Insect Biochemistry and Molecular Biology 152 (2023) 103879

Contents lists available at ScienceDirect

Insect Biochemistry and Molecular Biology

journal homepage: www.elsevier.com/locate/ibmb

Ancient origin and conserved gene function in terpene pheromone and defense evolution of stink bugs and hemipteran insects

Zarley Rebholz [a,1], Jason Lancaster [a,1], Hailey Larose [a,1], Ashot Khrimian [b], Katrin Luck [c], Michael E. Sparks [b], Kerry L. Gendreau [a], Leena Shewade [d], Tobias G. Köllner [c], Donald C. Weber [b], Dawn E. Gundersen-Rindal [b], Paul O'Maille [d], Alexandre V. Morozov [e], Dorothea Tholl [a,*]

[a] Department of Biological Sciences, Virginia Tech, Latham Hall, 220 Ag Quad Lane, Blacksburg, VA, 24061, USA
[b] Invasive Insect Biocontrol and Behavior Laboratory, USDA Agricultural Research Service, Beltsville, MD, 20705, USA
[c] Department of Biochemistry, Max Planck Institute for Chemical Ecology, Hans-Knöll-Strasse 8, D-07745, Jena, Germany
[d] SRI International, Biosciences Division, 333 Ravenswood Avenue, Menlo Park, CA, 94025-3493, USA
[e] Department of Physics & Astronomy and Center for Quantitative Biology, Rutgers University, 136 Frelinghuysen Rd., Piscataway, NJ, 08854-8019, USA

ARTICLE INFO

Keywords:
Pheromone
Terpene
Terpene synthase
Isoprenyl diphosphate synthase
Pentatomid
Hemiptera

ABSTRACT

Insects use diverse arrays of small molecules such as metabolites of the large class of terpenes for intra- and inter-specific communication and defense. These molecules are synthesized by specialized metabolic pathways; however, the origin of enzymes involved in terpene biosynthesis and their evolution in insect genomes is still poorly understood. We addressed this question by investigating the evolution of isoprenyl diphosphate synthase (IDS)-like genes with terpene synthase (TPS) function in the family of stink bugs (Pentatomidae) within the large order of piercing–sucking Hemipteran insects. Stink bugs include species of global pest status, many of which emit structurally related 15-carbon sesquiterpenes as sex or aggregation pheromones. We provide evidence for the emergence of IDS-type TPS enzymes at the onset of pentatomid evolution over 100 million years ago, coinciding with the evolution of flowering plants. Stink bugs of different geographical origin maintain small IDS-type families with genes of conserved TPS function, which stands in contrast to the diversification of TPS genes in plants. Expanded gene mining and phylogenetic analysis in other hemipteran insects further provides evidence for an ancient emergence of IDS-like genes under presumed selection for terpene-mediated chemical interactions, and this process occurred independently from a similar evolution of IDS-type TPS genes in beetles. Our findings further suggest differences in TPS diversification in insects and plants in conjunction with different modes of gene functionalization in chemical interactions.

1. Introduction

Chemical interactions are ubiquitous among all forms of life. Organisms use blends of small molecules or semiochemicals to interact with individuals of their own or other species (Junker et al., 2018). Insects employ a large number of diverse semiochemicals for defense or as pheromones in intraspecific communication (Ishikawa, 2020; Tabata, 2017). Yet, knowledge of the biosynthesis and evolution of these specialized metabolites is still limited. One of the best studied groups of insect pheromones are hydrocarbons, alcohols, aldehydes, and esters (Blomquist and Ginzel, 2021; Castillo et al., 2012; Jurenka, 2021; Keeling et al., 2021; Montagne and Wicker-Thomas, 2021). These compounds are derived from fatty acids and, thus, evolved from precursors in core carbon metabolism. Whether other types of pheromone or defense compounds are equally derived from primary metabolic pathways is not well understood. Various insect taxa use compounds in the large class of terpenes (isoprenoids) as sex, aggregation, alarm and trail pheromones or as anti-aphrodisiacs and defense compounds (Havlíčková et al., 2019; Keeling et al., 2021; Krall et al., 1997; Pickett et al., 2013; Schulz et al., 2008; Weber et al., 2018). Whereas the evolution of terpene specialized metabolites related to chemical interactions is well documented in plants and microbes (e.g. Ballal et al.,

* Corresponding author.
E-mail address: tholl@vt.edu (D. Tholl).
[1] Zarley Rebholz, Jason Lancaster, and Hailey Larose contributed equally to the work.

https://doi.org/10.1016/j.ibmb.2022.103879
Received 4 September 2022; Received in revised form 24 November 2022; Accepted 25 November 2022
Available online 2 December 2022

2020; Chen et al., 2011; Degenhardt et al., 2009; Fischer et al., 2015; Salmon et al., 2015; Wei et al., 2020; Zi et al., 2014), their biosynthetic origin in insects has remained largely obscure. This lack of knowledge is partly due to assumptions that insects obtain terpene metabolites from their host plants. For example, male neotropical orchid bees (Euglossini) collect volatile terpenes and other volatiles from flowers to attract conspecific females (Pokorny et al., 2017). On the other hand, applications of labeled pheromone precursors as performed in the pine engraver beetle *Ips pini* (Curculionidae: Scolytinae) (Seybold et al., 1995) repeatedly indicated that terpene pheromones can be synthesized *de novo* via endogenous enzymatic pathways.

In central terpene metabolism, isoprenyl diphosphate synthases (IDSs) catalyze the condensation of the 5-carbon terpene precursors isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP) to form different size prenyl diphosphate intermediates including 10-carbon geranyl diphosphate (GPP), 15-carbon (*E*,*E*)-farnesyl diphosphate (FPP), which is the central precursor in insect juvenile hormone biosynthesis, and 20-carbon all-*trans* geranylgeranyl diphosphate (GGPP) (Noriega, 2014; Tholl, 2015). There is growing evidence that insects have evolved *trans*-IDS-type enzymes, which possess terpene synthase (TPS) activity and convert GPP or FPP into 10-carbon monoterpene and 15-carbon sesquiterpene pheromones or their respective precursors. For example, *I. pini* uses a bifunctional IDS/TPS enzyme that makes myrcene, the precursor of its aggregation pheromone ipsdienol, from GPP (Gilg et al., 2005, 2009). In males of the striped flea beetle *Phyllotreta striolata* (Chrysomelidae: Galerucinae), four members of a clade of six IDS-type enzymes, derived from FPP synthases, were found to function as sesquiterpene synthases (Beran et al., 2016), one of which (*Ps*TPS1) catalyzes the conversion of (*Z*,*E*)-FPP to the sesquiterpene himachala-9,11-diene as the major aggregation pheromone component. Moreover, in the nymphalid butterfly *Heliconius melpomene* a TPS has been identified that is related to *trans*-GGPP synthases but uses GPP as substrate to produce the monoterpene anti-aphrodisiac β-ocimene (Darragh et al., 2021).

Despite these findings, insufficient genomic information has largely prevented a more detailed investigation of the biosynthetic evolution of terpene semiochemicals in insects at the levels of gene structure, expression, and family organization. Pheromone evolution has been associated with the expansion and differential expression of biosynthetic gene families as has been shown, for example, in the formation of long chain hydrocarbon and fatty alcohol pheromones in *Drosophila* and bumble bees (Fang et al., 2009; Shirangi et al., 2009; Tupec et al., 2019). To what extent TPS genes and their functions have diversified or remained conserved in TPS gene families throughout insect evolution is unknown. To address these questions, we investigate the evolution of terpene biosynthesis in stink bugs (Hemiptera: Pentatomidae) and hemipteran insects in general. The order Hemiptera represents a large and diversified group of terrestrial and aquatic piercing-sucking insects, which undergo incomplete metamorphosis and retain many ancestral insect states (Panfilio and Angelini, 2018). Hemipteran insects include true bugs such as pentatomids in the suborder Heteroptera, aphids, whiteflies, and scale insects (suborder Sternorrhyncha), cicadas and planthoppers (suborder Auchenorrhyncha), and moss bugs (suborder Coleorrhyncha) (Johnson et al., 2018). Many species, especially in the Heteroptera and Sternorrhyncha, have been documented for their terpene pheromones and defense secretions (Table S1 and references therein). The occurrence of terpenes in these groups and even older insect orders such as termites and cockroaches (Blattodea) and stick insects (Phasmatodea) (Persoons et al., 1979; Piskorski et al., 2007; Smith et al., 1979) suggests an emergence of terpene specialized metabolism in insects as early as 300–350 million years ago (mya) during the radiation of seed plants.

Stink bugs are represented by a diverse group of species that include herbivores and carnivores as well as specialists and generalists with effective mechanisms of chemical communication (Čokl et al., 2017). Their ability to easily adapt to different environmental conditions has caused several species to become economically important pests with severe impact on many agricultural crops in the Neotropics and worldwide (Weber et al., 2018). Mature males of many pentatomids emit sesquiterpenes as sex or aggregation pheromones (Tholl, 2021; Weber et al., 2018). For instance, the southern green stink bug *Nezara viridula*, a globally invasive pest, releases the aggregation pheromone *trans-/-cis*-(*Z*)-α-bisabolene epoxide (Aldrich et al., 1987; Harris and Todd, 1980) (Fig. 1). We previously discovered an *N. viridula* IDS-type TPS enzyme, which catalyzes the conversion of (*E*,*E*)-FPP to (+)-(*S*,*Z*)-α-bisabolene as the likely precursor of the pheromone (Lancaster et al., 2019) (Fig. 1). In addition, our investigation of the harlequin bug, *Murgantia histrionica*, a crucifer specialist native to Central America (tribe Strachiini), led to the identification of an IDS-type TPS, which makes (1*S*,6*S*,7*R*)-1,10-bisaboladien-1-ol (sesquipiperitol) as a terpene intermediate in the biosynthesis of its aggregation pheromone 10,11-epoxy-1-bisabolen-3-ol named murgantiol (Khrimian et al., 2014a; Lancaster et al., 2018) (Fig. 1).

Interestingly, other pentatomid species release structurally identical or similar cyclic bisabolane-type terpene pheromones despite their different geographic origin or tribe-specific association. For instance, the aggregation pheromone of the brown marmorated stink bug, *Halyomorpha halys* (Stål) (tribe Cappaeini), an invasive generalist native to Asia, is structurally identical to murgantiol although of a different stereoisomeric composition (Khrimian et al., 2014a, 2014b) (Fig. 1). In addition, several other stink bug species of severe pest status in South America, the rice stink bugs *Oebalus poecilus* and *Mormidea v-luteum* and the rice stalk stink bug *Tibraca limbativentris* (Borges et al., 2006; de Oliveira et al., 2013; Moliterno et al., 2021) use sesquipiperitol or putative intermediates of the murgantiol pathway as pheromone constituents (Fig. 1). In this study, we seek not only to understand the biosynthetic origin of these overlapping pheromone chemistries to provide information on the functional evolution of IDS-type TPS genes in these stink bug species, but also to shed light on the more general question as to what extent pheromone biosynthetic enzymes diversify or remain conserved in insect or animal speciation in comparison to functionally related enzymes in plants and other organisms.

We show, based on genome analysis of IDS-type TPS gene families and phylogenetic comparison, that pentatomid IDS-type TPS genes evolved by an early paralogous divergence of two TPS clades with clade-specific conservation of TPS gene function upon speciation, albeit with different tissue specificity. Our results suggest a limited degree of functional divergence in terpene pheromone evolution in an insect family, which stands in contrast to the high level of TPS gene diversification in plants (Chen et al., 2011) and the dynamics of other gene families in pheromone evolution. Gene mining and phylogenetic analysis of IDS-like genes in other hemipteran lineages further provide evidence for an ancient divergence of IDS-type TPS genes from IDS (FPPS) genes in correlation with the evolution of terpene semiochemicals in hemipteran insects.

2. Materials and Methods

2.1. Chemicals

Isoprenyl diphosphate substrates were from Isoprenoids.com. Unless otherwise specified, all other reagents and solvents were purchased from Sigma-Aldrich.

2.2. Insects

A greenhouse colony of *H. halys* was started from late instar nymphs collected from insecticide-free vegetable plots in Beltsville, MD, USDA. Insects were reared in ventilated plastic cylinders (21 cm × 21 cm o. d.) on a diet of organic green beans and shelled raw sunflower seeds and buckwheat seeds (2:1 w/w), glued onto squares of brown wrapping paper with wheat-based wallpaper paste and distilled water supplied in

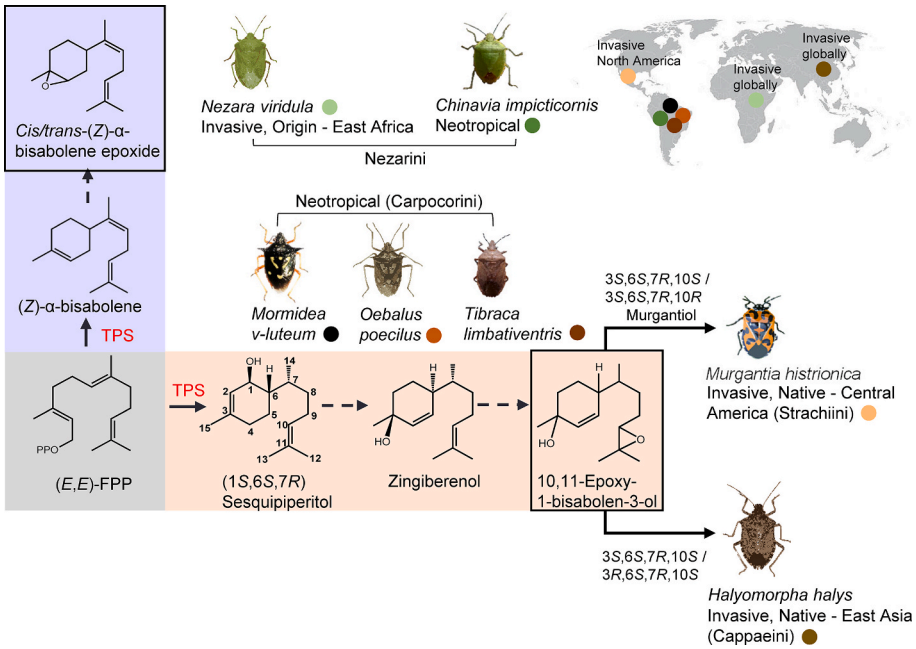

Fig. 1. Sesquiterpene aggregation pheromones and their precursors in stink bugs (pentatomids) with pest status in the Neotropics or worldwide. The harlequin bug *Murgantia histrionica* and the brown marmorated stink bug *Halyomorpha halys* share the same pheromone compound (10,11-epoxy-1-bisabolen-3-ol, boxed) although with different stereoisomeric composition. The precursor sesquipiperitol is made by an IDS-type terpene synthase (TPS) enzyme from the central intermediate (*E,E*)-FPP. Depicted species in the tribe Carpocorini release pheromones of the sesquipiperitol/10,11-epoxy-1-bisabolen-3-ol (murgantiol) complex including zingiberenol, a putative pathway intermediate. The southern green stink bug *Nezara viridula* and its relative *Chinavia impicticornis* emit *cis/trans*-(*Z*)-α-bisabolene epoxide isomers as major pheromone constituents (boxed). The precursor (*Z*)-α-bisabolene is produced by an IDS-type TPS from (*E,E*)-FPP in *N. viridula*. Dashed arrows indicate putative enzymatic steps. The inset shows native regions and current distributions of species. Depicted insects are modified from images by Gabor Vetek (*N. viridula*), Douglas Napier (*C. impicticornis*), Insetologia (*M. v-luteum*), Natasha Wright, Braman Termite & Pest Elimination, Bugwood.org (*O. poecilus*), Lucas Rubio (*T. limbativentris*), and Daniel Frank (*M. histrionica*).

two cotton-stopped 7 cm × 2 cm o. d. Shell vials held together with a rubber band. Insects were reared in a climate controlled growth chamber (25 ± 5 °C, 16:8 h L:D, 65% RH). Eggs were collected weekly and hatched in plastic Petri dishes with a water vial, and after molting to second-instars, the nymphs were transferred to the larger rearing containers as described above for the remaining four instars. Newly eclosed adults were removed from cages three times weekly and moved to new cages, isolating males and females. Insects were kept until the immature (2–3 days post molt) or mature (14–15 days post molt) adult stage.

