# What goes in must come out? The metabolic profile of plants and caterpillars, frass, and adults of asota (erebidae: aganainae) feeding on ficus (moraceae) in New Guinea

by SEGAR, S., FONTANILLA, A.M., AUBONA, G., SISOL, M., KUUKKANEN, I., SALMINEN, J.P., MILLER, S.E., HOLLOWAY, J.D., NOVOTNY, V. and VOLF, M.

# WHAT GOES IN MUST COME OUT? THE METABOLIC PROFILE OF PLANTS AND CATERPILLARS, FRASS, AND ADULTS OF *ASOTA* (EREBIDAE: AGANAINAE) FEEDING ON *FICUS* (MORACEAE) IN NEW GUINEA

ALYSSA M. FONTANILLA[1,2 ♀], GIBSON AUBONA[3 ♀], MENTAP SISOL[3], ILARI KUUKKANEN[4], JUHA-PEKKA SALMINEN[4], SCOTT E. MILLER[5], JEREMY D. HOLLOWAY[6], VOJTECH NOVOTNY[1,2], MARTIN VOLF[1], AND SIMON T. SEGAR[7*]

[1]*Biology Centre, Institute of Entomology, the Czech Academy of Sciences, České Budějovice, Czech Republic*

[2]*Faculty of Science, University of South Bohemia, České Budějovice, Czech Republic*

[3]*New Guinea Binatang Research Center, Madang, Papua New Guinea*

[4]*Natural Chemistry Research Group, Department of Chemistry, University of Turku, Turku, Finland*

[5]*National Museum of Natural History, Smithsonian Institution, Washington DC, USA*

[6]*Department of Life Sciences, The Natural History Museum, London, UK*

[7]*Agriculture and Environment Department, Harper Adams University, Newport, UK*

♀*Joint first authors*

Corresponding author:
Simon T. Segar
email: ssegar@harper-adams.ac.uk
A.M.F. ORCID: 0000-0002-4707-4918
J.P.S. ORCID: 0000-0002-2912-7094
S.E.M. ORCID: 0000-0002-4138-1378
V.N. ORCID: 0000-0001-7918-8023
M.V. ORCID: 0000-0003-4126-3897
S.T.S. ORCID: 0000-0001-6621-9409

**Acknowledgements**

We thank the assistants, taxonomists, and the staff of the New Guinea Binatang Research Centre in Papua New Guinea. We also thank the Papua New Guinea Forest Research Institute and the Department of Environment and Conservation for their help in getting plant and insect export permits, and the lab members of the Natural Chemistry Research Group at the University of Turku, Finland for their assistance with chemical analyses. A.M.F., V.N., M.V. and S.T.S. acknowledge the Grant Agency of the Czech Republic (grant number 19-28126X). V.N. acknowledges support from the ERC grant no. 669609. S.T.S. acknowledges departmental support from Harper Adams University. J-P.S. acknowledges funding from the Academy of Finland. This material is based upon work supported by the U.S. National Science Foundation under grants DEB 9707928, 0211591 and 0515678, 0816749 and 0841885. DNA barcoding was provided by the Biodiversity Institute of Ontario, University of Guelph, with funding from Genome Canada and the Ontario Genomics Institute to the International Barcode of Life Project. We dedicate this manuscript to our late colleagues Stewart Wossa, Simon Leather and Freddy Pius.

**Abstract-** Insect herbivores have evolved a broad spectrum of adaptations in response to the diversity of chemical defences employed by plants. Here we focus on two species of New Guinean *Asota* and determine how these specialist moths deal with the leaf alkaloids of their *Ficus* fig hosts. As each focal *Asota* species is restricted to one of three chemically distinct species of *Ficus*, we also test whether these specialised interactions lead to similar alkaloid profiles in both *Asota* species. We reared *Asota* caterpillars on their respective *Ficus* hosts in natural conditions and analysed the alkaloid profiles of leaf, frass, caterpillar, and adult moth samples using UHPLC–MS/MS analyses. We identified 43 alkaloids in our samples. Leaf alkaloids showed various fates. Some were excreted in frass or found in caterpillars and adult moths. We also found two apparently novel indole alkaloids—likely synthesised *de novo* by the moths or their microbiota—in both caterpillar and adult tissue but not in leaves or frass. Overall, alkaloids unique or largely restricted to insect tissue were shared across moth species despite feeding on different hosts. This indicates that a limited number of plant compounds have a direct ecological function that is conserved among the studied species. Our results provide evidence for the importance of phytochemistry and metabolic strategies in the formation of plant–insect interactions and food webs in general. Furthermore, we provide a new potential example of insects acquiring chemicals for their benefit in an ecologically relevant insect genus.

**Key Words-** Alkaloids, plant–insect interactions, food-webs, biodiversity, host-specificity, herbivores.

# INTRODUCTION

To this day, more than 200,000 specialized plant metabolites have been described, many of which have known or assumed anti-herbivore functions (Kessler and Kalske 2018). In response to this diversity of plant chemical defences, herbivorous insects have evolved various physiological adaptations to circumvent and/or exploit the specialized metabolites produced by their host plants (Petschenka and Agrawal 2016; Dussourd 2017).

Adaptations to plant chemical defences by herbivores can be specific to a restricted set of compounds in a given class (Lindigkeit et al. 1997), and are typically conserved among closely related herbivore species (Nallu et al. 2018; Allio et al. 2021). Plant chemistry thus often predicts interactions among plants and their specialized herbivores. In some cases, it can be a better predictor of host-use that the relatedness and phylogeny of the host plants (Becerra 1997; Endara et al. 2017). Nymphalid Melitaeii butterflies, for instance, are chemical specialists that can exploit distantly related but chemically similar plants containing iridoid glycosides (Wahlberg 2001). In turn, host plant chemistry can predict host-use by herbivores among both distantly or closely related plants even in extremely diverse systems, such as tropical forests that harbour myriad insect herbivores and plants. For example, polyphenol oxidative activity in leaves of 88 plant species could predict their use by geometrid moths in hyper-diverse food webs in tropical lowland forests in New Guinea (Segar et al. 2017). Such trends occur even among congeneric hosts, with high elevation populations of New Guinean *Ficus* that possess characteristic alkaloid profiles harbouring caterpillar communities distinct from their lowland relatives (Volf et al. 2020). But while such findings underscore the crucial role of plant chemistry in shaping host preference in insect herbivores, we rarely understand the physiological counteradaptations by herbivores behind host use patterns.

Physiological counteradaptations in insects broadly fall into four groups, which can be characterized as either relying on avoiding prolonged contact with potentially toxic compounds, or contributing to

defence (Trigo 2011) and communication (Wink 2019). Some plant metabolites can be excreted without modification such that potential toxicity is avoided rather than exploited. Furthermore, plant metabolites can be excreted after detoxification via chemical modification (e.g., glycosylation) (Salminen et al. 2004). Targeted comparisons between leaf tissue and insect frass in these cases will reveal either a complete overlap in chemical composition or partial overlap with modifications of some compounds. Insect tissue composition will be largely independent of leaf composition. Insects are, however, not limited to excretion and avoidance. Secondary metabolites can also be sequestered, in such cases chemicals are usually used for defence against predators and parasitoids (e.g., cardenolides by Monarch butterflies) or modified for other uses, such as defence or mating (e.g., pyrrolizidine alkaloids in certain arctiine moths) (Weller et al. 1999; Salminen et al. 2004; Petschenka and Agrawal 2016; Heckel 2018). These last two cases generate their own expectations, but for both examples similarity will exist between plant tissue and insect tissue, while frass will show little chemical overlap with either. Thus, we have a set of expectations on dissimilarity among frass, insect and plant tissue related to feeding ecology.

