A radiation of *Psylliodes* flea beetles on Brassicaceae is associated with the evolution of specific detoxification enzymes

Matilda W. Gikonyo[1], Seung-Joon Ahn[1], Maurizio Biondi[2], Frank Fritzlar[3], Yu Okamura[4], Heiko Vogel[4], Tobias G. Köllner[5], Ismail Şen[6], David Hernández-Teixidor[7], Chi-Feng Lee[8], Harald Letsch[9*] and Franziska Beran[1*]

[1]Research Group Sequestration and Detoxification in Insects, Max Planck Institute for Chemical Ecology, Jena, Germany

[2]Department of Life, Health and Environmental Sciences, University of L'Aquila, L'Aquila, Italy

[3]Thüringer Landesamt für Umwelt, Bergbau und Naturschutz, Weimar, Germany

[4]Department of Entomology, Max-Planck Institute for Chemical Ecology, Jena, Germany

[5]Department of Natural Product Biosynthesis, Max-Planck Institute for Chemical Ecology, Jena, Germany

[6]Faculty of Technology, Isparta University of Applied Sciences, Isparta, Turkey

[7]Island Ecology and Evolution Research Group, Instituto de Productos Naturales y Agrobiología (IPNA-CSIC), La Laguna, Tenerife, Canary Islands, Spain

[8]Applied Zoology Division, Taiwan Agricultural Research Institute, Taichung, Taiwan

[9]Division of Tropical Ecology and Animal Biodiversity, University of Vienna, Vienna, Austria

*Corresponding

Harald Letsch, Division of Tropical Ecology and Animal Biodiversity, University of Vienna, Vienna, Austria. Email: harald.letsch@univie.ac.at

Franziska Beran, Research Group Sequestration and Detoxification in Insects, Max Planck Institute for Chemical Ecology, Jena, Germany. Email: fberan@ice.mpg.de

Abstract

Flea beetles of the genus *Psylliodes* have evolved specialized interactions with plant species belonging to several distantly related families, mainly Brassicaceae, Solanaceae, and Fagaceae. This diverse host use indicates that *Psylliodes* flea beetles are able to cope with different chemical defense metabolites, including glucosinolates, the characteristic defense metabolites of Brassicaceae. Here we investigated the evolution of host use and the emergence of a glucosinolate-specific detoxification mechanism in *Psylliodes* flea beetles. In phylogenetic analyses, *Psylliodes* species clustered into four major clades, three of which contained mainly species specialized on either Brassicaceae, Solanaceae, or Fagaceae. Most members of the fourth clade have broader host use, including Brassicaceae and Poaceae as major host plant families. Ancestral state reconstructions suggest that *Psylliodes* flea beetles were initially associated with Brassicaceae and then either shifted to Solanaceae or Fagaceae, or expanded their host repertoire to Poaceae. Despite a putative ancestral association with Brassicaceae, we found evidence that the evolution of glucosinolate-specific detoxification enzymes coincides with the radiation of *Psylliodes* on Brassicaceae, suggesting that these are not required for using Brassicaceae as hosts but could improve the efficiency of host use by specialized *Psylliodes* species.

Introduction

Most herbivorous insect species are specialists that feed on only a few plant species within one plant genus or family (Forister et al., 2015). This dietary specialization is often accompanied by the evolution of specific resistance mechanisms that allow the insect to overcome the host plant's characteristic chemical defense or enable the recruitment of plant toxins for the insect's own protection (Beran & Petschenka, 2022; Heckel, 2014; Jeschke et al., 2016). Indeed, a number of studies revealed an important impact of plant defense compounds on insect host use (Becerra, 1997; Berenbaum, 2001; Edger et al., 2015), which is conserved in many insect clades at least at the plant family level (Ehrlich & Raven, 1964; Jolivet & Hawkeswood, 1995; Robinson et al., 2023). On the other hand, there are striking examples of host-use variability, where closely related insect species specialize on distantly related plants that contain different types of defense compounds (Gikonyo et al., 2019; Jolivet & Hawkeswood, 1995; Salvi et al., 2019). However, not much is known about how resistance mechanisms evolve in insect clades with diverse host use.

Leaf beetles (Chrysomelidae) are associated with more than 100 different plant families and the majority of the roughly 40,000 species are considered as specialists (Jolivet & Hawkeswood, 1995; Mitchell, 1981; Rheinheimer & Hassler, 2018). According to phylogenetic studies of several different chrysomelid genera, the evolutionary history of host associations in leaf beetles is remarkably complex and highly taxon-specific. The proposed scenarios range from almost parallel diversification of beetles and their host plants in the genus *Phyllobrotica* (Farrell & Mitter, 1990) to dynamic host shifts between different plant families in the genera *Oreina* and *Gonioctena* (Dobler et al., 1996; Mardulyn et al., 1997). Plant defense compounds are also thought to have played an important role in the evolution of host use in leaf beetles, for example, in the genera *Blepharida*, *Phratora*, and *Phyllotreta*, which specialize mainly on the plant families Burseraceae, Salicaceae, and Brassicaceae, respectively (Becerra, 1997; Gikonyo et al., 2019; Köpf et al., 1998).

In contrast to the previously mentioned genera, which are mainly associated with one plant family, the leaf beetle genus *Psylliodes*, with about 200 described species, is associated with more than 20 different plant families (Gikonyo, Biondi, & Beran, 2019; Jolivet & Hawkeswood, 1995). The major host plant families Brassicaceae, Solanaceae, Poaceae, and Fagaceae are not only distantly related, but also contain very different types of chemical defense compounds such as glucosinolates, alkaloids, benzoxazinoids, and phenolic compounds (Blažević et al., 2020; Burlacu et al., 2020; Chowanski et al., 2016; Niculaes et al., 2018). *Psylliodes* species must thus be able to cope with diverse plant defense compounds. While most

3

Psylliodes species are specialized on plant species belonging to a single plant family, some species are known to have a broader host repertoire that often includes at least one of the major host plant families mentioned above.

Based on morphology, *Psylliodes* species have been classified into five different sub-genera and within the largest subgenus *Psylliodes sensu stricto* into several species groups (Doguet, 1994; Leonardi, 1970; Nadein, 2006, 2007a, 2007b). Assignment to species groups is not always clear and depends on the morphological characters considered, so that the composition of the species groups partly overlaps. It is, however, interesting to note that species that have been assigned to a certain group are often also associated with similar host plants; for example, members of the *Psylliodes napi* species group mainly feed on Brassicaceae, while several species assigned to the *Psylliodes picina* group are associated with Fagaceae (reviewed in Gikonyo et al. (2019)). Although this suggests that host use is to some extent phylogenetically conserved, it remains unknown how *Psylliodes* species have evolved specialized interactions with such diverse plant families.

The major host plant family of *Psylliodes*, the family Brassicaceae, is well-known for its characteristic chemical defense, the glucosinolate-myrosinase system. Glucosinolates are amino acid-derived secondary metabolites that are hydrolyzed by myrosinase enzymes upon tissue damage to release deterrent and reactive isothiocyanates and other biologically active hydrolysis products (Blažević et al., 2020). Both specialist and generalist herbivores from different insect orders and feeding guilds are able to overcome this defense and were shown to employ an astonishing variety of resistance strategies targeting either the glucosinolate itself, the process of glucosinolate hydrolysis, or the breakdown product (Friedrichs et al., 2020; Jeschke et al., 2016; Malka et al., 2020). Previous research on the specialist cabbage stem flea beetle, *Psylliodes chrysocephala*, revealed that adult beetles combine at least three different strategies to cope with the glucosinolate-myrosinase system in their brassicaceous host plants (Ahn et al., 2019; Beran et al., 2018). Two of these prevent glucosinolate hydrolysis either by sequestration or by enzymatic conversion to desulfo-glucosinolates. A third strategy prevents isothiocyanate toxicity through the conjugation to glutathione, followed by metabolism of the isothiocyanate-glutathione conjugate via the mercapturic acid pathway. Considering that studies on other insect species usually found one major detoxification pathway (Friedrichs et al., 2020; Gloss et al., 2014; Jeschke et al., 2017), this combination of several different resistance mechanisms in *P. chrysocephala* is unusual and raises questions about the evolutionary history of these different mechanisms in the genus *Psylliodes*.

4

Molecular studies focused on the identification of glucosinolate-specific sulfatase (GSS) enzymes in *P. chrysocephala* (Ahn et al., 2019). Comparative transcriptome analyses combined with biochemical studies revealed a specific expansion of arylsulfatases in this species that was linked to the evolution of GSS activity. In total, five GSS with different substrate specificities have been identified, which are derived from a conserved coleopteran arylsulfatase gene (*Sulf4*).

The goal of this study is to gain insights into the evolution of host use and a glucosinolate-specific resistance mechanism in the genus *Psylliodes*. To this end, we inferred the phylogenetic relationships between 60 *Psylliodes* species and reconstructed the ancestral host associations using models that account for changes in the host repertoire. Our taxon sampling covers most major species groups (i.e., 13 out of 16 groups comprising at least three species) that have been proposed based on morphological features (Leonardi, 1970; Nadein, 2006, 2007a). The host use of *Psylliodes* species assigned to the three major species groups that were not sampled (*P. glaber* group, *P. montana* group, *P. saulcyi* group) is largely unknown (reviewed in Gikonyo et al. (2019)).

Based on the obtained phylogenetic framework, we investigated the molecular evolution of *GSS* in *Psylliodes*. Specifically, we sought to understand whether the ability to detoxify glucosinolates by desulfation represents an ancestral trait of P*sylliodes* or whether GSS have evolved in within this genus. To address this question, we searched for GSS candidates in the transcriptomes of several *Psylliodes* species representing most of the major clades recovered in the species phylogeny, and analyzed their enzymatic activities and gene diversification patterns.

Materials and Methods

Taxon sampling

For phylogenetic analyses, we obtained adult specimens of 52 *Psylliodes* species of which 50 could be determined to species level based on morphological characters (Doguet, 1994; Nadein, 2007b; Nadein & Lee, 2012; Warchałowski, 2010). The identification of the species was verified by specialists in flea beetle taxonomy (Maurizio Biondi, Frank Fritzlar, Ismail Şen, Ali Gök, Konstantin Nadein). From *P. chrysocephala*, a species known to be morphologically highly variable (Doguet, 1994; Warchałowski, 2010), we included two specimen, the typical dark metallic form, and *P. chrysocephala* var. *collaris*, a form with red pronotum and green or blue elytra. Specimens were preserved in 100% ethanol and stored at -20°C until DNA extraction. Six additional *Psylliodes* species were obtained as DNA samples. Two other *Psylliodes* species were included using COI sequences that have been deposited in NCBI.

5

Except for *P. wollastoni*, which belongs to the subgenus *Eupus* (Nadein, 2007a), all sampled *Psylliodes* species belong to the subgenus *Psylliode*s sensu stricto. These species cover 17 out of 27 taxonomic species groups that have been established based on morphological features (reviewed in Gikonyo et al. (2019)). However, several species groups comprise only one or two species, so when considering only groups with at least three species, 13 out of 16 species groups were included in our taxon sampling. Collection data, host use, and species group information for all *Psylliodes* species included in this study are summarized in Supplementary Table S1. As outgroups, 40 species of Phytophaga beetles comprising 19 other Chrysomelidae, 16 Cerambycidae, 4 Curculionidae, and 1 Brentidae, were included (Supplementary Table S2).

For transcriptome sequencing, we collected adult specimens of six *Psylliodes* species from host plants or using sweep nets in the field (Supplementary Table S3). In addition, we sampled larvae and pupae of *P. chrysocephala* from our laboratory colony (described in Ahn et al. (2019)) for transcriptome sequencing.

RNA isolation and cDNA synthesis

Total RNA was extracted from beetle samples using the InnuPREP RNA Mini Kit (Analytik Jena AG, Jena, Germany). Detailed information on the samples is provided in Supplementary Table S3. Remaining DNA was removed using an on-column DNase treatment (Qiagen) according to the manufacturer's protocol. RNA quality was determined using an Agilent 2100 Bioanalyzer. Total RNA was used to synthesize first-strand cDNA using the SuperScript® IV First-Strand Synthesis System (Invitrogen), and 5'- and 3'-rapid amplification of cDNA ends (RACE)-ready cDNA using the SMARTer RACE cDNA Amplification Kit (Clontech) according to the manufacturer's instructions, respectively.

Transcriptome sequencing and assembly

RNA-Seq libraries were prepared and sequenced at the Max Planck Genome Centre (Cologne, Germany). Poly-A RNA was enriched from up to 1 µg of total RNA using the NEBNext® Poly(A) mRNA Magnetic Isolation Module. RNA-Seq libraries were prepared as described in NEBNext Ultra™ II Directional RNA Library Prep Kit for Illumina (New England Biolabs). A total of eleven cycles were applied to enrich library concentration. Sequencing-by-synthesis was performed on a HiSeq 2500 Sequencing System (Illumina, San Diego, CA, USA) using paired-end read technology (2 × 100 bp or 2 × 150 bp). Illumina sequencing resulted in 100-200 million reads for each library (Table S3). Filtering, removal of adaptor sequences, trimming of reads, and *de novo* transcriptome assemblies were performed using CLC Genomics Workbench version 7.0.1. The following assembly parameters were applied: automatic word size = yes; bubble size = 250; minimum contig length = 200; auto-detect paired distances = yes; perform scaffolding

6

= yes; map reads back to contigs = yes, with mismatch cost = 2; insertion and deletion cost = 3, length fraction = 0.8, and similarity fraction = 0.9. Scaffolding was selected and conflicts among individual bases were resolved in all assemblies by voting for the base with the highest frequency. Short contigs (< 200 bp) were removed from the final datasets. The completeness of the assembled transcriptomes was evaluated using Benchmarking Universal Single-Copy Orthologs (BUSCO) version 3.0.2 (Seppey, Manni, & Zdobnov, 2019) using the Arthropoda odb 9 gene set (1066 genes; Supplementary Table S3). RNA-Seq data have been deposited in GenBank under the BioProject accession number PRJNA768846.