2.3. Crude protein extracts of H. halys

Fifteen-day old virgin male and female *H. halys* were used to prepare crude protein extracts. Whole bugs were frozen in liquid nitrogen, pulverized with a mortar and pestle, and suspended in assay buffer (25 mM HEPES, 5 mM $MgCl_2$, 10% glycerol, 0.5 mM PMSF, 1 mM DTT, pH 7). Samples were centrifuged for 15 min at 16000×*g* and the supernatant collected. Protein concentration was determined with a Bradford Assay (Bio-Rad) according to the manufacturer's instructions.

2.4. Identification and cloning of IDS-type TPS genes from H. halys and N. viridula

A search for putative IDS-type genes of *H. halys* was performed by a tBLASTn search in NCBI with organism restricted to *H. halys* and using the amino acid sequences of *Mh*TPS (MG662378.1) and *Mh*FPPS (MG662379.1) from *M. histrionica* (Lancaster et al., 2018), the GPPS/TPS (AAX55632.1) and FPPS (AAX55631.1) sequences from *I. pini*, and the TPS1 (KT959248) sequence from *P. striolata* (Beran et al., 2016; Gilg et al., 2009) as queries. Sequences of *HhIDS1-7* (Table S2, Fig. S1) were identified in the *H. halys* genome assembly (GCA_000696795.1; official gene set halhal_OGSv1.1 - https://data.nal.usda.gov/dataset/halyomorpha-halys-official-gene-sets-v10-and-v11) and transcriptome (SRA acc. no. SRP040652) (Sparks et al., 2014, 2020). Total RNA was extracted from individual mature (14–15 days post molt) *H. halys* males and females using Trizol reagent (Invitrogen, Thermo Fisher Scientific) according to the manufacturer's protocol. RNA was DNase treated with RQ1 DNase I (Promega, Madison, WI) and purified using the RNeasy Plant Mini Kit (Qiagen). Full length cDNAs for *HhIDS1*, *HhIDS2* and *HhIDS7* were generated from total male RNA using GoScript reverse transcriptase (Promega) and primers listed in Table S3. Target sequences were amplified using Q5 proofreading DNA polymerase (New England Biolabs) and ligated into the pGEM-T Easy vector (Promega). Sequences were verified before cloning into expression vectors. cDNAs for *HhIDS3*, *HhIDS4*, *HhIDS5* and *HhIDS6* could not be amplified from male or female RNA. Sequences for *HhIDS1* (*HhTPS2*), *HhIDS2* (*HhFPPS*) and *HhIDS7* (*HhTPS1*) were verified by Sanger sequencing and have been deposited in the GenBank database under accession numbers MG870387 (*HhTPS1*), MG917093 (*HhTPS2*) and MG870389 (*HhFPPS*).

A putative sesquipiperitol synthase was identified from *N. viridula* using publicly available transcriptome data with tBLASTn searches and the functionally characterized sesquipiperitol synthase *Mh*TPS (MG662378.1) from *M. histrionica* as the query sequence. Total RNA was extracted from fat body tissue of five mature (14–15 days post molt) *N. viridula* males using Trizol Reagent according to the manufacturer's protocol. RNA was DNase treated with RQ1 DNase I (Promega, Madison, WI) and purified using the RNeasy Plant Mini Kit (Qiagen). Full length *NvTPS2* cDNA was generated from total RNA using the SuperScript II Reverse Transcription Kit (Invitrogen) and Q5 proofreading DNA polymerase. Amplification was performed with primers carrying attb sites for Gateway cloning (Table S3). The resulting amplicon was transferred by Gateway BP reaction (Invitrogen) into the donor vector pDONR/Zeo (Invitrogen) and sequenced. The *NvTPS2* sequence has been deposited in the GenBank database under accession number ON934605.

2.5. Heterologous expression of recombinant HhFPPS

Full-length *HhFPPS* (*HhIDS2*) was amplified from the pGEM-T Easy construct with Q5 DNA polymerase and cloned into the pEXP5-NT/TOPO expression vector containing an N-terminal 6x histidine tag (Invitrogen). The resulting construct was transformed into *Escherichia coli* strain BL21(DE3)pLysS (Life Technologies). *E. coli* cultures (50 mL) were grown at 37 °C and 220 rpm and induced with 1 mM IPTG after reaching an OD_{600} of 0.6. Cells were cultivated at 18 °C for another 18 h, centrifuged for 15 min at 5,000×*g* at 4 °C, and resuspended in 2 mL extraction buffer (50 mM Tris HCl [pH 7.5], 20 mM imidazole, 300 mM NaCl, 10% glycerol (v/v), 5 mM $MgCl_2$, 2 mM DTT) supplemented with 0.3 mg/mL lysozyme (AppliChem), 2.5 U/mL benzonase (Novagen) and proteinase inhibitors (Protease Inhibitor Mix HP, SERVA). Following

incubation at 4 °C for 30 min on ice, cells were disrupted with a sonicator (Bandelin UW2070, 50%, 4 × 30 s treatment) and lysates centrifuged at 4 °C for 30 min at 15000×*g* to obtain soluble fractions. Recombinant proteins were purified using Ni-NTA Spin Columns (Qiagen) according to the manufacturer's instructions. In preparation for enzymatic assays, the buffer was exchanged with 25 mM 3-(N-morpholino)-2-hydroxypropanesulfonic acid (MOPSO, pH 7.2, 10% [v/v] glycerol, 1 mM DTT, 5 mM $MgCl_2$) using PD-10 Desalting Columns (GE Healthcare Life Sciences).

2.6. Heterologous expression of recombinant HhTPS1, HhTPS2, and NvTPS2

Due to difficulties in the cloning process, *HhTPS1* (*HhIDS7*) *and HhTPS2* (*HhIDS1*) were synthesized by GenScript (Piscataway, NJ) with codon optimization for *E. coli* and cloned into the pET19b bacterial expression vector with fusion to an N-terminal histidine tag. The vector construct was transformed into *E. coli* BL21(DE3)pLysS cells (Invitrogen) and single colonies were selected at 18 °C on LB with ampicillin (100 μg/mL) and chloramphenicol (34 μg/mL). Cells were grown overnight in 5 mL LB plus antibiotics at 37 °C prior to inoculation of 200 mL of the same medium. Following cultivation at 18 °C for 4–8 h to an OD_{600} of 0.50, protein expression was induced with 0.5 mM IPTG and the cells were cultivated for another 48 h. Cells were then washed in 100 mL wash buffer (20 mM Tris-HCl, 50 mM KCl, pH 7), pelleted and resuspended in 15 mL cell lysis buffer (50 mM NaH_2PO_4, 300 mM NaCl, 5 mM imidazole, 0.5 mM PMSF, 2 mM DTT, pH 8) followed by sonication on ice for 2 × 30 s (1 min interval) at 20% amplitude (Branson Digital Sonifier). The supernatant was purified with Ni-NTA agarose (Qiagen) using three washes with wash buffer (50 mM Tris-HCl, 500 mM NaCl, 10% glycerol, 1 mM DTT, pH 8) containing 30 mM imidazole and the target protein was eluted in a 1 mL fraction with elution buffer (50 mM Tris-HCl, 500 mM NaCl, 250 mM imidazole, 10% glycerol, 1 mM DTT, pH8) prior to desalting into TPS assay buffer (25 mM HEPES, 10 mM $MgCl_2$, 10% glycerol, pH 7) with PD MiniTrap G-25 desalting columns (GE Healthcare).

NvTPS2 was cloned into the pDEST-His-MBP expression vector (Addgene plasmid # 11,085) via Gateway LR cloning (Invitrogen) and transformed into *E. coli* BL21(DE3)pLysS cells (Nallamsetty et al., 2005). Single colonies were selected at 37 °C on LB with ampicillin (100 μg/mL) and chloramphenicol (34 μg/mL). Expression cultures (50 mL of auto-induction media, MilliporeSigma) were started from 5 mL overnight LB cultures and incubated at 18 °C for 72 h. Cells were pelleted and washed in 50 mL pellet wash buffer (see above), pelleted, and resuspended in 15 mL cell lysis buffer (50 mM Tris-HCl, 500 mM NaCl, 20 mM imidazole, 10% glycerol, 1% Tween-20, 0.5 mM PMSF, 1 mM DTT, pH 8). Cells were then sonicated as described. The recombinant protein was purified with Ni-NTA agarose using two washes of wash buffer containing 20 mM imidazole and eluted as 500 μL fractions with elution buffer as described above.

2.7. IDS activity assay and analysis

To determine IDS activity of *Hh*FPPS, enzyme assays were performed with 96 μl of Ni-NTA purified protein mixed with 2 μl 50 μM isopentenyl diphosphate (IPP) and 2 μl 50 μM dimethylallyl diphosphate (DMAPP) and incubated at 30 °C for 2 h. IDS enzyme products were analyzed according to Beran et al. (2016) with an Agilent 1260 HPLC system (Agilent Technologies) coupled to an API 5000 triple-quadrupole mass spectrometer (AB Sciex Instruments).

2.8. TPS activity assay

To determine TPS activity in crude protein extracts from whole male and female *H. halys* bugs, extracts (100 μg total protein) were incubated with 100 μM (*E,E*)-FPP in 25 mM HEPES, 5 mM $MgCl_2$, 10% glycerol, pH 7, at a final volume of 200 μL. Assays were overlaid with an equal volume of hexane and incubated for 6 h at 30 °C. Terpene products were extracted by vigorous mixing for 15 s and the organic phase was separated by centrifugation at 4000×*g* for 10 min prior to drying over $MgSO_4$. One μl of sample was used for GC-MS analysis. Assays with the recombinant *Hh*TPS proteins were performed in assay buffer (see above) in a total volume of 100 μl with Ni-NTA purified protein, 1 mM DTT, and 50 μM of (*E,E*)-, (*Z,E*)- or (*Z,Z*)-FPP. Assays were incubated at 30 °C for 1 h with a 100 μL hexane overlay followed by product extraction and analysis as described above. TPS activity of purified, recombinant *Nv*TPS2 protein was determined in assay buffer (25 mM HEPES, 10% glycerol, pH 7) with 1 mM DTT, 10 mM $MgCl_2$ and 60 μM of (*E,E*)-, (*Z,E*)- or (*Z,Z*)-FPP in a total volume of 1 mL. Reactions were incubated at 30 °C for 30 min and assay products adsorbed in the headspace by automated solid phase microextraction (SPME) prior to GC-MS analysis.

2.9. Quantitative (real Time)-Reverse transcription PCR (qRT-PCR) analysis

For sex-specific gene expression analysis from whole bugs, mature (15 days post molt) male and female insects were fixed by adding hexane to a sealed jar and then frozen in liquid nitrogen. Single males and females were used as individual replicates. For tissue-specific gene expression analysis, male insects were killed with hexane vapor in a screw top jar and dissected in PBS. For each replicate, five insects were dissected with the tissues frozen in liquid nitrogen between dissections. Each replicate contained the pooled tissue of five individuals. Tissue groups were midgut, abdominal soft tissue minus midgut, and abdominal cuticle including attached epithelial cells. Samples were kept at −80 °C prior to RNA extraction. RNA was extracted from whole bugs or pooled tissue samples with the Trizol reagent as described above. cDNAs were generated from total RNA using the SuperScript II Reverse Transcription Kit (Invitrogen). Relative transcript abundance was measured by qRT-PCR using the delta-delta Ct method and normalized to the 30 S ribosomal protein S4 (RpS4) (Livak and Schmittgen, 2001). Primers were designed to amplify a fragment of approximately 100 bp using Geneious (v. 2022.0.1) and tested for non-specific binding. Reaction plates contained 2 μl cDNA (5 ng/μl), 0.6 μL of each primer (300 nM final concentration), 6.8 μl dH_2O and 10 μl PowerSYBR Green PCR Master Mix (Applied Biosystems, Thermo Fisher Scientific) per well. The samples were analyzed using Applied Biosystems QuantStudio6 Pro with default settings (50 °C 2 min, 95 °C 10 min followed by 40 cycles of 95 °C 15 s, 60 °C 1 min). Primers were tested for non-specific amplification by analyzing the dissociation curve after PCR. Significance was measured using Student's t-test or one-way ANOVA.

2.10. Gas chromatography-mass spectrometry (GC-MS) analysis of enzyme products

Enzymatic products obtained from TPS assays with crude protein extract or purified recombinant enzymes of *H. halys* were analyzed by GC-MS applying liquid injection at 240 °C in a splitless mode (crude protein) or in a split 5 mode (purified protein). SPME-adsorbed products from assays with recombinant *Nv*TPS2 were thermally desorbed at 240 °C and analyzed in splitless mode. Compounds were separated on a GC-2010 gas chromatograph (Shimadzu) using a 30 m × 0.25 mm i. d. X 0.25 μm film Zebron ZB-XLB column (Phenomenex) coupled with a QP2010S mass spectrometer (Shimadzu). The GC program was as follows: 40 °C with 2 min hold, then raised to 220 °C at 5 °C/min, then raised to 240 °C at 70 °C/min followed by a 2 min hold time. Mass spectrometry was performed with an ion source temperature of 240 °C, interface temperature of 280 °C, electron ionization (EI) potential of 70 eV and scan range of 50–400 atomic mass units. Helium was used as a carrier gas at 1.4 or 1.9 mL/min. Terpene olefin and alcohol products were identified by comparison of retention times and mass spectra with those of authentic standards.

2.11. Determination of the stereospecificity of the HhTPS1 sesquipiperitol product

To identify the stereospecific configuration of the sesquipiperitol product of *Hh*TPS1, several stereoisomers were produced for chemical correlations as described previously by Lancaster et al. (2018). Oxidation of the *Hh*TPS1 product to sesquipiperitone and conversion to bisabolanes by dehydration/hydrogenation was performed following the procedures described by Lancaster et al. (2018). Gas chromatographic comparison of the bisabolane derivatives of *Hh*TPS1-produced sesquipiperitol with bisabolane standards obtained by dehydration/hydrogenation from stereo-defined zingiberenols was performed on a Hydrodex β-6TBDM column (25 m × 0.25 mm ID) under the following conditions: 50 °C (5 min) to 90 °C at 10 °C/min; H_2 was the carrier gas at a flow rate of 1.5 mL/min; the injection temperature was 200 °C; the detector temperature (FID) was 200 °C. GC-MS analyses of synthetic *R*,*S*,*R*-sesquipiperitol and *S*,*S*,*R*-sesquipiperitol and the *Hh*TPS1-derived sesquipiperitol were conducted in electron impact (EI) ionization mode at 70 eV with an Agilent Technologies 5973 mass selective detector interfaced with a 6890 N GC system and equipped with a 30 m × 0.25 mm x 0.25 μm Rtx 1701 column. The column temperature was maintained at 40 °C for 5 min, then raised to 240 °C at 7 °C/min, then raised to 270 °C at 15 °C/min. Injections were done splitless at 260 °C. Helium was used as a carrier gas at 1 mL/min. Analyses of sesquipiperitones were conducted on the GC-MS instrument mentioned above using a 30 m × 0.25 mm x 0.25 μm HP-5 GC column. The column temperature was maintained at 40 °C for 5 min, then raised to 240 °C at 7 °C/min, then raised to 270 °C at 15 °C/min. Injections were done splitless at 260 °C. Helium was used as a carrier gas at 1 mL/min.

2.12. Homology modeling and residue analysis

Homology models of FPPS and TPS protein structures were generated by submitting full-length amino-acid sequences to the SWISS-MODEL server (Bertoni et al., 2017; Bienert et al., 2017; Guex et al., 2009; Studer et al., 2020; Waterhouse et al., 2018). Model quality was evaluated based on the QMEANDisCo Global score (Studer et al., 2020). Homology models were visually inspected and analyzed using UCSF Chimera (Pettersen et al., 2004) by super positioning models on available crystallographic structures of IDS enzymes and complexes from the RCSB Protein Data Bank (Berman et al., 2000). Substrates were docked by manual positioning of ligands (e.g. FPP) from superimposed crystallographic complexes. Substrate binding residues (and positions) were identified using UCSF Chimera in tandem with Cytoscape (Shannon et al., 2003). In brief, network models of crystallographic complexes were created using the Residue Interaction Network Generator (RING) server (Clementel et al., 2022), and imported into Cytoscape for analysis. Substrate binding residues were extracted from network models as a list of first neighbor nodes that directly contact the substrate. Pyrophosphate and isoprenyl tail-binding designations were attributed by visual inspection of protein complexes and comparing interatomic distances. All structural figures were rendered using UCSF Chimera.

