

# Evolutionary history of tyrosine-supplementing endosymbionts in pollen-feeding beetles

Jürgen C. Wierz<sup>1,2</sup>, Matthew L. Gimmel<sup>3</sup>, Selina Huthmacher<sup>2</sup>, Tobias Engl<sup>1,2</sup>, Martin Kaltenpoth<sup>1,2,\*</sup>

<sup>1</sup> Department of Insect Symbiosis, Max Planck Institute for Chemical Ecology, Jena, Germany

<sup>2</sup> Department of Evolutionary Ecology, Institute of Organismic and Molecular Evolution, Johannes Gutenberg University, Mainz, Germany

<sup>3</sup> Department of Invertebrate Zoology, Santa Barbara Museum of Natural History, Santa Barbara, CA, USA

\* Corresponding author: Martin Kaltenpoth, Department of Insect Symbiosis, Max Planck Institute for Chemical Ecology, Hans-Knöll-Str. 8, D-07745 Jena Germany, kaltenpoth@ice.mpg.de

© The Author(s) [2024]. Published by Oxford University Press on behalf of the International Society for Microbial Ecology

## Abstract

Many insects feeding on nutritionally challenging diets like plant sap, leaves, or wood engage in ancient associations with bacterial symbionts that supplement limiting nutrients or produce digestive or detoxifying enzymes. However, the distribution, function, and evolutionary dynamics of microbial symbionts in insects exploiting other plant tissues or relying on a predacious diet remain poorly understood. Here, we investigated the evolutionary history and function of the intracellular gamma-proteobacterial symbiont '*Candidatus Dasytiphilus stammeri*' in soft-winged flower beetles (Melyridae, Dasytinae) that transition from saprophagy or carnivory to palinivory (pollen-feeding) between larval and adult stage. Reconstructing the distribution of the symbiont within the Dasytinae phylogeny unraveled a long-term coevolution, originating from a single acquisition event with subsequent host-symbiont codiversification, but also several independent symbiont losses. The analysis of 20 different symbiont genomes revealed that their genomes are severely eroded. However, the universally retained shikimate pathway indicates that the core metabolic contribution to their hosts is the provisioning of tyrosine for cuticle sclerotization and melanization. Despite the high degree of similarity in gene content and order across symbiont strains, the capacity to synthesize additional essential amino acids and vitamins and to recycle urea is retained in some but not all symbionts, suggesting ecological differences among host lineages. This report of tyrosine-provisioning symbionts in insects with saprophagous or carnivorous larvae and pollen-feeding adults expands our understanding of tyrosine supplementation as an important symbiont-provided benefit across a broad range of insects with diverse feeding ecologies.

Keywords:

Melyridae Dasytinae, pollen feeding, *Dasytiphilus stammeri*, bacteriome, mutualism, tyrosine supplementation, genome erosion, host symbiont co-evolution, cocladogenesis

## Introduction

Beetles (Coleoptera) are the most speciose of all insect orders and show a remarkable diversity in physiological and morphological traits (1). Their defining characteristic is the modification of the anterior pair of wings into protective wing-cases, called elytra, which provide efficient protection against predators and various environmental stressors such as desiccation (2,3), allowing the colonization of diverse habitats. However, the high investment into the elytral cuticle comes at a cost, as beetles require high amounts of tyrosine during development. This aromatic amino acid is the precursor for 3,4-dihydroxyphenylalanine (DOPA), a central metabolite in the biosynthesis of melanin that serves for cuticle tanning, as well as phenolic compounds used for cuticle hardening (sclerotization) (4). Insects in general lack the shikimate pathway for the biosynthesis of aromatic compounds and therefore are not able to produce tyrosine *de novo*. Hence, they rely on taking up aromatic amino acids via the food or acquiring them with the help of microbial symbionts.

Symbiotic interactions between insects and bacteria are widespread (5,6), and symbiont-provided nutritional supplements enable many insect taxa to thrive in otherwise inhospitable ecological niches (7–10). Consequently, symbiont acquisition or loss events can shape the ecology of the hosts. A new symbiont acquisition could allow a host to transition between ecological niches and compensate for a change in nutrient availability. Symbionts producing tyrosine or its precursors have been described in herbivorous beetles of the families Curculionidae, Silvanidae, and Bostrichidae (11–14), as well as in certain ants (15,16). Experimental deprivation of tyrosine-

supplementing symbionts can result in a lighter, softer cuticle (11,12,17), which reduces the protection from desiccation (12,18) as well as from predation and pathogen infestation (19).

Intimate long-term associations, including those with tyrosine-provisioning microorganisms, can have a severe impact on the symbionts' genome evolution. As vertically transmitted symbionts permanently live in the stable environment of the host, many genes experience relaxed selection and consequently deteriorate (20). Such genome reduction usually eliminates genes that are neither required by the symbiont nor beneficial to the host (21), whereas symbiont capabilities that are advantageous to the host are maintained (22). As a result, the gene repertoire gets fine-tuned according to the hosts' needs, and even in close related co-evolving symbiotic systems the symbionts' capabilities can differ in key elements depending on the ecology of the host (23). A common adaptation to compensate for the loss of genes and thus reduce the effects of genome erosion are multifunctional symbiont enzymes (24). Furthermore, hosts can sometimes compensate for symbiont gene losses, completing otherwise fragmentary pathways (25). For example, some weevils encode a gene for the final step in the tyrosine biosynthesis pathway, i.e. tyrosine transaminase, utilizing precursors that are supplied by their *Nardonella* symbionts, which encode the shikimate pathway but lack the transaminase (11).

Soft-winged flower beetles (Melyridae) form the most speciose beetle family within the polyphagous superfamily Cleroidea, with over 6,000 described species in about 300 genera and three subfamilies worldwide (26–28). Many species may play important roles in pollination, as adults are found on flowers, often in large aggregations, and feed on pollen (29,30). By contrast, larvae have been recorded in leaf litter and dead wood and may be xylophagous, saprophagous, predacious, or scavenge on dead arthropods (26,31,32). Moreover, it seems that larvae of some species switch from being scavengers to herbivores during their development (33). Whereas a diet composed of arthropod prey or carcasses usually is nutrient rich, a pollen-based diet can be scarce in certain nutrients (34,35). Adults of both sexes from at least some species within the family carry as-yet functionally uncharacterized bacteriome-localized symbionts (9,36). More

specifically, *Dasytes virens* and *Dasytes plumbeus*, members of the subfamily Dasytinae, harbor a single intracellular *Enterobacteriaceae* symbiont called ‘*Candidatus Dasytiphilus stammeri*’ (henceforth “*Dasytiphilus*”) (36). However, no bacteriome-housed symbionts have been reported for species in other Melyridae subfamilies. Moreover, a different lifestyle has been observed in other subfamilies, where the adults can be aggressive and prey on small insects, as well as on each other (37).

Beyond its morphological and taxonomic description, little is known about this symbiotic system, which is surprising for a beetle family of this size and potential influence on ecosystem services. Hence, we investigated 58 species of Melyridae for the presence of bacterial symbionts by a combination of high-throughput 16S rRNA gene based microbiota profiling, diagnostic PCRs, metagenome sequencing, and fluorescence *in situ* hybridization. We reconstructed a molecular phylogeny of the hosts to elucidate the origin and evolutionary dynamics of the association with nutritional symbionts, supporting a single acquisition of bacteriome-localized symbionts in Dasytinae and multiple loss events. In order to gain functional insights, we assembled the symbiont genomes of 20 host species, revealing tyrosine biosynthesis as a conserved pathway across all investigated *Dasytiphilus* strains, with additional amino acid and vitamin biosynthetic pathways being ancestrally encoded and retained in genomes of certain symbiont subclades but lost in others. These findings elucidate the function and dynamic evolutionary history of intracellular symbionts in pollen-feeding Dasytinae beetles and highlight the importance of tyrosine supplementation for the ecology and evolution across beetle families with diverse feeding ecologies.

## Material and methods

### Melyridae species collection

Melyridae beetles were collected during the summers of 2011 to 2022 from various locations (Table 1). Individuals were stored either in ethanol or dry at -20°C or -80°C, depending on future use. Species from the USA and *Malachius bipustulatus* were identified based on morphology, species from Europe by barcoding of the CO1 gene (see below for details).

### DNA extraction

Head and leg tissue samples taken from the insects were extracted for the identification of host species. For diagnostic PCRs for the presence of *Dasytiphilus*, bacterial community profiling, and genome sequencing, either dissected bacteriomes or, if not possible, entire host abdomens were extracted in order to obtain bacterial symbiont DNA (Figure S1). Extractions were done using the Epicentre MasterPure™ Complete DNA and RNA Purification Kit (Illumina Inc., Madison, WI, USA) following the manufacturer's instructions, including RNase digestion. DNA of samples intended for genome sequencing had its purity assessed using a NanoPhotometer P330 (Implen, München, Germany) and DNA quantity was measured using a Qubit dsDNA BR assay kit on a Qubit 2.0 Fluorometer (Invitrogen by Thermo Fisher Scientific, USA, MA).

### Diagnostic PCR and Sanger sequencing to assess the presence of

#### *Dasytiphilus*

General eubacterial and diagnostic PCRs were performed to assess *Dasytiphilus* presence and were done on a T-Professional-Gradient Thermocycler (Biometra, Göttingen, Germany) using reaction mixes of 6.9 µL ultrapure H<sub>2</sub>O, 1.25 µL of 10x reaction buffer, 0.25 µL 25 mM MgCl<sub>2</sub>, 1.5 µL 2 mM dNTPs, 1 µL of both forward and reverse primer (each 10 pmol/µl), 0.1 µL of 5 U/µL Taq

DNA polymerase and 1  $\mu$ L template. Alternatively, a Mastercycler EP Gradient S Thermocycler (Eppendorf AG, Hamburg, Germany) was used with a reaction mix containing 9.5  $\mu$ L ultrapure H<sub>2</sub>O, 12.5  $\mu$ L of Q5<sup>®</sup> High-Fidelity 2X Master Mix (NEB), 1  $\mu$ L of both forward and reverse primer (each 10 pmol/ $\mu$ l), and 1  $\mu$ L template. Primers used for *Dasytiphilus* identification were the eubacterial primer pairs fD1 and rP2 as well as slightly modified versions fD1\_Mely and rP2\_Mely. Moreover, the symbiont-specific primer pair Dasy\_Sym\_fwd3 and Dasy\_Sym\_rev3 was utilized (Table S1).

To guarantee that *Dasytiphilus* sequences were amplified, Sanger sequencing was conducted following general eubacterial and diagnostic PCRs. Products were purified using the innuPREP PCRpure Kit (Analytik Jena, Jena, Germany) according to the manufacturer's instructions, and were subsequently sequenced by a commercial service (StarSeq, Mainz, Germany). Alternatively, products were purified with the Zymo Research DNA Clean & Concentrator-5 kit (Zymo Research, Irvine, CA, USA) following the manufacturer's protocol and were sequenced in house with an AB Hitachi 3730XL DNA Analyzer (Applied Biosystems by Thermo Fisher Scientific, USA, MA).

### **Microbial community profiling by 16S rRNA gene amplicon sequencing**

Bacterial community composition was evaluated via high-throughput amplicon sequencing of the 16S rRNA gene. Samples were sequenced in a paired end approach with read lengths of 300nt by a commercial provider (StarSeq, Mainz, Germany) on an Illumina MiSeq platform (Illumina Inc., San Diego, CA, USA) using V3 reagent and 25% PhiX to balance the composition of bases. Amplified regions were either V3-V4 region with primers 341f and 806bR, V4 region with primers 515F and 806bR or V4-V5 region with primers 515F and 909R (Table S1). This was due to temporal differences in sample availability and constraints in assay design due to concurrent sequencing projects. A comprehensive list of all sequenced of Melyridae samples with the

amplified region can be found in Table S2. No-template DNA extractions were sequenced as negative controls. Based on the sequenced Illumina reads, amplicon sequence variants (ASVs) were obtained after read trimming depending on the amplified region, quality filtering, dereplicating, and chimera removal in R utilizing the package DADA2 (38). Subsequently, the pre-trained classifier Silva 138.1 was used to assign taxonomy (39,40). Finally, reads identified as chloroplast or mitochondria were removed prior to analysis.

