetic properties of MhTPS were examined in assays using [1-3H](E,E)-FPP as substrate. Developmental, sex- and tissue-specific expression of the MhTPS transcript or MhTPS activity were determined by qRT-PCR and TPS activity assays of crude lysates. MhIDS-1 and MhTPS involvement in pheromone production was verified by injecting young males of M. histrionica with dsRNA of MhIDS or la3c (control) followed by qRT-PCR and pheromone analysis in the headspace. Phylogenetic analysis was performed using maximum likelihood. MhTPS and MhFPPS protein variants were generated by site-directed mutagenesis, expressed in E. coli, and analyzed for TPS and FPPS activity. Further details on organisms, experimental and analytical procedures, data analysis, phylogenetic analysis, protein homology modeling and substrate docking are provided in SI Appendix, SI Materials and Methods.

Acknowledgments
We thank Megan V. Herlihy and Jeremy K. Turner for assisting with harlequin bug rearing and volatile collection and Filadelfo Garay for assisting with these analyses. This work was supported by the Department of Agriculture National Institute of Food and Agriculture Grant 2016-6703-24759 (to D.T., D.E.G.-R., A.K., T.P.K., C.T., and D.C.W.) and the Virginia Polytechnic Institute and State University Departments of Biochemistry and Biological Sciences and the Translational Plant Program.

1. Gherselli J, Dudaeva N (2007) The function of terpene natural products in the natural world. Nat Chem Biol 3:408–416.
2. Osbourn A, Goss RJM, Field RA (2011) The saponins: Polar isoprenoids with important and diverse biological activities. Nat Prod Rep 28:1261–1268.
3. Tholl D (2015) Biosynthesis and biological functions of terpenoids in plants. Biochemistry of Isoprenoids. Advances in Biochemical Engineering/Biotechnology, eds Schrader J, Bohmman J (Springer, New York), Vol 148, pp 63–106.
4. Junker RR, et al. (March 3, 2017) Covariation and phenotypic integration in chemical communication displays: Biosynthetic constraints and eco-evolutionary implications. New Phytol. 10.1111/nph.14005.
5. Huang M, et al. (2012) The major volatile organic compound emitted from Arabidopsis thaliana flowers, the sesquiterpene (E)-caryophyllene, is a defense against a bacterial pathogen. New Phytol 193:997–1008.
6. Ragujo RA (2016) More lessons from linalool: Insights gained from a ubiquitous floral volatile. Cur Opin Plant Biol 21:32–36.
7. Rasmann S, et al. (2005) Recruitment of entomopathogenic nematodes by insect-damaged maize roots. Nature 434:732–737.
8. Robert C, et al. (2012) Herbivore-induced plant volatiles mediate host selection by a root herbivore. New Phytol 194:1061–1069.
9. Vaughan MM, et al. (2013) Formation of the unusual semivolatile diterpene rhizalene by the Arabidopsis class I terpene synthase TPS08 in the root stele is involved in defense against belowground herbivory. Plant Physiol 155:1108–1125.
10. Ditengou FA, et al. (2015) Volatile signalling by sesquiterpenes from ectomycorrhizal fungi reprogram root architecture. Nat Commun 6:6279.
11. Muller M, Buchbauer G (2011) Essential oil components as pheromones. A review. Flavour Fragrance J 26:357–377.
12. Sobotnik J, Jiroušová A, Hanuš R (2010) Chemical warfare in termites. J Insect Physiol 56:1012–1021.
13. Burse A, et al. (2009) Always being well prepared for defense: The production of deterrents by juvenile Chrysomelidae beetles (Chrysomelidae). Phytochemistry 70:1899–1909.
14. Pickett JA, Allemann RA, Tittiger C (2016) Pheromone production in bark beetles. Insect Biochem Mol Biol 40:699–712.
15. Brown AE, Riddick EW, Aldrich JR, Holmes WE (2006) Identification of (E)- beta-caryophyllene as a gender-specific terpene produced by the multicolored Asian lady beetle. J Chem Ecol 32:2489–2499.
16. Degenhardt J, Köllner TG, Gershenzon J (2009) Monoterpene and sesquiterpene synthases and the origin of terpene skeletal diversity in plants. Phytochemistry 70:1621–1637.
17. Chen F, Tholl D, Bohmnan J, Pichersky E (2011) The family of terpene synthases in plants: A mid-size family of genes for specialized metabolism that is highly diversified throughout the kingdom. Plant J 66:212–229.
18. Li J, Mafu S, Peters RJ (2014) To gibberellins and beyond! Surveying the evolution of diterpene metabolism. Annu Rev Plant Biol 65:259–286.
19. Jia Q, et al. (2016) Microbial-type terpene synthase genes occur widely in nonseed land plants, but not in seed plants. Proc Natl Acad Sci USA 113:12328–12333.
20. Moraes MCB, Pareja M, Laumann RA, Borges M (2008) The chemical volatiles (semi-chemicals) produced by neotropical stink bugs (Hemiptera: Pentatomidae). Neotrop Entomol 37:489–505.
21. Beran F, et al. (2016) Novel family of terpene synthases evolved from trans-isoprenyl diphosphate synthases in a flea beetle. Proc Natl Acad Sci USA 113:2922–2927.
22. Khmirin A, et al. (2014) Discovery of the aggregation pheromone of the brown marmorated stink bug (Halyomorpha halys) through the creation of stereosymmetric libraries of 1-bisabol-3-ols. J Nat Prod 77:1708–1717.
23. Aldrich JR, et al. (1993) Artifacts and pheromone blends from Nezara spp. and other stink bugs (Hemiptera, Pentatomidae). Z Naturforsch C Biosci 48:73–79.
24. Borges M, et al. (2006) Sex attractant pheromone from the rice stalk stink bug, Thrips tabaci (Hemiptera: Thripidae). J Chem Ecol 32:2749–2761.
25. Murgantia histrionica: A pheromone blend with multiple redundant components. J Chem Ecol 33:1787–1798.
26. Weber DC, Khmirin A, Blasiolli-Moraes MC, Millar JG (2018) Semiochemistry of Pentatomidae: Invasive Stink Bugs and Related Species (Pentatomidae): Biology, Higher Systematics, Semiochemistry, and Management, ed McPherson JE (CRC, Boca Raton, Fl), pp 677–725.
27. Khmirin A, et al. (2014) Determination of the stereochemistry of the aggregation pheromone of harlequin bug, Murgantia histrionica. J Chem Ecol 40:1260–1268.
28. Weber DC, et al. (2014) Attractiveness of harlequin bug, Murgantia histrionica, aggregation pheromone: Field response to isomers, ratios, and dose. J Chem Ecol 40:1251–1259.
29. Zahn DK, Moreira JA, Millar JG (2008) Identification, synthesis, and bioassay of a male-specific aggregation pheromone from the harlequin bug, Murgantia histrionica. J Chem Ecol 34:238–251.
30. Sparks ME, et al. (2017) A transcriptome survey spanning life stages and sexes of the harlequin bug, Murgantia histrionica. Insects 8:55.
31. Hoylmann F, Tsarvka E, Jakupovic J (1984) Sesquiterpenes and acetylceines from Aryantheum aduumum sp. jacobaeifolium. Phytochemistry 23:1103–1104.
32. Liu YK, Brown GD (1997) Oxygenated bisabolanes from Alpinia densinervia. Phytochemistry 45:537–544.
33. Cool LG (1996) Sesquiterpene alcohols from foliation of Fitzroya cupressoides. Phytochemistry 42:1015–1017.
34. Tai Y, et al. (2002) A DNA clone for beta-caryophyllene synthase from Artemisia annua. Phytochemistry 61:523–529.
35. Frick S, et al. (2013) Metal ion control product specificity of isoprenyl diphosphate synthases in the insect terpene pathway. Proc Natl Acad Sci USA 110:4194–4199.
36. Zhang ZQ, Zhang HE, Shelby KS, Zhang D, Gunderson-Rindal DE (2014) Transcriptome of the invasive brown marmorated stink bug, Halyomorpha halys (Stål) (Hemiptera: Pentatomaevae). PLoS One 9:e111646.
50. Aldrich JR, Oliver JE, Lusby WR, Kochansky JP, Lockwood JA (1987) Pheromone strains of the cosmopolitan pest, Nezara viridula (Heteroptera, Pentatomidae). J Exp Zool 244:171–175.