2.13. Phylogeny of IDS (FPPS) and TPS proteins from pentatomids and other functionally characterized insect IDS and IDS-type TPS enzymes

Protein sequences of the *H. halys*, *M. histrionica*, and *N. viridula* IDS-type families were aligned with previously characterized IDSs and IDS-type TPSs from other insects and a four vertebrate FPPS outgroup using Clustal Omega (v1.2.2) (Sievers and Higgins, 2018) with automatically adjusted settings. To reduce noise in the phylogeny resulting from poorly aligned and/or saturated sites, the alignment was trimmed using Gblocks (v0.91) (Castresana, 2000) with the following parameters: a maximum number of contiguous non-conserved positions of 9, a minimum block length of 6, and half allowed gap positions. A maximum likelihood phylogeny was inferred using IQTree (v2.0.4) with aBayes, SH-aLRT (x10,000), and UFBoot2 (x10,000) support (Anisimova et al., 2011; Diep Thi et al., 2018; Guindon et al., 2010; Minh et al., 2020). An additional phylogeny was inferred with Bayesian methodology using MRBAYES (v3.2.5) (Huelsenbeck and Ronquist, 2001) with 10, 000, 000 Markov chain Monte Carlo generations run over four chains and sampling every 100 iterations with the first 25% of samples discarded as burn-in. The JTT + I + G substitution model was used for both maximum likelihood and Bayesian phylogenetic inference and was selected with Bayesian Information Criterion using Modelfinder (Kalyaanamoorthy et al., 2017), limited to models supported by both IQTree and MRBAYES. The resulting consensus phylogenies were visualized and edited using iTOL (v6.5.2) (Letunic and Bork, 2021).

2.14. Expanded Hemiptera IDS-like protein phylogenies

Potential hemipteran IDS-like gene transcripts were mined using a tBLASTn-based approach. Only FPPS-like genes were considered excluding putative GGPP synthase-like genes and long chain polyprenyl diphosphate synthases. All *H. halys* IDS and IDS-type TPS protein sequences were used as search queries in tBLASTn searches of the NCBI Nucleotide collection and Transcriptome Shotgun Archive. The search was restricted to sequences from hemipteran species with a maximum number of 5000 target sequences and other parameters set to default. All obtained sequences (as of November 2021) were downloaded and duplicate sequences by accession number were removed. Open reading frames beginning with ATG and above 900 bp in length were predicted using the "Find ORFs" function in Geneious Prime (v.2021.1.1), extracted, and translated. Predicted coding sequences that did not contain any IDS characteristic motifs were removed as were duplicates or near-duplicates from the same species, with preference given to the longer sequence and/or sequences obtained from the NCBI Nucleotide collection in the case when sequences were near the same length. The translated sequences were aligned and trimmed with Gblocks as described above with the inclusion of an outgroup of five characterized coleopteran FPPS proteins generating a final alignment of 306 sequences, containing 241 aligned positions. The alignment was used to infer a maximum likelihood phylogeny using IQTree (v2.0.4) with aBayes, SH-aLRT (x1,000), and UFBoot2 (x1,000) support and applying the LG + R6 substitution model selected by Modelfinder by Bayesian Information Criterion. The consensus phylogeny was visualized and annotated in iTOL.

2.15. Selection analysis of IDS and IDS-type TPS genes

A codon alignment of pentatomid IDS-type TPS genes and hemipteran IDS (FPPS) genes with known or inferred function was generated using Clustal Omega with default parameters in conjunction with the Translation Align function in Geneious Prime (v 2021.1.1). The alignment was trimmed using Gblocks with the parameters described above. A consensus maximum likelihood phylogeny was inferred using IQTree with SH-aLRT (x1000), UFboot (x1000), and aBayes support using the codon substitution model, MG+F3X4+R3, selected by Modelfinder. To detect and compare differences in gene-wide substitution ratios (ω) (i.e., dn:ds, the ratio of substitution rates at non-synonymous sites to those at synonymous sites) between IDS and IDS-type TPS clades in the phylogeny, branch models were fit using PAML (Yang, 2007) utilizing the graphical user interface, EasyCodeML (v1.4) (Gao et al., 2019). The likelihood of a two-ratio branch model (in which the average gene-wide ω of the TPS clade was permitted to vary from ω of the IDS background) was compared to that from a one-ratio model (in which a uniform gene-wide ω was estimated for all branches) using a likelihood ratio test (LRT). Branch-site models were also fit with codeML (PAML v4.9j) to compare site-specific selection pressures across branches in the phylogeny and to test for evidence of positive selection that may be obscured in gene-wide ω estimations. Evidence for positive selection was statistically assessed with LRTs comparing H_0 models (fix_omega = 1,

fixing foreground ω_2 values equal to 1 as for the background) to H_1 models (fix_omega = 0, in which foreground ω_2 was permitted to be estimated ≥1, indicative of positive selection). Multiple-testing corrections were applied using the Bonferroni correction method according to Anisimova and Yang (2007).

3. Results

3.1. Identification and functional characteristics of an IDS-like gene family in H. halys

For a deeper investigation of the evolution and function of IDS-like TPS genes involved in pentatomid pheromone biosynthesis, we first mined IDS-like genes in the available *H. halys* genome assembly and transcriptome (Sparks et al., 2014, 2020) using query sequences of the functionally characterized IDS-type TPS and FPPS proteins from *M. histrionica* (Lancaster et al., 2018), the bifunctional GPPS/TPS from *I. pini*, and TPS1 from *P. striolata* (Beran et al., 2016; Gilg et al., 2009). Mining of the *H. halys* genome resulted in the identification of seven IDS-like sequences (*HhIDS1-7*, Table S2). Transcriptomic data obtained from different sexes and developmental stages showed two of the putative genes were highly expressed in both males and females (*HhIDS2* and *HhIDS7*, RNA IDs XM_014420915.1 and XM_014420697.1, respectively), one gene showed low but proportionally higher expression in males than females (*HhIDS1*, RNA ID XM_014433717.1), and a fourth (*HhIDS6*, RNA ID XM_014433739.1) was weakly expressed in both sexes (Sparks et al., 2014, 2017). No transcripts were reported for sequences *HhIDS3*, *HhIDS4*, and *HhIDS5* although full length sequences of these genes had been retrieved from the *H. halys* genome. In accordance with the observed transcript abundances, full length cDNAs could only be amplified for *HhIDS1*, *HhIDS2*, and *HhIDS7* from RNA extracted from mature *H. halys* males, while no amplicons were obtained from the remaining genes. The corresponding *Hh*IDS1 protein (43.78 kDa) contains 377 amino acids, while the *Hh*IDS2 protein (46.31 kDa) and the *Hh*IDS7 protein (41.98 kDa) contain 403 and 368 amino acids, respectively (Fig. S1).

Full length clones of these genes were then expressed in *E. coli* for functional characterization. Hexane extraction and GC-MS analysis of the volatile products released from recombinant *Hh*IDS7 protein incubated with (*E*,*E*)-FPP showed the presence of large amounts of sesquipiperitol (Fig. 2A and B, Fig. S2). The same product, although at lower amounts, was made in the presence of (*Z*,*E*)-FPP while no product was detected with (*Z*,*Z*)-FPP as the substrate (Fig. S3A). Assays with *Hh*IDS7 protein also showed trace amounts of the sesquiterpene olefins γ-curcumene, α-zingiberene and β-sesquiphellandrene (Fig. 2A; Fig. S2), which are likely thermal dehydration products of sesquipiperitol at high temperature (240 °C) of the GC injection port as was found by Lancaster et al. (2018). We did not test whether changes in cofactor type and modifications of pH conditions would alter the enzymatic product profile of the *H. halys* sesquipiperitol synthase as was reported previously for insect IDS enzymes (Frick et al., 2013). However, such modifications did not alter product specificity of the related *M. histrionica* sesquipiperitol synthase (Lancaster et al., 2018) suggesting a similar outcome for the *H. halys* enzyme.

To compare the configuration of the sesquipiperitol product of *Hh*IDS7 with that of the sesquipiperitol made by the *M. histrionica* TPS1 enzyme and the configuration of the *H. halys* pheromone components, we first determined the relative configuration at C-6 and C-7. Oxidation of the enzymatic product to sesquipiperitone defined a relative 6*S*,7*R* or 6*R*,7*S* configuration (Fig. S4). By converting sesquipiperitol to bisabolanes, we further determined the stereochemistry at C-7 to be *R*, which resulted in an absolute configuration of 6*S*, 7*R* (Fig. S5). Subsequent comparison of the *Hh*TPS7 product with 1*S*,6*S*,7*R* and 1*R*,6*S*,7*R* sesquipiperitols on a non-chiral Rtx 1701 column determined an *S*-

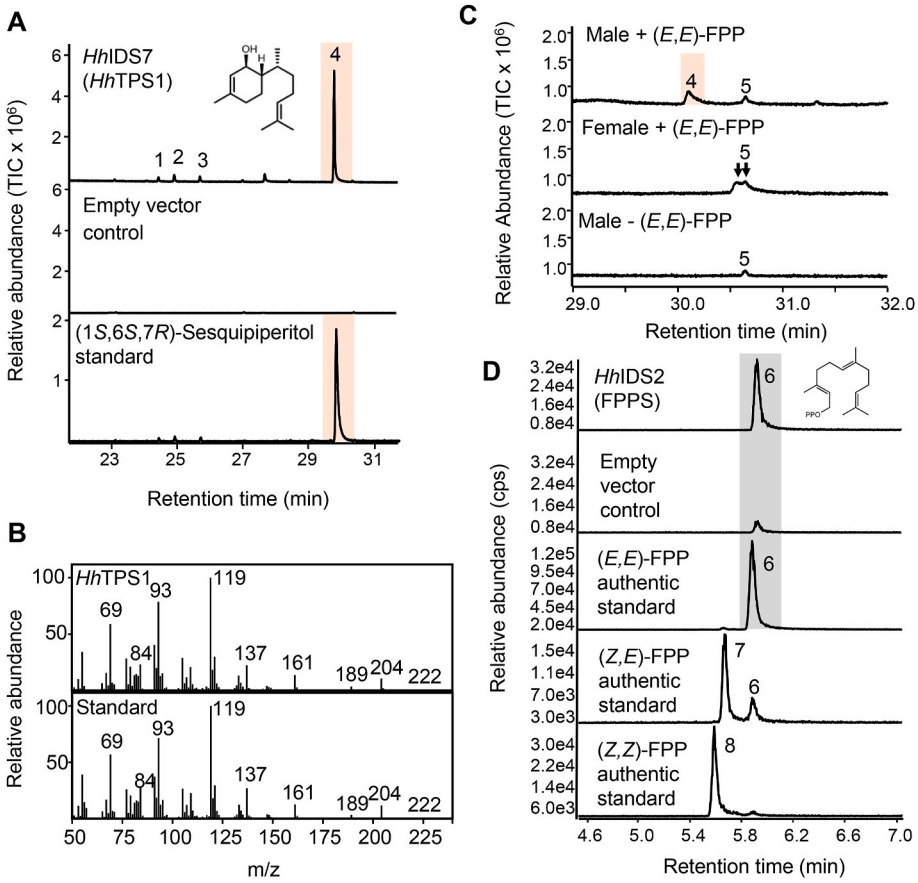

Fig. 2. Functional characterization of *Hh*IDS7 (*Hh*TPS1) and *Hh*IDS2 (*Hh*FPPS) from *H. halys*. (*A*) GC-MS analysis of products from an assay with recombinant *Hh*IDS7 in the presence of (*E*,*E*)-FPP. (*B*) Mass spectra of enzyme product 4 and synthetic (1*S*,6*S*,7*R*)-sesquipiperitol. Compounds 1–3 are considered thermal rearrangements products of sesquipiperitol. (*C*) Terpene synthase activity in crude protein extracts from male and female *H. halys*. Protein lysates were assayed with 100 μM (*E*,*E*)-FPP. Volatile products were extracted with hexane and analyzed by GC-MS. (*D*) LC-MS analysis of FPP products from assays with recombinant *Hh*IDS2 protein in the presence of 50 μM IPP and 50 μM DMAPP. Authentic FPP isomers were used for comparison. 1, γ-curcumene; 2, α-zingiberene; 3, β-sesquiphellandrene; 4, sesquipiperitol isomer; 5, farnesol isomers; 6, (*E*,*E*)-FPP; 7, (*Z*,*E*)-FPP; 8, (*Z*,*Z*)-FPP.

configuration at C-1 and thus an overall 1*S*,6*S*,7*R* stereochemistry (Fig. S6). This absolute configuration is identical to that of the sesquipiperitol produced by the *M. histrionica* TPS enzyme and the 6*S*,7*R* configuration matches those of the *H. halys* (and *M. histrionica*) pheromone products (Khrimian et al., 2014a, 2014b). We further confirmed the presence of sesquipiperitol synthase activity in crude protein lysates from whole mature male *H. halys* bugs (Fig. 2C). An enzymatic product with the retention time and mass spectrum identical to that of sesquipiperitol was identified in crude protein extracts of males but not females where only putative farnesol isomers were detected.

Expression of the recombinant *Hh*IDS1 protein and examination of TPS activity showed that the protein produced several sesquiterpene olefins and alcohols at low abundance when exposed to either (*E*,*E*)-FPP or (*Z*,*E*)-FPP as substrates, the most prominent being identified putatively as elemol (Fig. S3B). This product was, however, not detected in crude protein lysates (Fig. 2C). Recombinant *Hh*IDS2 protein did not make a terpene product when assayed with different FPP isomers. Instead, the *Hh*IDS2 enzyme converted the prenyl diphosphate precursors IPP and DMAPP to (*E*,*E*)-FPP but no other FPP isomer (Fig. 2D). Due to the TPS activities of *Hh*IDS7 and *Hh*IDS1 we designate these enzymes hereafter as *Hh*TPS1 and *Hh*TPS2, respectively. Furthermore, we designate *Hh*IDS2 as *Hh*FPPS because of its ability to function as an