### **Symbiont localization via fluorescence *in situ* hybridization (FISH)**

To localize the symbionts in multiple species of Melyridae beetles, Fluorescence *in situ* hybridization (FISH) was performed on specimens of 13 different Dasytinae species and *Malachius bipustulatus* as an outgroup (Table 1). Whole beetles or abdomens were individually fixed in Carnoy's solution (67% ethanol, 25% chloroform, and 8% glacial acetic acid), dehydrated in a series of ascending concentration (30, 50, 70, 80, 90, and 96%) of n-Butanol and subsequently embedded in Technovit 8100 (Heraeus Kulzer) following the manufacturer's instructions. Transversal or sagittal histological sections of 8µm thickness were prepared utilizing a glass knife on a Leica RM 2245 microtome and transferred to microscope slides. To each slide, 100–150µl hybridization mix was applied and then covered with a glass cover slip. The hybridization mix contained hybridization buffer (0.9 M NaCl, 0.02 M Tris/HCl (pH=8), 0.01% SDS), 0.5 µM of the fluorescently labelled oligonucleotide probes EUB338-Cy3 and *Dasy\_ent\_cy5* (Table S1) and 0.5 mg/ml DAPI for host cell counterstaining. Slides were hybridized over-night at 50°C in a humid chamber. Afterwards, the glass cover slip was removed and the slides were washed at 50°C for two hours submerged in wash buffer, followed by an additional washing step in distilled water at 50°C for 20 minutes. The wash buffer consisted of 0.1 M NaCl, 0.02 M Tris/HCl (pH=8), 5 mM EDTA, and 0.01% SDS. After washing, 30 µl of VectaShield was applied to each slide upon which a glass cover slip was then placed. Samples



were observed with a Thunder Imaging System (Leica, Wetzlar, Germany). The signals for cy3, cy5, background, and DAPI were acquired with the 555-nm, 635-nm, 475-nm, and 390-nm EFW LED8 respectively, at 50% (cy3, cy5, background) power or 5% power (DAPI), the DFT51010 filter cube, and a 590-nm, 642-nm, 535-nm, and 460-nm fast emission filter, respectively. Finally, images were processed in the Leica Application Suite X software (Leica, Wetzlar, Germany) with the instant and small volume computational clearing algorithms.

### **Symbiont genome sequencing and assembly**

For symbiont genome reconstruction, short-read shotgun library preparation and sequencing was performed by the Max Planck Genome Center (Cologne, Germany) utilizing a HiSeq 3000 or NextSeq 2000 system (Illumina Inc., San Diego, CA, USA), or by a commercial service (CeGaT GmbH, Tübingen, Germany) using a NovaSeq 6000 platform (Illumina Inc., San Diego, CA, USA) (Table S3). Obtained paired sequence reads of 150 bp each were uploaded to KBase (41) and read quality was assessed with FastQC v0.11.5-v0.11.9. The reads were trimmed using Trimmomatic v0.36 (42), and trimmed reads were assembled with metaSPAdes v3.13.0-v3.15.3 (43) and MEGAHIT v1.2.9 (44).

For *Listrus* sp. 02, long reads were used in addition to the Illumina reads to assemble the symbiont genome. For this, high molecular weight (HMW) genomic DNA was selectively extracted with the Short Read Eliminator Kit XS (Circulomics, USA, MD), to enrich for fragments longer than 10 kb. The size-selected HMW DNA was used as the starting material for the preparation of an ONT library, following the manufacturer's guidelines for the Ligation Sequencing Kit SQK-LSK110 (Oxford Nanopore Technologies, Oxford, United Kingdom), and sequenced on R9.4.1 (FLO-MIN106) flow cells using the MinION sequencing device with MinKNOW Software version 22.10.7. Super High-accuracy (SUP) base-calling was performed using GUPPY version 5.0.14. Flye (v2.9.1) assembler (45,46) in meta-mode was used to generate a de novo genome assembly

from the ONT data, followed by four iterations of polishing using Racon, and one round of error correction using Medaka. For final error correction of the ONT-based assembly, we used ntEdit with paired-end (2x 150 bp) Illumina data generated from the same HMW DNA used for ONT sequencing. To remove duplications (heterozygous regions) and generate haploid genome(s) for further downstream analysis, we used purge\_haplotigs.

### **Symbiont genome analysis**

After assembly, the received contigs were annotated with Prokka v1.14.5 (47) and Rast v1.073 (48). Symbiont contigs were identified in Geneious (Geneious Prime various versions between 2019 and 2023) based on length, GC content, read coverage and gene synteny. Contigs from both assembly methods were de-novo assembled, resulting in either closed genomes or high-quality draft genomes. Minor polishing was done on selected genomes by read mapping and pilon genome improvement (49). Genomes were analyzed with the help of KEGG: Kyoto Encyclopedia of Genes and Genomes (50–52). Gene synteny was visualized with Clinker (53), showing only the highest similarity links between genes. When analyzing specific symbiont capabilities, we focused on pathways that could be directly relevant to the interaction with the host.

### **Symbiont phylogenetic reconstruction**

The symbiont phylogeny was reconstructed in KBase (41) with the plugin “Insert Set of Genomes Into SpeciesTree – v2.2.0”. The plugin took user-provided genomes and constructed alignments of 49 core, universal genes defined by COG (Clusters of Orthologous Groups) gene families with publicly available genomes of closely related bacteria. Subsequently, it used an approximately-maximum-likelihood algorithm to create a phylogenetic tree (54).

### *dN/dS of lysine pathway genes*

We evaluated the selective forces acting on lysine pathway genes of *Listrus* spp. more closely by examining the ratio of synonymous vs non-synonymous substitutions in order to assess whether substitution patterns differed from those of lysine pathway genes of *Dasytastes* clade symbionts. To infer orthologous genes, orthogroups were reconstructed using OrthoFinder v2.5.4 (55) with the MSA option enabled using muscle v5.1 (56) and FastTree v2.11.1 (54). Orthologs for each of the candidate genes were extracted from the phylogenetic hierarchical orthogroups results and used for downstream analysis. For each candidate gene, orthologous amino acid sequences were aligned with muscle v5.1 with default parameters. Then, the corresponding nucleotide sequences were codon-aligned based on the protein alignments using the pal2nal v14 script (57). Phylogenetic trees were reconstructed using FastTree v2.11.1 based on codon alignments. Both tree and alignment were used to detect changes in the selective pressures acting on *Listrus* spp. genes compared to the outgroup. Branches corresponding to *Listrus* spp. were marked using the vision.hyphy.org tool and the “select descendant branches” enable. Eventual relaxation or intensification of the selection strength was tested using RELAX from HYPHY package v2.5.33 (58). Additionally, a test for positive selection was conducted using aBSREL (59) from HYPHY package. The *p* value threshold for considering tests as significant was set at 0.05.

## **Host identification and phylogenetic reconstruction**

### *Host species identification*

To robustly link host species with symbiont presence/absence, we attempted to identify each analyzed host taxon. Host species that could not be identified based on morphology were identified via DNA barcoding, i.e. sequencing of the cytochrome c oxidase subunit I gene (CO1). In this gene, more than 2% sequence divergence is usually recorded between closely related animal species, so the mutation rate is high enough to distinguish between species (60). The CO1

gene was amplified using PCR with primers LepF1 and LepR1 (Table S1) and the product was purified and subjected to Sanger sequencing (see above for details). However, several species could not be identified to a described species because no conspecific reference sequence was available and the necessary modern taxonomic revisions have not been completed. For the genus *Listrus*, in particular, species are poorly characterized and undescribed species certainly exist; as a consequence, only morphospecies designations were given for members of this genus.

### *Host phylogenetic reconstruction*

To infer the evolutionary history of the symbiosis with *Dasytiphilus*, we reconstructed a phylogeny of Dasytinae beetles and outgroup taxa based on the sequences of up to four marker genes: 18S rRNA gene, 28S rRNA gene, CO1 and Cytochrome b gene (cytB). Where possible, sequences were obtained from the assembled shotgun sequencing data, otherwise the respective regions were amplified by PCR and Sanger sequenced. The partial 18S rRNA gene (~1,400 bp) was amplified with primer pair 18S\_ai and 18S\_3'l. Three individual pieces were sequenced with primers 18S\_ai, 18S\_a1.0 and 18S\_3'l, respectively, and subsequently assembled to cover the complete amplified region. The partial 28S rRNA gene (~600 bp) was amplified with primer pair 28Sff and 28Srr and sequenced with 28Sff. For the CO1 gene, either one long sequence (~1,400 bp) was amplified with primer pair LepF1 and Pat, or two smaller overlapping fragments were amplified, the first one with primer pair LepF1 and LepR1 and the second one with primer pair Melyridae\_CO1\_F1 and Pat. Following this, three pieces were sequenced with primers LepR1, Melyridae\_CO1\_F1 and Pat and assembled to obtain a contiguous sequence. The partial cytB gene (~300 bp) was amplified with primer pair Sytb\_F and Sytb\_R and sequenced with Sytb\_F. Additionally, previously published sequences for some species were used (28). For each gene, an alignment was created using Geneious Prime 2023.1.2. The four alignments were concatenated and used for phylogenetic reconstruction via Bayesian analysis (61) utilizing the

MrBayes plugin v3.2.6 in Geneious Prime 2023.1.2. Through the custom command block, the alignment was partitioned by gene, and for the protein-coding genes CO1 and cytB, also by codon position (codon positions 1 and 2 in one partition, position 3 in another partition). From a random starting tree, two runs with four chains were run over 20 million generations, until the standard deviation of split frequencies was consistently below 0.01, using a 4by4 GTR model and gamma-shaped rate variation. Sampling was done every 2,000 generations, and a consensus tree was created after 25% burn-in.

## Results

### ***Dasytiphilus* is widespread but not omnipresent in Dasytinae beetles**

We screened 58 Melyridae species (Table 1), mostly from the subfamily Dasytinae, for the specific symbiont *Dasytiphilus*, confirming the presence of *Dasytiphilus* in 27 Dasytinae species via multiple different methods, whereas no symbionts were detected in 29 other Dasytinae species nor in the two investigated species of the subfamily Malachiinae. Additionally, *Dasytiphilus* was also not detected in *Eronyxa pallida* from the closely related beetle family Lophocateridae.

The 16S rRNA gene amplicon sequencing analysis of a subset of the taxa revealed that *Dasytiphilus* was present in 13 and absent in 30 Melyridae species (Figure 1, Figure S2). When symbionts were present, they always constituted over 50% of the total microbial community (range: 56% - 99%), thus making false-negative results unlikely.

## ***Dasytiphilus* is present in distinct host clades and codiversified with its**

### **hosts**

On the basis of four marker genes, we reconstructed a phylogeny of Dasytinae beetles and representative taxa from the subfamilies Melyrinae and Malachiinae (Figure 2). In our analysis, the Dasytinae formed a monophyletic clade, slightly differing from a previously published topology (28), in which the Melyrinae were nested within Dasytinae. Five distinct clades were evident in Dasytinae: 1) the “*Ameocerus*-clade”, represented by five Southern Hemisphere genera; 2) the “*Dasytastes*-clade”, represented by seven North American genera (and including species currently misplaced in *Dasytes*) and one South American genus (*Listrocerus*); 3) the “*Danacea*-clade”, represented by three Palearctic genera; 4) the “*Listrus*-clade”, represented by nine nominal North American genera; and 5) the “*Dasytes*-clade”, represented by six Palearctic genera. These five clades correspond to putative natural tribes within Dasytinae based on morphology, and contradict current nominal tribal assignments in several cases (e.g., four of the *Ameocerus*-clade genera are currently placed with *Danacea* in the Danaceini). Historical generic and tribal placement based on superficial morphological traits often does not reflect the actual relationships within the Dasytinae. The subfamily is currently undergoing revision and reclassification (for example see (62)). Hence, several genera represented here will be either transferred to new or different tribes (e.g., all *Ameocerus*-clade and *Dasytastes*-clade genera) or synonymized with other genera (e.g., *Bytuosomus*, *Cradytes*, *Eudasytes*, *Listropsis*). Furthermore, certain species in the phylogeny will either be transferred (e.g., the ‘*Dasytes*’ species in the *Dasytastes*-clade) or are as yet undescribed species (e.g., *Pseudasydates* sp.n., *Trichochrous* sp.n.). Beetles belonging to the *Dasytastes*-clade and the *Listrus*-clade were collected in North America, whereas beetles in the *Danacea*-clade and the *Dasytes*-clade originated from Europe. Species harboring *Dasytiphilus* clustered in the last four of the aforementioned clades, although no *Ameocerus*-clade species were screened. All analyzed

European Dasytinae species harbored *Dasytiphilus*. In contrast, we could not detect the symbiont in several North American taxa.

