

School of Doctoral Studies in Biological Sciences  
University of South Bohemia in České Budějovice  
Faculty of Science  
Department of Ecosystem Biology

**With a little help from my friends:  
“understanding the roles and importance of the  
millipede gut microbiome”**

Ph.D. Thesis

**Julius Eyiuche Nweze**

Supervisor: Dr Roey Angel

Institute of Soil Biology and Biogeochemistry, Biology Centre  
CAS

České Budějovice 2024





### **This thesis should be cited as:**

Nweze, J. E., 2024: With a little help from my friends: “understanding the roles and importance of the millipede gut microbiome”. Ph.D. Thesis Series, No. z. University of South Bohemia, Faculty of Science, Department of Ecosystem Biology, School of Doctoral Studies in Biological Sciences, České Budějovice, Czech Republic, yy pp.

### **Annotation**

This research aimed to elucidate the role and importance of the millipede gut microbiome in cellulose digestion by using inhibitors to disrupt potential host-symbiosis and assessing their effects on millipede digestion and overall health. It involved the first comprehensive profiling of microbial communities within the hindgut and faeces of two distinct millipede species: *Epibolus pulchripes*, a tropical species found on the East African coast, and *Glomeris connexa*, a temperate species native to Central Europe. Although both species share a similar detritivorous lifestyle, they differ in size and gut redox conditions, with *G. connexa* being smaller (10-17 mm) than *E. pulchripes* (130-160 mm). The study also revealed the potential of the hindgut bacterial community in breaking down complex polysaccharides and recycling nutrients. It described the active bacterial community vital for certain processes and the extent of the millipedes' dependence on them. Additionally, the research provided a comprehensive investigation of viral communities in the hindguts of the two millipedes and their role in enhancing metabolism and modulating microbial composition. Furthermore, it introduced a new perspective that millipedes primarily ingest litter to gain access to microbial biomass (primarily fungal), which they and their gut microbiota consume.

## **Declaration**

I hereby declare that I have written the presented thesis by myself using the literature provided in the list of references.

.....

Julius Eyiuche Nweze

České Budějovice, 15.01.2024

This thesis originated from a partnership of the Faculty of Science, University of South Bohemia, and Institution of Soil Biology and Biogeochemistry, Biology Centre CAS, supporting doctoral studies in the Ecosystem Biology program.



Přírodovědecká  
fakulta  
Faculty  
of Science

Jihočeská univerzita  
v Českých Budějovicích  
University of South Bohemia  
in České Budějovice



BIOLOGICKÉ  
CENTRUM  
AV ČR, v. v. i.



Institute of Soil Biology and Biogeochemistry

### **Financial support**

The research presented in this Ph.D. thesis was supported by by a Junior Grant of the Czech Science Foundation (GA ČR grant no. 19-24309Y).

### **Acknowledgements**

I want to express my deepest gratitude to my supervisor, Dr. Roey Angel. His teachings, support, patience, and personalised approach have been invaluable, particularly the opportunity he provided to develop my bioinformatics skills. I greatly thank everyone who guided me through my PhD study, offering advice, assistance, and steadfast support when necessary. Special thanks are due to Eva Petrová and Dr. Ana C. Lara-Rodriguez, who welcomed me to the laboratory and imparted molecular microbiology skills to me. Further appreciation is extended to Lucie Faktorová, Dr. Terézia Horváthová, and

Šárka Otáhalová for their assistance with millipede collection and maintenance.

I am also grateful to Dr. Shruti Gupta, with whom cooperation was a genuine pleasure. Her deep involvement in all my experiments helped ensure the journey was not overly arduous. Sincere thanks are also extended to Dr. Michaela Salcher for instructing me in the principles of FISH and CARD-FISH. Dr. Vladimír Šustr deserves particular mention for his help with millipede dissection and manuscript corrections.

I am grateful to the Biology Centre CAS (grant: IBERA) and the European Community Action Scheme for the Mobility of University Students (grant: CZ CESKE01) for providing grants for my stay at Prof Andreas Brune's lab at the Max Plank Institute of Terrestrial Microbiology in Marburg, Germany. Profound thanks go to Prof Brune for the opportunity to work under his guidance on my data analysis, results and first manuscript. The internship was an important and enjoyable experience that significantly facilitated my professional development.

Thanks are also due to the Department of Ecosystem Biology and the Faculty of Science members for maintaining a warm and welcoming atmosphere. The annual faculty conferences were events I always eagerly anticipated.

Finally, I cannot overstate my gratitude to my family for their unwavering support and continual encouragement throughout my years of study. Their love and encouragement were instrumental in keeping me focused on my path.

## List of publications and author's contribution

The thesis is based on the following manuscripts:

I. **Julius E. Nweze**, Vladimír Šustr, Andreas Brune, Roey Angel (2024). Functional similarity despite taxonomical divergence in the millipede gut microbiota points to a common trophic strategy. *Microbiome* 12, 16. <https://doi.org/10.1186/s40168-023-01731-7>

IF 15.5

JN was responsible for the collection of millipede samples, execution of laboratory experiments, and the completion of bioinformatics analyses. Furthermore, JN wrote the draft manuscript (contribution 70%).

II. **Julius E. Nweze**, Johannes Sergej Schweichhart, Roey Angel (2024). Viral communities in millipede guts: Insights into the diversity and potential role in modulating the microbiome. *Environmental Microbiology*, 26(2),e16586. <https://doi.org/10.1111/1462-2920.16586>

IF 5.5

JN was responsible for bioinformatics analyses and writing the draft manuscript (contribution 60%).

III. **Julius E. Nweze**, Shruti Gupta, Michaela Salcher, Vladimír Šustr, Terézia Horváthová, Roey Angel (2023). Cellulose fermentation by the gut microbiome is likely not essential for the nutrition of millipedes (Manuscript).

JN undertook the responsibility of millipede collection, checking the effect of antibiotics and methane inhibitor, bioinformatics analyses and writing the manuscript (contribution 60%).

## Co-author agreement

Roey Angel, the supervisor of this Ph.D. thesis and the corresponding author of the listed manuscripts acknowledge the contribution of JN as stated above

.....  
Roey Angel, Ph.D.

Vladimír Šustr, the co-author of paper I acknowledge the contribution of JN as stated above:

.....  
Vladimír Šustr, Ph.D.

Prof. Andreas Brune, the co-author of paper I acknowledge the contribution of JN as stated above:

.....  
Andreas Brune, Ph.D.

Johannes Schweichhart, the co-author of paper II acknowledge the contribution of JN as stated above:

.....  
Johannes Schweichhart, M.Sc.

Shruti Gupta, the co-author of paper III acknowledge the contribution of JN as stated above:

.....  
Shruti Gupta, Ph.D.

Michaela Salcher, the co-author of paper III acknowledge the contribution of JN as stated above:

.....  
Michaela Salcher, Ph.D.

Terézia Horváthová, the co-author of paper III acknowledge the contribution of JN as stated above:

.....  
Terézia Horváthová, Ph.D.

## Table of Contents

1 Introduction.....	1
1.1 General morphology and characteristics of millipedes.....	1
1.2 Millipede taxonomy and features.....	2
1.3 Millipedes in terrestrial habitats.....	4
1.4 General feeding biology.....	5
1.5 General ecological functions.....	6
1.5.1 Millipedes as efficient detritivores.....	6
1.5.2 The contribution of millipedes to the soil carbon cycle.....	7
1.5.3 The contribution of millipedes to the soil nitrogen cycle.....	8
1.5.4 The contribution of millipedes to the soil phosphorus cycle.....	9
1.5.5 Greenhouse gas emissions.....	9
1.6 Gut structure.....	10
1.6.1 Millipede guts as microhabitats.....	11
1.6.2 Physicochemical factors.....	13
1.7 Microorganisms.....	15
1.7.1 Bacteria.....	16
1.7.2 Archaea.....	18
1.7.3 Yeasts.....	19
1.7.4 Other fungi.....	20
1.7.5 Millipede gut flagellates and ciliates.....	20
1.7.6 Nematodes.....	21
1.7.7 Viruses.....	22
2 Objectives.....	24
3 Results and Discussion.....	25
3.1 Summary of the results derived from Paper No. 1.....	25
3.2 Summary results derived from Paper No. 2.....	28
3.3 Summary results derived from Manuscript No. 3.....	30
4 Conclusion and future prospects.....	33
4.1 Conclusions derived from Paper No. 1.....	33
4.2 Conclusions derived from Paper No. 2.....	34
4.3 Conclusions derived from Paper No. 3.....	35
4.4 The millipede microbiome – future research directives.....	36
References.....	37
5 Attached publications.....	47
5.1 Paper No. 1.....	47
5.2 Paper No. 2.....	75
5.3 Paper No. 3.....	106
6 Curriculum vitae.....	159

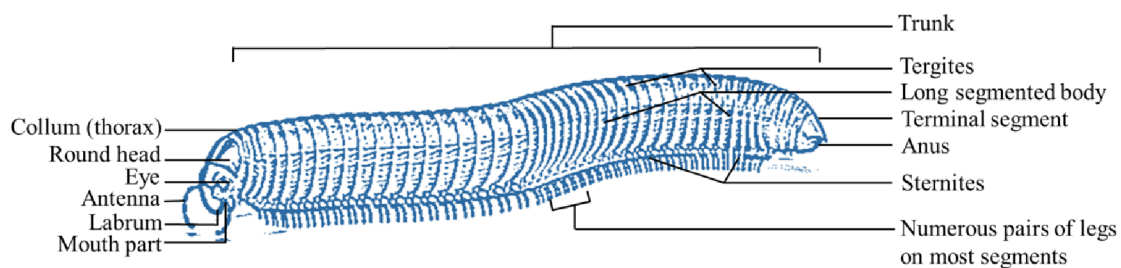




# 1 Introduction

## 1.1 General morphology and characteristics of millipedes

Millipedes are myriapods with short heads, long segmented bodies, and numerous pairs of legs, their most obvious feature (Chitty, 2022). The heads are flattened below and rounded above; the first three body segments comprise the thorax. The rest is the abdomen (Fig. 1). Two body segments fused to form diplo-segments, containing double pairs of spiracles with internal pouch openings linked to the trachea (Hopkin & Read, 1992). The eyes comprise several simple flat lens ocelli arranged in a group on the front/side of the head (Müller et al., 2007). They move using their antennae, which continually taps the ground as millipede moves along. Behind their antenna is also a pair of oval-shaped sensory organs, probably used to measure the humidity in their surroundings (Minelli & Golovatch, 2001).



*Fig. 1: The external body parts of a millipede.*

All millipedes share the general morphology described above but can nevertheless be subdivided based on their body types, including the wedge type, globular/roller type, bulldozer/rammer type, borer type, and soft bark/bristle type (Fig. 2) (Hopkin & Read, 1992; Sridhar & Ashwini, 2016).

For instance, the rammer or bulldozer body type is found in *Epibolus pulchripes* (Spirobolida) (Koch, 2015), while *Glomeris connexa* possess globular or roller body type (Glomerida) (Rosenberg, 2006). They have an unfastened heart that passes through the entire body and aorta, stretching into the head (Rajulu, 1971). Also, they have two pairs of excretory organs called malpighian tubules, which are present in the mid-part of the guts (Farquharson, 1974). Gonophores and vulvae are the sex organs in the male and female, respectively (Minelli & Golovatch, 2013).

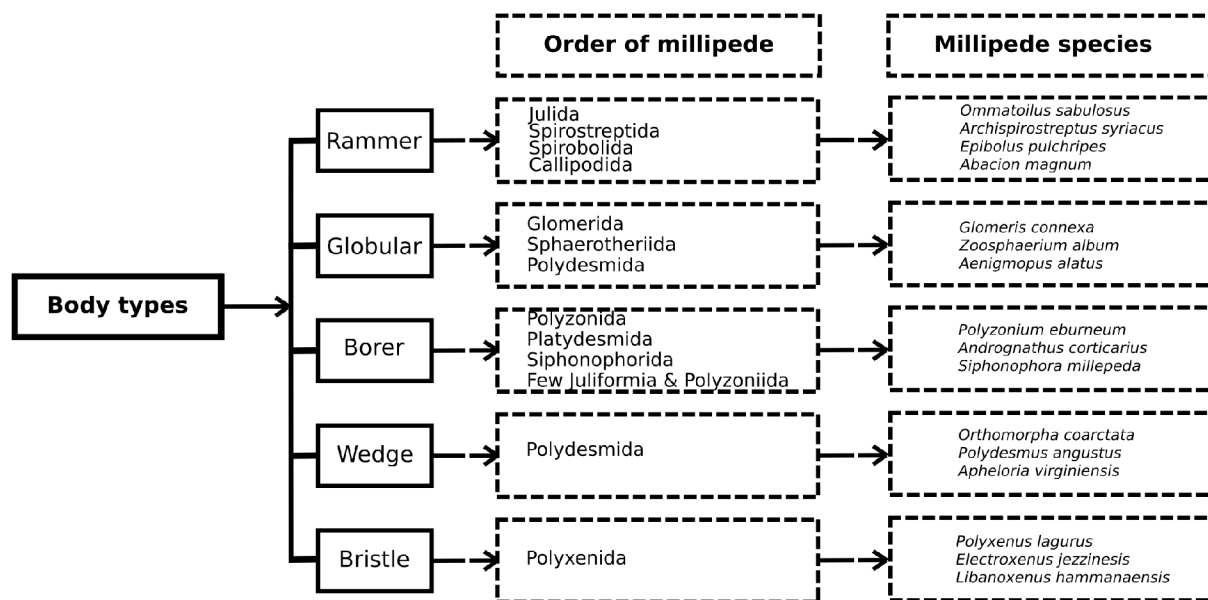


Fig. 2: Different body types in millipedes.

## 1.2 Millipede taxonomy and features

Millipedes comprise the most populous class within the subphylum Myriapoda and the phylum Arthropoda, boasting approximately 12,000 described species (Brewer & Bond, 2013; Sierwald & Bond, 2007). The class Diplopoda is divided into two subclasses, 16 orders and 145 families (Fig. 3). Fifty-three

families, comprising approximately 217 genera and 915 nominal species, are present in North America (which includes the US and Canada). However, numerous additional families are yet to be described (Shelley, 2003). While there is one single order, Polyxenida (bristle millipedes), in the basal subclass Penicillata, others belong to the subclass Chilognatha containing two infraclasses: Pentazonia (short-bodied pill millipedes) and Helminthomorpha (worm-like millipedes) (Shear et al., 2011). Pentazonia have a relatively compact body size and encompass the superorder Oniscomorpha (capable of rolling into a ball) as well as the order Glomeridesmida (which lacks the ability to roll into a ball) (Ax, 2000; Shelley, 2011).

Pentazonia resides within the prominent millipede subclass Chilognatha, distinguished by its calcified exoskeleton (Enghoff & Minelli, 1990). Colobognatha is a clade of Helminthomorpha comprising four orders that exhibit several traits in common, such as two pairs of simple leg-like gonopods, tubular defensive glands, a narrow head, and no more than two pairs of ocelli (Shear, 2011; Blanke & Wesener, 2014; Koch, 2015). Another noteworthy superorder within the clade Helminthomorpha is Juliformia, comprising three extant orders: Julida, Spirobolida, and Spirostreptida. These orders exhibit elongated cylindrical bodies with sclerites fused into complete rings (Wilson, 2006). Though much work has been done in this area over the years, the taxonomy of the millipedes is still somewhat controversial (Brewer et al., 2012). For instance, certain authors lump species into as few as three genera (Shear, 2011), revealing one facet of the taxonomic debate surrounding millipedes. This diversity within millipede classifications highlights the contentious nature of their taxonomy, accentuating an array of higher-level groups. The challenges persist in estimating species diversity due to inconsistent taxonomic efforts spanning temporal, geographic, and

phylogenetic scales. Knowledge gaps among millipede groups worsen classification controversies (Brewer et al., 2012).

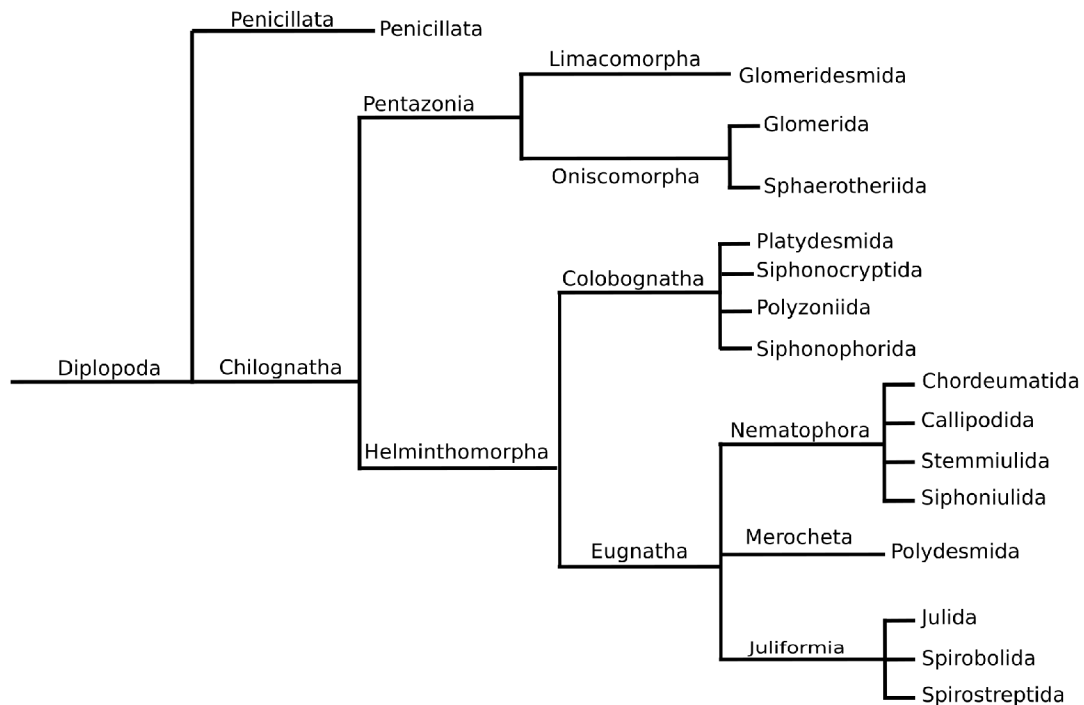


Fig. 3: Strict consensus of the millipede orders. Adapted from (Shear, 2011).

### 1.3 Millipedes in terrestrial habitats

Millipedes inhabit every continent except Antarctica (Mbenoun Masse et al., 2018). They have a wide distribution across diverse terrestrial ecosystems, spanning forests, grasslands, farmlands, urban green spaces, and residential areas (Hashimoto et al., 2004). Few millipede species exhibit extensive natural distributions; the majority are often local endemics found in single caves, islands, valleys, or mountains. This contrasts with the vast diversity of Diplopoda, estimated at over 80,000 species, primarily concentrated in tropical countries. There are few locations worldwide where the local diplopod

population surpasses two dozen species; one such instance is a patch of rainforest in central Amazonia where 33 millipede species coexist. (Golovatch & Kime, 2009; Alagesan, 2016).

Millipedes typically inhabit forest floors, where they can find ample food, shelter, and moderate moisture levels, which helps them keep their water balance (Bogyó et al., 2015). Noteworthy millipede species are found in tropical and temperate terrestrial habitats, including *Epibolus pulchripes* (Spirobolida) along the East African coast (Enghoff, 2010) and *Glomeris connexa* (Glomerida), a common species in Central Europe (Hoess & Scholl, 2001).

## **1.4 General feeding biology**

Most terrestrial millipedes feed on decomposing vegetation, faeces, or organic matter mixed with soil (Coulis et al., 2013). They can consume approximately 10% to 20% of conifer litter daily, potentially accounting for up to 36% of the annual litter fall (Cárcamo et al., 2000). Millipedes are selective feeders, showing a preference for specific types of leaf litter, such as litter with high calcium contents, while avoiding those high in polyphenols. This behaviour contributes to the processing of approximately 15%–20% of the calcium input into hardwood forest floors (Benckiser, 1997; Coleman et al., 2004). Assimilation efficiency varies across studies, ranging from 5% to 50% (David, 2014). Consequently, millipedes prioritise non-structural plant compounds during the early stages of digestion, leading to a significant portion of the ingested plant material remaining undigested (Cárcamo et al., 2000; Gillon & David, 2001; Rawlins et al., 2006). Millipedes break down plant litter in their guts and excrete it as round pellets of leaf fragments, bacteria, fungi, and algae, aiding in microorganism decomposition (Hopkin & Read, 1992).

Millipedes exhibit coprophagia, which involves consuming faecal matter, including their own, from other species or individuals (Weiss, 2006). Certain millipedes are obligate coprophages, believed to be closely linked with essential microorganisms for food digestion, resulting in faeces with elevated pH, moisture content, and bacterial counts compared to un-ingested leaf litter (McBrayer, 1973). Coprophagy may be linked to increased microbial activity and decomposition (Hashimoto et al., 2004). It has been suggested that geophagy, the consumption of soil, contributes to coprophagy in specific millipede species (Mwabvu, 1998). Coprophagy is considered a survival strategy among millipedes; in the case of cockroaches and termites, it is associated with hindgut fermentation systems (Nalepa et al., 2001).

## **1.5 General ecological functions**

### **1.5.1 Millipedes as efficient detritivores**

Millipedes play important roles in soil and litter ecosystems in tropical and temperate regions, contributing to the breakdown and decomposition of leaf litter, thus facilitating nutrient cycling in the soil (Crawford, 1992; Alagesan, 2016). Their importance in soil processes has been acknowledged for decades. Notably, litter fragmentation correlates with the size and structure of mandibles; millipede species possessing large mandibles can graze on larger litter particles, whereas others can only feed on finer fragments (Kaneko, 1988; Kheirallah, 1990). Detailed observations of millipedes' mouthparts reveal their capacity to mechanically break down the plant cells of the ingested litter, including the microorganisms associated with the litter (David, 2015). This process increases the surface area of the litter, providing microorganisms easier access to their food sources and thereby accelerating decomposition (Toyota et al., 2006). Furthermore, litter breakdown leads to increased

availability of glucose and other substances, resulting in elevated early respiration rates of microorganisms in the soil (Suzuki et al., 2013). The leaf litter consumed by millipedes is digested within their gut and secreted as pellets containing leaf waste, bacteria, fungi, and algae, supporting decomposition processes carried out by microorganisms (Alagesan, 2016).

### **1.5.2 The contribution of millipedes to the soil carbon cycle**

Soil carbon storage is essential to maintaining ecosystem functions and mitigating climate change (Bot & Benites, 2005). In millipedes, the carbon cycling process is influenced by the digestion, absorption, and excretion of substantial amounts of faecal pellets obtained from plant material consumption. There are two contrasting perspectives on how millipedes impact the soil carbon cycle (Wang et al., 2018). The conventional viewpoint suggests that millipedes crush plant material, transforming it into faecal pellets, thereby increasing the specific surface area for microbial activity. This was believed to accelerate the process of carbon mineralisation (Scheu & Wolters, 1991). However, recent research conducted over the last decade contradicts this notion. It revealed that this conversion does not expedite carbon mineralisation but contributes to soil carbon stability. Furthermore, observations indicate that faecal pellets' decomposition rate is slower than litter's (Suzuki et al., 2013).

Millipedes alter soil structure, organic matter, and mineral composition through their locomotion and burrowing activities. These actions increase soil permeability, improve aeration and water-retention capacity, facilitate root penetration, and prevent surface crusting and topsoil leaching (Chakravarthy & Sridhara, 2016). Additionally, millipede faecal pellets form soil aggregates and humus, enhancing soil quality and nutrient retention while promoting the mixing of mineral and organic soil fractions (Culliney, 2013).

### **1.5.3 The contribution of millipedes to the soil nitrogen cycle**

Millipedes are also believed to play a role in the nitrogen cycle (Cortes et al., 2018). They thrive on nitrogen-poor leaf litter with a high carbon-to-nitrogen ratio and likely struggle to obtain enough nitrogen from their diets. They are suggested to acquire additional nitrogen from symbiotic microbes in their hindguts to supplement their diet (Nardi et al., 2002). Hence, the decrease in the carbon-to-nitrogen ratio in faecal pellets in contrast to litter is credited to the decomposition and absorption of soluble compounds in the gut, alongside a several-hundredfold rise in bacterial (both dead and living) presence in the faeces compared to leaf litter (David, 2014).

A study suggests that millipedes at different developmental stages have diverse effects on the nitrogen cycle. The larvae boost nitrogen leaching from the soil, whereas adults augment the soil's nitrogen content. This is possibly due to adult millipedes encouraging the integration of fragmented litter into the soil through gut processes linked to feeding on litter and minimal assimilation (Toyota & Kaneko, 2012). Millipedes return substantial nitrogen to the soil through ingestion and secretion, primarily as ammonia (Bocock, 1963). They enhance soil nitrification by releasing ammonia-rich faecal material (Fujimaki et al., 2010). Their activities influence nitrogen mobilisation, especially in nitrogen-limited ecosystems (Symstad et al., 1998). While millipedes predominantly excrete ammonia and uric acid, the role of uric acid is still debated, and the potential contribution of symbiotic microbes to nitrogen cycling remains understudied (Nardi et al., 2002). Further research into millipedes' unquantified nitrogen cycle contributions could enhance our understanding of soil nitrification and improve nitrogen management in ecosystems (Cortes et al., 2018).



#### **1.5.4 The contribution of millipedes to the soil phosphorus cycle**

In addition to their role in litter decomposition and soil structure maintenance, millipedes impact the phosphorus cycle in ecosystems (Smit et al., 2001). However, research on the effects of millipedes on phosphorus conversion is relatively limited compared to their influence on the nitrogen cycle (da Silva et al., 2017). Notably, high population densities of millipede species like *Glyphiulus granulatus* were shown to increase soil phosphorus availability, thereby also accelerating the release of essential elements such as magnesium, potassium, nitrogen, and carbon in ecosystems, particularly in areas undergoing vegetation regeneration (Smit et al., 2001).

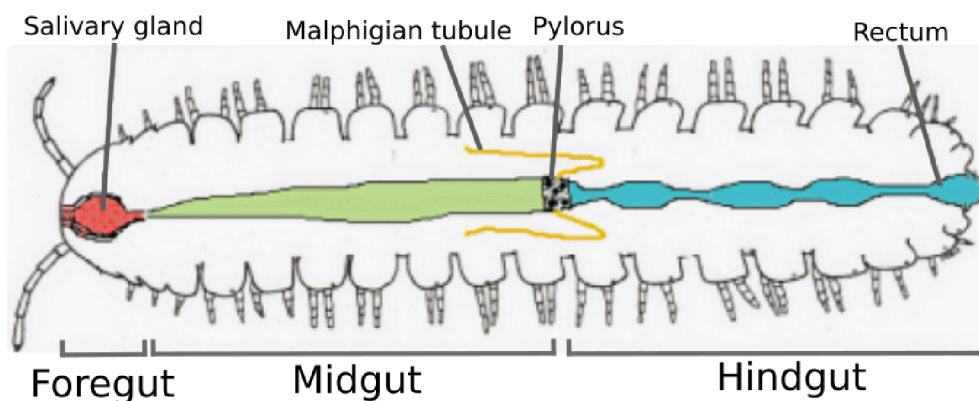
#### **1.5.5 Greenhouse gas emissions**

The soil fauna holds the potential to significantly impact the spatial and temporal variations of soil greenhouse gas sources and sinks (Šustr et al., 2020). Studies on the contribution of millipedes to greenhouse gas emissions are limited. Research by Šustr *et al.* (2020) has shed light on the greenhouse gas emissions associated with millipede activities. While most millipede species produce CO<sub>2</sub>, some are capable of CH<sub>4</sub> emission. This behaviour appears taxon-specific and is particularly common among tropical millipedes in Spirobolida and Spirostreptida order. According to the authors, the emission of these gases can be influenced by factors such as leaf quality and feeding regime, while CO<sub>2</sub> production primarily reflects the metabolic processes of millipedes. Nitrous oxide (N<sub>2</sub>O), a potent greenhouse gas, is emitted in trace amounts, mainly by members of the Glomeridae family, and its production appears to correlate with the nitrogen content of their food (Šustr et al., 2020). These findings suggest that millipedes, particularly those in tropical regions, may play a role in the global methane budget, but the exact contribution

remains uncertain. For example, tropical species like *Epibolus pulchripes* have been estimated to produce substantial amounts of CH<sub>4</sub>. However, varying population densities among habitats and challenges in maintaining certain species under laboratory conditions make it difficult to provide precise global estimates of millipede-related CH<sub>4</sub> emissions. Consequently, any assessment of the overall impact of millipedes on the global methane budget remains somewhat speculative (Šustr & Šimek, 2009).

## 1.6 Gut structure

The millipede's alimentary canal is a straight tube starting from the mouth and ending at the anus (Fig. 4), divided into foregut, midgut, and hindgut (Nunez & Crawford, 1977; Shukla & Shukla, 1980). Except for the midgut, assumed to be the primary site of digestion, the entire tract has an internal cuticular lining (Nardi, Bee, et al., 2016). The foregut is flanked by salivary glands, which produce lubricating secretions with digestive enzymes (Moreira-de-Sousa et al., 2016). The junction between the midgut and hindgut features the malpighian tubules (Moreira De Sousa & Silvia Fontanetti, 2012). The hindgut is internally lined with cuticles, regionally differentiated, and potentially allows exchange between the millipede's hemocoel and the lumen (Nardi, Bee, et al., 2016). The hindgut is basic (pH 8.0–9.0), the midgut is acidic (pH 5.0–6.0), and the salivary glands and foregut have a slightly basic pH (6.5–7.5 and 7.0–7.5 respectively) (Nunez & Crawford, 1976). The midgut is believed to be the site of digestion, with enzymes secreted by epithelial cells and possibly by microorganisms in the lumen (Hopkin & Read, 1992; Nunez & Crawford, 1977).



*Fig. 4: A typical digestive system of millipede.*

It may also play a role in synthesising compounds like protein, lipids, and calcium transport (Moreira-de-Sousa et al., 2016). The hindgut has a cuticular surface lined with polarised scales (Nardi, Bee, et al., 2016).

### **1.6.1 Millipede guts as microhabitats**

The alimentary tracts of millipedes are small ecosystems that provide hospitable and multifaceted environments for diverse assemblages of microorganisms (Innsbruck, 1992). The host's enzymes, microorganisms, and their extracellular enzymes can intricately interact in the guts. Millipedes can select from their gut and breed rare soil microorganisms that differ from the soil and leaf litter microbiome (Nardi, Bee, et al., 2016). A microbial survey from the intestine has found a diverse microbiota that includes trichomycetes, other fungi, bacteria, yeasts, and archaea on the cuticular surface secreted by the hindgut epithelium (Byzov, 2006; Rosenberg, 2006). Regional variations in the surface topography within the hindgut of a given millipede are reflected in differing and diverse microbial assemblages (Nardi, Bee, et al., 2016).

Microorganisms poorly populate the foregut of millipedes. Conversely, the midgut is believed to constitute the absorptive surface where the semi-

permeable peritrophic membrane is continually secreted. Lined with cuticles, the hindgut is highly developed and bears flat cuticular surfaces and multiple-shaped spines that provide microbial colonisation sites (Crawford et al., 1983). Nardi, Bee, and Taylor (2016) observed that microorganisms in two millipede species they studied were found in the gut lumina along the entire digestive tract, with the highest microbial densities occurring in the hindguts. Trichomycetes inhabit only the anterior third of the hindgut, and the posterior third is occupied by scattered clusters of filamentous bacteria together with their lower adherent microbes (Wright, 2011). The densest microbial communities inhabit the hindgut's core region. Whereas microbial films are adherent to the cuticle that lines the hindgut, the foregut and midgut are mostly inhabited by unattached microbes (Shukla & Shukla, 1980). Furthermore, the identification of COG (Clusters of Orthologous Groups) associated with lipid transport and metabolism in the millipede hindgut (Sardar et al., 2022) suggests the presence of genes supporting the absorption of fat-soluble nutrients. It has been suggested that millipedes are not adequately equipped with specialised enzymes for digesting lignocellulose. Therefore, to digest their food, they are thought to depend on their gut microflora (Dhivya & Alagesan, 2017). Moreover, they are known to be poor assimilators and generate copious amounts of faecal matter rich in organic material with a C/N ratio lower than that in the undigested litter (Ambarish & Sridhar, 2016).

### **1.6.2 Physicochemical factors**

The alimentary systems of millipedes are small ecosystems with a broad variety of microhabitats that differ in their abiotic and biotic environment (Dhivya & Alagesan, 2017). Many of the environmental characteristics are

intrinsic to the gut, whereas others result from the physiological activities of the host or the microbial residents in the respective locations. Both the biotic and the abiotic environments affect the physicochemical condition in the various gut compartments (Rosenberg, 2006). Studies have shown that the gut environment of well-studied arthropods like termites is not strictly anaerobic but rather characterised by moderately reduced conditions (microaerophilic) (Boga et al., 2003; Wertz et al., 2012). The redox potential of the various microhabitats in millipedes is modulated by their oxygen status, the production of redox-active compounds such as hydrogen or ferrous iron (as in soil-feeding termites), or intestinal pH variations (Byzov, 2006). The gut microbial communities maintain an oxic-anoxic gradient with lowered pH and redox potential. A low redox potential environment means that the use of microbe-accessible carbohydrate is via the relatively energetically inefficient fermentative metabolism. At a minimum, the change in physicochemical conditions of gut compartments will be selective for particular species (Paoletti et al., 2013; Šustr, Stingl, et al., 2014).

Measurements of the oxygen concentration or redox conditions in the gut of millipedes are rare. Bignell (1984) measured the mean redox potential in the gut of the pill millipede, *G. marginata* (temperate species; body mass approximately 0.2 g) from +267 to +307 mV in the midgut, and +167 to +277 mV in the hindgut, which corresponds to oxidative conditions. Within the gut lumen of two tropical millipede species, *Archispirostreptus gigas* and *Epibolus pulchripes*, notably reducing conditions were observed (ranging between -114 and -243 mV) throughout the entire intestinal tract. The redox potential was recorded as the lowest in the posterior midgut (-242 and -243 mV for *E. pulchripes* and *A. gigas*, respectively), steadily increasing along the hindgut section (Horváthová et al., 2021).

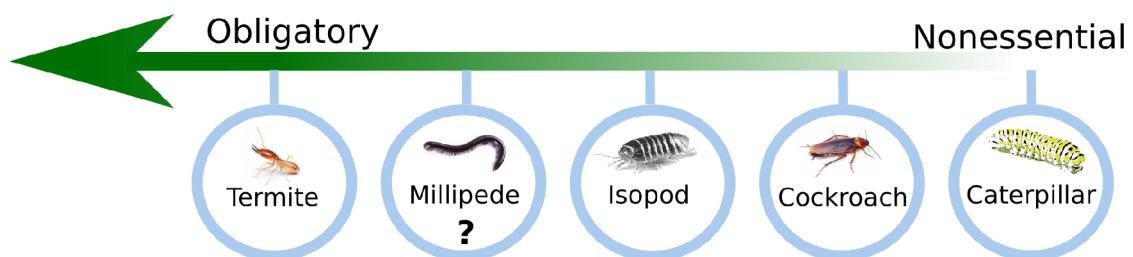
Variations in physicochemical conditions within the lumen of different gut compartments can also be attributed to pH differences (Engel & Moran, 2013). For instance, the intestinal pH profiles of *E. pulchripes* and *A. gigas* showed acidity in the midgut (pH 4.4 to 6.1), shifting to a slightly alkaline state (pH 7.3 to 7.9) in the hindgut after passing through the pyloric region (Horváthová et al., 2021).

Host secretions, nutrients absorbed by the millipede midgut's epithelial cells and stored as glycogen in nearby hepatic cells, are believed to influence the microbiota in different gut compartments (Nardi, Miller, et al., 2016). For example, the glucose concentrations in the midgut and hindgut of *Pachyiulus flavipes* are 1.1 g/l and 2.3 g/l, respectively. In *Rossiulus kessleri*, the glucose concentration in the midgut fluid is 1.28 g/l. Such relatively elevated glucose levels are similar to those of the culture media and could suggest that carbohydrates are actively being hydrolysed in the digestive tracts of the millipede (Byzov, 2006). In the anterior part of the guts, a study suggested that the digestive enzymes in saliva and midgut secretions not only provide sugars or amino acids as substrates for the resident microbiota but also digest microbial biomass (Byzov et al., 1998). The data on millipede digestive tract enzymatic activities have been reviewed (Hopkin & Read, 1992), and many researchers have found in millipede enzymes that can digest lipids, proteins, and simple carbohydrates (Guru et al., 2013; Kaplan & Hartenstein, 1978; Marcuzzi & Lafisca, 1975; Nunez & Crawford, 1976).

## **1.7 Microorganisms**

All other arthropods, including millipedes, are known to host a diverse community of microorganisms (Degli Esposti & Martinez Romero, 2017), but the degree of dependency differs widely among distinct groups (Fig. 5). For

example, a well-studied insect in the infraorder Isoptera (termite) have been recognised as being essentially dependent on the microorganisms in their guts, especially for cellulose fermentation or humus digestion, and without which they cannot develop or survive (Brune, 2014). On the other hand, the caterpillar larvae (Lepidoptera) have been reported to entirely lack gut microorganisms and can completely survive or develop when it is removed by antibiotic treatment (Hammer et al., 2017). However, millipede dependency on their microbiome has been understudied.



*Fig. 5: Dependence of the representative arthropods on their gut microbiota.*

Understanding the potential symbiotic relationship between millipedes and their microbial communities is currently limited, and the specific functions of individual microbial species within the host (Fig. 6) remain unclear. Previous studies on microbial communities in millipede digestive tracts have primarily relied on conventional cultivation methods and microscopic observations. Achieving a comprehensive understanding of the structure and dynamics of the microbial community in the invertebrates' intestinal tract necessitates a holistic approach that transcends single methodologies. Consequently, it would be helpful to conduct cross-analyses that integrate results from various methods to shed light on this complex relationship (da Silva Correia et al., 2018).

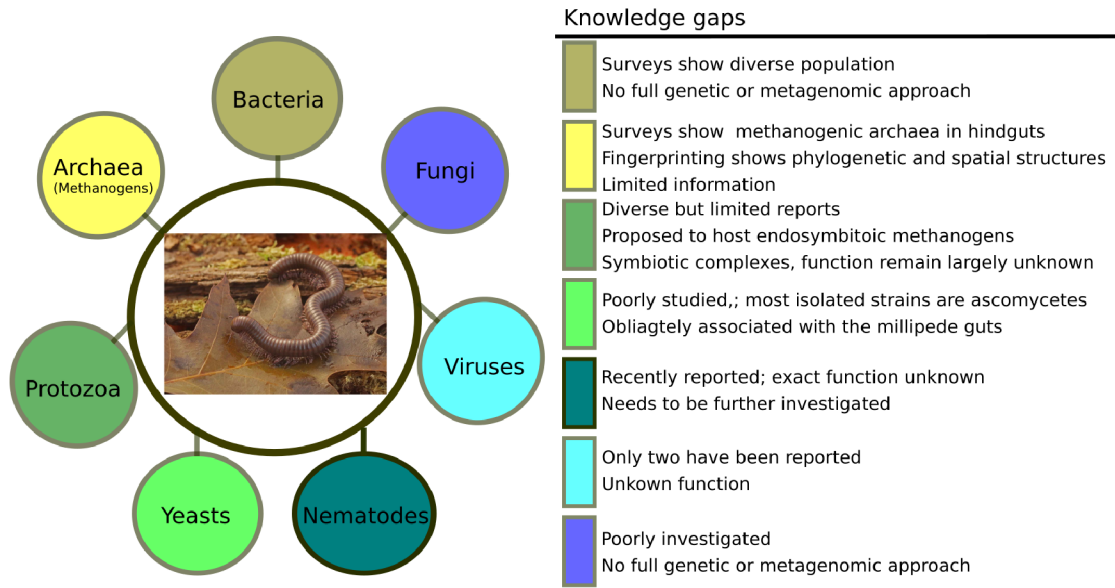


Fig. 6: Available information and gaps in our knowledge regarding the microbial communities that reside in the guts of millipedes.

### 1.7.1 Bacteria

Though millipedes are known to have diverse bacterial populations through culture-dependent methods, a full genetic or metagenomic approach revealing their activities in the gut has yet to take off. The only available reports have been a few microbial community surveys of the gut of millipedes (Degli Esposti & Martinez Romero, 2017). According to one of the surveys, the most dominant bacteria found dwelling in most millipede guts are members of the facultatively anaerobic bacteria family, *Enterobacteriaceae*. The genera in this family include *Escherichia*, *Enterobacter*, *Klebsiella*, *Vibrio*, *Plesiomonas*, *Erwinia* and *Salmonella* (Byzov, 2006; Konig, 2006). Other bacterial groups associated with the guts are the Gram-positive group of *Actinobacteria* and *Firmicutes*, the *Spirochetes* and the *Flavobacterium/Bacteroides* branch (Konig, 2006). Consistent with many authors, the bacterial counts (Table 1) have shown that bacterial growth is enhanced both in the guts and faecal



pellets of millipedes, which act as habitats for soil bacteria (Dhivya & Alagesan, 2017).

Table 1: Bacterial counts in the guts of some millipede species.

Millipede Species	Bacterial counts (cfu/g)	Authors
<i>Glomeris marginata</i>	23.4x10 <sup>7</sup>	(Anderson & Bignell, 1980)
<i>Glomeris marginata</i>	22.8x10 <sup>8</sup>	(Ineson & Anderson, 1985)
<i>Spinotarsus colosseus</i>	4.7x10 <sup>11</sup>	(Ramanathan & Alagesan, 2012)
<i>Arthrosphaera magna</i>	3.8x10 <sup>11</sup>	(Ramanathan & Alagesan, 2012)
<i>Aulocobolus newtoni</i>	3.1x10 <sup>11</sup>	(Ramanathan & Alagesan, 2012)
<i>Jonespeltis splendidus</i>	153.5x10 <sup>4</sup>	(Bano, Bagyaraj & Krishnamoorthy, 1976)
<i>Ommatoiulus sabulosus</i>	3.84x10 <sup>6</sup>	(Jarosz & Kania, 2000)
<i>Xenobolus carnifex</i>	1.4x10 <sup>8</sup>	(Alagesan & Muthukrishnan, 2005)
<i>Spinotarsus colosseus</i>	3.4x10 <sup>6</sup>	(Dhivya & Alagesan, 2018)

Table 1 shows gut bacterial counts in selected millipedes. Numerous researchers have successfully isolated various bacterial species from the guts of millipedes (Dhivya & Alagesan, 2018; Ineson & Anderson, 1985; Jarosz & Kania, 2000; Kania & Kłapeć, 2012; Oravec, 2002; Ramanathan & Alagesan, 2012; Soil, 2005). Since many host-associated microorganisms cannot be grown outside their hosts, our knowledge remains limited. Recent studies have reported the most prevalent taxa in the millipede species *Anadenobolus monilicornis* (Geli-Cruz et al., 2019) (only a pre-print of a metagenomic study

available) and *Telodeinopus aoutii* (Sardar et al., 2022) (only transcriptome data).

### 1.7.2 Archaea

Only limited information about the presence of archaea in millipede guts is available. As mentioned, it has been reported that certain species from the millipede orders Julida, Spirobolida, and Spirostreptida shelter a community of methanogenic archaea in their guts that may be contributing to the well-being of these groups (Sridhar & Kadamannaya, 2011; Šustr, Chroňáková, et al., 2014). The potential benefits for millipedes are yet to be fully explored. Untargeted microscopic examination of microbial populations cannot easily distinguish archaea from bacteria, but molecular techniques can discern the presence of archaea alongside bacteria (Nardi, Bee, et al., 2016). These methods encompass 16S rRNA gene (Kim & Chun, 2014) or metagenomic sequencing (Liu et al., 2022), fluorescence *in situ* hybridization (FISH) (Garimberti & Tosi, 2010) or catalysed reporter deposition-fluorescence *in situ* hybridization (CARD-FISH) (Wilhartitz et al., 2007) and PCR (B. A. White et al., 1999). The initial comprehensive examination of methanogenic diversity within millipede digestive tracts identified the presence of archaeal orders *Methanomicrobiales*, *Methanosarcinales*, *Methanobacteriales*, and some unclassified groups (Šustr, Chroňáková, et al., 2014). Lang and Brune (2014) isolated a hydrogenotrophic methanogen, *Methanoplasmatales* (now referred to as *Methanomassiliicoccales*), from the hindgut of *Anadenobolus sp.*

Molecular approaches for analysing hindgut endosymbiotic methanogens in ciliates or flagellates are very limited. However, Van Hoek *et al.* (2000) were able to identify one archaeal sequence from an undetermined millipede symbiotic ciliate, while Paul *et al.* (2012) identified *Methanomassiliicoccales*

in ciliates from a tropical millipede, *Anadenobolus sp.* In addition to their role in reducing H<sub>2</sub> partial pressure through methane production, similar to termites (Brune, 2010), the specific functions of methanogens in different millipede species have not been thoroughly assessed.

### **1.7.3 Yeasts**

The community of yeasts in the millipede guts have been poorly studied and has only been looked at using cultivation methods. The most commonly isolated strains are ascomycetes. This has been demonstrated in *Pachyiulus flavipes* where the predominating species are *Pichia membranaefaciens*, *Debaryomyces hansenii*, *Torulaspota delbrueckii*, and *Williopsis californica* (Byzov, Vu Nguyen Thanh, et al., 1993). According to these researchers, these yeasts can be considered symbionts and have proven to be obligately associated with the millipede guts. The structure and composition of the community are constant and remain unaltered under different feeding and rearing conditions such as the feeding of sterile substrates, long-term starvation, and relatively low temperatures (Byzov, Thanh, et al., 1993).

Research using scanning electron microscopy has revealed the distribution of yeast community in *Megaphyllum projectum*, *Glomeris connexa*, and *Leptoiulus polonicus*. The results show that yeasts mostly colonise the hindgut of freshly collected diplopods with densities of about 10<sup>3</sup> cells/mm<sup>2</sup> (10<sup>4</sup> cells/gut) while only a few cells could be found in the midgut (Byzov, Vu Nguyen Thanh, et al., 1993). Also, yeast-like fungi have been isolated from the gut content of the millipede species, *Ommatoiulus sabulosus*, but they occurred at low population densities and the species was not identified (Jarosz & Kania, 2000).

#### **1.7.4 Other fungi**

Other fungi are believed to play a role in making available to the millipedes essential amino acids and vitamins that they are unable to synthesise themselves, particularly in species living in the desert, such as *Orthoporus ornatus* (Ambarish & Sridhar, 2016; D. Bignell, 1989). However, the population of yeasts and filamentous fungi in the gut tends to decline when compared to bacterial community (Farfan, 2010) and has been poorly investigated. An example is Zygomycota (Trichomycetes), obligate fungal symbionts that reside in the digestive systems of millipedes and various arthropods (M. M. White et al., 2000), especially in the hindgut (Lichtwardt, 1996). The interactions between Trichomycetes and their hosts are usually antagonistic or commensalistic and mutualistic in some cases, depending on developmental and environmental circumstances (Contreras & Cafaro, 2013). Nardi, Bee and Taylor (2016) also confirmed the occupation of a section of the anterior hindgut of *Cylindroiulus caeruleocinctus* by Trichomycetes. A strange fungus showing characteristics of *Antennopsis*, *Hormiscioideus* and *Coreomycetopsis* from two species of Danish millipedes (Julida) has been reported (Enghoff & Reboleira, 2017). They also recorded peculiar structures, tentatively referred to as fungi, from several millipede orders, where they occur between micro-scutes of the external cuticle.

#### **1.7.5 Millipede gut flagellates and ciliates**

In arthropods like termites, the importance of the flagellate and ciliates is reflected by their high abundance in this micro-ecosystem (Brune & Dietrich, 2015). According to Hongoh and Ohkuma (2010), one feature of microbiota is the cellular association of the gut flagellates with bacteria and methanogenic archaea. The bacterial and methanogenic symbionts are observed both inside

and on the surface of the host flagellate cells (Tokura et al., 2000) or ciliates (Van Hoek et al., 2000). Several other related protozoa have been isolated from different millipede species (Lalpotu, 1980; Bhandari, 2010; Ganapati & Narasimhamurti, 1960; Paul et al., 2012; Šustr, Chroňáková, et al., 2014).

*Eccrinales* under the sub-order *Ecrinaceae* are another morphologically diverse eukaryotic genera found in Diplopoda. They include four species of *Eccrinoides*, 21 species of *Enterobryus*, and one species of *Eccrinidus*. Although *Eccrinales* have been considered members of the Trichomycetes for the last 50 years, ribosomal gene (18S and 28S) sequence analyses have shown no close relationship to the Trichomycetes or other fungi. Instead, *Eccrinales* belong to the protist class *Mesomycetozoa* (animal-fungi boundary) and share a common ancestor with the *Amoebidiales* (Cafaro, 2005).

### **1.7.6 Nematodes**

Nematodes exhibit surprising adaptability, forming mutualistic or commensal relationships with host animals in unexpected environments (Vlaar et al., 2021). Phillips *et al.* (2019) reported eight nematode species coexisting in the same segment of millipede intestines, suggesting a commensal relationship. Nematodes enter millipede intestines when their eggs are ingested along with plant material. In the intestines, they survive and develop by feeding on pre-digested food from their millipede hosts, without causing apparent harm. Apart from consuming pre-digested food, nematodes often feed on intestinal bacteria (Phillips et al., 2016). However, parasitic nematodes from the infra-orders *Rhigonematomorpha* and *Oxyuridomorpha* have also been isolated from millipedes (Phillips, 2017).

In Japan, a recent study identified 73 nematode species belonging to the genus *Rhigonema* in the hindguts of the millipede *Riukiaria sp.* It was observed that these nematodes harboured segmented filamentous bacteria, which were later identified as members of the *Lachnospiraceae* family (Kitagami et al., 2019). According to the authors, these segmented filamentous bacteria may be specific to *Riukiaria sp.*, as they were not found in all millipede hindguts. Another report identified a *Rhigonema* species, *R. naylae*, in the hindgut of the polydesmid millipede *Parafontaria laminata*, also in Japan (Morffe & Hasegawa, 2017). These symbiotic nematodes found in millipedes do not pose any risk to other animals, including humans, as most of them play beneficial roles in nutrient recycling (Phillips et al., 2019). However, some nematodes, particularly in the genus *Coronostoma*, have been reported to prey on other nematodes (Phillips et al., 2016). Further research is needed to identify the millipede species hosting nematodes, determine the nematode species present in millipedes, and understand their precise roles.