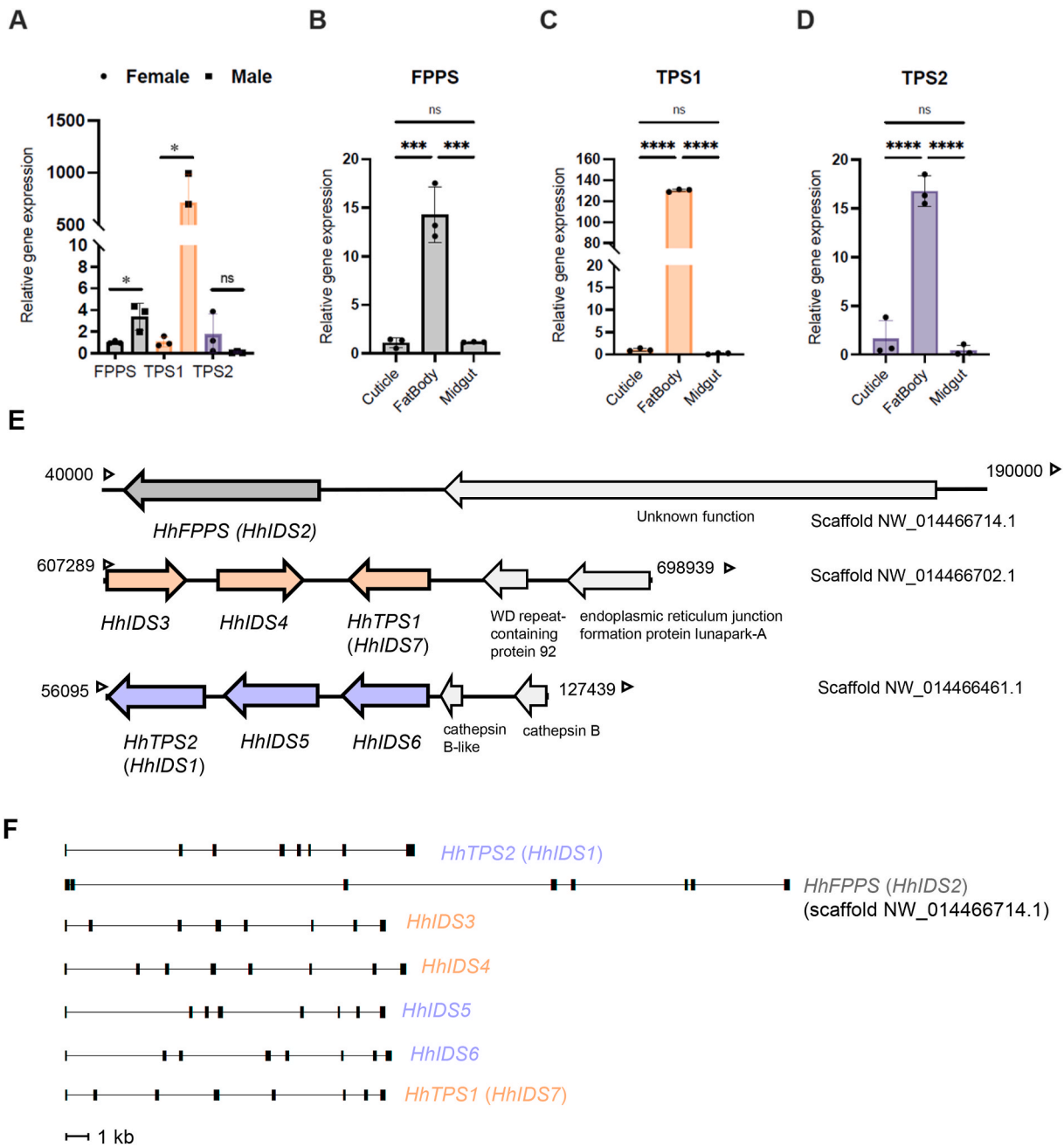

Fig. 3. Expression, genomic organization and structure of genes of the *H. halys* IDS/TPS family. (*A*) Transcript abundance of *HhFPPS, HhTPS1, HhTPS2* as determined by qRT-PCR in mature *H. halys* females and males (n = 3, ±SD) and (*B-D*) in different tissue types of mature *H. halys* males (n = 3, ±SD). Significance was determined using (*A*) student's t-test and (*B–C*) one-way analysis of variance (ANOVA) and means grouped by Tukey's HSD; *P < 0.05, ***P < 0.001, ****P < 0.0001. *H. halys* gene expression was normalized against the *RpS4* housekeeping gene and transcript abundance is shown relative to that in female or the cuticle tissue as determined by qRT-PCR. (*E*) Genomic organization of *H. halys* FPPS and IDS-type genes. Colored arrows indicate clustered gene positions on three different genome scaffolds. (*F*) Exon-intron structures of *H. halys* IDS-type genes. Exons are depicted by boxes and introns are depicted by lines.

IDS enzyme and lack of TPS activity. Cross-species sequence comparison showed an expectedly high degree of sequence similarity between *Hh*FPPS and canonical FPPSs of other pentatomids (*Mh*FPPS, *Nv*FPPS) (Fig. S1, Table S4). In addition, *Hh*TPS1 was found to be highly identical to the sesquipiperitol synthase from *Murgantia histrionica* (*Mh*TPS1), which indicates a high level of functional gene/protein conservation in the pheromone biosynthetic pathway of both species.

3.2. Sex and tissue specific expression of H. halys TPS and FPPS genes

To determine the sex-specificity of *HhTPS1, HhTPS2*, and *HhFPPS* expression, the transcript abundance of these genes was examined in whole mature males and females (Fig. 3A). *HhTPS1* was found to be expressed in males at high levels, whereas transcript abundances of *HhTPS2* and *HhFPPS* were comparatively low in both sexes (Fig. 3A, Fig. S7). Transcript levels of *HhTPS2* were found to be slightly elevated in females (Fig. S7) suggesting that this TPS could be involved in female specific terpene metabolism. However, no terpene synthase activity other than the formation of sesquipiperitol was detected in crude extracts of *H. halys* males and females, which provides little current evidence for an *in vivo* function of *HhTPS2*.

To measure the tissue-specific expression of *HhTPS1, HhTPS2*, and *HhFPPS* in males, mature males were dissected into the midgut, the fat body along with other abdominal soft tissues, and the abdominal cuticle including sternites with attached epithelial cells. Transcript abundance of *HhTPS1* was highest in the fat body compared to the mid gut and cuticle-associated tissue and similar expression profiles, although at lower levels, were observed for the other two genes (Fig. 3B–D). These tissue-specific expression patterns indicate a higher degree of terpene metabolic activity in the male fat body of *H. halys*.

3.3. Genomic organization, gene structures, and protein structural changes of the H. halys IDS/TPS genes

Localization of the TPS and IDS-like genes on the *H. halys* genome revealed the presence of two gene clusters with *HhIDS3*, *HhIDS4*, and *HhTPS1* forming one cluster and *HhTPS2*, *HhIDS5*, and *HhIDS6* representing the second cluster (Fig. 3E). Interestingly, the *HhFPPS* gene does not cluster with any other gene of the *H. halys* IDS-like gene family. All genes consist of eight exons and seven introns (Fig. 3F). However, several introns of the *HhFPPS* gene are expanded in size compared to those of the clustered *HhIDS/TPS* genes. Amino acid sequence identities of the proteins encoded by genes of the same cluster range between 60 and 89% suggesting their emergence by tandem gene duplications, whereas amino acid sequences between the two clusters share only 40% or lower identity (Table S4). *Hh*FPPS represents the most distant member of the family with only 20–23% amino acid sequence identity relative to other *Hh*IDS/TPS proteins.

We further determined the position of introns and corresponding intron phases in the genes of *H. halys* in comparison to those of FPP synthase and IDS-type TPS genes from other insects (Fig. S8). With exception of the first intron, all genes of the *H. halys* IDS-type gene family share the same position and intron phases of their six other introns. These intron positions and phases were also found to be conserved in FPPS genes of aphids (*Acyrthosiphon pisum*) and cockroaches (*Blattella germanica*), which supports the notion that Hemiptera and Blattodea share the same ancestral exon-intron structure of FPPS genes and this gene architecture has been conserved in the IDS-type TPS genes in *H. halys* and likely other pentatomids (Fig. S8). Further comparison of the *H. halys* genes with FPPS genes from representatives of Lepidoptera (*Bombyx mori*), Diptera (*Drosophila melanogaster*) and Coleoptera (*Dendroctonus ponderosae*, *Tribolium castaneum*, *Phyllotreta striolata*) showed an overlap in two to four intron positions (Fig. S8). Notably, the number of introns in the FPPS genes of *T. castaneum* and *P. striolata* has been reduced to three and their positions have been, similarly to those in *H. halys*, retained in the IDS-type TPS genes of *P. striolata*. The fact that the TPS genes of *H. halys* and *P. striolata* share the same number of introns and identical intron positions with their respective FPPS genes in the same species and other species of the same order supports an independent emergence of TPS genes from FPPS progenitors in these lineages.

To identify potential structural modifications of the *H. halys* TPS proteins in comparison to bonafide IDS enzymes, we compared the amino acid sequences from *Hh*TPS1, *Hh*TPS2 and the other *H. halys* IDS-type proteins with those from *Hh*FPPS and other stink bug FPPSs (Fig. S1). We identified several residue substitutions indicative for the neofunctionalization of IDS-type TPS enzymes such as the replacement of aromatic amino acid residues in position 4 and 5 upstream of the first aspartate rich motif with smaller, aliphatic residues (Fig. S1). These non-aromatic residues are likely critical for TPS activity by facilitating proper positioning of the FPP substrate and subsequent cyclization as was assumed from substitutions of these residues in the *M. histrionica* TPS protein (Lancaster et al., 2018). Moreover, all *H. halys* IDSs or TPSs exhibit substitutions to several putative IPP binding residues, which we predicted to be present in FPPS proteins based on crystallographic structures of IDS enzymes (Fig. S1, Fig. S9). Structural analysis of *Hh*FPPS and *Hh*TPS1 models suggests that substitutions at the IPP binding residue positions presumably prevent proper binding of the diphosphate moiety and positioning of the prenyl tail of IPP, thereby facilitating binding and conversion of a single allylic substrate (FPP) (Fig. S9). Mutational analysis will be required to corroborate changes of enzymatic activity resulting from these substitutions.

3.4. Identification of a sesquipiperitol synthase and genomic localization of IDS-type genes in N. viridula and Euschistus heros

Unexpectedly, BLAST searches of more recent publicly available genome (https://www.ncbi.nlm.nih.gov/assembly/GCA_928085145.1/#/st) and transcriptome resources of the southern green stink bug *N. viridula* identified an as-yet-uncharacterized IDS-like gene, which shares 84% amino acid sequence identity with *Hh*TPS1 (Table S4). This finding suggested a conserved sesquipiperitol synthase gene function in *N. viridula* different from *Nv*TPS1, which makes the precursor (*Z*)-α-bisabolene for the bisabolene epoxide pheromone in this species (Fig. 1, Fig. S1). While no transcript of this identified *N. viridula* gene was present in a previously obtained transcriptome of male cuticle associated tissue (Lancaster et al., 2019), we were able to clone a corresponding cDNA from male fat body tissue. TPS assays with the recombinant protein showed that it indeed converted (*E*,*E*)-FPP to sesquipiperitol as the main product (Fig. 4A). We therefore named this enzyme *Nv*TPS2. The *Nv*TPS2 protein had limited sesquipiperitol synthase activity with other FPP isomers but produced several sesquiterpene olefins with (*Z*,*Z*)-FPP (Figs. S10A and C). The recombinant enzyme also converted GPP into several monoterpenes including small amounts of piperitol, a C10 analog of sesquipiperitol (Fig. S10B).

We further determined the chromosome specific localization of all genes of the IDS-type family of *N. viridula*. Similar as with *H. halys*, we found the previously characterized *NvFPPS* gene (Lancaster et al., 2019) to be positioned on a separate chromosome independently of any IDS/TPS type gene clusters. The α-bisabolene synthase gene *NvTPS1* clusters with the uncharacterized *NvIDS3* gene (Lancaster et al., 2019) and another putative IDS-like gene (*NvIDS4*) on chromosome 6, whereas *NvTPS2* is located on chromosome 4 without any IDS-like genes in its proximity (Fig. 4B). While these findings indicate different degrees of gene duplication between the *N. viridula* and *H. halys* gene families they also suggest the presence of a sesquipiperitol synthase gene in a common ancestor of both species.

An additional search of the genome of the neotropical brown stink bug, *Euschistus heros*, a species which does not emit terpene pheromones (Weber et al., 2018), found tandemly arranged IDS-like homologs on two separate contigs. Despite incomplete sequence information of these genes, one set is more closely related to that of the *H. halys TPS1* cluster

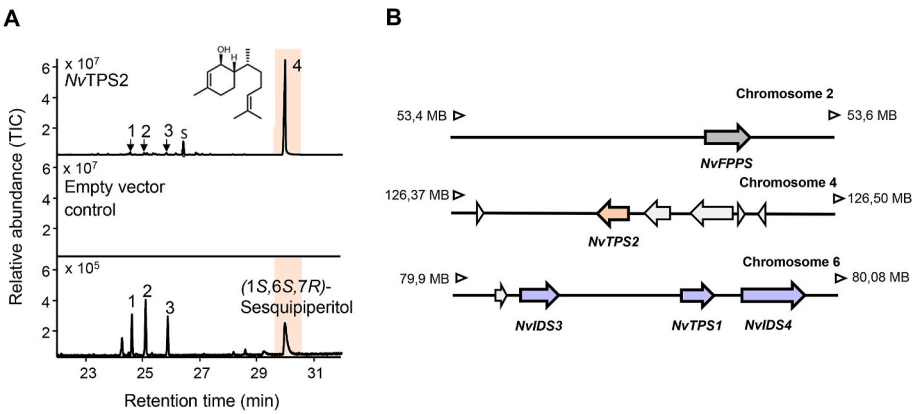

Fig. 4. Functional characterization of *Nv*TPS2 and genomic organization of the *N. viridula* IDS/TPS family. (*A*) GC-MS analysis of products from an assay with recombinant *Nv*TPS2 in the presence of (*E*,*E*)-FPP. *Nv*TPS2 converts (*E*,*E*)-FPP to sesquipiperitol (4) along with thermal rearrangement products (1–3). Lower panel: sesquipiperitol standard and thermal rearrangement products. (*B*) Chromosomal localization of *NvTPS2* and other members of the *N. viridula* IDS/TPS family. Genes in light grey are of unknown function.

(app. 70% amino acid sequence identity; *Euschistus heros* isolate 2730 contig00097, accession RCWM01000097.1), while the other shows higher sequence similarity (53–61% amino acid sequence identity) with genes of the *H. halys TPS2* cluster (*Euschistus heros* isolate 2730 contig00020, accession RCWM01000020.1) suggesting gene duplication events in the ancestor of both pentatomid species. These findings provide further evidence for the presence of IDS-like gene clusters in pentatomid genomes with perhaps additional functions independent of terpene pheromone biosynthesis.

3.5. Phylogenetic comparison of IDS-like proteins of the Pentatomidae and other species in the Heteroptera and the order Hemiptera

Identification of the *H. halys* and *N. viridula* IDS-type family allowed us to refine the evolutionary relationship of IDS-like proteins both within Pentatomidae and among related proteins of other insect species. A phylogenetic tree inferred from maximum-likelihood analysis (see "Phylogeny of FPPS and TPS proteins from Pentatomids" in Materials and Methods) indicates a paralogous division of two IDS-like clades with one clade containing the *H. halys* sesquipiperitol synthase (*Hh*TPS1) and its related *Hh*IDS3 and *Hh*IDS4 sequences as well as the *M. histrionica* and *N. viridula* sesquipiperitol synthase homologs, while the other clade contains the *N. viridula* α-bisabolene synthase (*Nv*TPS1), *Hh*TPS2 and its related *Hh*IDS5 and *Hh*IDS6 sequences (Fig. 5). The same phylogenetic relationship was inferred from a Bayesian analysis (Fig. S11) (see "Phylogeny of FPPS and TPS proteins from Pentatomids" in Materials and Methods). We refer to the two IDS-type clades from hereon as TPS-a and b clades, respectively. The progenitor of these IDS-type clades

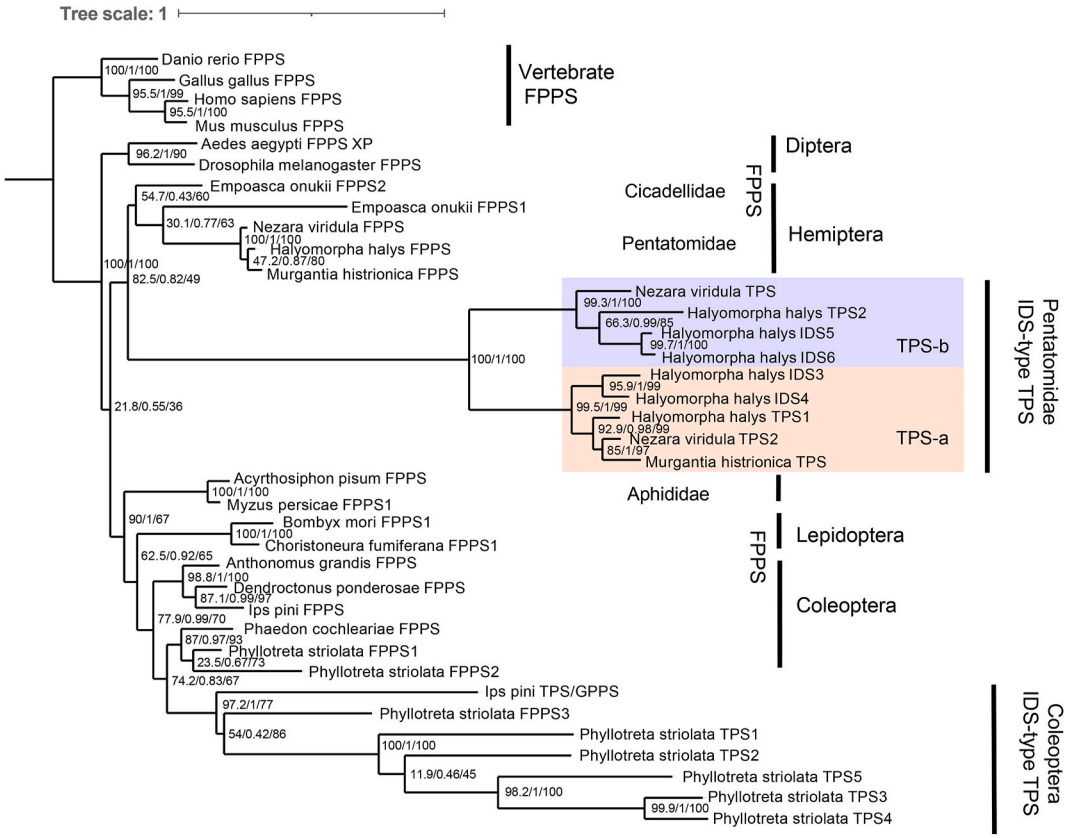

Fig. 5. Phylogram of functionally characterized insect FPPS and IDS-type TPS proteins inferred from maximum likelihood analysis (IQTree v2.0.4). Node values indicate SH-aLRT (x10,000), aBayes, and UFBoot (x10,000) support values, respectively. Branch lengths are scaled proportionately to the number of substitutions per site. The phylogeny was rooted to a clade of vertebrate FPPS proteins. Gene accession numbers are given in Fig. S11.

diverged from canonical FPP synthases in the pentatomids and other Hemiptera (Fig. 5, Fig. S11). A phylogenetic comparison of pentatomid TPSs with coleopteran FPPS and IDS/TPS proteins of *P. striolata* (Beran et al., 2016) further supports the finding that IDS-like genes with TPS function emerged independently from FPPS progenitors in these different lineages (Fig. 5, Fig. S11).