The symbiont phylogeny was created based on 49 marker genes extracted from whole genomes or almost complete draft genomes (Figure 2, Figure S3). The symbionts' closest relatives were '*Candidatus* Ishikawaella capsulata', *Buchnera aphidicola* and '*Candidatus* Blochmannia' endosymbionts. The different *Dasytiphilus* strains together formed a well-supported monophyletic clade, and the symbionts' phylogenetic relationships broadly corresponded to those of their hosts, with minor discrepancies within the four major clades. The 16S rRNA gene nucleotide similarity between different symbionts within each clade was high (within *Dasytastes*-clade 97.8%-99.3%; within *Danacea*-clade 96.0%-99.7%; within *Listrus*-clade 97.3%-99.8%; within *Dasytes*-clade 95.2%-99.9%), whereas the similarity between clades was in the range of 90-94%, regardless which clades were compared. Based on this 16S rRNA gene similarity the symbionts from different clades could be considered distinct species, with multiple strains within each clade. However, given their similar ecological niche within Dasytinae beetle hosts and to avoid confusion, we recommend using the previously published name '*Candidatus* Dasytiphilus stammeri' (36) for all *Dasytiphilus* symbionts of Dasytinae beetles. Host species affiliation can then be indicated by strain names using a four-letter code, consisting of the first letter of the host genus name and the first three letters of the host species epithet (e.g. '*Candidatus* Dasytiphilus stammeri' DPLU for the *Dasytiphilus* symbiont of *Dasytes plumbeus*) (see (13)). The similarity of symbiont sequences was also reflected in the high gene synteny between the different *Dasytiphilus* strains, with symbionts within each clade exhibiting perfect gene synteny (Figure S4), whereas individual rearrangements were observed between clades (Figure 3).

## ***Dasytiphilus* is located intracellularly in bacteriomes**

By performing FISH using symbiont-specific and general eubacterial probes on histological sections of adult beetles, *Dasytiphilus* was located intracellularly in bacteriomes in various Dasytinae (Figure 4, Figure S5). The bacteriomes are located in close proximity to the gut and the Malpighian tubules at the midgut-hindgut junction in adults of all analyzed species. Outside of the bacteriomes, no *Dasytiphilus* were found, including the ovaries (Figure S6), with the exception of a single event in which the symbionts were localized in the gut (Figure S7). Furthermore, *Dasytiphilus* could not be detected in any of the analyzed non-Dasytinae Melyridae, nor in several Dasytinae belonging to clades in which the symbionts were also not detected by PCR and sequencing (Figure S8).

## **Symbiont functional capabilities and genome evolution**

To gain insights into the functional capabilities of *Dasytiphilus* and its evolution, the genomes of symbionts associated with 20 different host species were sequenced, assembled and analyzed. Symbiont genome sizes varied between 437kb and 601kb, and GC content ranged from 28.3% to 33.8% (Figure S9), and both correlated strongly (Spearman rank correlation:  $S = 220.66$ ,  $\rho = 0.83$ ,  $p$  value  $< 0.0001$ ) (Figure S10). Coding density ranged from 84% to 93% and fell into the range of other endosymbionts (22). Differences in symbiont genome lengths were also reflected in their metabolic capabilities (Figure 5, Figure S11); however, some capabilities were universally retained. All *Dasytiphilus* strains retained glycolysis and pentose phosphate pathways, both utilizing  $\beta$ -D-Fructose-6P as the starting metabolite. However, the citrate cycle was incomplete, as the symbionts only encode genes for the enzymatic steps from 2-oxoglutarate to oxaloacetate via succinate. Similarly, the genetic repertoire for housekeeping, that is DNA replication and repair, transcription, and translation, was also reduced (Table S4).



Even though the *Dasytiphilus* strains featured heavily eroded genomes, several potentially host-relevant capabilities were retained (Supplementary File 02). All *Dasytiphilus* strains encoded genes for the complete shikimate pathway and subsequent tyrosine biosynthetic pathway (*aroF/G/H*, *aroB*, *aroQ*, *aroE*, *aroK/L*, *aroA*, *aroC*, *tyrA* and *aspC*) (Figure S11). Furthermore, all *Dasytastes*-clade symbionts and several *Listrus*-clade symbionts had almost complete lysine biosynthesis diaminopimelate pathways. Even though none of the strains had a fully complete lysine pathway (Figure S11), we postulate that it is indeed functional in all *Dasytastes*-clade symbionts and several *Listrus*-clade symbionts, hypothesizing that the missing catalytic steps encoded by *argD/dapC* (missing in *Dasytastes*-clade and *Listrus*-clade symbionts), *lysA*, and *dapE* (both absent in *Listrus*-clade symbionts) are likely taken over by multifunctional enzymes or encoded by the host. For example, a phosphoserine aminotransferase encoded by the *serC* gene can compensate for the catalytic step usually performed by the enzyme encoded by the *argD/dapC* gene (63). Concordantly, the *serC* gene was present in all analyzed *Dasytastes* and *Listrus* symbionts which otherwise had a nearly complete lysine synthesis pathway, whereas it was missing in *Danacea* and *Dasytes* symbionts that lacked the lysine pathway. The absence of *lysA* and *dapE* is likely compensated for by other symbiont or host enzymes, as hypothesized for other insect endosymbionts (64–66) (see Supplementary Results for more detailed information). Symbionts from the *Dasytastes*-clade also encoded genes for the pathways of the essential amino acids histidine, methionine, and threonine.

Beyond amino acids, *Dasytastes*-clade symbionts encoded genes for incomplete but likely functional biosynthetic pathways of the B-vitamins riboflavin (B2), pyridoxine (B6), and folate (B9).

A missing step in the enzymatic pathway for riboflavin synthesis can be taken over by an enzyme encoded by *yigL* (67), which was found in all *Dasytastes*-clade symbionts but missing in all other *Dasytiphilus*. Additionally, symbionts from the *Dasytastes*-clade are likely able to synthesize vitamin B6, and the absence of *dxs* could be compensated for by the presence of *dxr*, which

encodes DXP reductoisomerase (68). Moreover, symbionts of the *Dasytastes*-clade, with the exception of *Dasytastes* sp. 02, are presumably able to synthesize the vitamin B9 folate (see Supplementary Results for more detailed information).

Besides biosynthetic pathways encoding metabolites that can be beneficial for the host, further differences between the symbiont strains existed. Symbionts from the *Dasytastes*-clade encoded genes for urease (encoded by *ureA*, *ureB*, and *ureC*) and its auxiliary proteins (encoded by *ureD/ureH*, *ureF* and *ureG*), which catalyzes the hydrolysis of urea (69). A missing auxiliary protein encoded by *ureE* is not essential for the functioning of the urease (70,71). The urease-catalyzed hydrolysis of urea by *Dasytastes*-clade symbionts likely provides ammonia that the bacterium could use to synthesize glutamine with the help of glutamine synthetase encoded by *glnA* (72,73). Subsequently, glutamine could be further processed to carbamoyl phosphate by an enzyme complex encoded by *carA* and *carB* (74). Furthermore, the symbiont strains differed in their capabilities to synthesize metabolites important for the cell envelope. Symbiont strains of the *Dasytastes*-clade were the only strains able to produce cardiolipin. Additionally, symbionts of *P. viridicoerulea*, *D. aeratus*, *D. plumbeus* and *D. virens* lost the pathway to synthesize peptidoglycan.

#### *Evolutionary trajectories of lysine pathway genes*

The presence or absence of most biosynthetic pathways and other capabilities were homogeneous within the different clades. However, there were sporadic exceptions, for example the symbiont of *Listrus* sp. 01 lost 12 NADH-quinone oxidoreductase subunit genes. The lysine biosynthesis pathway was in varying states in the *Listrus*-clade symbionts: partial but functional in most *Listrus*-clade symbionts, non-functional because of pseudogenized genes in *Listrus* sp. 06, or completely absent in *Listrus* sp. 01. In order to gain insights into the evolutionary trajectories

of these genes in *Listrus*-clade symbionts, we performed a dN/dS analysis on genes involved in lysine biosynthesis. Several genes (*dapA*, *dapE*, *dapF*, and *thrA*) displayed significant relaxation of purifying selection in *Listrus* symbionts compared to *Dasytastes*-clade symbionts (Supplementary File 03), however, those genes were still under purifying selection.

## Discussion

Mutualistic insect-bacterial symbioses are common, with symbionts providing a large array of benefits to their hosts. Here, we report on a widespread bacterial endosymbiont within the Dasytinae subfamily of the Melyridae and elucidate its evolutionary history and dynamics. Host and symbiont phylogenies indicate long-term coevolution that originated from a single symbiont acquisition, but also several independent loss events. A comparison of 20 symbiont genomes revealed that extensive genome erosion occurred early in the evolutionary history of the symbiosis, but also identified core genes retained by all symbiont strains. The symbionts' main contribution to their hosts is likely the synthesis and provision of tyrosine to support cuticle formation, which is to our knowledge the first reported case of tyrosine supplementation in insects with pollen-feeding adults. Other potential supplementation in the form of essential amino acids, B-vitamins and nitrogen recycling likely play a role in some but not all of the taxa.

### Evolution of the symbiosis with *Dasytiphilus*

Comparing host and symbiont phylogenies provides strong evidence for co-diversification between the two symbiotic partners (Figure 2). Furthermore, it is likely that *Dasytiphilus* was acquired from a single event, considering the monophyly of the symbiont clade and the high degree of gene synteny between the different symbiont strains (Figure 3, Figure S4). The symbiont was not found outside of the Melyridae nor in the examined Malachiinae beetles

*Malachius viridulus* and *Malachius bipustulatus*. Furthermore, Buchner (9) did not report that any Melyridae outside of the Dasytinae have a bacteriome, even though it is very likely that he had access to common species from other subfamilies. Thus, the symbiosis with *Dasytiphilus* likely originated in the ancestor of the Dasytinae, or slightly later after the separation of the *Ameocercus*-clade, the earliest-branching clade of Dasytinae. Critical to determining the origin will be the screening of species from the *Ameocercus*-clade, as well as members of the subfamily Melyrinae, which were recovered as the sister-group to the *Ameocercus*-clade in previously published work (28) but sister to Malachiinae + Dasytinae in this study, and the other families of the melyrid lineage (Prionoceridae, Rhadalidae, Mauroniscidae, Phycosecidae). Since the origin of the symbiosis, multiple North American lineages lost *Dasytiphilus* (Figure 2). It is conceivable that symbiont acquisition as well as symbiont losses coincided with lifestyle changes. Adult beetles from the subfamily Dasytinae are known to feed on pollen (29), whereas the larvae are reported to be scavengers or predators (26,31,32). Conversely, members of other Melyridae subfamilies feed on arthropods throughout their entire life, sometimes being active hunters (37). Although some pollen offer a wide range of amino acids, some lack certain amino acids like tyrosine (34,35), and plant material is generally lower in nitrogenous compounds than animal tissues (75). The acquisition of *Dasytiphilus* may have facilitated a shift from predatory or scavenging lifestyle to a diet enriched in plant material by compensating nutritional deficiencies via the supplementation of amino acids and vitamins as well as the capacity to recycle urea. Unfortunately, detailed information on the natural history of North American Dasytinae is lacking, so the ecological factors explaining the loss of the symbionts and compensating for the ensuing nutritional deficiencies remain unknown. We have found no evidence of symbiont replacement or acquisitions of co-obligate symbionts, but individual events cannot be completely ruled out. Furthermore, the gut microbiota can supplement the nutritional needs of hosts, and may make a specialized symbiont obsolete. Future research is needed to expand the understanding of potential roles of other Dasytinae-associated microorganisms.