DNA extraction

Genomic DNA was extracted from one or two hind legs using the DNeasy Blood and Tissue Kit (Qiagen) according to the manufacturer's protocol. Legs were first crushed in extraction buffer using a plastic pestle, then homogenized with three metal beads (2.4 mm diameter, Askubal, Korntal-Münchingen, Germany) using a TissueLyzer II (Qiagen) at 30Hz for 10 min, and incubated with Proteinase K at 56 ºC overnight. The concentration of the extracted DNA was measured using a NanoDrop™ ND-1000 Spectrophotometer (Thermo Fisher Scientific). Whole genome amplification was performed for DNA samples with concentrations below 5 ng/µl using the GenomePlex® Complete Whole Genome Amplification Kit (Sigma-Aldrich, Germany) to obtain sufficient amounts of DNA.

Selection of marker genes

To identify suitable molecular markers for phylogenetic analyses, we randomly selected 72 candidate genes from the list of coleopteran single-copy genes provided in Misof et al. (2014). We first identified these genes in the annotated genomes of *Tribolium castaneum* and *Dendroctonus ponderosae* (Richards et al. 2008; Keeling et al. 2013) and used the corresponding amino acid sequences for reciprocal tBLASTn searches against the *Psylliodes* transcriptomes with default settings (BLOSUM62 matrix, E-value cut-off of 10). We recovered a total of 43 of these candidate genes as full length transcripts in the de novo assembled transcriptomes of all eight *Psylliodes* species. To facilitate analyses of marker gene sequences to be amplified from genomic DNA, we then focused on exon sequences. Therefore, we analyzed the exon-intron structure of *T. castaneum* and *D. ponderosae* candidate genes using Splign (Kapustin et al., 2008), and identified genes with at least one conserved exon with a minimum length of 500 bp. This step reduced the number of candidate genes to 25. To predict the corresponding exon sequences in *Psylliodes* genes, we aligned the gene sequences to those of *T. castaneum* and *D. ponderosae* in MAFFT (Katoh et al., 2019) using default settings. We further assessed the suitability of these genes for phylogenetic analyses by performing preliminary tests using nucleotide alignments of each individual gene and its

7

predicted longest conserved exon, respectively. Genes were considered as potentially suitable if they produced a phylogenetic tree with a majority of the nodes having bootstrap support values of ≥50, and if the branching patterns obtained using the full gene sequence and the longest conserved exon were comparable. This step reduced the number of candidate genes to 22. We designed degenerate primers for these 22 candidate genes and tested them using genomic DNA extracted from *P. chrysocephala* as a template. Sequencing of the sub-cloned PCR products revealed that 16 candidates were successfully amplified; however, six candidates were excluded because they contained an intron. Based on amplification success of the remaining ten candidate genes from genomic DNA of other *Psylliodes* species, we finally used eight single-copy nuclear genes together with the mitochondrial COI as molecular markers in this study. The primer sequences are provided in Supplementary Table S4.

Markers were amplified from genomic DNA using a proofreading DNA polymerase, either Advantage 2 Polymerase Mix (Takara Bio) or Phusion™ High-Fidelity DNA Polymerase (Thermo Fisher Scientific). PCR products were visualized by gel electrophoresis using a 1% agarose gel stained with Invitrogen SYBR® Safe DNA gel stain (Thermo Fisher Scientific). Target PCR products were purified from the agarose gel using the Zymoclean Gel DNA Recovery Kit (Zymo Research), cloned into the pCR4-TOPO vector (Invitrogen) and Sanger-sequenced. The sequences have been deposited in GenBank (MW254364 - MW254893) and are additionally provided in Supplementary Table S5. The results of the analysis of the COI sequences obtained in this study are described in the SI Results and Discussion. Marker gene sequences from outgroup species were obtained using tBLASTn searches in the corresponding transcriptome assemblies using protein sequences of *T. castaneum* as queries.

Phylogenetic reconstruction of *Psylliodes* species relationships

Nucleotide sequences of marker genes were aligned according to codon positions using MUSCLE (Edgar, 2004) in MEGA X (Kumar et al., 2018) and concatenated in Geneious 11.1.5. We tested for potential substitutional saturation using "Xia's test" (Xia & Lemey, 2009; Xia et al., 2003) in the program DAMBE v7.2 (Xia, 2018), and found the third codon position of all marker genes to be saturated (Supplementary Table S6). We thus partitioned the alignment according to gene and the first two codon positions using ModelFinder implemented in IQ-TREE (Kalyaanamoorthy et al., 2017; Nguyen et al., 2015), resulting in 8,808 aligned nucleotides and 11 partitions. We restricted the model search to those supported by the Bayesian Inference (BI) software package BEAST (Drummond et al., 2012) for subsequent analyses. Phylogenetic analyses were conducted using Maximum likelihood (ML) methods implemented in IQ-TREE 1.6.11 (Nguyen et al., 2015). Node support was calculated using corrected Ultrafast Bootstrapping (Hoang

8

et al., 2018) in 1,000 replicates with the "bnni" option to reduce the risk of overestimating branch support (Hoang et al., 2018), and an increased maximum number of iterations to stop (-nm 10,000). To avoid local optima, we repeated the ML analysis 50 times and selected the best tree by log likelihood.

We additionally used amino acid sequences to reconstruct the phylogenetic relationships among *Psylliodes* spp. Amino acid sequences were aligned using the Geneious alignment tool hosted in Geneious 11.1.5, with default settings (BLOSUM62 matrix; gap opening cost: 12, gap extension cost: 3) and subsequently concatenated. The alignment was partitioned according to gene using ModelFinder implemented in IQ-TREE (Kalyaanamoorthy et al., 2017; Nguyen et al., 2015), resulting in 2,936 aligned amino acids and 9 partitions. ML analysis was conducted in IQ-TREE 1.6.11 (Nguyen et al., 2015). Node support was calculated using Ultrafast Bootstrapping (Hoang et al., 2018) with 1,000 replicates corrected with the "bnni" option. All sequence alignments are provided in Edmond, the data repository of the Max Planck Society, under https://doi.org/10.17617/3.UYO1D2.

Divergence time estimation

Divergence time estimates were calculated in BEAST 1.10.4 using BI with MCMC chains of 50,000,000 generations, and a sampling frequency of 5,000 generations. The analysis was conducted using the tree topology obtained from the best IQ-TREE analysis as a fixed starting tree. The partitioning scheme and models of nucleotide substitution were defined according to the results of the ModelFinder analyses in IQ-TREE and models were unlinked among partitions. For the clock model priors, we used the uncorrelated lognormal relaxed clock (UCLN) model (Drummond et al., 2006). We tested the impact of i) different tree priors, ii) different fossil calibration prior densities, as well as iii) different fossil calibration schemes. To do so, we first applied different tree models, using either a Yule (pure-birth) tree prior or a birth-death tree prior. To calibrate the divergence time estimates, five fossil calibration points were applied, under three different prior distributions. In a first setup, all fossils were applied as uniform priors, with their age as the lower bound and the upper bound defined by the age of oldest polyphagan beetle †*Leehermania prorova* (223 Ma; (Chatzimanolis et al., 2012)), except for the root fossil, whose maximum age was defined by the oldest definite beetle †*Coleopsis archaica* (293.7 Ma; (Kirejtshuk et al., 2014)). To infer if the estimated ages are conditioned by the priors, we used a second setup, where the upper bound of the fossil calibration priors was defined by the age of the oldest hexapod †*Rhyniella praecusor* (412 Ma; (Wolfe et al., 2016). Additionally, we applied a third setup, where all fossils were applied as lognormal priors, with hard lower bounds according to the first setup and a soft upper bound, so that 95% of the distribution lay between the fossil age and 223 Ma. To test the fit of all different parameter settings, we

9

used Bayes factors (BF), obtained by marginal likelihood estimations (MLE) of all six analyses, using the path sampling (PS) and stepping-stone sampling (SS) methods in BEAST with default parameter settings (Baele et al., 2013). Convergence of the posterior distribution of parameter values was evaluated by the effective sample sizes (ESS) of all estimated parameters and branch lengths being greater than 200 in Tracer 1.7 (Rambaut et al., 2018). The first ten percent of generations were discarded and post burn-in trees were combined and used to construct the maximum clade credibility tree with median node heights in TreeAnnotator 1.10.4 (Suchard et al., 2018). Results are summarized in Supplementary Supplementary Table S7.

To further test the impact of fossil calibration priors on divergence times, we conducted additional BEAST analyses using the best of the above outlined setups and excluded †*Taimyraltica calcarata* (Fossil Group 2, FG2), and/or †*Crepidocnema yantarica* (Fossil Group 3, FG3). Additional information on the fossil priors used in this study are provided in Supplementary Table S8. To infer the impact of prior settings on the divergence time estimation, we additionally ran an analysis with an empty alignment, using the 'prior-only' option.

Assessment of phylogenetic conservatism of host plant use

Based on the obtained *Psylliodes* species tree and available host plant use data (Supplementary Table S1), we tested whether host use is phylogenetically constrained using delta statistics, an optimized method for detecting phylogenetic signals on categorical data (Borges et al., 2019). For each *Psylliodes* species, we assigned a binary state of host use for each of the four major host plant families, *i.e.,* Brassicaceae, Solanaceae, Fagaceae and Poaceae (used = 1, not used = 0). To infer significance, delta statistics calculated for each plant family were compared to the distribution of delta statistics based on randomized trait states (100 iterations), respectively. We performed this analysis in R version 4.1.0 using the script available under https://github.com/mrborges23/delta_statistic.

Ancestral host reconstruction

To reconstructed the evolutionary history of host use in *Psylliodes,* we used the MultiState module in BayesTraits v4.0.0 (Meade & Pagel, 2022; Pagel & Meade, 2006), which allows the use of multiple character states for each taxon. We used the best time-calibrated tree of the divergence time analyses in BEAST and pruned all taxa except *Psylliodes*. Host plant families according to Supplementary Table S1 (both major and potential host plants) were selected as character states. We tested the fit of two different models, one in which all transition rates between sets were constrained to be equal, and another in which the transition rates were free to vary. Models were fitted using a reversible-jump Markov chain Monte

10

Carlo (rjMCMC) analysis to derive posterior distributions of the ancestral state and transition rates. An exponential reversible-jump hyperprior (0 100) was specified for the rate parameter distributions. For each analysis, two independent Markov chains were run for 10,000,000 generations, with a burn-in of 1,000,000 and sampling every 1,000 generations to achieve adequate mixing and stationarity. We examined the parameter traces in Tracer version 1.7 to ensure convergence and effective sample sizes across the two independent runs. To compare model fitting, we used Bayes factors (BF), as obtained by MLE for each model via the stepping-stone sampler in BayesTraits, using 1,000 stones, each run for 100,000 iterations.

We additionally inferred the evolution of host repertoires in *Psylliodes* using a model of insect-host coevolution (Braga et al., 2020). Host repertoires cover both potential hosts and actual hosts, i.e., the fundamental host repertoire, presuming that potential hosts can be acquired or lost, thus developing into actual hosts or vice versa. The model requires both the phylogenetic relationships of insects and host plants. For this purpose, we used the phylogenetic tree of angiosperm families published by Magallon et al. (2015) and pruned this tree so that all 13 host plant families in the data set were represented. Hosts were coded according to Supplementary Table S1 as 0 (nonhost), 1 (potential host plant family or refugial host plant family), or 2 (major host plant family). Host repertoire reconstruction analyses were conducted in the software RevBayes (Höhna et al., 2016), using Bayesian Markov chain Monte Carlo (MCMC) inference. We ran two independent chains for 100,000 cycles, sampling every 50 cycles, and discarding the first 20,000 cycles as burn-in. Convergence and effective sample sizes across the runs were examined in Tracer version 1.7. To visualize the ancestral host reconstructions, the dimensionality of the host repertoire is reduced by assigning hosts to so-called modules, which are based on actual insect–host interactions. Modules are defined as groups of plants and beetles that interact more with each other than with other taxa (Braga et al., 2020). Results were displayed using the *plot_matrix_phylo* function of the evolnets package (Braga et al., 2020) in R (R Development Core Team, 2022), with modules defined with the *set.seed* (seed value: 111) and *mycomputeModules* functions in R. The posterior probability above which the ancestral states were shown, was set to 90% (default).

Identification of candidate *GSS* genes in beetle transcriptomes

Candidate *GSS* genes were identified in transcriptomes of different *Psylliodes* species using amino acid sequences of GSS from *P. chrysocephala* (KX986118 − KX986122) and GSS1 from *Plutella xylostella* (CAD33828.1) as queries in tBLASTn searches with default settings (BLOSUM62 matrix, E-value cut-off of 10). For partial transcripts, full-length open reading frames (ORFs) were obtained by RACE-PCR using the

11

SMARTer RACE cDNA Amplification Kit (Clontech) according to the manufacturer's instructions. Gene sequences were confirmed by amplification of full-length ORFs from cDNAs, cloning into the pCR4-TOPO vector (Invitrogen) and Sanger-sequencing. For one gene from *P. hospes*, only a partial sequence was amplified because several attempts to recover the full-length sequence by RACE-PCR failed. In the transcriptome of *P. kiesenwetteri*, we additionally found two pseudogenes with sequence similarity to arylsulfatases/*GSS*. Partial sequences of both pseudogenes were amplified from cDNAs and their sequences were verified by Sanger-sequencing. Sequences of primers used to amplify arylsulfatase/GSS genes from *Psylliodes* spp. are provided in Supplementary Table S9. Gene sequences have been submitted to GenBank under accession numbers MW179460 - MW179496 and are provided in Supplementary Table S10.