We further determined whether different selection regimes occurred in pentatomid TPS gene and hemipteran FPPS gene evolution by performing a selection analysis in PAML (Fig. S12). An extended phylogeny was generated including additional hemipteran putative FPPSs and IDS-like sequences from other pentatomid species, which were obtained from expanded gene mining as described below. Comparing different branch models (see Materials and Methods), a significant difference was detected in the gene-wide ratios of non-synonymous/synonymous substitutions (ω values), with relaxed evolutionary constraint on the TPS clade compared to the hemipteran FPPS background (Figs. S12A and B). We further employed branch-site models (see Materials and Methods) to test for site-specific positive selection underlying gene functionalization. Evidence for positive selection was detected on branches of initial TPS functional divergence from the FPPS clade (T0), and the originating branches of the TPS-a and TPS-b clades (T1 and T2) (Figs. S12A and C). As expected, we did not find significant evidence for positive selection for the functionally conserved FPPS clade (branch I0). Moreover, there was no significant evidence for positive selection acting on branches within the TPS-a and TPS-b clades (branches T3 and T4; Figs. S12A and C) despite a large proportion of codons being partitioned into relaxed site classes, suggesting that relaxed constraint rather than positive selection contributed to more recent evolutionary changes in TPS coding sequences (Fig. S12C). Overall, our results support the notion that pentatomid TPS genes originally diverged under positive selective pressure but underwent limited interspecific and intraspecific diversification following clade-specific divergence.

To gain a better understanding of the broader occurrence of terpenes in the Hemiptera, we performed a literature search of terpenes in different hemipteran sub- and infraorders (Table S1 and references therein). Among the Heteroptera and more specifically the pentatomids, sesquiterpene aggregation pheromones predominantly occur in several tribes of the Pentatominae. In addition, a number of species in the pentatomid subfamily Asopinae (predatory stink bugs) release monoterpene alcohol pheromones including piperitol, which is the C10 analog of sesquipiperitol. Monoterpene secretions with functions in attraction or defense have also been reported from several other families in the heteropteran infraorders Pentatomomorpha and Cimicomorpha such as the Acanthosomatidae (shield bugs), Cimicidae (bed bugs), Cycnidae (burrowing bugs), Miridae (plant bugs), Lygaeidae (seed bugs), Pyrrhocoridae (red bugs), Tingidae (lace bugs) and others (Table S1 and references therein). In the hemipteran suborder Sternorrhyncha, aphids are well-known for their emission of the sesquiterpene alarm pheromone β-farnesene (Pickett et al., 2013; Vandermoten et al., 2012). In the same suborder, cyclobutane type monoterpenes and sesquiterpenes and several hemiterpene-, monoterpene-, and norterpene esters occur as sex pheromones in scale insects while sesquiterpene acids are present in the hardened excretions of lac insects (infraorder Coccomorpha) (Table S1 and references therein). By contrast, we did not find any evidence of specialized terpene metabolites in cicadas, plant hoppers and moss bugs (suborders Auchenorrhyncha and Coleorrhyncha).

To investigate the evolution of IDS-like genes that may correlate with and give rise to terpenes in pentatomids and the various suborders of the Hemiptera, we mined hemipteran IDS-like genes from NCBI nucleotide and transcriptome shotgun assembly databases using tBLASTn searches and the *H. halys* IDS/TPS protein sequences as search queries. We focused our search on FPPS-like genes excluding GGPP synthase-like sequences. FPPS sequence homology was determined through the clustering with functionally characterized FPPS-type TPS proteins in neighbor joining phylogenies, the presence of aspartate-rich motifs characteristic of IDS enzymes, and sequence similarity to other insect FPPS-like proteins in tBLASTn searches of NCBI databases. Our search resulted in a collection of 299 unique proteins with a length of at least 300 amino acids. The mined sequences and a functionally validated coleopteran FPPS outgroup were used to infer a maximum likelihood phylogeny of hemipteran IDS (FPPS)-like proteins (see "Expanded Hemiptera IDS and TPS Phylogenies" in Materials and Methods). We found a substantial number of genes in all hemipteran suborders with putative FPPS function (Fig. 6; Fig. S13). However, gene mining also resulted in the identification of additional IDS-like genes in the Pentatomidae, which clustered closely with the genes of the TPS-a and TPS-b clades and thus may harbor TPS function (Fig. 6; Fig. S13). These include genes of the predatory spined soldier bug, *Podisus maculiventris*, which are closely related to the characterized sesquipiperitol synthases. Similar genes positioned in the type-a and –b clades were from the green stink bug *Chinavia impicticornis* and its relative *Chinavia* (formerly *Acrosternum*) *hilaris*. Mining of NCBI transcriptome shotgun assembly databases further found IDS-like genes with expression in several other species of the infraorder Pentatomomorpha including those with known monoterpene defense constituents such as the burrower bug *Sehirus cinctus*, the boxelder bug *Boisea trivittata*, and the firebug *Pyrrhocoris apterus* (Farine et al., 1992; Krall et al., 1997; Palazzo and Setzer, 2009). Notably, sequences from *S. cinctus* are more closely aligned with those of the pentatomid TPS-a clade (Fig. S13), which suggests that the divergence of the type-a and b clades probably occurred already in the superfamily of the Pentatomoidea. Moreover, IDS-like transcripts were present in the hemipteran infraorder Cimicomorpha, and in mealybugs such as the cotton mealybug, *Phenacoccus solenopsis*, and the lac insect *Kerria lacca* of the Sternorrhyncha, for which cyclobutane monoterpenes and cyclic sesquiterpene acids have been documented as sex pheromone and lac components, respectively (Kandasamy et al., 2021; Tabata and Ichiki, 2016). FPPS-like genes with possible TPS function were absent in aphids despite the presence of the alarm pheromone β-farnesene in these species. Also, these genes were not identified in the suborders Auchenorrhyncha and Coleorrhyncha in agreement with the lack of terpenes in these orders. Taken together, we found evidence for the evolution of IDS (FPPS)-like TPS genes in two main suborders of the Hemiptera, the Heteroptera (true bugs) and the Sternorrhyncha (Figs. 6 and 7, Fig. S13). Moreover, our phylogeny supports the emergence of hemipteran IDS-like genes from FPPS genes in this order. Although it could not be fully inferred from our analysis whether FPPSs and IDS-like genes diverged independently in the two suborders or in a progenitor of these lineages, a parsimony argument suggests this more likely occurred within a shared ancestral lineage (Figs. 6 and 7). Notably, in our large-scale phylogeny IDS-like sequences of the Heteroptera and Sternorrhyncha are more closely positioned to FPPS sequences of the Sternorrhyncha (Fig. 6A and B), which supports the notion of an emergence of IDS-type TPS genes at an early stage of hemipteran evolution.

4. Discussion

Terpene pheromone and defense secretions are common among insects but their biosynthetic origin and evolution are not well understood. Our study of pheromone biosynthesis in stink bugs has provided further evidence that terpene pheromone precursors can be made by IDS-type TPS enzymes. Stink bug TPS families emerged in paralogous clades at the onset of pentatomid evolution. Homologs in these clades show functional conservation and a limited degree of divergence in contrast to those in plant or microbial TPS families and in accordance with a smaller number of terpene isomers used in chemical communication. We further found that TPS genes in Pentatomidae and probably in several lineages of terpene releasing hemipteran insects have evolved from IDS genes with FPPS function by ancient gene duplication. This process seems to have occurred independently in hemipteran and coleopteran insects and probably in other insect orders.

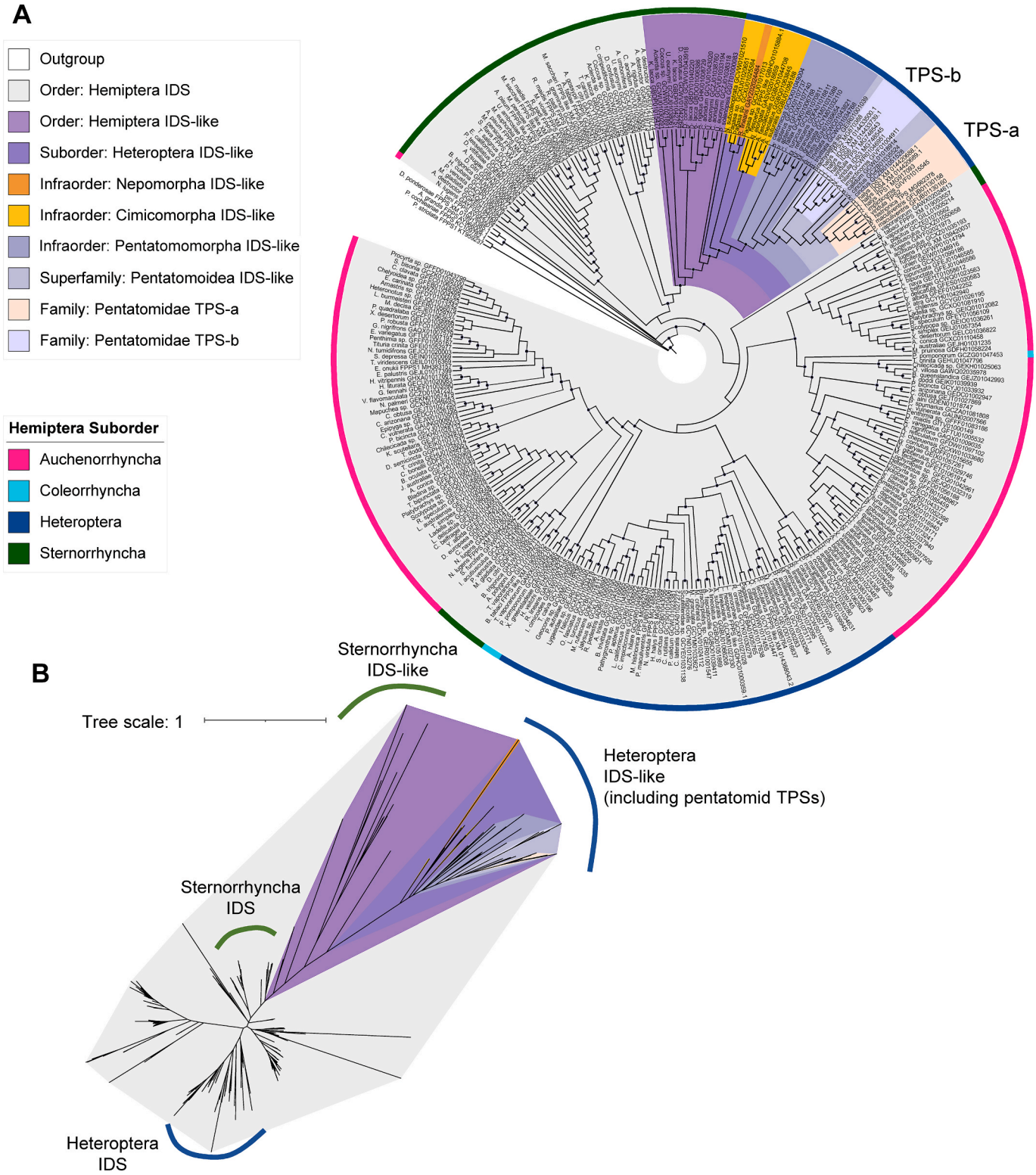

Fig. 6. Phylogeny of characterized and putative IDS (FPPS) and IDS-like proteins in the Hemiptera. (*A*) A cladogram is depicted inferred from maximum likelihood analysis. Well-supported nodes (>80% SH-aLRT (x1,000), >80% UFBoot (x1,000), and >0.80 aBayes support) are marked with black circles. The phylogeny was rooted to a Coleopteran FPPS outgroup. The same phylogeny displayed as phylogram with individual support values is depicted in Fig. S13. Species names and sequences are listed in clockwise order in Table S5. (*B*) Phylogeny in (*A*) shown as an unrooted phylogram. Branch lengths represent the number of amino acid substitutions per site. Heteropteran IDS-like proteins belong to a monophyletic lineage together with sternorrhynchan IDS and IDS-like proteins suggesting a common origin. Heteropteran IDSs form a more distantly related lineage of proteins.

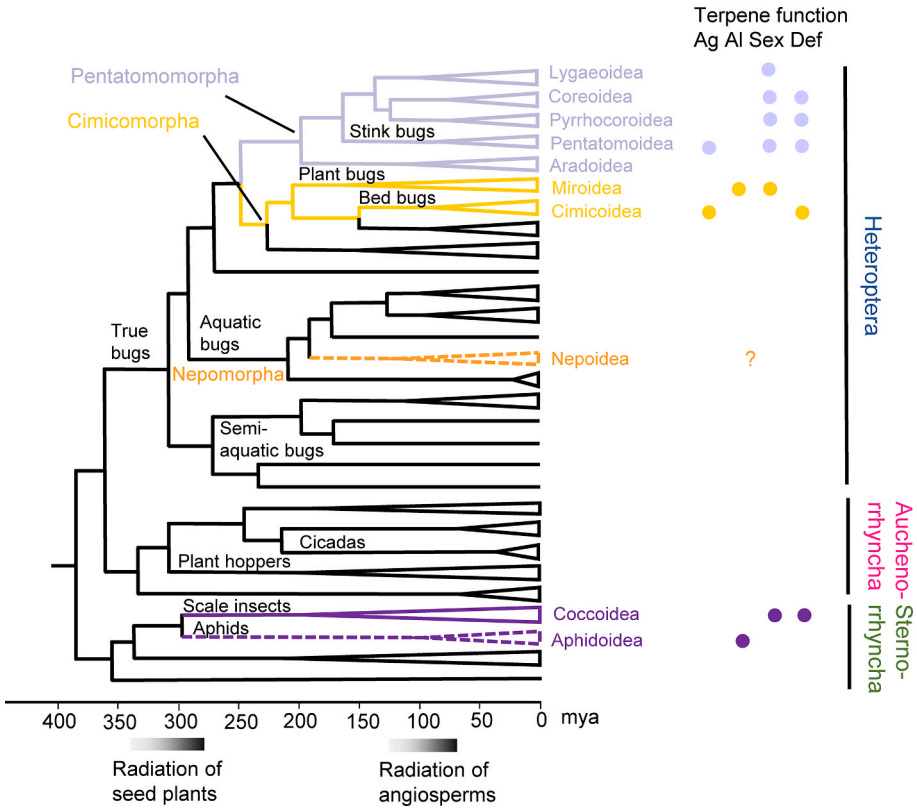

Fig. 7. Occurrence and possible time of evolution of terpene semiochemicals and IDS-like genes in the Hemiptera. Solid colored branches indicate the presence of IDS-like genes in different superfamilies according to the phylogeny in Fig. 6. A single IDS-like gene was found in a species in the Nepoidea (orange, dashed branch), but no terpenes have been described in this group. No IDS-like genes were identified in aphids (purple, dashed branch) despite the presence of sesquiterpene alarm pheromones. Ag, aggregation pheromone; Al, alarm pheromone; Sex, sex pheromone; Def, defensive or protective function. The tree is redrawn from Johnson et al. (2018) with modifications.