## Tyrosine provisioning as a conserved function of *Dasytiphilus*

All analyzed *Dasytiphilus* strains retained the capability to synthesize tyrosine, so it seems likely that the provisioning of this aromatic amino acid is the core function of the symbionts. Tyrosine is a pivotal precursor for metabolites necessary for the melanization and sclerotization of the cuticle (76,77). Consequently, the availability of tyrosine directly impacts the physicochemical properties of the cuticle (78). A lack of tyrosine can lead to a formation of a thinner, softer cuticle that is impaired in providing protection against biotic and abiotic stresses (11,12,17,19). Given that Dasytinae often occur in arid and semi-arid regions (26), a thick cuticle is likely important to protect the beetles against desiccation. Furthermore, their cuticle could be an effective defense or deterrent against predation, whereas symbiont-free Malachiinae evolved chemical defenses and aposematic coloration instead of relying on their poorly sclerotized and much weaker cuticle (79,80). The observation that in some *Dasytes* and *Psilothrix* species the bacteriomes seem to regress with age in adults (81) supports the hypothesis that the main contribution of the symbionts is completed as soon as the cuticle is formed (17). The tyrosine demands of a beetle peak during (late) larval and pupal stages, before the adult cuticle is about to be developed. Unfortunately, knowledge of the larval diet, and thus, dietary tyrosine availability, is scarce or absent for most genera of Dasytinae. Only for some Palearctic *Dasytes* and a *Psilothrix* species has it been reported that the larvae are predators and scavengers, or change from scavengers to herbivores as they develop, respectively (31,33,82). It is unlikely that any Dasytinae larvae have a pollen-based diet, based on the few circumstances of larval collection (in organic ground litter), and lack of larval collection in flowers despite countless collection events of adults in flowers (M.L. Gimmel, pers. obs.). However, it appears that the potentially tyrosine-rich diet of arthropod prey or carcasses still needs symbiont supplementation for the formation of the adult cuticle. Concordantly, omnivorous *Camponotus* ants utilize endosymbionts that nutritionally upgrade their

diet by supplementing tyrosine, and it is speculated that the symbiont acquisition was a major evolutionary step for this group of ants (83). Similarly, it is conceivable that this tyrosine supplementation is also needed in some Dasytinae despite the presumably carnivorous larval diet. Besides its relevance for the cuticle, tyrosine and its derivatives are also required for various other functions in insects, such as neurotransmission and immune reactions, but likely in lower quantities (84). Furthermore, there might be additional functions of tyrosine precursors in *Dasytastes*-clade symbionts. The intermediate chorismate can also be used as a precursor for folate (vitamin B9) biosynthesis, and additional amino acids and vitamins are likely supplied by the symbiont to its host.

Tyrosine-provisioning symbionts have been reported for several beetles and ants, and their contribution to host fitness is based on enhanced desiccation resistance and protection against predators and pathogens (11–16). The presence of tyrosine-providing symbionts in Dasytinae beetles indicates that tyrosine is a limiting resource across many taxa with diverse feeding habits, and the acquisition of nutritional endosymbionts is a widespread strategy to compensate for this deficiency.

### **Symbiont functions beyond tyrosine provisioning**

*Dasytastes*-clade strains retain a more diverse metabolic portfolio including several amino acid and vitamin biosynthesis pathways, as well as a urease. Thus, in addition to tyrosine and phenylalanine, *Dasytastes*-clade strains can synthesize the essential amino acids histidine, lysine, methionine, and threonine (Figure 5, Figure S11), deficiencies of which can severely lower insect fitness (85). This hints at additional supplementary needs of the *Dasytastes*-clade hosts, which therefore exhibit species-specific selective pressures on their respective symbionts. Most likely selection for the maintenance of these biosynthetic pathways weakened in non-*Dasytastes*-clade strains and thus they were rapidly lost, as the biosynthesis of some of these amino acids is especially costly (86). Comparable patterns are found in many symbiotic systems, such as in

Donaciinae (Chrysomelidae) beetles, whose symbionts retain genes encoding pectinolytic enzymes only in host species that feed on pectin-rich plants (23).

The presence of a urease suggests that symbiont-harboring beetles from the *Dasytastes*-clade are confronted with a particularly nitrogen-deficient diet. Urea is an abundant waste product of insect nitrogen metabolism, and urea recycling symbionts have been described in several insect taxa (87). For example, the *Blattabacterium* endosymbiont of cockroaches recycles host-derived urea and uses it as a nitrogen source for amino acid biosynthesis (88), and urease genes have recently been described for the symbionts of silvanid and bostrichid beetles (13,14), suggesting a role in nitrogenous waste recycling in beetles feeding on wood or stored grain products. Similarly, the urease-catalyzed hydrolysis of urea by *Dasytastes*-clade symbionts may provide ammonia that the bacterium could use to synthesize glutamine, aspartate and glutamate, important precursors for the biosynthesis several amino acids, as well as, in case of glutamate, a by-product of tyrosine biosynthesis. Alternatively, the initially synthesized glutamine might be transported directly to the host and fuel its amino acid metabolism.

Beyond essential amino acid biosynthesis and nitrogen recycling, B vitamins are another group of compounds that most insects cannot synthesize (89), yet they are pivotal for development, adult survival, and reproduction (90,91). Many insects feeding on B-vitamin deficient diets rely on endosymbionts that provide these compounds. Examples include the blood-feeding tsetse flies (92) and bedbugs (93), but also the seed-feeding cotton stainer *Dysdercus fasciatus* (94). Symbionts of *Dasytastes*-clade hosts can synthesize and potentially supplement riboflavin (B2), pyridoxine (B6) and folate (B9) to their hosts (Figure 5, Figure S11). Riboflavin can be crucial for insects during development as well as for adult survival (91). Furthermore, it is a precursor of flavin mononucleotide (FMN) and adenine dinucleotide (FAD), important cofactors for flavoproteins and flavoenzymes (95). Even though the enzymatic steps for FMN and FAD synthesis are missing in *Dasytiphilus*, they can be performed by a plethora of organisms (96).

Vitamin B6 is a relevant cofactor for many enzymes in insects (91). One of its vitamers, pyridoxal 5'-phosphate (PLP), is involved in the conversion of DOPA into dopamine (97), an essential step for cuticle sclerotization and melanization. Furthermore, PLP was shown to be involved in amino acid metabolism, cofactor biosynthesis, and cell wall metabolism, among other processes, in *Bacillus subtilis* (98). Folate is involved in the metabolism of amino acids and nucleic acids and can be crucial during insect development (91).

In addition to nutritional deficiencies in the diet, many herbivorous insects are confronted with the challenge of breaking down the plant cell wall. Palynivorous (i.e., pollen-feeding) species can utilize several mechanisms to overcome the recalcitrant pollen wall and gain access to the nutritious content (35). Additionally, cell wall polymers can be degraded using digestive enzymes provided by bacterial symbionts (23,99). Based on the pollen-feeding habit of many adult Dasytinae and the close proximity of the bacteriomes to the gut, we previously speculated that *Dasytiphilus* may contribute to the break-down of the pollen (36). However, the genomic analysis of *Dasytiphilus* genomes did not reveal any genes encoding for plant cell wall-degrading enzymes, so it seems unlikely that the symbionts contribute to the breakdown of pollen. Moreover, based on our detection of pollen in guts of Dasytinae adults from species with and without *Dasytiphilus* (Figure S8), these beetles seem to be able to access pollen nutrients without the help of *Dasytiphilus*. It is possible that Melyridae beetles access the pollen content by using osmotic shock and inducing pseudogermination, or by enlisting the help of gut bacteria that supply pectin-degrading enzymes, strategies that are also used by bees (100). Alternatively, they might encode pectinolytic enzymes in their own genomes to break open the pollen (101).



## Differences in cell envelope biosynthesis between *Dasytiphilus* strains

Our genomic analysis revealed that several symbiont strains lost the ability to produce certain cell envelope components (Figure 5), i.e., peptidoglycan and cardiolipin, a lipid that is used in the inner membrane. There are multiple potential explanations for these losses, which indicate that these deficient strains have further adapted to their life in the host environment. A complex cell envelope might not be essential anymore if the hosts provide a suitable environment. It is known that the intracellular environments of endosymbionts are often isotonic with the symbiont cytoplasm, diminishing the importance of a cell wall for maintaining turgor pressure (102). Alternatively, hosts sometimes contribute to their endosymbionts' cell wall formation with the help of horizontally acquired genes (25,103). These genes can be highly expressed in bacteriocytes harboring the endosymbiont, strongly suggesting a supportive role in, but also active control over symbiont cell envelope construction (104). Even though uncommon, it is not possible to rule out similar mechanisms in the Dasytinae symbiosis without information on the host genomes.

## *Dasytiphilus* genome evolution

Similar to other nutritional endosymbionts, *Dasytiphilus* strains exhibit eroded genomes of very small size with low GC content. The largest changes in the symbiont genomes likely occurred in the early stages of the symbiotic association. This is reflected in the high degree of genomic similarity and synteny between clades as well as a small number of inversions of parts of the symbiont genomes between clades (Figure 3). Genome size, gene content and order within each clade is almost perfectly conserved (Figure 5, Figure S4, Figure S10), providing additional evidence that most of the genome erosion happened early in the history of the symbiosis. However, the current genomes are still subject to ongoing erosion, which is likely facilitated by the reduced repertoire of DNA repair genes (Table S4). An example that might foreshadow the evolutionary trajectory of the *Dasytiphilus* genomes in closely related symbiont strains is the lysine

biosynthesis pathway in *Listrus*-clade symbionts. The different states of this pathway could indicate reduced selective pressures on some *Listrus* symbionts for the supplementation of lysine. Interestingly, the dN/dS analysis revealed that in the *Listrus*-clade strains that retained the lysine pathway genes, these genes were still under purifying selection, albeit less strongly than in *Dasytastes* strains. Therefore, the loss of these genes in these *Listrus* spp. symbionts seems unlikely, indicating that loss patterns of certain capabilities alone do not necessarily predict future gene losses in closely related strains.

### **Bacteriome structure and symbiont transmission**

Like many nutritional endosymbionts in Hemiptera, Coleoptera, Diptera, and Hymenoptera, *Dasytiphilus* is localized in specialized bacteriomes. The bacteriomes of Dasytinae are located in very close proximity to the gut. In multiple Dasytinae species accessory organs of the Malpighian tubules have been described that are touching the bacteriomes and are connected to the gut (81,105). However, these accessory organs and thus the connections from the bacteriome to the gut are absent in Dasytinae adult males (81). This sex-specific difference raises the possibility of a vertical transmission route via the gut and the surface on the eggs, as is observed for symbionts located in the gut or gut-associated organs of many Heteroptera and some Coleoptera (106,107). In accordance with an extracellular transmission route via the gut, we never observed *Dasytiphilus* in the ovaries by FISH (Figure S6); however, one *D. plumbeus* individual contained bacteria labeled with the *Dasytiphilus*-specific probe in the gut (Figure S7). A potential behavioral indicator for a transmission via the gut from the egg surface has been described for young larvae of Dasytinae beetles, which burst open the egg, then stay motionless in the egg for a few days (31,33). By doing so, they may be acquiring the symbionts from the egg surface and allowing time for the colonization of the bacteriomes via the gut. This feature is unique to Dasytinae beetles and differs from the behavior of beetles from other Melyridae subfamilies and the closely related

Rhadalidae (108). Concordantly, mature *P. viridicoerulea* larvae of both sexes have a direct connection from the gut to the bacteriome (81). This untested hypothesis of the symbiont transmission route is an intriguing scenario, as it would – to our knowledge – provide the only known example of a transmission route that involves a phase outside the insect body for a bacteriome-localized symbiont. If true, the Dasytinae symbiosis might offer unprecedented insights into the evolutionary origin of bacteriomes. To test our hypothesized transmission route, future investigations should include the localization of the symbiont during the egg stage. Furthermore, manipulations during this life stage, like surface sterilization of the eggs, could lead to aposymbiotic larvae / adults, experimentally testing for an extracellular phase of the symbionts.

## Conclusion

By investigating the taxonomic distribution, tissue localization, and putative function of the *Dasytophilus* symbiont within Dasytinae beetles, we elucidated the evolutionary history and dynamics of this symbiosis. Following a single acquisition event, four distinct host clades retained *Dasytophilus* and show clear signs of host-symbiont coevolution, whereas multiple symbiont losses occurred in independent host lineages. The bacteriome-localized symbiont contains an eroded genome, and its main contribution to the host's fitness likely resides in the supplementation of tyrosine that supports cuticle sclerotization and melanization. However, the symbionts of some host taxa probably also complement the host's diet by producing additional essential amino acids and vitamins, as well as by recycling urea. The host ecological traits correlating with the differences in symbiont-provided benefits or the complete losses of symbionts remain unknown, due to the scarcity of knowledge on the natural history and feeding ecology of Dasytinae. The close proximity of the Dasytinae bacteriomes to the digestive tract, the occasional presence of the symbionts in the gut, and their absence from the ovaries raises the intriguing possibility of an extracellular transmission route of the bacteriome-localized symbiont. As such,

the Dasytinae symbiosis provides an interesting system to study the evolutionary origin of bacteriome-localized symbionts as well as the ecological correlates of changes in symbiont presence and function, and it extends our understanding of tyrosine-supplementation as an important symbiont-provided benefit across a broad range of insect taxa with diverse nutritional ecologies.