Phylogenetic analysis of coleopteran arylsulfatase-like genes

Nucleotide sequences of arylsulfatase-like genes from *Psylliodes* spp. were aligned together with those previously identified in *P. chrysocephala* (KX986114 - KX986122), *Phyllotreta striolata* (KX086123 - KX086126), *Phyllotreta armoraciae* (KX986127 - KX986130), and *GSS1* from *P. xylostella* (*PxGSS1*, MG022096.1) as outgroup using the ClustalW 2.1 algorithm implemented in Geneious 11.1.5, with a gap opening cost of 10 and a gap extension cost of 0.1. Phylogenetic relationships were inferred by ML using the TIM3+F+I+G4 substitution model as selected by ModelFinder implemented in IQ-TREE with 1,000 bootstrap replicates, and by BI in MrBayes 3.2.7 hosted by CIPRES (Miller, Pfeiffer, & Schwartz, 2010). The MCMC analysis was conducted in two runs consisting of 100,000,000 generations with a sampling frequency of 1,000 and discarded the first 25% of the resulting trees as burn-in. Trees were rooted using *PxGSS1*.

Sulf4-like gene tree and *Psylliodes* species tree reconciliation

To predict *Sulf4*-like gene duplications and losses in *Psylliodes*, we reconciled the gene tree and the species tree using Notung 2.9.1.5 (Durand, Halldorsson, & Vernot, 2006). We pruned the *Psylliodes* species tree (Figure 1) to the seven *Psylliodes* species for which we generated transcriptomes by using the *drop.tip* function in the ape package (Paradis & Schliep, 2019) in R (R Development Core Team, 2022). The same method was used to prune the arylsulfatase gene tree described in the previous section to *Sulf4*-like genes from *Psylliodes* spp. We rearranged the gene tree using the -rearrange option in Notung with 75% bootstrap threshold to avoid false inference from weakly-supported nodes, and inferred putative *Sulf4*-like gene duplication and loss events using the -reconciliation option.

12

Analysis of predicted protein sequences

The molecular mass and isoelectric point of the putative arylsulfatases were predicted using the ProtParam tool (Gasteiger et al., 2005). The presence of a signal peptide and transmembrane domains were predicted using SignalP 4.1 (Petersen et al., 2011) and Phobius 1.01 (Kall et al., 2007), respectively. Sulfatase signature motifs were predicted using PROSITE (Sigrist et al., 2002). Subcellular protein localization was predicted using DeepLoc 1.0 (Almagro Armenteros et al., 2017) and the number of *N*-glycosylation sites were predicted using NetNGlyc 1.0 (Gupta & Brunak, 2002). Results are summarized in Supplementary Table S10.

Functional characterization of *Sulf4-like* genes from *Psylliodes* spp.

Since *GSS* have evolved by duplications of the conserved coleopteran *Sulf4* gene in *P. chrysocephala* (Ahn et al., 2019), we selected *Sulf4-like* genes identified by phylogenetic analyses in other *Psylliodes* spp. for functional characterization. *Sulf4* gene sequences excluding the C-terminal transmembrane domain were amplified from the pCR4-TOPO vector containing the full-length ORFs (described above) using gene-specific primers (Supplementary Table S9) and cloned into the pIEx-4 expression vector (Merck Chemicals GmbH, Darmstadt, Germany). Expression constructs were verified by Sanger-sequencing and used to transfect *Sf*9 insect cells using FuGENE HD Transfection Reagent (Promega GmbH, Mannheim, Germany) as described in Beran et al. (2014). After incubation at 27°C for 72 h, the cell culture medium was collected and centrifuged (4°C, 10 min, 21,000 × g) to remove cell debris, and proteinase inhibitors (EDTA-free Protease Inhibitor Cocktail Set V; Merck Chemicals GmbH, Darmstadt, Germany) were added to the supernatant. Successful expression was verified by Western blotting using an anti-His (C-terminal)-HRP-antibody (dilution 1:10,000; Invitrogen).

Arylsulfatase activity assays were performed by incubating 50 µl of culture medium containing the recombinant enzymes with 50 µl of 2 mM 4-nitrocatechol sulfate (4NCS) (Sigma-Aldrich) in 50 mM MES buffer with 250 mM NaCl (pH 5.2) at 35°C for 1 h. Control assays were performed with culture medium that was boiled at 95°C for 5 min. All assays were performed in triplicates. After incubation, 100 µl 0.2 M NaOH were added to stop the reaction and 4-nitrocatechol (4NC) (Sigma-Aldrich) was measured at 515

13

nm in a Tecan Infinite 200 spectrophotometer. The mean of the background controls was subtracted from each value, and the amounts the hydrolysis product were calculated using an external calibration curve prepared from 4-nitrocatechol.

In GSS activity assays, 50 µl culture medium containing the recombinant enzymes were incubated with 50 µl of a glucosinolate mixture (2 mM each) in 50 mM MES buffer supplemented with 250 mM NaCl (pH 5.2) at 35°C for 2 hrs. Glucosinolates were purchased from Phytoplan (Heidelberg, Germany) or Carl Roth GmbH (Karlsruhe, Germany), except for 4-hydroxybenzyl glucosinolate, which was extracted from *Sinapis alba* seeds following Thies (1988). One mixture contained aliphatic glucosinolates (2-propenyl glucosinolate, 2-hydroxy-3-butenyl glucosinolate, 4-methylsulfinylbutyl glucosinolate, 4-methylthiobutyl glucosinolate), the other mixture contained benzenic and indolic glucosinolates (benzyl glucosinolate, 2-phenylethyl glucosinolate, indol-3-ylmethyl glucosinolate). Control assays were performed with culture medium that was boiled at 95°C for 5 min. Assays were stopped by adding 200 µl of pure methanol. Residual intact glucosinolates were removed using 200 µl of 10% (w/v) diethylaminoethyl (DEAE)-Sephadex A-25 (GE Healthcare Life Science, Pittsburgh, PA, USA). Samples were filtered through a 0.2 µm syringe filter (Pall Corporation, Dreieich, Germany) and dried using a speed vacuum concentrator (Eppendorf, Hamburg, Germany). Assay products were dissolved in 90 µl ultrapure water and incubated with 0.0625 U (625 µg) myrosinase (Sigma-Aldrich) in 50 mM MES buffer (pH 6.5) and 25% glycerol at 25ºC overnight. Samples were stored at -20ºC until liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis.

LC-MS/MS

Desulfo-glucosinolates were analyzed by LC-MS/MS using an Agilent 1200 HPLC system (Agilent, Santa Clara, CA, USA) connected to an API3200 tandem mass spectrometer (AB Sciex Germany GmbH, Darmstadt, Germany). Desulfo-glucosinolates were separated on a Nucleodur Sphinx RP column (250 × 4.6 mm, 5 µm particle size; Macherey-Nagel) using a mobile phase consisting of 0.2% (v/v) formic acid in ultrapure water as solvent A and acetonitrile as solvent B, at a flow rate of 1 ml/min. The gradient was as follows: 1.5% (v/v) B (1 min), 1.5–5% (v/v) B (5 min), 5–7% (v/v) B (2 min), 7–13.3% (v/v) B (4.5 min), 13.3–100% (v/v) B (0.1 min), 100% B (0.9 min), 100–1.5% (v/v) B (0.1 min) and 1.5% (v/v) B (3.9 min). Desulfo-

14

glucosinolates were detected in positive ionization mode. Ionspray voltage was maintained at 5,000 eV. Gas temperature was set to 700°C, nebulizing gas and drying gas to 60 psi, curtain gas to 35 psi, and collision gas to 5 psi. Data analysis was performed using Analyst Software 1.6 Build 3773 (AB Sciex). To monitor analyte parent ion to product ion formation we used multiple-reaction monitoring (Supplementary Table S11). Desulfo-glucosinolates were quantified via external standard calibration curves prepared from desulfation of the eight glucosinolates using sulfatase enzyme solution from *Helix pomatia* prepared according to Graser et al. (2001).

Test for signatures of positive selection on *GSS* genes

To test for signatures of positive selection on *GSS*, we performed branch-site model tests (Zhang, Nielsen, & Yang, 2005) and additionally estimated the ratio of nonsynonymous (dN) to synonymous (dS) substitutions (ω ratio) for each branch in a *Sulf4/GSS* gene tree. Nucleotide sequences of *Sulf4/GSS* genes from *Psylliodes* spp. and *Phyllotreta* spp. excluding the pseudogenes and *PhosGSS-like 2* were aligned by codons using the Geneious alignment tool hosted in Geneious 11.1.5, applying the Blosum62 matrix (gap opening cost: 13, gap extension cost: 3). *PhosGSS-like 2* was excluded because it reduced the sites included in the analysis from 1665 bp to only 726 bp. The sequence alignment was used to construct a ML tree using the TN+F+I+G4 model as selected in IQ-TREE using ModelFinder, with 1,000 bootstrap replicates. The branching pattern of *Sulf4/GSS* genes from Brassicaceae-feeding *Psylliodes* species was congruent with that recovered in the complete gene tree (described above), whereas the branching pattern of *Sulf4* genes from other *Psylliodes* species differed between both analyses. The ω ratio for each branch was estimated using CODEML integrated in PAML (Yang, 2007). For this analysis, we collapsed nodes with bootstrap support values ≤ 50. In the branch-site model, we only inspected *Sulf4/GSS*-branches of Brassicaceae-feeding *Psylliodes* species using the following parameters: model = 2, NSsites = 2, and cleandata = yes in CODEML. We ran two different models, a null model, which fixed the non-synonymous substitution (dN) to synonymous substitution (dS) ratios (dN/dS) at the focal branch as 1 (with fixed_omega (ω) = 1 option), and an alternative model, which allowed the branch to have varied dN/dS ratios across sites. To determine whether the two models differ significantly, we used a likelihood-ratio test. *P*-values from each analysis were adjusted with false discovery rates (FDRs) to account for multiple testing (Benjamini & Hochberg, 1995). Amino acid sites with signatures of positive selection were inferred using the Bayes empirical Bayes analysis (BEB; 0.95 cut-off). To visualize the position of sites under putative positive selection, we modelled the protein structure of *Pchr*GSS2 as a representative of the

Sulf4/GSS clade in SWISS-MODEL (http://www.expasy.org/swissmod/SWISS-MODEL.html; (Schwede et al., 2003)) using the N-acetylgalactosamine-4-sulfatase from *Homo sapiens* (pdb 1fsu.1.A) as template. The model was visualized and analyzed using PyMOL (The PyMOL Molecular Graphics System, Version 1.2r3pre, Schrödinger, LLC).

Results

Phylogenetic relationships and divergence times in the genus *Psylliodes*

To investigate the evolution of host use in *Psylliodes* flea beetles, we inferred the relationships between 60 *Psylliodes* species using the maximum likelihood method based on the first and second codon position of the nucleotide (nt) data set. The results of the phylogenetic tree reconstructions and divergence time estimations are shown in Figure 1 and Supplementary Figure S1. Branch lengths and node ages refer to the median ages estimated using the best performing prior model for divergence time estimation in BEAST (Supplementary Table S7). Based on BF comparisons between the analyses with different tree priors and fossil calibration prior densities, the model with a uniform prior with younger maximum ages represented the best fit (Supplementary Table S7). In general, age estimations were largely congruent among all models, including those with reduced fossil calibration setups (Table 1). In contrast, divergence times clearly diverged from the complete analyses in an analysis using exclusively priors and an empty alignment, indicating that the BI analyses are not mainly driven by the priors. We also compared the result of the phylogenetic analysis based on the nt data set with that based on amino acid (aa) data. Both phylogenetic trees were generally congruent, with only minor differences between both (Supplementary Figures S2 and S3).

In Galerucinae, the Alticini appear monophyletic as the sister group to Galerucini (nt and aa bootstrap support ($BS_{nt/aa}$) = 100). Divergence time estimations propose the origin of Galerucinae about 74 Ma (median age; 95% Highest Posterior Density (HPD) = 63.1 – 91.5 Ma) in the Late Cretaceous, and the origin of Alticini about 66 Ma (95% HPD = 55.8 – 81.3 Ma) in the Early Paleocene. Within Alticini ($BS_{nt/aa}$ = 100), *Phyllotreta* was recovered as the sister group to *Psylliodes* ($BS_{nt/aa}$ = 97/95), with their divergence time