### **1.7.7 Viruses**

It is widely acknowledged that viruses, particularly those infecting prokaryotes and unicellular eukaryotes, represent the most prevalent type of biological entity on Earth (Koonin, 2010). Although some studies have focused on prokaryotic communities in millipede guts, our understanding of viruses in this context remains extremely limited. However, recent advancements in RNA sequencing have led to the discovery of approximately 12 new viruses in millipedes. Among these, the Wuhan Millipede Virus 1 was identified in two species of Polydesmida and one unidentified species. These viruses exhibited similarities to phleboviruses, but their exact abundance and functions within millipedes are still unknown (C.-X. Li et al., 2015).

Additionally, a virus known as millipede-associated circular virus 1 has been reported in the millipede species *Oxidus sp.* (Rosario et al., 2018).

Prokaryotes have evolved diverse mechanisms to control the dissemination of viruses. Among these mechanisms, the CRISPR-Cas system (Clustered Regularly Interspaced Short Palindromic Repeats and CRISPR-associated genes) offers a unique sequence-specific defence system that can be dynamically updated to combat novel threats, functioning as an adaptive immune system (Garrett, 2021; Koonin & Makarova, 2019). However, the prevalence of CRISPR-Cas systems within the prokaryotic community in millipede guts and the potential identification of viral hosts have not yet been investigated.

Viruses have also been found to play a significant role in shaping microbial communities in various environments. They can modulate the activity of these communities by influencing the relative abundance of different microbial members through predator-prey dynamics and lysogenic conversion (Luo et al., 2022; Rosenwasser et al., 2016). Moreover, viruses contribute to the adaptation of prokaryotes to challenging conditions by carrying a diverse range of auxiliary metabolic genes (AMGs) alongside their core viral genes. These AMGs are involved in central metabolic processes such as energy acquisition, degradation of xenobiotics, and stress tolerance (Sun et al., 2023; Zheng et al., 2022). However, the abundance of AMGs in arthropod-associated viruses has not been explored yet. Besides containing viruses affecting the nutrient cycling and controlling microbial community, arthropod guts can also contain pathogenic arthropode-borne viruses (arboviruses) infecting animals and humans (Mairuhu et al., 2004; Musso & Gubler, 2016) or plant-pathogens (N. Li et al., 2020). Due to the way millipedes interact with other arthropods, animals, plants and even humans, it has been suggested that millipede-

associated viruses might have clinical and economic importance (C.-X. Li et al., 2015). It should be given proper attention. Metagenomic and metatranscriptomic techniques have emerged as the gold standard for investigating the composition of viral communities, addressing a notable gap in microbial ecology research.



## 2 Objectives

The aims of this PhD study were to:

- Identify the microbial communities that inhabit the millipede guts
- Investigate whether and to what degree do millipedes depend on their microbiome
- Resolve anaerobic processes and associated taxa that can potentially drive methanogenesis in the millipede guts
- Understand the biological agents carrying out nitrogen fixation in the millipede guts, and determine the biological significance of N<sub>2</sub> fixation by bacteria in the guts
- Understand the potential functional roles of the millipede gut microbiota

## 3 Results and Discussion

### 3.1 Summary of the results derived from Paper No. 1

In this study, we used comparative metagenomics and metatranscriptomics to explore the hindgut microbiota of two millipede species: the tropical millipede *Epibolus pulchripes*, known for its significant methane emissions, and the temperate millipede *Glomeris connexa*, which does not emit methane (Šustr, Chroňáková, et al., 2014). Both millipedes were subjected to the same dietary conditions. Our primary objective was to uncover the metabolic potential of these species and gain insights into the trophic niche of these critical detritivores.

Our investigation unveiled notable distinctions between the two millipede species. *E. pulchripes* exhibited a considerably larger and more diverse microbial population in its hindgut than *G. connexa*. Additionally, the analysis of bacterial communities using 16S rRNA sequencing revealed distinct compositions, with *Bacteroidota* (*Bacteroidetes*) being the dominant group in *E. pulchripes* and *Proteobacteria* (*Pseudomonadota*) prevailing in *G. connexa*. The taxonomic classification of the contigs from the assembled metagenomic and metatranscriptomic reads closely mirrored the composition obtained from the amplicon sequencing. Despite equal sequencing effort, our *de novo* metagenomic assembly and binning yielded 282 metagenome-assembled genomes (MAGs) from *E. pulchripes* and 33 from *G. connexa*, including 90 novel bacterial taxa (81 in *E. pulchripes* and 9 in *G. connexa*). As anticipated from both libraries, methane-producing *Euryarchaeota* (orders *Methanobacteriales*, *Methanomassiliicoccales*, and *Methanosarcinales*) were found in *E. pulchripes*. Still, surprisingly, they were also present in the non-methane-emitting *G. connexa*. In both millipede species, the dominant fungal

phylum was *Ascomycota*. *E. pulchripes* also hosted significant numbers of *Nematoda* and *Ciliophora*, while *G. connexa* was characterised by *Apicomplexa* and *Metamonada*. *Eccrinales*, a protist order commonly found in millipedes (Cafaro, 2005), were rare, with minimal representation in both species and absent from the metatranscriptome. The metatranscriptome profile closely matched the metagenome in taxonomy but with differing relative abundances.

Nonetheless, even with this taxonomic divergence, most of the functions, such as carbohydrate hydrolysis, sulfate reduction, and nitrogen cycling, were shared by both species. Annotation of the predicted amino acid sequences in MAGs revealed a repertoire of carbohydrate-degrading enzymes (CAZymes) necessary to break down complex polysaccharides in plant litter. In *E. pulchripes*, members of the *Bacteroidota* were the primary contributors to complex carbon degradation, while in *G. connexa*, members of *Proteobacteria* dominated this process. The most abundant and expressed carbohydrate-active enzymes (CAZymes) were glycoside hydrolases (GHs), capable of breaking down various components, including fungal cell walls, hemicellulose, pectin, cellulose, starch, algal cell walls, and bacterial cell walls. This indicates the importance of fungal biomass in the millipedes' diet. Members of *Desulfobacterota* were the potential sulfate-reducing bacteria in *E. pulchripes*. *Actinobacteriota* (*E. pulchripes*) and *Proteobacteria* (both species) displayed the capacity for dissimilatory nitrate reduction, while only *Proteobacteria* possessed the capacity for denitrification (both species). However, certain functions were exclusive to *E. pulchripes*. These include reductive acetogenesis, acting as a sink for excess hydrogen produced during fermentation, and were found in members of *Desulfobacterota* and *Firmicutes* (*Bacillota*) in *E. pulchripes*. Considering that millipedes consume nitrogen-

deficient diets, the presence and expression of Molybdenum-dependent nitrogenases (*nifD**HK*) in a *Proteobacteria* MAG (*Pantoea cypripedii*) of *E. pulchripes* suggest that diazotrophs play a role in nitrogen fixation, as validated by the acetylene reduction assay. These findings provide the first comprehensive understanding of the genomic potential of the hindgut microbial community in millipedes and enhance our knowledge of the ecophysiology of these essential detritivores.

## 3.2 Summary results derived from Paper No. 2

This addresses the previously overlooked subject of DNA and RNA viral diversity within the hindguts of *E. pulchripes* and *G. connexa*. In addition to charting the viral diversity, this study revealed the prevalence of CRISPR-Cas loci in the prokaryotic communities of these millipede species by using the two sequencing libraries and metagenome-assembled genomes (MAGs). Predictive methods were employed to identify potential hosts and lifestyles for the detected viruses. Furthermore, the research evaluated the abundance of viral auxiliary metabolic genes (AMGs).

The findings, based on metagenomic and metatranscriptomic assembled viral genomes (MAVGs), demonstrated significant differences in viral communities that exhibit preferences for infecting the most abundant prokaryotic taxa. *E. pulchripes* harboured high-quality MAVGs, consisting of 253 free viruses and 45 proviruses, while *G. connexa*'s metagenome-derived high-quality viral genomes comprise 52 free viruses and 3 proviruses. DNA viruses are primarily classified into *Caudoviricetes* (dsDNA), *Cirivirales* (ssDNA), and *Microviridae* (ssDNA), while RNA viruses include *Leviviricetes* (ssRNA), *Potyviridae* (ssRNA), and eukaryotic viruses.

Lifestyle predictions indicated that the majority of MAVGs from the metagenome and prophage from MAGs in both species are lysogenic (temperate), while most MAVGs from the metatranscriptome are lytic (virulent). The *Caudoviricetes* class dominates the predicted viral genomes in both millipede species, displaying a mixture of lysogenic and virulent lifestyles. Putative viral hosts are successfully assigned to 141 MAVGs from *E. pulchripes* and 28 from *G. connexa*, with *Bacteroidota* being the most frequently predicted host for *E. pulchripes*, followed by *Bacillota*,

*Pseudomonadota*, and *Desulfobacterota*. In *G. connexa*, *Pseudomonadota* is the most commonly predicted host, followed by *Bacteroidota* and *Bacillota*. Low virus-to-microbe-ratios (mVMR) and a prevalence of lysogenic viruses suggest a "Piggyback-the-Winner" dynamic in both hosts.

The analysis of CRISPR-Cas gene abundance within MAGs revealed the presence of 61 arrays in *E. pulchripes* and 8 in *G. connexa*; primarily composed of subtypes I-C, I-B, and II-C CRISPR-Cas systems. These CRISPR-Cas systems were found in both species, with the main contributors originating from the taxonomic groups *Pseudomonadota*, *Bacteroidota*, and *Bacillota*. Additionally, 135 auxiliary metabolic genes (AMGs) from the class *Caudoviricetes* are identified in both millipede species, playing roles in chitin degradation, vitamins and amino acid biosynthesis, as well as sulfur metabolism.

### 3.3 Summary results derived from Manuscript No. 3

In the third manuscript, we conducted an extensive investigation to determine whether millipedes rely on their gut microbiota for cellulose digestion in plant litter through fermentation, similar to termites. We disrupted the microbiota of juvenile *E. pulchripes* and *G. connexa* using inhibitors and identified the active prokaryotes in their hindguts with the help of <sup>13</sup>C-labelled leaf litter and RNA-SIP.

Feeding millipedes with either sterile feed or food treated with an antibacterial and antifungal mixture led to a significant reduction in faecal production and microbial load, with minimal weight loss and no significant impact on survival or overall weight. This highlights quite well that gut microbiota don't have a "crucial" role in cellulose digestion in the gut. A substantial reduction in bacterial load resulted in a significant decrease in methane (CH<sub>4</sub>) emissions in *E. pulchripes*, while no CH<sub>4</sub> emissions were detected in *G. connexa*. Feeding *E. pulchripes* with litter treated with an inhibitor for methanogenesis (Na-2-BES) almost entirely halted CH<sub>4</sub> production within 14 days, with no impact on weight or faecal production. However, CH<sub>4</sub> production resumed after returning to normal feeding.

The use of antibiotics revealed that bacterial diversity remained largely unchanged, with *Bacteroidota* dominant in the hindgut of *E. pulchripes* and *Pseudomonadota* in *G. connexa*, as well as in the faeces of both species. Multiple-pairwise analysis showed differential abundance in a few microbial species between treatments, with no significant differences in alpha and beta diversity in the hindguts or faeces. In the hindgut of *E. pulchripes*, various protists, nematodes, and rotifers were present. Surprisingly, even after CH<sub>4</sub>

inhibition with Na-2-BES, methanogens in the orders *Methanobacteriales* and *Methanomassilliicoccales*, which are associated with protists, could still be detected using CARD-FISH.

The labelling of microbiota during  $^{13}\text{C}$ -labelled litter feeding was limited, indicating suboptimal assimilation efficiency. This can be attributed to the digestive characteristics of millipedes, which prioritise the digestion of non-structural compounds released during the early stages of litter decomposition, or the fact that millipedes mostly digest fungi. The labelled microbiota primarily belonged to the *Bacillota*, *Bacteroidota*, and *Pseudomonadota* phyla. A noteworthy observation was the changing dominance of *Pseudomonadota* in *E. pulchripes* from day 3 to day 21, while *Bacillota* became increasingly prevalent in *G. connexa* over the same period, ultimately becoming the most abundant phylum by day 21.



## 4 Conclusion and future prospects

### 4.1 Conclusions derived from Paper No. 1

Millipedes, important detritivores, are thought to rely on their gut microbiome for the digestion of plant material, similar to many other arthropods. This plant material is often low in nitrogen but rich in complex polysaccharides. However, the identity and nutritional contribution of their microbiome were until now largely unknown. This study represents the first effort to explore the metabolic diversity in millipedes, an important group of detritivores on our planet.

We observed significant variations in both the abundance and diversity of the gut microbial community between the two millipede species. These species differ in size, habitat, and gut redox conditions, but they share the same diet and lifestyle. Many functions carried by the gut microbiota were found in the metagenome-assembled genomes (MAGs) of both species, including the ability to break down complex carbohydrates. While lignin-modifying enzymes were scarce, there was substantial gene expression related to chitin degradation, indicating the potential significance of fungal biomass in the millipede diet, which may even surpass the importance of plant polymers.

Fermentative lineages, such as *Clostridiales* and *Bacteroidales*, were notably abundant in the larger *E. pulchripes*. However, clear evidence for reductive acetogenesis was lacking. Instead, we discovered strong indications of hydrogenotrophy, nitrogen recycling, and diazotrophy. These findings provide a foundation for future research to investigate these hypotheses concerning the trophic role of millipedes.

## **4.2 Conclusions derived from Paper No. 2**

Previous research has primarily focused on prokaryotes in the millipede gut, leaving the virome largely unexplored. Our study marks the first comprehensive investigation into the DNA and RNA viral communities inhabiting the hindguts of millipedes. Arthropods, despite their ecological significance, still lack in-depth virome research, presenting substantial potential for discovering novel viral lineages.

The microbiome plays a pivotal role in the ecological functions of arthropods, particularly detritivores. Understanding how the virome can influence microbial composition in arthropod guts may provide insights into the potential role of viruses in biogeochemical cycling. Previous studies have indicated that temperate phages can protect bacterial hosts from other phages by conferring superinfection immunity. Given that millipedes inhabit soil environments, their gut bacteria are likely exposed to various environmental phages. Investigating whether millipede gut phages offer superinfection immunity to safeguard the gut microbiota from environmental bacteriophages would be an intriguing avenue of exploration. In the future, millipedes could serve as a valuable model system for investigating the intricate interplay between bacteria, phages, and intestinal protozoa, particularly in detritivores.

## **4.3 Conclusions derived from Manuscript No. 3**

Millipedes, as crucial detritivores, have not previously been demonstrated to rely on their gut microbiota for cellulose digestion in plant litter via fermentation, similar to termites. Our study provides valuable insights into the

intricate interaction between millipedes and their gut microbiota, elucidating their digestion processes and methane emission dynamics. We noted a considerable reduction in faecal production and bacterial load, with insignificant changes in weight. This highlights the potential involvement of microbiota in breaking down fungal biomass, indicated by the prevalence of bacterial-derived chitinases. This might influence feeding preferences.

Our findings emphasise the resilience of microbial communities within millipedes' hindguts and highlight the central role of microbiota as a primary food source for these arthropods. Furthermore, our research underscores the limited efficiency of millipedes in assimilating cellulose, revealing that a considerable portion of ingested litter remains undigested, aligning with their preference for easily digestible non-structural compounds during initial litter decomposition.

The use of  $^{13}\text{C}$ -RNA-SIP to trace active microbiota has enhanced our comprehension of microbial community dynamics within millipedes' guts. Additionally, we identified substantial differences in gut microbial composition between the two millipede species and identified key phyla likely involved in the degradation and assimilation of non-structural polysaccharides or structural polysaccharides with a significant microbial component. These findings significantly advance our understanding of the intricate relationships between millipedes and their gut microbiomes, offering insights into the ecological roles of these arthropods and their associated microorganisms.

## **4.4 The millipede microbiome – future research directives**

Future research into the millipede gut microbiome should encompass multiple facets. Firstly, elucidate whether their survival depends on fungi consumption or exclusive feeding on fungi-associated litter, revealing preferred fungal classes and specific dietary habits. Secondly, efforts should focus on characterising metabolically important microbiota within millipede guts, similar to the methodologies employed in termite gut microbiome studies. These include targeting microbial groups involved in essential metabolic activities such as polysaccharide degradation, diazotrophy, nitrogen recycling, sulfate reduction, and detoxification, while also investigating the source of sulfate in the millipede gut. Simultaneously, investigating the role and characterising hindgut-dwelling protists, nematodes, and rotifers, akin to cellulolytic flagellated protists in termites, will uncover their contributions to millipede digestive processes, along with exploring their relationships with endosymbionts. Furthermore, exploring the intricate interplay between millipede gut prokaryotes and gut phages to unravel potential superinfection immunity mechanisms against environmental bacteriophages remains pivotal for understanding gut microbiota safeguarding. Lastly, a comprehensive analysis of complex interactions among bacteria, phages, intestinal protozoa, and millipedes will enrich our understanding of detritivorous ecosystem dynamics, offering insights into the intricate relationships shaping these ecosystems.

## References

- Alagesan, P. (2016). Millipedes: Diversity, distribution and ecology. In A. K. Chakravarthy & S. Sridhara (Eds.), *Arthropod diversity and conservation in the tropics and subtropics* (pp. 119–137). Springer Singapore. [https://doi.org/10.1007/978-981-10-1518-2\\_7](https://doi.org/10.1007/978-981-10-1518-2_7)
- Ambarish, C. N., & Sridhar, K. R. (2016). Chemical and Microbial Characterization of Feed and Faeces of Two Giant Pill-Millipedes from Forests in the Western Ghats of India. *Pedosphere*, 26(6), 861–871. [https://doi.org/10.1016/S1002-0160\(15\)60091-1](https://doi.org/10.1016/S1002-0160(15)60091-1)
- Ax, P. (2000). Pentazonia—Helminthomorpha. *Multicellular Animals: The Phylogenetic System of the Metazoa. Volume II*, 239–260.
- Benckiser, G. (1997). *Fauna in soil ecosystems: Recycling processes, nutrient fluxes, and agricultural production*. CRC Press.
- Bhandari, J. C. (2010). Studies on some ciliate and gregarine parasites of annelida and arthropoda. *University*. <http://hdl.handle.net/10603/78964>
- Bignell, D. (1989). Relative assimilations of <sup>14</sup>C-labelled microbial tissues and <sup>14</sup>C-plant fibre ingested with leaf litter by the millipede *Glomeris marginata* under experimental conditions. *Soil Biology and Biochemistry*, 21(6), 819–827. [https://doi.org/10.1016/0038-0717\(89\)90176-4](https://doi.org/10.1016/0038-0717(89)90176-4)
- Bignell, D. E. (1984). Direct potentiometric determination of redox potentials of the gut contents in the termites *Zootermopsis nevadensis* and *Cubitermes severus* and in three other arthropods. *Journal of Insect Physiology*, 30(2), 169–174. [https://doi.org/10.1016/0022-1910\(84\)90122-7](https://doi.org/10.1016/0022-1910(84)90122-7)
- Blanke, A., & Wesener, T. (2014). Revival of forgotten characters and modern imaging techniques help to produce a robust phylogeny of the Diplopoda (Arthropoda, Myriapoda). *Arthropod Structure & Development*, 43(1), 63–75. <https://doi.org/10.1016/j.asd.2013.10.003>
- Bocock, K. L. (1963). The digestion and assimilation of food by *Glomeris*. In *Soil organisms*.
- Boga, H. I., Ludwig, W., & Brune, A. (2003). *Sporomusa aerivorans* sp. Nov., an oxygen-reducing homoacetogenic bacterium from the gut of a soil-feeding termite. *International Journal of Systematic and Evolutionary Microbiology*, 53(5), 1397–1404. <https://doi.org/10.1099/ijs.0.02534-0>
- Bogyó, D., Magura, T., Nagy, D. D., & Tóthmérész, B. (2015). Distribution of millipedes (Myriapoda, Diplopoda) along a forest interior—Forest edge—Grassland habitat complex. *ZooKeys*, 510, 181–195. <https://doi.org/10.3897/zookeys.510.8657>
- Bot, A., & Benites, J. (2005). *The importance of soil organic matter: Key to drought-resistant soil and sustained food production*. Food and Agriculture Organization of the United Nations.
- Brewer, M. S., & Bond, J. E. (2013). Ordinal-level phylogenomics of the arthropod class Diplopoda (millipedes) based on an analysis of 221 nuclear protein-coding loci generated using next-generation sequence analyses. *PLoS ONE*, 8(11). <https://doi.org/10.1371/journal.pone.0079935>
- Brewer, M. S., Sierwald, P., & Bond, J. E. (2012). Millipede taxonomy after 250 years: Classification and taxonomic practices in a mega-diverse yet understudied arthropod group. *PloS One*, 7(5). <https://doi.org/10.1371/journal.pone.0037240>

- Brune, A. (2010). Methanogens in the Digestive Tract of Termites. In J. H. P. Hackstein (Ed.), *(Endo)symbiotic Methanogenic Archaea* (Vol. 19, pp. 81–100). Springer Berlin Heidelberg. [https://doi.org/10.1007/978-3-642-13615-3\\_6](https://doi.org/10.1007/978-3-642-13615-3_6)
- Brune, A. (2014). Symbiotic digestion of lignocellulose in termite guts. *Nature Reviews Microbiology*, 12(3), Art. 3. <https://doi.org/10.1038/nrmicro3182>
- Brune, A., & Dietrich, C. (2015). The gut microbiota of termites: Digesting the diversity in the light of ecology and evolution. *Annual Review of Microbiology*, 69(1), 145–166. <https://doi.org/10.1146/annurev-micro-092412-155715>
- Byzov, B. A. (2006). Intestinal microbiota of millipedes. In H. König & A. Varma (Eds.), *Intestinal microorganisms of termites and other invertebrates* (Vol. 6, pp. 89–114). Springer-Verlag. [https://doi.org/10.1007/3-540-28185-1\\_4](https://doi.org/10.1007/3-540-28185-1_4)
- Byzov, B. A., Thanh, V. N., Bab'Eva, I. P., Tretyakova, E. B., Dyvak, I. A., & Rabinovich, Y. M. (1998). Killing and hydrolytic activities of the gut fluid of the millipede *Pachyiulus flavipes* C.L. Koch on yeast cells. *Soil Biology and Biochemistry*, 30(8–9), 1137–1145. [https://doi.org/10.1016/S0038-0717\(97\)00190-9](https://doi.org/10.1016/S0038-0717(97)00190-9)
- Byzov, B. A., Thanh, V. N., & Babjeva, I. P. (1993). Yeasts associated with soil invertebrates. *Biology and Fertility of Soils*, 16(3), 183–187. <https://doi.org/10.1007/BF00361405>
- Byzov, B. A., Vu Nguyen Thanh, & Babjeva, I. P. (1993). Interrelationships between yeasts and soil diplopods. *Soil Biology and Biochemistry*, 25(8), 1119–1126. [https://doi.org/10.1016/0038-0717\(93\)90160-D](https://doi.org/10.1016/0038-0717(93)90160-D)
- Cafaro, M. J. (2005). Eccrinales (Trichomycetes) are not fungi, but a clade of protists at the early divergence of animals and fungi. *Molecular Phylogenetics and Evolution*, 35(1), 21–34. <https://doi.org/10.1016/j.ympev.2004.12.019>
- Cárcamo, H. A., Abe, T. A., Prescott, C. E., Holl, F. B., & Chanway, C. P. (2000). *Influence of millipedes on litter decomposition, N mineralization, and microbial communities in a coastal forest in British Columbia, Canada*. 30, 10.
- Chakravarthy, A. K., & Sridhara, S. (2016). Economic and ecological significance of arthropods in diversified ecosystems: Sustaining regulatory mechanisms. In *Economic and Ecological Significance of Arthropods in Diversified Ecosystems: Sustaining Regulatory Mechanisms*. <https://doi.org/10.1007/978-981-10-1524-3>
- Chitty, J. R. (2022). Myriapods (centipedes and Millipedes). In *Invertebrate Medicine* (pp. 399–412). John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781119569831.ch17>
- Coleman, D. C., Crossley, D. A., & Hendrix, P. F. (2004). Secondary Production: Activities of Heterotrophic Organisms—The Soil Fauna. *Fundamentals of Soil Ecology*, 79–185. <https://doi.org/10.1016/b978-012179726-3/50005-8>
- Contreras, K., & Cafaro, M. J. (2013). Morphometric Studies in *Enterobryus luteovirgatus* sp. Nov. (Ichthyosporea:Eccrinales) Associated with Yellow-banded Millipedes in Puerto Rico. *Acta Protozoologica*, 52, 291–297.
- Cortes, C., VandeVoort, A. R., & Snyder, B. (2018). Soil Nitrification Analysis and Millipede Contribution. *Georgia Journal of Science*, 76(1), 117.
- Coulis, M., Hättenschwiler, S., Fromin, N., & David, J. F. (2013). Macroarthropod-microorganism interactions during the decomposition of Mediterranean shrub litter at different moisture levels. *Soil Biology and Biochemistry*, 64, 114–121. <https://doi.org/10.1016/j.soilbio.2013.04.012>
- Crawford, C. S. (1992). Millipedes as Model Detritivores. *Ber. Nat.-Med. Verein Innsbruck*, 12.
- Crawford, C. S., Minion, G. P., & Boyers, M. D. (1983). Intima morphology, bacterial morphotypes, and effects of annual molt on microflora in the hindgut of the desert millipede, *Orthoporus ornatus* (Girard) (Diplopoda: Spirostreptidae). *International*

- Journal of Insect Morphology and Embryology*, 12(5–6), 301–312. [https://doi.org/10.1016/0020-7322\(83\)90025-9](https://doi.org/10.1016/0020-7322(83)90025-9)
- Culliney, T. (2013). Role of Arthropods in Maintaining Soil Fertility. *Agriculture*. <https://doi.org/10.3390/agriculture3040629>
- da Silva Correia, D., Passos, S. R., Proença, D. N., Morais, P. V., Xavier, G. R., & Correia, M. E. F. (2018). Microbial diversity associated to the intestinal tract of soil invertebrates. *Applied Soil Ecology*, 131, 38–46. <https://doi.org/10.1016/j.apsoil.2018.07.009>
- da Silva, V. M., Antonioli, Z. I., Jacques, R. J. S., Ott, R., Rodrigues, P. E. da S., Andrade, F. V., Passos, R. R., & Mendonça, E. de S. (2017). Influence of the tropical millipede, *Glyphiulus granulatus* (Gervais, 1847), on aggregation, enzymatic activity, and phosphorus fractions in the soil. *Geoderma*. <https://doi.org/10.1016/j.geoderma.2016.11.031>
- David, J.-F. (2014). The role of litter-feeding macroarthropods in decomposition processes: A reappraisal of common views. *Soil Biology and Biochemistry*, 76, 109–118. <https://doi.org/10.1016/j.soilbio.2014.05.009>
- David, J.-F. (2015). Diplopoda—Ecology. In *Treatise on Zoology-Anatomy, Taxonomy, Biology. The Myriapoda, Volume 2* (pp. 303–327). Brill.
- Degli Esposti, M., & Martinez Romero, E. (2017). The functional microbiome of arthropods. *PLOS ONE*, 12(5), e0176573. <https://doi.org/10.1371/journal.pone.0176573>
- Dhivya, A., & Alagesan, P. (2017). *Millipedes as Host for Microbes—A Review*. 8(1), 19–24. <https://doi.org/10.5829/idosi.ijmr.2017.19.24>
- Dhivya, A., & Alagesan, P. (2018). Isolation and identification of microbial load in the gut and faeces of millipede *Spinotarsus colosseus*. *World Journal of Zoology*, 13(2), 04–09. <https://doi.org/10.5829/IDOSI.WJZ.2018.04.09>
- Engel, P., & Moran, N. A. (2013). The gut microbiota of insects – diversity in structure and function. *FEMS Microbiology Reviews*, 37(5), 699–735. <https://doi.org/10.1111/1574-6976.12025>
- Enghoff, H. (2010). Revision of the East African millipede genus *Epibolus* Cook, 1897. *Insect Systematics & Evolution*, 8(1), 1–8. <https://doi.org/10.1163/187631277x00017>
- Enghoff, H., & Minelli, A. (1990). The ground-plan of chilognathan millipedes (external morphology). *Proceedings of the 7th International Congress of Myriapodology*. EJ Brill, Leiden, 1–21.
- Enghoff, H., & Reboleira, A. (2017). Diversity of non-Laboulbenialean fungi on millipedes. *Studies in Fungi*, 2(1), 130–137. <https://doi.org/10.5943/sif/2/1/15>
- Farfan, M. A. (2010). *Some Aspects of the Ecology of Millipedes (Diplopoda)* [Masters Degree, Ohio State University]. [https://etd.ohiolink.edu/apexprod/rws\\_etd/send\\_file/send?accession=osu1279635751&disposition=inline](https://etd.ohiolink.edu/apexprod/rws_etd/send_file/send?accession=osu1279635751&disposition=inline)
- Farquharson, P. A. (1974). A study of the Malpighian tubules of the pill millipede, *Glomeris marginata* (Villers) III. The permeability characteristics of the tubule. *Journal of Experimental Biology*, 60(1), 41–51.
- Fujimaki, R., Sato, Y., Okai, N., & Kaneko, N. (2010). The train millipede (*Parafontaria laminata*) mediates soil aggregation and N dynamics in a Japanese larch forest. *Geoderma*, 159(1–2), 216–220. <https://doi.org/10.1016/j.geoderma.2010.07.014>
- Ganapati, P. N., & Narasimhamurti, C. C. (1960). On a new haplosporidian (protozoa) *Nephridiophaga xenoboli* in the gut of a millipede *Xenobolus carnifex*. *Parasitology*, 50(3–4), 581–585. <https://doi.org/10.1017/S0031182000025646>

- Garimberti, E., & Tosi, S. (2010). Fluorescence in situ hybridization (FISH), basic principles and methodology. *Methods in Molecular Biology (Clifton, N.J.)*, 659, 3–20. [https://doi.org/10.1007/978-1-60761-789-1\\_1](https://doi.org/10.1007/978-1-60761-789-1_1)
- Garrett, S. C. (2021). Pruning and Tending Immune Memories: Spacer Dynamics in the CRISPR Array. *Frontiers in Microbiology*, 12. <https://doi.org/10.3389/fmicb.2021.664299>
- Geli-Cruz, O., Cafaro, M. J., Santos-Flores, C. J., Ropelewski, A. J., & Van Dam, A. R. (2019). *Taxonomic survey of Anadenobolus monilicornis gut microbiota via shotgun nanopore sequencing* [Preprint]. Genomics. <https://doi.org/10.1101/560755>
- Gillon, D., & David, J.-F. (2001). The use of near infrared reflectance spectroscopy to study chemical changes in the leaf litter consumed by saprophagous invertebrates. *Soil Biology and Biochemistry*, 33(15), 2159–2161. [https://doi.org/10.1016/S0038-0717\(01\)00139-0](https://doi.org/10.1016/S0038-0717(01)00139-0)
- Golovatch, S. I., & Kime, R. D. (2009). *Millipede ( Diplopoda ) distributions: A review*. 81(3), 565–597.
- Guru, A. D., Lahir, Y. K., & Mohite, V. T. (2013). Protease activity in the alimentary canal of millipede Harpappe haydeniana (Wood). *Journal of Experimental Zoology, India*, 16(2), 653–656.
- Hammer, T. J., Janzen, D. H., Hallwachs, W., Jaffe, S. P., & Fierer, N. (2017). Caterpillars lack a resident gut microbiome. *Proceedings of the National Academy of Sciences*, 114(36), 9641–9646. <https://doi.org/10.1073/pnas.1707186114>
- Hashimoto, M., Kaneko, N., Ito, M. T., & Toyota, A. (2004). Exploitation of litter and soil by the train millipede Parafontaria laminata (Diplopoda: Xystodesmidae) in larch plantation forests in Japan. *Pedobiologia*, 48(1), 71–81. <https://doi.org/10.1016/j.pedobi.2003.09.001>
- Hoess, R., & Scholl, A. (2001). Allozyme and Literature Study of *Glomeris guttata* Risso, 1826, and *G. connexa* Koch, 1847, a Case of Taxonomic Confusion (Diplopoda: Glomeridae). *Zoologischer Anzeiger - A Journal of Comparative Zoology*, 240(1), 15–33. <https://doi.org/10.1078/0044-5231-00003>
- Hongoh, Y., & Ohkuma, M. (2010). *Termite Gut Flagellates and Their Methanogenic and Eubacterial Symbionts* (pp. 55–79). [https://doi.org/10.1007/978-3-642-13615-3\\_5](https://doi.org/10.1007/978-3-642-13615-3_5)
- Hopkin, S. P., & Read, H. J. (1992). The biology of millipedes. *Choice Reviews Online*, 30(08), 30-4383-30–4383. <https://doi.org/10.5860/choice.30-4383>
- Horváthová, T., Šustr, V., Chroňáková, A., Semanová, S., Lang, K., Dietrich, C., Hubáček, T., Ardestani, M. M., Lara, A. C., Brune, A., & Šimek, M. (2021). Methanogenesis in the Digestive Tracts of the Tropical Millipedes *Archispirostreptus gigas* (Diplopoda, Spirostreptidae) and *Epibolus pulchripes* (Diplopoda, Pachybolidae). *Applied and Environmental Microbiology*, 87(15), e00614-21. <https://doi.org/10.1128/AEM.00614-21>
- Ineson, P., & Anderson, J. M. (1985). Aerobically isolated bacteria associated with the gut and faeces of the litter feeding macroarthropods *Oniscus asellus* and *Glomeris marginata*. *Soil Biology and Biochemistry*, 17(6), 843–849. [https://doi.org/10.1016/0038-0717\(85\)90145-2](https://doi.org/10.1016/0038-0717(85)90145-2)
- Innsbruck, V. (1992). *Interactions among Millipedes ( Diplopoda ) and their Intestinal Bacteria*. April, 289–296.
- Jarosz, J., & Kania, G. (2000). The question of whether gut microflora of the millipede *Ommatoiulus sabulosus* could function as a threshold to food infections. *Pedobiologia*, 44(6), 705–708. [https://doi.org/10.1078/S0031-4056\(04\)70083-9](https://doi.org/10.1078/S0031-4056(04)70083-9)



- Kaneko, N. (1988). Feeding habits and cheliceral size of oribatid mites in cool temperate forest soils in Japan. *Rev Ecol Biol Sol*, 25, 353–363.
- Kania, G., & Kłapeć, T. (2012). Seasonal activity of millipedes (Diplopoda) – their economic and medical significance. *Annals of Agricultural and Environmental Medicine*, 19(4), 5.
- Kaplan, D. L., & Hartenstein, R. (1978). Studies on monooxygenases and dioxygenases in soil macroinvertebrates and bacterial isolates from the gut of the terrestrial isopod, *Oniscus asellus* L. *Comparative Biochemistry and Physiology – Part B: Biochemistry And*. [https://doi.org/10.1016/0305-0491\(78\)90025-1](https://doi.org/10.1016/0305-0491(78)90025-1)
- Kheirallah, A. M. (1990). Fragmentation of leaf litter by a natural population of the millipede *Julus scandinavicus* (Latzel 1884). *Biology and Fertility of Soils*, 10, 202–206. <https://doi.org/10.1007/BF00336137>
- Kim, M., & Chun, J. (2014). Chapter 4—16S rRNA Gene-Based Identification of Bacteria and Archaea using the EzTaxon Server. In M. Goodfellow, I. Sutcliffe, & J. Chun (Eds.), *Methods in Microbiology* (Vol. 41, pp. 61–74). Academic Press. <https://doi.org/10.1016/bs.mim.2014.08.001>
- Kitagami, Y., Kanzaki, N., & Matsuda, Y. (2019). First Report of Segmented Filamentous Bacteria Associated with *Rhigonema* Sp. (Nematoda: Rhigonematidae) Dwelling in Hindgut of *Riukiaria* Sp. (Diplopoda: Xystodesmidae). *Helminthologia*, 56(3), 219–228. <https://doi.org/10.2478/helm-2019-0018>
- Koch, M. (2015). Diplopoda—General morphology. In *Treatise on Zoology—Anatomy, Taxonomy, Biology. The Myriapoda, Volume 2* (pp. 7–67). Brill. [https://doi.org/10.1163/9789004188273\\_003](https://doi.org/10.1163/9789004188273_003)
- Konig, H. (2006). *Bacillus* species in the intestine of termites and other soil invertebrates. *Journal of Applied Microbiology*, 101(3), 620–627. <https://doi.org/10.1111/j.1365-2672.2006.02914.x>
- Koonin, E. V. (2010). The wonder world of microbial viruses. *Expert Review of Anti-Infective Therapy*, 8(10), 1097–1099. <https://doi.org/10.1586/eri.10.96>
- Koonin, E. V., & Makarova, K. S. (2019). Origins and evolution of CRISPR-Cas systems. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 374(1772), 20180087. <https://doi.org/10.1098/rstb.2018.0087>
- Lalpotu, P. A. (1980). Studies on Ciliates of the Genus *Nyctotherus* Leidy, 1849: III. Parasites of Millipedes. *Archiv Fur Protistenkunde*, 123(3), 261–266. [https://doi.org/10.1016/S0003-9365\(80\)80010-8](https://doi.org/10.1016/S0003-9365(80)80010-8)
- Lang, K., & Brune, A. (2014). Diversity, ultrastructure, and comparative genomics of “Methanoplasmatales”, the seventh order of methanogens. *Philipp-Universität Marburg Biologie*, 1–18.
- Li, C.-X., Shi, M., Tian, J.-H., Lin, X.-D., Kang, Y.-J., Chen, L.-J., Qin, X.-C., Xu, J., Holmes, E. C., & Zhang, Y.-Z. (2015). Unprecedented genomic diversity of RNA viruses in arthropods reveals the ancestry of negative-sense RNA viruses. *ELife*, 4, e05378. <https://doi.org/10.7554/eLife.05378>
- Li, N., Yu, C., Yin, Y., Gao, S., Wang, F., Jiao, C., & Yao, M. (2020). Pepper Crop Improvement Against Cucumber Mosaic Virus (CMV): A Review. *Frontiers in Plant Science*, 11. <https://doi.org/10.3389/fpls.2020.598798>
- Lichtwardt, R. W. (1996). Trichomycetes and the Arthropod Gut. In *Human and Animal Relationships*. [https://doi.org/10.1007/978-3-662-10373-9\\_16](https://doi.org/10.1007/978-3-662-10373-9_16)
- Liu, S., Moon, C. D., Zheng, N., Huws, S., Zhao, S., & Wang, J. (2022). Opportunities and challenges of using metagenomic data to bring uncultured microbes into cultivation. *Microbiome*, 10(1), 76. <https://doi.org/10.1186/s40168-022-01272-5>

- Luo, X.-Q., Wang, P., Li, J.-L., Ahmad, M., Duan, L., Yin, L.-Z., Deng, Q.-Q., Fang, B.-Z., Li, S.-H., & Li, W.-J. (2022). Viral community-wide auxiliary metabolic genes differ by lifestyles, habitats, and hosts. *Microbiome*, *10*(1), 190. <https://doi.org/10.1186/s40168-022-01384-y>
- Mairuhu, A. T. A., Wagenaar, J., Brandjes, D. P. M., & van Gorp, E. C. M. (2004). Dengue: An arthropod-borne disease of global importance. *European Journal of Clinical Microbiology & Infectious Diseases: Official Publication of the European Society of Clinical Microbiology*, *23*(6), 425–433. <https://doi.org/10.1007/s10096-004-1145-1>
- Marcuzzi, G., & Lafisca, M. T. (1975). Observations on the digestive enzymes of some litter-feeding animals. In *Progress in Soil Zoology* (pp. 593–598). Springer Netherlands. [https://doi.org/10.1007/978-94-010-1933-0\\_65](https://doi.org/10.1007/978-94-010-1933-0_65)
- Mbenoun Masse, P. S., Nzoko Fiemapong, A. R., VandenSpiegel, D., & Golovatch, S. I. (2018). Diversity and distribution of millipedes (Diplopoda) in the Campo Ma'an National Park, southern Cameroon. *African Journal of Ecology*, *56*(1), 73–80. <https://doi.org/10.1111/aje.12418>
- McBrayer, J. F. (1973). Exploitation of deciduous leaf litter by *Apheloria montana* (Diplopoda: Eurydesmidae). *Pedobiologia*.
- Minelli, A., & Golovatch, S. I. (2001). *Alessandro Minelli\* and Sergei I. Golovatch †*. *4*, 291–303.
- Minelli, A., & Golovatch, S. I. (2013). Myriapods. In *Encyclopedia of Biodiversity: Second Edition*. <https://doi.org/10.1016/B978-0-12-384719-5.00208-2>
- Moreira De Sousa, C., & Silvia Fontanetti, C. (2012). Structure and function of the foregut and salivary glands of the synanthropic diplopod *Urostreptus atrobrunneus* (Spirostreptidae). *Animal Biology*, *62*(4), 493–504. <https://doi.org/10.1163/157075612X650168>
- Moreira-de-Sousa, C., Iamonte, M., Fontanetti, C. S., Moreira-de-Sousa, C., Iamonte, M., & Fontanetti, C. S. (2016). Midgut of the diplopod *Urostreptus atrobrunneus*: Structure, function, and redefinition of hepatic cells. *Brazilian Journal of Biology*, *77*(1), 132–139. <https://doi.org/10.1590/1519-6984.11715>
- Morffe, J., & Hasegawa, K. (2017). *Rhigonema naylae* n. sp. (Rhigonematomorpha: Rhigonematidae) a new parasitic nematode from a Japanese polydesmid millipede (Polydesmida: Xystodesmidae). *Zootaxa*, *4269*(2), 277–286. <https://doi.org/10.11646/zootaxa.4269.2.6>
- Müller, C. H. G., Sombke, A., & Rosenberg, J. (2007). The fine structure of the eyes of some bristly millipedes (Penicillata, Diplopoda): Additional support for the homology of mandibulate ommatidia. *Arthropod Structure & Development*, *36*(4), 463–476. <https://doi.org/10.1016/j.asd.2007.09.002>
- Musso, D., & Gubler, D. J. (2016). Zika Virus. In *Clinical Microbiology Reviews* (Vol. 29, Issue 3, pp. 487–524). <https://doi.org/10.1128/CMR.00072-15>
- Mwabvu, T. (1998). Food preference and coprophagy in a tropical millipede. *Journal of African Zoology*.
- Nalepa, C. A., Bignell, D. E., & Bandi, C. (2001). Detritivory, coprophagy, and the evolution of digestive mutualisms in Dictyoptera. In *Insectes Sociaux*. <https://doi.org/10.1007/PL00001767>
- Nardi, J. B., Bee, C. M., & Taylor, S. J. (2016). Compartmentalization of microbial communities that inhabit the hindguts of millipedes. *Arthropod Structure & Development*, *45*(5), 462–474. <https://doi.org/10.1016/j.asd.2016.08.007>

- Nardi, J. B., Mackie, R. I., & Dawson, J. O. (2002). Could microbial symbionts of arthropod guts contribute significantly to nitrogen fixation in terrestrial ecosystems? *Journal of Insect Physiology*, *48*(8), 751–763. [https://doi.org/10.1016/S0022-1910\(02\)00105-1](https://doi.org/10.1016/S0022-1910(02)00105-1)
- Nardi, J. B., Miller, L. A., & Bee, C. M. (2016). A novel arrangement of midgut epithelium and hepatic cells implies a novel regulation of the insulin signaling pathway in long-lived millipedes. *Journal of Insect Physiology*, *91*, 76–83. <https://doi.org/10.1016/j.jinsphys.2016.06.011>
- Nunez, F. S., & Crawford, C. S. (1976). Digestive enzymes of the desert millipede *Orthoporus ornatus* (Girard) (Diplopoda: Spirostreptidae). *Comparative Biochemistry and Physiology – Part A: Physiology*, *55*(2), 141–145. [https://doi.org/10.1016/0300-9629\(76\)90082-7](https://doi.org/10.1016/0300-9629(76)90082-7)
- Nunez, F. S., & Crawford, C. S. (1977). Anatomy and histology of the alimentary tract of the desert millipede *Orthoporus ornatus* (Girard) (Diplopoda: Spirostreptidae). *Journal of Morphology*, *151*(1), 121–130. <https://doi.org/10.1002/jmor.1051510107>
- Oravec, O. (2002). A molecular approach in the analysis of the faecal bacterial community in an African millipede belonging to the family Spirostreptidae (Diplopoda). *European Journal of Soil Biology*, *38*(1), 67–70. [https://doi.org/10.1016/S1164-5563\(01\)01128-1](https://doi.org/10.1016/S1164-5563(01)01128-1)
- Paoletti, M. G., Mazzon, L., Martinez-Sañudo, I., Simonato, M., Beggio, M., Dreon, A., Pamio, A., Brilli, M., Dorigo, L., Summers Engel, A., Tondello, A., Baldan, B., Concheri, G., & Squartini, A. (2013). A unique midgut-associated bacterial community hosted by the cave beetle *Cansiliella servadeii* (Coleoptera: Leptodirini) reveals parallel phylogenetic divergences from universal gut-specific ancestors. *BMC Microbiology*, *13*(1), 129. <https://doi.org/10.1186/1471-2180-13-129>
- Paul, K., Nonoh, J. O., Mikulski, L., & Brune, A. (2012). “Methanoplasmatales,” Thermoplasmatales-Related Archaea in Termite Guts and Other Environments, Are the Seventh Order of Methanogens. *Applied and Environmental Microbiology*, *78*(23), 8245–8253. <https://doi.org/10.1128/AEM.02193-12>
- Phillips, G. (2017). *Life where you least expect it: Biodiversity, abundance and prevalence of kleptoparasitic nematodes living inside the gastrointestinal tract of North American diplopods.* <https://www.semanticscholar.org/paper/Life-where-you-least-expect-it-%3A-Biodiversity%2C-and-Phillips/1ac751a87e6727b054a78ea2ff718737153a20b9>
- Phillips, G., Bernard, E. C., Pivar, R. J., Moulton, J. K., & Shelley, R. M. (2016). *Coronostoma claireae* n. sp. (Nematoda: Rhabditida: Oxyuridomorpha: Coronostomatidae) from the Indigenous Millipede *Narceus gordanus* (Chamberlain, 1943) (Diplopoda: Spirobolida) in Ocala National Forest, Florida. *Journal of Nematology*, *48*(3), 159–169. <https://doi.org/10.21307/jofnem-2017-023>
- Phillips, G., Yates, D. I., Shelley, R. M., Ortstadt, P. R., & Bernard, E. C. (2019). Investigating Commensal Relationships of Nematodes in Millipedes: Life in Unexpected Places. *The American Biology Teacher*. <https://doi.org/10.1525/abt.2019.81.4.278>
- Rajulu, G. S. (1971). Cardiac physiology of a millipede *Cingalobolus bugnioni* Carl (Myriapoda Diplopoda). *Monitore Zoologico Italiano-Italian Journal of Zoology*, *5*(1), 39–52.
- Ramanathan, B., & Alagesan, P. (2012). *Isolation, characterization and role of gut bacteria of three different millipede species.* 7.
- Rawlins, A., Bull, I., Poirier, N., Ineson, P., & Evershed, R. (2006). The biochemical transformation of oak (*Quercus robur*) leaf litter consumed by the pill millipede

- (*Glomeris marginata*). *Soil Biology and Biochemistry*, 38(5), 1063–1076. <https://doi.org/10.1016/j.soilbio.2005.09.005>
- Rosario, K., Mettel, K. A., Benner, B. E., Johnson, R., Scott, C., Yusseff-Vanegas, S. Z., Baker, C. C. M., Cassill, D. L., Storer, C., Varsani, A., & Breitbart, M. (2018). Virus discovery in all three major lineages of terrestrial arthropods highlights the diversity of single-stranded DNA viruses associated with invertebrates. *PeerJ*, 6, e5761. <https://doi.org/10.7717/peerj.5761>
- Rosenberg, F. (2006). Intestinal Microorganisms of Termites and Other Invertebrates. *Microbe Magazine*, 1(8), 389–389. <https://doi.org/10.1128/microbe.1.389.1>
- Rosenwasser, S., Ziv, C., Creveld, S. G. van, & Vardi, A. (2016). Virocell Metabolism: Metabolic Innovations During Host-Virus Interactions in the Ocean. *Trends in Microbiology*, 24(10), 821–832. <https://doi.org/10.1016/j.tim.2016.06.006>
- Sardar, P., Šustr, V., Chroňáková, A., Lorenc, F., & Faktorová, L. (2022). *De novo* metatranscriptomic exploration of gene function in the millipede holobiont. *Scientific Reports*, 12(1), Art. 1. <https://doi.org/10.1038/s41598-022-19565-y>
- Scheu, S., & Wolters, V. (1991). Influence of fragmentation and bioturbation on the decomposition of <sup>14</sup>C-labelled beech leaf litter. *Soil Biology and Biochemistry*, 23(11), 1029–1034. [https://doi.org/10.1016/0038-0717\(91\)90039-M](https://doi.org/10.1016/0038-0717(91)90039-M)
- Shear, W. A. (2011). Class Diplopoda de Blainville in Gervais, 1844. In: Zhang, Z.-Q. (Ed.) Animal biodiversity: An outline of higher-level classification and survey of taxonomic richness. *Zootaxa*, 3148(1), 159. <https://doi.org/10.11646/zootaxa.3148.1.32>
- Shear, W. A., Jones, T., & Wesener, T. (2011). Glomerin and homoglomerin from the North American pill millipede *Onomeris sinuata* (Loomis, 1943) (Diplopoda, Pentazonia, Glomeridae). *International Journal of Myriapodology*, 4, 1–10. <https://doi.org/10.3897/ijm.4.1105>
- Shelley, R. M. (2003). A revised, annotated, family-level classification of the Diplopoda. *Arthropoda Selecta*, 11(3), 187–207.
- Shelley, R. M. (2011). *The milliped order Glomeridesmida (Diplopoda: Pentazonia: Limacomorpha) in Oceania, the East Indies, and southeastern Asia; first records from Palau, the Philippines, Vanuatu, New Britain, the Island of New Guinea, Cambodia, Thailand, and Borneo and Sulawesi, Indonesia.*
- Shukla, G. S., & Shukla, S. C. (1980). Morphology of the alimentary canal of Millipede *Trigoniulus lumbricinus* (GERSTÄCKER) Diplopoda). *Deutsche Entomologische Zeitschrift*, 27(1–3), 51–55. <https://doi.org/10.1002/mmnd.19800270106>
- Sierwald, P., & Bond, J. E. (2007). Current Status of the Myriapod Class Diplopoda (Millipedes): Taxonomic Diversity and Phylogeny. *Annual Review of Entomology*. <https://doi.org/10.1146/annurev.ento.52.111805.090210>
- Smit, A. M., Van Aarde, R. J., A.-M., S., R. J., V. A., Smit, A. M., & Van Aarde, R. J. (2001). The influence of millipedes on selected soil elements: A microcosm study on three species occurring on coastal sand dunes. *Functional Ecology*, 15(1), 51–59. <https://doi.org/10.1046/j.1365-2435.2001.00493.x>
- Soil, I. (2005). Chapter – III MICRO FLORA AND MESOFAUNA OF MILLIPEDES. 1–50.
- Sridhar, K. R., & Ashwini, K. M. (2016). Diversity, restoration and conservation of millipedes. In *Biodiversity in India* (Vol. 5, pp. 1–38). <https://doi.org/10.13140/RG.2.1.3683.2889>
- Sridhar, K. R., & Kadamannaya, B. S. (2011). Pill millipedes—An overview. In *Organic Farming: Methods, Economics and Structure*.