4.1. Stink bug TPS genes evolved from FPPS progenitors in two clades with preserved gene function

Our analysis of the IDS-like gene family in *H. halys* identified six genes positioned in two clusters, which likely emerged by gene duplications. Each cluster was found to contain a functionally active TPS gene, while the function of the other genes remains uncertain. The canonical *H. halys* FPPS gene resides in a genomic position independent of the clustered genes. Its low sequence similarity with the IDS-like genes and extended length of introns suggest this gene to be derived from a more ancient duplication event. Phylogenetic comparison with IDS-type genes of other stink bugs and related species provided evidence that each cluster of *H. halys* IDS-type genes belongs to a paralogous clade named TPS-a and TPS-b, respectively. These clades most likely evolved from a common ancestor of the Pentatomidae or possibly the Pentatomoidea approximately more than 100 mya given the presence of putative IDS-like genes with close similarity to type-a TPS homologs in species of burrower bugs such as *Sehirus cinctus* (Fig. 6, Fig. S13). We found additional evidence for tandemly-arranged or clustered IDS-type genes of the TPS-a or b types in the genomes of *N. viridula* and *Euschistus heros* (Fig. 8A), even though no terpene pheromones are emitted by the latter species. The extent of gene cluster formation within the TPS type-a and -b clades differs between *H. halys* and the other species suggesting varying degrees of gene duplication or loss prior to or during speciation (Fig. 8A).

Functional characterization and phylogenetic comparison of the pentatomid TPS genes investigated in this work and in our previous studies show that TPS gene function has been conserved in each of the TPS clades. The TPS-b clade contains the (*Z*)-α-bisabolene synthase of *N. viridula* (*Nv*TPS1), which is believed to make the precursor of the bisabolene epoxide pheromone in this species (Lancaster et al., 2019). Other genes positioned in this clade are from *N. viridula* relatives in the genus *Chinavia* (or the separate genus *Acrosternum*), which are known as green stink bugs (Fig. 1). Since *C. impicticornis* and its relatives share

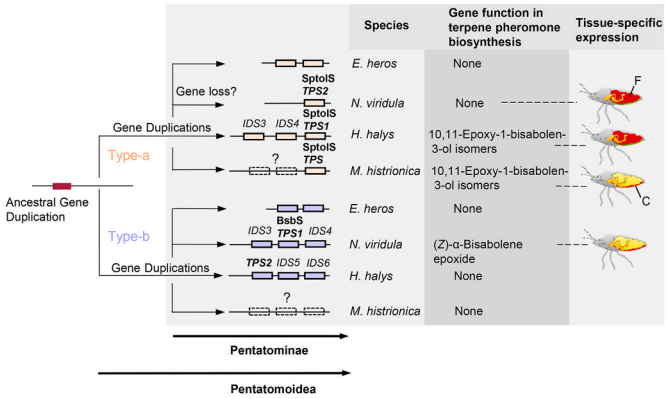

Fig. 8. TPS gene evolution and functionalization in pentatomids. Model of TPS type-a and type-b gene emergence in genomes of pentatomids, their function in pheromone biosynthesis and tissue-specific localization. Genes with identified function are in bold. Gene clusters in *M. histrionica* are unknown in the absence of genomic information. Colored boxes indicate different TPS subfamilies. SptolS, sesquipiperitol synthase; BsbS, (*Z*)-α-bisabolene synthase; C, cuticle-associated tissue; F, fat body.

(*Z*)-α-bisabolene type pheromones with *N. viridula* (Blassioli-Moraes et al., 2012) (Fig. 1, Table S1), it can be assumed that their IDS-like genes have a similar function in making (*Z*)-α-bisabolene. Within the TPS-a clade the sesquipiperitol synthase gene function has been conserved in at least three species (*M. histrionica*, *H. halys* and *N. viridula*) independent of their subfamily status and geographical origin. The presence of sesquipiperitol, zingiberenol and 10,11-epoxy-1-bisabolen-3-ol as pheromone constituents of several other neotropical stink bugs (Fig. 1) further suggests that the same sesquiterpene synthase function has been preserved in these species. Given the absence of murgantiol-like isomers in *N. viridula*, the identification of a

functional sesquipiperitol synthase (*Nv*TPS2) was unexpected. Previous analysis did not reveal any sesquipiperitol synthase activity in crude protein extracts of *N. viridula* males (Lancaster et al., 2019). Organic solvent extraction of adult males or females also did not provide evidence for detectable amounts of sesquipiperitol or possible terpene alcohol ester derivatives. Therefore, the *in vivo* function of *Nv*TPS2 remains uncertain at this time. Overall, the presence of sesquipiperitol synthases in the genomes of several pentatomid species supports a conserved biochemical function of this gene in a common progenitor in association with chemical attraction and/or perhaps a different role as may be suspected in *N. viridula*. Semiochemicals released by insects can have more than one function and may serve as pheromones and defensive compounds. Examples include defense secretions that appear to have evolved secondary roles as alarm, sex or aggregation pheromones (e.g. Sobotnik et al., 2008; Weiss et al., 2013). Therefore, it has been hypothesized that pheromonal functions may originate from defensive activities (Stökl and Steiger, 2017). Perhaps, sesquipiperitol and its derived compounds have evolved in similar ways.

Our phylogenetic analysis suggests that TPS-a type genes of other pentatomids such as the predatory stink bug, *P. maculiventris* may encode enzymes with similar activity as sesquipiperitol synthase (Fig. 6, Fig. S13). Males of *P. maculiventris* emit a monoterpene alcohol pheromone blend including small amounts of piperitol (Aldrich et al., 1986), which makes it likely that the identified genes encode enzymes that make monoterpene alcohols from the substrate GPP equivalent to the formation of sesquipiperitol from FPP. It should be noted that information on the evolutionary history of pentatomids is still limited because of the lack of DNA-supported phylogenies (Genevcius et al., 2021). For instance, predatory stink bugs such as *P. maculiventris* and herbivorous stink bugs are currently positioned in different subfamilies of the Asopinae and Pentatominae (Grazia et al., 2008). However, a recent DNA-based phylogeny by Genevcius et al. (2021) positions the Asopinae within the Pentatominae, which is consistent with our finding of a close relationship of terpene alcohol synthases in the TPS-a clade of the representative species of these groups.

Further analysis of the pentatomid sesquipiperitol synthases showed that the genes exhibit differences in tissue-specific expression with *HhTPS2* and *NvTPS2* being expressed in the fat body and transcript levels of *MhTPS* being highest in the tissue associated with the abdominal cuticle (Lancaster et al., 2018) (Fig. 8). Neither males of *H. halys* nor *M. histrionica* possess pheromone glands, in contrast to male *N. viridula*, which emit the bisabolene epoxide pheromone from glandular cells at the ventral abdomen (Cribb et al., 2006). The difference in tissue-specific gene expression suggests an evolution of different modes of spatial organization of terpene pheromone biosynthesis in these pentatomid species. While the formation of sesquipiperitol in the cuticle-associated tissue of *M. histrionica* is most likely coupled with its conversion to the pheromone and subsequent release from the same tissue, sesquipiperitol produced in the fat body of *H. halys* likely needs to be transported through the hemolymph to cuticle-associated cells, where the biosynthetic pathway is completed. A similar transport mediated by lipophorin has been described for hydrocarbon sex pheromones, e.g. in the spongy moth *Lymantria dispar* (Lepidoptera: Erebidae) (Jurenka et al., 2003).

4.2. Pentatomid IDS/TPS gene families show limited expansion and diversification compared to plants

Selection analysis of pentatomid IDS-type TPSs and hemipteran FPPSs indicated that the TPS gene clades are under relaxed evolutionary constraint in comparison to the FPPS clade and that periods of positive selection may have followed past gene duplication events, which is in accordance with the expansion of gene families resulting in novel functions (Lespinet et al., 2002) (Fig. S12). Functional expansion and diversification of gene families have been shown to play important roles in pheromone evolution as has been demonstrated, for example, for gene families of fatty acyl reductases in the Hymenoptera (Tupec et al., 2019). Nevertheless, it appears that, after an initial expansion, the genes within both pentatomid TPS clades have undergone limited interspecific and intraspecific diversification, potentially due to more recent relaxed purifying selection and/or neutral evolution, suggesting that small-size gene families are sufficient to generate precursors of terpene pheromones in pentatomid evolution. Species-dependent functionalization occurred among this small set of genes in the type-a and type-b clades resulting in a conserved formation of sesquiterpene alcohol or olefin intermediates such as sesquipiperitol and (*Z*)-α-bisabolene. These intermediates require only a few steps of isomerization and/or epoxidation to be converted to the final pheromone products. Thus, limited functional diversification of TPS genes in conjunction with a small number of derivatization steps has been successful in generating species-specific pheromone isomeric blends. Species specificity is further accomplished by the addition of other molecules to the pheromone mixture, such as fatty acid derivatives (Weber et al., 2018). There may be other constraints that play a role in a limited diversifying selection of TPS enzymes in pentatomids and other insects and animals in general. These may include adverse physiological effects of additional terpene products, energetic constraints of an expanded terpene metabolism, and possible competition with primary terpene metabolic pathways. It is also possible that more complex terpene blends may affect proper perception of the pheromone signal due to overlap with compounds emitted by plants or other organisms. Interestingly, Lassance et al. (2021) reported that the evolution of the signature fatty acid-derived pheromone in olethreutine moths is based on the function of a single fatty acid desaturase (FAD) gene, which emerged by an ancient gene duplication. Members of other FAD subfamilies do not appear to play a role in pheromone biosynthesis, which indicates a distinct selection and preservation of FAD genes involved in moth chemical communication similar to individual TPS genes involved in terpene pheromone biosynthesis in stink bugs. A similar scenario has been observed in the small IDS-type TPS gene family of the flea beetle *Phyllotreta striolata*, where only a single male-expressed TPS appears to be involved in pheromone biosynthesis (Beran et al., 2016).

The selection of individual genes in pheromone biosynthesis is supported by a reduced expression of the other members of the gene family. In *H. halys*, four out of the six IDS-type genes are not or only lowly expressed. Similarly, with the exception of *PsTPS1*, all other members of the TPS gene family in *P. striolata* are expressed at low levels even though three of these genes still encode functionally active enzymes (Beran et al., 2016). It is possible that some of the non-transcribed *H. halys* genes are also still functionally active as can be assumed from their coding sequence and residue substitutions characteristic for IDS-type TPS proteins (Fig. S9). A loss of expression of functionally active genes has also been reported for other pheromone biosynthetic genes such as for desaturases involved in the formation of cuticular hydrocarbon sex pheromones in *Drosophila* (Shirangi et al., 2009). Therefore, differential expression of gene family members likely underlies the evolutionary selection and specification of pheromone composition.

The restricted level of TPS gene evolution in insects stands in contrast to the species- and family-specific diversification of TPS genes in the form of large gene families in plants (Chen et al., 2011; Kumar et al., 2016; Li et al., 2012; Muchlinski et al., 2019). As in pentatomids, related plant species may share TPS orthologs of identical function albeit with a higher degree of sequence divergence. TPS gene families in higher and lower land plants often comprise genes of several TPS subfamilies (Chen et al., 2011). These more complex families seemingly have evolved over a longer time scale than those in insects given the emergence of the first land plants more than 500 mya; nevertheless, when compared at the same time scale, TPS genes in pentatomids never diversified to the same extent as those in flowering plants, which appeared at approximately the same time as pentatomids more than 100 mya (Fig. 7). The overall higher diversity of plant TPS genes is associated with the production of

more complex mixtures of terpenes (Degenhardt et al., 2009) that are believed to be adaptations to multifaceted inter-specific interactions in attraction or defense, in contrast to a smaller number of metabolites required in intra-specific communication. Therefore, the degree of TPS gene family expansion in insects and plants appears to be a consequence of specificity, mode of chemical interactions, and perhaps other constraints.

4.3. IDS-like gene evolution correlates with the occurrence of terpene pheromones and defense compounds in several hemiptera

Our search for terpene semiochemicals in the Hemiptera revealed several other species with the ability to release monoterpenes or sesquiterpenes as pheromones or defense compounds, especially in groups of the suborders Heteroptera and Sternorrhyncha (Table S1). Among these, sesquiterpenes are common constituents of pheromones in pentatomids but also occur in aphids and the excretions of lac insects (Morgan, 2010; Pickett et al., 2013). Monoterpenes are represented in the heteropteran infraorders Pentatomomorpha and Cimicomorpha, and irregular cyclobutane type monoterpenes and sesquiterpenes occur as sex pheromones of scale insects and mealybugs (Coccomorpha) (Table S1). Cyclobutane monoterpenes such as lavendulol are also known to be made in plants from DMAPP by action of TPS enzymes that catalyze an irregular coupling between isoprenoid units (Rivera et al., 2001). Whether this type of reaction is catalyzed by IDS-like enzymes encoded by transcripts identified in scale insects is currently unknown. With the exception of aphids, we found IDS-like genes in species of all other terpene releasing groups we identified (Figs. 6 and 7, Fig. S13), which suggests a possible role of these genes in *de novo* terpene biosynthesis. By contrast, there is currently no evidence for IDS-like genes in the suborders Auchenorrhyncha, Sternorrhyncha, and Coleorrhyncha within the of Hemiptera, in agreement with the lack of terpenes in these groups (Fig. 7). It should be noted though that our gene mining was focused on FPPS-like genes and that additional GGPP synthase-like genes can be found in transcriptomes of several hemipteran species, including aphids such as *A. pisum* and lac insects. A recent finding of a GGPP synthase-like protein with monoterpene synthase activity in the butterfly *Heliconius melpomene* (Darragh et al., 2021) suggests that a role of this family of IDS-like enzymes in the formation of specialized terpenes in insects cannot be entirely excluded. Furthermore, one should consider that some of the terpenes associated with several hemipteran species may be sequestered from host plants or derived from endosymbionts without or in addition to their possible biosynthesis by insect-specific pathways as has been found in bark beetles such as *Ips pini* (Chiu et al., 2018). Functional characterization of IDS-like genes in these species will give more insight into the extent of terpene biosynthetic evolution in hemipteran insects.

Our phylogenetic analysis of IDS-like genes in the Hemiptera supports an emergence of these genes from FPP synthases (Fig. 6, Fig. S13). IDS-type TPSs may have been selected for defensive and pheromonal functions and potentially facilitated a transition between inter- and intra-specific interactions. It is possible that the divergence of IDS-like genes occurred already in a progenitor of the Heteroptera and Sternorrhyncha more than 350 mya (Fig. 7) and that these genes were lost or became pseudogenes in the other hemipteran suborders because of the selection of other compounds or the supersedure by other modes of intra-specific communication such as vibration signals for mate localization (Cocroft and Rodriguez, 2005). Alternatively, IDS-like genes may have emerged independently in the distinct terpene emitting/secreting lineages of the hemipteran suborders, although this seems less likely in terms of parsimony and the observed phylogenetic relationships of IDS and IDS-like sequences in these groups (Fig. 6). However, at the scale of different insect orders, a phylogenetic comparison of functional IDS and TPS proteins from pentatomids with those from coleopteran species such as *P. striolata* strongly suggests an independent evolution of TPSs from FPPS progenitors in these lineages. Together, our study demonstrates that *de novo* biosynthesis of terpenes is common in pentatomids and might occur more broadly in the Hemiptera and other insect lineages.

Availability of data and materials

The data underlying this article are available in the article, its online material, and in publicly available repositories with URLs provided in the article. All gene sequences used for phylogenetic analysis are accessible through the GenBank database at NCBI (https://www.ncbi.nlm.nih.gov/genbank/) with accession numbers provided in the article. The sequence of the *Nv*TPS2 cDNA amplified in this work is accessible through accession number ON934605.