#### ACKNOWLEDGMENTS & FUNDING

We would like to thank Benjamin Weiss for preparing the histological sections, Jean Keller for helping with the dN/dS analysis, Sonja Gerwin for support with the symbiont localization via FISH, Ana Simão Pinto de Carvalho for helping with COG analysis, and Gianfranco Liberti and the German Barcode of Life project (GBOL) for providing insect specimens. This research was funded by a Consolidator Grant of the European Research Council (ERC CoG 819585 “SYMBEETLE” to MK) and by the Max Planck Society. MLG’s work was supported in part by the Schlinger Foundation.

#### COMPETING INTERESTS

The authors declare no competing interests.

#### DATA AVAILABILITY STATEMENT

The datasets presented in this study are deposited in online repositories. The BioProject PRJNA1068458 on NCBI contains the Metagenome-Assembled Genomes (MAG) for all 20 assembled *Dasytiphilus* strains, as well as the underlying raw sequences in a SRA database (SRP485589). In addition, the annotated symbiont genomes are available in Edmond, which is a research data repository for Max Planck researchers (<https://doi.org/10.17617/3.PFIWPC>).

### Supplementary files

Supplement 01 – File including more detailed results; a figure with information on bacteriome dissection; an additional figure about the bacterial community composition in Melyridae beetles; a phylogenetic reconstruction of the different *Dasytiphilus* strains based whole (draft-) genomes; a gene synteny plot comparing the gene order between the genomes of *Dasytiphilus* symbionts falling in the *Dasytes*-clade; additional FISH pictures including the *Dasytiphilus* free ovaries, as well as *Dasytiphilus* in the gut of *Dasytes plumbeus*; a representation of the circular genome of *Dasytiphilus* of *Dasytes niger*; a figure of the correlation of genome length with GC content of *Dasytiphilus* symbionts; a detailed figure showcasing the absence and presence of individual genes from selected amino acid and B-vitamin pathways; a table containing the diagnostic and quantitative PCR primers, 16S rRNA gene microbial community amplicon primers, and FISH probes; a table giving information about the regions of the 16S rRNA gene that were amplified during microbial community analysis for each sample; a table informing about the sequencing systems and services used for metagenome sequencing of individual taxa; a table giving an

overview of genes with DNA repair function and their presence in *Dasytiphilus* symbionts; and a table informing about the number of individuals that were screened for each analyzed species.  
Supplementary file 02 – Overview of hierarchical orthologues and respective KEGG annotation  
Supplementary file 03 – Detailed results of the dN/dS analysis

### **AUTHOR CONTRIBUTIONS**

JW and MK conceived the study and designed the experiments.

JW was responsible for the execution of the experiments, as well as the data analysis.

MG provided the majority of beetle specimens and contributed his expertise on Melyridae taxonomy and phylogenetic relationships.

SH contributed to beetle dissections, DNA extraction, diagnostic PCRs, Sanger sequencing, microbial community profiling, and FISH.

MK and TE offered supervision on the experimental design and data analysis.

JW and MK drafted the manuscript and all authors participated in editing the final version.

All authors contributed to the article and approved the submitted version.

### **COMPETING INTERESTS**

The authors declare no competing interests.

## References

1. McKenna DD, Shin S, Ahrens D, Balke M, Beza-Beza C, Clarke DJ, et al. The evolution and genomic basis of beetle diversity. *Proc Natl Acad Sci USA*. 2019;116(49):24729–37.
2. Goczał J, Beutel RG. Beetle elytra: evolution, modifications and biological functions. *Biol Lett*. 2023 Mar;19(3):20220559.
3. Linz DM, Hu AW, Sitvarin MI, Tomoyasu Y. Functional value of elytra under various stresses in the red flour beetle, *Tribolium castaneum*. *Sci Rep*. 2016;6:34813.
4. Arakane Y, Noh MY, Asano T, Kramer KJ. Tyrosine metabolism for insect cuticle pigmentation and sclerotization. In: Cohen E, Moussian B, editors. *Extracellular Composite Matrices in Arthropods*. Cham: Springer International Publishing; 2016. p. 165–220.
5. Douglas AE. Multiorganismal insects: diversity and function of resident microorganisms. *Annu Rev Entomol*. 2015;60:17–34.
6. Gilbert SF, Sapp J, Tauber AI. A symbiotic view of life: we have never been individuals. *Q Rev Bio*. 2012;87(4):325–41.
7. Akman Gündüz E, Douglas AE. Symbiotic bacteria enable insect to use a nutritionally inadequate diet. *Proc R Soc B*. 2009 Mar 7;276(1658):987–91.
8. Brownlie JC, Johnson KN. Symbiont-mediated protection in insect hosts. *Trends Microbiol*. 2009;17(8):348–54.
9. Buchner P. *Endosymbiosis of animals with plant microorganisms*. New York, NY: Interscience Publ; 1965. 909 p.
10. Moran NA, Telang A. Bacteriocyte-associated symbionts of insects. *BioScience*. 1998;48(4):295–304.
11. Anbutso H, Moriyama M, Nikoh N, Hosokawa T, Futahashi R, Tanahashi M, et al. Small genome symbiont underlies cuticle hardness in beetles. *Proc Natl Acad Sci USA*. 2017 Oct 3;114(40).
12. Engl T, Eberl N, Gorse C, Krüger T, Schmidt THP, Plarre R, et al. Ancient symbiosis confers desiccation resistance to stored grain pest beetles. *Mol Ecol*. 2018 Apr;27(8):2095–108.
13. Kiefer JST, Bauer E, Okude G, Fukatsu T, Kaltenpoth M, Engl T. Cuticle supplementation and nitrogen recycling by a dual bacterial symbiosis in a family of xylophagous beetles. *ISME J*. 2023 Apr 21
14. Kiefer JST, Batsukh S, Bauer E, Hirota B, Weiss B, Wierz JC, et al. Inhibition of a nutritional endosymbiont by glyphosate abolishes mutualistic benefit on cuticle synthesis in *Oryzaephilus surinamensis*. *Communications biology*. 2021;4(1):554.
15. Jackson R, Monnin D, Patapiou PA, Golding G, Helanterä H, Oettler J, et al. Convergent evolution of a labile nutritional symbiosis in ants. *ISME J*. 2022 Sep;16(9):2114–22.

16. Klein A, Schrader L, Gil R, Manzano-Marín A, Flórez L, Wheeler D, et al. A novel intracellular mutualistic bacterium in the invasive ant *Cardiocondyla obscurior*. ISME J. 2016 Feb;10(2):376–88.
17. Vigneron A, Masson F, Vallier A, Balmand S, Rey M, Vincent-Monégat C, et al. Insects recycle endosymbionts when the benefit is over. Curr Biol. 2014;24(19):2267–73.
18. Kanyile SN, Engl T, Heddi A, Kaltenpoth M. Endosymbiosis allows *Sitophilus oryzae* to persist in dry conditions. Front Microbiol. 2023 Jul 11;14:1199370.
19. Kanyile SN, Engl T, Kaltenpoth M. Nutritional symbionts enhance structural defence against predation and fungal infection in a grain pest beetle. J Exp Biol. 2022 Jan 1;225(1):jeb243593.
20. Moran NA. Accelerated evolution and Muller's ratchet in endosymbiotic bacteria. Proc Natl Acad Sci USA. 1996;93(7):2873–8.
21. Pettersson ME, Berg OG. Muller's ratchet in symbiont populations. Genetica. 2007;130(2):199–211.
22. McCutcheon JP, Moran NA. Extreme genome reduction in symbiotic bacteria. Nat Rev Microbiol. 2012 Jan;10(1):13–26.
23. Reis F, Kirsch R, Pauchet Y, Bauer E, Bilz LC, Fukumori K, et al. Bacterial symbionts support larval sap feeding and adult folivory in (semi-)aquatic reed beetles. Nat Commun. 2020 Jun 11;11(1):2964.
24. Kelkar YD, Ochman H. Genome reduction promotes increase in protein functional complexity in bacteria. Genetics. 2013;193(1):303–7.
25. Smith TE, Lee M, Person MD, Heseck D, Mobashery S, Moran NA. Horizontal-acquisition of a promiscuous peptidoglycan-recycling enzyme enables aphids to influence symbiont cell wall metabolism. mBio. 2021;12(6):e0263621.
26. Lawrence JF, Leschen RAB. 9.11. Melyridae Leach, 1815. In: Kükenthal W, Leschen RAB, Beutel RG, Lawrence JF, Kükenthal W, Leschen RAB, et al., editors. Coleoptera, Beetles, Volume 2, Morphology and Systematics (Elateroidea, Bostrichiformia, Cucujiformia partim). Berlin, New York: DE GRUYTER; 2010.
27. Mayor AJ. 74. Melyridae Leach, 1815. In: Arnett RH, Thomas MC, Skelley PE, Frank JH, editors. American beetles, Volume II Polyphaga: Scarabaeoidea through Curculionoidea. Boca Raton: CRC Press; 2002. p. 281–303.
28. Gimmel ML, Bocakova M, Gunter NL, Leschen RAB. Comprehensive phylogeny of the Cleroidea (Coleoptera: Cucujiformia). Syst Entomol. 2019;44(3):527–58.
29. Mawdsley JR. The Importance of species of Dasytinae (Coleoptera: Melyridae) as pollinators in western North America. Coleopt Bull. 2003;57(2):154–60.

30. Napoles JR, Lopez-Buenfil JA. *Listrus senilis* (LeConte) (Coleoptera: Melyridae), a little-known pollinator species associated to wild and cultivated *Dahlia* (Asteraceae) in Mexico. *Am J Ent.* 2022;6(2):22.
31. Asano M. Egg Hatch and Early Instars of *Dasytes* (*Mesodasytes*) *vulgaris* Nakane (Coleoptera: Melyridae: Dasytinae) and comparison with the related subfamily Malachiinae. *Coleopt Bull.* 2016;70(2):288–96.
32. Harde KW. A field guide in colour to beetles. Hammond PM, editor. London: Octopus Books; 1984. 333 p.
33. Fiori G. Contributi alla conoscenza morfologica ed etologica dei Coleotteri. IX. Studi Sassaresi. 1971;19:2–70.
34. Johri BM, Vasil IK. Physiology of pollen. *Bot Rev.* 1961 Jul;27(3):325–81.
35. Roulston TH, Cane JH. Pollen nutritional content and digestibility for animals. In: Dafni A, Hesse M, Pacini E, editors. *Pollen and Pollination*. Vienna: Springer Vienna; 2000. p. 187–209.
36. Weiss B, Kaltenpoth M. Bacteriome-localized intracellular symbionts in pollen-feeding beetles of the genus *Dasytes* (Coleoptera, Dasytidae). *Front microbiol.* 2016;7:1486.
37. Asano M. Morphology and biology of *Malachius prolongatus* (Coleoptera: Melyridae: Malachiinae). *Acta Entomologica Musei Nationalis Pragae.* 2017;57(2):603–15.
38. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods.* 2016 Jul;13(7):581–3.
39. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucl Acids Res.* 2012 Nov 27;41(D1):D590–6.
40. Yilmaz P, Parfrey LW, Yarza P, Gerken J, Pruesse E, Quast C, et al. The SILVA and “All-species Living Tree Project (LTP)” taxonomic frameworks. *Nucl Acids Res.* 2014 Jan;42(D1):D643–8.
41. Arkin AP, Cottingham RW, Henry CS, Harris NL, Stevens RL, Maslov S, et al. KBase: The United States department of energy systems biology knowledgebase. *Nat Biotechnol.* 2018 Aug;36(7):566–9.
42. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 2014 Aug 1;30(15):2114–20.
43. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol.* 2012 May;19(5):455–77.
44. Li D, Liu CM, Luo R, Sadakane K, Lam TW. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct *de Bruijn* graph. *Bioinformatics.* 2015 May 15;31(10):1674–6.