Aposematic tiger moths (woolly bears and allies) (Erebidae: Arctiinae) have long been the topic of studies on chemical defence strategies in Lepidoptera. These moths have an intimate relationship with their hosts and can exploit pyrrolizidine alkaloids for defence (as larvae) and courtship (as adults) (Weller et al. 1999; Hartmann et al. 2005). Their developmental variation in chemical ecology highlights the longstanding interaction between plant and insect that has minimised costs to the insect (Cogni et al. 2012; Zaspel et al. 2014) and the need to study multiple life stages. The tribe Lithosinii is also intimately connected to its lichen hosts from which it sequesters phenolic compounds (Scott Chialvo et al. 2018). Indeed, Anderson et al. (2017) found that *Crambidia cephalica* had a distinct metabolomic profile from its *Physcia* host despite sequestering lichen secondary metabolites. Furthermore, they observed differences across *C. cephalica* life stages, with the larval and pupal stages having the highest chemical similarity. This body of results highlights how metabolomic approaches can be applied to gain a deeper understanding of host preference and chemistry-based adaptations in insects, especially when sampling across tissue types and life stages.

Here we explore the metabolomes of New Guinean species of *Asota* (Erebidae: Aganainae), an aposematic moth genus closely related to arctiine moths (Zahiri et al. 2012). Unlike the Arctiinae, *Asota* and their chemistry have been surprisingly understudied despite being associated with severe skin irritation and fever outbreaks in humans (Wills et al. 2016). The toxicity of caterpillar and adult *Asota* fluids as well as that of adult scales (Sourakov and Emmel 2001; Wills et al. 2016) are strongly indicative of a physiological adaptation to host plant metabolites. Widescale sampling of lowland and highland food webs in New Guinea revealed that *Asota* is a *Ficus* specialist (Fig. S1) feeding in larger numbers on *Ficus* species with high alkaloid content and diversity (Novotny et al. 2010; Volf et al. 2018). We use the *Ficus–Asota* system to quantify alkaloids from the leaves, frass, caterpillars and adult moths to determine their fates and therefore possible ecological functions. We expect sequestration of

the alkaloids in the aposematic caterpillars and adults. Furthermore, given the conservatism in adaptations to host plant chemical defences among closely related herbivores, we expect convergence of alkaloid profiles as modified by the insects between *Asota* species feeding on different host plants. Through the use of a metabolomic approach, we connect the alkaloids found in *Ficus* and *Asota* to see whether the maxim "you are what you eat" applies also to caterpillars.

## METHODS AND MATERIALS

*Study system.*

We studied two species of *Asota* (*A. eusemioides* and *A. heliconia*) across three species of *Ficus* (*F. septica*, *F. pachyrrhachis* and *F. hispidioides*). Because of the complexity of the taxonomy of *Asota* (Holloway 2022), we became aware that *A. heliconia* was in fact a duplex of two cryptic species that are partially overlapping across the wide range of what was originally considered one species (Holloway 1988, Supplementary Information). However, these two species show segregation with altitude in New Guinea. This enabled us to establish that we were dealing with only the lowland member of the pair in Madang. A more detailed account is presented in Supporting Information. Our field collections also showed that caterpillars from each of the two *Asota* species were largely restricted to only one or two *Ficus* species. We found *A. eusemioides* only on *F. hispidioides* and *F. pachyrrhachis,* while almost all *A. heliconia* caterpillars collected for our experiment were from *F. septica*. Only two *A. heliconia* caterpillars were found on another *Ficus* species (*F. hispidioides)*; they were excluded from our study.

*Field experiment.*

We conducted a two-part experiment in Ohu, Madang Province, Papua New Guinea (−5.140° 145.410° 200 m) from February to April 2018. Because conditions and rearing approach were consistent across the entire study period, we combined experimental data from the first and second parts for analysis so that four individual trees from each *Ficus* species were involved (n=12). While a larger host sample size would have been preferable, replication across insects on the same host plant individual is required in our study. Metabolic analysis of resource and herbivores can only establish compound fate if all resource compounds have equal possibility to be found further down the trophic chain. And as we were mainly interested in how insects vary in their responses to a common resource, maximising variation at the resource level is not desirable. Moreover, designing complex experiments in remote tropical rainforests are often shaped by logistical limitations. Thus, given our research focus and the resources at hand, we decided to maximise replication at the level of the organism being studied (i.e., the insect) and control for repeated measures at all steps.

*Larval development in* Asota. The first part of our experiment was aimed at rearing caterpillars on each host to adulthood; this allowed accurate identification of species and developmental stage, and

confirmation that body size increased predictably across instars (Dyar 1890). We selected two individual trees from each host species (n=6) and attempted to rear five individual first instar caterpillars (collected from the appropriate host species in the surrounding forest) to adulthood (n=30). Individual caterpillars were placed on a single branch of the *Ficus* species from which they were sampled and enclosed in a fine breathable mesh for rearing (Fig. S2). Larvae were reared on young, fully expanded and healthy leaves until their last instar, and frass was regularly removed (at least daily) to keep conditions as clean as possible. Caterpillars were followed from 13/02/2018 until 01/03/18, by which time 21 out of 30 larvae had entered pupation. Eleven larvae were reared to adulthood and stored at −20°C before being freeze dried. Caterpillar body length was measured to the nearest 0.1 cm on a daily basis. We also collected 10 leaves from a separate branch of the same tree before placing the caterpillars (n=48; leaves from one tree were excluded and one tree was sampled for eight leaves). These leaves were placed in an ice box and transferred to a −20°C freezer before being freeze dried at the Binatang Research Centre.

*Chemical variation across sample types*. For the second part of our experiment, we selected an additional two individual trees from each host species, for a total of 12 trees. For each tree, we aimed to place 15 individual caterpillars of the appropriate *Asota* species on individual branches covered by mesh bags. Due to the local availability of first instar caterpillars, however, we ended up with 29 caterpillars on *F. septica*, 32 on *F. hispidioides* and 31 on *F. pachyrrhachis*, for a total of 92 caterpillars. As with the first part, caterpillars were allowed to feed only on young and healthy leaves.

Once caterpillars had moulted into the last instar, ca. one half (n=44) were freeze dried and collected as larval samples while the other half (n=48) were reared to adults. The former were starved for six hours prior to freeze drying to ensure that there was no contamination from plant metabolites. Caterpillars in the rearing treatment were allowed to feed for 24 hours after their final moult before being placed into pots with fresh leaves. Their frass was collected every 45 minutes for four and a half hours. The caterpillars were then returned to their host plants to pupate naturally, and adult moths were freeze dried after emergence. A total of 17 caterpillars from the rearing treatment died in the last instar and were removed; 15 of the 17 deaths were due to parasitoid attack. We also removed four caterpillars and one adult (and their frass samples) due to possible degradation to give a total of 70 insect samples.

Our aim was to collect six frass samples per caterpillar, for a total of 30 frass samples per tree. Because some samples were pooled in the chemical analysis due to low mass, we ended up with 170 frass samples instead of 180. In these cases, samples from the same individual—but not across individuals—were pooled. Frass was frozen immediately upon collection, as were the leaves upon which caterpillars had fed, and were later freeze dried. As the leaves on which the caterpillars fed were in generally poor condition and potentially influenced by caterpillar feeding itself, we again collected and freeze dried 10 leaves from a separate branch of the same tree to serve as baseline comparison.