16

estimated at about 45 Ma (95% HPD = 37.3 – 56.6 Ma) in the Eocene. Emergence of *Psylliodes* is dated at about 24 Ma (95% HPD = 19.3 – 30.6 Ma) in the Late Oligocene. Within *Psylliodes*, *P. attenuata* appeared as sister group to all other species ($BS_{nt/aa}$ = 100), which clustered into four major clades characterized by similar host plant use (Fig 1; clades A – D). Clade A ($BS_{nt/aa}$ = 71/82) emerged about 16 Ma (95% HPD 13.6 – 20.5) and comprises species that mainly feed on Brassicaceae. This clade was the sister group to clade B ($BS_{nt/aa}$ = 87/96), which emerged around 15 Ma (11.8 – 18.9 Ma) and comprises species associated with Solanaceae. Several Solanaceae-feeding species were also placed within the Brassicaceae-feeding clade A, but their position in the tree was not well supported ($BS_{nt/aa}$ = 69/60). The monophyletic clades A and B ($BS_{nt/aa}$ = 99) were sister to the monophyletic clades C and D ($BS_{nt/aa}$ = 100). Clade C emerged about 10 Ma (95% HPD = 8.2 – 13.6 Ma) and comprised species that mainly feed on Fagaceae. Clade D ($BS_{nt/aa}$ = 89/83) emerged about 9 Ma (95% HPD = 6.8 – 11.5 Ma) and comprised mainly species with a broader host repertoire including Poaceae in addition to Brassicaceae.
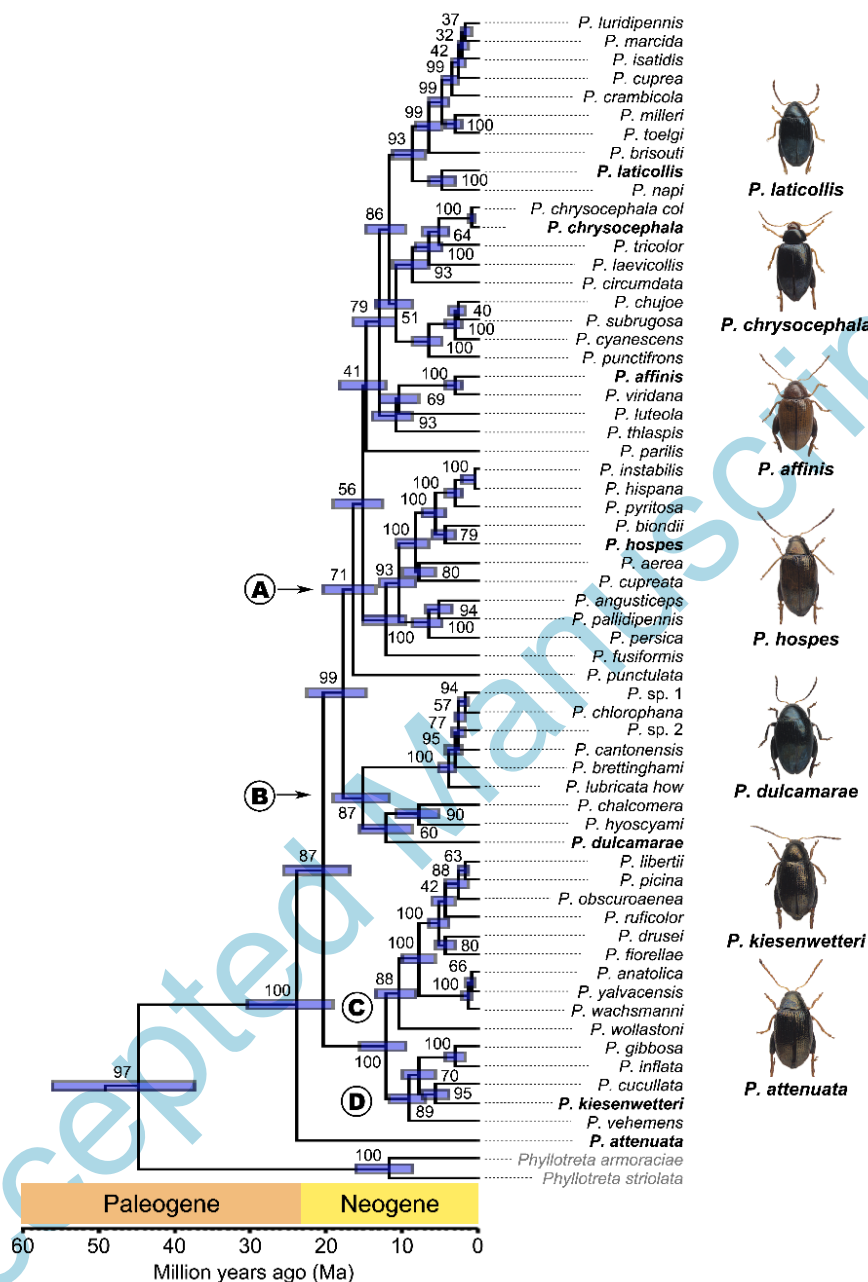
Figure 1. Phylogeny and divergence times of *Psylliodes* flea beetles. Maximum clade credibility tree with median ages in million years ago (Ma) from Bayesian unlinked uniform analysis in BEAST. The phylogeny was constructed from a combined nucleotide dataset using the first two codon positions of eight single-copy nuclear genes and the mitochondrial *cytochrome c oxidase I (COI)*. Ultrafast Bootstrap support values (1,000 replicates) are indicated on each node. The node bars indicate 95% highest probability density (HPD) values on node height (age). Chrysomelidae, Cerambycidae and Curculionidae species were used as outgroups (Figures S1 and S2). Letters A − D represent major *Psylliodes* clades. *Psylliodes* species with transcriptome data are written in bold. *P. chrysocephala col*, *P. chrysocephala* var. *collaris; P. lubricata how, P. lubricata* var. *howensis.* Beetle photos: Maurizio Biondi.

Table 1. Summary of tree reconstruction supports and divergence time estimates. Ultrafast bootstrap support values (BS) of major clades recovered in Maximum Likelihood analyses based on nucleotide (nt) or amino acid (aa) sequences of eight single-copy nuclear genes and *cytochrome c oxidase I* (1,000 replicates each). Median age estimations in million years ago (Ma) from three different fossil group (FG) combinations with upper and lower 95% highest posterior density values are in parenthesis; FG1: Five fossils; FG2: four fossils excluding *Taimyraltica calcarata*; FG3: four fossils excluding *Crepidocnema yantarica*.

Clade	Major host plant families	Support (BS$_{nt/aa}$)	Median age estimations in million years ago (Ma)		
			FG1	FG2	FG3
Chrysomeloidea	-	100/100	130.1 (122.5-158.3)	130.0 (122.5-160.4)	129.5 (122.5-155.3)
Chrysomelidae	-	100/100	118.5 (106.2-145.7)	118.3 (105.1-148.2)	118.0 (105.7-143.0)
Galerucinae	-	100/100	74.1 (63.1-91.5)	74.4 (63.7-94.1)	74.2 (63.7-91.7)
Alticini	-	100/100	65.6 (55.8-81.3)	65.9 (56.9-83.0)	66.3 (56.8-84.2)
Phyllotreta + Psylliodes	-	97/95	45.0 (37.3-56.6)	45.7 (37.2-58.5)	45.4 (37.1-58.4)
Psylliodes	-	100/100	24.0 (19.3-30.6)	24.6 (19.8-32.0)	24.5 (19.3-33.7)
Clade A	Brassicaceae	71/82	16.4 (13.6-20.5)	17.0 (13.6-22.9)	16.9 (14.0-21.9)
Clade B	Solanaceae	87/96	15.0 (11.8-18.9)	15.7 (12.0-20.3)	15.5 (11.3-21.1)
Clade C	Fagaceae	88/89	10.4 (8.0-13.6)	10.8 (8.2-14.1)	10.8 (8.1-14.0)
Clade D	Brassicaceae and Poaceae	89/83	8.8 (6.8-11.5)	9.1 (6.8-12.0)	9.1 (6.6-12.6)

To compare the inferred phylogenetic relationships among *Psylliodes* species with previous phylogenetic hypotheses based on morphological data (Leonardi, 1970; Nadein, 2007a), we mapped the species group information onto the phylogenetic tree (Supplementary Figure S4). We found that most species that have been assigned to a specific species group also formed well-supported clades in the phylogeny. One exception are species assigned to the *P. chrysocephala* group, which were distributed across several sub-clades in clade A. In addition, one sub-clade within clade A comprised four species, each of which has been assigned to a different species group.

Phylogenetic conservatism of host use

Because the major clades in the phylogeny of *Psylliodes* spp. were each associated with one or two particular plant families, we used delta statistics to test for phylogenetic conservatism of host use at the plant family level. Results from this analysis revealed strong phylogenetic signals for the use of both Brassicaceae and Poaceae, ($P > 0.001$), and weaker phylogenetic signals for the use of Solanaceae ($P = 0.02$) and Fagaceae ($P = 0.032$), indicating that associations with these plant families are indeed phylogenetically conserved in the genus *Psylliodes*.

Ancestral host plant reconstructions

To find out how host use evolved in the genus *Psylliodes*, we reconstructed ancestral host associations using two different methods (Figure 2 and Supplementary Figure S5). In BayesTraits, the BF comparisons between the model with independent transition rates among different hosts and the one with equal rates showed minimal differences (log BF < 0.5; (Kass & Raftery, 1995)). We thus discuss the results of the simpler model with equal rates (Figure 2). No particular plant family could be unambiguously assigned as ancestral host to the most recent common ancestor (MRCA) of the genus *Psylliodes*, but Brassicaceae showed the highest probability for this node (45%). An ancestral association with Brassicaceae is further predicted with higher probabilities for most backbone nodes within *Psylliodes* (75 – 99%). In addition to Brassicaceae, the ancestor of clades A and B has a higher probability of association with Solanaceae (clade names refer to Figure 1). Shifts to Solanaceae are predicted both within clade A and for the MRCA of clade B. Host use of the ancestor of clade C and D is again ambiguous, with similar probabilities for Brassicaceae, Poaceae, and Fagaceae. A shift to Fagaceae is predicted within clade C, whereas an expansion of host use from Brassicaceae to Poaceae is predicted within clade D. These results were largely consistent with results of the host repertoire analyses in RevBayes (Supplementary Figure S5). This analysis proposed that (i) the MRCA of *Psylliodes* used Brassicaceae and Cannabaceae, (ii) the ancestor of clade B used Solanaceae, (iii) an early ancestor within clade C used Fagaceae, (iv) the ancestor clade D used Brassicaceae, and (v) host use was expanded in clade D to Poaceae.
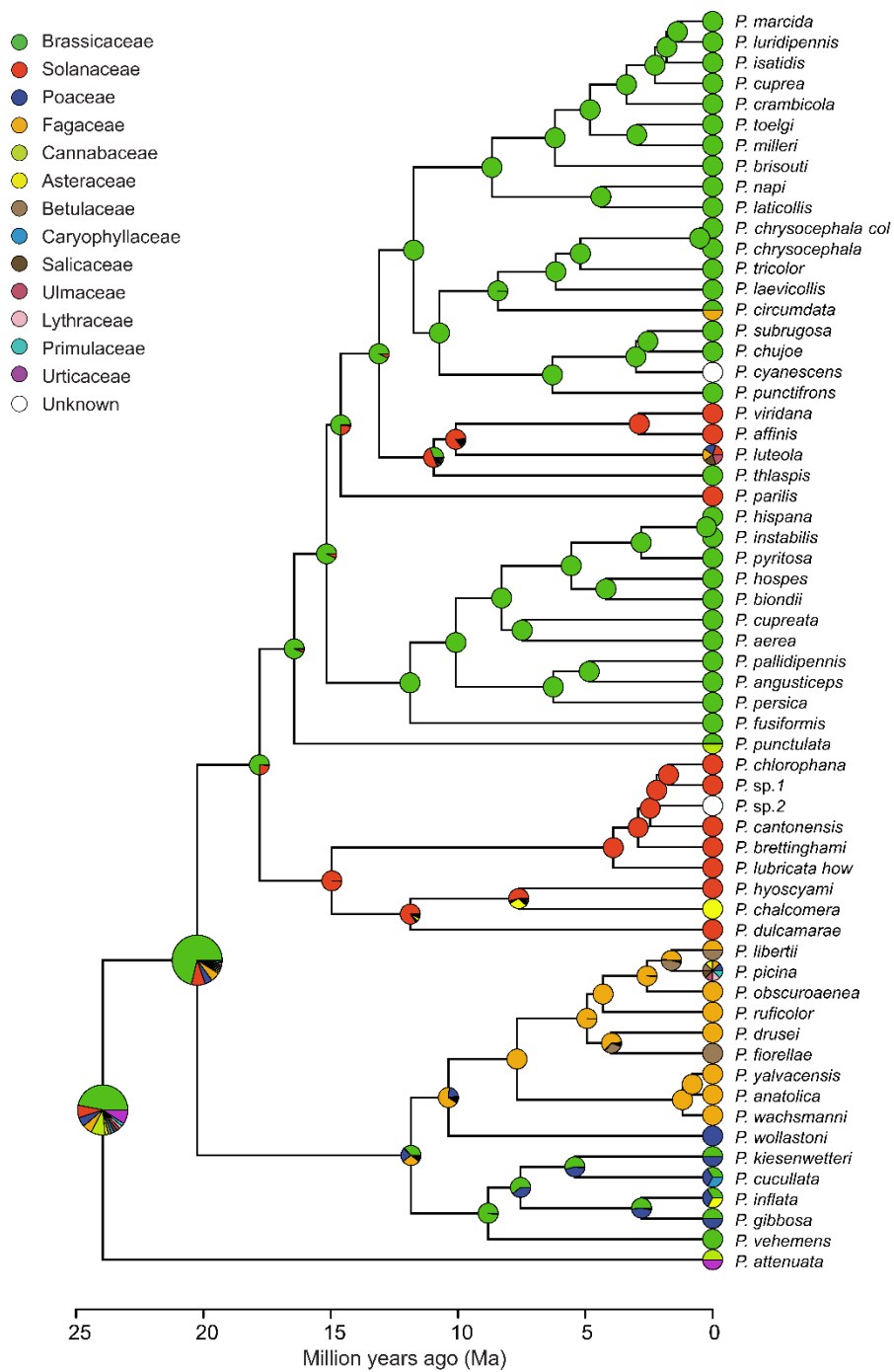
Figure 2. Bayesian ancestral state reconstruction of host plant use by Psylliodes flea beetles. Ancestral host use was reconstructed in BayesTraits using rjMCMC analysis. The proportions of the posterior probabilities are displayed as pie charts at the internal nodes of the tree. *P. chrysocephala col*, *P. chrysocephala* var. *collaris; P. lubricata how, P. lubricata* var. *howensis.*

21

Identification of candidate *GSS* in *Psylliodes* spp.

To investigate the molecular evolution of *GSS* in *Psylliodes*, we annotated all putative arylsulfatase genes in transcriptomes of six *Psylliodes* species including *P. attenuata* (the sister species to all other *Psylliodes* species in the phylogeny) and members of all major clades except clade C. Beetle transcriptomes harbored between four and eight arylsulfatase-like genes, which were recovered as full-length transcripts by RACE-PCR, except for one gene from *P. hospes*, for which only a partial sequence could be obtained (*PhosGSS-like 2*; Supplementary Table S8). The transcriptome of *P. kiesenwetteri* additionally harbored two pseudogenes with sequence similarity to arylsulfatases (Supplementary Figure S6). In addition, we found a putative arylsulfatase gene in a transcriptome of *P. chrysocephala* larvae and pupae (*PchrGSS6*), which was not present in the adult transcriptome analyzed earlier (Ahn et al., 2019).