Author contributions

ZR participated in the conceptualization of the study, performed all phylogenetic analyses and gene selection analysis, interpreted the data, wrote parts of the manuscript, and edited the manuscript. JL participated in the conceptualization of the study, performed cloning, expression, and functional characterization of *H. halys* TPS genes, wrote parts of the manuscript, and edited the manuscript. HL performed cloning, expression, and functional characterization of *N. viridula* TPS2, qRT-PCR analysis of *H. halys* genes, wrote parts of the manuscript, and edited the manuscript. AK determined the stereospecificity of the *Hh*TPS1 product, wrote parts of the manuscript, and edited the manuscript. KL performed expression and functional characterization of *H. halys* FPPS. MES performed sequence analysis and edited the manuscript. KG supported the gene selection analysis and edited the manuscript. LS supported homology modeling and structural analysis. TGK performed LC-MS/MS analysis of HhFPPS enzyme products and edited the manuscript. DCW provided *H. halys* for protein extraction, gene cloning and qRT-PCR and edited the manuscript. DEG-R contributed to discussions on the original development of the research. PEO performed homology modeling and structural analysis and edited the manuscript. AVM contributed to research discussions and edited the manuscript. DT conceived the study, supervised the project, interpreted the data, designed figures, and wrote the manuscript. All authors read and approved the final manuscript.

Declaration of competing interest

The authors declare that they have no conflict of interest.

Data availability

Data availability is shared in manuscript

Acknowledgements

This work was supported by the National Science Foundation (MCB1920914, MCB1920922, MCB1920925) to AVM, P.O., and D.T., respectively, and the US Department of Agriculture National Institute of Food and Agriculture (2016-67013-24759) to DT, AK, DEG-R, and DCW. Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture. The USDA is an equal opportunity provider and employer.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ibmb.2022.103879.

References

Aldrich, J.R., Lusby, W.R., Kochansky, J.P., 1986. Identification of a new predaceous stink bug pheromone and its attractiveness to the eastern yellowjacket. Experientia 42, 583–585. https://doi.org/10.1007/bf01946714.

Aldrich, J.R., Oliver, J.E., Lusby, W.R., Kochansky, J.P., Lockwood, J.A., 1987. Pheromone strains of the cosmopolitan pest, *Nezara viridula* (Heteroptera, Pentatomidae). J. Exp. Zool. 244, 171–175.

Anisimova, M., Gil, M., Dufayard, J.-F., Dessimoz, C., Gascuel, O., 2011. Survey of branch support methods demonstrates accuracy, power, and robustness of fast likelihood-based approximation schemes. Syst. Biol. 60, 685–699. https://doi.org/10.1093/sysbio/syr041.

Anisimova, M., Yang, Z., 2007. Multiple hypothesis testing to detect lineages under positive selection that affects only a few sites. Mol. Biol. Evol. 24, 1219–1228. https://doi.org/10.1093/molbev/msm042.

Ballal, A., Laurendon, C., Salmon, M., Vardakou, M., Cheema, J., Defernez, M., O'Maille, P.E., Morozov, A.V., 2020. Sparse epistatic patterns in the evolution of terpene synthases. Mol. Biol. Evol. 37, 1907–1924. https://doi.org/10.1093/molbev/msaa052.

Beran, F., Rahfeld, P., Luck, K., Nagel, R., Vogel, H., Wielsch, N., Irmisch, S., Ramasamy, S., Gershenzon, J., Heckel, D.G., Köllner, T.G., 2016. Novel family of terpene synthases evolved from trans-isoprenyl diphosphate synthases in a flea beetle. Proc. Natl. Acad. Sci. USA 113, 2922–2927. https://doi.org/10.1073/pnas.1523468113.

Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., Bourne, P.E., 2000. The protein Data Bank. Nucleic Acids Res. 28, 235–242. https://doi.org/10.1093/nar/28.1.235.

Bertoni, M., Kiefer, F., Biasini, M., Bordoli, L., Schwede, T., 2017. Modeling protein quaternary structure of homo- and hetero-oligomers beyond binary interactions by homology. Sci. Rep. 7 https://doi.org/10.1038/s41598-017-09654-8.

Bienert, S., Waterhouse, A., de Beer, T.A., Tauriello, G., Studer, G., Bordoli, L., Schwede, T., 2017. The SWISS-MODEL Repository-new features and functionality. Nucleic Acids Res. 45, D313–d319. https://doi.org/10.1093/nar/gkw1132.

Blassioli-Moraes, M.C., Laumann, R.A., Oliveira, M.W.M., Woodcock, C.M., Mayon, P., Hooper, A., Pickett, J.A., Birkett, M.A., Borges, M., 2012. Sex pheromone communication in two sympatric neotropical stink bug species *Chinavia ubica* and *Chinavia impicticornis*. J. Chem. Ecol. 38, 836–845.

Blomquist, G.J., Ginzel, M.D., 2021. Chemical ecology, biochemistry, and molecular biology of insect hydrocarbons. In: Douglas, A.E. (Ed.), Annu Rev Entomol, 66, pp. 45–60.

Borges, M., Birkett, M., Aldrich, J.R., Oliver, J.E., Chiba, M., Murata, Y., Laumann, R.A., Barrigossi, J.A., Pickett, J.A., Moraes, M.C.B., 2006. Sex attractant pheromone from the rice stalk stink bug, *Tibraca limbativentris* Stal. J. Chem. Ecol. 32, 2749–2761.

Castillo, C., Chen, H., Graves, C., Maisonnasse, A., Le Conte, Y., Plettner, E., 2012. Biosynthesis of ethyl oleate, a primer pheromone, in the honey bee (*Apis mellifera* L.). Insect Biochem. Mol. Biol. 42, 404–416. https://doi.org/10.1016/j.ibmb.2012.02.002.

Castresana, J., 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol. Biol. Evol. 17, 540–552. https://doi.org/10.1093/oxfordjournals.molbev.a026334.

Chen, F., Tholl, D., Bohlmann, 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. https://doi.org/10.1111/j.1365-313X.2011.04520.x.

Chiu, C.C., Keeling, C.I., Bohlmann, J., 2018. Monoterpenyl esters in juvenile mountain pine beetle and sex-specific release of the aggregation pheromone trans-verbenol. Proc. Natl. Acad. Sci. U.S.A. 115, 3652–3657. https://doi.org/10.1073/pnas.1722380115.

Clementel, D., Del Conte, A., Monzon, A.M., Camagni, Giorgia F., Minervini, G., Piovesan, D., Tosatto, S.C.E., 2022. Ring 3.0: fast generation of probabilistic residue interaction networks from structural ensembles. Nucleic Acids Res. https://doi.org/10.1093/nar/gkac365.

Cocroft, R.B., Rodriguez, R.L., 2005. The behavioral ecology of insect vibrational communication. Bioscience 55, 323–334. https://doi.org/10.1641/0006-3568(2005)055[0323:Tbeoiv]2.0.Co;2.

Čokl, A., Dias, A.M., Moraes, M.C.B., Borges, M., Laumann, R.A., 2017. Rivalry between stink bug females in a vibrational communication network. J. Insect Behav. 30, 741–758. https://doi.org/10.1007/s10905-017-9651-z.

Cribb, B.W., Siriwardana, K.N., Walter, G.H., 2006. Unicellular pheromone glands of the pentatomid bug *Nezara viridula* (Heteroptera: insecta): ultrastructure, classification, and proposed function. J. Morphol. 267, 831–840. https://doi.org/10.1002/jmor.10442.

Darragh, K., Orteu, A., Black, D., Byers, K.J.R.P., Szczerbowski, D., Warren, I.A., Rastas, P., Pinharanda, A., Davey, J.W., Fernanda Garza, S., Abondano Almeida, D., Merrill, R.M., McMillan, W.O., Schulz, S., Jiggins, C.D., 2021. A novel terpene synthase controls differences in anti-aphrodisiac pheromone production between closely related *Heliconius* butterflies. PLoS Biol. 19, e3001022 https://doi.org/10.1371/journal.pbio.3001022.

de Oliveira, M.W.M., Borges, M., Andrade, C.K.Z., Laumann, R.A., Barrigossi, J.A.F., Blassioli-Moraes, M.C., 2013. Zingiberenol, (1*S*,4*R*,1'*S*)-4-(1',5'-Dimethylhex-4'-enyl)-1-methylcyclohex-2-en-1-ol, identified as the sex pheromone produced by males of the rice stink bug *Oebalus poecilus* (Heteroptera: Pentatomidae). J. Agric. Food Chem. 61, 7777–7785. https://doi.org/10.1021/jf402765b.

Degenhardt, J., Köllner, T.G., Gershenzon, J., 2009. Monoterpene and sesquiterpene synthases and the origin of terpene skeletal diversity in plants. Phytochemistry 70, 1621–1637. https://doi.org/10.1016/j.phytochem.2009.07.030.

Diep Thi, H., Chernomor, O., von Haeseler, A., Minh, B.Q., Le Sy, V., 2018. UFBoot2: improving the ultrafast bootstrap approximation. Mol. Biol. Evol. 35, 518–522. https://doi.org/10.1093/molbev/msx281.

Fang, S., Ting, C.T., Lee, C.R., Chu, K.H., Wang, C.C., Tsaur, S.C., 2009. Molecular evolution and functional diversification of fatty acid desaturases after recurrent gene duplication in *Drosophila*. Mol. Biol. Evol. 26, 1447–1456. https://doi.org/10.1093/molbev/msp057.

Farine, J.P., Bonnard, O., Brossut, R., Le Quere, J.L., 1992. Chemistry of defensive secretions in nymphs and adults of fire bug, *Pyrrhocoris apterus* L. (Heteroptera, Pyrrhocoridae). J. Chem. Ecol. 18, 1673–1682. https://doi.org/10.1007/bf02751094.

Fischer, M.J.C., Rustenhloz, C., Leh-Louis, V., Perriere, G., 2015. Molecular and functional evolution of the fungal diterpene synthase genes. BMC Microbiol. 15 https://doi.org/10.1186/s12866-015-0564-8.

Frick, S., Nagel, R., Schmidt, A., Bodemann, R.R., Rahfeld, P., Pauls, G., Brandt, W., Gershenzon, J., Boland, W., Burse, A., 2013. Metal ions control product specificity of isoprenyl diphosphate synthases in the insect terpenoid pathway. Proc. Natl. Acad. Sci. USA 110, 4194–4199. https://doi.org/10.1073/pnas.1221489110.

Gao, F., Chen, C., Arab, D.A., Du, Z., He, Y., Ho, S.Y.W., 2019. EasyCodeML: a visual tool for analysis of selection using CodeML. Ecol. Evol. 9, 3891–3898. https://doi.org/10.1002/ece3.5015.

Genevcius, B.C., Greve, C., Koehler, S., Simmons, R.B., Rider, D.A., Grazia, J., Schwertner, C.F., 2021. Phylogeny of the stink bug tribe Chlorocorini (Heteroptera, Pentatomidae) based on DNA and morphological data: the evolution of key phenotypic traits. Syst. Entomol. 46, 327–338. https://doi.org/10.1111/syen.12464.

Gilg, A.B., Bearfield, J.C., Tittiger, C., Welch, W.H., Blomquist, G.J., 2005. Isolation and functional expression of an animal geranyl diphosphate synthase and its role in bark beetle pheromone biosynthesis. Proc. Natl. Acad. Sci. USA 102, 9760–9765. https://doi.org/10.1073/pnas.0503277102.

Gilg, A.B., Tittiger, C., Blomquist, G.J., 2009. Unique animal prenyltransferase with monoterpene synthase activity. Naturwissenschaften 96, 731–735. https://doi.org/10.1007/s00114-009-0521-1.

Grazia, J., Schuh, R.T., Wheeler, W.C., 2008. Phylogenetic relationships of family groups in Pentatomoidea based on morphology and DNA sequences (Insecta: Heteroptera). Cladistics 24, 932–976. https://doi.org/10.1111/j.1096-0031.2008.00224.x.

Guex, N., Peitsch, M.C., Schwede, T., 2009. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: a historical perspective. Electrophoresis 30 (Suppl. 1), S162–S173. https://doi.org/10.1002/elps.200900140.

Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., Gascuel, O., 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. 59, 307–321. https://doi.org/10.1093/sysbio/syq010.

Harris, V.E., Todd, J.W., 1980. Male-mediated aggregation of male, female and 5th-instar southern green stink bugs, *Nezara-viridula* (l) (Hemiptera, Pentatomidae) and concomitant attraction of a tachinid parasite, *Trichopoda pennipes* (Diptera, Tachinidae). Entomol. Exp. Appl. 27, 117–126. https://doi.org/10.1111/j.1570-7458.1980.tb02955.x.

Havlíčková, J., Dolejšová, K., Tichý, M., Vrkoslav, V., Kalinová, B., Kyjaková, P., Hanus, R., 2019. (3*R*,6*E*)-nerolidol, a fertility-related volatile secreted by the queens of higher termites (Termitidae: syntermitinae). Z. Naturforsch. C Biosci. 74, 251–264. https://doi.org/10.1515/znc-2018-0197.

Huelsenbeck, J.P., Ronquist, F., 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17, 754–755. https://doi.org/10.1093/bioinformatics/17.8.754.

Ishikawa, Y., 2020. Insect Sex Pheromone Research and beyond: from Molecules to Robots. Springer, Singapore.

Johnson, K.P., Dietrich, C.H., Friedrich, F., Beutel, R.G., Wipfler, B., Peters, R.S., Allen, J. M., Petersen, M., Donath, A., Walden, K.K.O., Kozlov, A.M., Podsiadlowski, L., Mayer, C., Meusemann, K., Vasilikopoulos, A., Waterhouse, R.M., Cameron, S.L., Weirauch, C., Swanson, D.R., Percy, D.M., Hardy, N.B., Terry, I., Liu, S., Zhou, X., Misof, B., Robertson, H.M., Yoshizawa, K., 2018. Phylogenomics and the evolution of hemipteroid insects. Proc. Natl. Acad. Sci. USA 115, 12775–12780. https://doi.org/10.1073/pnas.1815820115.

Junker, R.R., Kuppler, J., Amo, L., Blande, J.D., Borges, R.M., van Dam, N.M., Dicke, M., Dotterl, S., Ehlers, B.K., Etl, F., Gershenzon, J., Glinwood, R., Gols, R., Groot, A.T., Heil, M., Hoffmeister, M., Holopainen, J.K., Jarau, S., John, L., Kessler, A., Knudsen, J.T., Kost, C., Larue-Kontic, A.A.C., Leonhardt, S.D., Lucas-Barbosa, D., Majetic, C.J., Menzel, F., Parachnowitsch, A.L., Pasquet, R.S., Poelman, E.H., Raguso, R.A., Ruther, J., Schiestl, F.P., Schmitt, T., Tholl, D., Unsicker, S.B., Verhulst, N., Visser, M.E., Weldegergis, B.T., Köllner, T.G., 2018. Covariation and phenotypic integration in chemical communication displays: biosynthetic constraints and eco-evolutionary implications. New Phytol. 220, 739–749. https://doi.org/10.1111/nph.14505.

Jurenka, R.A., 2021. Lepidoptera: female sex pheromone biosynthesis and its hormonal regulation. In: Blomquist, G.J., Vogt, R.G. (Eds.), Insect Pheromone Biochemistry and Molecular Biology, second ed. Academic Press, San Diego, pp. 13–65.

Jurenka, R.A., Subchev, M., Abad, J.L., Choi, M.Y., Fabrias, G., 2003. Sex pheromone biosynthetic pathway for disparlure in the gypsy moth, *Lymantria dispar*. Proc. Natl. Acad. Sci. USA 100, 809–814. https://doi.org/10.1073/pnas.0236060100.

Kalyaanamoorthy, S., Bui Quang, M., Wong, T.K.F., von Haeseler, A., Jermiin, L.S., 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat. Methods 14, 587–589. https://doi.org/10.1038/nmeth.4285.

Kandasamy, T., Ekbal, S., Kumari, K., 2021. SSH reveals different functional categories of putative resin synthesizing genes in the industrial insect, *Kerria lacca* (Kerr). J. Asia-Pacif. Entomol. 24, 1187–1193. https://doi.org/10.1016/j.aspen.2021.11.001.