After accounting for excluded specimens, a total of 362 samples were used for subsequent analyses. The number of samples for each tissue type and adult body part are reported in Table 1. All materials were sent to the University of Turku for chemical analyses.

*Chemical analyses.*

We first grouped our samples based on tissue type (i.e., leaves, frass, caterpillars and adults). Adult samples were then further dissected based on body parts into three subgroups: i) body (which, for this study, was composed of the head, thorax and abdomen), ii) wings, and iii) legs and antennae (Table 1). All samples were ground into fine powder by a ball mill. We macerated 10 mg with 1800 µl methanol overnight in cold room, and then extracted via sonication in a water bath for 30 min, centrifuged at 14,000 rpm for 10 minutes. Supernatants were decanted to new Eppendorf tubes and the methanol was evaporated in an Eppendorf concentrator. Samples were then dissolved in 1000 µl of 5 mM aq. HCl, filtered through 0.20 µm PTFE filters and pipetted into an UHPLC wellplate prior to the UHPLC–MS/MS analyses.

UHPLC–MS/MS analyses were conducted on an Acquity UPLC system coupled with a DAD detector (Waters Corporation, Milfor, MA, USA) and a hybrid quadrupole-Orbitrap mass detector (Q Exactive™, Thermo Fisher Scientific GMbH, Bremen, Germany) via HESI source (H-ESI II, Thermo Fisher Scientific GmbH, Bremen, Germany). The column was Acquity UPLC BEH Phenyl (30 × 2.1 mm i.d.; 1.7 µm; Waters Corporation, Wexford, Ireland). The mobile phases consisted of acetonitrile (A) and 0.1% formic acid (B). The eluent profile was as follows: 0–0.1 min, 3% A in B (isocratic); 0.1–3.0 min, 3–45% A in B (linear gradient); 3.0–3.1 min; 45–90% A in B (linear gradient), 3.1–4.0 min 90% A in B (isocratic); 4.0–4.1 min; 90-3% A in B (linear gradient), 4.1–4.2 min; 3% A in B (isocratic). The eluent flow rate was 0.65 ml/min and injected volume 5 µl. The mass spectrometer was operated in a positive ionization mode with mass range of *m/z* 120-1000 and with lock mass. The following parameters were used for the positive ionization: spray voltage set at 3.0 kV, $N_2$ sheath gas flow rate at 60 arbitrary units, $N_2$ aux gas flow rate at 20 arbitrary units, capillary temperature 380 °C and S-lens RF level at 60. Orbitrap resolution for full scan was 70,000 (full MS) with an automatic gain of $3\times10^6$. Data dependent MS/MS-spectra was obtained with a resolution of 17,500.

Post analysis data handling was done using Thermo Xcalibur Qual Browser software (Version 3.0.63, Thermo Fisher Scientific Inc., Waltham, MA, USA), Compound Discoverer 3.1 (Thermo Fisher Scientific Inc., Waltham, MA, USA) and MZmine version 2.53 (Katajamaa et al. 2006; Pluskal et al. 2010) to achieve quantitation of all possible ions (as the extracted ion area / mg dry weight of tissue). Compound Discoverer utilized the following parameters: untargeted metabolomics workflow template with mass tolerance of 5 ppm, intensity threshold of 30%, S/N threshold of 3, minimum peak intensity of $1\times10^6$, and maximum element count of C×100, H×200, Cl×4, N×10, Na×4, O×100, P×3 and S×5.

For peak detection the following parameters were used: filter peaks true, maximum peak width of 0.5 min, remove singlets true, minimum # scans per peak 5 and minimum # isotopes 1.

Tentative identification of the detected molecules was based on the use of the GNPS analysis environment and its feature based molecular networking (FBMN) (Horai et al. 2010; Pence and Williams 2010; Wang et al. 2016; Nothias et al. 2020). All UHPLC-MS/MS datafiles were converted to mzXML format using ProteoWizard MSConvertGUI (version 3.0.19316). LC-MS feature detection and alignment for GNPS was done using the following MZmine methods and parameters:

- Mass detection: $MS^1$ noise level of $1\times10^5$ and $MS^2$ noise level of 0
- ADAP chromatogram builder: minimum group size in # of scan of 5, group intensity threshold of $3.0\times10^5$, minimum highest intensity threshold of $1.0\times10^5$ and *m/z* tolerance of 0.01 Da
- Chromatogram deconvolution: local minimum search algorithm with chromatographic threshold of $3.0\times10^5$, search minimum in RT range of 0.03 min, minimum relative height of 4.0%, minimum absolute height of $1.0\times10^5$, minimum top peak ratio of 1.4 and peak duration range of 0.01-0.50 min. Chromatogram deconvolution also used $MS^2$ scan pairing at *m/z* range of 0.01 Da and RT range of 0.05 min.
- Isotopic peak grouper: *m/z* tolerance of 0.01 Da, RT tolerance of 0.05 min and maximum charge of 3
- Join aligner: *m/z* tolerance of 0.01 Da, weight for *m/z* at 75, retention time tolerance of 0.05 min and weight of RT at 25
- Gap filling: peak finder (multithreaded) method with intensity tolerance of 25.0%, *m/z* tolerance of 0.001 Da and retention time tolerance of 0.01 min
- Feature list row filter: minimum peaks in an isotope patter selected with value of 2 and keep only peaks with $MS^2$ scan (GNPS) selected

GNPS parameters used for the FBMN are as follows: parent mass tolerance of 0.01 Da, ion tolerance of 0.01 Da, minimum pairs cos of 0.7, minimum matched peaks of 6, network TopK of 10, minimum cluster size of 2 and maximum connected component size of 100. The FBMN was additionally visualized with Cytoscape (version 3.8.1).

Alkaloids found to be indicative of sample type (see *Statistical analyses* for details on alkaloid selection) were further analysed by MS/MS to elucidate their structures on the basis of accurate masses of the molecules and fragments, and their corresponding double bond equivalents. Potential alkaloid-like compounds were statistically analyzed and 17 focal compounds were selected for closer manual examination of their molecular ions (based on CCA scores on the first two axes, please see below for more detail), MS/MS -spectra and corresponding double bond equivalent (DBE) which led to structural characterization of an isoquinoline alkaloid (A5), three pyridoindoles (A9, A16 and A28), four *seco*-phenantroindolizidines (A17, A22, A24, A35), three *seco*-phenantroindolizidines (A25, A36, A43), two

290 seco-phenantroindolizidine-*N*-oxides (A27 and A32), one phenantroindolizidine (A30) and three
291 alkaloid-like compounds (A37, A38 and A41) that we could not classify further.

292 A5, characterized as dihydro-dimethoxy-dihydroisoquinolinium, showed a molecular ion at *m/z*
293 224.09176 and a matching molecular formula $C_{11}H_{13}NO_4$. UV absorption maxima were observed at
294 204, 231 and 340 nm. Highly conjugated A30 was characterized as dimethoxy-dihydro-dibenzo-
295 pyrroloisoquinolinol at *m/z* 346.14421, molecular formula $C_{22}H_{25}NO_3$ and UV absorption maxima at
296 259, 281 and 340 nm.

297 A9, A16 and A28 showed molecular ions at *m/z* 232.10796, 215.08168 and 231.07640, respectively,
298 and molecular formulas $C_{12}H_{13}N_3O_2$, $C_{12}H_{10}N_2O_2$ and $C_{12}H_{10}N_2O_3$. A9 was characterized as amino-
299 tetrahydro-pyridoindole-carboxylic acid with UV absorption maxima observed at 225, 256 and 362 nm.
300 A16 was characterized as dihydro-pyridoindole-carboxylic acid with UV absorption maxima observed
301 at 250, 283 and 361 nm. A28 was characterized as hydroxy-dihydro-pyridoindole-carboxylic acid.