In both the ML and BI analyses, the identified arylsulfatase-like genes grouped into four well-supported clades (Supplementary Figure S6) corresponding to the four clades of coleopteran arylsulfatases identified in Ahn et al. (2019). The *Sulf1*- and *Sulf2*-clades each contained one transcript per *Psylliodes* species, whereas species-specific gene duplications were observed within both the *Sulf3* and *Sulf4* clades. Because *Sulf3* genes were duplicated in both Brassicaceae-feeding and non-Brassicaceae-feeding species and previous research has shown that *GSS* have evolved by duplications of *Sulf4* in *P. chrysocephala* (Ahn et al., 2019), we focused only on the *Sulf4* clade. *Sulf4*-like gene duplications were found in three *Psylliodes* species, with four paralogs found in each *P. laticollis* and *P. hospes*, and three paralogs in *P. kiesenwetteri* (Supplementary Figure S6). Sequence analyses revealed that two of the three paralogs in *P. kiesenwetteri* represent pseudogenes (Supplementary Figure S7). The additional gene identified in *P. chrysocephala* also clustered in the *Sulf4* clade, bringing the total number of *Sulf4*-like genes in this species to seven.

We translated the arylsulfatase gene sequences into the corresponding amino acid sequences and analyzed the predicted properties of the resulting enzymes (Supplementary Table S10). Except one enzyme (PhosGSS-like 1), all Sulf4-like enzymes were predicted to be extracellular and anchored in the membrane with a single transmembrane domain at the C-terminus (Supplementary Tables S10 and S12). All *Sulf4*-like genes encode a functional sulfatase signature motif I (CXPXR, where X can be any amino acid). In sulfatase signature motif II, a conserved and catalytically active histidine residue was substituted by an asparagine residue in several predicted enzymes (Supplementary Table S10).

22

Functional characterization of *Sulf4-like* genes from *Psylliodes* spp.

Ten out of 12 *Sulf4*-like genes were successfully expressed without the predicted C-terminal transmembrane domain in *Sf*9 insect cells (Supplementary Figure S8). Results of activity assays using different glucosinolates and 4NCS as substrates were combined with data from previous activity assays performed with recombinant Sulf4/GSS enzymes from *P. chrysocephala* (Ahn et al. 2019) and are shown in Figure 3. A total of five recombinant enzymes showed GSS activity under our assay conditions, PlatGSS1 and PlatGSS2 from *P. laticollis*, PhosGSS1 and PhosGSS2 from *P. hospes*, and PchrGSS6 from *P. chrysocephala*. Genes that could not be functionally characterized were named according to their position in the phylogenetic tree as "*GSS-like*" (Figure 3 and Supplementary Figure S6). All remaining genes were named "*Sulf4*".

The recombinant GSSs differed widely regarding their substrate specificities. While PlatGSS1 and PhosGSS2 each hydrolyzed only one out of nine tested glucosinolates, PlatGSS2, PhosGSS1, and PchrGSS6 showed a broader substrate specificity and hydrolyzed several aliphatic, benzenic, and indolic glucosinolates (Figure 3B). The remaining Sulf4-like enzymes, except for PattSulf4, hydrolyzed the general arylsulfatase substrate 4NCS. Among the recombinant enzymes that showed GSS activity, only PhosGSS2 also hydrolyzed 4NCS.

Molecular evolution of GSS

In ML and BI analyses, *GSS* genes clustered into three well-supported clades. GSS clade 1 comprised only genes from *P. hospes*, GSS clade 2 comprised two *GSS* genes from *P. chrysocephala* and *PlatGSS-like*, and GSS clade 3 comprised *GSS* and *GSS-like* genes from all three species (Supplementary Figure S6). To gain more insights in the diversification of *Sulf4*-like genes in *Psylliodes*, we reconciled the *Sulf4*-like gene tree and the *Psylliodes* species tree using Notung. This analysis inferred independent duplications of the conserved Sulf4 gene in clades A and D that in clade A were followed by lineage-specific gene duplications and losses. In total, 11 duplication and 9 losses were inferred (Supplementary Figure S9).

To investigate the possible role of positive selection in the evolution of *GSS* genes in the genus *Psylliodes*, we estimated the ω ratio (ratio of nonsynonymous (dN) to synonymous (dS) substitutions) for each branch in a phylogeny of *Sulf4/GSS* genes from *Psylliodes* spp. (except the partial *PhosGSS-like* 2). The ω ratios ranged between 0.001 and 17.737, showing the rates of molecular evolution to differ widely among *Sulf4* and *GSS* genes in *Psylliodes* spp. With few exceptions, *Sulf4* branches had lower ω ratios than *GSS* branches, suggesting that overall, *GSS* genes are under more relaxed selection compared to *Sulf4* genes

23

(Supplementary Figure S10). Elevated ω ratios > 1, which indicate positive selection, were estimated for two branches associated with GSS clades 1 (ω = 1.292) and 3 (ω = 17.737).
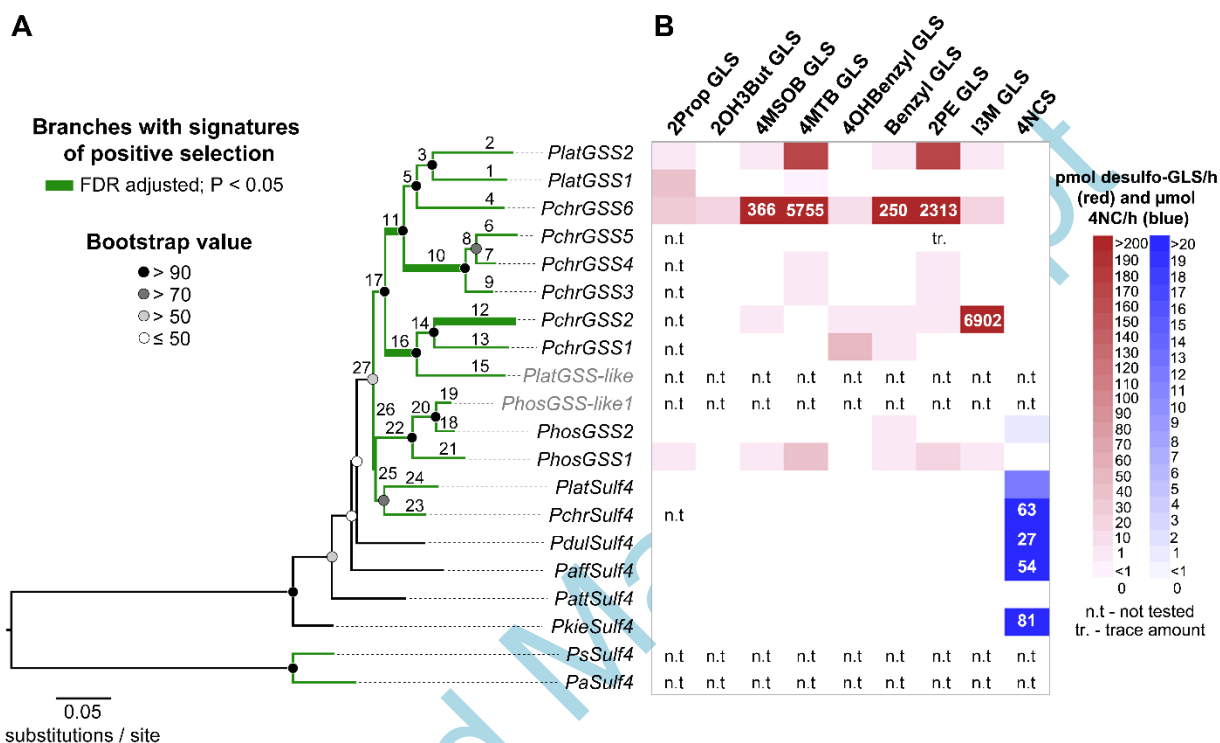

Figure 3. Test for signatures of positive selection on *Sulf4/GSS* genes from *Psylliodes* spp. and substrate specificities of the corresponding recombinant enzymes. (**A**) Maximum likelihood tree of *Sulf4/GSS* genes from *Psylliodes* spp. and *Phyllotreta* spp. used for branch model tests of positive selection. *Sulf4* sequences from *Phyllotreta* spp. were included as outgroup. Branches of genes from Brassicaceae-feeding species are coloured in green. Numbers on branches show the branches that were tested for positive selection, with branches in bold showing signatures signatures of positive selection (Supplementary Table S13). Dots at nodes represent ultrafast bootstrap support values above 50 (1,000 replicates). (B) Activity of recombinant enzymes towards glucosinolates (GLS) and the general substrate 4-nitrocatechol sulfate (4NCS). *PlatGSS-like* and *PhosGSS-like1* (names written in grey font colour) could not be tested because their expression level was below the detection limit (Figure S6). Results from enzyme assays were combined with previously published results of *P. chrysocephala* Sulf4/GSS (Ahn et al., 2019). The heat map was generated with a maximum cut-off of 200 pmol/hr desulfo-GLS and 20 μmol/hr 4-nitrocatechol (4NC). Higher values are provided as numbers in the heatmap. 2Prop, 2-propenyl; 2OH3But, 2-hydroxy-3-butenyl; 4MSOB, 4-methylsulfinylbutyl; 4MTB, 4-methylthiobutyl; 4OHBenzyl, 4-hydroxybenzyl; 2PE, 2-phenylethyl; I3M, indol-3-ylmethyl. *Pchr, Psylliodes chrysocephala; Plat, Psylliodes laticollis; Phos, Psylliodes hospes; Paff, Psylliodes affinis; Pdul, Psylliodes dulcamarae; Pkie, Psylliodes. kiesenwetteri; Patt, Psylliodes attenuata; Ps, Phyllotreta striolata; Pa, Phyllotreta armoraciae*.

We then inspected all branches of the *Sulf4/GSS* clade with branch-site codon substitution models. These analyses revealed signatures of positive selection at four branches associated with GSS clades 2 and 3, but not GSS clade 1 (Figure 3A, Supplementary Table S13). Specifically, a signature of positive selection was detected before diversification of *GSS* in clades 2 (branch 11, p = 0.005) and 3 (branch 16, p = 0.035), and in addition within both clades (Figure 3A, Supplementary Table S13). BEB analyses inferred a total of nine sites under putative positive selection (Supplementary Tables S12 and S13), of which four amino acid residues are predicted to be localized in the active site (Supplementary Figure S11). One of these sites is the catalytically active histidine residue, which is substituted with an asparagine residue in members of GSS clade 3 (Supplementary Table S12).

Discussion

To investigate the evolution of host use and an associated resistance trait in *Psylliodes* flea beetles, we reconstructed the phylogeny of 60 *Psylliodes* species, which represent most major taxonomic groups within the subgenus *Psylliodes* s. str. (reviewed in Gikonyo et al. (2019)). Our approaches could not unambiguously resolve all deeper relationships within *Psylliodes*, but recovered the position of *P. attenuata* as a sister group to the remaining species with maximum support. All other *Psylliodes* species clustered into two well-supported groups, each with two moderately supported clades (Figure 1). The largest of these clades (clade A) is the one with least support, but comprises several well-supported sub-clades of *Psylliodes* species that are specialized on Brassicaceae. With few exceptions, these sub-clades are concordant with species groups that have been proposed based on morphological features (Supplementary Figure S4). This also holds true for sub-clades within the three other major clades, which are mainly associated with either Solanaceae, Fagaceae, or with Brassicaceae and Poaceae, respectively. Consistent with these host use patterns, a test for phylogenetic conservatism provided strong evidence that associations with these four major host plant families are phylogenetically constrained in the genus *Psylliodes*.

In addition to Brassicaceae-specialists, clade A comprises three species that are specialized on Solanaceae, and *P. luteola*, a species with broader host use that includes fagaceous plants. The positions of all these species are weakly supported by both our molecular data and from a morphological perspective. For instance, *P. viridana*, a species that feeds on Solanaceae, has been assigned to the *P. brettinghami* group (Nadein, 2007b), whereas *P. brettinghami* clusters in clade B together with other species that feed on Solanaceae. Similarly, *P. luteola* has been proposed to belong to the *P. picina* group (Leonardi, 1970),

25

which represents the major clade within clade C. The position of *Psylliodes* species that do not feed on Brassicaceae in clade A can thus not be considered reliable.

Evolution of host use in the genus *Psylliodes*

The reconstruction of ancestral host use suggests that major host shifts from Brassicaceae-feeding ancestors to the distantly related plant families Solanaceae and Fagaceae have occurred early in the diversification of *Psylliodes* species and were each followed by species radiations. Within these groups, a few *Psylliodes* species have again shifted to other plant families, e.g., from Solanaceae to Asteraceae (clade B: *P. chalcomera*), and from Fagaceae to the closely related plant family Betulaceae (clade C: *P. fiorellae*). Whether additional independent host plant shifts from Brassicaceae to Solanaceae have occurred within clade A needs to be determined based on a more robust phylogenetic hypothesis. Within clade D, a host range expansion from Brassicaceae to Poaceae was predicted. In addition to this major host range expansion, there are other cases in which the host range of individual species has expanded. These species are scattered across the phylogenetic tree, suggesting that host range expansions (and contractions) have frequently occurred in the evolutionary history of *Psylliodes*.

Host shifts of herbivorous insects between distantly related plant families have been explained mainly by two different scenarios. In one scenario, host shifts may be facilitated by chemical similarity of the colonized plant families (Ehrlich & Raven, 1964). This scenario is unlikely in *Psylliodes*, as the major host plant families are known to contain very different types of characteristic defense compounds: glucosinolates in Brassicaceae, alkaloids in Solanaceae, phenolic metabolites in Fagaceae, and benzoxazinoid glucosides in Poaceae (Blažević et al., 2020; Chowanski et al., 2016; Niculaes et al., 2018). However, the expansion of the host range of *P. libertii* from Fagaceae to Betulaceae, and the host shift from Fagaceae to Betulaceae in *P. fiorellae* could be explained by this mechanism because both plant families are closely related and contain similar defense compounds.