Keeling, C.I., Tittiger, C., MacLean, M., Blomquist, G.J., 2021. Pheromone production in bark beetles. In: Blomquist, G.J., Vogt, R.G. (Eds.), Insect Phermone Biochemistry and Molecular Biology. Academic Press, San Diego, pp. 123–154.

Khrimian, A., Shirali, S., Vermillion, K.E., Siegler, M.A., Guzman, F., Chauhan, K., Aldrich, J.R., Weber, D.C., 2014a. Determination of the stereochemistry of the aggregation pheromone of harlequin bug, *Murgantia histrionica*. J. Chem. Ecol. 40, 1260–1268. https://doi.org/10.1007/s10886-014-0521-2.

Khrimian, A., Zhang, A.J., Weber, D.C., Ho, H.Y., Aldrich, J.R., Vermillion, K.E., Siegler, M.A., Shirali, S., Guzman, F., Leskey, T.C., 2014b. Discovery of the aggregation pheromone of the brown marmorated stink bug (*Halyomorpha halys*) through the creation of stereoisomeric libraries of 1-bisabolen-3-ols. J. Nat. Prod. 77, 1708–1717. https://doi.org/10.1021/np5003753.

Krall, B.S., Zilkowski, B.W., Kight, S.L., Bartelt, R.J., Whitman, D.W., 1997. Chemistry and defensive efficacy of secretion of burrowing bug (*Sehirus cinctus cinctus*). J. Chem. Ecol. 23, 1951–1962. https://doi.org/10.1023/b:joec.0000006482.12576.90.

Kumar, S., Kempinski, C., Zhuang, X., Norris, A., Mafu, S., Zi, J.C., Bell, S.A., Nybo, S.E., Kinison, S.E., Jiang, Z.D., Goklany, S., Linscott, K.B., Chen, X.L., Jia, Q.D., Brown, S. D., Bowman, J.L., Babbitt, P.C., Peters, R.J., Chen, F., Chappell, J., 2016. Molecular diversity of terpene synthases in the liverwort *Marchantia polymorpha*. Plant Cell 28, 2632–2650. https://doi.org/10.1105/tpc.16.00062.

Lancaster, J., Khrimian, A., Young, S., Lehner, B., Luck, K., Wallingford, A., Ghosh, S.K. B., Zerbe, P., Muchlinski, A., Marek, P.E., Sparks, M.E., Tokuhisa, J.G., Tittiger, C., Köllner, T.G., Weber, D.C., Gundersen-Rindal, D.E., Kuhar, T.P., Tholl, D., 2018. De novo formation of an aggregation pheromone precursor by an isoprenyl diphosphate synthase-related terpene synthase in the harlequin bug. Proc. Natl. Acad. Sci. USA 115, E8634–E8641. https://doi.org/10.1073/pnas.1800008115.

Lancaster, J., Lehner, B., Khrimian, A., Muchlinski, A., Luck, K., Köllner, T.G., Weber, D. C., Gundersen-Rindal, D.E., Tholl, D., 2019. An IDS-type sesquiterpene synthase produces the pheromone precursor (*Z*)-α-bisabolene in *Nezara viridula*. J. Chem. Ecol. 45, 187–197. https://doi.org/10.1007/s10886-018-1019-0.

Lassance, J.M., Ding, B.J., Lofstedt, C., 2021. Evolution of the codling moth pheromone via an ancient gene duplication. BMC Biol. 19 https://doi.org/10.1186/s12915-021-01001-8.

Lespinet, O., Wolf, Y.I., Koonin, E.V., Aravind, L., 2002. The role of lineage-specific gene family expansion in the evolution of eukaryotes. Genome Res. 12, 1048–1059. https://doi.org/10.1101/gr.174302.

Letunic, I., Bork, P., 2021. Interactive Tree of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res. 49, W293–W296. https://doi.org/10.1093/nar/gkab301.

Li, G.L., Köllner, T.G., Yin, Y.B., Jiang, Y.F., Chen, H., Xu, Y., Gershenzon, J., Pichersky, E., Chen, F., 2012. Nonseed plant *Selaginella moellendorfii* has both seed plant and microbial types of terpene synthases. Proc. Natl. Acad. Sci. USA 109, 14711–14715. https://doi.org/10.1073/pnas.1204300109.

Livak, K.J., Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$. Method Methods 25, 402–408.

Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A., Lanfear, R., 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol. Biol. Evol. 37, 1530–1534. https://doi.org/10.1093/molbev/msaa015.

Moliterno, A.A.C., De Melo, D.J., Zarbin, P.H.G., 2021. Identification of zingiberenol and murgantiol as components of the aggregation-sex pheromone of the rice stink bug, *Mormidea v-luteum* (Heteroptera: Pentatomidae). J. Chem. Ecol. 47, 1–9. https://doi.org/10.1007/s10886-020-01231-0.

Montagne, J., Wicker-Thomas, C., 2021. Drosophila pheromone production. In: Blomquist, G.J., Vogt, R.G. (Eds.), Insect Biochem. Mol. Biol. . Academic Press, San Diego, pp. 163–176.

Morgan, E.D., 2010. Biosynthesis in Insects, Advanced Edition ed. RSC Publishing, Cambridge, UK.

Muchlinski, A., Chen, X., Lovell, J.T., Köllner, T.G., Pelot, K.A., Zerbe, P., Ruggiero, M., Callaway, L., Laliberte, S., Chen, F., Tholl, D., 2019. Biosynthesis and emission of stress-induced volatile terpenes in roots and leaves of switchgrass (*Panicum virgatum* L.). Front. Plant Sci. 10 https://doi.org/10.3389/fpls.2019.01144.

Nallamsetty, S., Austin, B.P., Penrose, K.J., Waugh, D.S., 2005. Gateway vectors for the production of combinatorially-tagged His(6)-MBP fusion proteins in the cytoplasm and periplasm of *Escherichia coli*. Protein Sci. 14, 2964–2971. https://doi.org/10.1110/ps.051718605.

Noriega, F.G., 2014. Juvenile hormone biosynthesis in insects: what is new, what do we know, and what questions remain? ISRN Zoology, 967361.

Palazzo, M.C., Setzer, W.N., 2009. Monoterpene hydrocarbons may serve as antipredation defensive compounds in *Boisea trivittata*, the boxelder bug. Nat. Prod. Commun. 4, 457–459.

Panfilio, K.A., Angelini, D.R., 2018. By land, air, and sea: hemipteran diversity through the genomic lens. Curr. Opin. Insect Sci. 25, 106–115. https://doi.org/10.1016/j.cois.2017.12.005.

Persoons, C.J., Verwiel, P.E.J., Talman, E., Ritter, F.J., 1979. Sex-pheromone of the american cockroach, *Periplaneta americana* - isolation and structure elucidation of periplanone B. J. Chem. Ecol. 5, 221–236. https://doi.org/10.1007/bf00988237.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., Ferrin, T.E., 2004. UCSF chimera - a visualization system for exploratory research and analysis. J. Comput. Chem. 25, 1605–1612. https://doi.org/10.1002/jcc.20084.

Pickett, J.A., Allemann, R.K., Birkett, M.A., 2013. The semiochemistry of aphids. Nat. Prod. Rep. 30, 1277–1283. https://doi.org/10.1039/c3np70036d.

Piskorski, R., Hanus, R., Vasickova, S., Cvacka, J., Sobotnik, J., Svatos, A., Valterova, I., 2007. Nitroalkenes and sesquiterpene hydrocarbons from the frontal gland of three *Prorhinotermes* termite species. J. Chem. Ecol. 33, 1787–1794. https://doi.org/10.1007/s10886-007-9341-y.

Pokorny, T., Vogler, I., Losch, R., Schlütting, P., Juarez, P., Bissantz, N., Ramírez, S.R., Eltz, T., 2017. Blown by the wind: the ecology of male courtship display behavior in orchid bees. Ecology 98, 1140–1152. https://doi.org/10.1002/ecy.1755.

Rivera, S.B., Swedlund, B.D., King, G.J., Bell, R.N., Hussey, C.E., Shattuck-Eidens, D.M., Wrobel, W.M., Peiser, G.D., Poulter, C.D., 2001. Chrysanthemyl diphosphate synthase: isolation of the gene and characterization of the recombinant non-head-to-tail monoterpene synthase from *Chrysanthemum cinerariaefolium*. Proc. Natl. Acad. Sci. USA 98, 4373–4378.

Salmon, M., Laurendon, C., Vardakou, M., Cheema, J., Defernez, M., Green, S., Faraldos, J.A., O'Maille, P.E., 2015. Emergence of terpene cyclization in *Artemisia annua*. Nat. Commun. 6, 6143. https://doi.org/10.1038/ncomms7143.

Schulz, S., Estrada, C., Yildizhan, S., Boppre, M., Gilbert, L.E., 2008. An antiaphrodisiac in *Heliconius melpomene* butterflies. J. Chem. Ecol. 34, 82–93. https://doi.org/10.1007/s10886-007-9393-z.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., Ideker, T., 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 13, 2498–2504. https://doi.org/10.1101/gr.1239303.

Shirangi, T.R., Dufour, H.D., Williams, T.M., Carroll, S.B., 2009. Rapid evolution of sex pheromone-producing enzyme expression in *Drosophila*. PLoS Biol. 7, e1000168 https://doi.org/10.1371/journal.pbio.1000168.

Seybold, S.J., Quilici, D.R., Tillman, J.A., Vanderwel, D., Wood, D.L., Blomquist, G.J., 1995. De novo biosynthesis of the aggregation pheromone components ipsenol and ipsdienol by the pine bark beetles *Ips paraconfusus* Lanier and *Ips pini* (Say) (Coleoptera: scolytidae). Proc. Natl. Acad. Sci. U. S. A. 92, 8393–8397. https://doi.org/10.1073/pnas.92.18.8393.

Sievers, F., Higgins, D.G., 2018. Clustal Omega for making accurate alignments of many protein sequences. Protein Sci. 27, 135–145. https://doi.org/10.1002/pro.3290.

Smith, R.M., Brophy, J.J., Cavill, G.W.K., Davies, N.W., 1979. Iridodials and nepetalactone in the defensive secretion of the coconut stick insects, *Graeffea crouani*. J. Chem. Ecol. 5, 727–735. https://doi.org/10.1007/BF00986557.

Sobotnik, J., Hanus, R., Kalinova, B., Piskorski, R., Cvacka, J., Bourguignon, T., Roisin, Y., 2008. (*E,E*)-alpha-farnesene, an alarm pheromone of the termite *Prorhinotermes canalifrons*. J. Chem. Ecol. 34, 478–486. https://doi.org/10.1007/s10886-008-9450-2.

Sparks, M.E., Bansal, R., Benoit, J.B., Blackburn, M.B., Chao, H., Chen, M., Cheng, S., Childers, C., Dinh, H., Doddapaneni, H.V., Dugan, S., Elpidina, E.N., Farrow, D.W., Friedrich, M., Gibbs, R.A., Hall, B., Han, Y., Hardy, R.W., Holmes, C.J., Hughes, D.S. T., Ioannidis, P., Cheatle Jarvela, A.M., Johnston, J.S., Jones, J.W., Kronmiller, B.A., Kung, F., Lee, S.L., Martynov, A.G., Masterson, P., Maumus, F., Munoz-Torres, M., Murali, S.C., Murphy, T.D., Muzny, D.M., Nelson, D.R., Oppert, B., Panfilio, K.A., Paula, D.P., Pick, L., Poelchau, M.F., Qu, J., Reding, K., Rhoades, J.H., Rhodes, A., Richards, S., Richter, R., Robertson, H.M., Rosendale, A.J., Tu, Z.J., Velamuri, A.S., Waterhouse, R.M., Weirauch, M.T., Wells, J.T., Werren, J.H., Worley, K.C., Zdobnov, E.M., Gundersen-Rindal, D.E., 2020. Brown marmorated stink bug, *Halyomorpha halys* (Stål), genome: putative underpinnings of polyphagy, insecticide resistance potential and biology of a top worldwide pest. BMC Genom. 21 https://doi.org/10.1186/s12864-020-6510-7.

Sparks, M.E., Rhoades, J.H., Nelson, D.R., Kuhar, D., Lancaster, J., Lehner, B., Tholl, D., Weber, D.C., Gundersen-Rindal, D.E., 2017. A transcriptome survey spanning life stages and sexes of the harlequin bug, *Murgantia histrionica*. Insects 8, 55. https://doi.org/10.3390/insects8020055.

Sparks, M.E., Shelby, K.S., Kuhar, D., Gundersen-Rindal, D.E., 2014. Transcriptome of the invasive brown marmorated stink bug, *Halyomorpha halys* (Stal) (Heteroptera: Pentatomidae). PLoS One 9. https://doi.org/10.1371/journal.pone.0111646.

Stökl, J., Steiger, S., 2017. Evolutionary origin of insect pheromones. Curr. Opin. Insect. Sci. 24, 36–42. https://doi.org/10.1016/j.cois.2017.09.004.

Studer, G., Rempfer, C., Waterhouse, A.M., Gumienny, R., Haas, J., Schwede, T., 2020. QMEANDisCo-distance constraints applied on model quality estimation. Bioinformatics 36, 1765–1771. https://doi.org/10.1093/bioinformatics/btz828.

Tabata, J., 2017. Chemical ecology of insects: Applications and associations with plants and microbes. CRC Press, Boca Raton, FL.

Tabata, J., Ichiki, R.T., 2016. Sex pheromone of the cotton mealybug, *Phenacoccus solenopsis*, with an unusual cyclobutane structure. J. Chem. Ecol. 42, 1193–1200. https://doi.org/10.1007/s10886-016-0783-y.

Tholl, D., 2015. Biosynthesis and biological functions of terpenoids in plants. In: Schrader, J., Bohlmann, J. (Eds.), Biotechnology of Isoprenoids. Springer International Publishing, pp. 63–106.

Tholl, D., 2021. Biosynthesis of terpene pheromones in hemiptera/stink bugs. In: Blomquist, G.J., Vogt, R.G. (Eds.), Insect Pheromone Biochemistry and Molecular Biology, second ed. Academic Press, San Diego, pp. 269–281.

Tupec, M., Buček, A., Janoušek, V., Vogel, H., Prchalová, D., Kindl, J., Pavlíčková, T., Wenzelová, P., Jahn, U., Valterová, I., Pichová, I., 2019. Expansion of the fatty acyl reductase gene family shaped pheromone communication in Hymenoptera. Elife 8, e39231. https://doi.org/10.7554/eLife.39231.

Vandermoten, S., Mescher, M.C., Francis, F., Haubruge, E., Verheggen, F.J., 2012. Aphid alarm pheromone: an overview of current knowledge on biosynthesis and functions. Insect Biochem. Mol. Biol. 42, 155–163. https://doi.org/10.1016/j.ibmb.2011.11.008.

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F. T., de Beer, T.A.P., Rempfer, C., Bordoli, L., Lepore, R., Schwede, T., 2018. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res. 46, W296–w303. https://doi.org/10.1093/nar/gky427.

Weber, D.C., Khrimian, A., Blassioli-Moraes, M.C., Millar, J.G., 2018. Semiochemistry of Pentatomoidea. In: McPherson, J.E. (Ed.), Invasive stink bugs and related species (Pentatomoidea): Biology, higher systematics, semiochemistry, and management. CRC Press, Boca Raton, Florida, pp. 677–725.

Wei, G., Eberl, F., Chen, X., Zhang, C., Unsicker, S.B., Köllner, T.G., Gershenzon, J., Chen, F., 2020. Evolution of isoprenyl diphosphate synthase-like terpene synthases in fungi. Sci. Rep. 10 https://doi.org/10.1038/s41598-020-71219-z.

Weiss, I., Rössler, T., Hofferberth, J., Brummer, M., Ruther, J., Stökl, J., 2013. A nonspecific defensive compound evolves into a competition– avoidance cue and a female sex-pheromone. Nat. Commun. 4, 2767. https://doi.org/10.1038/ncomms3767.

Yang, Z., 2007. Paml 4: phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. 24, 1586–1591. https://doi.org/10.1093/molbev/msm088.

Zi, J., Mafu, S., Peters, R.J., 2014. To gibberellins and beyond! Surveying the evolution of (di)terpenoid metabolism. Annu. Rev. Plant Biol. 65, 259–286.