- Sun, M., Yuan, S., Xia, R., Ye, M., & Balcázar, J. L. (2023). Underexplored viral auxiliary metabolic genes in soil: Diversity and eco-evolutionary significance. *Environmental Microbiology*, 25(4), 800–810. <https://doi.org/10.1111/1462-2920.16329>
- Šustr, V., Chroňáková, A., Semanová, S., Tajovský, K., & Šimek, M. (2014). Methane production and methanogenic archaea in the digestive tracts of millipedes (Diplopoda). *PLoS ONE*, 9(7). <https://doi.org/10.1371/journal.pone.0102659>
- Šustr, V., & Šimek, M. (2009). Methane release from millipedes and other soil invertebrates in Central Europe. *Soil Biology and Biochemistry*, 41(8), 1684–1688. <https://doi.org/10.1016/j.soilbio.2009.05.007>
- Šustr, V., Šimek, M., Faktorová, L., Macková, J., & Tajovský, K. (2020). Release of greenhouse gases from millipedes as related to food, body size, and other factors. *Soil Biology and Biochemistry*, 144, 107765. <https://doi.org/10.1016/j.soilbio.2020.107765>
- Šustr, V., Stingl, U., & Brune, A. (2014). Microprofiles of oxygen, redox potential, and pH, and microbial fermentation products in the highly alkaline gut of the saprophagous larva of *Penthetria holosericea* (Diptera: Bibionidae). *Journal of Insect Physiology*, 67, 64–69. <https://doi.org/10.1016/J.JINSPHYS.2014.06.007>
- Suzuki, Y., Grayston, S. J., & Prescott, C. E. (2013). Effects of leaf litter consumption by millipedes (*Harpaphe haydeniana*) on subsequent decomposition depends on litter type. *Soil Biology and Biochemistry*, 57, 116–123. <https://doi.org/10.1016/j.soilbio.2012.07.020>
- Symstad, A. J., Tilman, D., Willson, J., & Knops, J. M. H. (1998). Species Loss and Ecosystem Functioning: Effects of Species Identity and Community Composition. *Oikos*. <https://doi.org/10.2307/3547058>
- Tokura, M., Ohkuma, M., & Kudo, T. (2000). Molecular phylogeny of methanogens associated with flagellated protists in the gut and with the gut epithelium of termites. *FEMS Microbiology Ecology*, 33(3), 233–240. <https://doi.org/10.1111/j.1574-6941.2000.tb00745.x>
- Toyota, A., & Kaneko, N. (2012). Faunal stage-dependent altering of soil nitrogen availability in a temperate forest. *Pedobiologia*, 55(3), 129–135. <https://doi.org/10.1016/j.pedobi.2011.10.007>
- Toyota, A., Kaneko, N., & Ito, M. T. (2006). Soil ecosystem engineering by the train millipede *Parafontaria laminata* in a Japanese larch forest. *Soil Biology and Biochemistry*, 38(7), 1840–1850. <https://doi.org/10.1016/j.soilbio.2005.12.015>
- Van Hoek, A. H. A. M. A. M., Van Alen, T. A., Sprakel, V. S. I. I., Leunissen, J. A. M. M., Brigge, T., Vogels, G. D., Hackstein, J. H. P. P., Van Alen, T. A., Sprakel, V. S. I. I., Leunissen, J. A. M. M., Brigge, T., Vogels, G. D., & Hackstein, J. H. P. P. (2000). Multiple acquisition of methanogenic archaeal symbionts by anaerobic ciliates. *Molecular Biology and Evolution*, 17(2), 251–258. <https://doi.org/10.1093/oxfordjournals.molbev.a026304>
- Vlaar, L. E., Bertran, A., Rahimi, M., Dong, L., Kammenga, J. E., Helder, J., Goverse, A., & Bouwmeester, H. J. (2021). On the role of dauer in the adaptation of nematodes to a parasitic lifestyle. *Parasites & Vectors*, 14(1), 554. <https://doi.org/10.1186/s13071-021-04953-6>
- Wang M., Fu S., Xu H., Wang M., Shi L., & College of Environment and Planning, Henan University, Kaifeng, Henan 475004. (2018). Ecological functions of millipedes in the terrestrial ecosystem. *Biodiversity Science*, 26(10), 1051–1059. <https://doi.org/10.17520/biods.2018086>

- Weiss, M. R. (2006). Defecation behavior and ecology of insects. *Annu. Rev. Entomol.*, *51*, 635–661. <https://doi.org/10.1146/annurev.ento.49.061802.123212>
- Wertz, J. T., Kim, E., Breznak, J. A., Schmidt, T. M., & Rodrigues, J. L. (2012). Genomic and physiological characterization of the Verrucomicrobia isolate *Diplosphaera colitermitum* gen. Nov., sp. Nov., reveals microaerophily and nitrogen fixation genes. *Applied and Environmental Microbiology*, *78*(5), 1544–1555. <https://doi.org/10.1128/AEM.06466-11>
- White, B. A., Cann, I. K. O., Kocherginskaya, S. A., Aminov, R. I., Thill, L. A., Mackie, R. I., & Onodera, R. (1999). Molecular analysis of Archaea, Bacteria and Eucarya communities in the rumen-review. *Asian-Australasian Journal of Animal Sciences*, *12*(1), 129–138. <https://doi.org/10.5713/ajas.1999.129>
- White, M. M., Cafaro, M. J., & Lichtwardt, R. W. (2000). Arthropod gut fungi from Puerto Rico and summary of tropical trichomycetes worldwide. *Caribbean Journal of Science*.
- Wilhartitz, I., Mach, R. L., Teira, E., Reinthaler, T., Herndl, G. J., & Farnleitner, A. H. (2007). Prokaryotic community analysis with CARD-FISH in comparison with FISH in ultra-oligotrophic ground- and drinking water. *Journal of Applied Microbiology*, *103*(4), 871–881. <https://doi.org/10.1111/j.1365-2672.2007.03319.x>
- Wilson, H. M. (2006). Juliformian millipedes from the Lower Devonian of Euramerica: Implications for the timing of millipede cladogenesis in the Paleozoic. *Journal of Paleontology*, *80*(4), 638–649. [https://doi.org/10.1666/0022-3360\(2006\)80\[638:JMFTLD\]2.0.CO;2](https://doi.org/10.1666/0022-3360(2006)80[638:JMFTLD]2.0.CO;2)
- Wright, K. A. (2011). Trichomycetes and Oxyuroid Nematodes in the Millipede , *Narceus annularis* In 1853 Joseph Leidy described the enteric flora and fauna of some arthropods , including those of the millipede *Julus marginatus*. At that time he elegantly illustrated the assoc. *Proceedings of the Helminthological Society of Washington*, *46*(2), 213–223.
- Zheng, X., Jahn, M. T., Sun, M., Friman, V.-P., Balcazar, J. L., Wang, J., Shi, Y., Gong, X., Hu, F., & Zhu, Y.-G. (2022). Organochlorine contamination enriches virus-encoded metabolism and pesticide degradation associated auxiliary genes in soil microbiomes. *The ISME Journal*, *16*(5), Art. 5. <https://doi.org/10.1038/s41396-022-01188-w>







## 5 Attached publications

### 5.1 Paper No. 1

**Nweze JE**, Šustr V, Brune A, Angel R. Functional similarity despite taxonomical divergence in the millipede gut microbiota points to a common trophic strategy (2024). *Microbiome* 12, 16. <https://doi.org/10.1186/s40168-023-01731-7> (IF 15.5).



RESEARCH

Open Access



# Functional similarity, despite taxonomical divergence in the millipede gut microbiota, points to a common trophic strategy

Julius Eyiuche Nweze<sup>1,2</sup>, Vladimír Šustr<sup>1</sup>, Andreas Brune<sup>3</sup> and Roey Angel<sup>1,2\*</sup>

## Abstract

**Background** Many arthropods rely on their gut microbiome to digest plant material, which is often low in nitrogen but high in complex polysaccharides. Detritivores, such as millipedes, live on a particularly poor diet, but the identity and nutritional contribution of their microbiome are largely unknown. In this study, the hindgut microbiota of the tropical millipede *Epibolus pulchripes* (large, methane emitting) and the temperate millipede *Glomeris connexa* (small, non-methane emitting), fed on an identical diet, were studied using comparative metagenomics and metatranscriptomics.

**Results** The results showed that the microbial load in *E. pulchripes* is much higher and more diverse than in *G. connexa*. The microbial communities of the two species differed significantly, with *Bacteroidota* dominating the hindguts of *E. pulchripes* and *Proteobacteria* (*Pseudomonadota*) in *G. connexa*. Despite equal sequencing effort, de novo assembly and binning recovered 282 metagenome-assembled genomes (MAGs) from *E. pulchripes* and 33 from *G. connexa*, including 90 novel bacterial taxa (81 in *E. pulchripes* and 9 in *G. connexa*). However, despite this taxonomic divergence, most of the functions, including carbohydrate hydrolysis, sulfate reduction, and nitrogen cycling, were common to the two species. Members of the *Bacteroidota* (*Bacteroidetes*) were the primary agents of complex carbon degradation in *E. pulchripes*, while members of *Proteobacteria* dominated in *G. connexa*. Members of *Desulfobacterota* were the potential sulfate-reducing bacteria in *E. pulchripes*. The capacity for dissimilatory nitrate reduction was found in *Actinobacteriota* (*E. pulchripes*) and *Proteobacteria* (both species), but only *Proteobacteria* possessed the capacity for denitrification (both species). In contrast, some functions were only found in *E. pulchripes*. These include reductive acetogenesis, found in members of *Desulfobacterota* and *Firmicutes* (*Bacillota*) in *E. pulchripes*. Also, diazotrophs were only found in *E. pulchripes*, with a few members of the *Firmicutes* and *Proteobacteria* expressing the *nifH* gene. Interestingly, fungal-cell-wall-degrading glycoside hydrolases (GHs) were among the most abundant carbohydrate-active enzymes (CAZymes) expressed in both millipede species, suggesting that fungal biomass plays an important role in the millipede diet.

**Conclusions** Overall, these results provide detailed insights into the genomic capabilities of the microbial community in the hindgut of millipedes and shed light on the ecophysiology of these essential detritivores.

**Keywords** Polysaccharide degradation, Hindgut microbiota, Millipede holobiont, Symbiosis, Glycoside hydrolases, Nutrient cycling, Acetogens, Ecosystem engineers

\*Correspondence:

Roey Angel  
roey.angel@bc.cas.cz

Full list of author information is available at the end of the article



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

## Introduction

Plant litter is the primary source of food and shelter for detritivorous animals [1], of which millipedes are one of the largest and most diverse members [2]. However, detritivores generally lack enzymes to digest complex polysaccharides [3, 4], which make up most of the plant litter biomass [5]. Instead, many rely on their gut microbiome to break down various hydrocarbon substrates [6] and release simple sugars or short-chain fatty acids that the host can absorb [7, 8]. In arthropods, the gut microbiome plays an important role in the development and adaptation of the host to its trophic niche [9–11]. Like all other soil arthropods, millipedes (class: Diplopoda) host a diverse community of microorganisms in their guts, which may be essential to the host's nutrition [12, 13]. In millipedes, the midgut and hindgut compartments are colonized by a dense population of aerobic and anaerobic bacteria, with the highest microbial density found in the hindgut [12].

Unlike the microbiome of other important detritivores, primarily termites [14] and earthworms [15], the millipede microbiome has received little attention so far. Only a handful of prokaryotic surveys were conducted, mostly using basic culture-dependent and molecular fingerprinting techniques [16]. Since many host-associated microorganisms cannot be grown outside their hosts, our knowledge remains limited. Recent studies have reported the most prevalent taxa in the millipede species *Anadenobolus monilicornis* [12] (only a preprint of a metagenomic study is available) and *Telodeinopus aoutii* [17] (only transcriptome data).

Freshly fallen leaf litter or wood bark contains mainly pectin, starch, cellulose, hemicellulose, and lignin [5]. The latter three are insoluble and chemically recalcitrant due to their dense structure [18] and are typically only hydrolyzed by microorganisms [19]. Indeed, gut extracts and even cultivated aerobes from several millipedes were shown to hydrolyze cellulose, hemicelluloses, and pectin [20–23]. These reports were supported by a recent metatranscriptomic study, where bacteria were shown to be the primary producers of hydrolytic enzymes in the tropical millipede *Telodeinopus aoutii* [24].

However, whether millipedes—like termites [25]—benefit directly from the lignocellulolytic activity of their gut microbiota or even rely on it as a primary source of nutrition remains an open question. Several researchers have hypothesized in the past that millipedes ingest litter primarily as a means of providing a substrate for microorganisms (bacteria, fungi, and lichens), which in turn serve as their food source [26, 27]. Accordingly, the central role of the millipede is to mix the litter layers, mechanically fragment the plant material, and inoculate the pieces with gut bacteria and fungi. If correct, we expect to see

an expression of glucanase and chitinase genes related to fungal cell wall degradation [28].

Despite the progress made in understanding the eco-physiology of the millipede holobiont, it remains unclear whether millipedes rely on fermentative degradation of cellulose to generate volatile fatty acids for their nutrition. Despite their common detritivorous lifestyle, some species were shown to be CH<sub>4</sub> emitters, while others were not, which has been attributed to differences in size and the resulting redox conditions in their digestive tracts [29, 30]. Since CH<sub>4</sub> is an end product of the cellulose degradation cascade under anaerobic conditions, a lack of methane production could indicate differences in the underlying microbial fermentations.

In addition to providing the enzymes required for the digestion of lignocellulose, millipede gut bacteria may also play other nutritional roles, such as fixing nitrogen and recycling nutrients, that compensate for their nitrogen-poor diets [31, 32]. However, it is unknown if millipedes can fix and recycle nitrogen, and a comprehensive molecular approach is needed to provide answers to these questions.

In this study, we used metagenomic and metatranscriptomic sequencing of the gut microbiome of two millipede model species to shed light on their metabolic potential and better understand the trophic niche of these keystone detritivores. We analyzed individuals of lab-maintained *Epibolus pulchripes* (order: Spirobolida) and *Glomeris connexa* (order: Glomerida). Both feed on senescent leaves but differ in size and habitat. *E. pulchripes* is a fairly large (130–160 mm) tropical millipede, widely spread along the East African coast [33], which has been shown to be a strong methane emitter [29]. *G. connexa* is a small (10–17 mm) species common to Central Europe [34] that was shown to be a non-methane emitter [30]. Our analysis covered genes involved in carbon, sulfur, and nitrogen cycling. In particular, we focused on carbohydrate-active enzymes (CAZymes) with secretion signal peptides targeting substrates from plant, fungi, and microbial origin.

## Methods

### Millipede sources and rearing conditions

Juvenile individuals of the tropical millipede *Epibolus pulchripes* were obtained from a breeding colony maintained in our lab. The animals are kept in a plastic terrarium (60×30×20 cm) on a forest floor substrate with peat, rotten wood, and a blend of leaf litter from maple, oak, Canadian poplar, and beech trees. The environment was maintained at 25 °C and subjected to a 12-h photoperiod under controlled conditions. Moisture was maintained by regularly spraying with tap water. The temperate *Glomeris connexa* was collected from a forest

near the Helfenburk castle near Bavorov (49° 8′ 10.32″ N, 14° 0′ 24.21″ E) in the Czech Republic. The collection of this species required no special permission. Both millipedes were identified down to the species level based on morphological features ([35, 36]; data not shown).

The animals were maintained in the laboratory for 25 days before dissection. Both species were kept in the lab in plastic terraria with aeration holes. The boxes contained commercial fine sand and *Populus x canadensis* (Canadian poplar) leaf litter. High humidity was maintained by spraying with tap water every other day. Both species were kept at near-optimal temperatures: *E. pulchripes* was kept at 25 °C in a light-regulated room, with a maximum of one individual in a box (19.3 × 13.8 × 5 cm). For *G. connexa*, five individuals were kept in a box (15 × 10 × 4 cm) at 15 °C in an incubator.

#### Acetylene reduction assay

ARA was performed as previously described [37] by placing a single millipede in 100 ml Schott DURAN borosilicate glass bottles, with or without leaf litter and supplementing the headspace with 4% acetylene (final conc.). Ethylene accumulation was measured at 0, 4, and 6 h by directly injecting 500 µl headspace gas into a GC (HP 5890 Series II equipped with a Porapak N column and an FID detector, Hewlett Packard).

#### Bacterial counts

Three pellets of fresh feces were collected from the millipede boxes at once using sterilized tweezers, suspended in 1 ml of phosphate buffer (pH 7.4) and plated 20 µl in triplicates on Lysogeny broth (LB) agar, and incubated at 25 °C. After 16 h, the colonies from each pellet were counted.

#### Nucleic acid extraction

Three replicates were analyzed for each millipede species. Because of the difference in body size, a single individual of *E. pulchripes* and five of *G. connexa* (from the same rearing box) were considered technical replicates. Animals were dissected, according to Sardar et al. [17]. The intact hindguts were separated and stored at −20 °C until nucleic acid extraction. The total nucleic acids (TNA) were extracted from the hindguts and feces, purified, and quantified according to Angel et al. [38]. Briefly, each sample (0.677–1.108 g for *E. pulchripes* and 0.083–0.092 g for *G. connexa*) was subjected to 3-consecutive bead beating rounds (Lysing Matrix E tubes; MP Biomedicals™) in a FastPrep-24™ 5G (MP Biomedicals™) in the presence of CTAB, phosphate buffer (pH 8.0), and phenol. The extract was then purified using phenol–chloroform–isoamyl alcohol (25:24:1; Thermo Scientific™), precipitated using a PEG solution with Invitrogen™

UltraPure™ Glycogen (Thermo Fisher Scientific) as a co-precipitant and purified using OneStep™ PCR Inhibitor Removal Kit (Zymo Research). The complete protocol is available online [39]. The quantity and quality of the DNA were determined using the Quant-iT™ PicoGreen HS Assay Kit (Thermo Fisher Scientific™) and the Agilent 2100 Bioanalyzer (Agilent). RNA was purified from the TNA extracts using TURBO™ DNase and the GeneJET RNA Cleanup and Concentration Micro Kit (Thermo Fisher Scientific). The RNA was quantified using the Quant-it RiboGreen RNA Assay Kit (Thermo Fisher Scientific). The quality of the RNA was evaluated by Novogene Sequencing – Europe (Cambridge, UK) using agarose gel electrophoresis and Agilent 2100 Bioanalyzer.

#### Amplicon library preparation, gene quantification, and sequencing

The bacterial diversity in the hindgut compartments from the two millipede species was analyzed by paired-end sequencing of the V4 region of the 16S rRNA genes on an Illumina MiniSeq platform (2 × 250 cycle configuration; V2 reagent kit; Illumina) at the DNA Services Facility at the University of Illinois, Chicago, USA (Table S1), following Naqib et al. [40]. After quantifying the DNA with PicoGreen, the samples were diluted to a final concentration of 10 ng µl<sup>−1</sup>. For PCR and library preparation, the primers 515F\_mod and 806R\_mod [41] were used to amplify the V4 region of the 16S rRNA gene. For gene quantification, the template DNA was diluted to 0.01 ng µl<sup>−1</sup>, and 2 µl was used per reaction with primers 338F–805R (0.5 µM), the 516P FAM/BHQ1 probe (0.2 µM) together with the digital droplet PCR Supermix for probes (Bio-Rad), and quantified on a QX200 AutoDG Droplet Digital PCR System (ddPCR; Bio-Rad). The full protocol can be found online [42]. The copy numbers of 16S rRNA were normalized for 1 ng of total DNA.

#### Library preparation and sequencing for metagenome and metatranscriptome

Library preparations, sequencing of the metagenomes and metatranscriptomes (see below), and quality control were provided by Novogene (UK) Company Limited. Metagenomic libraries were prepared using the same DNA preparations described above. Sequencing libraries were generated using NEBNext® Ultra™ DNA Library Prep Kit by Illumina (NEB, USA) following the manufacturer's recommendations, and index codes were added to attribute sequences to each sample. The libraries were pooled and sequenced on an Illumina NovaSeq PE 150 platform, generating an average of 50.3 G for *E. pulchripes* and 41.3 G base pairs for *G. connexa*.

Metatranscriptomic libraries were prepared using quality-controlled RNA preparations at the Novogene



(UK) Company Limited. The RNA was sequenced on an Illumina NovaSeq PE150 platform and generated 321.8 G reads. Briefly, three sample quality control methods were used: nanodrop, Agarose Gel Electrophoresis, and Agilent Bioanalyzer 2100. The rRNA was depleted using the Ribo-Zero kit (Thermo Fisher Scientific). Sequencing libraries were generated using the NEB-Next® Ultra™ RNA Library Prep Kit for Illumina® (NEB, Ipswich, MA, USA) following the manufacturer's instructions, and index codes were added to assign sequences to specific samples. The quality control for the library preparation included quantification and integrity evaluation using Qubit 2.0 (Thermo), Agilent 2100 Bioanalyzer System (Agilent Technologies), and qPCR to exclude DNA contamination.

#### Reconstruction of metagenome-assembled genomes

The raw sequence reads (from the metagenome and metatranscriptome libraries) were quality-filtered using Trimmomatic v0.39 [43]. Quality-filtered metagenomic reads from both millipede species were uploaded into *anvi'o* v7 metagenomic workflow [44], co-assembled (de novo) with MEGAHIT v1.2.9 [45], and assembled contigs < 1 kbp were removed. Bowtie2 V2.3.4.3 [46] was used for mapping the quality trimmed reads to the initial co-assembled contigs (before removing potential eukaryotic contigs) and SAMtools v2.4.2 [47] to sort the output SAM files into BAM files. We used the *anvi-display-contigs-stats* function to get a summary of contigs statistics from the co-assembly of each millipede separately and both species together. The open reading frames were identified with Prodigal v2.6.3 [48] and single-copy core genes (SCG) with HMMER v3.3.2 [49]. The gene-level taxonomy was predicted using Centrifuge v1.03-beta [50] and annotated functions using the NCBI's Clusters of Orthologous Groups (COG) [51] and KEGG Orthologs (KOs) databases [52]. Both Metabat2 v2.12.1 [53] and CONCOCT V0.38 [54] were used to create contigs clusters (bins) and the *anvi'o* interactive interface to refine the bins manually. Comparing the two methods, Metabat2 yielded higher quality MAGs, while many CONCOCT MAGs suffered from high contamination levels and taxonomic misclassification (data not shown). Therefore, only the Metabat2 MAGs were kept for downstream analysis. We retained all prokaryotic metagenome-assembled genomes (MAGs) with more than 50% completion and redundancy in SCG below 10% based on CheckM [55]. The *anvi-gen-phylogenomic-tree* function was used to plot a phylogenomic tree by concatenating 39 single-copy genes from *Bacteria\_71* (ribosomal proteins) from the recovered MAGs.

#### Taxonomic classification of sequence data

Unless mentioned otherwise, all data processing steps and plotting were done in R [56]. After amplicon sequencing, the 16S rRNA reads were demultiplexed using *cutadapt* V3.5 [57]. The raw reads were processed, assembled, and filtered using DADA2 v1.26, with the standard filtering parameters, according to Callahan [58]. Unique sequences were identified and clustered into amplicon sequence variants (ASV). Chimaeras were removed with the *removeBimeraDenovo* function. The quality-filtered pair-end reads were classified to the genus level using the Genome Taxonomy Database (GTDB) [59]. The resulting tables were merged into a *Phyloseq* object [60]. Decontamination was done using *decontam* v1.18 [61]. After rarefying the dataset without replacement, we calculated the Principal Coordinates Analysis (PCoA) using the unweighted UniFrac as distance [62].

To profile the prokaryotic community in the metagenome, the quality-filtered metagenomic reads from each millipede species were again co-assembled (de novo) with MEGAHIT v1.2.9 [45], and assembled contigs < 1 kbp were removed. For metatranscriptome, remnant rRNA reads were removed using *SortMeRNA* v4.3.4 [63], and the resulting non-rRNA reads for each millipede species were co-assembled (de novo) using *Trinity* v2.13.2 [64]. Next, all contigs sized < 500 bp were discarded. Potential eukaryotic contigs from both library types were removed using *Whokaryote* [65], which does not consider contigs with less than two genes. The prokaryotic contigs were taxonomically classified using the *Contig Annotation Tool (CAT)* v5.2.3 [66] based on the GTDB. Unclassified clades and clades with < 200 contigs at the phylum level were ignored. The taxonomy of the MAGs was inferred using *GTDB-Tk* v2.0.0 [67], which uses a 95% ANI cutoff for the species boundary to determine the phylogenetic placement and relative evolutionary divergence (RED) values of query genomes in the GTDB reference tree [58]. Genomes were defined as novel genera (all MAGs clustered at 60% AAI [68] without a genus GTDB-Tk assignment), novel species (GTDB-Tk ANI output < 95%), and novel strains (GTDB-Tk ANI output < 99%) [69]. The eukaryotic community structure in the metagenomes was determined using *METAXA2* [70], extracting SSU and LSU rRNA sequences. These were attributed to various origins. Non-bacterial rRNA sequences were validated via a blast analysis [71]. Metatranscriptomic reads were taxonomically classified using local *blastn* against the *nodes.dmp* and *names.dmp* database files.

#### Functional annotation

The prokaryotic co-assembled reads from metagenome and MAGs were profiled for functional traits and

metabolism using the METABOLIC v4.0 pipeline with default parameters [72]. We predicted the genes for carbohydrate-active enzymes (CAZymes), such as glycoside hydrolases (GHs), carbohydrate-binding modules (CBMs), polysaccharide lyases (PLs), and carbohydrate esterases (CEs), based on the dbCAN2 meta server. Along with this, signal peptide predictions were based on the SignalP 4.1 databases [73]. Further screening of CAZyme genes was performed manually, and CAZymes were defined as those predicted by at least two tools. Lastly, the putative substrates for the glycoside hydrolases were predicted based on information from the literature.

The gene homologs for acetogenesis, hydrogenases, nitrogen, and sulphur cycling were also predicted. Six genes for reductive acetogenesis absent in the pipeline were annotated with blastp [74] at an *e* value of  $1e-30$ . Blastp [71] was also used to confirm all the predictions from METABOLIC (maximum of five target sequences). The resulting data were imported and plotted in the R packages ggplot2 [75], circlize [76], and iTOL [77].

**Relative abundance of MAGs and gene families**

To determine the relative abundance of the MAGs in both metagenome and metatranscriptome, we mapped the reads from both library types to the MAGs. Each MAG's sample-specific mean coverage was used to calculate its relative abundance using CoverM v0.6.1 (<https://github.com/wwood/CoverM>) with default parameters of the coverm-genome function. CoverM used Minimap2 [78] for mapping and calculating read coverage per

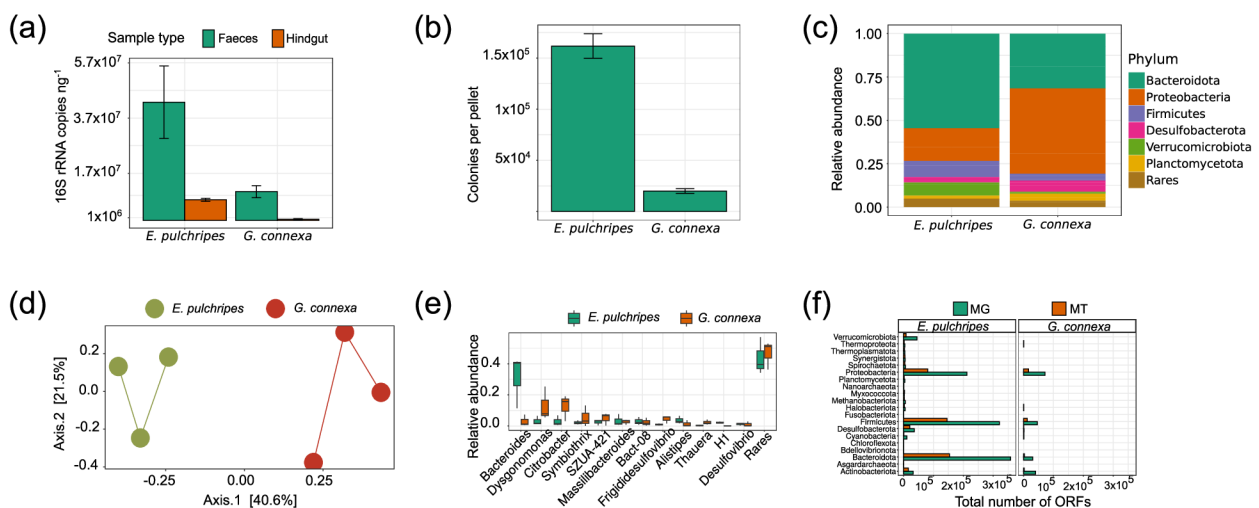
genome (relative abundance) with a  $-min-read-percent-identity$  of 90%.

For the abundance of the gene families in both metagenome and metatranscriptome, the reads from both library types were mapped to each gene, and the mean coverage was used to estimate the relative abundance in Transcripts per million (TPM) using CoverM within contig with  $-min-read-percent-identity$  of 90% to allow comparing the datasets. CoverM used the bwa-mem aligner [79]. TPM in gut metagenomes reflects the relative abundance of a gene in the bacterial community. Genes with a zero TPM value were removed, and the values were converted to  $\log(TPM + 1)$  for plotting. The quantification of SSU/LSU rRNA sequences in the metagenome and eukaryotic contigs in the metatranscriptome followed a similar methodology.

**Results**

**Bacterial load and 16S rRNA gene diversity in the millipede guts**

Quantification of the 16S rRNA gene copies in the hindgut using ddPCR yielded  $0.74 \times 10^7$  in *E. pulchripes* and  $0.39 \times 10^7$  per ng DNA in *G. connexa* (Fig. 1a; Table S1). The respective microbial load in the feces was 6 and 26 times higher than in the hindguts. Comparing these numbers with the number of viable bacteria in the feces (using a number of colonies per fecal pellet) revealed similar differences in microbial load and that a large proportion of the bacteria remained uncultured (Fig. 1b, Table S1).



**Fig. 1** Microbial load and community composition of the gut microbiome in *Epibolus pulchripes* and *Glomeris connexa*. **a** 16S rRNA gene copies in the hindgut and faecal samples from *E. pulchripes* and *G. connexa*. **b** Average colony counts from faeces samples grown on LB agar media. **c** Relative abundance of bacteria from the hindgut at phylum level based on Illumina sequencing of the 16S rRNA gene. **d** PCoA-analysis based on unweighted UniFrac distances between the microbial communities in the hindguts of both millipede species. **e** The relative abundance of dominant bacteria at the genus level. **f** Taxonomic classification of the prokaryotic community in the assembled metagenomes (MG) and metatranscriptomes (MT) from hindgut samples of *E. pulchripes* and *G. connexa* based on the total ORFs

The amplicon sequencing of the V4 region of the bacterial and archaeal 16S rRNA gene amplified from the millipede hindguts yielded an average of 38,317 high-quality reads. Of the 16 phyla represented in the dataset, the majority of sequences in *E. pulchripes* and *G. connexa* datasets were *Bacteroidota* (54.6% and 31.6%) and *Proteobacteria* (*Pseudomonadota*; 18.8% and 49.2%), followed by *Verrucomicrobiota* (7.6% and 1.1%), *Firmicutes* (*Bacillota*; 9.4% and 4.0%), *Desulfobacteriota* (3.1% and 6.4%), and *Plancomycetota* (1.8% and 4.1%; Fig. 1c, Table S2). However, despite the many shared phyla, the communities differed significantly on the genus level, with only 14.24% shared between the species (Fig. 1d). Following the trends on the phylum level, the two species differed in the relative abundance of these common genera. In *E. pulchripes*, the community was highly dominated by *Bacteroides* (*Bacteroidaceae*, 31.1%), followed by more minor members such as *Alistipes* (*Rikenellaceae*, 3.7%), *Massilibacteroides* (*Tannerellaceae*, 3.3%), *Dysgonomonas* (*Dysgonomonadaceae*, 3.1%), and others (Fig. 1e). In contrast, the distribution of genera in *G. connexa* was shallower and dominated by *Dysgonomonas* (12.9%) and *Citrobacter* (12.6%), and others.

#### Quality of metagenome and metatranscriptome assemblies

The six metagenomic libraries from the two millipede species yielded 0.63 G paired-end reads (Table S3). *E. pulchripes* samples were separately assembled with Megahit into 823.6 K contigs (total length – 2.9 Gb). The reads from *G. connexa* were assembled into 162.8 K contigs (total length – 0.5 Gb). Meanwhile, the metatranscriptomes yielded an average of 0.35 G paired-end. In *E. pulchripes*, the reads were assembled into 1.3 M contigs, while in *G. connexa*, the assembled reads constituted 136.0 K contigs (Table S4).

#### Microbial abundance in metagenomic and metatranscriptomic reads across hindgut samples

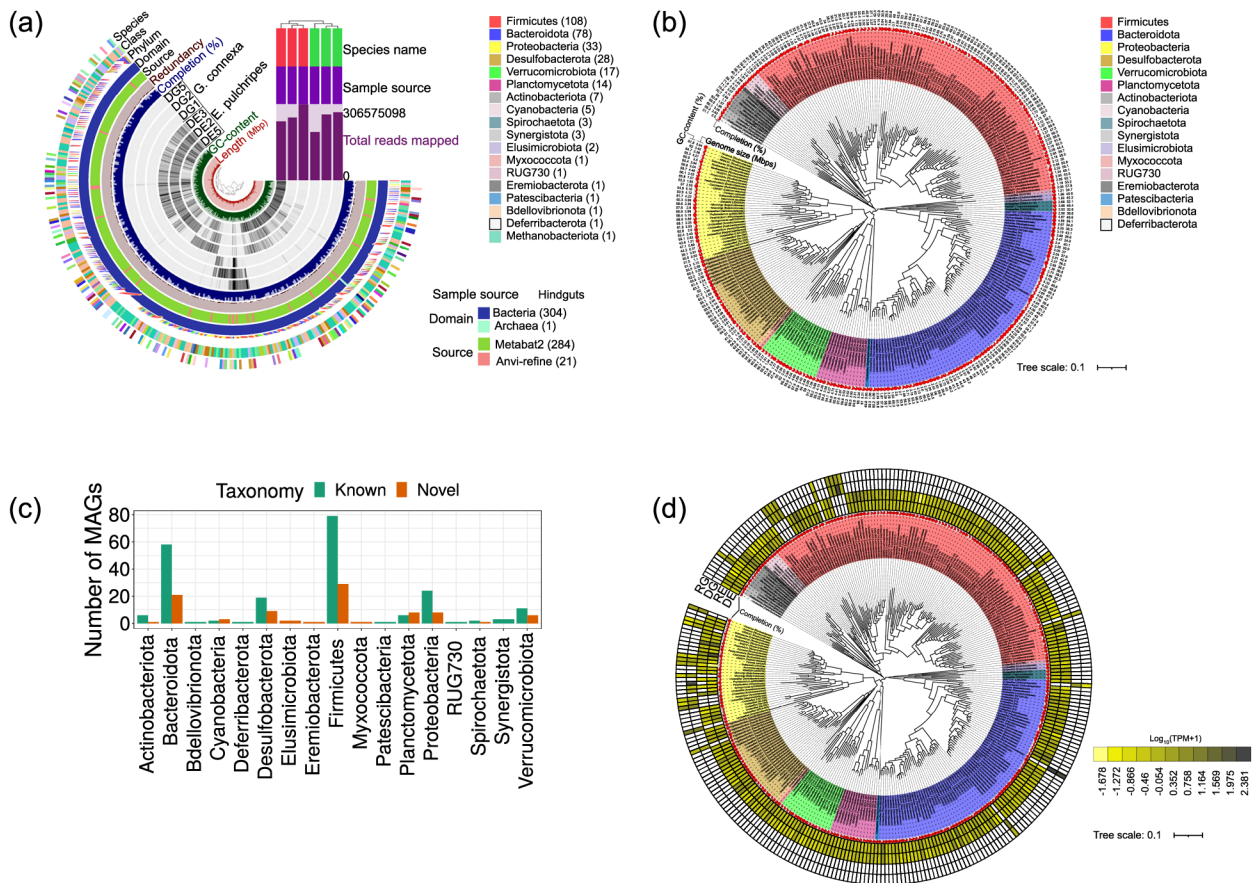
Filtering contigs of eukaryotic origin yielded 338,035 (41%) prokaryotic contigs for *E. pulchripes* and 62,892 (39%) for *G. connexa* (Table S5). For both species, 95% of the metagenomic contigs could be taxonomically assigned. In contrast, the metatranscriptomes contained only 17% and 7% of prokaryotes that also passed size and quality filtering. Of those, 58% in *E. pulchripes* and 74% in *G. connexa* could be taxonomically assigned, at least at the phylum level. The taxonomic classification of the contigs resembled the composition obtained from the amplicon sequencing, except for *Firmicutes*, which were under-represented in our amplicon library compared to the metagenome and metatranscriptome. Namely, in *E. pulchripes*, *Bacteroidota* was the most abundant and

active member of the community, with 34% of the total ORFs in both metagenomes and metatranscriptomes (Fig. 1e), followed by *Firmicutes* (30.3% and 32.1%), *Proteobacteria* (20% and 18%), *Verrucomicrobiota* (4.3%, 1.9%), *Desulfobacterota* (3.4% and 4.7%), and *Actinobacteriota* (3.1% and 3.6%). In samples from *G. connexa*, *Proteobacteria* was the highest-ranked taxon with 71,847 ORFs (37.1%) and 16,777 (47.1%) in metagenomes and metatranscriptomes, followed by *Firmicutes* (24% and 36%), *Actinobacteriota* (21% and 6.4%), *Bacteroidota* (16% and 7%), and *Desulfobacterota* (1% and 2.1%). In terms of non-bacterial diversity, as expected, methanogenic *Euryarchaeota* (mainly orders *Methanobacteriales*, *Methanomassiliococcales*, and *Methanosarcinales*) were detected in *E. pulchripes*, but also some in the non-CH<sub>4</sub>-emitting in *G. connexa* (Table S5). As for fungi, the phylum *Ascomycota* was found to be the most abundant among eukaryotes (>90% of the fungal contigs) in both millipede species. In addition, *E. pulchripes* also hosted Nematoda (42.8%) and *Ciliophora* (12.6%) in large numbers, while in *G. connexa*, *Apicomplexa* (74.2%), and *Metamonada* (18.2%) were the most dominant (Fig. S1; Table S5). The protist order *Eccrinales*, typically found microscopically in millipedes, was only represented by a single (*E. pulchripes*) or double (*G. connexa*) very rare contigs and was absent in the metatranscriptome (Table S5). The metatranscriptome profiling largely agreed with the metagenome regarding the taxonomic profile, but the relative abundances were significantly different, possibly due to the small size of the eukaryotic dataset (Fig. S1; Table S5).

#### De novo assembly of genomes and phylogenomic distribution

The metagenomic reads from both millipede species were co-assembled, binned, and refined, generating 305 MAGs, each with completeness >50% and redundancy <8.5% (Fig. 2a, Table S6). Notably, 47% of these MAGs exhibited a completeness of 90% or more, while 62% of the overall MAGs attained a completeness level of 80% or higher. One MAG was assigned to archaea and the rest to bacteria. The genome sizes ranged from 0.36 to 7.76 Mbp. We concatenated the amino acid sequences of the bacterial single-copy core genes from the MAGs (Table S6) and constructed a phylogenomic tree (Fig. 2b). After assigning taxonomy with GTDB-Tk (Tables S7 and S8), 108 MAGs (35.5%) were placed into the phylum *Firmicutes*, including the families of *Lachnospiraceae*, *Ruminococcaceae*, CAG-74 and some unclassified groups. The phylum *Bacteroidota* followed with 79 MAGs (26%), represented mainly by the families of *Tannerellaceae*, *UBA932*, *Bacteroidaceae*, *Azobacteroidaceae*, and *Rikenellaceae*. Thirty-two of the





**Fig. 2** Taxonomic composition of the recovered MAGs. **a** Static images from anvi's interactive display for recovered MAGs from the hindguts of *E. pulchripes* (DE, 3, 5) and *G. connexa* (DG1, 2, 5). The tree (dendrogram) at the central section of the anvi's interactive image shows the hierarchical clustering of MAGs based on their sequence composition and their distribution across samples. From inner to outer layers: length layer (shows the actual length in Mbps), GC-content, four view layers with information about MAGs across samples (mean coverage), completion, redundancy, source (automatically binned with MetaBAT2 and manually refined), domain of the MAGs (archaea or bacteria), genome phylum, class and species based on GTDB-Tk. The bars show the total number of reads mapped, sample source (hindgut) and sample names (*E. pulchripes* and *G. connexa*) **(b)** Phylogenomic tree based on 39 concatenated bacterial single copy gene (ribosomal proteins, see Table S8). **c** Potentially novel species from the hindgut of *E. pulchripes* and *G. connexa* identified with GTDB-Tk based on relative evolutionary distance. **d** The relative abundance of MAGs in the metagenomic and metatranscriptomic read samples, estimated for each sample replicate and the average was used in the plotting. DE and RE indicate the relative abundance of MAGs in the metagenome and metatranscriptomes of *E. pulchripes*, whereas DG and RG indicate the relative abundance of MAGs in the metagenome and metatranscriptomes of *G. connexa*

MAGs (10.5%) were placed into the *Proteobacteria*. In particular, we identified the *Alphaproteobacteria*, *Rs-D84*, *Beijerinckiaceae*, *Acetobacteraceae*, and the *Gammaproteobacteria*, *Enterobacteriaceae*, *Rhodocyclaceae*, and *Burkholderiaceae*. Other core phyla were *Desulfobacterota* (28 MAGs, 9.2%), *Verrucomicrobiota* (17 MAGs, 6%), *Planctomycetota* (14 MAGs, 5%), and *Actinobacteriota* (7 MAGs, 2.30%). Notably, ninety-one MAGs representing potentially novel bacterial species (81 in *E. pulchripes* and 10 in *G. connexa*) were identified based on relative evolutionary distance (RED) (Fig. 2c). The novel species belonged mainly to the *Firmicutes* and *Bacteroidota* phyla.

### Distribution of MAGs in metagenomes and metatranscriptomes

Of the retrieved MAGs, 272 (92%) were represented in *E. pulchripes* and only 23 in *G. connexa* (Fig. 2a and d). Ten of the MAGs were present in both species (Table S8). The MAGs were mapped to the quality-filtered metagenomic and non-rRNA pair-end reads to calculate the variation in abundance for each MAG across the samples. Additionally, the number of MAGs in each sample was estimated considering as “absent” those with abundances < 0.001% (Table S9). This analysis revealed that 66–75% and 36–44% of the metagenomic reads remained unmapped in *E. pulchripes* and *G. connexa*. In

metatranscriptomes, 71–86% and 82–85% of reads were unmapped in the samples from *E. pulchripes* and *G. connexa*. Two *Proteobacteria* MAGs from *E. pulchripes* and six MAGs from *G. connexa* (3 *Proteobacteria*, 2 *Firmicutes*, and 1 *Actinobacteriota*) remained unmapped in the metatranscriptomic samples.

### The repertoire of bacterial carbohydrate-degrading enzymes

The degradation of plant litter requires the concerted work of carbohydrate-active enzymes (CAZymes), including glycoside hydrolases (GHs), carbohydrate-binding modules (CBMs), carbohydrate esterases (CEs), glycosyltransferases (GTs), polysaccharide lyases (PLs), auxiliary activities (AAs), and S-layer homology modules (SLHs). We analyzed the MAGs for the presence of such CAZymes, focusing on proteins with secretion signal sequences (SSPs), which are secreted or targeted to other locations, such as the periplasmic space or bacterial cytoplasmic membrane [80]. However, we acknowledge that some bacterial proteins have been found to be secreted without any apparent signal peptide [81]. Annotation of the predicted amino acid sequences revealed 24,690 and 2042 CAZymes in the MAGs from *E. pulchripes* and *G. connexa*, respectively. Of these, 7721 and 352 had SSPs (Table S10). Among the potentially secreted CAZymes in *E. pulchripes*, GHs were the most abundant (82.3%), followed by CEs (8.8%), PLs (6.2%), CBMs (3.4%), GTs (1.1%), AAs (0.2%), and SLHs (0.03%). Also in *G. connexa*, GHs were the most abundant (73.6%) among the potentially secreted CAZymes. GHs were also the most abundant of all expressed CAZymes in *E. pulchripes* (80.3% of 6199) and *G. connexa* (73.4% of 215).

GHs (glycoside hydrolases) were classified into 127 families by the CAZyme database, and their substrate specificity can be predicted based on this structure (Table S11). Our annotation results showed that some GHs are located on the same MAGs with one or more CBMs, GTs, CEs, PLs, or other GHs (Table S12), suggesting that the CAZymes involved in polysaccharides degradation are organized in clusters.

In *E. pulchripes*, the majority of secreted GHs (6199) belonged to *Bacteroidota* (64%; presented in TPM), *Verrucomicrobiota* (12.2%), and *Firmicutes* (9.1%) (Fig. 3a). These same phyla also expressed the highest amount of GHs (6199), with *Bacteroidota* contributing the most (64%; presented in TPM) (Fig. 3b; Table S12). Based on the predicted substrate specificity, the secreted GHs from the MAGs assigned to *Bacteroidota* (4153) had the capability for the degradation of fungal cell walls (25%; presented in TPM), hemicellulose (17%), pectin (16%), pectin-hemicellulose (13.2%), pectin-hemicellulose-cellulose (13%), starch (6.5%), algal cell wall (4.2%),

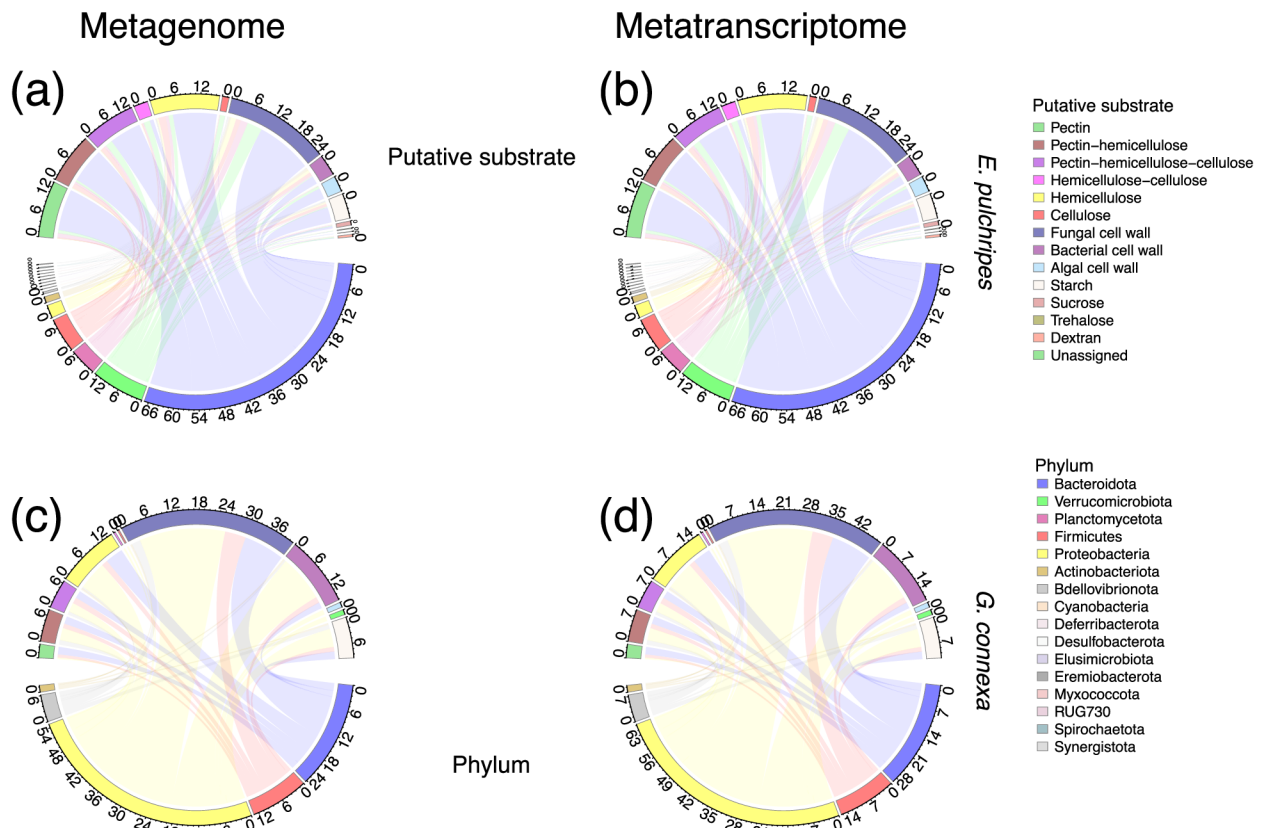
and bacterial cell wall (4.2%). The same pattern was also observed in the expressed GHs, although the relative abundance of the transcripts was lower than those in the metagenomes. *Bacteroidota*'s capabilities were mainly contributed by the families of *Bacteroidaceae* (17%; presented in TPM for metagenome), *Azobacteroidaceae* (9%), *Rikenellaceae* (16%), *UBA4181* (10.1%), *Tannerellaceae* (6.1%), and *UBA932* (3.9%). The same pattern was followed in expressing these genes (Fig. S2a and b).