302 A17 and A22 shared the molecular formula $C_{23}H_{27}NO_5$ with molecular ions at *m/z* 398.19639 and
303 398.19640, respectively. These alkaloids were characterized as isomers of dihydroxy-trimethoxy-*seco*-
304 phenantroindolizidine. MS/MS-spectra obtained for A17 and A22 had major fragments at *m/z* 70
305 matching dihydropyrrole and at *m/z* 218 matching dimethoxyphenol-dihydropyridine. UV absorption
306 maxima obtained for 17 at 241, 266 and 343 nm and for 22 at 235 and 269 nm. A24 was characterized
307 as dihydroxy-*seco*-phenantroindolizidine with molecular ion at *m/z* 308.16440 and with molecular
308 formula $C_{20}H_{21}NO_2$. For A24, MS/MS-spectra showed major fragments at *m/z* 188, 186, 70, 107 and
309 212, bearing similarities with other *seco*-phenantroindolizidines. A35 had molecular ion at *m/z*
310 350.21112 and was characterized as dimethoxy-methyl-*seco*-phenantroindolizidine with molecular
311 formula $C_{23}H_{27}NO_2$.

312 A27 and A32 had both major MS/MS fragment at *m/z* 86, characterized as N-hydroxydehydropyrrol,
313 indicating presence of N-oxide moiety in the structure. A27 was characterized as hydroxy-
314 dimethoxyphenyl-methoxyphenyl-hexahydro-indolizine-4-oxide with a molecular formula $C_{23}H_{27}NO_5$,
315 molecular ion at *m/z* 398.19605 and UV absorption maxima at 245, 261 and 343 nm. A32 was
316 characterized as dimethoxyphenyl-methyloxoniophenyl-hexahydro-indolizine-4-oxide with matching
317 molecular ion at *m/z* 382.20112 and molecular formula $C_{23}H_{27}NO_4$.

318 A25 was characterized as hydroxy-trimethoxy-*seco*-phenantroindolizidine at *m/z* 382.20120 with
319 molecular formula $C_{23}H_{28}NO_4$ and matching MS/MS-spectra had major fragments at *m/z* 70 and 218
320 (Lee et al. 2011). For A36 molecular ion was obtained at *m/z* 352.19038 and molecular formula
321 $C_{22}H_{25}NO_3$ with matching characterization of hydroxy-dimethoxy-*seco*-phenantroindolizidine (Stærk
322 et al. 2002). MS/MS fragmentation revealed several fragments at *m/z* 235, 121, 84, 266, 135, 125, 86,
323 334, 202 and 159. A43 was characterized as trimethoxy-*seco*-phenantroindolizidine with molecular ion

at *m/z* 366.20663 and molecular formula $C_{23}H_{27}NO_3$ (Stærk et al. 2002). MS/MS fragmentation revealed major fragments at *m/z* 70 and 202.

*Statistical analyses.*

*Larval development in* Asota. We used linear mixed models, implemented using the R package 'nlme' (Pinheiro et al. 2022), to test how caterpillar body size increased across instars. Body length was the response variable, instar and caterpillar species were the fixed categorical explanatory variables, and individual caterpillar was used as the random explanatory variable to account for multiple measures of the same individual dates as it grew. We also tested the size differences among instars of *A. eusemioides* across its two hosts (*F. hispidioides* and *F. pachyrrhachis*) using similar mixed models, but with host species instead of caterpillar species as the explanatory variable.

*Chemical variation across sample types.* Based on shared conditions for collection and rearing, we used the combined data from the first and second parts of the experiment in subsequent analyses. For the purposes of initial data exploration, we first ran a partial Principal Components Analysis (pPCA) to visualise the total variability in chemical composition between different tissue types and body parts. The effect of the individual was removed by conditioning the 'community' matrix on a vector coding for the individual sampled in cases of paired measures. Alkaloid concentration in area per g dry weight (DW) was log transformed. Mean values across individual host or insect were used when multiple samples of the same type were taken from the same individual.

Our experimental design necessitated the collection of frass from insects subsequently reared to adults, causing us to have some paired data points. To explore and control for this non-independence we ran four analyses. We first ran a standard CCA (i) in which the response variable was the same 'community' matrix dataset used in the pPCA, and the explanatory variables were tissue type and host *Ficus* species. Because *A. eusemioides* is found on both *F. hispidioides* and *F. pachyrrhachis* while *A. heliconia* is restricted to *F. septica*, the variable 'insect species' is collinear with 'host species'. Host species provides greater resolution and information content and is preferred for data exploration. An additional CCA (ii) included an extra explanatory variable, a vector called 'individual', was used to group any paired observations (e.g., frass and the insect from which it was collected). Next, we performed pCCA (iii) using the same set of variables from ii), with the effect of the individual removed by conditioning the 'community' matrix on a vector coding for the individual sampled. Finally, to further control for the possible influence of the individual on compound selection, we ran a standard CCA (iv; with the same formulation as above) on a reduced data set in which alternating adult or frass data points were removed when both were recorded from the same individual. For all CCAs model simplification proceeded through stepwise permutation tests (999 permutations) in both directions, and adjusted $R^2$ was used as the stopping criterion. The significance of explanatory variables were summarised as an

ANOVA table. All multivariate analyses were conducted in the R package 'vegan' (Oksanen et al. 2020) which implements the CCA following Legendre and Legendre (2012).

Additionally, sampling multiple body parts from the same adult allowed us to compare uptake of chemicals in the same individual. To control for the effect of individual, we visualised the adult body part dataset by using pCCA and adding a vector that grouped all samples taken from the same individual (as used in pPCA). Model simplification and significance followed that above as used for tissue type.

Differences in compound occurrence across tree species and sample groupings were initially tested using linear mixed models implemented in 'lme4' (Bates et al. 2015). These models were tested for dispersion, outliers, distribution of residuals and zero-inflation using the R package 'DHARMa' (Hartig 2020). Because our dataset comprised many zeros, we included a single zero-inflation parameter applying to all observations using the R package 'glmmTMB' (Brooks et al. 2017). A Gaussian distribution of errors was appropriate for all models. The response variable, alkaloid concentration (peak area/g DW), was log transformed, and the explanatory variables used for the models depended on which dataset was being used. For the tissue type dataset, we used host tree species and sample type (with adults split into each species) as explanatory variables. For the body part dataset, we used sample type (with adults split into each species) as the explanatory variable. We only split adult moths by species, assigning caterpillars and frass to species would create too many multiple comparisons for meaningful interpretation. We performed Tukey's HSD tests for linear combinations of each explanatory variable using the R package 'multcomp' (Hothorn et al. 2008). As our central question necessitated the use of the same individual across or within developmental stages, we included 'individual' as a random effect to all our mixed models.

## RESULTS

*Larval development in* Asota.

Caterpillar body size increased steadily and predictably, with little overlap in size among instars (Fig. S3). There was a significant difference in body size between instars ($\chi^2 = 1444.315$, $df = 1$, $P < 0.001$) but not between species ($\chi^2 = 1.6413$, $df = 1$, $P = 0.200$; Fig. S3). It was therefore possible to accurately follow larval development in these species through observation and measurement. Furthermore, there was a significant difference in size between instars in caterpillars of *A. eusemioides* ($\chi^2 = 904.221$, $df = 1$, $P < 0.001$) but not across *Ficus* hosts ($\chi^2 = 0.002$, $df = 1$, $P = 0.963$).