45. Kolmogorov M, Bickhart DM, Behsaz B, Gurevich A, Rayko M, Shin SB, et al. metaFlye: scalable long-read metagenome assembly using repeat graphs. *Nat Methods*. 2020 Nov;17(11):1103–10.
46. Kolmogorov M, Yuan J, Lin Y, Pevzner PA. Assembly of long, error-prone reads using repeat graphs. *Nat Biotechnol*. 2019 May;37(5):540–6.
47. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*. 2014 Jul 15;30(14):2068–9.
48. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, et al. The RAST Server: rapid annotations using subsystems technology. *BMC Genom*. 2008;9:75.
49. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, et al. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS one*. 2014;9(11):e112963.
50. Kanehisa M. Toward understanding the origin and evolution of cellular organisms. *Protein Sci*. 2019 Nov;28(11):1947–51.
51. Kanehisa M, Furumichi M, Sato Y, Kawashima M, Ishiguro-Watanabe M. KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Research*. 2023 Jan 6;51(D1):D587–92.
52. Kanehisa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucl Acids Res*. 2000 Jan 1;28(1):27–30.
53. Gilchrist CLM, Chooi YH. clinker & clustermap.js: automatic generation of gene cluster comparison figures. Robinson P, editor. *Bioinformatics*. 2021 Aug 25;37(16):2473–5.
54. Price MN, Dehal PS, Arkin AP. FastTree 2—approximately maximum-likelihood trees for large alignments. Poon AFY, editor. *PLoS ONE*. 2010 Mar 10;5(3):e9490.
55. Emms DM, Kelly S. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol*. 2019 Dec;20(1):238.
56. Edgar RC. Muscle5: High-accuracy alignment ensembles enable unbiased assessments of sequence homology and phylogeny. *Nat Commun*. 2022 Nov 15;13(1):6968.
57. Suyama M, Torrents D, Bork P. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucl Acids Res*. 2006 Jul 1;34(Web Server):W609–12.
58. Wertheim JO, Murrell B, Smith MD, Kosakovsky Pond SL, Scheffler K. RELAX: detecting relaxed selection in a phylogenetic framework. *Mol Biol Evol*. 2015 Mar;32(3):820–32.
59. Smith MD, Wertheim JO, Weaver S, Murrell B, Scheffler K, Kosakovsky Pond SL. Less is more: an adaptive branch-site random effects model for efficient detection of episodic diversifying selection. *Mol Biol Evol*. 2015 May;32(5):1342–53.
60. Hebert PDN, Ratnasingham S, deWaard JR. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proc Biol Sci*. 2003;270 Suppl 1:S96-9.

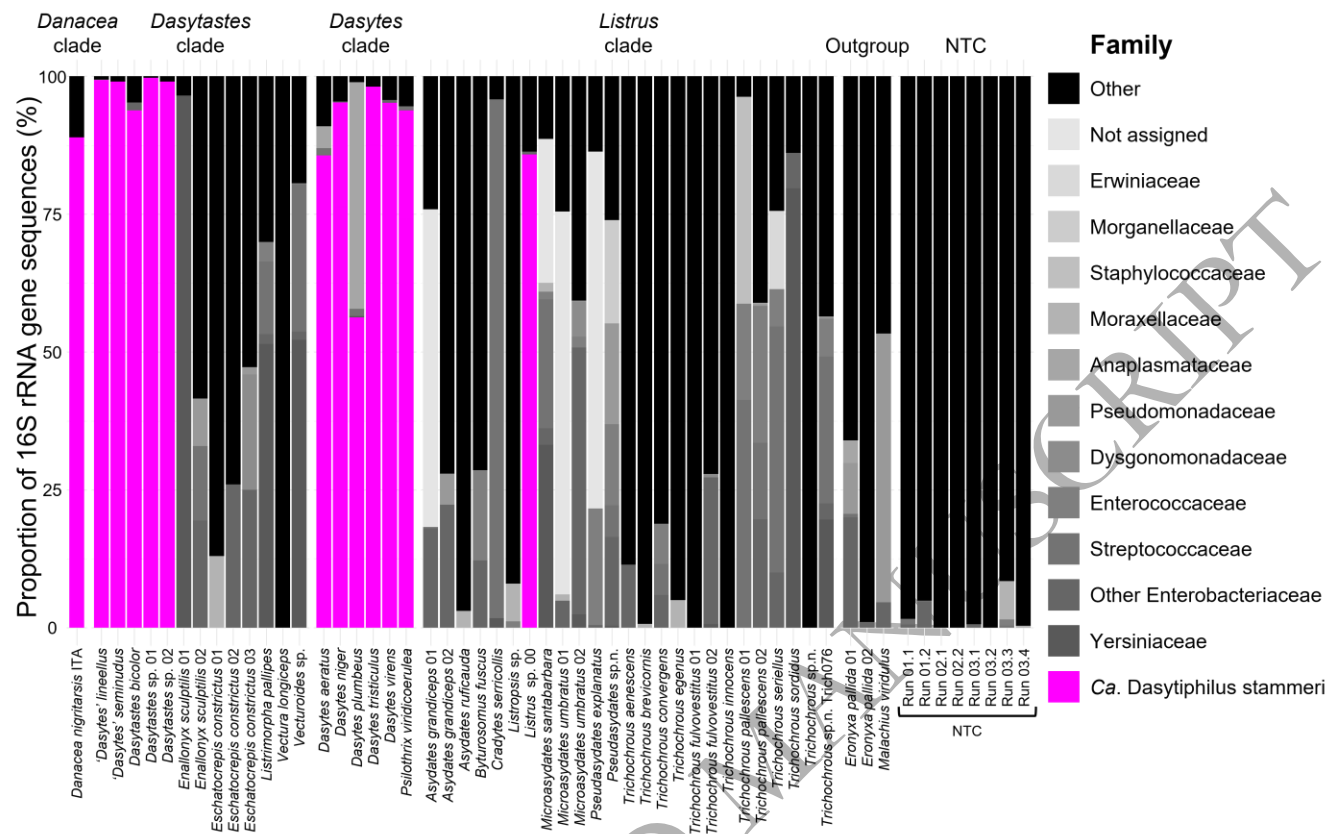
61. Huelsenbeck JP, Ronquist F. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics*. 2001;17(8):754–5.
62. Gimmel ML, Mayor AJ. On the North American species of *Dolichosoma* Stephens and *Psilothrix* Küster (Coleoptera: Melyridae: Dasytinae: Dasytini). *Coleopt Bull*. 2019;73(2):395.
63. Lal PB, Schneider BL, Vu K, Reitzer L. The redundant aminotransferases in lysine and arginine synthesis and the extent of aminotransferase redundancy in *Escherichia coli*: Aminotransferase redundancy. *Mol Microbiol*. 2014 Nov;94(4):843–56.
64. Bao XY, Yan JY, Yao YL, Wang YB, Visendi P, Seal S, et al. Lysine provisioning by horizontally acquired genes promotes mutual dependence between whitefly and two intracellular symbionts. Jiggins FM, editor. *PLoS Pathog*. 2021 Nov 29;17(11):e1010120.
65. Husnik F, Nikoh N, Koga R, Ross L, Duncan RP, Fujie M, et al. Horizontal gene transfer from diverse bacteria to an insect genome enables a tripartite nested mealybug symbiosis. *Cell*. 2013 Jun;153(7):1567–78.
66. Liao S, Poonpairaj P, Ko KC, Takatuska Y, Yamaguchi Y, Abe N, et al. Occurrence of Agmatine Pathway for Putrescine Synthesis in *Selenomonas ruminatum*. *Biosci Biotechnol Biochem*. 2008 Feb 23;72(2):445–55.
67. Manzano-Marín A, Simon JC, Latorre A. Reinventing the wheel and making it round again: evolutionary convergence in *Buchnera* – *Serratia* symbiotic consortia between the distantly related Lachninae aphids *Tuberolachnus salignus* and *Cinara cedri*. *Genome Biol Evol*. 2016 May;8(5):1440–58.
68. Takahashi S, Kuzuyama T, Watanabe H, Seto H. A 1-deoxy- D -xylulose 5-phosphate reductoisomerase catalyzing the formation of 2- C -methyl- D -erythritol 4-phosphate in an alternative nonmevalonate pathway for terpenoid biosynthesis. *Proc Natl Acad Sci USA*. 1998 Aug 18;95(17):9879–84.
69. Patra AK, Aschenbach JR. Ureases in the gastrointestinal tracts of ruminant and monogastric animals and their implication in urea-N/ammonia metabolism: A review. *J Adv Res*. 2018;13:39–50.
70. Brayman TG, Hausinger RP. Purification, characterization, and functional analysis of a truncated *Klebsiella aerogenes* UreE urease accessory protein lacking the histidine-rich carboxyl terminus. *J Bacteriol*. 1996 Sep;178(18):5410–6.
71. Lee MH, Pankratz HS, Wang S, Scott RA, Finnegan MG, Johnson MK, et al. Purification and characterization of *Klebsiella aerogenes* UreE protein: A nickel-binding protein that functions in urease metallocenter assembly. *Protein Sci*. 1993 Jun;2(6):1042–52.
72. Kumada Y, Benson DR, Hillemann D, Hosted TJ, Rochefort DA, Thompson CJ, et al. Evolution of the glutamine synthetase gene, one of the oldest existing and functioning genes. *Proc Natl Acad Sci USA*. 1993 Apr;90(7):3009–13.
73. Woolfolk CA, Shapiro B, Stadtman ER. Regulation of glutamine synthetase. *Arch Biochem Biophys*. 1966;116:177–92.

74. Rubino SD, Nyunoya H, Lusty CJ. In vivo synthesis of carbamyl phosphate from NH<sub>3</sub> by the large subunit of *Escherichia coli* carbamyl phosphate synthetase. *J Biol Chem*. 1987 Mar;262(9):4382–6.
75. Fagan WF, Siemann E, Mitter C, Denno RF, Huberty AF, Woods HA, et al. Nitrogen in insects: implications for trophic complexity and species diversification. *Am Nat*. 2002 Dec;160(6):784–802.
76. Brunet PCJ. The metabolism of the aromatic amino acids concerned in the cross-linking of insect cuticle. *Insect Biochem*. 1980 Jan;10(5):467–500.
77. Kramer KJ, Hopkins TL. Tyrosine metabolism for insect cuticle tanning. *Arch Insect Biochem Physiol*. 1987 Dec;6(4):279–301.
78. Hackman RH. Chemistry of the insect cuticle. In: *The Physiology of Insecta*. Elsevier; 1974. p. 215–70.
79. Majer K. A review of the classification of the Melyridae and related families (Coleoptera, Cleroidea). *Entomol Basil*. 1994;17:319–90.
80. Krzelj S. Structure anatomique comparée des élytres de Coléoptères. *Annales de la Société Royale Zoologique de Belgique*. 1969;99:85–109.
81. De Marzo L. Nuovi dati anatomici e citologici sul presunto simbiosoma dei Coleotteri Meliridi. *Entomologica*. 1994;28:77-91 Paginazione.
82. Constantin R. Descriptions des larves d' *Enicopus pyrenaicus* Fairmaire et de *Danacea pallipes* (Panzer). Contribution à l'étude de la biologie et de la systématique larvaire des Melyridae (Coleoptera). *Nouv Revue Ent*. 1989;6(4):387–405.
83. Feldhaar H, Straka J, Krischke M, Berthold K, Stoll S, Mueller MJ, et al. Nutritional upgrading for omnivorous carpenter ants by the endosymbiont *Blochmannia*. *BMC Biol*. 2007 Dec;5(1):48.
84. Nappi AJ, Christensen BM. Melanogenesis and associated cytotoxic reactions: applications to insect innate immunity. *Insect Biochem Mol Biol*. 2005 May;35(5):443–59.
85. Friend WG. Nutritional requirements of phytophagous insects. *Annu Rev Entomol*. 1958;3(1):57–74.
86. Kaleta C, Schäuble S, Rinas U, Schuster S. Metabolic costs of amino acid and protein production in *Escherichia coli*. *Biotechnol J*. 2013 Sep;8(9):1105–14.
87. Ren X, Guo R, Akami M, Niu C. Nitrogen acquisition strategies mediated by insect symbionts: a review of their mechanisms, methodologies, and case studies. *Insects*. 2022 Jan 12;13(1):84.
88. Sabree ZL, Kambhampati S, Moran NA. Nitrogen recycling and nutritional provisioning by *Blattabacterium*, the cockroach endosymbiont. *Proc Natl Acad Sci USA*. 2009 Nov 17;106(46):19521–6.
89. Douglas AE. The B vitamin nutrition of insects: the contributions of diet, microbiome and horizontally acquired genes. *Curr Opin Insect Sci*. 2017 Oct;23:65–9.