Another scenario explains host shifts between distantly related plant families due to proximity of potential hosts in a plant community in the insect's habitat (Pasteels & Rowell-Rahier, 1991). This scenario was considered the most likely to explain the evolution of host use in the leaf beetle genus *Gonioctena*, which, like *Psylliodes*, is associated with distantly related plant families without obvious chemical similarities. *Psylliodes* species are known to occur in a variety of different habitats (Gök & Aslan, 2006; Rheinheimer & Hassler, 2018), making it likely that the availability of plant species has been an important factor in the evolution of host use in *Psylliodes*.
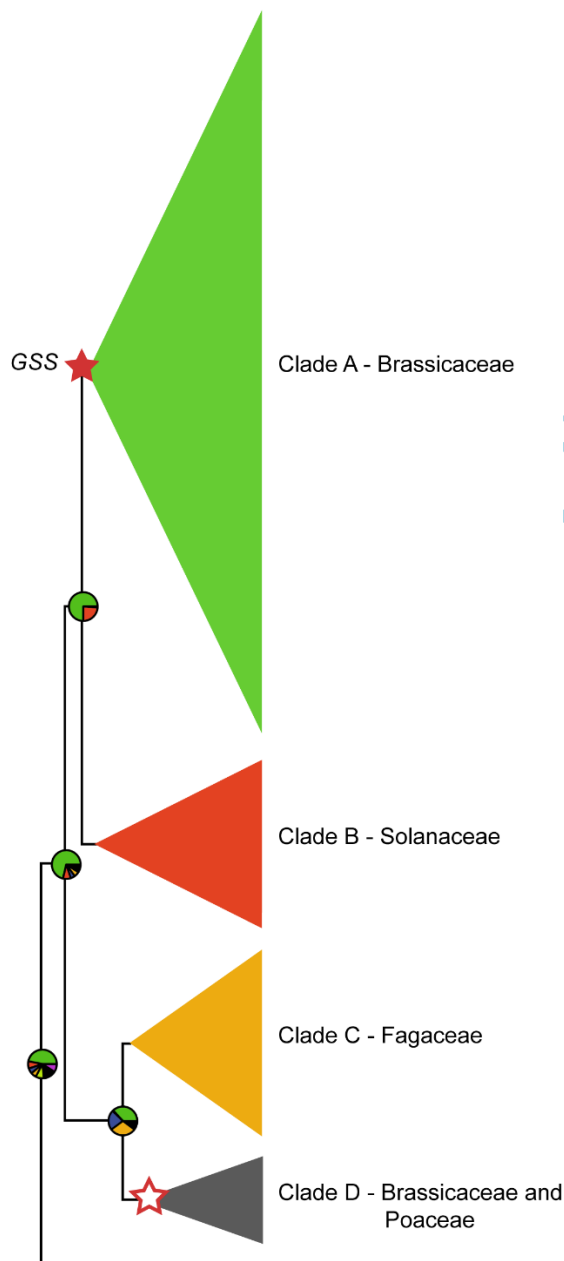
26

Figure 4. Evolution of host use and GSS in *Psylliodes* flea beetles. Species in the genus *Psylliodes* are associated with several distantly related plant families, mainly Brassicaceae, Solanaceae, Fagaceae, and Poaceae. Our results suggest an ancestral association of *Psylliodes* with Brassicaceae and two major host shifts, one to Solanaceae (clade B) and one to Fagaceae (in clade C). We also found evidence for a host range expansion from Brassicaceae to Poaceae (in clade D). Glucosinolate sulfatases (GSS), which have evolved by duplications of the conserved coleopteran arylsulfatase *Sulf4*, were found only in Brassicaceae-feeding species in clade A, and thus likely arose early during the radiation on Brassicaceae (red star). However, *Sulf4* gene duplications were also detected in a species belonging to clade D, but the duplicated genes are pseudogenes (unfilled star).

27

Evolution of host use in *Psylliodes* in a broader phylogenetic context

In recent phylogenetic studies on Alticini, the genus *Psylliodes* clustered together with the flea beetle genera *Phyllotreta*, *Epitrix*, *Chaetocnema*, and *Crepidodera* in a monophyletic clade (Douglas et al., 2023; Zhang et al., 2022). The genus *Phyllotreta* is also associated mainly with Brassicaceae and other closely related glucosinolate-containing plant families in the order Brassicales. However, despite sharing glucosinolate-containing plants as major hosts, *Psylliodes* and *Phyllotreta* were never recovered as sister genera (Douglas et al., 2023; Zhang et al., 2022), suggesting both have independently evolved to use Brassicaceae as a host. The different strategies that *Psylliodes* and *Phyllotreta* flea beetles use to cope with the glucosinolate-myrosinase defense system support this hypothesis (Beran et al., 2014; Beran et al., 2018). Instead, either *Epitrix* or *Crepidodera* were inferred as sister to *Psylliodes*, two genera associated mainly with Solanaceae or with Salicaceae and Rosaceae, respectively (Doguet, 1994; Jolivet & Hawkeswood, 1995; Rheinheimer & Hassler, 2018). The third closely related genus *Chaetocnema* is mainly associated with Poaceae, Amaranthaceae, and Polygonaceae (Rheinheimer & Hassler, 2018). The fact that the host use of *Psylliodes* partially overlaps with that of closely related flea beetle genera could suggest that the predicted host shifts from Brassicaceae-feeding ancestors to Solanaceae and Poaceae represent re-colonizations of ancestral host plant families that were kept in the fundamental host repertoire of this beetle lineage, a scenario predicted by the oscillation hypothesis where cycles of host expansion and specialization drive species diversification (Nylin et al., 2018).

Timing of diversification of *Psylliodes* species and their host plants

The divergence time of Galerucinae and Alticini recovered in our analyses are largely consistent with previous studies on these groups (Ge et al., 2011; Gómez-Zurita et al., 2007), whereas other studies yielded older ages for Galerucinae (Nie et al., 2021; Wang et al., 2014). However, in the latter works, Galerucinae and Alticini were represented by only a few taxa, making their age estimates less robust.

The emergence of *Psylliodes* at ca. 24 Ma and the major radiation on Brassicaceae at ca. 16 Ma occurred after the proposed radiation of Brassicaceae in the Northern hemisphere around 32 Ma (Edger et al., 2015; Huang et al., 2016). Brassicaceae probably originated in a warm and humid climate in the Eocene, but experienced their largest radiation as a cool-adapted family after the Eocene-Oligocene transition. This radiation was quickly tracked by *Ceutorhynchus* weevils (Letsch et al., 2018) and Pierinae butterflies (Edger et al., 2015), which both diversified on Brassicaceae in the Early Oligocene ~32 Ma. In contrast, the diversification of *Psylliodes* on Brassicaceae in the Miocene started considerably later than in other specialists.

28

Solanaceae most likely originated in South America and independently colonized Eurasia, Africa, and Australia several times (Dupin et al., 2017). In Eurasia, an early radiation of the Hyoscyameae tribe occurred at about 7 – 19 Ma in the Middle Miocene (Tu et al., 2010). This overlaps with the time estimations of the major radiation on Solanaceae in *Psylliodes* (clade B), which contemporaneously emerged in Europe and probably dispersed to Asia and Australia in the Pliocene ~4 Ma. This later radiation is again congruent with the dispersal of *Solanum* host plants in Eurasia and Australia (Dupin et al., 2017). A similar pattern was recovered for the radiation of *Psylliodes* on Fagaceae. Within Fagaceae, *Psylliodes* species (clade C) are mainly associated with oaks (*Quercus*). Recent genomic studies on the phylogeny and diversification of this large genus recovered a significant radiation of Eurasian oak species in the mid-to-late Miocene, after the Mid-Miocene climatic optimum (Hipp et al., 2020; Zhou et al., 2022). This radiation of *Quercus* species was quickly tracked by *Psylliodes* species feeding on them. The radiations of *Psylliodes* on both Fagaceae and Solanaceae therefore occurred almost simultaneously with that of their host plants.

Molecular evolution of *GSS* in *Psylliodes*

Previous research demonstrated that *P. chrysocephala* combines a conserved detoxification pathway with glucosinolate-specific detoxification strategies to overcome the characteristic defense of Brassicaceae. The latter includes the ability to convert glucosinolates to desulfo-glucosinolates, which are not substrates of plant myrosinases (Ettlinger et al., 1961). In quantitative feeding experiments, this pathway accounted for the metabolic fate of up to 25% of the total ingested glucosinolates, suggesting a less prominent role in overcoming this defense in the adult life stage (Ahn et al., 2019; Beran et al., 2018). In this study, we discovered an additional GSS in a transcriptome of *P. chrysocephala* larvae and pupae that showed much broader substrate specificity in activity assays compared to the previously identified GSS enzymes in this beetle (Figure 3). Whether there are indeed differences in the expression of *GSS* genes between larval and adult life stages, and consequently differences in the detoxification capacity, is an important question for future research. In addition, genetic manipulation of the GSS detoxification capacity in larvae and adults will help to understand the importance of this detoxification pathway for beetle performance and fitness.

Here we used comparative transcriptome analyses in combination with biochemical studies to investigate the molecular evolution of *GSS* genes in the genus *Psylliodes*. The main question we aimed to answer is when the GSS-based detoxification mechanism emerged in the evolutionary history of *Psylliodes*. In this study, *GSS* genes were only found in *Psylliodes* species that belong to clade A and are specialized on Brassicaceae, suggesting that the evolution of GSS genes coincides with this radiation of *Psylliodes* on

29

Brassicaceae. Our Notung analysis predicted several duplication events of the ancestral *Sulf4* gene in a common ancestor of *P. chrysocephala*, *P. laticollis*, and *P. hospes*, followed by subsequent species-specific gene duplication and loss events (Supplementary Figure S9). Because these three *Psylliodes* species belong to different well-supported sub-clades within clade A, this result indicates that GSS arose early during the radiation of specialized *Psylliodes* species.

The hypothesis that *GSS* evolved in *Psylliodes* clade A is also supported by the lack of functional GSS in *P. kiesenwetteri*, a species that also feeds on Brassicaceae, but belongs to clade D. Interestingly, *P. kiesenwetteri* adults expressed two pseudogenes that, like GSS, are also derived from *Sulf4*, but our analysis of the current GSS dataset suggests that these pseudogenes in *P. kiesenwetteri* and GSS originate from independent duplication events in clades A and D (Supplementary Figure S9). However, adults may not express all *GSS* genes encoded in the genome, so our current GSS data set may not be complete. Comparative genome analyses would therefore be the best strategy to a) determine whether GSS represent a derived trait in *Psylliodes* and b) predict the gene birth and death dynamics in this gene family. To date, the only *Psylliodes* species for which a genome assembly has recently become available is *P. chrysocephala* (King et al., 2023).

The diversification of GSS genes in *Psylliodes* spp. resulted in enzymes with different substrate specificities, which may correspond to the divergent glucosinolate profiles in the host plants of the respective beetle species. One example supporting this hypothesis is the substrate preference of PlatGSS2. This enzyme preferred 2-phenylethyl glucosinolate as a substrate in enzyme assays, the dominant glucosinolate in *Nasturtium* spp., the main host plants of *P. laticollis* (Jeon et al., 2017; Rheinheimer & Hassler, 2018). Another example is the broader substrate specificity of *Phos*GSS2, which is consistent with the broader food plant spectrum of *P. hospes* within Brassicaceae (Gikonyo et al. 2019). On the other hand, the substrate preferences of other GSSs cannot be directly linked to the known host use patterns of *Psylliodes* species, so their functional importance remains unclear.

Results obtained using branch-site model tests were consistent with positive selection acting on *GSS* both early during the diversification of *GSS* (branches 10, 11 and 16), and specifically on *PchrGSS2* (branch 12) (Figure 3, Supplementary Table S13). Four out of five sites with signatures of positive selection along these branches are located in the active site and could thus directly influence either substrate specificity or catalytic activity (Figure S11, Supplementary Table S12). One of these sites is a conserved histidine residue in sulfatase signature motif II (position 125), known to be involved in sulfate ester cleavage. A substitution in this position alone is known to have a drastic negative effect on catalytic activity. However, despite a

substitution in this position (H125N), PlatGSS2 and PchrGSS6 showed significant GSS activity under our assay conditions, which demonstrates that this substitution is not generally associated with weak enzyme activity. Other sites with signatures of positive selection along branches 10 (*PchrGSS3*, *PchrGSS4*, *PchrGSS5*) and 12 (*PchrGSS2*) are located outside of the active site, but may nevertheless influence catalytic activity by affecting overall protein conformation or stability (Supplementary Figure S11). To obtain more direct evidence for a possible role of positive selection on these sites, their functional importance should be investigated by site-directed mutagenesis.

The results of the branch-site model tests are also largely consistent with the ω ratios estimated for *Sulf4* and *GSS* branches, respectively, which suggested that *GSS* are generally under more relaxed constraints compared to the conserved *Sulf4* genes (Supplementary Figure S10). Although together these results generally support the hypothesis that the evolution of GSS has played an important role in the course of adaptation to Brassicaceae hosts in *Psylliodes*, more research is needed to understand the relationship between the molecular evolution of GSS and host use of *Psylliodes*.

Taken together, the results from our ancestral host reconstruction, predicting an ancestral association of *Psylliodes* flea beetles with Brassicaceae, and the putative emergence of GSS in clade A suggest that early *Psylliodes* species used other strategies to cope with the glucosinolate-myrosinase defense system (Figure 4). The most likely strategy is the detoxification of isothiocyanates via the conserved mercapturic acid pathway, which still plays an important role in *P. chrysocephala* (Beran et al., 2018). This pathway is considered to be a basal detoxification strategy for isothiocyanates, but is also associated with metabolic costs that can have a negative impact on insect performance (Gloss et al., 2014; Jeschke et al. 2016a, 2016b, 2017). This might explain why additional and presumably less costly detoxification mechanisms have arisen in specialized *Psylliodes* species.