*Chemical variation across sample types*.

The pPCA demonstrated a distinct clustering of sample types, with insect tissue broadly grouping together and frass and leaf samples both forming two distinct clusters (Fig. 1a). Several compounds are associated with specific sample types. The first two principal components explained 61% of the total inertia. Standard CCAs with 'individual' included as an explanatory variable (ii in *Methods and*

392 *Materials: Statistical Analyses*) explained an additional 6% of the inertia in comparison to CCAs
393 without the variable individual (i in *Methods and Materials: Statistical Analyses*), although individual
394 was not included as a significant variable following model selection. Chemical composition varied
395 significantly across sample types ($F$ = 36.294, $P$ = 0.001) and tree species ($F$ = 50.858, $P$ = 0.001)
396 respectively; the two variables together explained 68% of the total inertia. These results were also
397 reflected in both (iii) pCCA (Fig. 1b; sample type: $F$ = 30.075, $P$ = 0.001; tree species: $F$ = 25.023, $P$
398 = 0.001; 48% of inertia explained by constrained variables and 27% by conditional variables), and in
399 (iv) a simplified version of the analysis with replicated individuals removed (sample type: $F$ = 23.992,
400 $P$ = 0.001; tree species: $F$ = 40.127, $P$ = 0.001; 68% of the total inertia).

401 There is a much weaker relationship between chemical composition and body part (Figs. 1c and 1d),
402 although there is some degree of separation between body, wings and legs and antennae. While 36% of
403 the total inertia was explained by the first two principal components, the constrained variables in the
404 pCCA explained only 5% of the total inertia in the dataset. In contrast conditional variable contributed
405 61%. Only body part ($F$ = 5.5796, $P$ = 0.001) was significantly correlated to chemical composition,
406 which was collinear with tree species.

407 For sample type we selected 17 alkaloids (hereafter referred to as 'indicator compounds') that showed
408 the highest scores on the first two CCA axes (Table 2). Zero-inflated linear mixed models showed that
409 all indicator compounds except for A9, A16, A37 and 38 differed significantly in occurrence across tree
410 species (Table 3, Table S1). Four alkaloids which were either present in small amounts (A28 and A37)
411 or entirely absent in leaf tissue (A9 and A16) were found in high amounts in both *A. eusemioides* and
412 *A. heliconia* caterpillars and adults (Fig. 2a,b,c). A37 was also absent in *A. eusemioides* frass (Fig. 2a,b),
413 while A38 was present in very small amounts in frass regardless of *Asota* species. A43 was present in
414 leaves and frass but in low frequency in *A. eusemioides* caterpillars and adults (Table 3). Fewer
415 compounds differed significantly across insect tissue (7/17 compounds between adult *A. eusemoides*
416 and *A. heliconia*, with 6/17 and 7/17 for respective comparisons to caterpillars) than between adult and
417 frass samples (*A. eusemoides*: 11/17; *A. heliconia*: 11/17) and adult and leaf (*A. eusemoides*: 14/17; *A.*
418 *heliconia*: 13/17) (Table S1). Leaf tissue was more distinct from insect tissue than frass, while adults
419 and caterpillars most resembled each other in terms of chemical composition. Three indicator
420 compounds (A27, A32 and A37) showed trends across body parts (Fig. 1 d), but these differences were
421 non-significant (Table S2).
422
423 DISCUSSION

424 Our study demonstrates that specialised *Asota* moths not only bypass putative *Ficus* defences but derive
425 new and potentially more potent compounds from them. By identifying alkaloids that possibly shape
426 the interaction between *Asota* and *Ficus* plants, we also demonstrate that overall chemical composition

in adult insects can converge from contrasting chemical starting points. This suggests that these caterpillars filter, ingest and modify specialised plant metabolites to achieve a specific chemical phenotype. Our results thus imply that plant metabolites differ in their bioactivity, and both plants and insects may screen and seek, respectively, for the ones that play the strongest roles in their diverse interactions with other trophic levels (Wetzel and Whitehead 2020).

Caterpillar and adult profiles of each moth/host combination were broadly similar, forming one large but coherent cluster in comparison with the two distinct leaf and frass clusters. This pattern occurred despite the distinct host plant chemical profiles, with *Ficus septica* leaves standing out in particular. Indeed, *F. septica* has been identified to have largely unique defences also by our previous studies on the evolution of both chemical and physical defences among a broader set of New Guinean *Ficus* species (Volf et al. 2018). That *Asota* moths processed the alkaloids from this species and the other hosts studied here in a largely similar way supports our hypothesis of chemical convergence across *Asota* species, probably due to the conservatism in the metabolism of host plant defences among closely related herbivores (Nallu et al. 2018; Allio et al. 2021).

Additionally, our findings on the convergent alkaloid profiles among *Asota* species suggest that compounds relevant for moth fitness form a small and distinct subset of those present in *Ficus* tissue. In this case it is possible that *Asota* moths screen for active *Ficus* leaf compounds that they can potentially use to their own benefit. The properties of these key compounds, possibly anti-fungal or anti-predator, remain to be determined but they likely play a defensive role given the high toxicity of *Asota* (Wills et al. 2016). There are similar parallels from other insect systems where various related compounds are sequestered to produce a limited number of defensive metabolites. For example, *Chrysomela* leaf beetles sequester several but not all salicinoids to produce salicylaldehyde as a defence against invertebrate predators (Soetens et al. 1998). Since both *Asota* species studied here have brightly coloured bodies, perhaps the excretion of alkaloids from within the haemolymph is central to their defence (Sourakov and Emmel 2001). This could be supported by the diverse range of alkaloids found in the bodies of our focal taxa.

One key finding was the occurrence of two indole alkaloids (A9 and A16) in insect tissue despite being totally absent from leaves. As far as we are aware, these compounds have not been previously isolated from insect or plant tissue or synthesised artificially. Metabolism of indoles by insects is known; mirid bugs (Hemiptera: Heteroptera) apparently detoxify or modify indoles to prevent them entering the haemolymph (Hori 1979). Indoles are a class of alkaloids best avoided by invertebrate and vertebrate predators alike, as their effects are deleterious and can extend to behavioural modifications. Indole derivatives, for example, increase mortality (likely through neurotoxic activity) in ants (Costa et al. 2019), which are important predators of caterpillars in tropical forests (Sam et al. 2015). Indoles are also widely employed as signalling molecules and are able to counter viral and bacterial pathogens (Lee

et al. 2015). We suggest that the two novel indole compounds found in our study are either modified from other *Ficus* alkaloids or synthesised *de novo*, and possibly serve a defensive function in the focal system. Until further exploratory research is conducted, we also consider the possibility of these two indole derivatives being produced by endosymbionts. Indeed, the production of bioactive compounds by symbiotic microbes is not a rare occurrence among insects (Beemelmanns et al. 2016). More detailed examination (e.g., precursor feeding) are needed to confirm the molecular pathways involved, but our study opens up a number of opportunities for mechanistic elucidation by highlighting the importance of these particular indole alkaloids.

A large number of alkaloids seem to be passively taken up without differences in occurrence across plant, frass and insect tissue. In comparison to the alkaloids that differed in concertation between the various *Asota* samples and leaves, these alkaloids may play less important roles in *Ficus-Asota* interactions. Our previous studies have shown that various groups of insect herbivores show differential responses to chemical defences in *Ficus*. While *Asota* prefers to feed on highly alkaloidal *Ficus* species, alkaloids structure assemblages of other caterpillar groups in the community (Volf et al. 2018). The alkaloids that were passively taken up by *Asota* could thus be potentially involved in protecting *Ficus* trees from other insect herbivores.