90. Hickin ML, Kakumanu ML, Schal C. Effects of *Wolbachia* elimination and B-vitamin supplementation on bed bug development and reproduction. *Sci Rep*. 2022 Jun 17;12(1):10270.
91. Serrato-Salas J, Gendrin M. Involvement of microbiota in insect physiology: focus on B vitamins. Weiss B, Yount J, editors. *mBio*. 2023 Feb 28;14(1):e02225-22.
92. Michalkova V, Benoit JB, Weiss BL, Attardo GM, Aksoy S. Vitamin B<sub>6</sub> generated by obligate symbionts is critical for maintaining proline homeostasis and fecundity in tsetse flies. Drake HL, editor. *Appl Environ Microbiol*. 2014 Sep 15;80(18):5844–53.
93. Hosokawa T, Koga R, Kikuchi Y, Meng XY, Fukatsu T. *Wolbachia* as a bacteriocyte-associated nutritional mutualist. *Proc Natl Acad Sci USA*. 2010 Jan 12;107(2):769–74.
94. Salem H, Bauer E, Strauss AS, Vogel H, Marz M, Kaltenpoth M. Vitamin supplementation by gut symbionts ensures metabolic homeostasis in an insect host. *Proc R Soc B*. 2014 Dec 7;281(1796):20141838.
95. Massey V. The chemical and biological versatility of riboflavin. *Biochem Soc Trans*. 2000 Aug 1;28(4):283–96.
96. Abbas CA, Sibirny AA. Genetic control of biosynthesis and transport of riboflavin and flavin nucleotides and construction of robust biotechnological producers. *Microbiol Mol Biol Rev*. 2011 Jun;75(2):321–60.
97. Raboni S, Spyrakis F, Campanini B, Amadasi A, Bettati S, Peracchi A, et al. Pyridoxal 5'-Phosphate-dependent enzymes: catalysis, conformation, and genomics. In: *Comprehensive Natural Products II*. Elsevier; 2010. p. 273–350.
98. Richts B, Rosenberg J, Commichau FM. A survey of Pyridoxal 5'-Phosphate-dependent proteins in the gram-positive model bacterium *Bacillus subtilis*. *Front Mol Biosci*. 2019 May 10;6:32.
99. Salem H, Bauer E, Kirsch R, Berasategui A, Cripps M, Weiss B, et al. Drastic genome reduction in an herbivore's pectinolytic symbiont. *Cell*. 2017 Dec;171(7):1520-1531.e13.
100. Engel P, Martinson VG, Moran NA. Functional diversity within the simple gut microbiota of the honey bee. *Proc Natl Acad Sci USA*. 2012 Jul 3;109(27):11002–7.
101. Shin NR, Shin S, Okamura Y, Kirsch R, Lombard V, Svacha P, et al. Larvae of longhorned beetles (Coleoptera; Cerambycidae) have evolved a diverse and phylogenetically conserved array of plant cell wall degrading enzymes. *Syst Entomol*. 2021 Oct;46(4):784–97.
102. Otten C, Brill M, Vollmer W, Viollier PH, Salje J. Peptidoglycan in obligate intracellular bacteria. *Mol Microbiol*. 2018;107(2):142–63.
103. Bublitz DC, Chadwick GL, Magyar JS, Sandoz KM, Brooks DM, Mesnage S, et al. Peptidoglycan production by an insect-bacterial mosaic. *Cell*. 2019;179(3):703-712.e7.

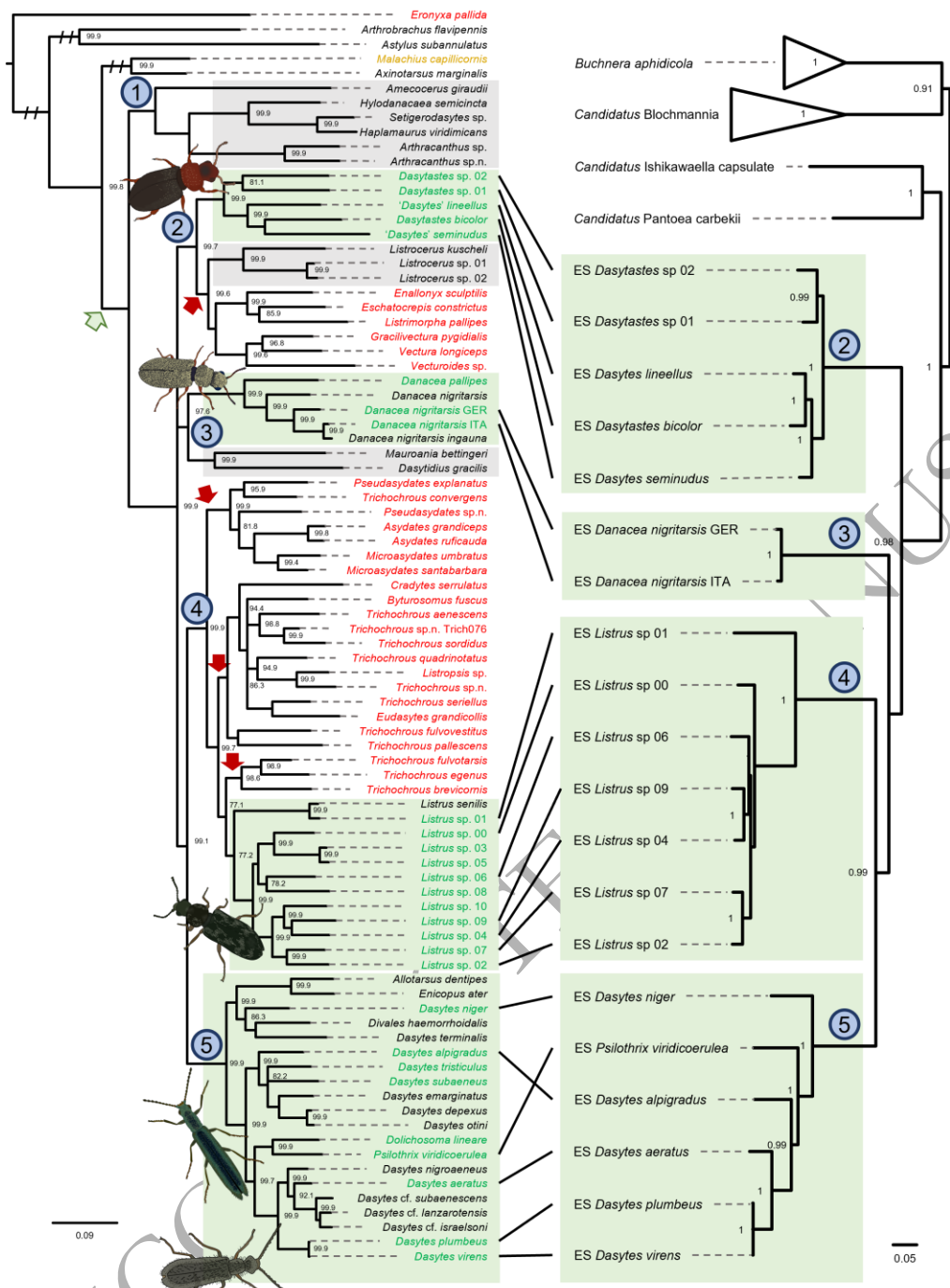
104. Smith TE, Li Y, Perreau J, Moran NA. Elucidation of host and symbiont contributions to peptidoglycan metabolism based on comparative genomics of eight aphid subfamilies and their *Buchnera*. *PLoS Genet*. 2022;18(5):e1010195.
105. Holmgren N. Über die Exkretionsorgane des *Apion flavipes* und *Dasytes niger*. *Anat Anz*. 1902;(22):225–39.
106. Salem H, Florez L, Gerardo N, Kaltenpoth M. An out-of-body experience: the extracellular dimension for the transmission of mutualistic bacteria in insects. *Proc R Soc B*. 2015;282(1804):20142957.
107. Salem H, Kaltenpoth M. Beetle–bacterial symbioses: endless forms most functional. *Annu Rev Entomol*. 2022 Jan 7;67(1):201–19.
108. Asano M. Egg hatch and first instar of *Semijulistus spectabilis* (Lewis) (Coleoptera: Rhadalidae) and diversification of the ontogenic development of the melyrid lineage. *Coleopt Bull*. 2019 Sep 22;73(3):501.

UNCORRECTED MANUSCRIPT



**Figure 1:** Bacterial community composition in Melyridae beetles given in relative abundance of bacterial amplicon sequence variants (ASVs) determined at family level by DADA2 analysis of Illumina 16S rRNA gene amplicons. Samples with less than 1,000 reads after removal of reads assigned to chloroplasts and mitochondria were excluded. Every bar represents a single individual, with DNA extracted from the whole body. The Dasytinae-specific symbiont *Dasytiphilus* is highlighted in magenta. Only the 50 most abundant bacterial ASVs are displayed with annotated family, remaining ASVs are grouped as “Other.” NTC = No template extraction controls.

UNCORRECTED

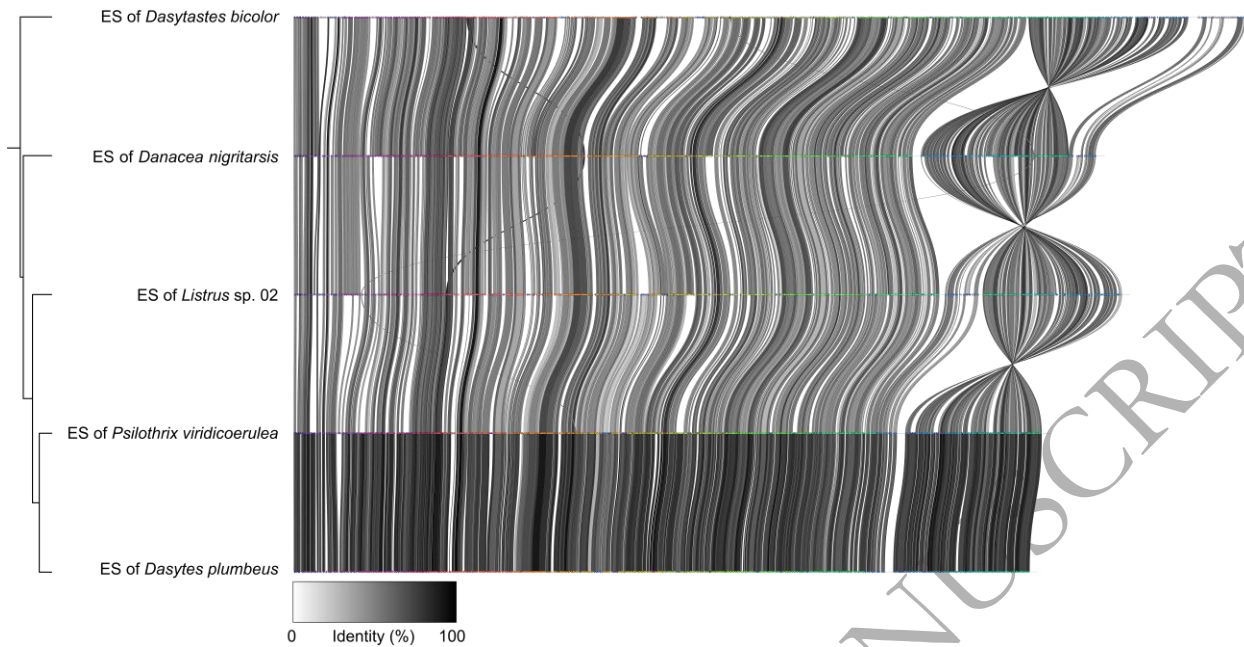


**Figure 2:** Evolutionary history of Dasytinae beetles and their endosymbiont *Dasytiphilus stammeri*. The phylogeny on the left shows the relationships of Melyridae beetles based on a partitioned Bayesian analysis of COI, cytb, 18S and 28S genes. Five distinct clades were evident in the Dasytinae: 1) the “*Amecocerus*-clade”; 2) the “*Dasytastes*-clade”; 3) the “*Danaceae*-clade”; 4) the “*Listrus*-clade”; and 5) the “*Dasytes*-clade”. Taxa in green font carry the Dasytinae-specific symbiont *Dasytiphilus*, whereas no evidence for the symbiont’s presence was found for taxa in red font. Taxa in black font were not screened for the symbiont, but were added to the phylogeny to better resolve phylogenetic relationships. Although *Malachius capillicornis* in particular was not screened for the symbiont, the orange font serves as a reminder that other *Malachius* species were investigated and showed no signs of *Dasytiphilus* presence. Green and red arrows indicate potential symbiont acquisition and loss events, respectively, during the evolutionary history of

the Melyridae. Green boxes highlight clades in which all screened species harbored *Dasytiphilus* and in which presumably all host taxa are infected. Grey boxes indicate clades for which the symbiont status of the host species is unknown. Node labels indicate support values, based on Bayesian posterior probabilities, with values below 75 removed. Beetle illustrations depict representative species of the symbiont-containing clades, i.e., *Dasytastes bicolor*, *Danacea nigratarsis*, *Listrus* sp., *Dolichosoma lineare*, and *Dasytes plumbeus* (from top to bottom). The right side depicts the phylogenetic relationships of the different *Dasytiphilus* strains based on a set of 49 COG. The phylogeny was reconstructed using an approximately-maximum-likelihood algorithm, and node labels indicate local support values. Outgroups are the close relatives '*Candidatus Pantoea carbekii*', '*Candidatus Ishikawaella capsulata*', *Buchnera aphidicola*, and '*Candidatus Blochmannia*' endosymbionts. Connections between the phylogenies highlight host-symbiont associations.