GH abundance was lower in *G. connexa* (259). Approximately 67.3% (presented in TPM) of the secreted GHs were encoded in *Proteobacteria*, followed by *Bacteroidota* (12.4%), *Firmicutes* (9.4%), and *Actinobacteriota* (9.9%) (Fig. 3c). The same trend was seen in the expression of these GHs (215), with *Proteobacteria* accounting for 62% of all GHs (Fig. 3d; Table S12). The secreted GHs from the MAGs assigned to *Proteobacteria* (142) had the highest capacity to degrade fungal cell-wall (36%; presented in TPM), bacterial cell-wall (21%), hemicellulose (13%), and starch (13.3%). *Bacteroidota* also possessed the same capability. The same pattern was also observed in the expressed GHs. The hydrolytic activities of *Proteobacteria* stemmed from the families of *Enterobacteriaceae* (11.2%; presented in TPM for metagenome), *Sphingomonadaceae* (22%), *Rhizobiaceae* (13%), *Microbacteriaceae* (10%), and *Aeromonadaceae* (6.6%) (Fig. S2c). The *Bacteroidota* family, *Dysgonomonadaceae*, also possessed a high GH abundance (9.2%). The same pattern was followed in expressing these genes (Fig. S2d).

Figure 4 presents the top 50 GHs belonging to various subfamilies and their putative substrate groups. The top five most prevalent glycoside hydrolase families in *E. pulchripes* were GH43, GH13, GH5, GH3, and GH23. The family GH43, with 316 GHs and 26 subfamilies, was the most abundant. The glycoside hydrolase families GH23, GH18, and GH92 for chitin degradation were among the most abundant GHs. Their abundance in the metatranscriptomes was lower but showed a similar trend. Among the most prevalent glycoside hydrolase (GH) families in *G. connexa* were GH23 and GH18, which break down chitin, and GH13, which break down starch. Others were GH3 (hemicellulose) and GH103 (peptidoglycan). Here as well, the metatranscriptomes showed a similar pattern.

The lignin-degrading CAZymes were scarcely present in both millipede species. The auxiliary activities group of CAZymes (AAs), affiliated in part with ligninolytic activity, made up only about 0.18% (14) of the total CAZymes with SSP (Table S12) in *E. pulchripes*. This group comprised ten AA1 families multicopper oxidases (4 *Bacteroidota*, 5 *Proteobacteria*, and 1 *Verrucomicrobiota*), and one AA3 (cellobiose dehydrogenase; from *Bacteroidota*), one AA5 (galactose oxidase; from *Myxococcota*), one AA10 (lytic chitin monooxygenase; from *Proteobacteria*),





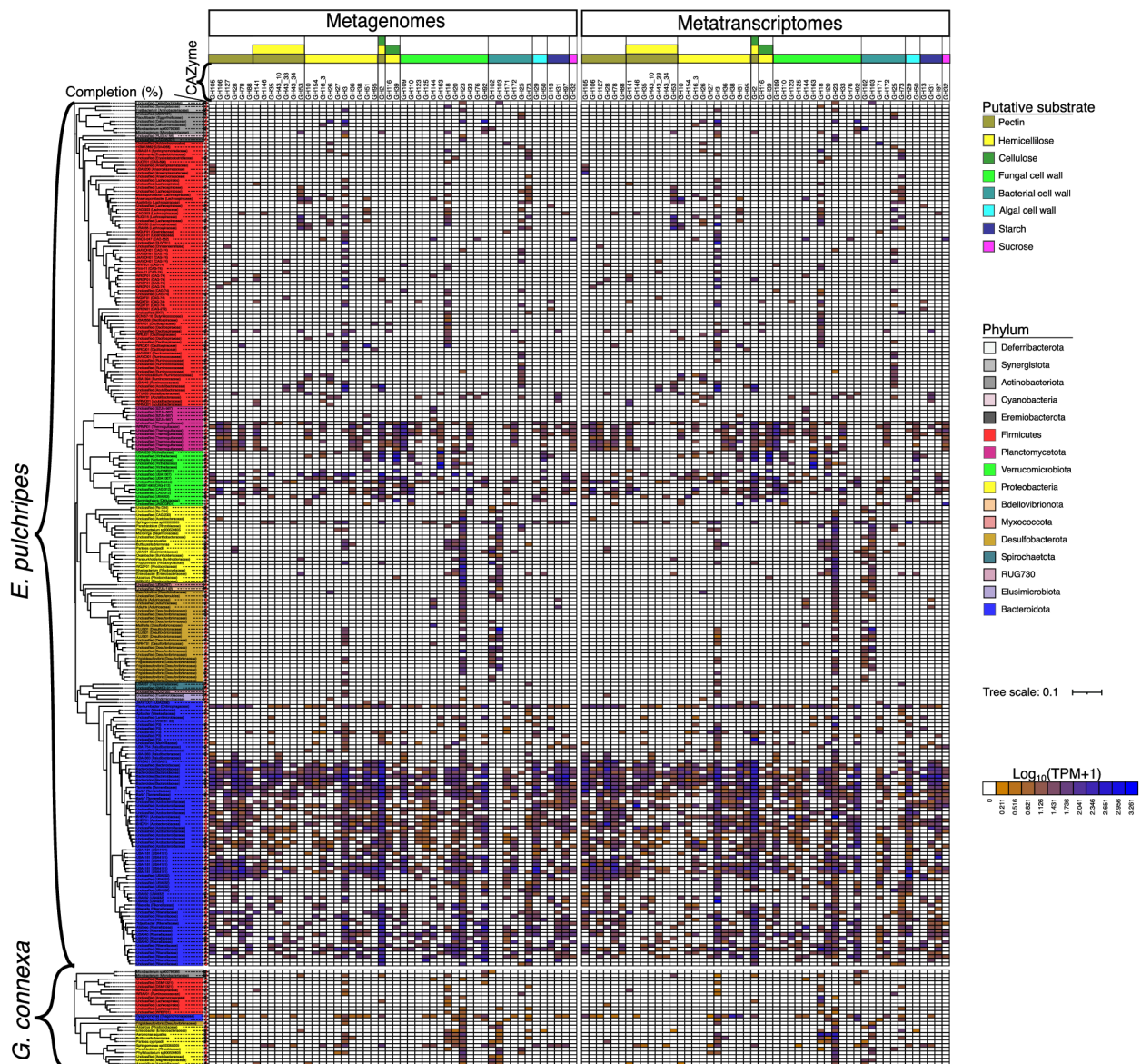
**Fig. 3** Relative abundance of glycoside hydrolases (GHs) with secretion signal peptides in metagenome-assembled genomes (MAGs) and their corresponding transcripts. The GHs were grouped at the family level and according to their putative substrates (top of the chord) and the taxa contributing to the GHs (bottom of the chord). Chord (a) displays the contribution of GHs from different phyla in metagenomes, while chord (b) shows its corresponding GH transcripts from the hindgut of *E. pulchripes*. Chord (c) shows the abundance of GHs at the phylum level in metagenomes, while chord (d) displays its corresponding GH transcripts from the hindgut of *G. connexa*. The pair-end reads of both library types were mapped to the genes to get the coverage and calculate the relative abundance in transcripts per million (TPM). The mean TPM was calculated from the three replicate samples and summed for each taxonomic level

and one AA12 (from *Bacteroidota*). In terms of relative abundance (TPM), the majority of the AAs (42.4%) were sourced from Proteobacteria, with Bacteroidota (30.5%), Myxococcota (24.4%), and Verrucomicrobiota (2.7%) contributing to a lesser extent. The corresponding transcripts also followed a similar pattern. In *G. connexa*, we found only seven AAs: six AA1 (5 *Proteobacteria*, 1 *Desulfobacterota*, and 1 *Firmicutes*) and one AA10 from *Proteobacteria*. In terms of relative abundance, the majority of the AA abundance (78%) was attributed to Proteobacteria, with Desulfobacterota (15%) and Firmicutes (8%) following behind in contribution. Once again, a similar pattern was observed for the corresponding transcripts.

**Acetogenesis in the millipede hindguts**

Acetogenesis can act as a sink for excess hydrogen produced during fermentation. We analyzed the community acetogenesis in the assembled reads and found that the key genes for heterotrophic acetogenesis were

present and expressed in the libraries of both millipede species (Fig. 5a; Tables S13 and S14). These include pyruvate:ferredoxin oxidoreductase (*porA*), phosphotransacetylase (*pta*), and acetate kinase (*ack*). A pathway involving *porA*, *pta*, *ack*, and the proteins, acetyl-CoA synthetase (ADP-forming, alpha domain) (*acdA*) and acetyl-CoA synthetase (*acs*), plays a vital role in the production and consumption of acetate through the acetate switch [82, 83]. In addition, the essential genes for reductive acetogenesis via the Wood Ljungdahl pathway (*fhs*, *folD*, *metF*, *fdhF*, and *acsABCDE*, *ack*, and *pta*) were also present and expressed in both species, except for methyltransferase (*acsE*), a subunit of acetyl-CoA synthase (*acsABCDE*) which was absent in *G. connexa*. The ability to perform heterotrophic acetogenesis was found in *Proteobacteria*, *Actinobacteria*, *Desulfobacterota*, and *Firmicutes* in *E. pulchripes*, *Proteobacteria*, and *Firmicutes* in *G. connexa*. The capacity for reductive acetogenesis was found in *Desulfobacterota*, *Actinobacteria*,



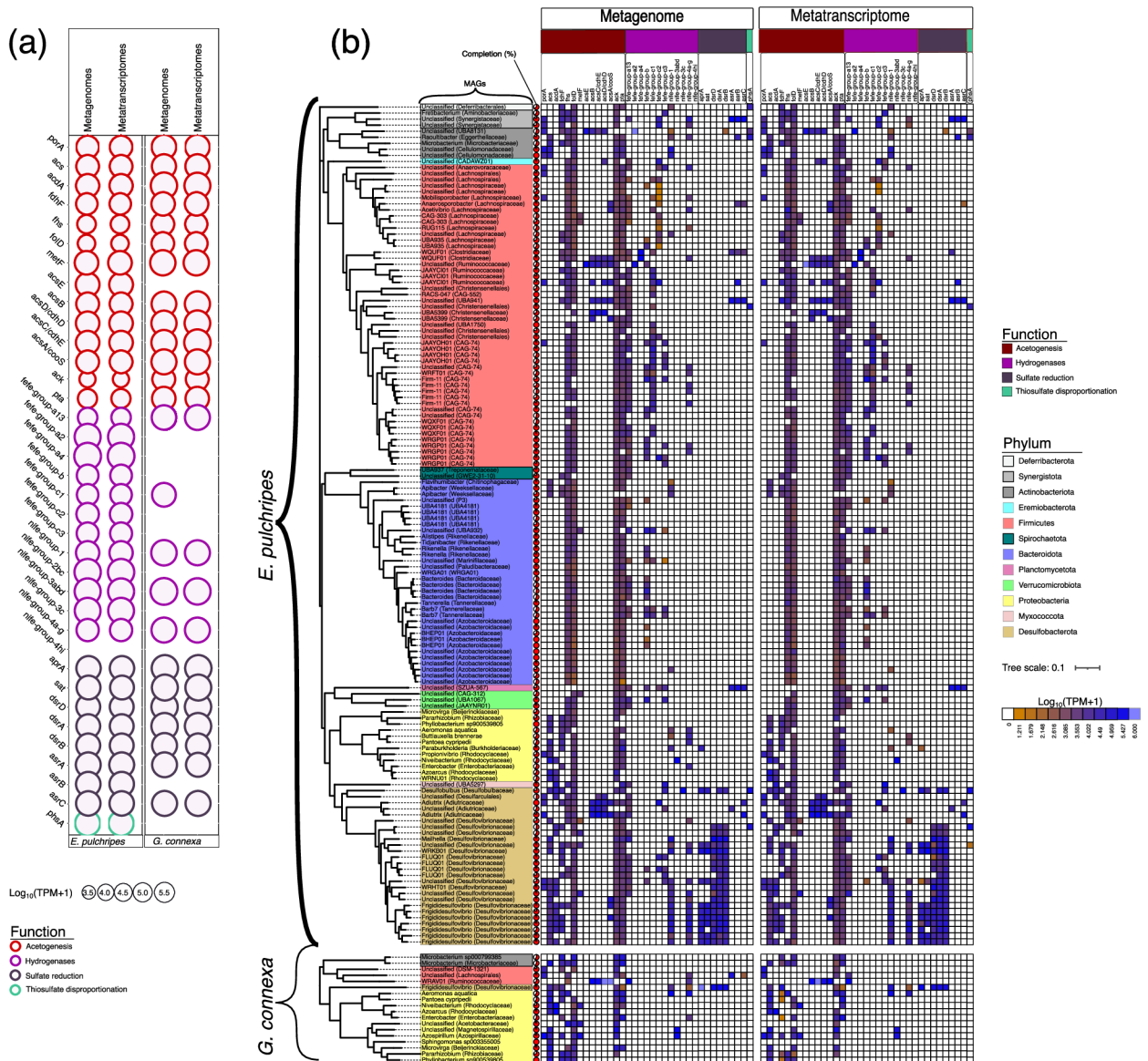
**Fig. 4** Glycoside hydrolase families and their taxonomic origin. The heatmap shows the relative abundance of the top 50 glycoside hydrolases (GHs) with a secretion signal peptide (SSP) from the MAGs and their corresponding transcripts. The GHs were grouped at the family level according to their putative substrates. The colour scale represents the log transformation of TPM +1. The tree was reconstructed using 39 concatenated bacterial single copy gene (ribosomal proteins) from our MAGs

and *Firmicutes* from *E. pulchripes*. *Proteobacteria* from *G. connexa* had all the genes for reductive acetogenesis apart from the *acsABCDE* (Fig. S3a and b).

In the MAGs from *E. pulchripes*, the genes for heterotrophic acetogenesis were encoded by a few MAGs belonging to the core phyla (Fig. 5b; Tables S13 and S14). The three critical genes for acetate production (*porA*, *pta*, and *ack*) were possessed and expressed by two *Actinobacteriota* MAGs, four *Firmicutes* MAG, four *Desulfobacterota* MAGs, one *Proteobacteria* MAG, and one

*Elusimicrobiota* MAG. Additionally, only two novel, unclassified *Desulfovibrionaceae* MAGs (*Desulfobacterota*) encoded and expressed the five genes for the production or consumption of acetate [83]. The *Firmicutes* MAGs from *G. connexa* also contained the three genes (*porA*, *pta*, and *ack*). However, these genes were not fully expressed in either of the MAGs.

We searched each MAG for the presence of seven enzymes associated with reductive acetogenesis via the Wood-Ljungdahl pathway (WLP). We found that



**Fig. 5** Abundance of gene functions involved in involved acetogenesis, hydrogenases and sulfur cycling pathways in metagenomic and metatranscriptomic libraries and MAGs **(a)** Relative abundance of genes and transcripts for acetogenesis, hydrogen sensing/evolution/bifurcation (hydrogenases) and sulfur cycling in the metagenomic (MG) and metatranscriptomic (MT) contigs from the hindguts of *E. Pulchripes* and *G. connexa*. Acetogenesis includes heterotrophic acetogenesis via a combination of glycolysis, pyruvate:ferredoxin oxidoreductase (*porA*), Phosphotransacetylase (*pta*) and acetate kinase (*ack*), and reductive acetogenesis via the Wood–Ljungdahl pathway (WLP). Full names of the gene families and their corresponding KEGG IDs are available in Table S13 **(b)** A heatmap showing the abundance of genes and transcripts in each MAG with at least four acetogenic genes or one of the sulfate-reduction genes. The tree was reconstructed using 39 concatenated bacterial single copy gene (ribosomal proteins) from the MAGs

most of the MAGs encoded and expressed *fdhF*, *fhs*, and *fold*, *ack*, *pta*, but the *metF* and *acsABCDE* were present in only a few MAGs (Fig. 5b; Tables S12 and S13). Of those MAGs, twenty in *E. pulchripes* and two in *G. connexa* encoded at least five of the gene subunits, but none of these MAGs contained the complete set of genes. In *E. pulchripes*, these MAGs included

Firmicutes (11 MAGs), *Desulfobacterota* (8 MAGs), *Bacteroidota* (1 MAG), and *Actinobacteriota* (2 MAGs). One additional *Adiutrix* MAG and two Firmicutes lacked only the genes *metF* and *acsE*, while another *Adiutrix* MAG was missing the genes *fdhF*, *metF*, and *acsE*. The MAG that encoded at least five of the genes in *G. connexa* belonged to *Actinobacteriota* and



*Firmicutes*. In both millipede species, at least five of the genes were also expressed.

#### Hydrogen metabolism in the millipede hindguts

Hydrogenases are required in various anaerobic pathways or H<sub>2</sub> uptake. Here, we also sought the community genes encoding Ni–Fe hydrogenase, Fe hydrogenase, and FeFe hydrogenase. All the hydrogenase genes (14 orthologs) identified in the community metagenomes were expressed in *E. pulchripes*, except for a [Ni–Fe] group 4 hydrogenase (Fig. 3a; Tables S12 and S13). The only subgroups present and expressed in *G. connexa* were one FeFe hydrogenase and three Ni–Fe hydrogenases.

We identified the genes encoding hydrogenases in the MAGs (Fig. 5b; Tables S13 and S14). Numerous MAGs from *E. pulchripes* encoded one or more types of hydrogenases. Groups A1 and A3 [FeFe] hydrogenases, which are common in many fermentative bacteria, were most prevalent in all MAGs, particularly in *Firmicutes*, *Bacteroidota*, *Planctomycetota*, *Desulfobacterota*, and *Verrucomicrobiota*. [FeFe] hydrogenases of Group A2 and A4 were present in a few *Firmicutes* MAGs and Group B in some *Firmicutes* and *Bacteroidota*. The second most abundant [FeFe] hydrogenases were from Group C1 and were found primarily in *Firmicutes*. [NiFe] hydrogenases were detected mostly in *Desulfobacterota* (Group 1 and Group 4) and a few *Firmicutes* (Group 4). In *G. connexa*, only two *Firmicutes* MAGs encoded a [FeFe] Group A1-hydrogenases. Other [FeFe] hydrogenases were absent. [NiFe] hydrogenases from Groups 1, 3 and 4 were found in a few MAGs of *Proteobacteria*, *Firmicutes*, and *Desulfobacterota*. [Fe] hydrogenases, which are restricted to methanogenic archaea, were absent from both millipede species. Many MAGs expressed multiple hydrogenases, including both [FeFe] and [NiFe] hydrogenases, sometimes up to three paralogues.

#### Sulfur metabolism in the millipede hindguts

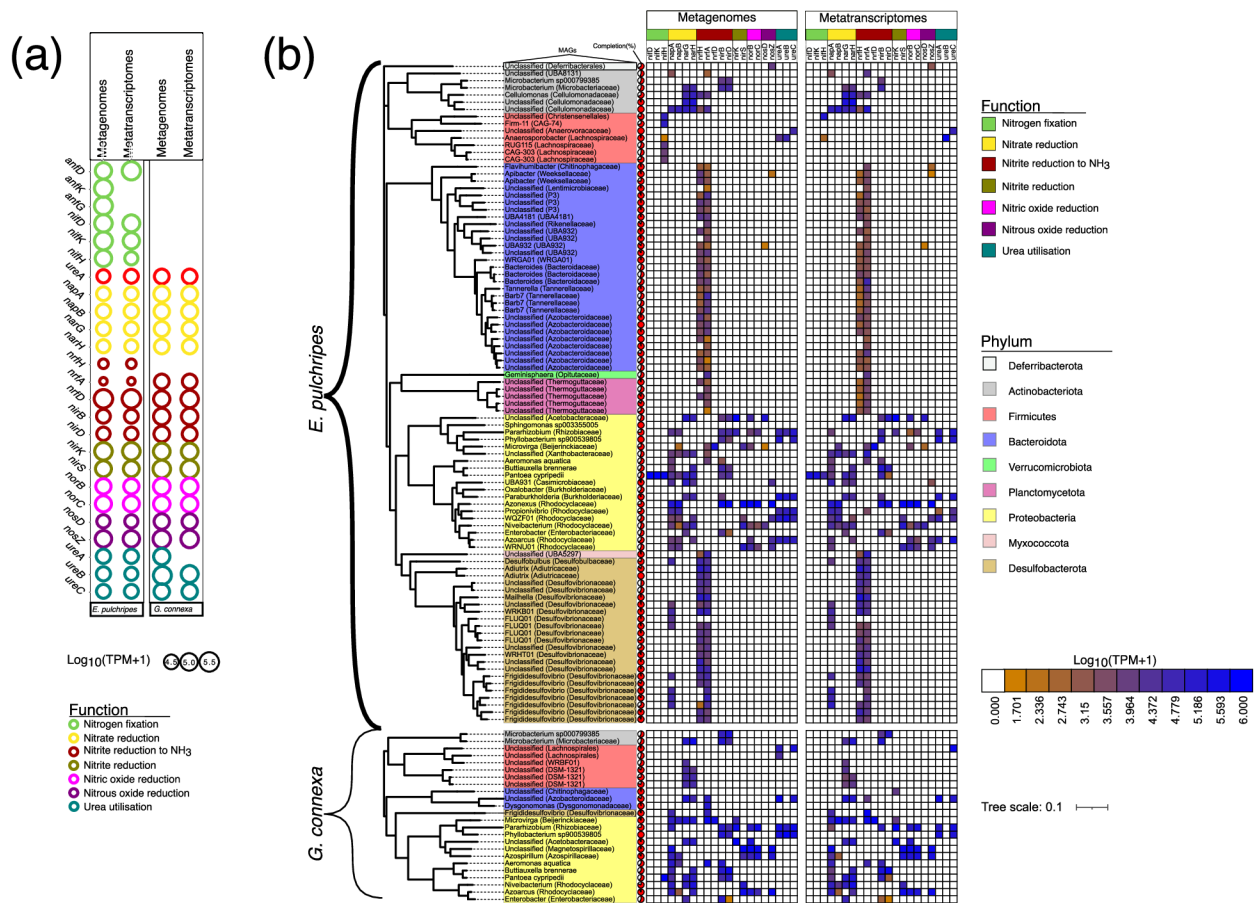
The prospect of sulfate as an alternative hydrogen sink to acetogenesis was assessed by searching the genes involved in dissimilatory sulfate reduction. We examined the occurrence and expression of key genes involved in sulfate reduction in metagenomes, metatranscriptomes, and MAGs. All the genes were present and expressed in metagenomes from both millipede species, except anaerobic sulfite reductase subunit B (*asrB*) and thiosulfate reductase/polysulfide reductase chain A (*phsA*) in *G. connexa* (Fig. 5a; Table S13). Genes encoding sulfate reductase (*dsrAB*) were present and expressed only in 21 out of the 28 *Desulfobacterota* MAGs from *E. pulchripes*; 19 of them possessed and also expressed *dsrD* (Fig. 5b; Table S14). In addition, we found that 10 MAGs possessed and expressed *aprA* and *sat*. However, both *dsr*,

*aprA*, and *sat* genes were absent from the three *Adiutricaceae* MAGs, which agrees with previous results [84]. Among the MAGs from *G. connexa*, only *Desulfovibrionaceae* possessed and expressed *dsrABD*. A thiosulfate reductase (*phsA*) gene involved in thiosulfate disproportionation was present and expressed in several MAGs from *E. pulchripes* (*Desulfovibrionaceae*, 3 MAGs; *Firmicutes*, 1 MAG; *Actinobacteriota*, 2 MAGs). The genes for anaerobic sulfite reduction (*asrABC*) were present and expressed in unclassified *Synergistota* (1 MAG) and *Planctomycetota* (1 MAG) from *E. pulchripes*. The *asr* genes were incomplete in the *Firmicutes* (2 MAGs) from *G. connexa*.

#### Nitrogen cycling by millipede hindgut bacteria

As described above, nitrogen fixation and cycling genes can help alleviate the nitrogen demands of detritivores and microbes living in litter. We investigated the presence and expression of key genes involved in nitrogen fixation and cycling in metagenomes, metatranscriptomes, and MAGs (Fig. 6a; Fig S4; Tables S15 and S16). The structural genes of Mo-nitrogenase (*nifDKH*) were present and expressed in the metagenome from *E. pulchripes*. Genes encoding the alternative, Fe–Fe nitrogenase (*anfDGGK*), were present, but only *anfD* was expressed. The second alternative, V-Fe nitrogenase (*vnfDKG* genes), was absent. Nitrogenase genes were absent from the metagenome of *G. connexa*, except *nifH*, which was detected in the assembly but was removed due to its short contigs. Genes for aerobic (*amoABC*) or anaerobic ammonium oxidation (*hzoAB*) were absent from the metagenomes. Still, we detected a nitrite oxidoreductase (*nxrAB*) in both millipede species, which may be involved in nitrite oxidation or nitrate reduction. We identified several other genes involved in various forms of nitrogen cycling in both species, including those for nitrate reduction (*napAB* and *narGH*), nitrite reduction to ammonia (*nrfADH* and *nirBD*), nitrite reduction (*nirKS*), nitric oxide reduction (*norBC*), nitrous oxide reduction (*nosZD*), and urea hydrolysis (*ureABC*). However, a portion of the *nrf* gene (*nrfH*) was absent in *G. connexa*. In addition, the urea-hydrolyzing genes (*ureABC*) were present and expressed in both species, except the *ureA* gene, which was not expressed in *G. connexa*.

The *nifDKH* genes in *E. pulchripes* were encoded and expressed by a MAG assigned to *Pantoea cypripedii* (*Proteobacteria*; Fig. 4b and Table S15). The same MAG was also present in *G. connexa*, but the *nifH* gene was not transcribed. In addition, six unclassified *Firmicutes* from *E. pulchripes* (4 *Lachnospiraceae*, 1 *Christensenellales*, and 1 *Oscillospirales* (CAG-74 family)) encoded only the *nifH* gene. The occurrence of active biological nitrogen in *E. pulchripes* but not in *G. connexa* was corroborated



**Fig. 6** Genes and transcripts involved in nitrogen fixation and nitrogen cycling pathways. **a** Relative abundance of genes and transcripts for nitrogen cycling in the in metagenomic (MG) and metatranscriptomic (MT) contigs from the hindguts of *E. Pulchripes* and *G. connexa*. Included are the genes involved in nitrogen fixation, nitrite oxidation, nitrate reduction, nitrite reduction to ammonia, nitrite reduction, nitric oxide reduction, nitrous oxide reduction, and urea utilisation. Full names of the gene families and their corresponding KEGG IDs are available in Table S13 **(b)** A heatmap showing the relative abundance of the genes and transcripts in each MAG with at least one of the genes. TPM+1. The tree was reconstructed using 39 concatenated bacterial single copy gene (ribosomal proteins) from our MAGs

using ARA, which showed ethylene accumulation only in *E. pulchripes* (with or without litter; Fig. S5).

Similarly to the case with sulfate, denitrification (nitrate and nitrite reduction) can serve as an alternative respiration pathway for bacteria in the absence of oxygen. For nitrate reduction, the membrane-bound nitrate reductases (*napAB*) were present and expressed in four MAGs assigned to *Rhodocyclaceae* (Proteobacteria) and one assigned to *Cellulomonadaceae* (Firmicutes) in *E. pulchripes* and two *Proteobacteria* MAGs from *G. connexa* (Fig. 4b). The soluble nitrate reductase genes (*narGH*) were encoded and expressed in four MAGs assigned to *Actinobacteriota* and six assigned *Proteobacteria* in *E. pulchripes*. In *G. connexa*, the genes were expressed in MAGs assigned to *Firmicutes* (3), *Bacteroidota* (1), and *Proteobacteria* (4). The second gene, *nxrAB*, identified as the nitrate reductase gene, was found in the same MAGs possessing *narGH*. The nitrite reductase genes (*nirKS*

and nitric oxide reductases (*norBC*) involved in nitrite and nitric oxide reduction through the Nir pathway were found to be encoded and expressed in *Proteobacteria* in both species of millipedes. Our findings showed that the nitrous oxidase accessory protein (*nosD*) was only present and expressed in the MAGs from *Firmicutes* and *Proteobacteria* in *E. pulchripes*. For the nitrous-oxide reductase (*nosZ*) gene, *Proteobacteria* possessed and expressed the gene in both species. Additionally, MAGs assigned to *Deferribacterota* (1) and *Bacteroidota* (2) also expressed the gene in *E. pulchripes*.

Nitrite reduction to ammonia through the Nrf pathway (*nrfAH*) was present and expressed in MAGs from *Desulfobacterota* (20), *Actinobacteriota* (2), *Verrucomicrobiota* (5), one *Myxococcota* (1), and *Bacteroidota* (23). The *nrfAD* genes were only present in one *Proteobacteria* MAG in both species. The *nirBD* genes were encoded and expressed in five *Proteobacteria* MAGs from *E.*



*pulchripes*, four *Proteobacteria*, and one *Actinobacteriota* MAGs from *G. connexa*.

Ureolytic bacteria are often present in environments where urea is constantly produced, such as the millipede gut [16]. Therefore, we also examined the MAGs of both species of millipedes and identified genes that may be involved in using urea (*ureABC*). These urease subunits were only found in *Proteobacteria* and *Firmicutes* in the *E. pulchripes* MAGs. Specifically, the complete set of ureases was encoded by five *Proteobacteria* MAGs in *E. pulchripes* and two *Proteobacteria* MAGs in *G. connexa*. However, none of the three gene subunits was expressed in either species of millipedes. The *ureAC* subunits were only expressed (Fig. 4b).

## Discussion

Although *E. pulchripes* and *G. connexa* were fed the same diet, they differed strongly in the composition and size of their microbiome. The differences in microbial load in the gut and feces of both species have also been reported for other millipede species using classical methods [85, 86]. Our findings using cultivation and ddPCR revealed variations in microbial abundance between the two species, while using amplicon sequencing and metagenomics, we showed differences in community composition. Despite similar sequencing efforts, we obtained roughly nine times more MAGs from *E. pulchripes* species than *G. connexa*. This variation in microbial concentration between even relatively closely related arthropod species has been documented before and could stem from different morphological or physicochemical gut conditions (pH and oxygen availability) [87].

The microbial community composition was in good agreement between all three profiling methods (amplicon sequencing, metagenomics, and metatranscriptomics), providing mutual support for the methods. Notably, however, *Firmicutes* and *Actinobacteriota* were underrepresented in our amplicon-based profiling. Comparing the two millipede species, they resembled their taxonomic composition on the phylum level, and many of the genera were shared. However, they differed remarkably in their relative abundances. While seldom tested directly, we assert that the redox state in the gut is one of the main forces shaping the community composition at higher taxonomic levels [also suggested in [87]. Accordingly, the phylum-level composition in *E. pulchripes*, whose gut redox potential is highly negative and hence reducing [29], was dominated by phyla representing many bacteria capable of anaerobic metabolism, such as *Bacteroidota*, *Firmicutes*, and *Verrucomicrobia*. In contrast, the much smaller *G. connexa*, with a typical positive and hence oxidative gut [16], comprised nearly 50% *Proteobacteria*. Similar taxonomic composition, dominated by

*Proteobacteria* with low proportions of *Bacteroidota*, is common to many arthropods [87], including other millipedes [88], terrestrial isopods [89, 90], beetles [91, 92], and in many bamboo-feeding Hemiptera, Orthoptera, Lepidoptera, and Coleoptera [11], where the redox conditions in the gut are expected to be positive. Conversely, termites and cockroaches, with typical anoxic guts and active fermentation, typically have lower proportions of *Proteobacteria* and are dominated by *Bacteroidota* and *Firmicutes*, similar to *E. pulchripes* [93–95]. However, three major phyla in termites, with significant importance to their metabolism, namely, *Spirochaetota*, *Fibrobacterota*, and *Elusimicrobiota*, were rare phyla in our datasets. *Spirochaetota* and *Fibrobacterota* are associated with wood-feeding termites and play an important role in cellulose degradation [96, 97]. At the genus level, *Bacteroides* (*Bacteroidota*) dominated in *E. pulchripes*. In contrast, *Dysgonomonas* (*Bacteroidota*) and *Citrobacter* (*Proteobacteria*) dominated in *G. connexa*. Similar to our results, the hindguts of cockroaches harbored mostly representatives of *Bacteroidaceae*, many of which remained unclassified at the genus level [95]. *Dysgonomonas* dominated in dung beetle larvae and pupa [92, 98].

Following taxonomy classification using GTDB-Tk, we identified potentially novel bacteria, primarily assigned to *Firmicutes* and *Bacteroidota*. Only 11% of the MAGs were classified to the species level, while the remaining 88% were not assigned to any known genera. Among the novel MAGs assigned to *Firmicutes*, only 12 were classified to the family level, and none were assigned to the genus level, except for one MAG assigned to *Holdemania* from the family *Erysipelotrichaceae*, which has only three published genomes in NCBI. We had only two MAGs assigned to the *Erysipelotrichaceae* family. Additionally, MAGs assigned to the families *Butyricoccaceae*, *Ruminococcaceae*, and *Acutalibacteraceae* have no published genomes in NCBI. In *Bacteroidota*, we classified 12 MAGs to the family level and 3 MAGs to the genus level. Among these were the genera *Rikenella* and *Tannerella*, with only three published genomes in NCBI. Furthermore, the family *Azobacteroidaceae* has no published genomes in NCBI. Our findings suggest that a substantial number of novel genera or higher taxonomic ranks were detected in these millipede species. The detection of methanogenic archaea in both millipedes was expected. While only *E. pulchripes* is considered CH<sub>4</sub>-emitting, molecular evidence for the presence of methanogenic DNA was also reported for *G. connexa* [30].

The most dominant fungal group, *Ascomycota*, is common in invertebrate gut microbiomes [85, 99], including other millipedes [17]. Since *Ascomycota* are dominant in leaf litter (especially in its early stages of decomposition



[100]), they are probably ingested with the leaves and are not residents of the gut microbiome. Rhabditid nematodes, another abundant eukaryote found in *E. pulchripes*, have been reported in the gut of various millipedes [101, 102]. However, their role in the gut microbiome remains unclear. Different protists were also among the dominant eukaryotic groups in both millipede species, though their abundance varied. Protists of the phylum *Ciliophora* (ciliates) are known to host and support symbiotic methanogens in termites thanks to their ability to generate hydrogen [103, 104]. It is therefore likely, though not yet shown, that this is also their role in millipedes. Lastly, *Eccrinales*, a protist order formally considered fungi and grouped together as *Trichomycetes*, are commonly found in many (but not all) millipedes and are considered a regular, non-pathogenic part of the gut microbiome [105]. In our dataset, they were only found as single or double very rare contigs (under 0.2% of the Eukaryotic abundance) and were absent in the metatranscriptome. Therefore, their role in millipedes remains elusive.

Plant material comprises structural components, including cellulose, hemicellulose, pectin, and lignin [106]. In many herbivores, the genomes of individual gut bacteria often encode hundreds of enzymes that help degrade complex plant polysaccharides [6]. Different substrates require different digestive enzymes, which occasionally work in concert [107]. Since polysaccharides typically cannot pass through the cell membrane, we focused only on the CAZymes possessing secretion signal peptides (SSP) that could be released into extracellular space and act on substrates outside the cells [80]. Most CAZymes with SSP were glycoside hydrolases (GHs), the primary enzyme families responsible for polysaccharide degradation [108]. Analysis of GHs with SSP in our MAGs showed that numerous predicted proteins were encoded and expressed. The majority of recent studies on millipedes [24] and other arthropods [14, 109, 110] also revealed significant levels of microbial GHs. However, these studies did not differentiate between GHs with and without signal peptides.

Both *E. pulchripes* and *G. connexa* possessed an abundance of secreted GHs that can degrade pectin (GH28, GH78, GH105, GH106, GH127, and GH88) and hemicellulose (GH2 and GH3), with pectin methylsterases (GH78, GH105, GH106, GH127, and GH88) being particularly prevalent. This could be because of the need to de-esterify homogalacturonan, a common component of plant cell walls [111]. The GH2 enzyme has multiple functions, such as  $\beta$ -galactosidases,  $\beta$ -glucuronidases,  $\beta$ -mannosidases, and *exo*- $\beta$ -glucosaminidases, which assist in the degradation of hemicellulose. Similarly, the

GH3 enzyme aids in plant and bacterial cell wall remodeling, cellulosic biomass degradation, energy metabolism, and pathogen defense [112]. Other abundant GHs include those that can degrade fungal cell walls, which comprise chitin, beta-glucans, and glycoproteins [28]. These were the most abundant GHs in *G. connexa* and the third most abundant in *E. pulchripes*. In addition to chitin, millipedes may also obtain macronutrients, including calcium (present as calcium oxalate), from feeding on fungi and thus support their diet [113]. Additionally, we found a high abundance of some GH families responsible for breaking down the carbohydrate backbone of bacterial peptidoglycans (GH73) and algal cell walls (GH29 and GH50). This suggests that leaf litter-colonizing microorganisms (primarily fungi and bacteria, and maybe also algae) are ingested with the food and serve as a carbon source. Both structural compounds and microorganisms are used as carbon and energy sources in some macroarthropods [114].

The secreted GHs were expressed by different taxa, reflecting the overall differences in community composition. In *E. pulchripes*, the dominant phylum was *Bacteroidota*, which had the highest abundance of GHs for complex carbon degradation at the genome-resolved levels. *Bacteroidota* is known for its ability to break down various complex polysaccharides [115] and is the most polysaccharolytic phylum in cockroaches [116]. On the other hand, *Proteobacteria* were the main source of GHs in *G. connexa*, similar to beetles (Coleoptera) [117]. Both *Bacteroidota* and *Proteobacteria* are agents of complex polysaccharide degradation in isopods [89]. There was also a high level of contributions from *Verrucomicrobiota*, *Firmicutes*, *Planctomycetota*, and *Proteobacteria* in *E. pulchripes* and *Bacteroidota* and *Firmicutes* in *G. connexa*. The success of *Bacteroidota* as a major polysaccharide degrader in *E. pulchripes* was linked to families of *Rikenellaceae*, *Bacteroidaceae*, *UBA4181*, *Azobacteroidaceae*, *Tannerellaceae*, and *UBA932*. Meanwhile, the families of *Sphingomonadaceae*, *Aeromonadaceae*, *Enterobacteriaceae*, and *Rhizobiaceae* significantly contributed to the hydrolytic activities of *Proteobacteria* in *G. connexa*. Similar families with such capabilities were present in omnivorous American cockroaches [116]. However, in contrast to past termite studies, we could not identify any contributions from the rare millipede phyla *Spirochaetota* and *Fibrobacterota* [14, 93].

In animals that rely on symbiotic digestion of (ligno) cellulose, the fermentation products of the bacterial symbionts fuel the carbon and energy metabolism of the host [25]. The hydrogen formed in the fermentations is either converted to methane or—in the case of termites—used for reductive acetogenesis [118]. In some large millipede

species, *Archispirostreptus gigas* and *Epibolus pulchripes*, acetate, and formate have been shown to accumulate in the gut, indicating bacterial fermentation activities in the digestive tracts [29]. In heterotrophic metabolism, the enzyme pyruvate:ferredoxin oxidoreductase (*porA*) oxidatively decarboxylates pyruvate to form acetyl-CoA and CO<sub>2</sub> [119]. Two molecules of acetyl-CoA are converted to two acetate molecules through phosphotransacetylase (*pta*) and acetate kinase (*ack*). In our study, we identified heterotrophic metabolism in some phyla from *E. pulchripes* that possessed and expressed the *porA*||*pta*||*ack* genes. The activity of *porA* has been identified as the sole site of energy conservation in the model acetogen, *Acetobacterium woodii* [118, 119]. However, it is worth noting that while acetate production is often used as a marker for acetogens, it is not definitive proof of acetogenesis, as other bacteria may also produce acetate. To be considered a true acetogen, a bacterium must be able to perform reductive acetogenesis, which involves using the Wood Ljungdahl Pathway (WLP) to convert two molecules of CO<sub>2</sub> produced by the oxidative decarboxylation of pyruvate into additional acetate [120]. Based on the seven key enzymes of reductive acetogenesis, we identified formate dehydrogenase H (*fdhF*), formate-tetrahydrofolate ligase/formyl tetrahydrofolate synthetase (*fhs*), methenyltetrahydrofolate cyclohydrolase (*fold*), methylenetetrahydrofolate reductase (*metF*), acetyl-CoA synthase (*acsABCDE*), phosphotransacetylase (*pta*), and acetate kinase (*ack*) [121] in our MAGs. In *E. pulchripes*, putative acetogens with near-complete pathways were found in MAGs assigned to *Firmicutes* (*Clostridaceae*) and *Desulfobacterota* (*Adiutricaceae*). The expression of *fdhF* and *acsABCDE* in these two phyla is a strong predictor for reductive acetogenesis. Two or three missing reductive acetogenic genes in *E. pulchripes* could be related to the incompleteness of our MAGs. The complete *acsABCDE* genes were lacking in MAGs from *G. connexa*, and other genes may not be sufficient to suggest that reductive acetogenesis is present in this species.

Some members of the family *Clostridaceae* (*Firmicutes*) are well-studied acetogens [121]. Microbiota studies have identified acetogenic bacteria belonging to *Firmicutes* [122, 123] from termites and particularly *Ruminococcaceae* in the rumen [120]. Similarly, *Adiutricaceae* in the phylum *Desulfobacterota* from termite guts have also been postulated to be putative acetogens [94, 124]. However, similar to our data, Arora et al. also reported a dominant *Adiutricaceae* MAG with an incomplete pathway [124]. In addition, although belonging to the phylum *Desulfobacterota* [108], which includes many sulfate reducers, none of the MAGs classified as *Adiutricaceae* encoded for the *dsrAB* genes required for dissimilatory sulfate reduction. Therefore, these organisms could also

be scavenging hydrogen. A similar observation was made in termites [84, 124].

In addition to genes involved in reductive acetogenesis found in the MAGs from *E. pulchripes*, the putative acetogens possessed and expressed one or more [FeFe] or [NiFe] hydrogenase subgroups. These FeFe hydrogenase subgroups [FeFe] in the Group A-C series are used for H<sub>2</sub>-evolution reaction/electron-bifurcation (*fefe-group-a1,3*), H<sub>2</sub>-uptake/electron-bifurcation (*fefe-group-a4*), H<sub>2</sub>-uptake (*fefe-group-b*), and H<sub>2</sub>-sensing (*fefe-group-c1-3*). The [NiFe] hydrogenase group found in the putative acetogens is for H<sub>2</sub>-uptake (*nife-group-1*) and evolution (*nife-group-4a-g*) [125].

Sulfate-reducing bacteria in the gut of millipedes have not been reported, but their consistent presence in the intestinal tract of many arthropods [126, 127] suggests that they may play a role either in the consumption of hydrogen produced by fermenting bacteria (which requires the presence of sulfate) or the production of hydrogen through fermentation [126, 128]. Although sulfate concentrations in millipede guts are most likely minuscule, similar to those in termites [129], the expression of the *sat*, *aprA*, and *dsrABD* genes and [NiFe] hydrogenases of Group 1 that are involved in H<sub>2</sub> uptake [126] indicate that the MAGs of *Desulfovibrionaceae* possess the ability to reduce sulfate. Another piece of evidence is the expression of the *acdA* and *acs* genes, which shows that *Desulfovibrionaceae* can use acetate to reduce sulfate since acetate is a competitive substrate for sulfate-reducing bacteria [130]. The expression of the *phsA* gene for thiosulfate disproportionation suggests that at least some of the sulfide produced in this process is reoxidized by the same organisms in the microoxic gut periphery [131], which likely provides the same microoxic conditions as in other arthropods [25].

Leaf litter has a notoriously high C:N ratio, and millipedes and their microbiome are likely permanently nitrogen-starved. Several arthropods living on an N-poor diet have been demonstrated to fix atmospheric nitrogen [132]. The presence and expression of Molybdenum-dependent nitrogenases (*nifDHK*) by *Pantoea cypripedii* (*Proteobacteria*) indicate that the gut microbiota of *E. pulchripes* contributes to dinitrogen reduction. Members of the genus *Pantoea* frequently form associations with various hosts, such as insects, plants, and humans, and are well known for their ability to fix nitrogen [133, 134]. The positive results from the ARA experiment demonstrate that biological nitrogen fixation is occurring in *E. pulchripes*.

As in termites, the gut microbiota of millipedes may also contribute to nitrogen metabolism by recycling uric acid or urea, which are waste products of the host [135, 136] or by reducing dietary nitrate [137]. We found that



the gut microbiota of both *E. pulchripes* and *G. connexa* expresses genes involved in urea oxidation, denitrification, and dissimilatory nitrate reduction to ammonia (DNRA). Denitrification is an important process in various soil fauna, including earthworms [138] and termites [137]. The most important contributors to these activities in *E. pulchripes* (Spirobolida) and *G. connexa* (Glomeridae) are *Actinobacteria*, *Firmicutes*, and *Proteobacteria*. Assuming that denitrification produces traces of N<sub>2</sub>O, the production of this greenhouse gas may not be restricted to the Glomeridae family, as previously thought [139]. DNRA activities have been documented by stable-isotope analyses in soil-feeding termites [137, 140] and several freshwater insects [140]. Key genes of DNRA are the nitrite reductases *nrfA* and *nirB* [141]. They were expressed by members of *Cellulomonadaceae* (*Actinobacteriota*) from *E. pulchripes* and members of *Proteobacteria* from *E. pulchripes* and *G. connexa*. Likewise, the presence of the *nrfAH* genes has been established in the gut microbiota of termites [124, 142] and aquatic insects [140].

Several MAGs of *Proteobacteria* from *E. pulchripes* and *G. connexa* also expressed ureases (*ureABC*), suggesting they contribute to ammonia production from urea. Urease activity is common in many bacteria from host-associated environments [143, 144]. *Proteobacteria* and *Actinobacteria* with urease activity have been isolated from millipede guts [86, 145].

## Conclusions

The data presented here is a comprehensive chart of the metabolic diversity in two millipede model species; one of Earth's most important groups of detritivores. We found substantial differences in both abundance and diversity of the gut microbial community between two millipede species that differ in their size, habitat, and gut redox conditions but share the same diet and lifestyle. Many functions encoded by the gut microbiota were present in the MAGs of both species, including the capacity to degrade complex carbohydrates. Lignin-modifying enzymes were very few, but a high expression of genes for chitin degradation indicates that fungal biomass may play an important role in the millipede diet, perhaps exceeding that of plant polymers. Fermentative lineages (*Clostridiales* and *Bacteroidales*) were particularly abundant in the large *E. pulchripes*, but clear evidence for reductive acetogenesis was lacking. Instead, we found strong evidence for hydrogenotrophy, nitrogen recycling, and diazotrophy. The results should serve as a roadmap for further studies to test these hypotheses regarding the trophic role of millipedes.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40168-023-01731-7>.

**Additional file 1: Fig. S1.** The relative abundance of eukaryotes in assembled metagenomic and metatranscriptomic reads from hindguts of *E. pulchripes* and *G. connexa*. (a) Abundance of fungi in (a) metagenome and (b) metatranscriptome. Abundance of algae and protists in (a) metagenome and (b) metatranscriptome. The paired-end reads of both library types were mapped to the genes/contigs to obtain the coverage and calculate the relative abundance in Transcript Per Million (TPM). The mean TPM was calculated from the three replicate samples and aggregated for each taxonomic level. **Fig. S2.** Relative abundance of glycoside hydrolases (GHs) with signal peptides in metagenome-assembled genomes (MAGs) and their corresponding transcripts. The GHs were grouped at the family level and according to their putative substrates (top of each chord) and the top 3 taxa (at the family level) contributing the GHs (bottom of each chord). Chord (a) displays the contribution of GHs from different families in metagenomes, while chord (b) shows its corresponding GH transcripts from the hindgut of *E. pulchripes*. Chord (c) shows the abundance of GHs at the family level in metagenomes, while chord (d) displays its corresponding GH transcripts from the hindgut of *G. connexa*. The pair-end reads of both library types were mapped to the genes to get the coverage and calculate the relative abundance in Transcript Per Million (TPM). The mean TPM was calculated from the three replicate samples and summed for each taxonomic level. **Fig. S3.** Relative abundance and taxonomic distribution of genes involved in acetogenesis, hydrogen sensing/evolution/bifurcation (hydrogenases) and sulfur cycling in the metagenomic (MG) and metatranscriptomic (MT) contigs from the hindguts of *E. pulchripes* and *G. connexa*. (a) Boxplots showing the relative abundance of the bacterial genes for a function within a phylum. (b) Taxonomic distribution of genes and transcripts at the phylum level. The pair-end reads from metagenomes and metatranscriptomes were mapped to all the genes to get their coverages and averaged to estimate their relative abundance in transcript per kilobase million (TPM). **Fig. S4.** Relative abundance and taxonomic distribution of genes involved in nitrogen fixation and recycling, and their corresponding transcripts. (a) Boxplots showing the relative abundance of the genes for a function within a phylum. (b) Taxonomic distribution of the genes and transcripts at the phylum level. The pair-end reads from metagenomes and metatranscriptomes were mapped to all the genes to get their coverages and averaged to estimate their relative abundance in transcript per kilobase million (TPM). **Fig. S5.** Functional assay for the activity of the N<sub>2</sub>-fixing nitrogenase enzyme in the reduction of acetylene to ethylene in *E. pulchripes*.