In contrast to the above subset of compounds, phenantroindolizidine alkaloids appear to be actively expelled through frass. While the effects of these specific compounds on insects have yet to be established, a phenantroindolizidine alkaloid isolated from *Cynanchum komarovii* (Asclepiadaceae) was reported to have negative effects on the feeding and growth of *Plutella xylostella* larvae (Guo et al. 2014). We therefore assume that the phenantroindolizidine alkaloids characterised in our study are excreted by *Asota* caterpillars to avoid harmful metabolic effects. However, the ability to deal with these toxic compounds without large differences in growth rates across caterpillar species may help explain specialised *Ficus* feeding in the genus *Asota*. While *Asota* can tolerate these compounds, most groups of insect herbivore avoid feeding on *Ficus* (Basset et al. 1997). Presumably this tolerance of a relatively restricted group of alkaloids trades off with an ability to feed more generally on other chemical classes. Additional studies with increased sampling at the host level are needed to study performance as well as the generality of our findings.

In conclusion, our results provide insights on how physiological adaptations could influence host use in a highly specialised tropical food web. Based on our findings and previous work on their taxonomy and biogeography, *Asota* are well suited for future chemistry- and evolution-based research on phenotypic plasticity and physiological trade-offs in Lepidoptera (Fordyce and Nice 2008). There are two major clades in a phylogeny of *Asota* that include New Guinea species, but *A. eusemioides* and the *A. heliconia* duplex fall into just one of these (Holloway 2022, Supplementary Information). Whilst *A. eusemioides* is restricted to New Guinea and nearby islands, the *A. heliconia* duplex participates in a

presumed complex hybrid system that involves several other species over western China and mainland Southeast Asia. A few widespread species and many more localised ones in the clade also occur between this area of the Asian mainland and Australasia. *Ficus* biogeography is similarly complex, offering a range of opportunities for expansion of a comparative approach to explore selection in geneflow in *Asota* and other *Ficus*-feeding Aganainae. *Asota* and *Ficus* would therefore make an excellent system for further studying the evolution of detoxification and other metabolic strategies, as well as a useful comparison to other insect groups as more powerful predictive methods are developed (Braga et al. 2021).

## REFERENCES


Allio R, Nabholz B, Wanke S, et al (2021) Genome-wide macroevolutionary signatures of key innovations in butterflies colonizing new host plants. Nat Commun 12:354. https://doi.org/10.1038/s41467-020-20507-3

Anderson TJ, Wagner DL, Cooper BR, et al (2017) HPLC-MS analysis of lichen-derived metabolites in the life stages of *Crambidia cephalica* (Grote & Robinson). J Chem Ecol 43:66–74. https://doi.org/10.1007/s10886-016-0799-3

Basset Y, Novotny V, Weiblen G (1997) *Ficus*: a resource for arthropods in the tropics, with particular reference to New Guinea. In: Watt AD, Stork NE, Hunter MD (eds) Forests and insects. Chapman & Hall, London, pp 341–361

Bates D, Mächler M, Bolker B, Walker S (2015) Fitting linear mixed-effects models using lme4. J Stat Soft 67:1–48. https://doi.org/10.18637/jss.v067.i01

Becerra JX (1997) Insects on plants: macroevolutionary chemical trends in host use. Science 276:253–256. https://doi.org/10.1126/science.276.5310.253

Beemelmanns C, Guo H, Rischer M, Poulsen M (2016) Natural products from microbes associated with insects. Beilstein J Org Chem 12:314–327. https://doi.org/10.3762/bjoc.12.34

Braga MP, Janz N, Nylin S, et al (2021) Phylogenetic reconstruction of ancestral ecological networks through time for pierid butterflies and their host plants. Ecol Lett 24:2134–2145. https://doi.org/10.1111/ele.13842

Brooks ME, Kristensen K, van Benthem KJ, et al (2017) glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. The R journal 9:400. https://doi.org/10.3929/ethz-b-000240890

Cogni R, Trigo JR, Futuyma DJ (2012) A free lunch? No cost for acquiring defensive plant pyrrolizidine alkaloids in a specialist arctiid moth (*Utetheisa ornatrix*). Mol Ecol 21:6152–6162. https://doi.org/10.1111/mec.12086

Costa ÂCF, Cavalcanti SCH, Santana AS, et al (2019) Formicidal activity of indole derivatives on *Atta opaciceps* (Borgmeier): lethal, behavioural and locomotive effects. J Appl Entomol 143:58–68. https://doi.org/10.1111/jen.12569

534 Dussourd DE (2017) Behavioral sabotage of plant defenses by insect folivores. Annu Rev Entomol
535 62:15–34. https://doi.org/10.1146/annurev-ento-031616-035030

536 Dyar HG (1890) The number of molts of lepidopterous larvae. Psyche 5:420–422.
537 https://doi.org/10.1155/1890/23871

538 Endara M-J, Coley PD, Ghabash G, et al (2017) Coevolutionary arms race versus host defense chase
539 in a tropical herbivore–plant system. Proc Natl Acad Sci USA 114:E7499–E7505.
540 https://doi.org/10.1073/pnas.1707727114

541 Fordyce JA, Nice CC (2008) Antagonistic, stage-specific selection on defensive chemical
542 sequestration in a toxic butterfly. Evolution 62:1610–1617. https://doi.org/10.1111/j.1558-
543 5646.2008.00388.x

544 Guo H, Yang M, Qi Q (2014) Insecticidal and antifeedant effects of two alkaloids from *Cynanchum*
545 *komarovii* against larvae of *Plutella xylostella* L. J Appl Entomol 138:133–140.
546 https://doi.org/10.1111/jen.12075

547 Hartig F (2020) DHARMa: residual diagnostics for hierarchical (multi-level/mixed) regression
548 models. R package version 0.3

549 Hartmann T, Theuring C, Beuerle T, et al (2005) Acquisition, transformation and maintenance of
550 plant pyrrolizidine alkaloids by the polyphagous arctiid *Grammia geneura*. Insect
551 Biochemistry and Molecular Biology 35:1083–1099.
552 https://doi.org/10.1016/j.ibmb.2005.05.011

553 Heckel DG (2018) Insect detoxification and sequestration strategies. Annual Plant Reviews online
554 77–114. https://doi.org/10.1002/9781119312994.apr0507

555 Holloway JD (1988) The moths of Borneo: family Arctiidae, subfamilies Syntominae, Euchromiinae,
556 Arctiinae; Noctuidae misplaced in Arctiidae (Camptoloma, Aganainae). Southdene Sdn Bhd,
557 Kuala Lumpur

558 Holloway JD (2022) Aspects of the Aganainae. In: Aganainae of the World (Lepidoptera, Erebidae).
559 Museum Witt, Munich, pp 279–293

560 Horai H, Arita M, Kanaya S, et al (2010) MassBank: a public repository for sharing mass spectral
561 data for life sciences. J Mass Spectrom 45:703–714. https://doi.org/10.1002/jms.1777

562 Hori K (1979) Metabolism of ingested indole-3-acetic acid in the gut of various heteropterous insects.
563 Appl Ent Zool 14:149–158. https://doi.org/10.1303/aez.14.149

564 Hothorn T, Bretz F, Westfall P (2008) Simultaneous inference in general parametric models. Biom J
565 50:346–363. https://doi.org/10.1002/bimj.200810425