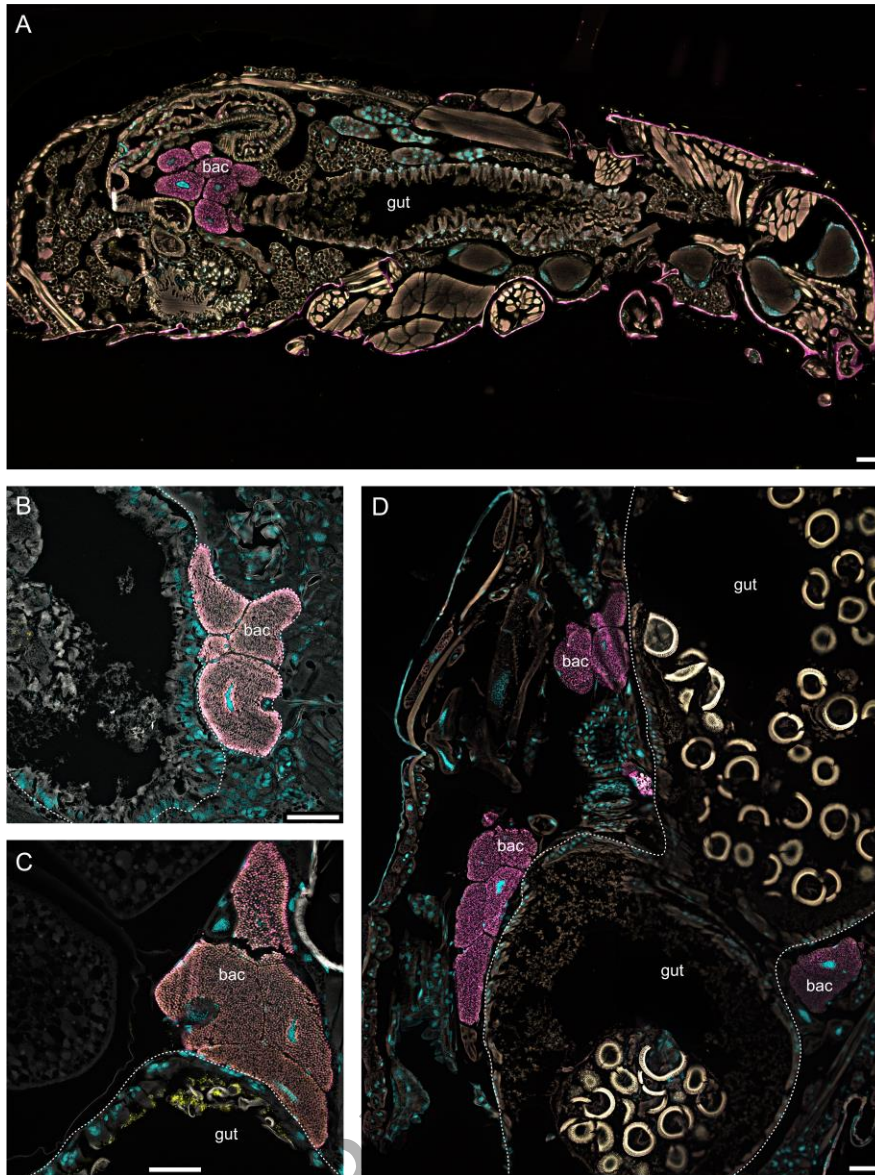
UNCORRECTED MANUSCRIPT





**Figure 3: Genome synteny of *Dasytaphilus* representatives from each infected host clade**

Gene synteny plot, comparing the gene order between the *Dasytaphilus* endosymbionts of *Dasytastes bicolor* (*Dasytastes*-clade), *Danacea nigritarsis* ITA (*Danacea*-clade), *Listrus* sp. 02 (*Listrus*-clade), *Psilothrix viridicoerulea* (*Dasytes*-clade), and *Dasytes plumbeus* (*Dasytes*-clade). The gene identity percentage of homologous proteins is based on amino acid sequences and indicated by different grey values. The phylogeny on the left is based on a set of 49 COG and was reconstructed using an approximately-maximum-likelihood algorithm.



**Figure 4:** Tissue localization of *Dasytiphilus* symbionts in the abdomen of *Listrus* sp. 01 (A), *Danacea nigritarsis* GER (B), *Dasytes plumbeus* (C) and *Dasytes tristiculus* (D), as revealed by fluorescence *in situ* hybridization sagittal sections. *Dasytiphilus* (in magenta) are densely packed in the bacteriocytes (bac) that were located in close proximity to the gut and the Malpighian tubules at the midgut-hindgut junction in adults of all analyzed species. General bacteria (in yellow) are sometimes visible in the gut (C). Background autofluorescence is given in white, and a general DNA counterstain in cyan (DAPI). Pictures shown are overlaps of all four channels. Consumed pollen was often visible in the gut (C and D, with a strong auto-fluorescence in the cy3 channel), confirming the pollen-feeding habit of the beetles. Scale bars = 50  $\mu$ m.



**Table 1:** List of Melyridae (and one Lophocateridae) species screened for the Dasytinae-specific symbiont *Dasytiphilus stammeri*. Taxa are ordered by clade in the host phylogeny, separated by empty grey rows, and then alphabetically. The presence of *Dasytiphilus* was assessed by PCR, Sanger sequencing, FISH, and Illumina 16S rRNA gene amplicon profiling; empty cells indicate that this method was not used to assess symbiont presence in the respective host species. “Genome info” gives information about the results of Illumina shotgun sequencing of the symbiont genome, with length in base pairs and the GC content.

Host		Collection	Symbiont presence confirmation method					Genome info		
Genus	Species	Country	Symbiont present	PCR	Sanger	FISH	Amplicon	Genome available	Genome length (bp)	GC %
<i>Eronyxa</i>	<i>pallida</i>	USA, CA	no	no	no		no			
<i>Malachius</i>	<i>bipustulatus</i>	Germany	no			no				
<i>Malachius</i>	<i>viridulus</i>	USA, CA	no				no			
<i>Dasytastes</i>	<i>bicolor</i>	USA, CA	yes	yes			yes	yes	589,892	33.0
<i>Dasytes</i>	<i>lineellus</i>	USA, CA	yes	yes	yes		yes	yes	585,061	32.6
<i>Dasytes</i>	<i>seminudus</i>	USA, CA	yes	yes	yes		yes	yes (draft)	~586,000	32.8
<i>Dasytastes</i>	sp. 01	USA, CA	yes	yes	yes		yes	yes	600,652	33.8
<i>Dasytastes</i>	sp. 02	USA, CA	yes	yes	yes		yes	yes	592,705	33.1
<i>Enallonyx</i>	<i>sculptilis</i>	USA, CA	no	no	no		no			
<i>Eschatocrepis</i>	<i>constrictus</i>	USA, CA	no	no	no	no	no			
<i>Gracilivectura</i>	<i>pygidialis</i>	USA, CA	no	no	no	no	no			
<i>Listrimorpha</i>	<i>pallipes</i>	USA, CA	no	no	no		no			
<i>Vectura</i>	<i>longiceps</i>	USA, CA	no	no	no		no			
<i>Vecturoides</i>	sp.	USA, CA	no				no			
<i>Danacea</i>	<i>nigritarsis</i> GER	Germany	yes	yes	yes	yes		yes	497,568	28.8
<i>Danacea</i>	<i>nigritarsis</i> ITA	Italy	yes	yes	yes	yes	yes	yes	497,621	28.8
<i>Danacea</i>	<i>pallipes</i>	Germany	yes	yes	yes					
<i>Asydates</i>	<i>grandiceps</i>	USA, CA	no	no	no		no			
<i>Asydates</i>	<i>ruficauda</i>	USA, CA	no	no	no		no			
<i>Byturosomus</i>	<i>fuscus</i>	USA, CA	no	no	no		no			
<i>Cradytes</i>	<i>serricollis</i>	USA, NM	no	no	no		no			
<i>Cradytes</i>	<i>serrulatus</i>	USA, AZ	no	no	no		no			
<i>Eudasytes</i>	<i>grandicollis</i>	USA, NV	no	no	no		no			
<i>Listropsis</i>	sp.	USA, CA	no	no	no		no			
<i>Listrus</i>	sp. 00	USA, CA	yes	yes	yes		yes	yes (draft)	~511,000	32.5
<i>Listrus</i>	sp. 01	USA, AZ	yes	yes		yes		yes	440,755	28.3
<i>Listrus</i>	sp. 02	USA, CA	yes	yes				yes	513,178	33.3
<i>Listrus</i>	sp. 03	USA, NV	yes	yes	yes					
<i>Listrus</i>	sp. 04	USA, UT	yes	yes				yes (draft)	~515,000	32.9
<i>Listrus</i>	sp. 05	USA, CA	yes	yes	yes					
<i>Listrus</i>	sp. 06	USA, CA	yes	yes				yes	509,826	32.6
<i>Listrus</i>	sp. 07	USA, CA	yes	yes				yes (draft)	~461,000	33.3
<i>Listrus</i>	sp. 08	USA, CA	yes	yes	yes					
<i>Listrus</i>	sp. 09	USA, OR	yes	yes				yes	518,020	32.9
<i>Listrus</i>	sp. 10	USA, OR	yes	yes	yes	yes				
<i>Microasydates</i>	<i>santabarbara</i>	USA, CA	no	no	no		no			
<i>Microasydates</i>	<i>umbratus</i>	USA, CA	no	no	no		no			
<i>Pseudasydates</i>	<i>explanatus</i>	USA, CA	no	no	no		no			
<i>Pseudasydates</i>	sp.n.	USA, NM	no	no	no		no			
<i>Trichochrous</i>	<i>aenescens</i>	USA, CA	no	no	no		no			
<i>Trichochrous</i>	<i>brevicornis</i>	USA, CA	no	no	no		no			
<i>Trichochrous</i>	<i>convergens</i>	USA, AZ	no	no	no		no			
<i>Trichochrous</i>	sp.n.	USA, CA	no	no	no		no			
<i>Trichochrous</i>	<i>egenus</i>	USA, CA	no	no	no		no			
<i>Trichochrous</i>	<i>fulvotarsis</i>	USA, CA	no	no	no		no			
<i>Trichochrous</i>	<i>fulvovestitus</i>	USA, CA	no	no	no		no			
<i>Trichochrous</i>	<i>pallescens</i>	USA, CA	no	no	no	no	no			
<i>Trichochrous</i>	<i>quadrinotatus</i>	USA, CA	no	no	no		no			
<i>Trichochrous</i>	<i>seriellus</i>	USA, UT	no	no	no		no			
<i>Trichochrous</i>	<i>sordidus</i>	USA, CA	no	no	no		no			
<i>Trichochrous</i>	sp.n. Trich076	USA, CA	no	no	no		no			

<i>Dasytes</i>	<i>aeratus</i>	Germany	yes	yes	yes		yes	yes	455,894	30.9
<i>Dasytes</i>	<i>alpigradus</i>	Italy	yes	yes	yes			yes	476,890	30.0
<i>Dasytes</i>	<i>niger</i>	Germany	yes	yes	yes	yes	yes	yes	486,045	30.9
<i>Dasytes</i>	<i>plumbeus</i>	Germany	yes	yes	yes	yes	yes	yes	456,731	30.4
<i>Dasytes</i>	<i>subaeneus</i>	Germany	yes	yes	yes					
<i>Dasytes</i>	<i>tristiculus</i>	Italy	yes	yes	yes	yes	yes			
<i>Dasytes</i>	<i>virens</i>	Germany	yes	yes	yes	yes	yes	yes	456,754	30.4
<i>Dolichosoma</i>	<i>lineare</i>	Germany	yes	yes	yes	yes				
<i>Psilothrix</i>	<i>viridicoerulea</i>	Germany	yes	yes	yes	yes	yes	yes	464,075	28.7

UNCORRECTED MANUSCRIPT