51. Cribb BW, Siriwardana KN, Walter GH (2006) unicellular pheromone glands of the pentatomid bug Nezara viridula (Heteroptera: Insecta): Ultrastructure, classification, and proposed function. J Morphol 267:831–840.

52. Qiu Y, et al. (2012) An insect-specific P450 oxidative decarbonylase for cuticular hydrocarbon biosynthesis. Proc Natl Acad Sci USA 109:14858–14863.

53. Martins GF, Ramalho-Ortigao JM (2012) Oenocytes in insects. Invert Surviv J 9:139–152.

54. Wallrapp FH, et al. (2013) Prediction of function for the polyisoprenyl transferase subgroup in the isoprenoid synthase superfamily. Proc Natl Acad Sci USA 110:E1196–E1202.

55. Lai C, McMahon R, Young C, Marky TFC, Langley CH (1998) quemao, a Drosophila bristle locus, encodes geranylgeranyl pyrophosphate synthase. Genetics 149:1051–1061.

56. Hosfield DJ, et al. (2004) Structural basis for bisphosphonate-mediated inhibition of isoprenoid biosynthesis. J Biol Chem 279:8526–8529.

57. Christianson DW (2017) Structural and chemical biology of terpenoid cyclases. Chem Rev 117:11570–11648.

58. Salmon M, et al. (2015) Emergence of terpene cyclization in Artemisia annua. Nat Commun 6:6143.

59. Le SQ, Gascuel O (2008) An improved general amino acid replacement matrix. Mol Biol Evol 25:1307–1320.