566 Katajamaa M, Miettinen J, Orešič M (2006) MZmine: toolbox for processing and visualization of
567 mass spectrometry based molecular profile data. Bioinformatics 22:634–636.
568 https://doi.org/10.1093/bioinformatics/btk039

569 Kessler A, Kalske A (2018) Plant secondary metabolite diversity and species interactions. Annu Rev
570 Ecol Evol Syst 49:115–138. https://doi.org/10.1146/annurev-ecolsys-110617-062406

571 Lee J-H, Wood TK, Lee J (2015) Roles of indole as an interspecies and interkingdom signaling
572 molecule. Trends Microbiol 23:707–718. https://doi.org/10.1016/j.tim.2015.08.001

Lee Y-Z, Huang C-W, Yang C-W, et al (2011) Isolation and biological activities of phenanthroindolizidine and septicine alkaloids from the Formosan *Tylophora ovata*. Planta medica 77:1932–1938

Legendre P, Legendre L (2012) Numerical ecology, 3rd edn. Elsevier

Lindigkeit R, Biller A, Buch M, et al (1997) The two faces of pyrrolizidine alkaloids: the role of the tertiary amine and its N-oxide in chemical defense of insects with acquired plant alkaloids. Eur J Biochem 245:626–636. https://doi.org/10.1111/j.1432-1033.1997.00626.x

Nallu S, Hill JA, Don K, et al (2018) The molecular genetic basis of herbivory between butterflies and their host plants. Nat Ecol Evol 2:1418–1427. https://doi.org/10.1038/s41559-018-0629-9

Nothias L-F, Petras D, Schmid R, et al (2020) Feature-based molecular networking in the GNPS analysis environment. Nat Methods 17:905–908. https://doi.org/10.1038/s41592-020-0933-6

Novotny V, Miller SE, Baje L, et al (2010) Guild-specific patterns of species richness and host specialization in plant–herbivore food webs from a tropical forest. J Anim Ecol 79:1193–1203. https://doi.org/10.1111/j.1365-2656.2010.01728.x

Oksanen J, Blanchet FG, Friendly M, et al (2020) vegan: Community Ecology Package. Version 2.5-7URL https://CRAN.R-project.org/package=vegan

Pence HE, Williams A (2010) ChemSpider: an online chemical information resource. J Chem Educ 87:1123–1124. https://doi.org/10.1021/ed100697w

Petschenka G, Agrawal AA (2016) How herbivores coopt plant defenses: natural selection, specialization, and sequestration. Curr Opin Insect Sci 14:17–24. https://doi.org/10.1016/j.cois.2015.12.004

Pinheiro J, Bates D, R Core Team (2022) nlme: linear and nonlinear mixed effects models. R package version 3.1-157, https://CRAN.R-project.org/package=nlme

Pluskal T, Castillo S, Villar-Briones A, Orešič M (2010) MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. BMC Bioinformatics 11:395. https://doi.org/10.1186/1471-2105-11-395

Salek RM, Steinbeck C, Viant MR, et al (2013) The role of reporting standards for metabolite annotation and identification in metabolomic studies. GigaScience 2(1):2047–217X. https://doi.org/10.1186/2047-217X-2-13

Salminen J-P, Lahtinen M, Lempa K, et al (2004) Metabolic modifications of birch leaf phenolics by an herbivorous insect: detoxification of flavonoid aglycones via glycosylation. Z Naturforsch C 59:437–444. https://doi.org/10.1515/znc-2004-5-627

Sam K, Koane B, Novotny V (2015) Herbivore damage increases avian and ant predation of caterpillars on trees along a complete elevational forest gradient in Papua New Guinea. Ecography 38:293–300. https://doi.org/10.1111/ecog.00979

Scott Chialvo CH, Chialvo P, Holland JD, et al (2018) A phylogenomic analysis of lichen-feeding tiger moths uncovers evolutionary origins of host chemical sequestration. Mol Phylogenet Evol 121:23–34. https://doi.org/10.1016/j.ympev.2017.12.015

611 Soetens P, Pasteels JM, Daloze D, Kaisin M (1998) Host plant influence on the composition of the
612 defensive secretion of *Chrysomela vigintipunctata* larvae (Coleoptera: Chrysomelidae).
613 Biochem Syst Ecol 26:703–712. https://doi.org/10.1016/S0305-1978(98)00039-8

614 Sourakov A, Emmel TC (2001) On the toxic diet of day-flying moths in the Solomon Islands
615 (Lepidoptera: Arctiidae). Tropical Lepidoptera Research 12:5–6

616 Stærk D, Lykkeberg AK, Christensen J, et al (2002) In vitro cytotoxic activity of
617 phenanthroindolizidine alkaloids from *Cynanchum vincetoxicum* and *Tylophora tanakae*
618 against drug-sensitive and multidrug-resistant cancer cells. J Nat Prod 65:1299–1302.
619 https://doi.org/10.1021/np0106384

620 Trigo JR (2011) Effects of pyrrolizidine alkaloids through different trophic levels. Phytochem Rev
621 10:83–98. https://doi.org/10.1007/s11101-010-9191-z

622 Volf M, Laitila JE, Kim J, et al (2020) Compound specific trends of chemical defences in *Ficus* along
623 an elevational gradient reflect a complex selective landscape. J Chem Ecol 46:442–454.
624 https://doi.org/10.1007/s10886-020-01173-7

625 Volf M, Segar ST, Miller SE, et al (2018) Community structure of insect herbivores is driven by
626 conservatism, escalation and divergence of defensive traits in *Ficus*. Ecol Lett 21:83–92.
627 https://doi.org/10.1111/ele.12875

628 Wahlberg N (2001) The phylogenetics and biochemistry of host-plant specialization in melitaeine
629 butterflies (Lepidoptera: Nymphalidae). Evolution 55:522–537.
630 https://doi.org/10.1111/j.0014-3820.2001.tb00786.x

631 Wang M, Carver JJ, Phelan VV, et al (2016) Sharing and community curation of mass spectrometry
632 data with Global Natural Products Social Molecular Networking. Nat Biotechnol 34:828–837.
633 https://doi.org/10.1038/nbt.3597

634 Weller SJ, Jacobson NL, Conner WE (1999) The evolution of chemical defences and mating systems
635 in tiger moths (Lepidoptera: Arctiidae). Biol J Linn Soc 68:557–578.
636 https://doi.org/10.1111/j.1095-8312.1999.tb01188.x

637 Wetzel WC, Whitehead SR (2020) The many dimensions of phytochemical diversity: linking theory
638 to practice. Ecol Lett 23:16–32. https://doi.org/10.1111/ele.13422

639 Wills PJ, Anjana M, Nitin M, et al (2016) Population explosions of tiger moth lead to lepidopterism
640 mimicking infectious fever outbreaks. PLoS ONE 11:e0152787.
641 https://doi.org/10.1371/journal.pone.0152787

642 Wink M (2019) Quinolizidine and pyrrolizidine alkaloid chemical ecology – a mini-review on their
643 similarities and differences. J Chem Ecol 45:109–115. https://doi.org/10.1007/s10886-018-
644 1005-6

645 Zahiri R, Holloway JD, Kitching IJ, et al (2012) Molecular phylogenetics of Erebidae (Lepidoptera,
646 Noctuoidea). Syst Entomol 37:102–124. https://doi.org/10.1111/j.1365-3113.2011.00607.x

647 Zaspel JM, Weller SJ, Wardwell CT, et al (2014) Phylogeny and evolution of pharmacophagy in tiger
648 moths (Lepidoptera: Erebidae: Arctiinae). PLoS ONE 9:e101975.
649 https://doi.org/10.1371/journal.pone.0101975

650

651

## Declarations

### Funding

A.M.F., V.N., M.V. and S.T.S. acknowledge the Grant Agency of the Czech Republic (grant number 19-28126X). V.N. acknowledges support from the European Research Council grant no. 669609. S.T.S. acknowledges departmental support from Harper Adams University. This material is based upon work supported by the U.S. National Science Foundation under grants DEB 9707928, 0211591 and 0515678, 0816749 and 0841885. DNA barcoding was provided by the Biodiversity Institute of Ontario, University of Guelph, with funding from Genome Canada and the Ontario Genomics Institute to the International Barcode of Life Project. J-P.S. acknowledges funding from the Academy of Finland.