**Additional file 2: Table S1.** Mean values for colony counts and 16S rRNA copies. **Table S2.** Read counts and taxonomic classification of 16S rRNA amplicon sequence. **Table S3.** Sample information for metagenomic and metatranscriptomic sequencing. **Table S4.** Contig stats for de novo co-assembled reads from *E. pulchripes*. **Table S5.** Removal of potential eukaryotic and unknown contigs with Whokaryote. **Table S6.** Contig stat for de novo co-assembled reads from *E. pulchripes* and *G. connexa*. **Table S7.** Taxonomic classification of MAGs with GTDB-Tk. **Table S8.** Taxonomic classification of MAGs from *E. pulchripes* and *G. connexa*. **Table S9.** MAG coverage in quality-filtered metagenomic pair-end reads. **Table S10.** Distributions of all CAZymes with or without signal peptides in bacterial MAGs from *E. pulchripes* and *G. connexa*. **Table S11.** Grouping of glycoside hydrolases at the family level. **Table S12.** Annotated Carbohydrate-active enzyme (CAZymes) from MAGs using dbCAN2 meta web server. **Table S13.** Community gene annotation of the metagenome and metatranscriptome samples (prokaryotic co-assemblies). **Table 14.** Community relative abundance of genes involved in acetogenesis, hydrogenases and sulfur cycling for *E. pulchripes* and *G. connexa*. **Table S15.** Community gene annotation of the metagenome and metatranscriptome samples (prokaryotic co-assemblies). **Table S16.** Community relative abundance of genes involved in nitrogen cycling. **Table S17.** NCBI BioProject PRJNA948469.

### Acknowledgements

We appreciate the support of Shruti Gupta, Lucie Faktorová, and Eva Petrová with the collection of *G. connexa* samples, Lucie Faktorová and Shruti Gupta with the upkeep of the *E. pulchripes* colony, and Shruti Gupta and Eva Petrová for the guidance and assistance with DNA and RNA extractions and quantification. We would also like to thank, in particular, reviewer 2 for the thorough review of the manuscript and for the insightful comments.

### Authors' contributions

RA and VS conceptualised the experiments and approach. JN collected the millipede samples and performed the lab experiments. RA and AB designed the data analyses. JN and RA performed the bioinformatics analyses. JN and RA wrote the manuscript, with significant contributions from AB and VS. All authors read and accepted the final version of this manuscript.

### Authors' information

Not applicable.

### Funding

JN and RA were supported by a Junior Grant of the Czech Science Foundation (GA ČR grant no. 19-24309Y). In addition, JN was supported by IBERA and Erasmus grants (CZ CESKE01) for a scientific stay at AB's lab.

### Availability of data and materials

The shotgun metagenomes, metatranscriptomes, and short-read amplicon sequencing data were deposited under the NCBI BioProject PRJNA948469 (see Table S17). For reproducibility, reusability, and transparency, we have also made available the FASTA file for the co-assembled metagenomes from *Epibolus pulchripes* and *Glomeris connexa* and the anvi'o merged profile database (doi: <https://doi.org/10.6084/m9.figshare.22336732>). The FASTA file and anvi'o merged profile database for *G. connexa* can be found at doi: [10.6084/m9.figshare.22339522](https://doi.org/10.6084/m9.figshare.22339522) and for *E. pulchripes* at doi: <https://doi.org/10.6084/m9.figshare.22337500>. The scripts used in this study are available on GitHub (<https://github.com/julipeale2001/Millipede-gut-microbiome-metomics-analysis>).

### Declarations

#### Ethics approval and consent to participate

The collection of *G. connexa* required no special permission. The maintenance of the millipedes and the experiments performed require no special permissions or ethical considerations.

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare no competing interests.

#### Author details

<sup>1</sup>Institute of Soil Biology and Biogeochemistry, Biology Centre CAS, České Budějovice, Czechia. <sup>2</sup>Faculty of Science, University of South Bohemia, České Budějovice, Czechia. <sup>3</sup>RG Insect Gut Microbiology and Symbiosis, Max Planck Institute for Terrestrial Microbiology, Marburg, Germany.

Received: 16 April 2023 Accepted: 22 November 2023

Published online: 29 January 2024

### References

- Alagesan P. Millipedes: diversity, distribution and ecology. In: Chakravathy AK, Sridhara S, editors. *Arthropod diversity and conservation in the tropics and sub-tropics*. Singapore: Springer Singapore; 2016. p. 119–37. Available from: [http://link.springer.com/10.1007/978-981-10-1518-2\\_7](http://link.springer.com/10.1007/978-981-10-1518-2_7). Cited 2019 Aug 20.
- Golovatch SI, Kirme RD. Millipede (Diplopoda) distributions: a review. *Soil Org.* 2009;81:565–97.
- Watanabe H, Tokuda G. Cellulolytic systems in insects. *Annu Rev Entomol.* 2010;55:609–32.
- Wybouw N, Pauchet Y, Heckel DG, Van Leeuwen T. Horizontal gene transfer contributes to the evolution of arthropod herbivory. *Genome Biol Evol.* 2016;8:1785–801.
- Kögel-Knabner I. The macromolecular organic composition of plant and microbial residues as inputs to soil organic matter. *Soil Biol Biochem.* 2002;34:139–62.
- Flint HJ, Bayer EA, Rincon MT, Lamed R, White BA. Polysaccharide utilization by gut bacteria: potential for new insights from genomic analysis. *Nat Rev Microbiol.* 2008;6:121–31.
- Douglas AE. The microbial dimension in insect nutritional ecology. *Funct Ecol.* 2009;23:38–47.
- Nicholson JK, Holmes E, Kinross J, Burcelin R, Gibson G, Jia W, et al. Host-gut microbiota metabolic interactions. *Science.* 2012;336:1262–7.
- Bouchon D, Zimmer M, Dittmer J. The terrestrial isopod microbiome: an all-in-one toolbox for animal–microbe interactions of ecological relevance. *Front Microbiol.* 2016;7:1472. Available from: <https://www.frontiersin.org/articles/10.3389/fmicb.2016.01472>. Cited 2022 Dec 31.
- Graf J. Lessons from digestive-tract symbioses between bacteria and invertebrates. *Annu Rev Microbiol.* 2016;70:375–93.
- Huang K, Wang J, Huang J, Zhang S, Vogler AP, Liu Q, et al. Host phylogeny and diet shape gut microbial communities within bamboo-feeding insects. *Front Microbiol.* 2021;12:633075. Available from: <https://www.frontiersin.org/article/10.3389/fmicb.2021.633075>. Cited 2022 Jun 17.
- Geli-Cruz O, Cafaro MJ, Santos-Flores CJ, Ropelewski AJ, Van Dam AR. Taxonomic survey of *Anadenobolus monilicornis* gut microbiota via shotgun nanopore sequencing. *Genomics.* 2019. Available from: <http://biorxiv.org/lookup/doi/10.1101/560755>. Accessed 30 Nov 2023.
- Nardi JB, Bee CM, Taylor SJ. Compartmentalization of microbial communities that inhabit the hindguts of millipedes. *Arthropod Struct Dev.* 2016;45:462–74.
- Brune A, Dietrich C. The gut microbiota of termites: digesting the diversity in the light of ecology and evolution. *Annu Rev Microbiol.* 2015;69:145–66.
- Sun M, Chao H, Zheng X, Deng S, Ye M, Hu F. Ecological role of earthworm intestinal bacteria in terrestrial environments: a review. *Sci Total Environ.* 2020;740:140008.
- Byzov BA. Intestinal microbiota of millipedes. In: König H, Varma A, editors. *Intestinal microorganisms of termites and other invertebrates*. Berlin/Heidelberg: Springer-Verlag; 2006. p. 89–114. Available from: [http://link.springer.com/10.1007/3-540-28185-1\\_4](http://link.springer.com/10.1007/3-540-28185-1_4). Cited 2019 Aug 20.
- Sardar P, Šustr V, Chroňáková A, Lorenc F, Faktorová L. De novo metatranscriptomic exploration of gene function in the millipede holobiont. *Sci Rep.* 2022;12:16173.
- Heinze T, Petzold-Welcke K, van Dam JEG. Polysaccharides: molecular and supramolecular structures. Terminology. The European Polysaccharide network of Excellence (EPNOE): Research Initiatives and results. 2012. p. 23–64.
- Warren RA. Microbial hydrolysis of polysaccharides. *Annu Rev Microbiol.* 1996;50:183–212.
- Taylor EC. Role of aerobic microbial populations in cellulose digestion by desert millipedes. *Appl Environ Microbiol.* 1982;44:281–91.
- Szabo IM, Nasser E-GA, Striganova B, Rakhmo YR, Jager K, Heydrich M, et al. Interactions among millipedes (Diplopoda) and their intestinal bacteria. 1990. p. 8.
- Ramanathan B, Alagesan P. Isolation, characterization and role of gut bacteria of three different millipede species. 2012. p. 7.
- Beck L, Friebe B. Verwertung von Kohlenhydraten bei *Oniscus asellus* (Isopoda) und *Polydesmus angustus* (Diplopoda). *Pedobiologia.* 1981.
- Sardar P, Šustr V, Chroňáková A, Lorenc F. Metatranscriptomic holobiont analysis of carbohydrate-active enzymes in the millipede *Telodeinopus aoutii* (Diplopoda, Spirostreptida). *Front Ecol Evol.* 2022;10:931986. Available from: <https://www.frontiersin.org/articles/10.3389/fevo.2022.931986>. Cited 2022 Sep 27.
- Brune A. Symbiotic digestion of lignocellulose in termite guts. *Nat Rev Microbiol.* 2014;12:168–80.
- Maraun M, Scheu S. Changes in microbial biomass, respiration and nutrient status of beech (*Fagus sylvatica*) leaf litter processed by millipedes (*Glomeris marginata*). *Oecologia.* 1996;107:131–40.

27. Bignell D. Relative assimilations of <sup>14</sup>C-labelled microbial tissues and <sup>14</sup>C-plant fibre ingested with leaf litter by the millipede *Glomeris marginata* under experimental conditions. *Soil Biol Biochem.* 1989;21:819–27.
28. Garcia-Rubio R, de Oliveira HC, Rivera J, Trevijano-Contador N. The fungal cell wall: *Candida*, *Cryptococcus*, and *Aspergillus* species. *Front Microbiol.* 2020;10:2993. Available from: <https://www.frontiersin.org/articles/10.3389/fmicb.2019.02993>. Cited 2023 Aug 18.
29. Horváthová T, Šustr V, Chroňáková A, Semanová S, Lang K, Dietrich C, et al. Methanogenesis in the digestive tracts of the tropical millipedes *Archispirostreptus gigas* (Diplopoda, Spirostreptidae) and *Epibolus pulchripes* (Diplopoda, Pachybolidae). *Appl Environ Microbiol.* 2021;87:e00614–e621.
30. Šustr V, Chroňáková A, Semanová S, Tajovský K, Šimek M. Methane production and methanogenic archaea in the digestive tracts of millipedes (Diplopoda). *PLoS One.* 2014;9:e102659.
31. Ceja-Navarro JA, Nguyen NH, Karaoz U, Gross SR, Herman DJ, Andersen GL, et al. Compartmentalized microbial composition, oxygen gradients and nitrogen fixation in the gut of *Odontotaenium disjunctum*. *ISME J.* 2014;8(1):6–18.
32. David J-F. The role of litter-feeding macroarthropods in decomposition processes: a reappraisal of common views. *Soil Biol Biochem.* 2014;76:109–18.
33. Enghoff H. East African giant millipedes of the tribe Pachybolini (Diplopoda, Spirobolida, Pachybolidae). *Zootaxa.* 2011;2753:1–41.
34. Hoess R, Scholl A. Allozyme and literature study of *Glomeris guttata* Risso, 1826, and *G. connexa* Koch, 1847, a case of taxonomic confusion (Diplopoda: Glomeridae). *Zoologischer Anzeiger.* 2001;240:15–33.
35. Gerstaecker A. Die Gliedertier - Fauna des Sansibar-Gebietes. [The arthropod fauna of the Zanzibar region]. Hansebooks; 2016.
36. Kocourek P, Tajovský K, Dolejš P. Mnohonozky České republiky—Příručka pro určování našich druhů [Millipedes of the Czech Republic—Guide for identification of our species].—Základní organizace Českého svazu ochránců přírody. Vlašim: English abstract; 2017. p. 256.
37. Unkovich M, Cadisch G, Australian Centre for International Agricultural Research. Measuring plant-associated nitrogen fixation in agricultural systems. Canberra: ACIAR; 2008.
38. Angel R, Claus P, Conrad R. Methanogenic archaea are globally ubiquitous in aerated soils and become active under wet anoxic conditions. *ISME J.* 2012;6:847–62.
39. Angel R, Petrova E, Lara-Rodriguez A. Total nucleic acids extraction from soil V.6. protocols.io. 2021;6. [cited 2022 May 3]. Available from: <https://www.protocols.io/view/total-nucleic-acids-extraction-from-soil-bi46k-gze>.
40. Naqib A, Poggi S, Green SJ. Deconstructing the polymerase chain reaction II: an improved workflow and effects on artifact formation and primer degeneracy. *PeerJ.* 2019;7:e7121.
41. Walters W, Hyde ER, Berg-Lyons D, Ackermann G, Humphrey G, Parada A, et al. Improved bacterial 16S rRNA gene (V4 and V4–5) and fungal internal transcribed spacer marker gene primers for microbial community surveys. *mSystems.* 2016;1:e00009-15.
42. Angel R, Petrova E, Lara A. qPCR: Bacterial SSU rRNA 338F-516P-805R v4. 2020. Available from: <https://www.protocols.io/view/qpcr-bacterial-ssu-rna-338f-516p-805r-bqx5mrxq6>.
43. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 2014;30:2114–20.
44. Eren AM, Kiehl E, Shaiber A, Veseli I, Miller SE, Schechter MS, et al. Community-led, integrated, reproducible multi-omics with anvi'o. *Nat Microbiol.* 2021;6:3–6.
45. Li D, Liu C-M, Luo R, Sadakane K, Lam T-W. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics.* 2015;31:1674–6.
46. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods.* 2012;9:357–9.
47. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. *Bioinformatics.* 2009;25:2078–9.
48. Hyatt D, Chen G-L, LoCascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics.* 2010;11:119.
49. Finn RD, Clements J, Eddy SR. HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res.* 2011;39:W29-37.
50. Kim D, Song L, Breitwieser FP, Salzberg SL. Centrifuge: rapid and sensitive classification of metagenomic sequences. *Genome Res.* 2016;26:1721–9.
51. Galperin MY, Wolf YI, Makarova KS, Vera Alvarez R, Landsman D, Koonin EV. COG database update: focus on microbial diversity, model organisms, and widespread pathogens. *Nucleic Acids Res.* 2020;49:D274–81.
52. Kanehisa M, Goto S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 2000;28:27–30.
53. Kang DD, Li F, Kirton E, Thomas A, Egan R, An H, et al. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ.* 2019;7:e7359.
54. Alneberg J, Bjarnason BS, de Bruijn I, Schirmer M, Quick J, Ijaz UZ, et al. Binning metagenomic contigs by coverage and composition. *Nat Methods.* 2014;11:1144–6.
55. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* 2015;25:1043–55.
56. RCore T. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2016.
57. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* 2011;17:10–2.
58. Callahan BJ, Sankaran K, Fukuyama JA, McMurdie PJ, Holmes SP. Bioconductor workflow for microbiome data analysis: from raw reads to community analyses. *F1000Res.* 2016;5:1492.
59. Parks DH, Chuvochina M, Waite DW, Rinke C, Skarshewski A, Chaumeil P-A, et al. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat Biotechnol.* 2018;36:996–1004.
60. McMurdie PJ, Holmes S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One.* 2013;8: e61217.
61. Davis NM, Proctor DM, Holmes SP, Relman DA, Callahan BJ. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome.* 2018;6:226.
62. Lozupone C, Knight R. UniFrac: a new phylogenetic method for comparing microbial communities. *Appl Environ Microbiol.* 2005;71:8228–35.
63. Kopylova E, Noé L, Touzet H. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics.* 2012;28:3211–7.
64. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. *Nat Biotechnol.* 2011;29:644–52.
65. Pronk LJU, Medema MH. Whokaryote: distinguishing eukaryotic and prokaryotic contigs in metagenomes based on gene structure. *Microb Genom.* 2022;8:mgen000823.
66. von Meijenfeldt FAB, Arkhipova K, Cambuy DD, Coutinho FH, Dutilh BE. Robust taxonomic classification of uncharted microbial sequences and bins with CAT and BAT. *Genome Biol.* 2019;20:217.
67. Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. GTDB-Tk: a toolkit to classify genomes with the genome taxonomy database. *Bioinformatics.* 2020;36:1925–7.
68. Luo C, Rodriguez-R LM, Konstantinidis KT. MyTaxa: an advanced taxonomic classifier for genomic and metagenomic sequences. *Nucleic Acids Res.* 2014;42:e73.
69. Glendinning L, Stewart RD, Pallen MJ, Watson KA, Watson M. Assembly of hundreds of novel bacterial genomes from the chicken caecum. *Genome Biol.* 2020;21:34.
70. Bengtsson-Palme J, Hartmann M, Eriksson KM, Pal C, Thorell K, Larsson DGJ, et al. METAXA2: improved identification and taxonomic classification of small and large subunit rRNA in metagenomic data. *Mol Ecol Resour.* 2015;15:1403–14.
71. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol.* 1990;215:403–10.
72. Zhou Z, Tran PQ, Breister AM, Liu Y, Kieft K, Cowley ES, et al. METABOLIC: high-throughput profiling of microbial genomes for functional traits, metabolism, biogeochemistry, and community-scale functional networks. *Microbiome.* 2022;10:33.
73. Drula E, Garron M-L, Dogan S, Lombard V, Henrissat B, Terrapon N. The carbohydrate-active enzyme database: functions and literature. *Nucleic Acids Res.* 2021;50:D571–7.



74. Sayers EW, Bolton EE, Brister JR, Canese K, Chan J, Comeau DC, et al. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* 2021;50:D20–6.
75. Wickham H. *ggplot2: elegant graphics for data analysis.* Springer-Verlag New York; 2016. Available from: <https://ggplot2.tidyverse.org>.
76. Gu Z, Gu L, Eils R, Schlesner M, Brors B. “Circlize” implements and enhances circular visualization in R. 2014.
77. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* 2021;49:W293–6.
78. Li H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics.* 2018;34:3094–100.
79. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv; 2013. [cited 2023 Nov 30]. Available from: <http://arxiv.org/abs/1303.3997>.
80. Paetzel M. Bacterial signal peptidases. In: Kuhn A, editor. *Bacterial cell walls and membranes.* Cham: Springer International Publishing; 2019. p. 187–219. [https://doi.org/10.1007/978-3-030-18768-2\\_7](https://doi.org/10.1007/978-3-030-18768-2_7). Cited 2022 Jun 16.
81. Rabouille C. Pathways of unconventional protein secretion. *Trends Cell Biol.* 2017;27:230–40.
82. Lin H, Castro NM, Bennett GN, San K-Y. Acetyl-CoA synthetase overexpression in *Escherichia coli* demonstrates more efficient acetate assimilation and lower acetate accumulation: a potential tool in metabolic engineering. *Appl Microbiol Biotechnol.* 2006;71:870–4.
83. Wolfe AJ. The acetate switch. *Microbiol Mol Biol Rev.* 2005;69:12–50.
84. Ikeda-Ohtsubo W, Strassert JFH, Köhler T, Mikaelyan A, Gregor I, McHardy AC, et al. ‘*Candidatus* Adiatrix intracellularis’, an endosymbiont of termite gut flagellates, is the first representative of a deep-branching clade of Deltaproteobacteria and a putative homoacetogen. *Environ Microbiol.* 2016;18:2548–64.
85. Dhivya A, Alagesan P. Isolation and identification of microbial load in the gut and faeces of millipede *Spinotarsus colosseus*. *World J Zool.* 2018;13:04–9.
86. Ineson P, Anderson JM. Aerobically isolated bacteria associated with the gut and faeces of the litter feeding macroarthropods *Oniscus asellus* and *Glomeris marginata*. *Soil Biol Biochem.* 1985;17:843–9.
87. Engel P, Moran NA. The gut microbiota of insects – diversity in structure and function. *FEMS Microbiol Rev.* 2013;37:699–735.
88. Knapp BA, Seeber J, Rief A, Meyer E, Insam H. Bacterial community composition of the gut microbiota of *Cylindroiulus fulviceps* (diplopoda) as revealed by molecular fingerprinting and cloning. *Folia Microbiol.* 2010;55:489–96.
89. Bredon M, Herran B, Bertaux J, Grève P, Moumen B, Bouchon D. Isopod holobionts as promising models for lignocellulose degradation. *Bio-technol Biofuels.* 2020;13:49.
90. Delhoumi M, Catania V, Zaabar W, Tolone M, Quatrini P, Achouri MS. The gut microbiota structure of the terrestrial isopod *Porcellionides pruinosus* (Isopoda: Oniscidea). *Eur Zool J.* 2020;87:357–68.
91. Xu L, Sun L, Zhang S, Wang S, Lu M. High-resolution profiling of gut bacterial communities in an invasive beetle using PacBio SMRT sequencing system. *Insects.* 2019;10:248.
92. Suárez-Moo P, Cruz-Rosales M, Ibarra-Laclette E, Desgarennes D, Huerta C, Lamelas A. Diversity and composition of the gut microbiota in the developmental stages of the dung beetle *Copris incertus* Say (Coleoptera, Scarabaeidae). *Front Microbiol.* 2020;11:1698. Available from: <https://www.frontiersin.org/article/10.3389/fmicb.2020.01698>. Cited 2022 Jun 17.
93. Berlanga M, Llorens C, Comas J, Guerrero R. Gut bacterial community of the xylophagous cockroaches *Cryptocercus punctulatus* and *Parasphera boleiriana*. *PLoS One.* 2016;11:e0152400.
94. Dietrich C, Köhler T, Brune A. The Cockroach origin of the termite gut microbiota: patterns in bacterial community structure reflect major evolutionary events. *Appl Environ Microbiol.* 2014;80:2261–9.
95. Lampert N, Mikaelyan A, Brune A. Diet is not the primary driver of bacterial community structure in the gut of litter-feeding cockroaches. *BMC Microbiol.* 2019;19:238.
96. He S, Ivanova N, Kirton E, Allgaier M, Bergin C, Scheffrahn RH, et al. Comparative metagenomic and metatranscriptomic analysis of hindgut paunch microbiota in wood- and dung-feeding higher termites. *PLoS One.* 2013;8:e61126. Korb J, editor.
97. Tokuda G, Mikaelyan A, Fukui C, Matsuura Y, Watanabe H, Fujishima M, et al. Fiber-associated spirochetes are major agents of hemicellulose degradation in the hindgut of wood-feeding higher termites. *Proc Natl Acad Sci U S A.* 2018;115:E11996–2004.
98. Egert M, Wagner B, Lemke T, Brune A, Friedrich MW. Microbial community structure in midgut and hindgut of the humus-feeding larva of *Pachnoda ephippiata* (Coleoptera: Scarabaeidae). *Appl Environ Microbiol.* 2003;69:6659–68.
99. Mohammed WS, Ziganshina EE, Shagimardanova EI, Gogoleva NE, Ziganshin AM. Comparison of intestinal bacterial and fungal communities across various xylophagous beetle larvae (Coleoptera: Cerambycidae). *Sci Rep.* 2018;8:10073.
100. Voříšková J, Baldrian P. Fungal community on decomposing leaf litter undergoes rapid successional changes. *ISME J.* 2013;7:477–86.
101. Carta LK, Thomas WK, Meyer-Rochow VB. Two nematodes (Nematoda: Diplogastridae, Rhabditidae) from the invasive millipede *Chamberlinius huaiensis* Wang, 1956 (Diplopoda, Paradoxosomatidae) on Hachijojima Island in Japan. *J Nematol.* 2018;50:479–86.
102. Nagae S, Sato K, Tanabe T, Hasegawa K. Symbiosis of the millipede parasitic nematodes Rhigonematoidea and Thelastomatoidea with evolutionary different origins. *BMC Ecol Evol.* 2021;21:120.
103. Purdy KJ. The distribution and diversity of euryarchaeota in termite guts. *Advances in applied microbiology.* Academic Press; 2007. p. 63–80. Available from: <https://www.sciencedirect.com/science/article/pii/S0065216407620036>. Cited 2023 Nov 9.
104. Hongoh Y, Ohkuma M. Termite gut flagellates and their methanogenic and eubacterial symbionts. 2010. p. 55–79.
105. Lichtwardt RW. Trichomycetes and the Arthropod Gut. In: Brakhage AA, Zipfel PF, editors. *Human and animal relationships.* Berlin: Springer; 2008. p. 3–19.
106. Gibson LJ. The hierarchical structure and mechanics of plant materials. *J R Soc Interface.* 2012;9:2749–66.
107. Wei H, Xu Q, Taylor LE, Baker JO, Tucker MP, Ding S-Y. Natural paradigms of plant cell wall degradation. *Curr Opin Biotechnol.* 2009;20:330–8.
108. Zhang H, Yohe T, Huang L, Entwistle S, Wu P, Yang Z, et al. dbCAN2: a meta server for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res.* 2018;46:W95–101.
109. Carvalho DB, Paixão DA, Persinoti GF, Cota J, Rabelo SC, Grandis A, et al. Degradation of sugarcane bagasse by cockroach consortium bacteria. *Bioenergy Res.* 2022;15:1144–56.
110. Ni J, Tokuda G. Lignocellulose-degrading enzymes from termites and their symbiotic microbiota. *Biotechnol Adv.* 2013;31:838–50.
111. Wormit A, Usadel B. The multifaceted role of pectin methylesterase inhibitors (PMEIs). *Int J Mol Sci.* 2018;19:2878.
112. The CAZyPedia Consortium. Ten years of CAZyPedia: a living encyclopedia of carbohydrate-active enzymes. *Glycobiology.* 2018;28:3–8.
113. Cromack K, Sollins P, Todd RL, Crossley DA, Fender WM, Fogel R, et al. Soil microorganism—arthropod interactions: fungi as major calcium and sodium Sources. In: Mattson WJ, editor, et al., *The role of arthropods in forest ecosystems.* Berlin, Heidelberg: Springer; 1977. p. 78–84.
114. Frouz J, Kristůfek V, Li X, Santrůčková H, Sustr V, Brune A. Changes in amount of bacteria during gut passage of leaf litter and during coprophagy in three species of *Bibionidae* (Diptera) larvae. *Folia Microbiol (Praha).* 2003;48:535–42.
115. McKee LS, La Rosa SL, Westereng B, Eijsink VG, Pope PB, Larsbrink J. Polysaccharide degradation by the Bacteroidetes: mechanisms and nomenclature. *Environ Microbiol Rep.* 2021;13:559–81.
116. Vera-Ponce de León A, Jahnes BC, Duan J, Camuy-Vélez LA, Sabree ZL. Cultivable, host-specific Bacteroidetes symbionts exhibit diverse polysaccharolytic strategies. *Appl Environ Microbiol.* 2020;86:e00091–20.
117. Bozorov TA, Rasulov BA, Zhang D. Characterization of the gut microbiota of invasive *Agrilus mali* Matsumura (Coleoptera: Buprestidae) using high-throughput sequencing: uncovering plant cell-wall degrading bacteria. *Sci Rep.* 2019;9:4923.
118. Pester M, Brune A. Hydrogen is the central free intermediate during lignocellulose degradation by termite gut symbionts. *ISME J.* 2007;1:551–65.
119. Ragsdale SW. Pyruvate ferredoxin oxidoreductase and its radical intermediate. *Chem Rev.* 2003;103:2333–46.
120. Gagen EJ, Padmanabha J, Denman SE, McSweeney CS. Hydrogenotrophic culture enrichment reveals rumen Lachnospiraceae and

- Ruminococcaceae acetogens and hydrogen-responsive Bacteroidetes from pasture-fed cattle. *FEMS Microbiol Lett.* 2015;362:fnv104.
121. Schuchmann K, Müller V. Autotrophy at the thermodynamic limit of life: a model for energy conservation in acetogenic bacteria. *Nat Rev Microbiol.* 2014;12:809–21.
  122. Kane MD, Brauman A, Breznak JA. *Clostridium mayombi* sp. nov., an H<sub>2</sub>/CO<sub>2</sub> acetogenic bacterium from the gut of the African soil-feeding termite *Cubitermes speciosus*. *Arch Microbiol.* 1991;156:99–104.
  123. Kane MD, Breznak JA. *Acetonema longum* gen nov. sp. nov., an H<sub>2</sub>/CO<sub>2</sub> acetogenic bacterium from the termite *Pterotermes occidentis*. *Arch Microbiol.* 1991;156:91–8.
  124. Arora J, Kinjo Y, Šobotník J, Buček A, Clitheroe C, Stiblik P, et al. The functional evolution of termite gut microbiota. *Microbiome.* 2022;10:78.
  125. Søndergaard D, Pedersen CNS, Greening C. HydDB: a web tool for hydrogenase classification and analysis. *Sci Rep.* 2016;6:34212.
  126. Martins M, Pereira IAC. Sulfate-reducing bacteria as new microorganisms for biological hydrogen production. *Int J Hydrogen Energy.* 2013;38:12294–301.
  127. Dröge S, Limper U, Emtiazi F, Schönig I, Pavlus N, Drzyzga O, et al. In vitro and in vivo sulfate reduction in the gut contents of the termite *Mastotermes darwiniensis* and the rose-chafer *Pachnoda marginata*. *J Gen Appl Microbiol.* 2005;51:57–64.
  128. Greening C, Geier R, Wang C, Woods LC, Morales SE, McDonald MJ, et al. Diverse hydrogen production and consumption pathways influence methane production in ruminants. *ISME J.* 2019;13:2617–32.
  129. Brune A, Ohkuma M. Role of the termite gut microbiota in symbiotic digestion. In: Bignell DE, Roisin Y, Lo N, editors. *Biology of termites: a modern synthesis*. Dordrecht: Springer Netherlands; 2011. p. 439–75. [https://doi.org/10.1007/978-90-481-3977-4\\_16](https://doi.org/10.1007/978-90-481-3977-4_16). Cited 2022 Nov 18.
  130. Goevert D, Conrad R. Carbon isotope fractionation by sulfate-reducing bacteria using different pathways for the oxidation of acetate. *Environ Sci Technol.* 2008;42:7813–7.
  131. Kuhnigk T, Branke J, Krekeler D, Cypionka H, König H. A feasible role of sulfate-reducing bacteria in the termite gut. *Syst Appl Microbiol.* 1996;19:139–49.
  132. Bar-Shmuel N, Behar A, Segoli M. What do we know about biological nitrogen fixation in insects? Evidence and implications for the insect and the ecosystem. *Insect Sci.* 2020;27:392–403.
  133. Loiret FG, Ortega E, Kleiner D, Ortega-Rodés P, Rodés R, Dong Z. A putative new endophytic nitrogen-fixing bacterium *Pantoea* sp. from sugarcane. *J Appl Microbiol.* 2004;97:504–11.
  134. Walterson AM, Stavriniades J. *Pantoea*: insights into a highly versatile and diverse genus within the Enterobacteriaceae. *FEMS Microbiol Rev.* 2015;39:968–84.
  135. Hongoh Y. Toward the functional analysis of uncultivable, symbiotic microorganisms in the termite gut. *Cell Mol Life Sci.* 2011;68:1311–25.
  136. Breznak JA, Switzer JM. Acetate synthesis from H<sub>2</sub> plus CO<sub>2</sub> by termite gut microbes. *Appl Environ Microbiol.* 1986;52:623–30.
  137. Ngugi DK, Ji R, Brune A. Nitrogen mineralization, denitrification, and nitrate ammonification by soil-feeding termites: a <sup>15</sup>N-based approach. *Biogeochemistry.* 2011;103:355–69.
  138. Horn MA, Mertel R, Gehre M, Kästner M, Drake HL. In vivo emission of dinitrogen by earthworms via denitrifying bacteria in the gut. *Appl Environ Microbiol.* 2006;72:1013–8.
  139. Šustr V, Šimek M, Faktorová L, Macková J, Tajovský K. Release of greenhouse gases from millipedes as related to food, body size, and other factors. *Soil Biol Biochem.* 2020;144:107765.
  140. Ayayee P, Bhattacharyya S, Arnold T, Werne J, Leff L. Experimental investigation of potential biological nitrogen provisioning by freshwater insect gut microbiomes using <sup>15</sup>N isotope analysis. *Preprints*; 2019. Available from: <https://www.preprints.org/manuscript/201908.0034/v1>. Cited 2023 Mar 22.
  141. Wang H, Gunsalus RP. The *nrfA* and *nirB* nitrite reductase operons in *Escherichia coli* are expressed differently in response to nitrate than to nitrite. *J Bacteriol.* 2000;182:5813–22.
  142. Kuwahara H, Yuki M, Izawa K, Ohkuma M, Hongoh Y. Genome of 'Ca. *Desulfovibrio trichonymphae*', an H<sub>2</sub>-oxidizing bacterium in a tripartite symbiotic system within a protist cell in the termite gut. *ISME J.* 2017;11:766–76.
  143. López-Sánchez MJ, Neef A, Peretó J, Patiño-Navarrete R, Pignatelli M, Latorre A, et al. Evolutionary convergence and nitrogen metabolism in *Blattabacterium* strain Bge, primary endosymbiont of the cockroach *Blattella germanica*. *PLoS Genet.* 2009;5:e1000721.
  144. Sabree ZL, Kambhampati S, Moran NA. Nitrogen recycling and nutritional provisioning by *Blattabacterium*, the cockroach endosymbiont. *Proc Natl Acad Sci.* 2009;106:19521–6.
  145. Ma Z, Liu H, Liu C, He H, Zhao J, Wang X, et al. *Streptosporangium sonchi* sp. nov. and *Streptosporangium kronopolitis* sp. nov., two novel actinobacteria isolated from a root of common sowthistle (*Sonchus oleraceus* L.) and a millipede (*Kronopolites svenhedind* Verhoeff). *Antonie Van Leeuwenhoek.* 2015;107:1491–9.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.





# **Functional similarity, despite taxonomical divergence in the millipede gut microbiota, points to a common trophic strategy**

**Julius E. Nweze<sup>1,2</sup>, Vladimír Šustr<sup>1</sup>, Andreas Brune<sup>2</sup>, Roey Angel<sup>1\*</sup>**

<sup>1</sup>Institute of Soil Biology and Biogeochemistry, Biology Centre CAS, České Budějovice, Czechia

<sup>2</sup>Department of Ecosystem Biology, Faculty of Science, University of South Bohemia in České Budějovice, Czechia

<sup>3</sup>RG Insect Gut Microbiology and Symbiosis, Max-Planck Institute for Terrestrial Microbiology, Marburg, Germany

\* Correspondence: roey.angel@bc.cas.cz

## **Supplementary Material**

### **Supplementary Tables**

See file: Nweze\_et\_al\_Microbiome.xlsx

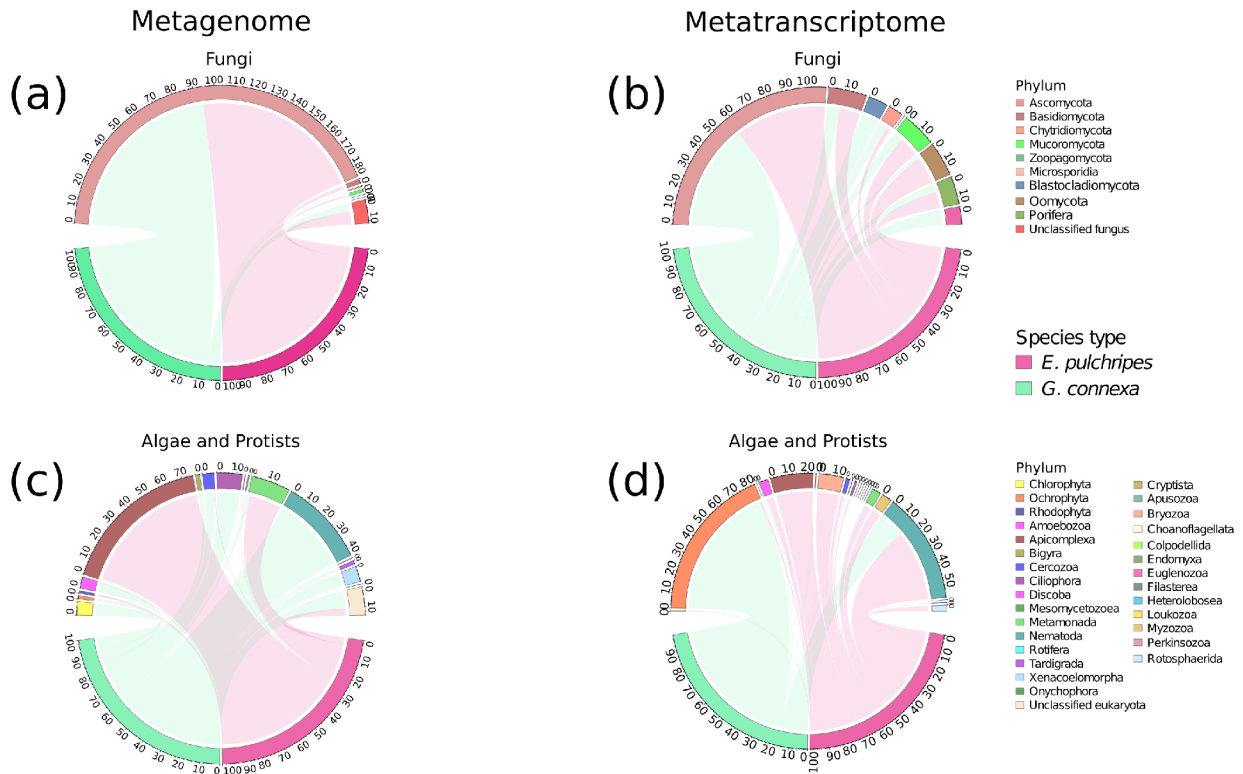


Fig. S1. The relative abundance of eukaryotes in assembled metagenomic and metatranscriptomic reads from hindguts of *E. pulchripes* and *G. connexa*. (a) Abundance of fungi in (a) metagenome and (b) metatranscriptome. Abundance of algae and protists in (a) metagenome and (b) metatranscriptome. The paired-end reads of both library types were mapped to the genes/contigs to obtain the coverage and calculate the relative abundance in Transcript Per Million (TPM). The mean TPM was calculated from the three replicate samples and aggregated for each taxonomic level.

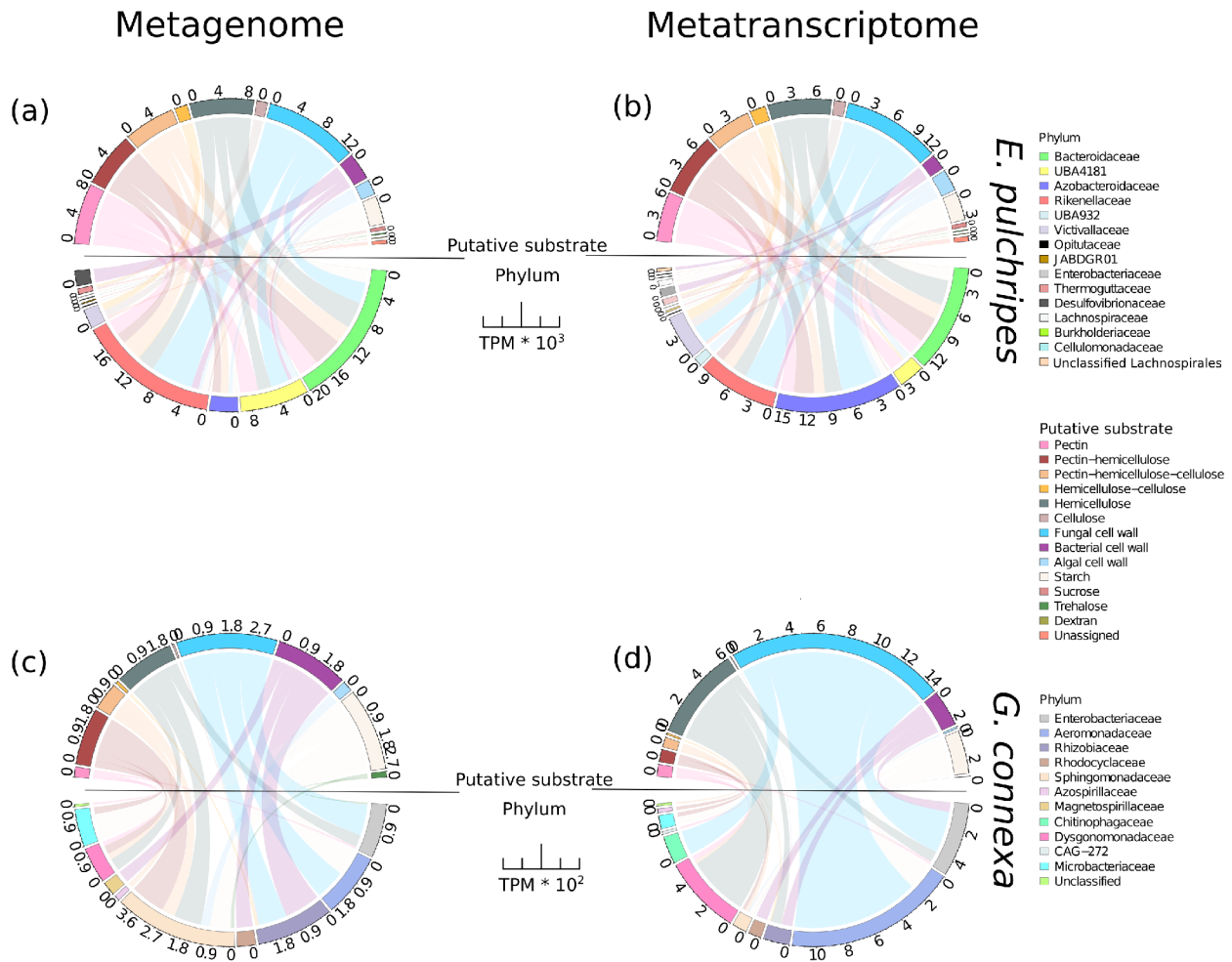


Fig. S2. Relative abundance of glycoside hydrolases (GHs) with signal peptides in metagenome-assembled genomes (MAGs) and their corresponding transcripts. The GHs were grouped at the family level and according to their putative substrates (top of each chord) and the top 3 taxa (at the family level) contributing the GHs (bottom of each chord). Chord (a) displays the contribution of GHs from different families in metagenomes, while chord (b) shows its corresponding GH transcripts from the hindgut of *E. pulchripes*. Chord (c) shows the abundance of GHs at the family level in metagenomes, while chord (d) displays its corresponding GH transcripts from the hindgut of *G. connexa*. The pair-end reads of both library types were mapped to the genes to get the coverage and calculate the relative abundance in Transcript Per Million (TPM). The mean TPM was calculated from the three replicate samples and summed for each taxonomic level.

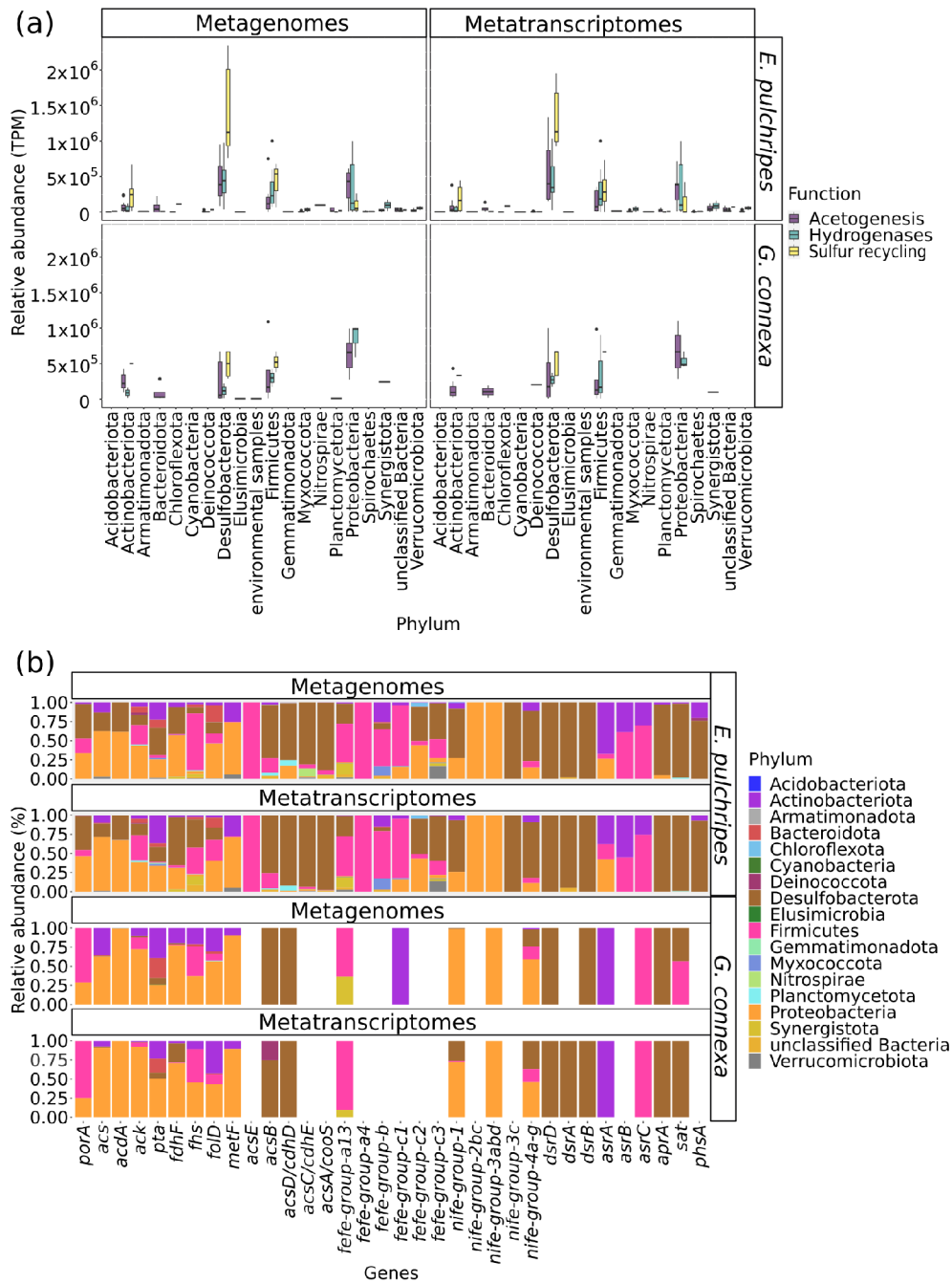


Fig. S3. Relative abundance and taxonomic distribution of genes involved in acetogenesis, hydrogen sensing/evolution/bifurcation (hydrogenases) and sulfur cycling in the metagenomic (MG) and metatranscriptomic (MT) contigs from the hindguts of *E. Pulchripes* and *G. connexa*. (a) Boxplots showing the relative abundance of the bacterial genes for a function within a phylum. (b) Taxonomic distribution of genes and transcripts at the phylum level. The pair-end reads from metagenomes and metatranscriptomes were mapped to all the genes to get their coverages and averaged to estimate their relative abundance in transcript per kilobase million (TPM).

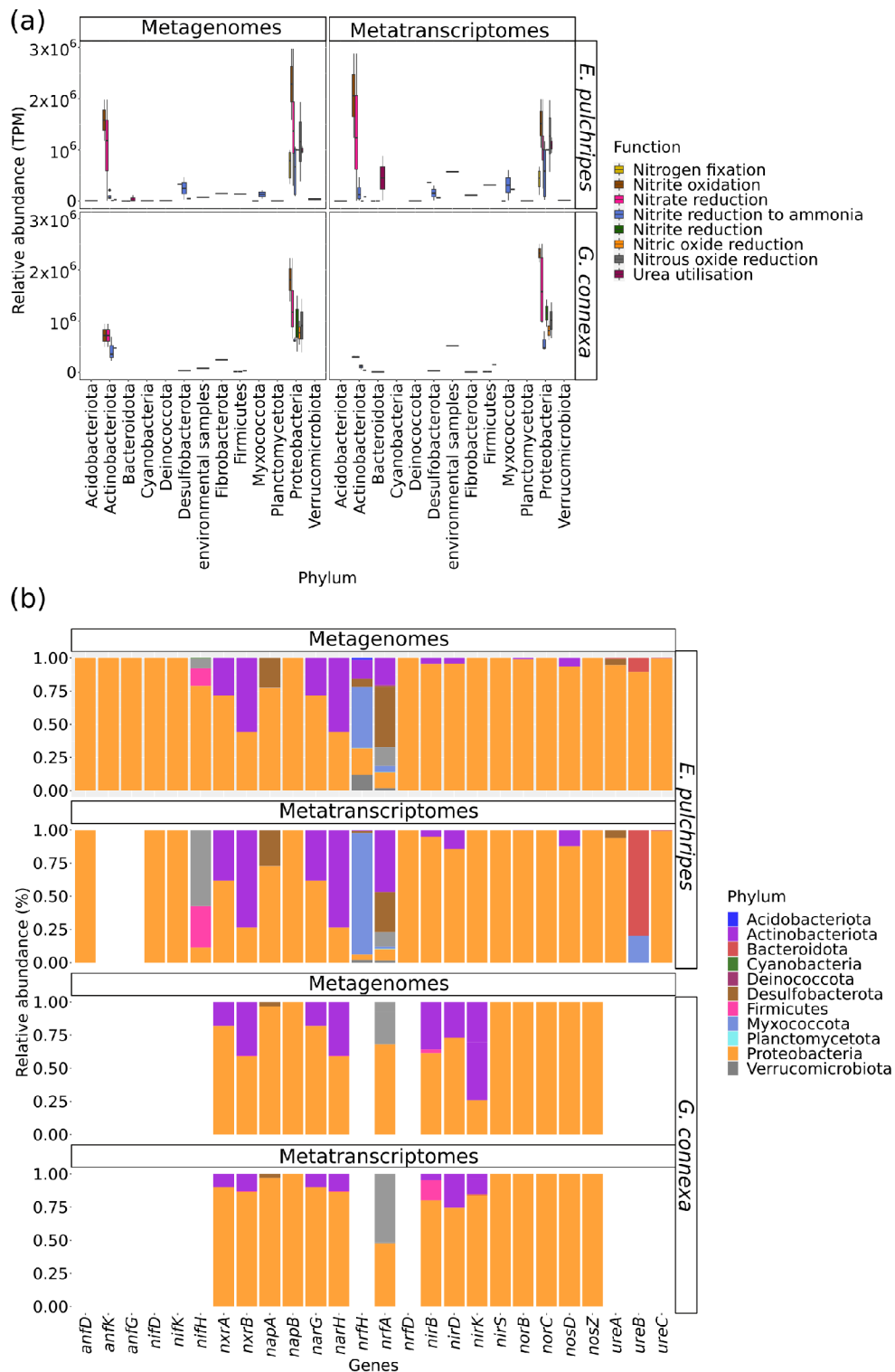


Fig. S4. Relative abundance and taxonomic distribution of genes involved in nitrogen fixation and recycling, and their corresponding transcripts. (a) Boxplots showing the relative abundance of the genes for a function within a phylum. (b) Taxonomic distribution of the genes and transcripts at the phylum level. The pair-end reads from metagenomes and metatranscriptomes were mapped to all the genes to get their coverages and averaged to estimate their relative abundance in transcript per kilobase million (TPM).