32

Literature cited

Ahn, S.-J., Betzin, F., Gikonyo, M. W., Yang, Z.-L., Köllner, T. G., & Beran, F. (2019). Identification and evolution of glucosinolate sulfatases in a specialist flea beetle. *Scientific Reports*, *9*(1), 15725. https://doi.org/10.1038/s41598-019-51749-x

Almagro Armenteros, J. J., Sonderby, C. K., Sonderby, S. K., Nielsen, H., & Winther, O. (2017). DeepLoc: prediction of protein subcellular localization using deep learning. *Bioinformatics*, *33*(21), 3387-3395. https://doi.org/10.1093/bioinformatics/btx431

Baele, G., Li, W. L., Drummond, A. J., Suchard, M. A., & Lemey, P. (2013). Accurate model selection of relaxed molecular clocks in bayesian phylogenetics. *Molecular Biology and Evolution*, *30*(2), 239-243. https://doi.org/10.1093/molbev/mss243

Becerra, J. X. (1997). Insects on plants: macroevolutionary chemical trends in host use. *Science*, *276*(5310), 253-256. https://doi.org/10.1126/science.276.5310.253

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B-Statistical Methodology*, *57*(1), 289-300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Beran, F., Pauchet, Y., Kunert, G., Reichelt, M., Wielsch, N., Vogel, H., Reinecke, A., Svatoš, A., Mewis, I., Schmid, D., Ramasamy, S., Ulrichs, C., Hansson, B. S., Gershenzon, J., & Heckel, D. G. (2014). *Phyllotreta striolata* flea beetles use host plant defense compounds to create their own glucosinolate-myrosinase system. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(20), 7349-7354. https://doi.org/10.1073/pnas.1321781111

Beran, F., & Petschenka, G. (2022). Sequestration of plant defense compounds by insects: from mechanisms to insect plant coevolution. *Annual Review of Entomology*, *67*, 163-180. https://doi.org/10.1146/annurev-ento-062821-062319

Beran, F., Sporer, T., Paetz, C., Ahn, S. J., Betzin, F., Kunert, G., Shekhov, A., Vassão, D. G., Bartram, S., Lorenz, S., & Reichelt, M. (2018). One pathway is not enough: the cabbage stem flea beetle *Psylliodes chrysocephala* uses multiple strategies to overcome the glucosinolate-myrosinase defense in its host plants. *Frontiers in Plant Science*, *9*, 1754. https://doi.org/10.3389/fpls.2018.01754

Berenbaum, M. R. (2001). Chemical mediation of coevolution: phylogenetic evidence for Apiaceae and associates. *Annals of the Missouri Botanical Garden*, *88*(1), 45-59. https://doi.org/10.2307/2666131

34

Blažević, I., Montaut, S., Burčul, F., Olsen, C. E., Burow, M., Rollin, P., & Agerbirk, N. (2020). Glucosinolate structural diversity, identification, chemical synthesis and metabolism in plants. *Phytochemistry*, *169*, 112100. https://doi.org/10.1016/j.phytochem.2019.112100

Borges, R., Machado, J. P., Gomes, C., Rocha, A. P., & Antunes, A. (2019). Measuring phylogenetic signal between categorical traits and phylogenies. *Bioinformatics*, *35*(11), 1862-1869. https://doi.org/10.1093/bioinformatics/bty800

Braga, M. P., Landis, M. J., Nylin, S., Janz, N., & Ronquist, F. (2020). Bayesian inference of ancestral host-parasite interactions under a phylogenetic model of host repertoire evolution. *Systematic Biology*, *69*(6), 1149-1162. https://doi.org/10.1093/sysbio/syaa019

Burlacu, E., Nisca, A., & Tanase, C. (2020). A comprehensive review of phytochemistry and biological activities of *Quercus* species. *Forests*, *11*(9). https://doi.org/10.3390/f11090904

Chatzimanolis, S., Grimaldi, D. A., Engel, M. S., & Fraser, N. C. (2012). *Leehermania prorova,* the earliest Staphyliniform beetle, from the late Triassic of Virginia (Coleoptera: Staphylinidae). *American Museum Novitates*, *3761*, 1-28. https://doi.org/10.1206/3761.2

Chowański, S., Adamski, Z., Marciniak, P., Rosiński, G., Büyükgüzel, E., Büyükgüzel, K., Falabella, P., Scrano, L., Ventrella, E., Lelario, F., & Bufo, S. A. (2016). A review of bioinsecticidal activity of Solanaceae alkaloids. *Toxins (Basel)*, *8*(3), 60. https://doi.org/10.3390/toxins8030060

Dobler, S., Mardulyn, P., Pasteels, J. M., & Rowell-Rahier, M. (1996). Host-plant switches and the evolution of chemical defense and life history in the leaf beetle genus *Oreina*. *Evolution*, *50*(6), 2373-2386. https://doi.org/10.1111/j.1558-5646.1996.tb03625.x

Doguet, S. (1994). *Coléoptère Chrysomelidae. Volume 2. Alticinae. Faune de France 80.* Federation Francaise des Societes de Sciences Naturelles.

Douglas, H. B., Konstantinov, A. S., Brunke, A. J., Moseyko, A. G., Chapados, J. T., Eyres, J., Richter, R., Savard, K., Sears, E., Prathapan, K. D., Ruan, Y. Y., & Dettman, J. R. (2023). Phylogeny of the flea beetles (Galerucinae: Alticini) and the position of *Aulacothorax* elucidated through anchored phylogenomics (Coleoptera: Chrysomelidae: Alticini). *Systematic Entomology*, *48*(3), 361-386. https://doi.org/10.1111/syen.12582

Drummond, A. J., Ho, S. Y., Phillips, M. J., & Rambaut, A. (2006). Relaxed phylogenetics and dating with confidence. *PLoS Biology*, *4*(5), e88. https://doi.org/10.1371/journal.pbio.0040088

Drummond, A. J., Suchard, M. A., Xie, D., & Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, *29*(8), 1969-1973. ttps://doi.org/10.1093/molbev/mss075

Dupin, J., Matzke, N. J., Särkinen, T., Knapp, S., Olmstead, R. G., Bohs, L., & Smith, S. D. (2017). Bayesian estimation of the global biogeographical history of the Solanacea. *Journal of Biogeography*, *44*(4), 887-899. https://doi.org/10.1111/jbi.12898

Durand, D., Halldórsson, B. V., & Vernot, B. (2006). A hybrid micro-macroevolutionary approach to gene tree reconstruction. *Journal of Computational Biology*, *13*(2), 320-335. https://doi.org/10.1089/cmb.2006.13.320

Edgar, R. C. (2004). MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, *32*(5), 1792-1797. https://doi.org/10.1093/nar/gkh340

Edger, P. P., Heidel-Fischer, H. M., Bekaert, M., Rota, J., Gloeckner, G., Platts, A. E., Heckel, D. G., Der, J. P., Wafula, E. K., Tang, M., Hofberger, J. A., Smithson, A., Hall, J. C., Blanchette, M., Bureau, T. E., Wright, S. I., dePamphilis, C. W., Schranz, M. E., Barker, M. S., Conant, G. C., Wahlberg, N., Vogel, H., Pires, J. C., & Wheat, C. W. (2015). The butterfly plant arms-race escalated by gene and genome duplications. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(27), 8362-8366. https://doi.org/10.1073/pnas.1503926112

Ehrlich, P. R., & Raven, P. H. (1964). Butterflies and plants: a study in coevolution. *Evolution*, *18*(4), 586-608. https://doi.org/10.1111/j.1558-5646.1964.tb01674.x

Ettlinger, M. G., Thompson, C. P., Dateo, G. P., Mabry, T. J., & Harrison, B. W. (1961). Vitamin C as a coenzyme: the hydrolysis of mustard oil glucosides. *Proceedings of the National Academy of Sciences of the United States of America*, *47*(12), 1875-1880. https://doi.org/10.1073/pnas.47.12.1875

Farrell, B., & Mitter, C. (1990). Phylogenesis of insect/plant interactions: have *Phyllobrotica* leaf beetles (Chrysomelidae) and the Lamiales diversified in parallel? *Evolution*, *44*(6), 1389-1403. https://doi.org/10.1111/j.1558-5646.1990.tb03834.x

Forister, M. L., Novotny, V., Panorska, A. K., Baje, L., Basset, Y., Butterill, P. T., Cizek, L., Coley, P. D., Dem, F., Diniz, I. R., Drozd, P., Fox, M., Glassmire, A. E., Hazen, R., Hrcek, J., Jahner, J. P., Kaman, O., Kozubowski, T. J., Kursar, T. A., Lewis, O. T., Lill, J., Marquis, R. J., Miller, S. E., Morais, H. C., Murakami, M., Nickel, H., Pardikes, N. A., Ricklefs, R. E., Singer, M. S., Smilanich, A. M., Stireman, J. O., Villamarín-Cortez, S., Vodka, S., Volf, M., Wagner, D. L., Walla, T., Weiblen, G. D., & Dyer, L. A. (2015). The global distribution of diet breadth in insect herbivores. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(2), 442-447. https://doi.org/10.1073/pnas.1423042112

36

Friedrichs, J., Schweiger, R., Geisler, S., Mix, A., Wittstock, U., & Müller, C. (2020). Novel glucosinolate metabolism in larvae of the leaf beetle *Phaedon cochleariae*. *Insect Biochemistry and Molecular Biology*, *124*. https://doi.org/10.1016/j.ibmb.2020.103431

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., & Bairoch, A. (2005). Protein identification and analysis tools on the ExPASy server. In M. J. Walker (Ed.), *The Proteomics Protocols Handbook* (pp. 571-607). Humana Press. https://doi.org/10.1385/1-59259-890-0:571

Ge, D. Y., Chesters, D., Gómez-Zurita, J., Zhang, L. J., Yang, X. K., & Vogler, A. P. (2011). Anti-predator defence drives parallel morphological evolution in flea beetles. *Proceedings of the Royal Society B-Biological Sciences*, *278*(1715), 2133-2141. https://doi.org/10.1098/rspb.2010.1500

Gikonyo, M. W., Biondi, M., & Beran, F. (2019). Adaptation of flea beetles to Brassicaceae: host plant associations and geographic distribution of *Psylliodes* Latreille and *Phyllotreta* Chevrolat (Coleoptera, Chrysomelidae). In: M. Schmitt, C. S. Chaboo, & M. Biondi (Eds.), *Research on Chrysomelidae 8*. *Zookeys*, *856*, 51-73. https://doi.org/10.3897/zookeys.856.33724

Gloss, A. D., Vassão, D. G., Hailey, A. L., Nelson Dittrich, A. C., Schramm, K., Reichelt, M., Rast, T. J., Weichsel, A., Cravens, M. G., Gershenzon, J., Montfort, W. R., & Whiteman, N. K. (2014). Evolution in an ancient detoxification pathway is coupled with a transition to herbivory in the Drosophilidae. *Molecular Biology and Evolution*, *31*(9), 2441-2456. https://doi.org/10.1093/molbev/msu201

Gök, A., & Aslan, E. G. (2006). Species composition and abundance of flea beetles (Coleoptera: Chrysomelidae) associated with moist habitats in Isparta and Burdur provinces, Turkey. *Proceedings of the Entomological Society of Washington*, *108*(3), 543-549.

Gómez-Zurita, J., Hunt, T., Kopliku, F., & Vogler, A. P. (2007). Recalibrated tree of leaf beetles (Chrysomelidae) indicates independent diversification of angiosperms and their insect herbivores. *Plos One*, *2*(4). https://doi.org/10.1371/journal.pone.0000360

Graser, G., Oldham, N. J., Brown, P. D., Temp, U., & Gershenzon, J. (2001). The biosynthesis of benzoic acid glucosinolate esters in *Arabidopsis thaliana*. *Phytochemistry*, *57*(1), 23-32. https://doi.org/10.1016/S0031-9422(00)00501-X

Gupta, R., & Brunak, S. (2002). Prediction of glycosylation across the human proteome and the correlation to protein function. *Pacific Symposium on Biocomputing*, *7*, 310-322. https://doi.org/10.1142/9789812799623_0029

Heckel, D. G. (2014). Insect detoxification and sequestration strategies. In C. Voelckel & G. Jander (Eds.), *Annual Plant Reviews: Insect-Plant Interactions* (Vol 47 pp. 77-114). Wiley. https://doi.org/10.1002/9781118829783.ch3

Hipp, A. L., Manos, P. S., Hahn, M., Avishai, M., Bodénès, C., Cavender-Bares, J., Crowl, A. A., Deng, M., Denk, T., Fitz-Gibbon, S., Gailing, O., Gonzalez-Elizondo, M. S., Gonzalez-Rodriguez, A., Grimm, G. W., Jiang, X. L., Kremer, A., Lesur, I., McVay, J. D., Plomion, C., . . . Valencia-Avalos, S. (2020). Genomic landscape of the global oak phylogeny. *New Phytologist*, *226*(4), 1198-1212. https://doi.org/10.1111/nph.16162

Hoang, D. T., Chernomor, O., Von Haeseler, A., Minh, B. Q., & Vinh, L. S. (2018). UFBoot2: Improving the ultrafast bootstrap approximation. *Molecular Biology and Evolution*, *35*(2), 518-522. https://doi.org/10.1093/molbev/msx281

Höhna, S., Landis, M. J., Heath, T. A., Boussau, B., Lartillot, N., Moore, B. R., Huelsenbeck, J. P., & Ronquist, F. (2016). RevBayes: bayesian phylogenetic inference using graphical models and an interactive model-specification language. *Systematic Biology*, *65*(4), 726-736. https://doi.org/10.1093/sysbio/syw021

Huang, C. H., Sun, R., Hu, Y., Zeng, L., Zhang, N., Cai, L., Zhang, Q., Koch, M. A., Al-Shehbaz, I., Edger, P. P., Pires, J. C., Tan, D. Y., Zhong, Y., & Ma, H. (2016). Resolution of Brassicaceae phylogeny using nuclear genes uncovers nested radiations and supports convergent morphological evolution. *Molecular Biology and Evolution*, *33*(2), 394-412. https://doi.org/10.1093/molbev/msv226

Jeon, J., Bong, S. J., Park, J. S., Park, Y. K., Arasu, M. V., Al-Dhabi, N. A., & Park, S. U. (2017). De novo transcriptome analysis and glucosinolate profiling in watercress (*Nasturtium officinale* R. Br.). *BMC Genomics*, *18*(1), 401. https://doi.org/10.1186/s12864-017-3792-5

Jeschke, V., Gershenzon, J., & Vassão, D. G. (2016a). Insect detoxification of glucosinolates and their hydrolysis products. In S. Kopriva (Ed.), *Advances in Botanical Research - Glucosinolates* (Vol. 80, pp. 199-245). Elsevier. https://doi.org/10.1016/bs.abr.2016.06.003.