### Conflict of interest

The authors have no conflict of interest to disclose.

### Data availability statement

Data will be made available from the Dryad Digital Repository on publication of the manuscript.

### Author contributions

J-P.S., V.N., M.V. and S.T.S, conceived the ideas and designed the methodology; G.A., M.S., I.K. and J-P.S. collected the data; S.T.S. analysed the data; A.M.F. and S.T.S. prepared the figures; A.M.F., G.A., S.M., J.H., M.V. and S.T.S. led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

## Tables

**Table 1.** A table of tissue types collected in the study. For adults we analysed three different sections of the body. Note that 17 caterpillars and their associated frass sampled were removed for statistical analysis (please see *Chemical variation across sample types*).

| Sample Type | n |
|---|---|
| Adult | 39 |
| Caterpillar | 31 |
| Leaf | 106 |
| Tree | 12 |
| Frass | 108 |

706 **Table 2** Molecular formulae, grouping and characterization of 17 compounds strongly indicative of
707 each sample type (full HPLC-MS/MS data available in the SI). Compounds marked with an asterisk
708 were found in Volf et al. (2018). Level of identification is given following Salek et al. (2013).
709

| Alkaloid group | Compound | Formula | Characterization | Level |
|---|---|---|---|---|
| isoquinoline alkaloid | A5 | $C_{11}H_{13}NO_4$ | dihydroxy-dimethoxy-dihydroisoquinolinium | 4 |
| pyridoindole | A9 | $C_{12}H_{13}N_3O_2$ | amino-tetrahydro-pyridoindole-carboxylic acid | 4 |
| | A16 | $C_{12}H_{10}N_2O_2$ | dihydro-pyridoindole-carboxylic acid | 4 |
| | A28 | $C_{12}H_{10}N_2O_3$ | hydroxy-dihydro-pyridoindole-carboxylic acid | 4 |
| *seco*-phenantroindolizidine | A17 | $C_{23}H_{27}NO_5$ | dihydroxy-trimethoxy-*seco*-phenantroindolizidine | 4 |
| | A22 | $C_{23}H_{27}NO_5$ | dihydroxy-trimethoxy-*seco*-phenantroindolizidine | 4 |
| | A24 | $C_{20}H_{21}NO_2$ | dihydroxy-*seco*-phenantroindolizidine | 3 |
| | A25* | $C_{23}H_{28}NO_4$ | hydroxy-trimethoxy-*seco*-phenantroindolizidine[1] | 2 |
| | A35 | $C_{23}H_{27}NO_2$ | dimethoxy-methyl-phenantroindolizidine | 4 |
| | A36* | $C_{22}H_{25}NO_3$ | hydroxy-dimethoxy-*seco*-phenantroindolizidine[2] | 3 |
| | A43* | $C_{23}H_{27}NO_3$ | trimethoxy-*seco*-phenantroindolizidine[2] | 4 |
| *seco*-phenantroindolizidine-*N*-oxide | A27 | $C_{23}H_{27}NO_5$ | hydroxy-dimethoxyphenyl-methoxyphenyl-hexahydro-indolizine-oxide | 4 |
| | A32* | $C_{23}H_{27}NO_4$ | dimethoxyphenyl-methyloxoniophenyl-hexahydro-indolizine-oxide | 4 |
| phenantroindolizidine | A30 | $C_{22}H_{25}NO_3$ | dimethoxy-dihydrodibenzopyrroloisoquinolinol | 4 |
| not classified | A37 | - | - | 4 |
| | A38 | - | - | 4 |
| | A41 | - | - | 4 |

710

711

712 **Table 3** Results of zero-inflated linear mixed models, including which sample types the concentration
713 of the 17 'indicator compounds' were strongly related to.

| Alkaloid group | Compound | Tree species ($df = 2$) | | Sample type ($df = 4$) | | Indicative sample type |
|---|---|---|---|---|---|---|
| | | $\chi^2$ | *P* | $\chi^2$ | *P* | |
| isoquinoline alkaloid | A5 | 21.266 | < 0.001 | 657.407 | < 0.001 | leaf, caterpillar and adult, |
| pyridoindole | A9 | 0.099 | 0.952 | 1589.940 | < 0.001 | caterpillar and adult |
| | A16 | 2.453 | 0.293 | 10243.693 | < 0.001 | caterpillar and adult |
| | A28 | 13.667 | 0.001 | 453.957 | < 0.001 | caterpillar, adult and frass |
| *seco*-phenantroindolizidine | A17 | 296.640 | < 0.001 | 74.802 | < 0.001 | caterpillar and adult |
| | A22 | 202.840 | < 0.001 | 256.520 | < 0.001 | leaf and frass |
| | A24 | 71.998 | < 0.001 | 13.562 | 0.009 | *Ficus septica* |
| | A25* | 130.060 | < 0.001 | 238.250 | < 0.001 | leaf and frass |
| | A35 | 91.683 | < 0.001 | 570.000 | < 0.001 | frass, *Ficus septica* |
| | A36* | 729.320 | < 0.001 | 979.060 | < 0.001 | leaf and frass, *Ficus septica* |
| | A43* | 151.150 | < 0.001 | 818.910 | < 0.001 | leaf and frass |
| *seco*-phenantroindolizidine-*N*-oxide | A27 | 24.382 | < 0.001 | 178.698 | < 0.001 | leaf and frass |
| | A32* | 106.810 | < 0.001 | 340.250 | < 0.001 | leaf and frass |
| phenantroindolizidine | A30 | 94.338 | < 0.001 | 102.512 | < 0.001 | caterpillar, adult and frass |
| not classified | A37 | 4.485 | 0.106 | 20.046 | < 0.001 | caterpillar and adult |
| | A38 | 4.693 | 0.096 | 88.210 | < 0.001 | caterpillar and adult |
| | A41 | 58.739 | < 0.001 | 29.058 | < 0.001 | caterpillar and adult |

## Figure captions

**Fig. 1** Biplots of the first two partial PCA (1a and 1c) and partial CCA (1b and 1d) axes displaying the distribution of sample types based on their alkaloid profiles. For (a) and (b), samples are classified based on the tissue type (i.e., leaf, caterpillar, frass and adult) and the *Ficus* or *Asota* species from which they were collected. For (c) and (d), samples are classified according to adult body parts (i.e., body, wings, and legs and antennae) and *Asota* species. Alkaloids (shown in grey text in the biplots) selected for further analyses are listed in Table 2.

**Fig. 2** Boxplots showing the concentrations (in log(area per g DW)) of 17 alkaloids indicative of tissue type in (a) *Asota eusemioides* and *Ficus hispidioides*, (b) *Asota eusemioides* and *Ficus pachyrrhachis,* and (c) *Asota heliconia* and *Ficus septica*.