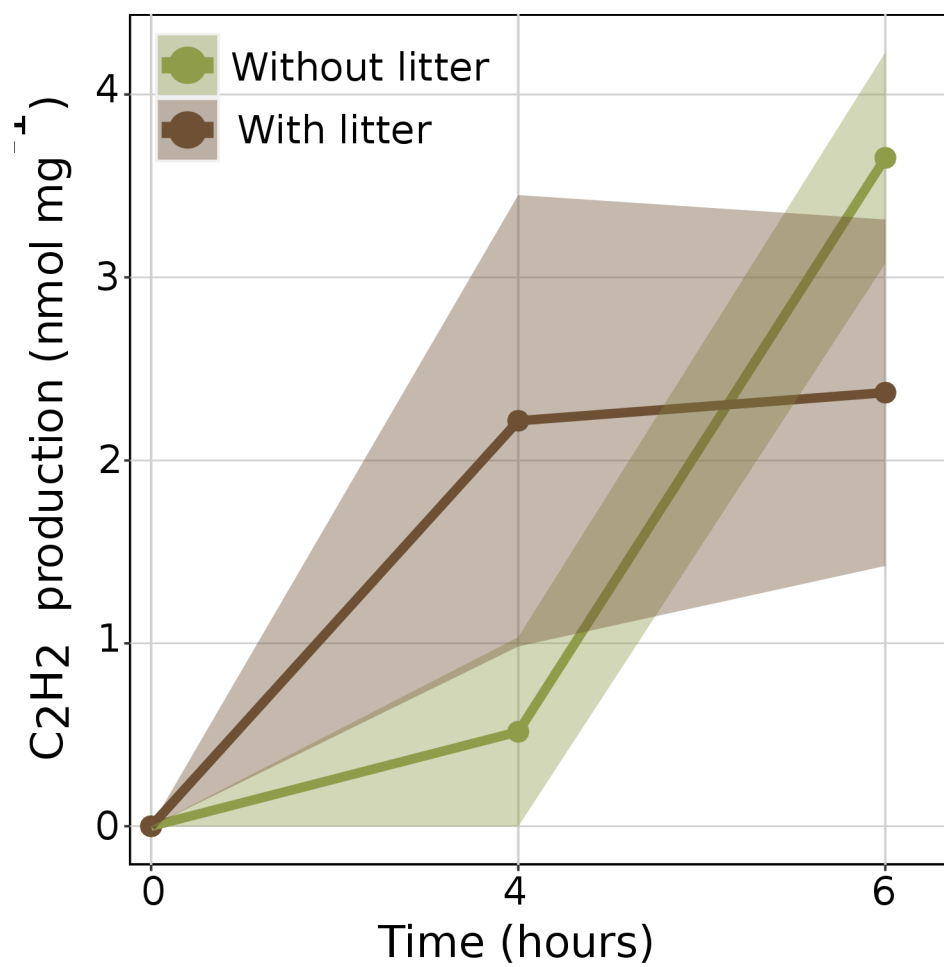


Fig. S5. Functional assay for the activity of the N<sub>2</sub>-fixing nitrogenase enzyme in the reduction of acetylene to ethylene in *E. pulchripes*.



## 5.2 Paper No. 2

**Julius Eyiuche Nweze**, Johannes Sergej Schweichhart, Roey Angel (2024).  
Viral communities in millipede guts: Insights into the diversity and potential  
role in modulating the microbiome. *Environmental Microbiology*,  
26(2),e16586. <https://doi.org/10.1111/1462-2920.16586> **(IF 5.5)**



# Viral communities in millipede guts: Insights into the diversity and potential role in modulating the microbiome

Julius Eyiuche Nweze<sup>1,2</sup> | Johannes Sergej Schweichhart<sup>1,2</sup> | Roey Angel<sup>1</sup> 

<sup>1</sup>Institute of Soil Biology and Biogeochemistry, Biology Centre CAS, České Budějovice, Czechia

<sup>2</sup>Department of Ecosystem Biology, Faculty of Science, University of South Bohemia in České Budějovice, České Budějovice, Czechia

## Correspondence

Roey Angel, Institute of Soil Biology and Biogeochemistry, Biology Centre CAS, České Budějovice, Czechia.

Email: [roey.angel@bc.cas.cz](mailto:roey.angel@bc.cas.cz)

## Funding information

Czech Science Foundation (GA ČR), Grant/Award Number: 21-04987S; Grantová Agentura České Republiky, Grant/Award Number: 19-24309Y

## Abstract

Millipedes are important detritivores harbouring a diverse microbiome. Previous research focused on bacterial and archaeal diversity, while the virome remained neglected. We elucidated the DNA and RNA viral diversity in the hindguts of two model millipede species with distinct microbiomes: the tropical *Epibolus pulchripes* (methanogenic, dominated by Bacillota) and the temperate *Glomeris connexa* (non-methanogenic, dominated by Pseudomonadota). Based on metagenomic and metatranscriptomic assembled viral genomes, the viral communities differed markedly and preferentially infected the most abundant prokaryotic taxa. The majority of DNA viruses were Caudoviricetes (dsDNA), Cirlivirales (ssDNA) and Microviridae (ssDNA), while RNA viruses consisted of Leviviricetes (ssRNA), Potyviridae (ssRNA) and Eukaryotic viruses. A high abundance of subtypes I-C, I-B and II-C CRISPR-Cas systems was found, primarily from Pseudomonadota, Bacteroidota and Bacillota. In addition, auxiliary metabolic genes that modulate chitin degradation, vitamins and amino acid biosynthesis and sulphur metabolism were also detected. Lastly, we found low virus-to-microbe-ratios and a prevalence of lysogenic viruses, supporting a *Piggyback-the-Winner* dynamic in both hosts.

## INTRODUCTION

With over 13,000 known species, millipedes are crucial detritivores that play a significant role in tropical and temperate ecosystems (Byzov, 2006; Crawford, 1992). Millipedes support the cycling of organic matter in terrestrial ecosystems by consuming large amounts of recalcitrant plant litter (Joly et al., 2020; Wang et al., 2018). Millipedes host a diverse array of intestinal organisms, including bacteria, archaea, fungi, nematodes, protists and viruses; many of which have unknown functions (Byzov, 2006). Among these, bacteria are implicated as the primary contributors to polysaccharide degradation, essential amino acid biosynthesis, short-chain fatty acid metabolism and fermentation in millipede guts (Koubová et al., 2023; Nweze et al., 2024; Sardar, Šustr, Chroňáková, & Lorenc, 2022; Sardar, Šustr, Chroňáková, Lorenc, & Faktorová, 2022; Taylor, 1982).

In most ecosystems, including animal guts, bacteria are the main hosts of viruses (phages) (Kirsch et al., 2021). Phages affect bacterial communities by changing their composition in a predator–prey dynamic, stimulating nutrient cycling, conferring immunity against other phages through superinfection exclusion, horizontally transferring genetic material, and modulating their host's metabolism via auxiliary metabolic genes (AMGs) (Díaz-Muñoz & Koskella, 2014; Mirzaei & Maurice, 2017; Quistad et al., 2017; Shkoporov & Hill, 2019). AMGs are notably prevalent in viral genomes, and metagenomic and viromic analyses have revealed a multitude of new AMGs over time (Chen et al., 2020; Emerson et al., 2018). Examining the viral AMG compositions in the millipede gut would provide us with valuable insights into the ecological roles of viruses.

Prokaryotes defend themselves against viruses by storing segments of viral DNA called spacers in the

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2024 The Authors. *Environmental Microbiology* published by Applied Microbiology International and John Wiley & Sons Ltd.

CRISPR locus. This system, known as clustered regularly interspaced short palindromic repeats associated proteins (CRISPR-Cas), acts as a memory of past infections. When a virus returns, it uses these spacers to produce guide RNAs (gRNAs) for precise targeting (Hille & Charpentier, 2016; Koonin & Makarova, 2009). Diverse *cas* gene variations and distinct CRISPR locus arrangements lead to multiple types of CRISPR-Cas systems, which can target DNA or RNA viruses (Watson et al., 2021). Understanding which hosts are abundant with CRISPR-Cas systems and what their target viral types are can inform us about the viral-host dynamics.

Two major models were proposed to explain virus-host dynamics in the environment: ‘Kill-the-Winner’ (KtW) predicts that lytic viruses will support high host diversity by suppressing the most abundant hosts (Thingstad, 2000; Winter et al., 2010), while ‘Piggy-back-the-Winner’ (PtW) suggests that decreasing virus-to-microbe-ratio (VMR) with higher host densities is due to selection favouring a lysogenic lifestyle (Knowles et al., 2016; Silveira & Rohwer, 2016). VMR is traditionally determined by counting virus-like particles and cells (Parikka et al., 2017), but this method has limitations (Danovaro & Middelboe, 2010; Holmfeldt et al., 2012; Kaletta et al., 2020). Metagenomics VMR (mVMR) offers an alternative and is effective, especially in complex environments like soils and gut ecosystems with prevalent lysogenic phages (López-García et al., 2023). Predicting the VMR in the millipede gut is important because it holds significance for both millipede health and ecology, serving as an indicator of interactions between viruses and bacteria.

Thus far, only a few studies on millipede viruses have been published. In a July 2023 Scopus database search for ‘millipede’ and ‘virus’ or ‘phages’, only 10 results were found, with just one relevant to millipede-associated viruses. This is surprising in light of several works (Li, Shi, et al., 2015), which revealed that arthropods, including millipedes, can host evolutionary predecessors of significant pathogenic RNA viruses in vertebrates, such as Influenza and Ebola, and even discovered entirely new RNA virus families (Li, Shi, et al., 2015; Kirsch et al., 2021; Shi et al., 2016). In addition to RNA viruses, a high diversity of ssDNA *Cressdnaviricota* was also seen in millipedes (Kirsch et al., 2021).

Recent advances in metagenomics, metatranscriptomics, and bioinformatic tools provide increasingly insightful means to study phages in the environment. These include assembly and analysis of phage and host genomes, inferring phage-host interactions (e.g., through CRISPR-Cas systems), potential effects of viral infection on the host metabolism through AMGs, and the detection of lysogenic phages (Coutinho et al., 2018; Puxty et al., 2018; Watson et al., 2021; Zhang et al., 2018).

Using a previously published metagenomics and metatranscriptomics dataset, we investigated the DNA and RNA virus diversity in the hindguts of two model millipede species: *Epibolus pulchripes* (order: Spirobolida) from the tropical East African coast (Enghoff, 2011) and *Glomeris connexa* (order: Glomerida) from temperate Central Europe. Despite sharing a similar detritivorous lifestyle, these species differ in size and gut redox conditions, with *G. connexa* being smaller (10–17 mm) compared to *E. pulchripes* (130–160 mm). As a result, *G. connexa* possesses a mostly oxic gut, which is overwhelmingly dominated by Pseudomonadota (*Proteobacteria*), while *E. pulchripes* has an anoxic, highly-reduced, and CH<sub>4</sub>-emitting gut, which is dominated by Bactroidota and hosts many Bacillota (Firmicutes). In addition, we investigated the abundance of CRISPR-Cas loci in the prokaryotic community using both libraries and metagenome-assembled genomes (MAGs) from both millipede species. We also employed predictive methods to identify potential hosts for the detected viruses. Lastly, we evaluated the abundance of viral AMGs.

## EXPERIMENTAL PROCEDURES

### Collection of metagenomic and metatranscriptomic data

Quality-filtered metagenomic and metatranscriptomic data from the hindguts of *E. pulchripes* and *G. connexa* were obtained from our previous study (NCBI BioProject PRJNA948469) (Nweze et al., 2024). The datasets were produced from the gut content of dissected millipedes. Juvenile individuals of the tropical millipede *E. pulchripes* were obtained from a breeding colony maintained in our lab. The temperate millipede *G. connexa* was collected from the wild. The animals were maintained in terraria in the laboratory for 25 days before dissection (Nweze et al., 2024). For each set of libraries, we used three replicates of paired-end reads for each millipede species. From the metagenomic dataset, 305 metagenome-assembled genomes recovered from the hindguts of both millipede species were recruited.

### De novo assembly of reads and viral identification

Quality-filtered metagenomic reads from three replicated samples of each millipede species were co-assembled using MEGAHIT v1.2.9 (Li, Liu, et al., 2015) with default settings. For each millipede species, the non-rRNA metatranscriptomic reads were also co-assembled using Trinity v2.13.2 (Grabherr et al., 2011). The contig information for both assemblies was retrieved by



creating a contig database in *anvi'o*-7.1 and running *anvi-display-contigs-stats* (Eren et al., 2021).

Viral sequences were retrieved from both libraries using *VirSorter2*, which is a multi-classifier that detects diverse dsDNA viruses, ssDNA viruses, dsDNA phages, RNA viruses, *Nucleocytoviricota* (NCLDV) and virophages (*Lavidaviridae*) (Guo et al., 2021). Following recent advances in the discovery of viruses with small genome sizes (see Discussion), the minimum contig length was set to 1 kb. To reduce the impact of false positives, identified viral contigs were further confirmed and classified using *geNomad* v1.5.0 (Camargo et al., 2023). The identified viruses are assigned to taxonomic lineages according to the International Committee on Taxonomy of Viruses (ICTV) (Lefkowitz et al., 2018). This software serves to distinguish viruses from plasmid genomes and subsequently assigns taxonomy to identified viral genomes. It predicts genes in input sequences via *Pyrodigal-gv* (Larralde, 2022) and assigns these predicted genes to specific marker protein families from a dataset of 227,897 profiles. This differentiation is achieved using *MMseqs2* (Steininger & Söding, 2017), which is specific to chromosomes, plasmids or viruses. As post-classification filters, sequences were classified as a plasmid or virus with a score of at least 0.7, and sequences shorter than 2.5 kb were required to encode at least one hallmark gene. Additionally, the potential host regions in any identified proviruses were retained.

### Quality of single-contig viral genomes

*CheckV* v1.0.1 (Nayfach et al., 2021) was used to evaluate the quality of single-contig viral genomes from *geNomad*. This included the identification and removal of the host (prokaryotes, eukaryotes or millipedes) contamination for integrated proviruses, estimation of completeness for genome fragments and identification of closed genomes. Based on these three criteria, the single-contig viral genomes were categorised into one of the five quality tiers, which conform to and extend the Minimum Information about an Uncultivated Virus Genome (MIUViG) quality standards: complete, high-quality (>90% completeness), medium-quality (50%–90% completeness), low-quality (<50% completeness) and undetermined quality (Roux, Adriaenssens, et al., 2019). We selected all genomes with ≥50% estimated completeness for further analysis.

### Relative abundance calculations

To determine the abundance of the identified viruses, the quality-filtered reads from both libraries were aligned to these sequences. The mean coverage obtained from the alignment was then used to estimate

the relative abundance, expressed in transcripts per million (TPM), using the *CoverM* tool, with a minimum read percent identity threshold of 90% (<https://github.com/wwood/CoverM>). The alignment process was conducted using the *bwa-mem* aligner (Li & Durbin, 2010). The relative abundances of the major capsid proteins (MCP), RNA-dependent RNA polymerases (RdRP) and cellular universal single-copy (USCG) in the metagenomic and metatranscriptomic reads were estimated using the same tool and were plotted using the *R ggplot2* package (Wickham, 2016), visualising the different viral taxa.

### Lifestyles and host prediction

The lifestyles (virulent or lysogenic) of the identified viral genomes were predicted using the *Phage TYPE* prediction tool (*PhaTYP*) with default settings (Shang et al., 2023). Host prediction was done using the integrated *Phage Host Prediction* (*iPhoP*) tool with default settings, as described by Roux et al. (2023). The host database was based on MAGs and the *iPhoP* database composed of the GTDB release 202 (47,894 genomes), Public IMG genomes not already in GTDB, as of 7 July 2021 (21,372 genomes), and from the GEM dataset (<https://portal.nersc.gov/GEM/>, 52,515 genomes). The taxonomic annotations for the 305 MAGs in our database were performed using *GTDB-Tk* v1.7.0 (GTDB release 202) (Chaumeil et al., 2020). The predicted host is determined based on the prokaryotic genomes with the highest probability.

From the 304 retrieved MAGs obtained from the hindguts of both millipede species, putative prophages were detected and annotated using *DBSCAN-SWA* (Gan et al., 2022). This tool efficiently predicts prophage regions within bacterial genomes, exhibiting quicker processing times compared to previous tools and demonstrating significant detection capability validated through an analysis involving 184 manually curated prophages. *CheckV* v1.0.1 was used to check the completeness of the identified prophages.

### The viral proteomic tree

For constructing the viral proteomic tree, we used *ViP-Tree* v3.6 (Nishimura et al., 2017), a tool that calculates global genome-wide similarities between viruses using *tBLASTx* after selecting a preset reference viral genomes stored in *Virus-Host DB* classified into six categories mainly based on their nucleic acid types: dsDNA, ssDNA, dsRNA, ssRNA, ssRNA-RT and dsRNA-RT. The  $S_G$  computation method follows Bhunchoth et al. (2016), employing phylogenies via the *ETE3* toolkit package v3.0.0b33 (Huerta-Cepas

et al., 2016). Sequences were aligned using MAFFT v6.861b (Kato & Toh, 2008) and trimmed with trimAl v1.4.rev6 (Capella-Gutiérrez et al., 2009) for improved trees, and ProtTest (Abascal et al., 2005) in pmodeltest v1.4 to select protein evolution models (WAG, VT, LG). Tree reconstruction used RaxML v8.1.20 (Stamatakis, 2014) with the chosen model and PROTGAMMA parameter, drawing branch supports from 100 bootstrap replicates. R package phangorn's midpoint function rooted the trees (Schliep, 2011). The reference viruses' genome and host information are based on the Virus-Host DB (Mihara et al., 2016). The scaling of branch lengths was performed using a log transformation.

## Identification of CRISPR systems

CRISPRCasTyper v1.8.0 (Russel et al., 2020) was used to detect CRISPR-Cas genes and arrays in the metagenome, MAGs and metatranscriptome. The prediction probability was set at 0.95. Predictions were based on the CRISPR-Cas, Cas operons, Cas operons orphan (those not in CRISPR\_Cas.tab), CRISPR, CRISPR near Cas and CRISPR orphan. Those designated as putative CRISPR-Cas, Cas operons and CRISPR were not included in the final plots.

## Identification of potential AMGs

The viral AMGs within the quality viral contigs were identified using geNomad (Camargo et al., 2023) and VIBRANT (Kieft et al., 2020). The Carbohydrate Active EnZyme (CAZy) annotation was further performed using the dbCAN2 server (<https://bcb.unl.edu/dbCAN2/>) with default settings. We manually examined the genomic context to validate the functional annotation of the candidate AMGs. Based on this, a gene was considered a high-confidence viral AMG if it was located between two viral hallmark genes or virus-like genes or positioned adjacent to a viral hallmark gene or a virus-like gene. A genomic map of single-viral contigs containing AMGs of interest was visualised using the gggenes R package (Wilkins, 2023).

## Virus-microbe ratio

Metagenomics virus-to-microbe ratios (mVMR) were calculated by searching for the MCP, RdRP and cUSCG genes as proxies for DNA viruses, RNA viruses and cell counts, respectively (López-García et al., 2023). Based on that, we considered an mVMR below 1 to be 'low', between 1 and 3 to be 'medium', and above that to be 'high'. USCGs classified as Streptophyta, *Arthropoda* and other invertebrates were not considered. The metagenome assembly was screened

for 132 MCPs against hidden Markov models (HMMs) using hmmsearch v3.3 (Wheeler & Eddy, 2013) with a significance threshold ( $E$ -value =  $-5$ ), using the preformed MK\_Selection.hmm profiles provided by the authors. The HMM recognises MCPs from virtually all known families of DNA viruses infecting bacteria, archaea and eukaryotes. For RdRPs in the metatranscriptome, a comprehensive dataset of HMM profiles of RdRp domains called NeoRdRp was used (Sakaguchi et al., 2022). In both library types, a similar hmmsearch was conducted for 40 universal single-copy genes (USCGs) from fetchMGs (<https://github.com/motu-tool/fetchMGs>). To reduce the variation in ribosomal protein genes, the average relative abundances of USCGs at the class level were used as proxies for cell counts. The mVMR was calculated by selecting the top 15 most abundant USCGs, following López-García et al. (2023). For MCPs and RdRp gene identification and taxonomy assignment, geNomad-annotate (Camargo et al., 2023) was used, employing a marker database for viral classification. USCGs were assigned to taxa via BLAST search against the NCBI nucleotide collection.

## RESULTS

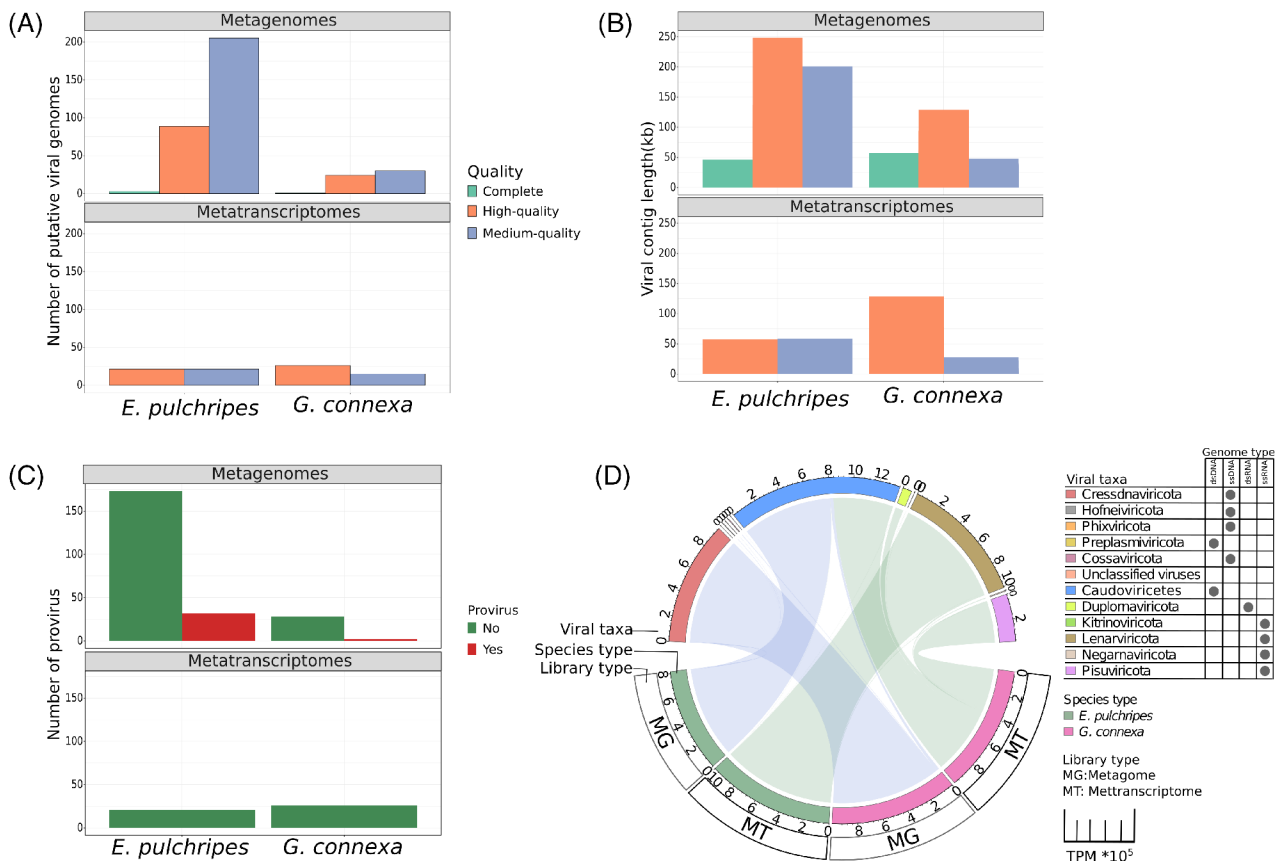
### The quality of metagenomic and metatranscriptomic assembled reads

Each millipede species was represented by three replicates of metagenomic and metatranscriptomic paired-end reads from their hindguts. Detailed assembly information for these libraries from *E. pulchripes* and *G. connexa* can be found in Nweze et al. (2024), which is summarised in Table S1.

### Identification of viral contigs and quality assessment

In total, we identified 4693 viruses (sequences with virus identity) in *E. pulchripes* from the metagenomic library and 1257 viruses from the metatranscriptomic library (Figure 1A; Table S2). Among the identified metagenome-assembled viral genomes (MAVGs) from the metagenomic library, three were determined to be complete, 89 were high quality and 206 were medium quality. The lengths ranged from 2.4 to 248 kb (Figure 1B; Table S2). Additionally, 3871 MAVGs were classified as low-quality, and 524 MAVGs were undetermined (no viral genes detected by CheckV). From the metatranscriptomic library, 21 MAVGs were high quality, 21 were medium quality, 1049 were classified as low quality, and 166 MAVGs were undetermined. The viral genome lengths ranged from 2.4 to 59 kb.

In *G. connexa*, we identified a total of 1048 MAVGs from the metagenomic library, with 1 classified as



**FIGURE 1** The quality of DNA and RNA viruses (MAVGs) in the assembled sequence reads. (A) The number of MAVGs (complete, high-quality and medium-quality) identified in metagenomes and metatranscriptomes from the hindgut of *Epibolus pulchripes* and *Glomeris connexa* and assessed by CheckV. (B) The average length of the MAVGs. (C) The number of viral genomes classified as free or provirus with evidence of host integration. (D) The relative abundance of the MAVGs (bottom arch) identified in metagenomes (MG) and metatranscriptomes (MT) from the hindgut of *E. pulchripes* and *G. connexa*. The pair-end metagenomic reads were mapped to the identified AVGs to get the coverage and calculate the relative abundance in transcript per million (TPM; top arch). The mean TPM was calculated from the three replicate samples and totalled for each taxonomic level (TPM  $\times 10^5$ ).

complete, 24 as high-quality, and 30 as medium-quality (Figure 1A). Like in *E. pulchripes*, none of the MAVGs in the metatranscriptomic library were classified as complete. Instead, 26 high-quality viruses and 15 medium-quality viruses were identified. Eight hundred and forty-six MAVGs in the metagenomic library and 133 in the metatranscriptomic library were of low quality (Table S3). Additionally, 147 MAVGs from the metagenomic library and 51 MAVGs from the metatranscriptomic library could not be determined or classified. The lengths of the complete, high-quality and medium-quality viral genomes ranged from 2.2 to 129.2 kb (metagenome) and 2.1 to 128.9 kb (metatranscriptome), respectively (Figure 1B).

## Viral diversity and abundance in millipede hindguts

The viral genomes analysed in this study using geNomad showed no significant gene similarities with

previously deposited reference viruses, making it challenging to assign them to specific families with confidence. In *E. pulchripes*, high-quality MAVGs grouped into 253 free viruses and 45 proviruses (i.e., the host region was found on both ends of viral genomes). Conversely, metagenome-derived high-quality viral genomes from *G. connexa* comprised 52 free viruses and 3 proviruses (Figure 1C). Following the ICTV guidelines for taxonomic classification of viruses from metagenomes (Simmonds et al., 2023), the viral genomes in the metagenome from *E. pulchripes* were classified into three realms (Table S4): *Duplodnaviria* (281), *Monodnaviria* (13) and *Varidnaviria* (2). In the metatranscriptome, two realms were found: *Riboviria* (27) and *Duplodnaviria* (15). Upon initial examination of their abundances, the class Caudoviricetes (98.1%) (Uroviricota) dominated the metagenome (Figure 1D), followed by Phixviricota (0.4%) and Preplasmiviricota (0.2%). The phylum Lenarviricota (98.9%) dominated the metatranscriptome, followed by Pisuviricota (0.7%). In the metagenome from *G. connexa*, 50 Duplodnaviria,



and 5 Monodnaviria MAVGs were the only realms found. It is worth noting that while the majority of MAVGs from this host species belonged to Duplodnaviria, a single Monodnaviria class (Arfiviricetes) made up nearly 97% of the relative abundance (TPN; Figure 1D). From the metatranscriptome, we identified Riboviria (32), classified into five phyla, and Duplodnaviria (9). The phylum Cressdnaviricota (96.6%) had the highest relative abundance in the metagenome, followed by Caudoviricetes (2.8%) and Cossaviricota (0.6%). In the metatranscriptome, Caudoviricetes (53.5%) was the most prevalent, followed by Pisuviricota (36.6%), Duplornaviricota (7.1%) and Lenarviricota (2.1%).

### Lifestyles and virus-host linkage prediction

Successful lifestyle predictions were made for 289 DNA MAVGs and 7 RNA MAVGs, while in *G. connexa*, it was 50 and 14 for DNA and RNA MAVGs, respectively (Figure 2A; Table S5). In *E. pulchripes*, 57% of viral genomes from the metagenome were lysogenic (temperate), while 86% from the metatranscriptome were lytic (virulent). Similarly, 58% of viral genomes from the metagenome in *G. connexa* were lysogenic, while 93% from the metatranscriptome were virulent. The class Caudoviricetes predominated among the predicted viral genomes in both millipede species, displaying a mix of lysogenic and virulent lifestyles (Figure 2B).

Putative viral hosts were successfully assigned to 141 viral genomes from *E. pulchripes* (ca. 42%) and 28 from *G. connexa* (ca. 29%; Table S6). The most frequently predicted host for the viral genomes from *E. pulchripes* was Bacteroidota (32.1%), followed by Bacillota (29.3%), Pseudomonadota (18.6%) and Desulfobacterota (8.57%; Figure 2C). In the case of *G. connexa*, the most commonly predicted host was Pseudomonadota (57.1%), followed by Bacteroidota (21.4%) and Bacillota (10.7%; Figure 2D). Additionally, other viruses with predicted hosts included Phixviricota (two Bacteroidota) and Cressdnaviricota (one Bacillota) in *E. pulchripes*, as well as Cossaviricota (one Bacteroidota) in *G. connexa*.

Within the MAGs from *E. pulchripes*, we identified 374 prophages (Figure 2E; Table S5), categorised as high-quality (4), medium-quality (20), low-quality (236) and not-determined (114). Similarly, in *G. connexa*, 89 prophages were identified, categorised as high-quality (4), low-quality (49), medium-quality (13) and not-determined (23). All prophages classified as high- and medium-quality were assigned to the class Caudoviricetes, comprising 24 prophages from *E. pulchripes* and 17 from *G. connexa*. These prophages originated from bacterial phyla, including Bacillota, Pseudomonadota, Desulfobacterota, Actinomycetota and Spirochaetota in

*E. pulchripes*, and Pseudomonadota and Bacillota in *G. connexa*.

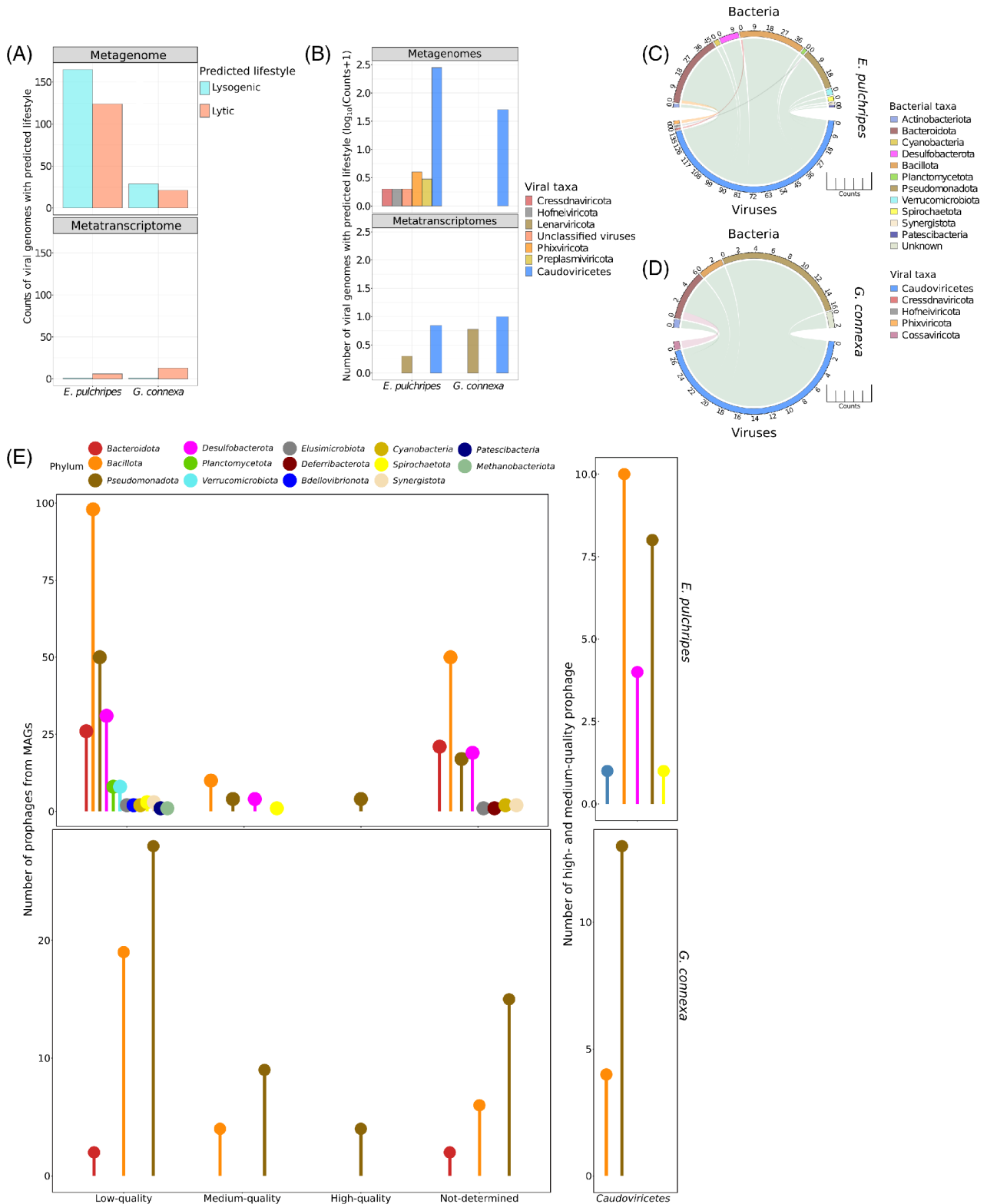
### Amino-acid-based viral phylogeny

Viruses lack universal genes that can be used to construct a unified phylogeny into which all viruses can be classified (Holmes, 2011; Rohwer & Edwards, 2002). We, therefore, reconstructed a proteomic tree based on translated nucleic acid sequence according to virus types (dsDNA/ssDNA/dsRNA/ssRNA).

For *E. pulchripes*, within the DNA viruses, the dominant class Caudoviricetes had MAVGs clustered with the family Autographiviridae, whose host groups were Cyanobacteria (1) and Pseudomonadota (3). The rest were clustered into unclassified viral families associated with hosts such as Pseudomonadota (19), Bacterioidota (15), Actinomycetota (10), Bacillota (33), Cyanobacteria (4) (Figure 3A; Table S7). Two MAVGs clustered in the phylum Preplasmiviricota within the family Tectiviridae. The host for this reference family was Pseudomonadota (Figure 3B; Table S8).

In the Cressdnaviricota, eight MAVGs clustered with the family Plectroviridae, whose host is Mycoplasmatota (Figure 3C; Table S8). One quality MAVG in the phylum Hofneiviricota was recovered and clustered as a distant clade of the Inoviridae family, whose host is Pseudomonadota (Figure 3D; Table S9). Two closely related and two distantly related Phixviricota MAVGs were assigned to the family Microviridae from the reference genomes (Figure 3E; Table S8), with Pseudomonadota and Bacteroidota as possible hosts.

In the RNA pool, we identified 15 MAVGs in the class Caudoviricetes. The majority of these MAVGs (9) were assigned to the Autographiviridae family with Pseudomonadota as hosts (Supplementary Figure 1A; Table S9). The remaining MAVGs (6) clustered with unclassified families, with Pseudomonadota (4) and Bacillota (2) as host groups. The three Kitrinoviricota MAVGs did not cluster with any known phages or viruses (Supplementary Figure 1B; Table S10). In the Lenarviricota, six MAVGs formed a distinct cluster separate from the reference phages and could not be assigned a family or host, while two MAVGs clustered with phages in the family Fiersviridae, with Pseudomonadota as the hosts (Supplementary Figure 1C; Table S10). Three MAVGs assigned to Duplornaviricota were found within a cluster of the viral family Cystoviridae, with Pseudomonadota as the host (Supplementary Figure 1F; Table S10). As for eukaryotic viruses, we found several MAVGs associated with the phyla Lenarviricota, Pisuviricota, Picornaviridae and Negarnaviricota, many of which are from unclassified viral families. The hosts for these viruses included Ascomycota (fungi) and Arthropoda (Supplementary Figure 1D,E,G; Table S10).

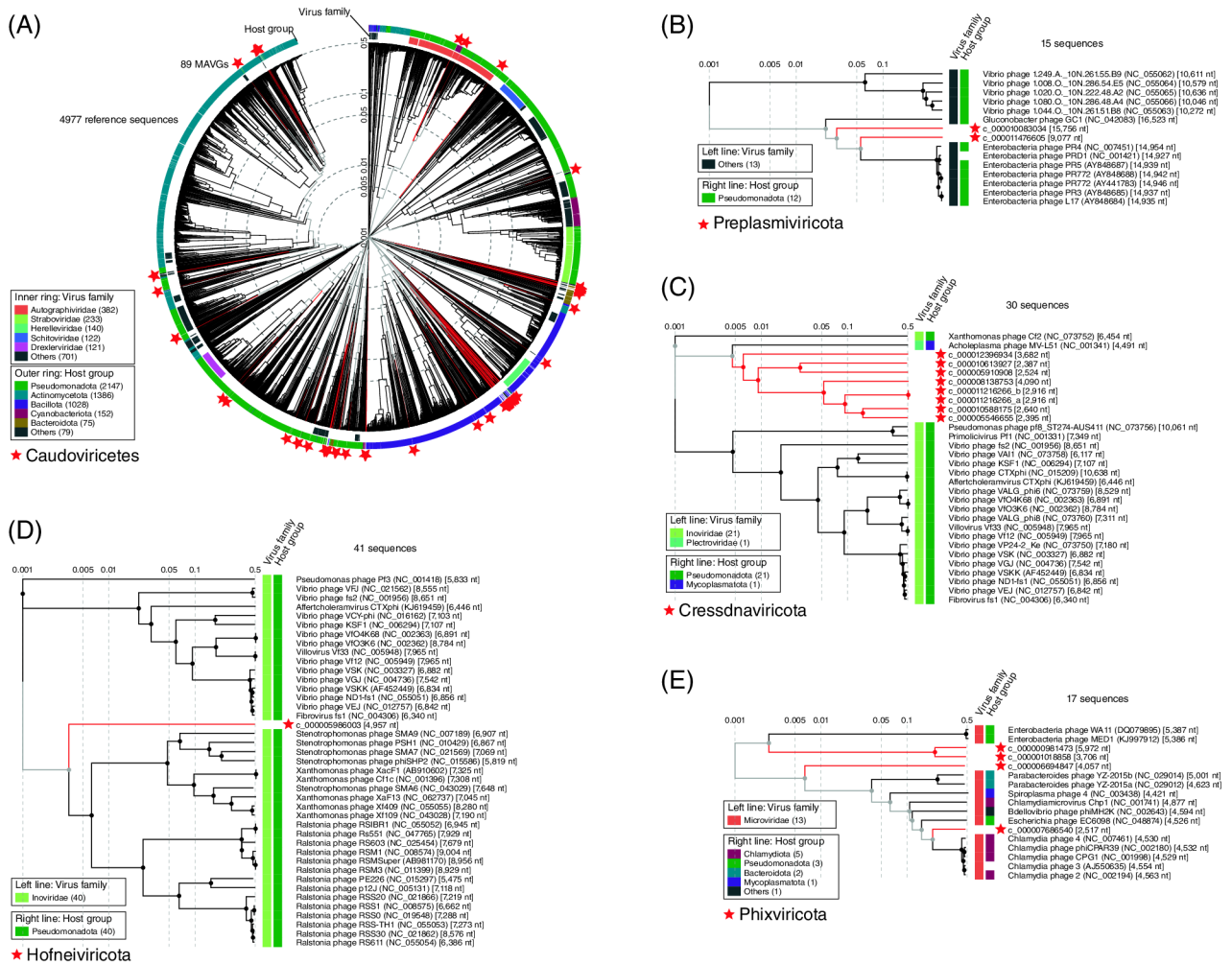


**FIGURE 2** Lifestyles and putative host assignment to viral genomes from the hindguts of *Epibolus pulchripes* and *Glomeris connexa*. (A) Predicted lifestyles of the identified viral genomes. (B) Taxonomic origin of the viral genomes with predicted lifestyles. Predicted host for the identified viral genomes from the hindguts of (C) *E. pulchripes* and (D) *G. connexa*. (E) Putative prophages recovered from metagenome-assembled genomes with their colour-coded bacterial and viral taxonomic origins.

In the DNA pool from *G. connexa*, proteomic analysis of 50 MAVGs belonging to Caudoviricetes revealed clustering in various unclassified families. The host

groups associated with these clusters included Pseudomonadota (16), Actinomycetota (14), Bacillota (9) and unknown (11) (Supplementary Figure 2A;





**FIGURE 3** Proteomic tree of viral genomes (MAVGs) in metagenome from *Epibolus pulchripes*. (A) Genome similarity in the dsDNA viruses between the putative viral genome classified in the class Caudoviricetes (89) and related reference viral genomes. (B) Genome similarity in the dsDNA viruses between the putative viral MAVGs classified in the phylum Preplasmiviricota (2) and related reference viral genomes. (C) Genome similarity in the ssDNA viruses between the putative viral single-genome classified in the phylum Cressdnaviricota (8) and related reference viral genomes. (D) Genome similarity in the ssDNA viruses between the putative viral single-genome classified in the phylum Hofneiviricota (1) and related reference viral genomes. (E) Genome similarity in the ssDNA viruses between the putative viral single-genome classified in the phylum Phixviricota (4) and related reference viral genomes. The red stars represent each MAVG in different phyla.

Table S11). Within the phylum Cressdnaviricota, two MAVGs were identified. One MAVG clustered with the family Circoviridae, with Mollusca as the host. The other MAVG clustered with an unclassified family, with Arthropoda identified as the host. (Supplementary Figure 2B; Table S12). The three MAVGs assigned to the phylum Cossaviricota formed a cluster with the family Parvoviridae, and the identified host for the reference family was Arthropoda. In the RNA pool, nine MAVGs were assigned to the order Caudoviricetes. Six of them clustered with the family Autographiviridae, and the remaining three clustered with unknown families (Supplementary Figure 3A; Table S12). Pseudomonadota served as their potential host. Within the phylum Lenarviricota, five MAVGs clustered with the family Fiersviridae, while three MAVGs clustered with the family Steitzviridae (Supplementary Figure 3C; Table S12). The identified host group for both families was

Pseudomonadota. However, the remaining five MAVGs did not form clusters with phages. Instead, they formed clusters with unclassified viral families with fungal hosts (Supplementary Figure 3D; Table S12).

The remaining viral phyla were classified as eukaryotic viruses. The three MAVGs within the phylum Kitrioviricota could also not be assigned to any known family but were identified as plant pathogens (Streptophyta) (Supplementary Figure 3B; Table S12). The five MAVGs assigned to the phylum Duplornaviricota clustered with the different viral groups within the family Totiviridae with different fungal hosts (Supplementary Figure 3E). Within the phylum Pisuviricota, six MAVGs clustered with the family Picornaviridae, with Chordata identified as the host group. The remaining four MAVGs formed a separate cluster with no clear hosts (Supplementary Figure 3F and Table S12). In the phylum Negarnaviricota, the MAVG clustered within the family

Rhabdoviridae, with Chordata identified as the host group (Supplementary Figure 3G; Table S12).

## Community abundance of CRISPR-Cas genes and transcripts

In the metagenomic analysis, a total of 110 CRISPR-Cas loci were identified in the gut of *E. pulchripes*, accompanied by additional 126 orphan *cas* operons and 445 orphan CRISPR arrays (those not present in CRISPR-Cas loci) (Supplementary Figure 4A, Table S13). In the metatranscriptome, only 19 CRISPR-Cas loci were observed, along with 83 orphan *cas* operons and 852 CRISPR arrays. Additionally, we found 24 putative CRISPR-Cas loci in the metagenome and 6 putative CRISPR-Cas loci in the metatranscriptome. These putative loci were classified as such because they consisted of lonely *Cas* genes adjacent to a CRISPR array. For *G. connexa*, 13 CRISPR-Cas loci were identified in the metagenome, along with 21 orphan *cas* operons and 99 CRISPRs. In the metatranscriptome, 2 orphan *cas* operons and 38 CRISPRs were found. Four putative CRISPR-Cas loci were also observed in the metagenome, while one putative CRISPR-Cas locus was identified in the metatranscriptome.

CRISPR-Cas systems use either multiple proteins (Class I systems) or a single multifunctional and multi-domain protein to target non-self genetic material (Class II systems) and are further divided into six types and over 30 subtypes (Makarova et al., 2018). Type 1 was the most abundant in both library types, followed by Type II (Supplementary Figure 4B). In the DNA pool of *E. pulchripes*, Class 1 subtype I-C was the most abundant (35), followed by subtype I-B (16), Class 2 subtype II-C (15), Class 1 subtype I-E (8) and Class 2 subtype III-E (8) (Supplementary Figure 4C). In the RNA pool, the CRISPR-Cas system was predominantly represented by subtype I-B (8) and subtype I-C (5). For *G. connexa*, the CRISPR-Cas system was only predicted in the RNA pool (13) but not in the DNA. The most abundant subtypes in the metagenome were subtypes II-C (4), I-E (3) and I-C (2). The arrangements of genes in the CRISPR-Cas system and the array direction are depicted in Supplementary Figure 5. The orphan *cas* operons (Supplementary Figure 4D) and orphan CRISPR arrays (Supplementary Figure 4E) were found in both millipede species.

## Abundance and phylogenetic origins of CRISPR-Cas arrays and associated genes in MAGs

Since CRISPR-Cas arrays cannot be taxonomically affiliated by themselves, we also investigated the abundance of CRISPR-Cas genes in MAGs to identify which

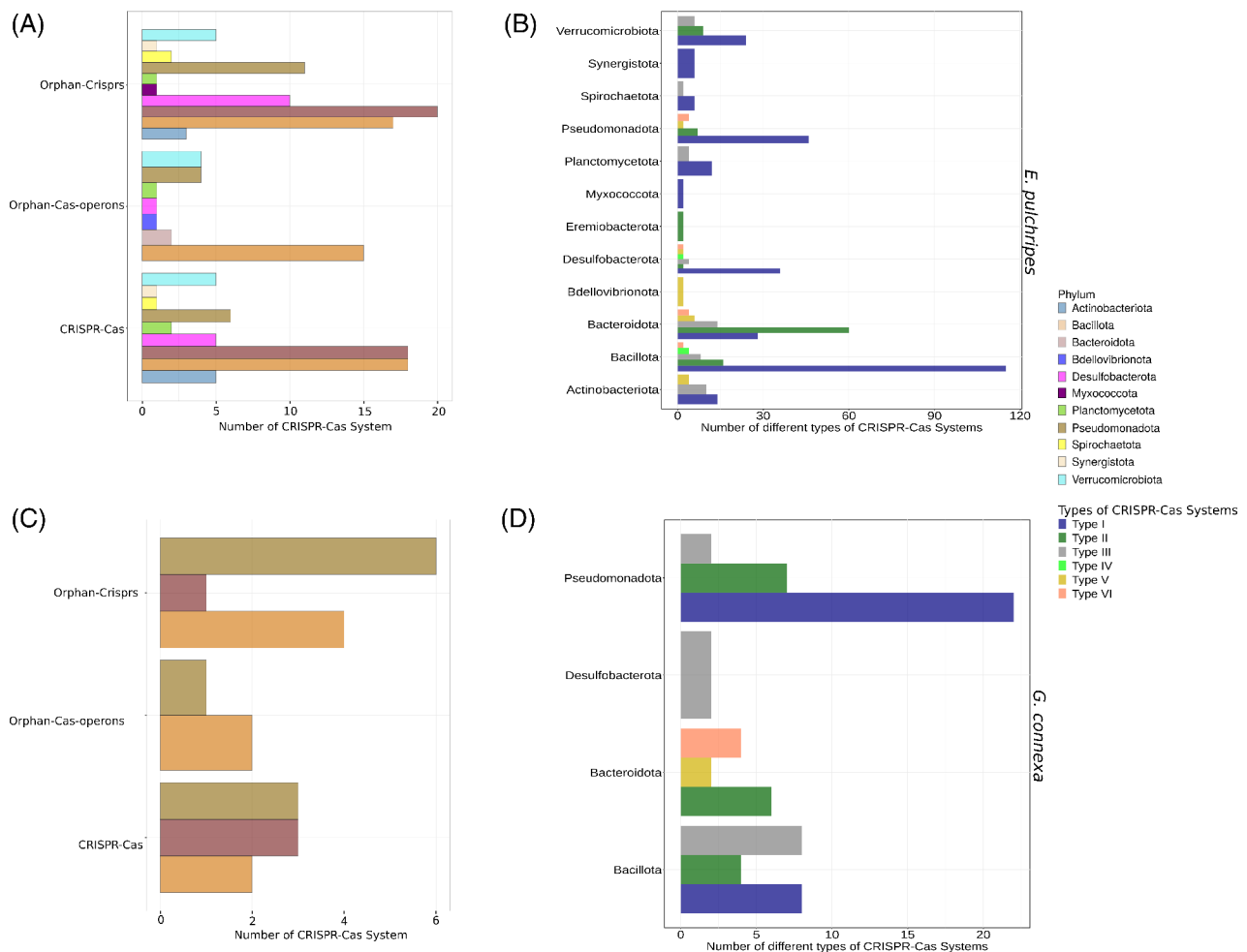
bacterial taxa possessed them (see Tables S14 and S15). From *E. pulchripes* MAGs, we identified a total of 61 CRISPR-Cas systems, along with 28 *cas* operons-orphan and 71 CRISPR orphans (Figure 4A). Among the MAGs, Bacteroidota (18) and Bacillota (18) exhibited the highest abundance of CRISPR-Cas arrays. Bacillota (15) had the highest number of orphan *cas* operons, while Bacteroidota (20) and Bacillota (17) possessed the highest number of orphan CRISPR arrays. Type I was the most abundant in most of the phyla, with the highest contribution from Bacillota, followed by Type II, with the highest contribution from Bacteroidota (Figure 4B). At the subtype level (Supplementary Figure 6A), Class 1 I-C was the most prevalent CRISPR-Cas system, followed by Class 2 II-C and Class 1 I-B. Bacillota and Desulfobacterota predominantly possessed subtype I-C, while Bacteroidota had a higher abundance of subtype II-C. Subtype I-B was found in Bacillota and Bacteroidota.