Jeschke, V., Gershenzon, J., & Vassão, D. G. (2016b). A mode of action of glucosinolate-derived isothiocyanates: detoxification depletes glutathione and cysteine levels with ramifications on protein metabolism in *Spodoptera littoralis*. *Insect Biochemistry and Molecular Biology*, *71*, 37-48. https://doi.org/10.1016/j.ibmb.2016.02.002.

Jeschke, V., Kearney, E. E., Schramm, K., Kunert, G., Shekhov, A., Gershenzon, J., & Vassão, D. G. (2017). How glucosinolates affect generalist lepidopteran larvae: growth, development and glucosinolate metabolism. *Frontiers in Plant Science*, *8*. https://doi.org/10.3389/fpls.2017.01995

Jolivet, P. H., & Hawkeswood, T. J. (1995). *Host-plants of Chrysomelidae of the world*. Backhuys Publishers.

Käll, L., Krogh, A., & Sonnhammer, E. L. (2007). Advantages of combined transmembrane topology and signal peptide prediction – the Phobius web server. *Nucleic Acids Research*, *35*(suppl_2), W429-W432. https://doi.org/10.1093/nar/gkm256

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., & Jermiin, L. S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods*, *14*(6), 587-589. https://doi.org/10.1038/nmeth.4285

Kapustin, Y., Souvorov, A., Tatusova, T., & Lipman, D. (2008). Splign: algorithms for computing spliced alignments with identification of paralogs. *Biology Direct*, *3*, 1-13. https://doi.org/10.1186/1745-6150-3-20

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*(430), 773-795. https://doi.org/10.1080/01621459.1995.10476572

Katoh, K., Rozewicki, J., & Yamada, K. D. (2019). MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Briefings in Bioinformatics*, *20*(4), 1160-1166. https://doi.org/10.1093/bib/bbx108

King, R., Buer, B., Davies, T. G. E., Ganko, E., Guest, M., Hassani-Pak, K., Hughes, D., Raming, K., Rawlings, C., Williamson, M., Crossthwaite, A., Nauen, R., & Field, L. (2023). The complete genome assemblies of 19 insect pests of worldwide importance to agriculture. *Pesticide Biochemistry and Physiology*, *191*. https://doi.org/10.1016/j.pestbp.2023.105339

Kirejtshuk, A. G., Poschmann, M., Prokop, J., Garrouste, R., & Nel, A. (2014). Evolution of the elytral venation and structural adaptations in the oldest Palaeozoic beetles (Insecta: Coleoptera: Tshekardocoleidae). *Journal of Systematic Palaeontology*, *12*(5), 575-600. https://doi.org/10.1080/14772019.2013.821530

Köpf, A., Rank, N. E., Roininen, H., Julkunen-Tiitto, R., Pasteels, J. M., & Tahvanainen, J. (1998). The evolution of host-plant use and sequestration in the leaf beetle genus *Phratora* (Coleoptera: Chrysomelidae). *Evolution*, *52*(2), 517-528. https://doi.org/10.1111/j.1558-5646.1998.tb01651.x

Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution*, *35*(6), 1547-1549. https://doi.org/10.1093/molbev/msy096

Leonardi, C. (1970). Materiali per uno studio filogenetico del genere *Psylliodes* (Coleoptera Chrysomelidae). *Atti della Società italiana di scienze naturali e del Museo civico di storia naturale di Milano*, *110*(3), 201-223.

Letsch, H., Gottsberger, B., Metzl, C., Astrin, J., Friedman, A. L. L., McKenna, D. D., & Fiedler, K. (2018). Climate and host-plant associations shaped the evolution of ceutorhynch weevils throughout the Cenozoic. *Evolution*, *72*(9), 1815-1828. https://doi.org/10.1111/evo.13520

Magallón, S., Gómez-Acevedo, S., Sánchez-Reyes, L. L., & Hernández-Hernández, T. (2015). A metacalibrated time-tree documents the early rise of flowering plant phylogenetic diversity. *New Phytologist*, *207*(2), 437-453. https://doi.org/10.1111/nph.13264

Malka, O., Easson, M. L. A. E., Paetz, C., Götz, M., Reichelt, M., Stein, B., Luck, K., Stanišić, A., Juravel, K., Santos-Garcia, D., Mondaca, L. L., Springate, S., Colvin, J., Winter, S., Gershenzon, J., Morin, S., & Vassao, D. G. (2020). Glucosylation prevents plant defense activation in phloem-feeding insects. *Nature Chemical Biology*, *16*(12), 1420-1426. https://doi.org/10.1038/s41589-020-00658-6

Mardulyn, P., Milinkovitch, M. C., & Pasteels, J. M. (1997). Phylogenetic analyses of DNA and allozyme data suggest that *Gonioctena* leaf beetles (Coleoptera; Chrysomelidae) experienced convergent evolution in their history of host-plant family shifts. *Systematic Biology*, *46*(4), 722-747. https://doi.org/10.1093/sysbio/46.4.722

Meade, A., & Pagel, M. (2022). Ancestral state reconstruction using BayesTraits. *Methods in Molecular Biology*, *2569*, 255-266. https://doi.org/10.1007/978-1-0716-2691-7_12

Miller, M. A., Pfeiffer, W., & Schwartz, T. (2010). *Creating the CIPRES Science Gateway for inference of large phylogenetic trees*. 2010 Gateway Computing Environments Workshop, GCE, New Orleans, LA, USA. https://doi.org/10.1109/GCE.2010.5676129

Mitchell, R. (1981). Insect behavior, resource exploitation, and fitness. *Annual Review of Entomology*, *26*(1), 373-396. https://doi.org/10.1146/annurev.en.26.010181.002105

Nadein, K. S. (2006). A significance of the tegmen structure for classification of the genus *Psylliodes* Latreille, 1829 (Coleoptera: Chrysomelidae: Psylliodina). *Proceedings of the Russian Entomological Society. St. Petersburg*, *77*, 250-254.

Nadein, K. S. (2007a). On the taxonomy and classification of the genus *Psylliodes* Latreille, 1825 (Coleoptera, Chrysomelidae, Galerucinae). *Entomologica Basiliensia et Collectionis Frey*, *29*, 307-332.

Nadein, K. S. (2007b). A review of the leaf-beetle genus *Psylliodes* Latreille (Coleoptera, Chrysomelidae) from Russia and neighboring countries: I. A key to subgenera, species-groups, and species. *Entomological Review*, *87*(3), 330-360. https://doi.org/10.1134/S0013873807030086

Nadein, K. S., & Lee, C.-F. (2012). New data about some Alticini from Taiwan with descriptions of two new species (Coleoptera: Chrysomelidae). *Bonn zoological Bulletin*, *61*(1), 41-48.

Nguyen, L. T., Schmidt, H. A., von Haeseler, A., & Minh, B. Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating Maximum-Likelihood phylogenies. *Molecular Biology and Evolution*, *32*(1), 268-274. https://doi.org/10.1093/molbev/msu300

Niculaes, C., Abramov, A., Hannemann, L., & Frey, M. (2018). Plant protection by benzoxazinoids - recent insights into biosynthesis and function. *Agronomy*, *8*(8). https://doi.org/10.3390/agronomy8080143

Nie, R., Vogler, A. P., Yang, X. K., & Lin, M. Y. (2021). Higher-level phylogeny of longhorn beetles (Coleoptera: Chrysomeloidea) inferred from mitochondrial genomes. *Systematic Entomology*, *46*(1), 56-70. https://doi.org/10.1111/syen.12447

Nylin, S., Agosta, S., Bensch, S., Boeger, W. A., Braga, M. P., Brooks, D. R., Forister, M. L., Hamback, P. A., Hoberg, E. P., Nyman, T., Schapers, A., Stigall, A. L., Wheat, C. W., Osterling, M., & Janz, N. (2018). Embracing colonizations: A new paradigm for species association dynamics. *Trends in Ecology and Evolution*, *33*(1), 4-14. https://doi.org/10.1016/j.tree.2017.10.005

Pagel, M., & Meade, A. (2006). Bayesian analysis of correlated evolution of discrete characters by reversible-jump Markov chain Monte Carlo. *The American Naturalist*, *167*(6), 808-825. https://doi.org/10.1086/503444

Paradis, E., & Schliep, K. (2019). ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, *35*(3), 526-528. https://doi.org/10.1093/bioinformatics/bty633

Pasteels, J. M., & Rowell-Rahier, M. (1991). Proximate and ultimate causes for host plant influence on chemical defense of leaf beetles (Coleoptera, Chrysomelidae). *Entomologia Generalis*, *15*(4), 227-235. https://doi.org/10.1127/entom.gen/15/1991/227

Petersen, T. N., Brunak, S., von Heijne, G., & Nielsen, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature Methods*, *8*(10), 785-786. https://doi.org/10.1038/nmeth.1701

R Development Core Team. (2022). *R: A Language and Environment for Statistical Computing.* In R Foundation for Statistical Computing. http://www.r-project.org/

Rambaut, A., Drummond, A. J., Xie, D., Baele, G., & Suchard, M. A. (2018). Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology*, *67*(5), 901-904. https://doi.org/10.1093/sysbio/syy032

Rheinheimer, J., & Hassler, M. (2018). *Die Blattkäfer Baden-Württembergs*. Kleinsteuber Verlag.

Robinson, G. S., Ackery, P. R., Kitching, I., Beccaloni, G. W., & Hernández, L. M. (2023). *HOSTS - a Database of the World's Lepidopteran Hostplants*. [Data set]. Natural History Museum. https://doi.org/10.5519/havt50xw

Salvi, D., D'Alessandro, P., & Biondi, M. (2019). Host plant associations in Western Palaearctic *Longitarsus* flea beetles (Chrysomelidae, Galerucinae, Alticini): a preliminary phylogenetic assessment. *Zookeys*(856). https://doi.org/10.3897/zookeys.856.32430

Schwede, T., Kopp, J., Guex, N., & Peitsch, M. C. (2003). SWISS-MODEL: An automated protein homology-modeling server. *Nucleic Acids Research*, *31*(13), 3381-3385. https://doi.org/10.1093/nar/gkg520

Seppey, M., Manni, M., & Zdobnov, E. M. (2019). BUSCO: assessing genome assembly and annotation completeness. *Methods in Molecular Biology*, *1962*, 227-245. https://doi.org/10.1007/978-1-4939-9173-0_14

Sigrist, C. J., Cerutti, L., Hulo, N., Gattiker, A., Falquet, L., Pagni, M., Bairoch, A., & Bucher, P. (2002). PROSITE: A documented database using patterns and profiles as motif descriptors. *Briefings in Bioinformatics*, *3*(3), 265-274. https://doi.org/10.1093/bib/3.3.265

Suchard, M. A., Lemey, P., Baele, G., Ayres, D. L., Drummond, A. J., & Rambaut, A. (2018). Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evolution*, *4*(1), vey016. https://doi.org/10.1093/ve/vey016

Thies, W. (1988). Isolation of sinigrin and glucotropaeolin from cruciferous seeds. *European Journal of Lipid Science and Technology*, *90*, 311–314. https://doi.org/10.1002/lipi.19880900806

Tu, T., Volis, S., Dillon, M. O., Sun, H., & Wen, J. (2010). Dispersals of Hyoscyameae and Mandragoreae (Solanaceae) from the New World to Eurasia in the early Miocene and their biogeographic diversification within Eurasia. *Molecular Biology and Evolution*, *57*(3), 1226-1237. https://doi.org/10.1016/j.ympev.2010.09.007

Wang, B., Ma, J. Y., McKenna, D. D., Yan, E. V., Zhang, H. C., & Jarzembowski, E. A. (2014). The earliest known longhorn beetle (Cerambycidae: Prioninae) and implications for the early evolution of Chrysomeloidea. *Journal of Systematic Palaeontology*, *12*(5), 565-574. https://doi.org/10.1080/14772019.2013.806602

Warchałowski, A. (2010). *The Palaearctic Chrysomelidae. Identification Keys* (D. lwan, Ed. Vol. 2). Natura optima dux Foundation.

Wolfe, J. M., Daley, A. C., Legg, D. A., & Edgecombe, G. D. (2016). Fossil calibrations for the Arthropod tree of life. *Earth-Science Reviews*, *160*, 43-110. https://doi.org/10.1016/j.earscirev.2016.06.008

Xia, X. (2018). DAMBE7: new and improved tools for data analysis in molecular biology and evolution. *Molecular Biology and Evolution*, *35*(6), 1550-1552. https://doi.org/10.1093/molbev/msy073

Xia, X., & Lemey, P. (2009). Assessing Substitution Saturation with DAMBE. In P. Lemey, M. Salemi, & A. Vandamme (Eds.), *The phylogenetic handbook: A practical approach to phylogenetic analysis and hypothesis testing* (pp. 615-630). Cambridge University Press.

Xia, X., Xie, Z., Salemi, M., Chen, L., & Wang, Y. (2003). An index of substitution saturation and its application. *Molecular Phylogenetics and Evolution*, *26*(1), 1-7. https://doi.org/10.1016/S1055-7903(02)00326-3

Yang, Z. (2007). PAML 4: phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution*, *24*(8), 1586-1591. https://doi.org/10.1093/molbev/msm088

Zhang, H., Song, N., & Yin, X. (2022). Higher-level phylogeny of Chrysomelidae based on expanded sampling of mitogenomes. *PLoS One*, *17*(1), e0258587. https://doi.org/10.1371/journal.pone.0258587

Zhang, J., Nielsen, R., & Yang, Z. (2005). Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. *Molecular Biology and Evolution*, *22*(12), 2472-2479. https://doi.org/10.1093/molbev/msi237

Zhou, B. F., Yuan, S., Crowl, A. A., Liang, Y. Y., Shi, Y., Chen, X. Y., An, Q. Q., Kang, M., Manos, P. S., & Wang, B. S. (2022). Phylogenomic analyses highlight innovation and introgression in the continental radiations of Fagaceae across the Northern Hemisphere. *Nature Communications*, *13*(1). https://doi.org/10.1038/s41467-022-28917-1