Different subtypes of (f) CRISPR-Cas loci, orphan *cas* genes and orphan CRISPR arrays were found in MAGs from *E. pulchripes* (Supplementary Figure 6B,C).

In the MAGs obtained from *G. connexa*, we identified a total of eight CRISPR-Cas systems, three *cas* operons-orphan, and 11 CRISPR orphans (Figure 4C). Type I was the most prevalent, mostly contributed by Pseudomonadota (Figure 4D). These CRISPR-Cas systems included subtypes I-C from Pseudomonadota and Bacillota, II-C from Bacteroidota and Pseudomonadota, and III-A and VI-B1 from Bacillota and Bacteroidota (Figure 6D). Similarly, these phyla also possessed a few orphan *cas* operons and CRISPR arrays (Figure 6E and F). The arrangements of genes in the CRISPR-Cas system and the array direction are depicted in Supplementary Figure 7.

## Metabolic pathways

A total of 135 AMGs were identified in both millipede species (Tables S17 and S18). These AMGs were composed of 32 distinct AMGs within 75 distinct MAVGs. In the metagenome libraries from *E. pulchripes*, 27 distinct AMGs (out of a total of 103) were found in 55 distinct MAVGs, while in the metagenome libraries from *G. connexa*, 12 distinct AMGs (out of a total of 14) were found in 9 distinct MAVGs. In the metatranscriptome from *E. pulchripes*, three distinct AMGs (out of a total of 11) were detected in eight distinct MAVGs. Similarly, in the metatranscriptome from *G. connexa*, four distinct AMGs (out of a total of seven) were found in three distinct MAVGs. Except for two MAVGs that were assigned to the phyla Phixviricota and Cressdnaviricota, which represent ssDNA viruses, the remaining MAVGs containing AMGs were assigned to the class Caudoviricetes (dsDNA viruses; Tables S16–S18). These AMGs participate in various



**FIGURE 4** Phylogenetic origins of CRISPR-Cas systems and associated genes in the metagenome-assembled genomes (MAGs) from *Epibolus pulchripes* and *Glomeris connexa*. (A) The total number of CRISPR-Cas loci, orphan cas genes, and CRISPR arrays in MAGs from *E. pulchripes* with colour-coded phyla. (B) Different types of all CRISPR-Cas systems in MAGs recovered from *E. pulchripes*. (C) A total number of CRISPR-Cas loci, orphan cas genes, and CRISPR arrays in MAGs from *G. connexa* with colour-coded phyla. (D) Different types of all CRISPR-Cas systems in MAGs recovered from *G. connexa*.

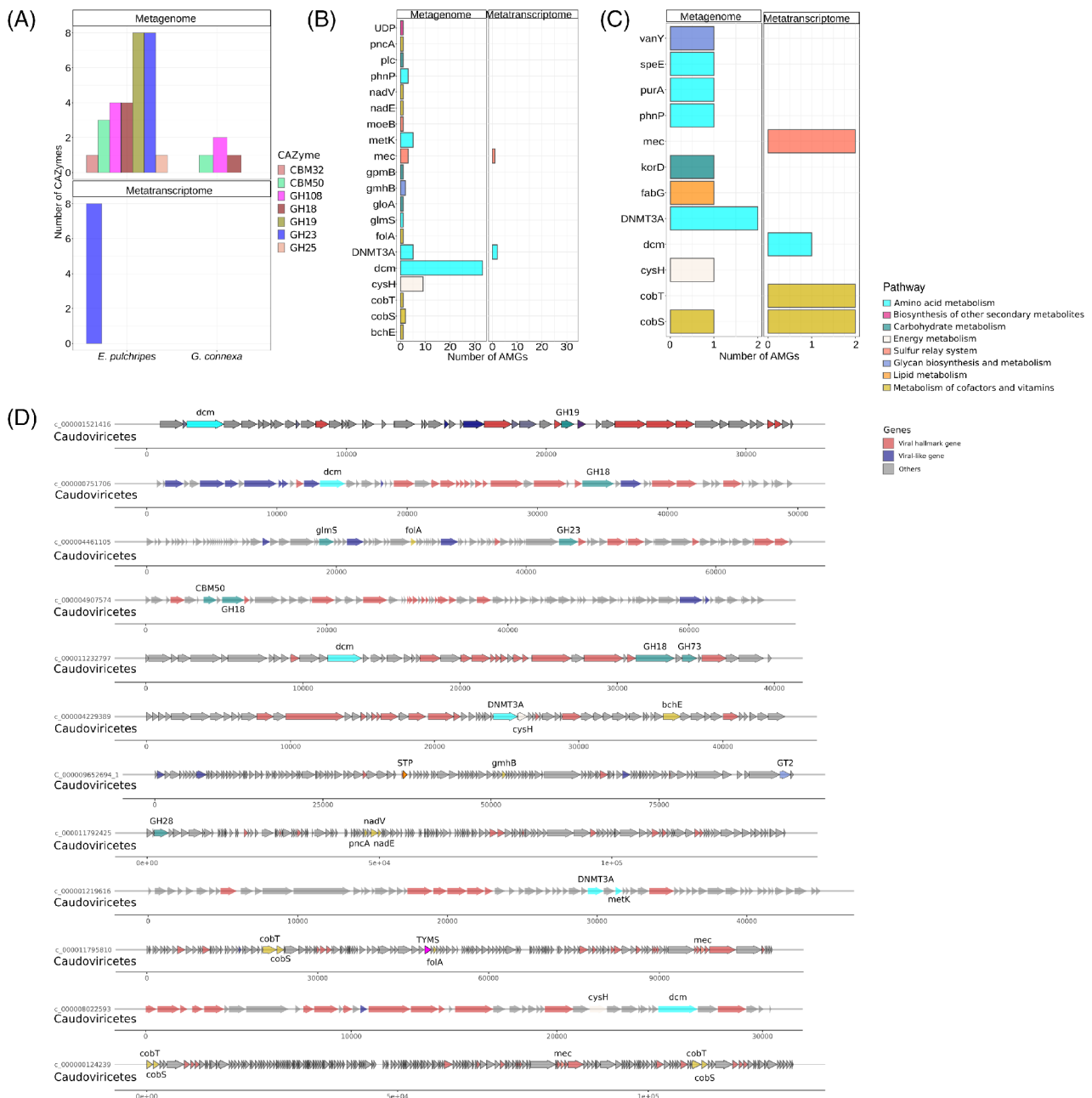
pathways associated with amino acid metabolism (55 AMGs), carbohydrate metabolism (46), metabolism of cofactors and vitamins (13), energy metabolism (10), sulphur relay system (7), glycan biosynthesis and metabolism (3), biosynthesis of other secondary metabolites (1) and lipid metabolism (1).

For carbohydrate metabolism, 42 AMGs were classified as Carbohydrate-Active enzymes (CAZymes), which included five carbohydrate-binding modules (CBMs) from the CBM32 (1) and CBM50 (4) families, as well as 40 glycosyl hydrolases (GHs) from the GH23 (15), GH19 (8), GH18 (5), GH25 (1) and GH108 (7) families (Figure 5A and Table S16). Most of these CAZymes (30) were obtained from the metagenomic MAVGs of *E. pulchripes*. Among them, 10 GHs (GH18 and GH19) were found to potentially target chitin, while 32 other CAZymes were associated with chitin or peptidoglycan as their potential substrates. CAZymes exclusively involved in peptidoglycan degradation were

excluded since they are viral hallmark genes. In addition, AMGs identified in carbohydrate metabolism included 2,3-bisphosphoglycerate-dependent phosphoglycerate mutase (*gpmB*), lactoylglutathione lyase (*gloA*), phospholipase C (*plc*) and 2-oxoglutarate ferredoxin oxidoreductase subunit delta (*korD*). Among these, *gpmB*, *gloA* and *plc* AMGs were derived from metagenomic MAVGs of *E. pulchripes* (Figure 5B), while the *korD* AMG belonged to metagenomic MAVGs of *G. connexa* (Figure 5C).

The most abundant AMGs were primarily involved in amino acid metabolism. Among them, the *dcm*, *DNMT3A*, and *metK* AMGs, which are involved in cysteine and methionine metabolism, were particularly abundant in metagenomic MAVGs from *E. pulchripes*. Only a few of these AMGs were found in the metatranscriptomic MAVGs of both millipede species (Figure 5B,C). The identification of four *phnP* indicated that these genes might contribute to phosphate



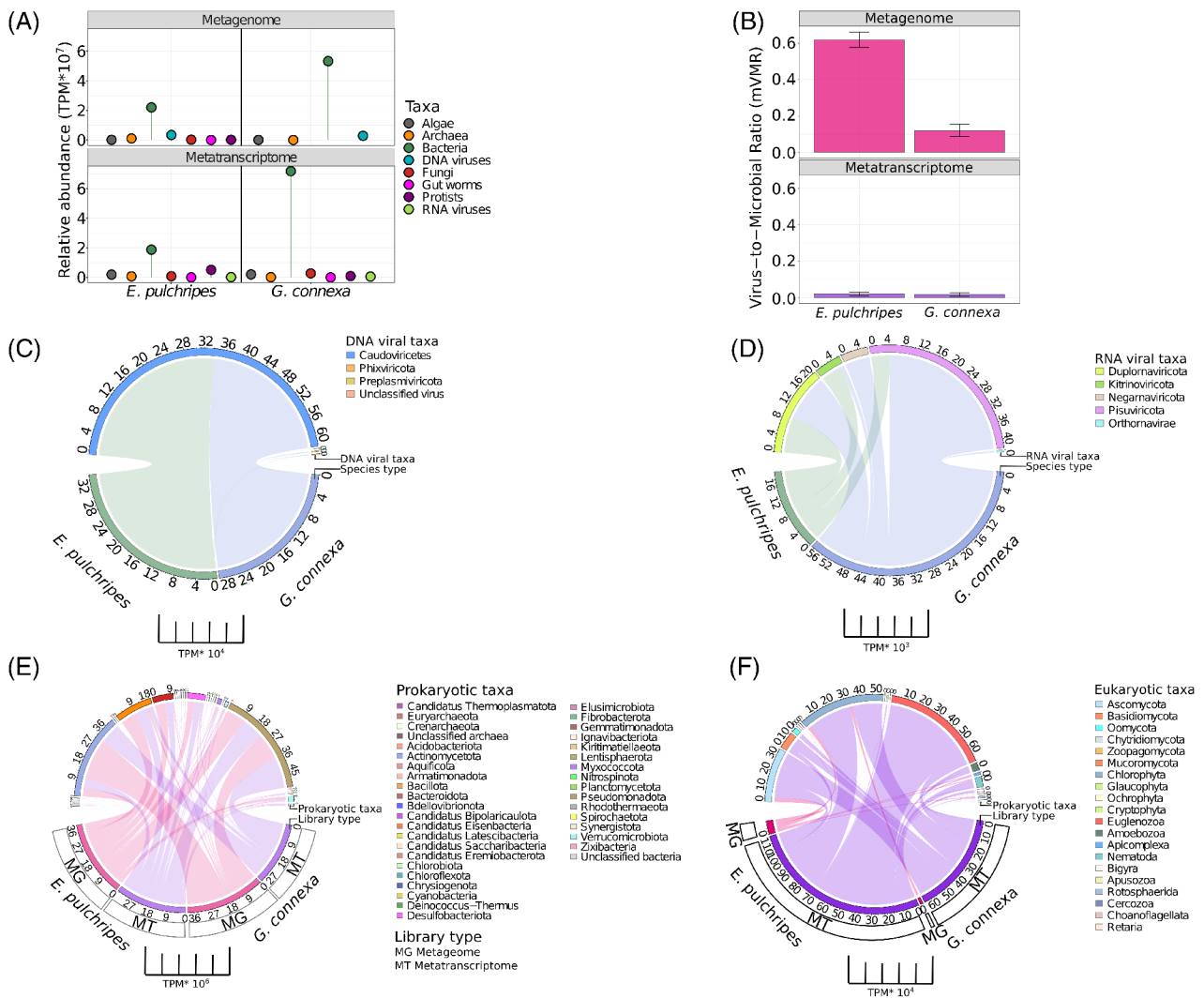


**FIGURE 5** Distribution of virus auxiliary metabolic genes (AMGs) in the metagenome-assembled viral genomes (MAVGs) from the hindguts of *Epibolus pulchripes* and *Glomeris connexa*. (A) The AMGs in the form of Carbohydrate-Active enZYmes (CAZymes) recovered in the metagenome and metatranscriptome from *E. pulchripes* and *G. connexa*. These CAZymes were annotated using the dbCAN3 server. The AMGs involved in various metabolisms and biosynthesis recovered in the metagenome and metatranscriptome from *E. pulchripes* (B) and *G. connexa* (C). (D) Representative arrow maps of the detected AMGs from MAVGs recovered. Annotation was done using geNomad and VIBRANT.

solubilisation by their host. The presence of seven distinct genes for the metabolism of cofactors and vitamins indicated the potential role of viruses in the biosynthesis of cobalamin (*cobS*, *cobT* and *bchE*), folate (*folA*) and niacin (*nadE*, *nadV* and *pncA*). Additionally, we also identified two AMGs that are likely involved in the sulphur relay system (*mec* and *moeB*). Figure 5D, Supplementary Figures 8–10 depict the genomic arrangement of the detected AMGs in different MAVGs.

## DNA and RNA virus-microbe ratio

The estimated viral-microbial ratio (mVMR) in both types of libraries was low for the two millipede species, ranging from 0.120 to 0.619 for DNA viruses and only 0.0163 to 0.0198 for RNA viruses (Figure 6A,B, Tables S19 and S20). When the relative abundance of USCGs was averaged at a lower taxonomic level, the estimated mVMR decreased by approximately twofold. As expected, most of the MCPs in the metagenome



**FIGURE 6** Virus-to-microbe ratio and the frequency of viral taxa and major microbes from the hindguts of *Epibolus pulchripes* and *Glomeris connexa*. (A) Normalised counts (TPM) of DNA viruses (MCPs), RNA viruses (RdRPs) and cells (USCGs) from the cellular domains (bacteria, archaea, fungi, algae, protists and gut worms). (B) The virus-to-microbe ratio was calculated by averaging the identified USCGs at the class level. (C) Taxa of DNA viruses based on phylogenetic assignment of MCPs. (D) Taxa of RNA viruses based on the phylogenetic assignment of RdRPs. (E) Bacteria cells based on the phylogenetic assignment of USCGs. (F) Eukaryotic cells based on the phylogenetic assignment of USCGs. Normalised counts (TPM) of DNA viruses (MCPs) were obtained from the metagenome, while RNA viruses (RdRPs) were obtained from the metatranscriptome. Cells (USCGs) were obtained from both the metagenomic and metatranscriptomic libraries to normalise with MCPs and RdRPs, respectively.

were classified as Caudoviricetes in both millipedes (Figure 6C). In contrast, the RdRP for RNA viruses showed higher diversity, with Duplornaviricota dominating *E. pulchripes* and Pisuviricota dominating *G. connexa* (Figure 6D). The bacterial taxa identified through USCGs were in agreement with the microbial composition obtained from the MAGs (Figure 6E). Archaea (Figure 6E) and microbial eukaryotes were also present (Figure 6F).

## DISCUSSION

Previous research on millipedes concentrated primarily on RNA viruses and on ssDNA viruses of the

Cressdnaviricota phylum (Li, Shi et al., 2015; Kirsch et al., 2021; Shi et al., 2016), partly due to the potential role of arthropods as carriers of plant and animal viruses. Conversely, phages, which constitute the large majority of DNA viruses in many environments (Kirsch et al., 2021), remain unexplored in millipedes. Based on MAVGs, we report on both DNA and RNA viruses and provide a detailed analysis of phage-host interactions and their suspected impact on millipede gut functioning.

Overall, we assembled over five times more MAVGs from *E. pulchripes* metagenomes than from *G. connexa*, although we encountered challenges in assembling complete viral genomes. This was not surprising, considering that the bacterial load and diversity

in *E. pulchripes* are much higher (Koubová et al., 2023). In contrast, the number of RNA MAVGs in both millipede species was on par. Overall, the differences in viral diversity between the two millipede species reflected the bacterial diversity found in each host species and matched it well according to host and AMG predictions in the viral MAVG dataset, and CRISPR-Cas classifications in the bacterial MAG dataset. Naturally, at this point, it cannot be excluded that some of those differences reflect the hosts' rearing conditions or stochastic colonisation processes rather than being part of the species' core gut virome.

Largely consistent with our classification of MCP sequences, >90% of DNA MAVGs belonged to the Caudoviricetes (previously Caudovirales). These viruses can be lytic or lysogenic and make up more than 85% of sequences in public genomic databases, dominating most metaviromes, including those found in termite guts (Dion et al., 2020; Li, Shi, et al., 2015; Kirsch et al., 2021). Existing reference datasets mainly contain members of the Caudoviricetes with genome sizes of 16–500 kb. This stems from their predominance in nature but also from previous recommendations to filter out genomes smaller than 10 kb (Eren et al., 2021). Other frequently overlooked viruses that have smaller genomes: Microviridae (starting from 4.4 kb; Ackermann, 1998), Leviviricetes (from 3.7 kb; Tikhe & Husseneder, 2018), Circoviridae (from 1.7 kb; Kirsch et al., 2021), Totiviridae (from 4.6 kb; Ghabrial et al., 2015), Potyviridae (from 8 kb; Pasin et al., 2022). These viruses were prominent members of the viral community in our dataset. Previous studies demonstrated that current procedures in viral ecology are biased towards Caudoviricetes and that size thresholds below 10 kb might be needed for the discovery of previously overlooked viruses (e.g., Kauffman et al., 2018; Roux, Krupovic, et al., 2019).

Moreover, similar to termite guts (Marynowska et al., 2020), we found that nearly all CRISPR sequences originated from Caudoviricetes, further demonstrating their dominance in this system. Based on this and the general importance of bacteria in millipede guts, we assume these viruses are the main viral players impacting the millipede gut functioning. However, we also recovered Microviridae (ssDNA) and Tecviridae (linear dsDNA) MAVGs from metagenomes of both millipede species. While generally low in abundance based on RdRP estimates, we found evidence that positive-sense RNA viruses of the class Leviviricetes play a role in *G. connexa*. Before 2021, Leviviricetes contained only four viruses, and these MAVGs represent a valuable addition to sequence repositories (Callanan et al., 2021). Concerning viruses infecting eukaryotes, it is most likely that fungal and protozoan viruses have the potential to influence millipede gut functioning. Interestingly, in *G. connexa*, the most prominent DNA MAVG belongs to the circular,

Rep-encoding single-stranded (CRESS) DNA viruses of the order Cirlivirales, which contains two families: Circoviridae, known to infect animals, and Vilyaviridae infecting *Giardia* (Krupovic et al., 2020). In *G. connexa*, this MAVG had about 150 times higher TPM value than the most abundant Caudoviricetes MAVG. At the level of RNA viruses and based on RdRP genes, viruses of the Partitiviridae, Endornaviridae and Totiviridae constituted the most abundant members of the RNA virus community, potentially infecting fungi and protozoa. These viral families were also well represented among the RNA MAVGs. Nweze et al. (2024) postulated that fungal biomass plays an important role in the diet of the millipede and its microbiome. Additional support for this comes from the high fraction of AMGs predicted to encode glycoside hydrolases involved in chitin degradation. Potyviridae (class Pisuviricota) encompasses over 30% of known plant viruses, many of which hold significant agricultural importance (Pasin et al., 2022; Riechmann et al., 1992). Whether they are resident members of the gut virome, or simply transiently introduced with the ingested plant material, remains to be studied. However, their presence in both species suggests that millipedes might be important vectors for plant viruses.

Except for two MAVGs, all predicted AMGs in this study were derived from Caudoviricetes. When actively reproducing, these phages are hypothesised to affect the metabolic functioning of their hosts. To which degree this is happening remains difficult to predict. The main hosts of phage MAVGs belong to the Bacteroidota and Pseudomonadota, which have recently been identified as the primary bacterial agents of complex polysaccharide degradation in *E. pulchripes* and *G. connexa* (Nweze et al., 2024; Winter et al., 2010). Our low mVMR estimates, combined with the prevalence of lysogenic viruses and the finding of a high proportion of type I and a low proportion of type III CRISPR-Cas systems, especially in *E. pulchripes*, suggests that the millipede guts follow a PtW dynamic (Goldberg et al., 2018; Nobrega et al., 2020; Rollie et al., 2020; Touchon et al., 2016; Watson et al., 2021). This would bring active viral reproduction and lysis (and thus AMG activity) under the governance of yet insufficiently understood controls of lysogenic to lytic switching (Howard-Varona et al., 2017; Knowles et al., 2016). However, the ecological framework of phage life strategies, like KtW and PtW, is still debated. The transition between strategies depends on factors such as nutrient availability, host physiology and phage type (Zhang et al., 2017).

Lysogenic phages can have subtle yet significant impacts on millipede hosts, affecting their immune responses, nutrition and microbial communities (Keen & Dantas, 2018). These phages can make non-pathogenic bacteria virulent (Wagner & Waldor, 2002), potentially influencing the millipede's immunity. Altered



bacterial metabolism due to lysogeny can indirectly affect millipede nutrition. Additionally, lysogenic phages can modify the gut microbiota, impacting food digestion. Some lysogenic phages carry resistance genes, potentially safeguarding millipedes from harmful pathogens and contributing to their overall health. Six out of the 10 most abundant Caudoviricetes DNA MAVGs in *E. pulchripes* were predicted to be lysogenic and, according to our metatranscriptomic data, were not actively reproducing at a high level. In *G. connexa*, Caudoviricetes DNA MAVGs were much less prominent (ca. 10 times less abundant than their counterparts in *E. pulchripes*) but, following our metatranscriptomic data, were actively reproducing at a high rate. The observed prophages from MAGs were predominantly defective or incomplete. While incomplete prophages may not enter the lytic cycle, potentially rendering their host susceptible to competition from related strains for space and nutrients, they can still provide crucial remnant genetic material to the host (Nepal et al., 2022). Nevertheless, it is important to acknowledge that the incompleteness of our MAVGs could have impacted our predictions regarding viral lifestyles and may have limited our ability to detect lysogenic genes.

Based on AMGs identified in this study, we hypothesise that during active replication, viruses in the hindgut of millipedes influence important metabolic functionalities, including chitin degradation, vitamin biosynthesis, amino acid and sulphur metabolism. However, the lower number of AMGs due to the absence of viral hallmark or viral-like genes, particularly in *G. connexa*, might stem from the incompleteness of our MAVGs.

Not surprisingly, the composition of bacteria harbouring CRISPR-Cas systems reflected the general microbial composition in the guts. Similar results were observed in an in-depth survey of the CRISPR-Cas systems of the human microbiome (Münch et al., 2021). Moreover, as in termite guts (Marynowska et al., 2020), we found that nearly all CRISPR sequences originated from Caudoviricetes, further demonstrating their dominance in this and other arthropod systems. Almost only CRISPR-Cas subsystems effective against DNA viruses (namely, I, II) were detected. While these are typically the more dominant ones in nature (Watson et al., 2021), the higher-than-expected proportion of subtype II CRISPR-Cas is in agreement with the recent claim regarding their prevalence in host-associated systems (Weissman et al., 2019).

## CONCLUSIONS

Our study represents the first comprehensive investigation of DNA and RNA viral communities in the hindguts of millipedes. Despite their importance, arthropods are

still understudied concerning their virome and bear considerable potential for discovering novel viral lineages. Many of the ecological functions of arthropods, detritivores, and others depend on their microbiome. Learning how the virome can modulate the microbial composition in arthropod guts may help us uncover the potential role of viruses in biogeochemical cycling. Earlier research (Hunter & Fusco, 2022) suggested that lysogenic phages can protect their bacterial hosts from other phages through superinfection immunity. Since millipedes inhabit soil, their gut bacteria likely encounter various environmental phages. It would be intriguing to explore whether millipede gut phages offer superinfection immunity to safeguard the gut microbiota from environmental bacteriophages. In the future, millipedes may serve as a model system to investigate the interplay between bacteria, phages and intestinal protozoa in detritivores.

## AUTHOR CONTRIBUTIONS

**Julius Eyiuche Nweze:** Conceptualization; methodology; software; data curation; formal analysis; writing – original draft; writing – review and editing; visualization. **Johannes Sergej Schweichhart:** Methodology; software; writing – review and editing; writing – original draft. **Roey Angel:** Conceptualization; methodology; investigation; supervision; funding acquisition; writing – original draft; writing – review and editing; project administration; resources; software.

## ACKNOWLEDGEMENTS

JEN and RA were supported by a Junior Grant from the Czech Science Foundation (GA ČR), grant number 19-24309Y. JSS was supported by a Standard Grant from the Czech Science Foundation (GA ČR), grant number 21-04987S. Open access publishing facilitated by Biologické centrum Akademie věd České republiky, as part of the Wiley - CzechELib agreement.

## CONFLICT OF INTEREST STATEMENT

The authors declare that they have no competing interests.

## DATA AVAILABILITY STATEMENT

The data and scripts that support the findings of this study are openly available on GitHub at <https://github.com/ISBB-anaerobic/Millipede-gut-viral-community-analysis>, and in Zenodo: <https://zenodo.org/doi/10.5281/zenodo.8328030>. We have also made available the FASTA files for the co-assembled metagenomes and metatranscriptomes from *Epibolus pulchripes* (<https://doi.org/10.6084/m9.figshare.22337500.v1> and <https://doi.org/10.6084/m9.figshare.22339522.v1>).

## ORCID

Roey Angel  <https://orcid.org/0000-0002-2804-9661>



## REFERENCES

- Abascal, F., Zardoya, R. & Posada, D. (2005) ProtTest: selection of best-fit models of protein evolution. *Bioinformatics*, 21, 2104–2105.
- Ackermann, H.-W. (1998) Tailed bacteriophages: the order Caudovirales. *Advances in Virus Research*, 51, 135–201.
- Bhunchoth, A., Blanc-Mathieu, R., Mihara, T., Nishimura, Y., Askora, A., Phironrit, N. et al. (2016) Two asian jumbo phages,  $\square$ RSL2 and  $\square$ RSF1, infect *Ralstonia solanacearum* and show common features of  $\square$ KZ-related phages. *Virology*, 494, 56–66.
- Byzov, B.A. (2006) Intestinal microbiota of millipedes. In: König, H. & Varma, A. (Eds.) *Intestinal microorganisms of termites and other invertebrates*. Berlin/Heidelberg: Springer-Verlag, pp. 89–114.
- Callanan, J., Stockdale, S.R., Adriaenssens, E.M., Kuhn, J.H., Rumnieks, J., Pallen, M.J. et al. (2021) Leviviricetes: expanding and restructuring the taxonomy of bacteria-infecting single-stranded RNA viruses. *Microbial Genomics*, 7, 000686.
- Camargo, A.P., Roux, S., Schulz, F., Babinski, M., Xu, Y., Hu, B. et al. (2023) Identification of mobile genetic elements with geNomad. *Nature Biotechnology*, 1–10.
- Capella-Gutiérrez, S., Silla-Martínez, J.M. & Gabaldón, T. (2009) trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*, 25, 1972–1973.
- Chaumeil, P.-A., Mussig, A.J., Hugenholz, P. & Parks, D.H. (2020) GTDB-Tk: a toolkit to classify genomes with the genome taxonomy database. *Bioinformatics*, 36, 1925–1927.
- Chen, L.-X., Méheust, R., Crits-Christoph, A., McMahon, K.D., Nelson, T.C., Slater, G.F. et al. (2020) Large freshwater phages with the potential to augment aerobic methane oxidation. *Nature Microbiology*, 5, 1504–1515.
- Coutinho, F.H., Gregoracci, G.B., Walter, J.M., Thompson, C.C. & Thompson, F.L. (2018) Metagenomics sheds light on the ecology of marine microbes and their viruses. *Trends in Microbiology*, 26, 955–965.
- Crawford, C.S. (1992) Millipedes as model detritivores. *Berichte des Naturwissenschaftlich-Medizinischen Vereins in Innsbruck*, 10, 277–288.
- Danovaro, R. & Middelboe, M. (2010) *Separation of free virus particles from sediments in aquatic systems*. Waco, TX: The American Society of Limnology and Oceanography, pp. 74–81. Manual of aquatic viral ecology (MAVE).
- Díaz-Muñoz, S.L. & Koskella, B. (2014) Bacteria–phage interactions in natural environments. *Advances in Applied Microbiology*, 89, 135–183.
- Dion, M.B., Oechslin, F. & Moineau, S. (2020) Phage diversity, genomics and phylogeny. *Nature Reviews Microbiology*, 18, 125–138.
- Emerson, J.B., Roux, S., Brum, J.R., Bolduc, B., Woodcroft, B.J., Jang, H.B. et al. (2018) Host-linked soil viral ecology along a permafrost thaw gradient. *Nature Microbiology*, 3, 870–880.
- Enghoff, H. (2011) East African giant millipedes of the tribe Pachybolini (Diplopoda, Spirobolida, Pachybolidae). *Zootaxa*, 2753, 1–41.
- Eren, A.M., Kiefl, E., Shaiber, A., Veseli, I., Miller, S.E., Schechter, M. S. et al. (2021) Community-led, integrated, reproducible multi-omics with anvio. *Nature Microbiology*, 6, 3–6.
- Gan, R., Zhou, F., Si, Y., Yang, H., Chen, C., Ren, C. et al. (2022) DBSCAN-SWA: an integrated tool for rapid prophage detection and annotation. *Frontiers in Genetics*, 13, 885048.
- Ghabrial, S.A., Castón, J.R., Jiang, D., Nibert, M.L. & Suzuki, N. (2015) 50-plus years of fungal viruses. *Virology*, 479–480, 356–368.
- Goldberg, G.W., McMillan, E.A., Varble, A., Modell, J.W., Samai, P., Jiang, W. et al. (2018) Incomplete prophage tolerance by type III-A CRISPR-Cas systems reduces the fitness of lysogenic hosts. *Nature Communications*, 9, 61.
- Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D. A., Amit, I. et al. (2011) Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. *Nature Biotechnology*, 29, 644–652.
- Guo, J., Bolduc, B., Zayed, A.A., Varsani, A., Dominguez-Huerta, G., Delmont TO et al. (2021) VirSorter2: a multi-classifier, expert-guided approach to detect diverse DNA and RNA viruses. *Microbiome*, 9, 1–13.
- Hille, F. & Charpentier, E. (2016) CRISPR-Cas: biology, mechanisms and relevance. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, 371, 20150496.
- Holmes, E.C. (2011) What does virus evolution tell us about virus origins? *Journal of Virology*, 85, 5247–5251.
- Holmfeldt, K., Odić, D., Sullivan, M.B., Middelboe, M. & Riemann, L. (2012) Cultivated single-stranded DNA phages that infect marine Bacteroidetes prove difficult to detect with DNA-binding stains. *Applied and Environmental Microbiology*, 78, 892–894.
- Howard-Varona, C., Hargreaves, K.R., Abedon, S.T. & Sullivan, M.B. (2017) Lysogeny in nature: mechanisms, impact and ecology of temperate phages. *The ISME Journal*, 11, 1511–1520.
- Huerta-Cepas, J., Serra, F. & Bork, P. (2016) ETE 3: reconstruction, analysis, and visualization of phylogenomic data. *Molecular Biology and Evolution*, 33, 1635–1638.
- Hunter, M. & Fusco, D. (2022) Superinfection exclusion: a viral strategy with short-term benefits and long-term drawbacks. *PLoS Computational Biology*, 18, e1010125.
- Joly, F.-X., Coq, S., Coulis, M., David, J.-F., Hättenschwiler, S., Mueller, C.W. et al. (2020) Detritivore conversion of litter into faeces accelerates organic matter turnover. *Communications Biology*, 3, 660.
- Kaletta, J., Pickl, C., Griebler, C., Klingl, A., Kurmayer, R. & Deng, L. (2020) A rigorous assessment and comparison of enumeration methods for environmental viruses. *Scientific Reports*, 10, 18625.
- Katoh, K. & Toh, H. (2008) Recent developments in the MAFFT multiple sequence alignment program. *Briefings in Bioinformatics*, 9, 286–298.
- Kauffman, K.M., Hussain, F.A., Yang, J., Arevalo, P., Brown, J.M., Chang, W.K. et al. (2018) A major lineage of non-tailed dsDNA viruses as unrecognized killers of marine bacteria. *Nature*, 554, 118–122.
- Keen, E.C. & Dantas, G. (2018) Close encounters of three kinds: bacteriophages, commensal bacteria, and host immunity. *Trends in Microbiology*, 26, 943–954.
- Kieft, K., Zhou, Z. & Anantharaman, K. (2020) VIBRANT: automated recovery, annotation and curation of microbial viruses, and evaluation of viral community function from genomic sequences. *Microbiome*, 8, 1–23.
- Kirsch, J.M., Brzozowski, R.S., Faith, D., Round, J.L., Secor, P.R., & Duerkop, B.A. (2021). Bacteriophage-bacteria interactions in the gut: from invertebrates to mammals. *Annual Review of Virology*, 8(1), 95–113.
- Knowles, B., Silveira, C., Bailey, B., Barott, K., Cantu, V., Cobián-Güemes, A. et al. (2016) Lytic to temperate switching of viral communities. *Nature*, 531, 466–470.
- Koonin, E.V. & Makarova, K.S. (2009) CRISPR-Cas: an adaptive immunity system in prokaryotes. *F1000 Biology Reports*, 1, 95.
- Koubová, A., Lorenc, F., Terézia Horváthová, A. & Chroňáková, V.Š. (2023) Millipede gut-derived microbes as a potential source of cellulolytic enzymes. *World Journal of Microbiology and Biotechnology*, 39, 169.
- Krupovic, M., Varsani, A., Kazlauskas, D., Breitbart, M., Delwart, E., Rosario, K. et al. (2020) Cressdnaviricota: a virus phylum unifying seven families of rep-encoding viruses with single-stranded, circular DNA genomes. *Journal of Virology*, 94, e00582-20. Available from: <https://doi.org/10.1128/jvi.00582-20>
- Larralde, M. (2022) Pyrodigal: python bindings and interface to prodigal, an efficient method for gene prediction in prokaryotes. *Journal of Open Source Software*, 7, 4296.
- Lefkowitz, E.J., Dempsey, D.M., Hendrickson, R.C., Orton, R.J., Siddell, S.G. & Smith, D.B. (2018) Virus taxonomy: the database

- of the international committee on taxonomy of viruses (ICTV). *Nucleic Acids Research*, 46, D708–D717.
- Li, C.-X., Shi, M., Tian, J.-H., Lin, X.-D., Kang, Y.-J., Chen, L.-J. et al. (2015) Unprecedented genomic diversity of RNA viruses in arthropods reveals the ancestry of negative-sense RNA viruses. Goff SP, editor. *eLife*, 4, e05378.
- Li, D., Liu, C.-M., Luo, R., Sadakane, K. & Lam, T.-W. (2015) MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics*, 31, 1674–1676.
- Li, H. & Durbin, R. (2010) Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics*, 26, 589–595.
- López-García, P., Gutiérrez-Preciado, A., Krupovic, M., Ciobanu, M., Deschamps, P., Jardillier, L. et al. (2023) Metagenome-derived virus-microbe ratios across ecosystems. *The ISME Journal*, 1–12, 1552–1563.
- Makarova, K.S., Wolf, Y.I. & Koonin, E.V. (2018) Classification and nomenclature of CRISPR–Cas systems: where from here? *The CRISPR Journal*, 1, 325–336.
- Marynowska, M., Goux, X., Sillam-Dussès, D., Rouland-Lefèvre, C., Halder, R., Wilmes, P. et al. (2020) Compositional and functional characterisation of biomass-degrading microbial communities in guts of plant fibre- and soil-feeding higher termites. *Microbiome*, 8, 96.
- Mihara, T., Nishimura, Y., Shimizu, Y., Nishiyama, H., Yoshikawa, G., Uehara, H. et al. (2016) Linking virus genomes with host taxonomy. *Viruses*, 8, 66.
- Mirzaei, M.K. & Maurice, C.F. (2017) Ménage à trois in the human gut: interactions between host, bacteria and phages. *Nature Reviews Microbiology*, 15, 397–408.
- Münch, P.C., Franzosa, E.A., Stecher, B., McHardy, A.C. & Huttenhower, C. (2021) Identification of natural CRISPR systems and targets in the human microbiome. *Cell Host & Microbe*, 29, 94–106.e4.
- Nayfach, S., Camargo, A.P., Schulz, F., Eloe-Fadrosh, E., Roux, S. & Kyrpides, N.C. (2021) CheckV assesses the quality and completeness of metagenome-assembled viral genomes. *Nature Biotechnology*, 39, 578–585.
- Nepal, R., Houtak, G., Wormald, P.-J., Psaltis, A.J. & Vreugde, S. (2022) Prophage: a crucial catalyst in infectious disease modulation. *The Lancet Microbe*, 3, e162–e163.
- Nishimura, Y., Yoshida, T., Kuronishi, M., Uehara, H., Ogata, H. & Goto, S. (2017) ViPTree: the viral proteomic tree server. *Bioinformatics*, 33, 2379–2380.
- Nobrega, F.L., Walinga, H., Dutilh, B.E. & Brouns, S.J.J. (2020) Prophages are associated with extensive CRISPR–Cas auto-immunity. *Nucleic Acids Research*, 48, 12074–12084.
- Nweze, J.E., Šustr, V., Brune, A. & Angel, R. (2024) Functional similarity despite taxonomical divergence in the millipede gut microbiota points to a common trophic strategy. *Microbiome*, 12, 16.
- Parikka, K.J., Le Romancer, M., Wauters, N. & Jacquet, S. (2017) Deciphering the virus-to-prokaryote ratio (VPR): insights into virus–host relationships in a variety of ecosystems. *Biological Reviews*, 92, 1081–1100.
- Pasin, F., Daròs, J.-A. & Tzanetakis, I.E. (2022) Proteome expansion in the Potyviridae evolutionary radiation. *FEMS Microbiology Reviews*, 46, fuac011.
- Puxty, R.J., Evans, D.J., Millard, A.D. & Scanlan, D.J. (2018) Energy limitation of cyanophage development: implications for marine carbon cycling. *The ISME Journal*, 12, 1273–1286.
- Quistad, S.D., Grasis, J.A., Barr, J.J. & Rohwer, F.L. (2017) Viruses and the origin of microbiome selection and immunity. *The ISME Journal*, 11, 835–840.
- Riechmann, J.L., Lain, S. & García, J.A. (1992) Highlights and prospects of potyvirus molecular biology. *Journal of General Virology*, 73, 1–16.
- Rohwer, F. & Edwards, R. (2002) The phage proteomic tree: a genome-based taxonomy for phage. *Journal of Bacteriology*, 184, 4529–4535.
- Rollie, C., Chevallereau, A., Watson, B.N.J., Chyou, T., Fradet, O., McLeod, I. et al. (2020) Targeting of temperate phages drives loss of type I CRISPR–Cas systems. *Nature*, 578, 149–153.
- Roux, S., Adriaenssens, E.M., Dutilh, B.E., Koonin, E.V., Kropinski, A.M., Krupovic, M. et al. (2019) Minimum information about an uncultivated virus genome (MIUViG). *Nature Biotechnology*, 37, 29–37.
- Roux, S., Camargo, A.P., Coutinho, F.H., Dabdoub, S.M., Dutilh, B.E., Nayfach, S. et al. (2023) iPHoP: an integrated machine learning framework to maximize host prediction for metagenome-derived viruses of archaea and bacteria. *PLoS Biology*, 21, e3002083.
- Roux, S., Krupovic, M., Daly, R.A., Borges, A.L., Nayfach, S., Schulz, F. et al. (2019) Cryptic inoviruses revealed as pervasive in bacteria and archaea across Earth’s biomes. *Nature Microbiology*, 4, 1895–1906.
- Russel, J., Pinilla-Redondo, R., Mayo-Muñoz, D., Shah, S.A. & Sørensen, S.J. (2020) CRISPRCasTyper: automated identification, annotation, and classification of CRISPR–Cas loci. *The CRISPR Journal*, 3, 462–469.
- Sakaguchi, S., Urayama, S., Takaki, Y., Hirotsuna, K., Wu, H., Suzuki, Y. et al. (2022) NeoRdRp: a comprehensive dataset for identifying RNA-dependent RNA polymerases of various RNA viruses from metatranscriptomic data. *Microbes and Environments*, 37, ME22001.
- Sardar, P., Šustr, V., Chroňáková, A. & Lorenc, F. (2022) Metatranscriptomic holobiont analysis of carbohydrate-active enzymes in the millipede *Telodeinopus aoutii* (Diplopoda, Spirostreptida). *Frontiers in Ecology and Evolution*, 10, 10.
- Sardar, P., Šustr, V., Chroňáková, A., Lorenc, F. & Faktorová, L. (2022) *De novo* metatranscriptomic exploration of gene function in the millipede holobiont. *Scientific Reports*, 12, 16173.
- Schliep, K.P. (2011) phangorn: phylogenetic analysis in R. *Bioinformatics*, 27, 592–593.
- Shang, J., Tang, X. & Sun, Y. (2023) PhaTYP: predicting the lifestyle for bacteriophages using BERT. *Briefings in Bioinformatics*, 24, bbac487.
- Shi, M., Lin, X.D., Tian, J.H., Chen, L.J., Chen, X., Li, C.X. et al. (2016) Redefining the invertebrate RNA virosphere. *Nature*, 540, 539–543.
- Shkoporov, A.N. & Hill, C. (2019) Bacteriophages of the human gut: the “known unknown” of the microbiome. *Cell Host & Microbe*, 25, 195–209.
- Silveira, C. & Rohwer, F. (2016) Piggyback-the-winner in host-associated microbial communities. *npj Biofilms and Microbiomes*, 2, 16010. Available from: <https://doi.org/10.1038/npjbiofilms.2016.10>
- Simmonds, P., Adriaenssens, E.M., Zerbini, F.M., Abrescia, N.G., Aiewsakun, P., Alfenas-Zerbini, P. et al. (2023) Four principles to establish a universal virus taxonomy. *PLoS Biology*, 21, e3001922.
- Stamatakis, A. (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30, 1312–1313.
- Steinberger, M. & Söding, J. (2017) MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nature Biotechnology*, 35, 1026–1028.
- Taylor, E.C. (1982) Role of aerobic microbial populations in cellulose digestion by desert millipedes. *Applied and Environmental Microbiology*, 41, 281–291.
- Thingstad, T.F. (2000) Elements of a theory for the mechanisms controlling abundance, diversity, and biogeochemical role of lytic bacterial viruses in aquatic systems. *Limnology and Oceanography*, 45, 1320–1328.
- Tikhe, C.V. & Husseneder, C. (2018) Metavirome sequencing of the termite gut reveals the presence of an unexplored bacteriophage community. *Frontiers in Microbiology*, 8, 2548.
- Touchon, M., Bernheim, A. & Rocha, E.P. (2016) Genetic and life-history traits associated with the distribution of prophages in bacteria. *The ISME Journal*, 10, 2744–2754.

- Wagner, P.L. & Waldor, M.K. (2002) Bacteriophage control of bacterial virulence. *Infection and Immunity*, 70, 3985–3993.
- Wang, M., Fu, S., Xu, H., Wang, M. & Shi, L. (2018) College of Environment and Planning, Henan University, Kaifeng, Henan 475004. Ecological functions of millipedes in the terrestrial ecosystem. *Biodiversity Science*, 26, 1051–1059.
- Watson, B.N., Steens, J.A., Staals, R.H., Westra, E.R. & van Houte, S. (2021) Coevolution between bacterial CRISPR-Cas systems and their bacteriophages. *Cell Host & Microbe*, 29, 715–725.
- Weissman, J.L., Laljani, R.M.R., Fagan, W.F. & Johnson, P.L.F. (2019) Visualization and prediction of CRISPR incidence in microbial trait-space to identify drivers of antiviral immune strategy. *The ISME Journal*, 13, 2589–2602.
- Wheeler, T.J. & Eddy, S.R. (2013) Nhmmer: DNA homology search with profile HMMs. *Bioinformatics*, 29, 2487–2489.
- Wickham, H. (2016) *ggplot2: elegant graphics for data analysis*. New York: Springer-Verlag.
- Wilkins, D. (2023) gggenes: draw gene arrow maps in 'ggplot2.' R package version 0. 5. 1.
- Winter, C., Bouvier, T., Weinbauer, M.G. & Thingstad, T.F. (2010) Trade-offs between competition and defense specialists among unicellular planktonic organisms: the “killing the winner” hypothesis revisited. *Microbiology and Molecular Biology Reviews*, 74, 42–57.
- Zhang, J., Gao, Q., Zhang, Q., Wang, T., Yue, H., Wu, L. et al. (2017) Bacteriophage–prokaryote dynamics and interaction within anaerobic digestion processes across time and space. *Microbiome*, 5, 5.
- Zhang, Y.-Z., Shi, M. & Holmes, E.C. (2018) Using metagenomics to characterize an expanding virosphere. *Cell*, 172, 1168–1172.

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Nweze, J.E., Schweichhart, J.S. & Angel, R. (2024) Viral communities in millipede guts: Insights into the diversity and potential role in modulating the microbiome. *Environmental Microbiology*, 26(2), e16586. Available from: <https://doi.org/10.1111/1462-2920.16586>



# **Viral communities in millipede guts: Insights into diversity and the potential role in modulating the microbiome**

**Julius E. Nweze<sup>1,2</sup>, Johannes Schweichhart<sup>1,2</sup>, Roey Angel<sup>1\*</sup>**

<sup>1</sup>Institute of Soil Biology and Biogeochemistry, Biology Centre CAS, České Budějovice, Czechia

<sup>2</sup>Department of Ecosystem Biology, Faculty of Science, University of South Bohemia in České Budějovice, Czechia

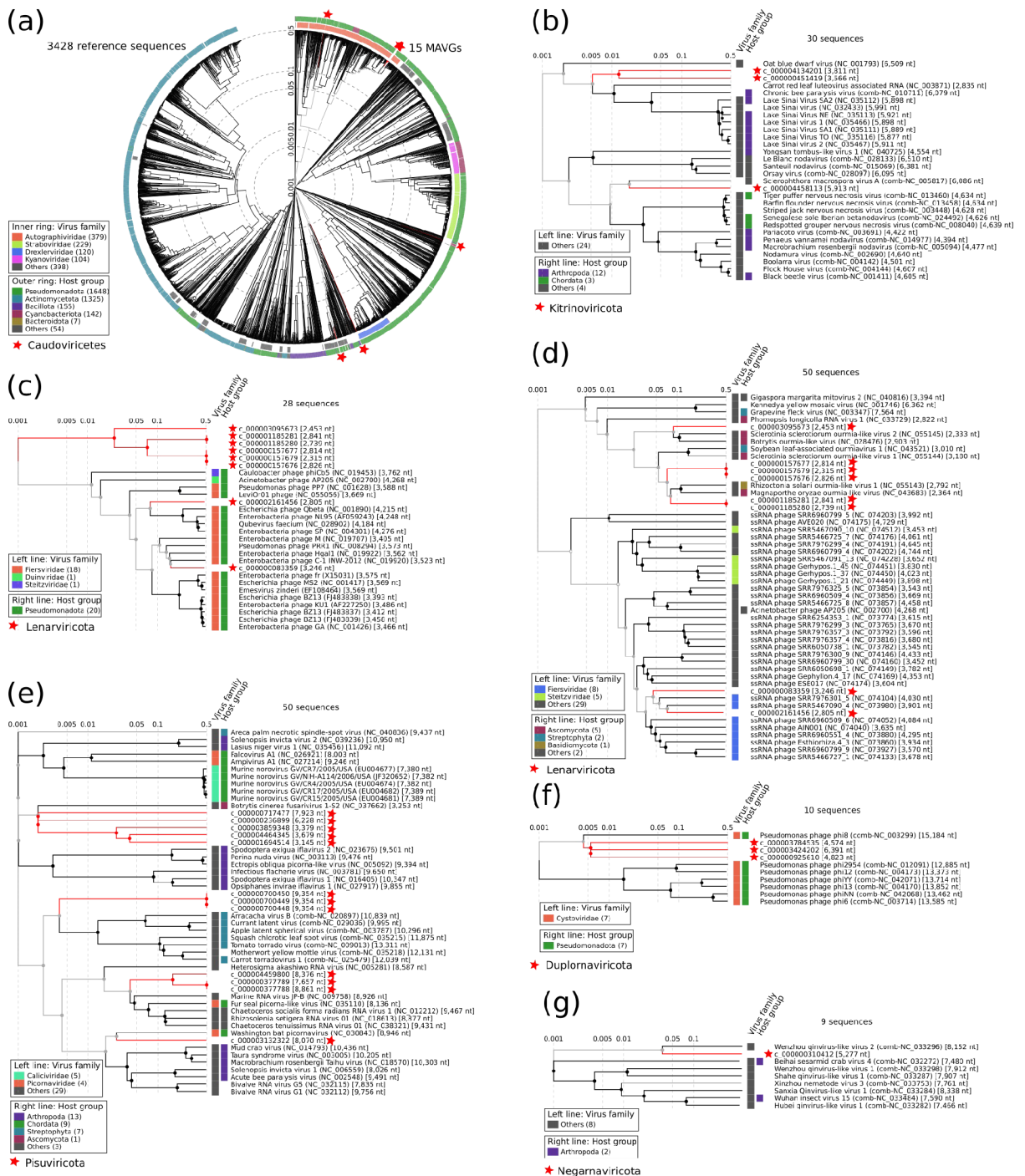
\* Correspondence: roey.angel@bc.cas.cz

## **Supplementary Material**

### **Supplementary Tables**

See file: Nweze\_et\_al\_Environmental\_Microbiology.xlsx



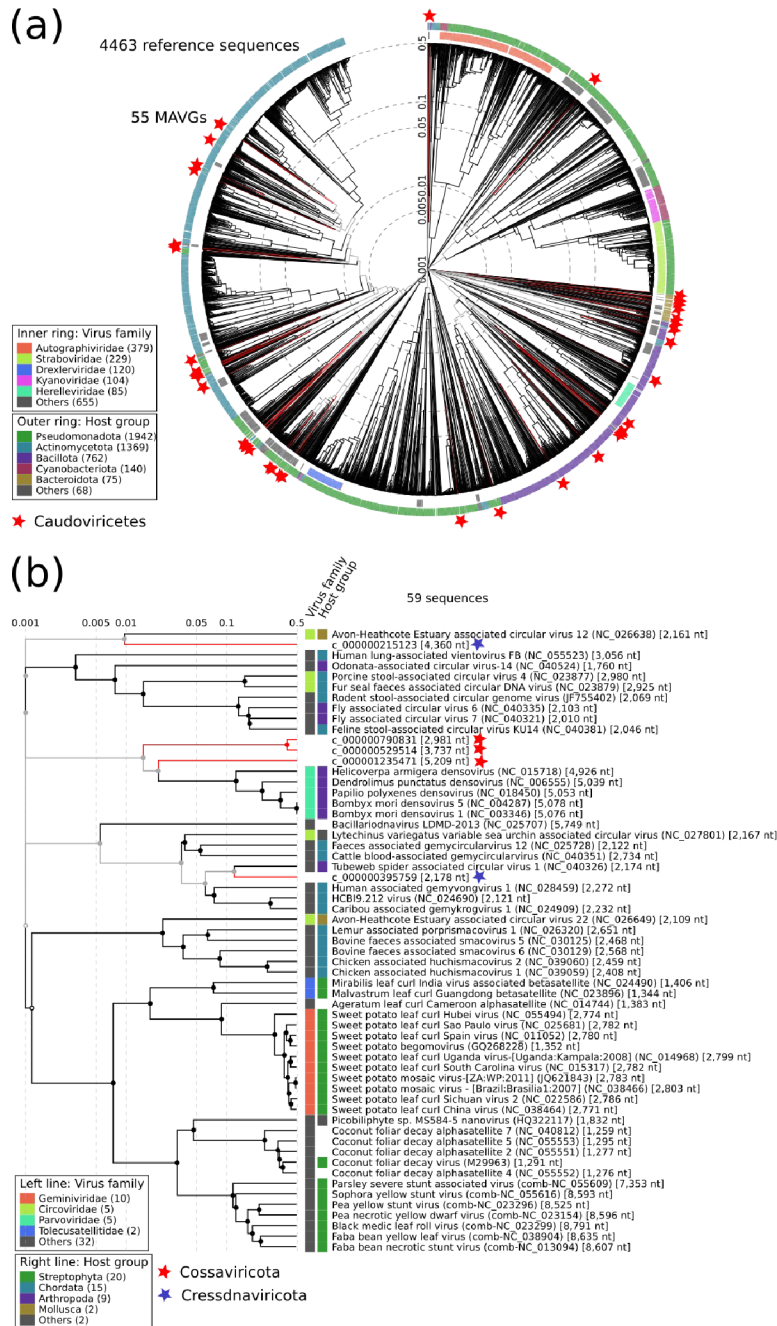


**Supplementary Fig. 1. Proteomic tree of viral genomes in metatranscriptome from *E. pulchripes*.**

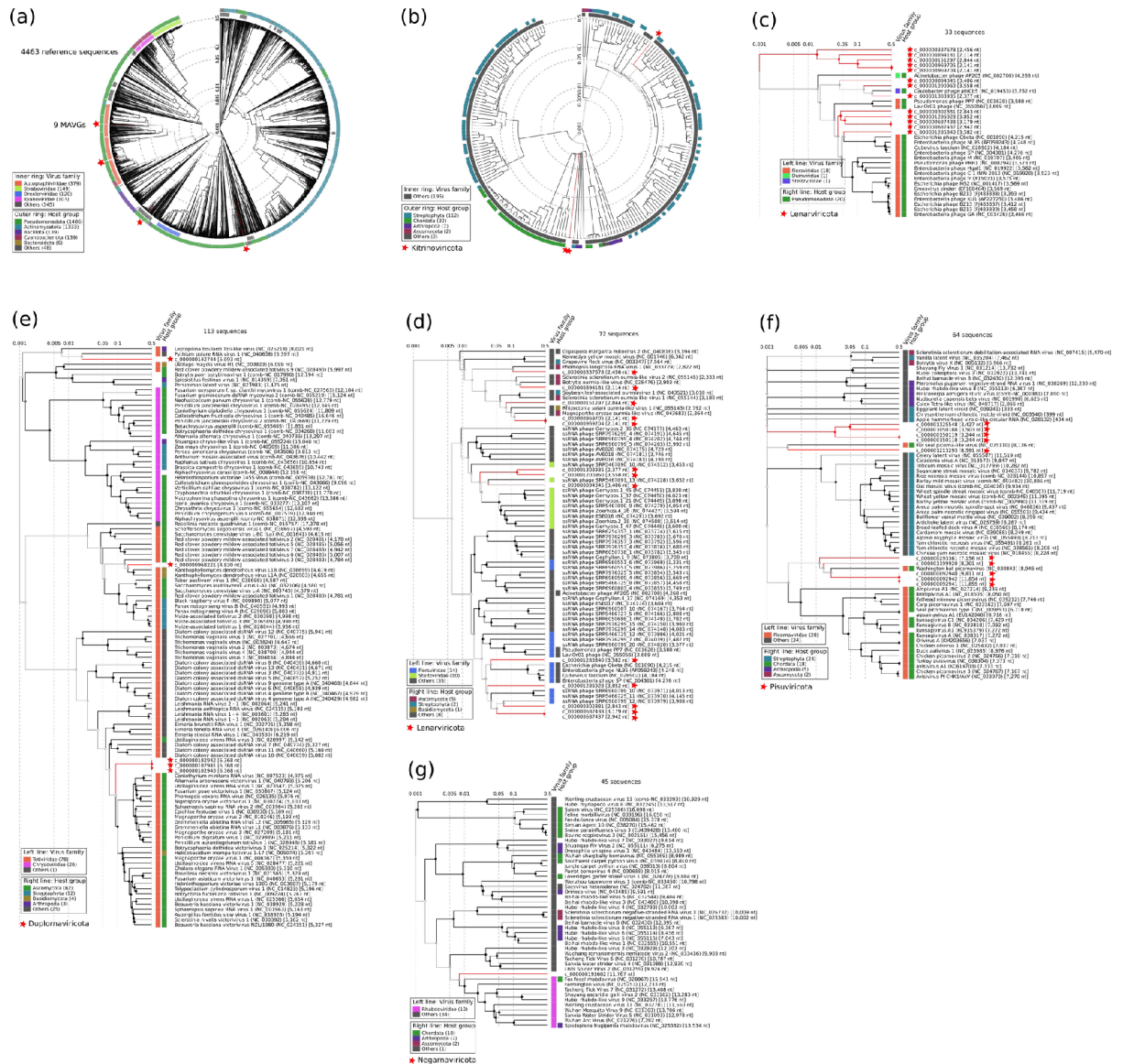
(a) Proteomic similarity in the dsDNA viruses between the metagenome-assembled viral genomes (MAVGs; 15) classified in the phylum *Uroviricota* (order *Caudoviricetes*) and related reference viral genomes. (b) Genome similarity in the ssRNA viruses between the MAVGs classified in the phylum *Ktirnoviricota* (3) and related reference viral genomes. (c) Genome

similarity in the ssRNA viruses between the MAVGs classified in the phylum *Lenarviricota* (8), prokaryotic and (d) eukaryotic reference viral genomes. (e) Genome similarity in the ssRNA viruses between the MAVGs classified in the phylum *Pisuviricota* (12) and related eukaryotic reference viral genomes. (f) Genome similarity in the dsRNA viruses between the MAVGs classified in the phylum *Duplornaviricota* (3) and related prokaryotic reference viral genomes. (g) Genome similarity in the dsRNA viruses between the MAVGs classified in the phylum *Negarnaviricota* (1) and related eukaryotic reference viral genomes. See **Fig. 2** and the **Materials and Methods** for how the tree was generated. The red stars represent each MAVG from our metatranscriptome.





**Supplementary Fig. 2. Proteomic tree of viral genomes in metagenome from *G. connexa*.** (a) Genomic similarity in the dsDNA viruses between the metagenome-assembled viral genomes (55 MAVGs) classified in the phylum *Uroviricota* (class *Caudoviricetes*) and 4463 related reference viral genomes. (b) Genomic similarity in the ssDNA viruses between the MAVGs classified in the phyla *Cossaviricota* and *Cressdnaviricota*, and related reference viral genomes. See **Fig. 2** and the **Materials and Methods** for how the tree was generated. The red stars represent each MAVG from our metagenome.



**Supplementary Fig. 3. Proteomic tree of viral genomes in the metatranscriptome from *G. connexa*.** (a) Genomic similarity in the dsDNA viruses between the MAVGs classified in the phylum *Uroviricota* (9) and 4463 related reference viral genomes. (b) Genome similarity in the ssRNA viruses between the MAVGs classified in the phylum *Kitriroviricota* (3) and 203 related reference viral genomes. (c) Genome similarity in the ssRNA viruses between the MAVGs classified in the phylum *Lenarviricota* (13) and 20 prokaryotic and (d) 59 eukaryotic reference viral genomes. (e) Genome similarity in the ssRNA viruses between the MAVGs classified in the phylum *Duplornaviricota* (5) and 108 related eukaryotic reference viral genomes. (f) Genome similarity in

the dsRNA viruses between the MAVGs classified in the phylum *Pisuviricota* (10) and 54 related eukaryotic reference viral genomes. (g) Genome similarity in the dsRNA viruses between the MAVGs classified in the phylum *Negarnaviricota* (1) and 44 related eukaryotic reference viral genomes. See **Fig. 2** and the Materials and Methods for how the tree was generated. The red stars represent each MAVG from our metatranscriptome.

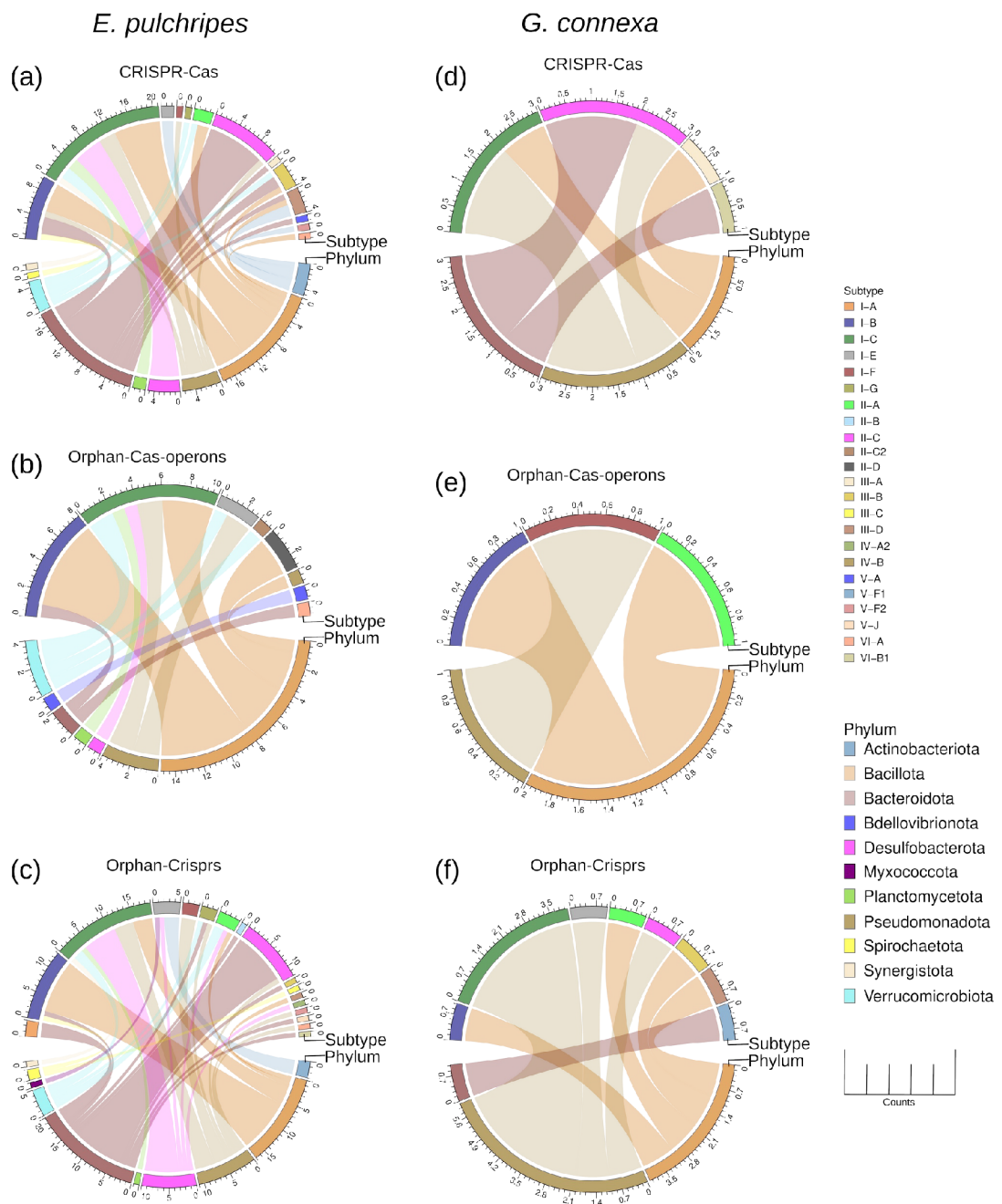


**Supplementary Fig. 4 CRISPR-Cas loci and the associated genes in the metagenomic and metatranscriptomic libraries from the hindguts of *E. pulchripes* and *G. connexa*.** (a) Total number of CRISPR-Cas loci, orphan Cas-genes and CRISPR arrays in both library types from the two millipede species. (b) Number of different type of CRISPR-Cas systems from both library from *E. pulchripes* and *G. connexa*. (c) The number of subtypes of CRISPR-Cas loci in both library types (lacking in the metatranscriptome from *G. connexa*). (d) The number of orphan Cas-operons (those not present in CRISPR-Cas loci) and (e) orphan CRISPR arrays in both library type.



**Supplementary Fig. 5. Arrangements and directions of the CRISPR-Cas systems identified in the metagenomes and metatranscriptomes from *E. pulchripes* and *G. connexa*.** (a) The CRISPR-Cas systems in the metagenome and (b) metatranscriptome from *E. pulchripes*. (c) The CRISPR-Cas systems in the metagenome from *G. connexa*.

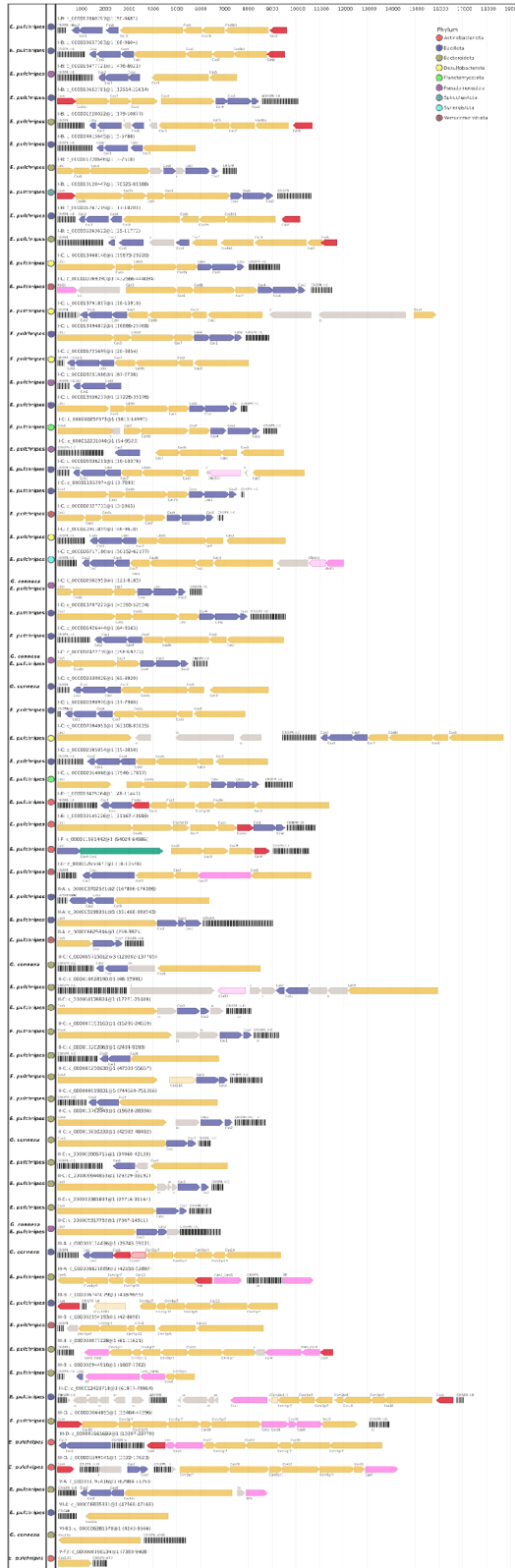




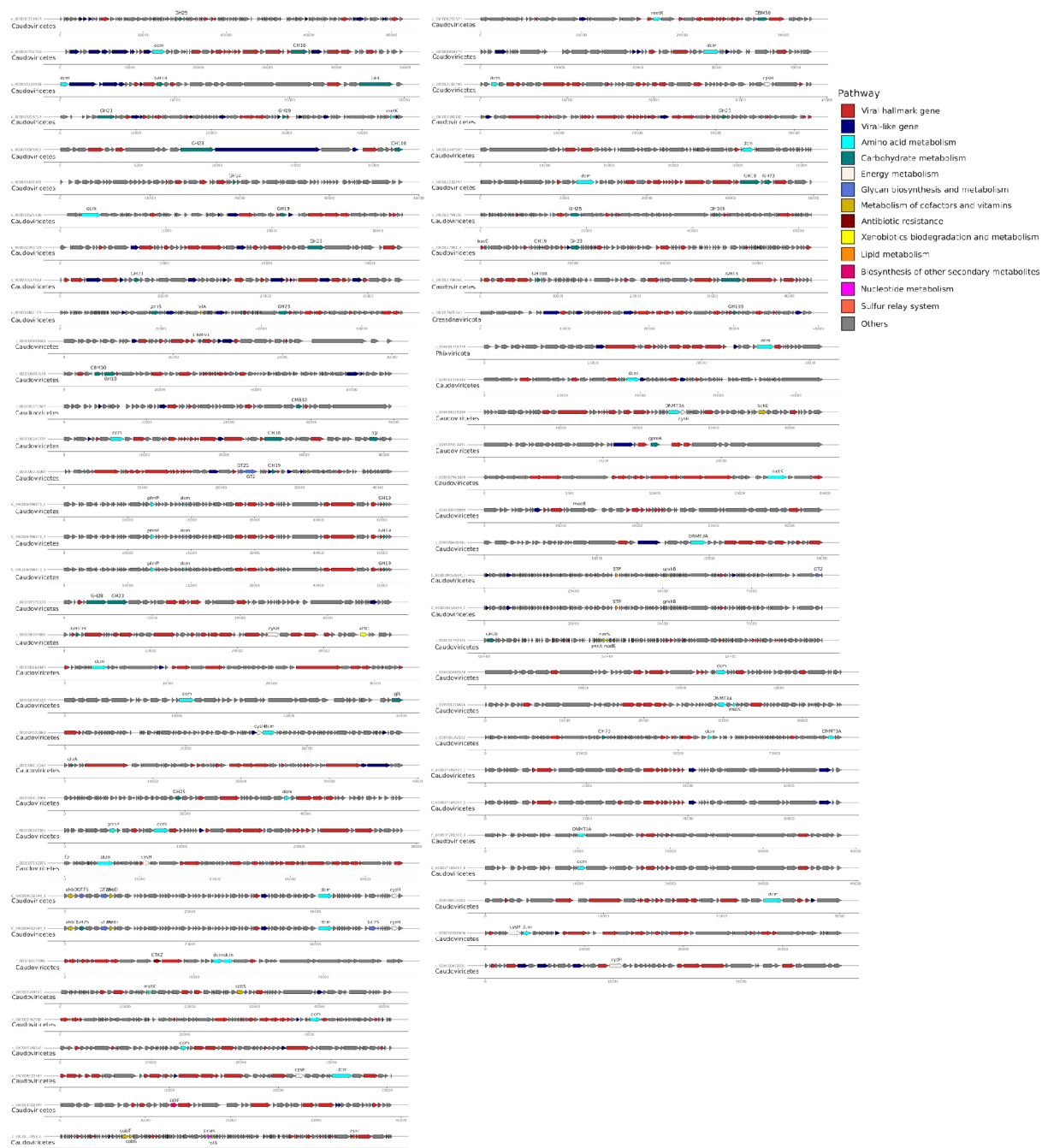
**Supplementary Fig. 6. Phylogenetic origins of CRISPR-Cas systems and associated genes in the MAGs from *E. pulchripes* and *G. connexa*.** (a) Total number of CRISPR-Cas loci, orphan Cas genes and CRISPR arrays in MAGs from *E. pulchripes* with colour-coded phyla. Different subtypes of (b) CRISPR-Cas loci (c) orphan Cas genes (d) orphan CRISPR arrays found in MAGs from *E. pulchripes*. (e) Total number of CRISPR-Cas loci, orphan Cas genes and CRISPR arrays in MAGs from *G. connexa*. Different subtypes of (f) CRISPR-Cas loci (g) orphan Cas genes (h) orphan

CRISPR arrays found in MAGs from *G. connexa*. Top arch - CRISPR-Cas subtypes. Bottom arch - phyla.

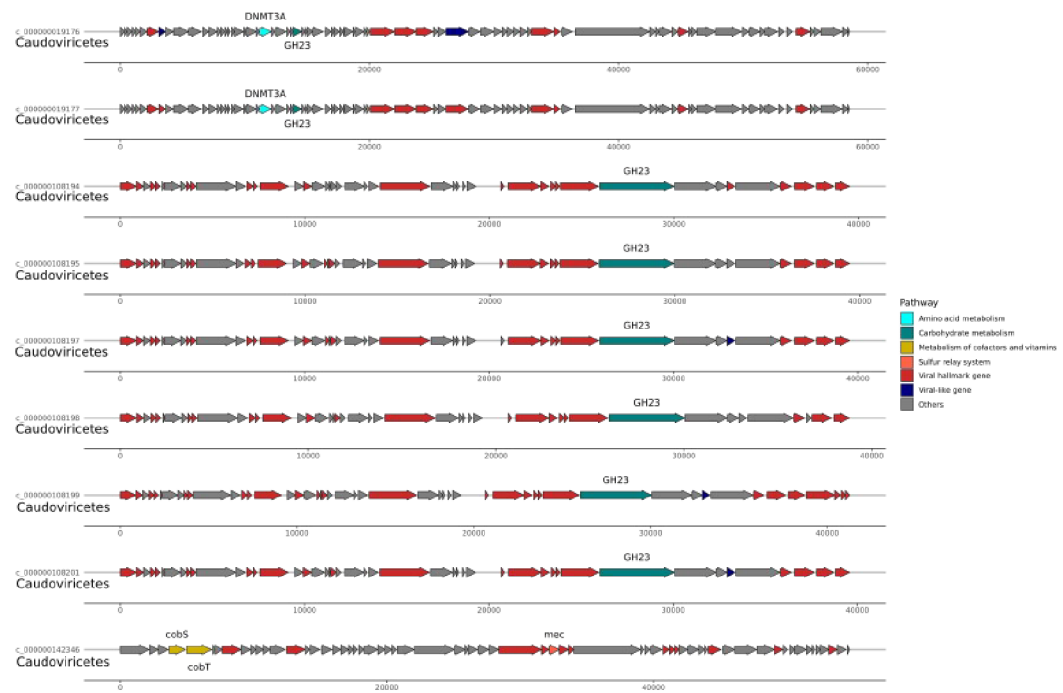




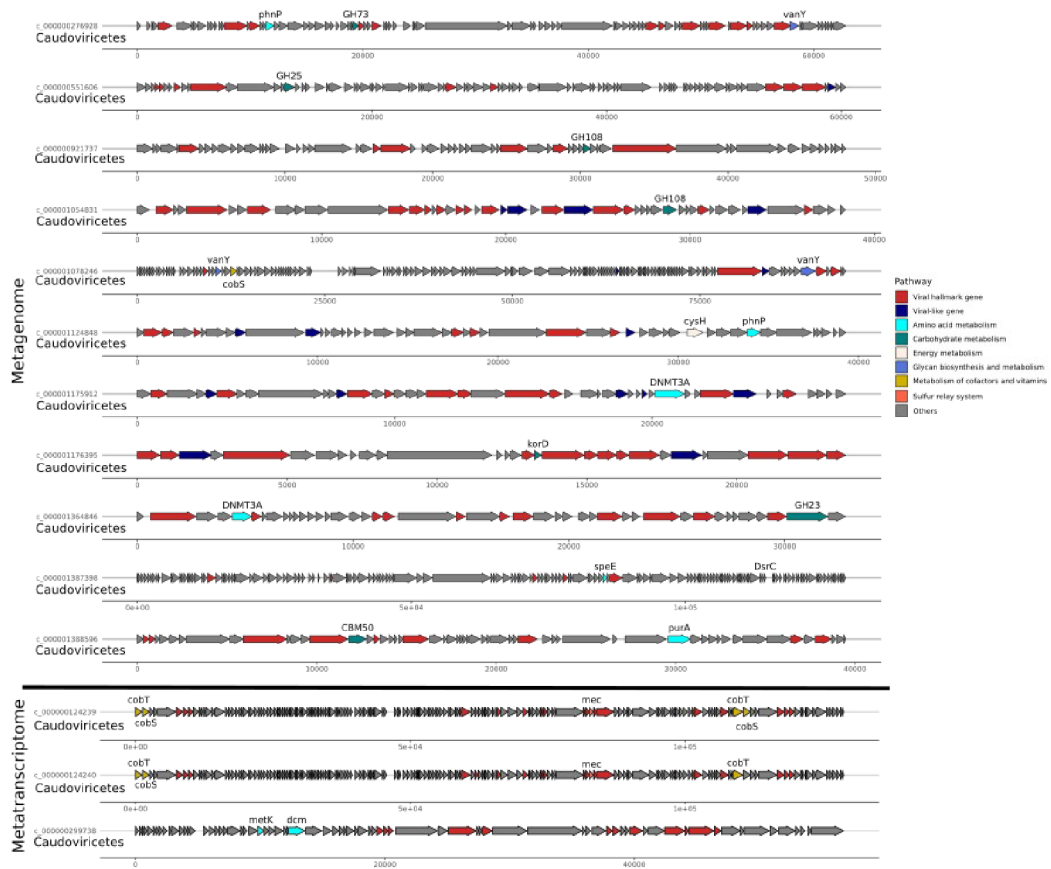
**Supplementary Fig. 7. Arrangements and directions of the CRISPR-Cas systems identified in the MAGs from *E. pulchripes* and *G. connexa*.** Each CRISPR-Cas system identified in our study was labelled with the specific phylum of the metagenome-assembled genomes (MAGs) from which it originated in each millipede species.



**Supplementary Fig. 8. Arrow maps of the detected virus auxiliary metabolic genes (AMGs) from the metagenome-assembled viral genomes (MAVGs) recovered in metagenome from the hindgut of *E. pulchripes*.**



**Supplementary Fig. 9. Arrow maps of the detected virus auxiliary metabolic genes (AMGs) from metagenome-assembled viral genomes (MAVGs) recovered in metatranscriptome from the hindgut of *E. pulchripes*.**



**Supplementary Fig. 10. Arrow maps of the detected virus auxiliary metabolic genes (AMGs) from metagenome-assembled viral genomes (MAVGs) recovered in metagenome and metatranscriptome from the hindgut of *G. connexa*.**

### **5.3 Manuscript No. 3**

**Julius Eyiuche Nweze**, Shruti Gupta, Michaela Salcher, Vladimír Šustr, Terézia Horváthová, Roey Angel (2024). Cellulose fermentation by the gut microbiome is likely not essential for the nutrition of millipedes (Manuscript).





# Cellulose fermentation by the gut microbiota is likely not essential for the nutrition of millipedes

Julius Eyiuche Nweze<sup>1,2</sup>, Shruti Gupta<sup>1</sup>, Michaela Salcher<sup>3</sup>, Vladimír Šustr<sup>1</sup>, Terézia Horváthová<sup>1</sup>, Roey Angel<sup>1,2\*</sup>

<sup>1</sup>Institute of Soil Biology and Biogeochemistry, Biology Centre CAS, České Budějovice, Czechia

<sup>2</sup>Faculty of Science, University of South Bohemia in České Budějovice, Czechia

<sup>3</sup>Institute of Hydrobiology, Biology Centre CAS, České Budějovice, Czechia

\* Correspondence: Roey Angel; Institute of Soil Biology and Biogeochemistry, Biology Centre CAS, České Budějovice, Czechia; +420 387775847; roey.angel@bc.cas.cz

## Abstract

Millipedes are believed to rely on their gut microbiome to process plant-litter-cellulose via fermentation, like many other arthropods. However, this belief needs more evidence. To examine this, we disrupted the gut microbiota of juvenile *Epibolus pulchripes* (tropical, methane-emitting) and *Glomeris connexa* (temperate, non-methane-emitting) using inhibitors and isotopic labelling. Feeding them sterile or antibiotics-treated litter notably reduced faecal production and microbial load without major impacts on survival or weight. Bacterial diversity stayed similar, with *Bacteriodota* dominant in *E. pulchripes* and *Pseudomonadota* in *G. connexa*. Sodium-2-bromoethanesulfonate treatment halted CH<sub>4</sub> emissions in *E. pulchripes* after 14 days, but it resumed after returning to normal feeding. Employing <sup>13</sup>C-labelled leaf litter and RNA-SIP showed a slow and gradual prokaryote labelling,

indicating a significant density shift only by day 21. Surprisingly, labelling of the fungal biomass was somewhat quicker. Our results suggest that the gut microbiota might not be essential for cellulose digestion.

## **Introduction**

Like most animals, invertebrates build complex partnerships with diverse microbial communities (Petersen & Osvatic 2018), contributing to their evolutionary and ecological success (Moran *et al.* 2019). The host-associated microorganisms serve multiple roles, such as supporting the host's nutritional needs (Douglas 2009), influencing sexual development (Perlmutter & Bordenstein 2020), and modulating the immune system (Hurst & Darby 2009). This concept gave rise to the notion of animals as "holobionts", where the host and its microbiota are considered a single ecological entity (Bordenstein & Theis 2015; Zilber-Rosenberg & Rosenberg 2008). Recent studies on microbiomes support the widespread prevalence of microbial partnerships across the animal kingdom (Russell *et al.* 2014; Vavre & Kremer 2014).

While most invertebrates have microbial associations, their reliance on them varies widely. Termites, for instance, depend entirely on their gut microbiota for nutrition (Brune 2014). Conversely, many other arthropods may lack a resident gut microbiota and develop fully even germ-free (Hammer *et al.* 2019). Most arthropods generally fall between these extremes, relying on their microbiota for some form of support (e.g. cockroaches (Mikaelyan *et al.* 2016; Tinker & Ottesen 2016) or isopods (Bouchon *et al.* 2016; Mattila *et al.* 2014). Detritivorous and xylophagous animals typically rely on gut microorganisms for cellulose digestion. Although animal cellulases are found in some gut systems (Watanabe & Tokuda 2001), (ligno)cellulolytic bacteria and fungi are

generally deemed necessary for hydrolysis and fermentation, releasing short-chain fatty acids, which get absorbed by the host (Schmidt & Engel 2021).

Millipedes (Diplopoda) are important detritivores widely distributed and abundant in many temperate and tropical ecosystems (Kime & Golovatch 2000). Despite their status as keystone species in tropical and temperate forests (Crawford 1992), they have been understudied compared to other detritivores, with little focus on their microbiome. Because of the nutrient-poor nature of plant litter, millipedes contend with low assimilation efficiencies through high ingestion rates to compensate (David 2014). Like other arthropods, millipedes host diverse gut microorganisms (Byzov 2006). In particular, the central hindgut was shown to host the highest density of microorganisms, which attach to its cuticle, while the foregut and midgut contain mostly transient inhabitants (Nardi *et al.* 2016). Various studies suggest that certain millipede gut bacteria possess enzymes for breaking down plant polysaccharides (Alagesan 2003; Koubová *et al.* 2023; Nweze *et al.* 2024; Ramanathan & Alagesan 2012; Sardar *et al.* 2022a; Taylor 1982). If millipedes rely on cellulose for their nutrition, then fermentation followed by methanogenesis should occur extensively in their guts, similar to ruminants or wood-feeding termites (Brune 2014). However, methanogenesis has only been observed in some millipede species, but not others, and its occurrence correlated to the millipede's size (Šustr *et al.* 2014a). Despite these findings, direct proof of gut microorganisms supporting the millipede's nutritional needs has not been demonstrated. An alternative hypothesis suggests millipedes foster microbial growth in litter, potentially digesting the resulting fungal and bacterial biomass (Bignell 1989).

To assess the role of the millipede gut microbiota, we conducted experiments using two model species: the CH<sub>4</sub>-emitting *Epibolus pulchripes* (Spirobolida) and the non-CH<sub>4</sub>-emitting *Glomeris connexa* (Glomerida), which do not emit

CH<sub>4</sub>. *E. pulchripes* is a large millipede (130–160 mm) common along the East African coast (Enghoff 2011), while *G. connexa* is smaller (10-17 mm) and native to Central Europe (Hoess & Scholl 2001). We examined the effects of inhibitors on body weight, survival, faecal bacterial load, gut bacterial composition, and CH<sub>4</sub> production. Additionally, we identified metabolically active hindgut prokaryotes using <sup>13</sup>C-RNA-SIP.

## **Materials and Methods**

### **Animal collection and maintenance**

We used juvenile *E. pulchripes* from our lab breeding colony and wild-caught *G. connexa* from Czechia (forest locality of Helfenburk near Bavorova; 49°8'10.32"N, 14°0'24.21"E). No specific permit was required for the collection. Species identification relied on morphological features (Gerstaecker 1873; Kocourek *et al.* 2017); data not shown). Before use, the animals were kept in the lab for several weeks. Both species were housed in perforated plastic terraria, filled with commercial sand as a substrate, broken terracotta pots for shelter, and locally collected or purchased (see below) Canadian poplar (*Populus x canadensis*) leaf litter (see below). Moisture was maintained by spraying with tap water every other day. Both species experienced a 12-hour photoperiod. *E. pulchripes* was housed individually in a box (19.3 x 13.8 x 5 cm) at 25 °C and in a climate-controlled room. Meanwhile, five *G. connexa* individuals were kept in each box (15 x 10 x 4 cm) in an incubator (TERMOBOX LBT 165, Vanellus s.r.o.) at a temperature of 15 °C.

### **Antibiotic curing**

Each millipede species comprised 40 individuals split into four groups of ten: Control, Sterile, diluted antibiotics (2X-Diluted in *E. pulchripes* and 5X-Diluted

in *G. connexa*) and undiluted antibiotics (Undiluted in *E. pulchripes* and 2X-Diluted in *G. connexa*). Briefly, the Control group was fed untreated, senesced leaves, the Sterile group was fed autoclaved leaves, and the antibiotics-treated groups were fed autoclaved leaves treated with antibiotics. *E. pulchripes* groups were fed around 2.4 g of litter, while *G. connexa* groups received 0.5 g. Just before feeding, the leaf litter was sprayed with 500  $\mu\text{l}$  of tap water (Control), sterile distilled water (Sterile), or antibiotics solution containing penicillin G: 10,000 units  $\text{ml}^{-1}$ , streptomycin sulfate: 10  $\mu\text{g ml}^{-1}$  and amphotericin B: 25  $\mu\text{g ml}^{-1}$  (Thermo Fisher Scientific), following Zimmer and Bartholme (2003). The terraria, sand, and litter were replaced every 7 days to maintain hygiene.

The animal fitness was followed for 42 days by aseptically measuring their weights on a Sartorius digital scale with an accuracy of 0.01g. During feeding, three pellets of fresh faeces (0.15–0.19 g for *E. pulchripes* and 0.01–0.02 g for *G. connexa*) were sampled from the millipede boxes, suspended in phosphate buffer (2mL; pH 7.4), plated in triplicates on Lysogeny broth (LB) agar plates and incubated at 25 °C. After 16 h, the colonies were counted and used to quantify the bacterial load. The remaining faecal material was counted and kept at -20 °C for DNA extraction. Methane emission was also monitored (see below).

## **Inhibition of methanogenesis**

Thirty *E. pulchripes* individuals were divided into three groups of ten. The Control group was fed on untreated litter, while the other two groups were fed litter treated with 5 mM (5mM-Na-BES) and 10 mM (10mM-Na-BES) of sodium 2-bromoethanesulfonate (Na-BES; Sigma-Aldrich) to inhibit methanogenesis. Moisture was maintained by spraying with sterile tap water or Na-BES solution every other day. The animals' weight and  $\text{CH}_4$  production

were regularly monitored for 64 days. Methane emission measurements were conducted by placing the millipedes in sealed glass bottles, with wet filter paper pieces, to maintain humidity (volume 130 ml for *E. pulchripes*; 30 ml for *G. connexa*; Thermo Fisher Scientific) for 4 h at 20 °C. The control was glass vessels without animals. Headspace samples (0.5 ml) were collected at the start and the end of incubation using a gas-tight syringe and analysed on a gas chromatograph (HP 5890 series II; Hewlett Packard, Palo Alto, CA, USA) equipped with a 2 m Porapak N column at 75 °C and an FID detector. The difference in CH<sub>4</sub> concentration between the two time points was used to calculate the production rate (in nmol mg<sup>-1</sup> d<sup>-1</sup>).

## **Identification and enumeration of protists and symbiotic methanogens**

Fourteen days post-CH<sub>4</sub>-inhibition, fresh *E. pulchripes* faecal pellets were crushed using a sterilised mortar and pestle, vortexed in 5 ml of 1X phosphate buffer saline (PBS) solution (pH 7.2), and then incubated at room temperature for 2–6 h to dissolve the aggregates. After spin-down, 2 µl of the supernatant was examined under a bright-field microscope using a Neubauer chamber (Sigma-Aldrich). Protists were identified and enumerated. Part of the supernatant was also fixed at 4 °C for 1.5 hours with 2% paraformaldehyde (PFA; Sigma-Aldrich), subjected to sequential vacuum filtration through 10 µm and 0.2 µm white polycarbonate filters (Sigma-Aldrich). These filters were air-dried and stored at -20 °C for Catalysed Reporter Deposition Fluorescence *in situ* Hybridization (CARD-FISH) analysis.

For CARD-FISH, specific HRP rRNA-targeting oligonucleotide probes were used (biomers.net). These included a universal probe for archaea (ARC915; Stahl 1991), *Methanobacteriales* (MB311; (Crocetti *et al.* 2006)) and



*Methanomascilliicoccales* (RC281r\_mod; modified from (Iino *et al.* 2013). A nonsense probe (NON-EUB338; (Wallner *et al.* 1993)) served as a negative control (Table S1). Probe coverage and specificity were assessed with TestProbe on ARB Silva (Quast *et al.* 2013). Hybridisation stringency was evaluated *in silico* using the Mathfish platform (Yilmaz *et al.*, 2011) and confirmed with optimised formamide concentration (Sigma-Aldrich; Table S1). The filters were prepared following the Piwosz *et al.* (Piwosz *et al.* 2021). See Supplementary material for further details. Filters were mounted on a glass slide and visualised using an OLYMPUS BX53 epifluorescence microscope (Olympus Optical Ltd.). Methanogens per ciliate were manually counted. Positive (using a general archaeal probe) and negative (no probe and a nonsense probe) control filters were also analysed.

## **Stable isotope labelling of RNA**

For the SIP experiment, three replicates from separate terraria were used for each species. *E. pulchripes* had one individual per replicate, while *G. connexa* had five to adjust for size differences. Millipedes were fed 99.9% <sup>13</sup>C-labelled Canadian-poplar leaves (IsoLife, Netherlands). Control groups were fed unlabelled leaves. Temperature (25 °C and 15 °C) and humidity (50-60%) were consistently maintained. Before the main experiment, a preliminary feeding test determined the ideal labelling duration and sampling intervals. Two individuals per species received 0.5 g and 0.05 g of labelled litter weekly for 14 days. Faecal samples were collected every 2 days for isotopic labelling analysis.

To quantify isotopic labelling before DNA sequencing, 1.9 g of faeces from each millipede species were dried in a SpeedVac DNA130 (Thermo Fisher Scientific) at 45 °C for 3 h. Dried samples were weighed, and 25 µg were transferred into triplicate tin capsules. Isotopic labelling was assessed at the

Stable Isotope Facility, Biology Centre CAS, using a Thermo Scientific™ 253 Plus™ 10 kV IRMS equipped with a SmartEA Isolink and GasBench II (Thermo Fisher Scientific). The  $^{13}\text{C}$  at% was calculated following Hayes (2004). Animals were sacrificed on days 3, 7, 14, and 21, dissected following Sardar et al. (2022b) and stored at  $-20\text{ }^{\circ}\text{C}$  for subsequent analysis.

## **Nucleic acid extraction and quantification**

DNA and RNA were immediately extracted from fresh hindgut and faeces samples, purified and quantified according to Angel *et al.* (2021). Hindgut samples from the SIP experiment measured 0.677–1.108 g for *E. pulchripes* and 0.083–0.092 g for *G. connexa*. Pooled faecal pellet samples from the antibiotics curing and inhibition of methanogenesis experiments were 0.43–0.59 g for *E. pulchripes* and 0.2–0.4 for *G. connexa*. See Supplementary material for further details.

## **Isopycnic ultra-centrifugation of $^{13}\text{C}$ labelled RNA**

Following RNA purification, density gradient centrifugation was performed in caesium trifluoroacetate (CsTFA) density gradients following a previously published protocol (Angel *et al.*, 2020). See Supplementary material for further details.

## **Gene quantification, amplicon library construction and sequencing**

DNA extracts from the antibiotics treatment experiment (24 samples per species) were subjected to 16S-rRNA-gene quantification using the QX200 AutoDG Droplet Digital PCR System (ddPCR; Bio-Rad), primers 338F—805R and the 516P FAM/BHQ1 probe (Yu *et al.* 2005). DNA extracts from the methanogenesis inhibition experiment were used for quantifying the *mcrA* gene as a marker for methanogens using primers *mlas\_mod* and *mcrA-rev*, according

to Angel *et al.* (Angel *et al.* 2011). Before sequencing, the cDNA from the SIP fractions (160 samples for each millipede species) was used for quantifying the 16S-rRNA of bacteria using the same method as mentioned above and the 18S copies of fungi using the FungiQuant system (Liu *et al.* 2012). For amplicon sequencing, the V4 region of the 16S rRNA gene was amplified and sequenced in a two-step protocol on an Illumina MiniSeq platform (2 × 250 cycle configuration; V2 reagent kit; Illumina) according to Naqib *et al.* (Naqib *et al.* 2019). PCR amplification was performed on 10 ng of DNA or 2 µl of cDNA with primers 515F\_mod and 806R (Walters *et al.* 2016), synthesised with the Fluidigm linkers CS1 and CS2 on their 5' end. Sequencing was performed at the DNA Services Facility at the University of Illinois, Chicago, USA.

## **Bioinformatic and statistical analyses**

Unless mentioned otherwise, all bioinformatic and statistical analyses were done in R V4.1.1 (RCore 2016). A linear mixed-effects model (Bates *et al.* 2015) was fitted to determine the effect of treatments and time on the millipede weight and microbial load. Differences between treatments in terms of total faecal pellet production, methane emission, mcrA and 16S rRNA copies were evaluated using an ANOVA model (Girden 1992) followed by Tukey's HSD test for pairwise comparisons (Keselman & Rogan 1977) or a linear regression model (Zou *et al.* 2003). Survival analysis of the animals was also computed using the Kaplan-Meier estimates (Goel *et al.* 2010).

Sequencing data was analysed as follows: primer and linker regions were removed from the raw amplicon reads using Cutadapt (V3.5; (Martin 2011)). The raw reads were processed, assembled and filtered using the R package DADA2 (V1.28) with the following non-standard filtering parameters: maxEE = c(2, 2) in the filterAndTrim function and pseudo pooling in the dada function

(Callahan *et al.* 2016). Chimaeras were removed with the `removeBimeraDenovo` option. The quality-filtered pair-end reads were classified to the genus level using SILVA (Quast *et al.* 2013), and those not classified as bacteria or archaea were filtered out. Heuristic decontamination was done using the `decontam` R package (Davis *et al.* 2018), and unique sequences were identified and clustered in an amplicon sequence variant (ASV) table. The resulting tables were imported into the R package `Phyloseq` (McMurdie & Holmes 2013). Read counts were normalised using median sequencing depth before plotting taxa abundance and after excluding ASVs without taxonomic assignments at the phylum level and those below a 5% prevalence threshold. Alpha diversity indices were computed using the `vegan` package on unfiltered and non-normalised data (Dixon 2003) and evaluated using the Kruskal-Wallis test (McKight & Najab 2010) and Dunn's test (Dinno & Dinno 2017). Corrections for multiple testing were made using the Benjamini-Hochberg (BH; (Benjamini & Hochberg 1995)) method. Values were compared and converted to a compact letter using the `cldList` function in the `rcompanion` package (Mangiafico & Mangiafico 2017). Beta diversity was calculated with a constrained analysis of principal coordinates (CAP; Anderson & Willis 2003). Lastly, a permutational multivariate ANOVA (Anderson 2001; function `vegan::adonis`) was conducted using the Bray-Curtis distance matrix and the `pairwise.adonis2` function (Martinez Arbizu 2020) to assess combined treatment and pairwise effects on the microbial community.

Differentially abundant genera were identified after sterile feeding or antibiotic treatment using ANCOM-BC2 (Lin & Peddada 2020). Before analysis, all ASVs not present in at least two samples or had an abundance of less than 2 were filtered. Only genera with adjusted P-values  $\leq 0.05$  were plotted.

Differentially abundant ASVs were subjected to a pseudo-count-addition sensitivity analysis.

Identification of isotopically labelled ASVs in the SIP experiment using differential abundance analysis followed Angel (Angel 2019). After decontaminating RNA-SIP reads, rare taxa (with <100 total reads, present in <2 fractions in a specific SIP gradient and its unlabelled counterpart). The DADA2 output sequences were aligned using mafft v7.505 (Kato *et al.* 2002), and a maximum-likelihood (ML) phylogenetic tree was constructed using IQ-TREE V2.1.1 (Minh *et al.* 2020) using the ‘-fast’ option. The 16S rRNA copies were plotted against the density ( $\text{g ml}^{-1}$ ) and used to calculate absolute ASV abundances. Fractions with densities  $>1.795 \text{ g ml}^{-1}$  (‘heavy’ fractions) from each labelled sample at each time point were compared against their unlabelled counterparts using DESeq2 V1.40.1 (Love *et al.* 2014), using the parametric fit type and the Wald significance test.  $\text{Log}_2$  fold change (LFC) shrinkage was applied using the function lfcShrink (Zhu *et al.* 2019), and the results were filtered to include only ASVs with a positive  $\text{log}_2$  fold change and a p-value  $<0.1$  (one-sided test).

## Results

### Antibiotic curing

Feeding millipedes with either sterile or treated feed (antibacterial and antifungal mixture) led to only negligible weight change in both species (Fig 1a; Table S2). In *E. pulchripes*, the control group showed a 5% increase over time, while the other treatments showed a 4-9% decrease in average weight with no significant trend. In contrast, *G. connexa* even showed a 3-8% increase in the treated groups but was also insignificant (Fig. 1b; Table S2). The treatment also did not significantly impact the millipedes' survival based on

Kaplan-Meier estimates (Fig. S1). Despite maintaining stable weight, faecal production decreased over time in response to antibiotics or sterile feed (Fig. 1c and d). While the number of faecal pellets generally declined in both species, there was a significant reduction under all treatments except for sterile feed in *G. connexa*. However, the differences between the treated groups were not statistically significant ( $P = <10^{-7} - 0.046$ ; Table S3).

Total faecal colony counts in both millipede species were consistently higher in the control group compared to the antibiotic-treated or sterile feeding groups at all time points (Fig. 1e and f; Table S4). After 35 days for *E. pulchripes* and 16 days for *G. connexa*, most animals in the treatment groups ceased faecal production, leading to cessation of plate count.. On day 35 in *E. pulchripes*, the control group significantly differed from the other treatments ( $P = 5.6 \times 10^{-4}$ ), but no significant differences existed between the sterile-fed group and the antibiotics-treated groups (Fig. 1e). For *G. connexa*, significant differences were noted on day 16 between the control and antibiotic-treated groups and between the sterile-fed and 2X-diluted groups (Fig. 1f;  $P = 2.4 \times 10^{-4}$ ). Faecal 16S rRNA gene copies in *E. pulchripes* were reduced by 46%–70% in the treated groups compared to the control group (Fig. 1g; Table S5). In *G. connexa*, 33.9%–40.6% reductions were observed in the sterile, 5X-diluted, and 2X-diluted groups, but these differences were not statistically significant.

After noting a substantial decrease in bacterial load, we measured CH<sub>4</sub> emission on day 35 (Fig. 1h; Table S6). As anticipated, CH<sub>4</sub> was present in *E. pulchripes* but absent in *G. connexa* (data not shown). The control groups displayed a significantly higher CH<sub>4</sub> production rate ( $284.1 \pm 58 \text{ nmol mg}^{-1} \text{ d}^{-1}$ ) than the other treatments ( $P = 0.0008$ ). However, the treated groups saw a 57–74% reduction in CH<sub>4</sub> production without significant differences between them.



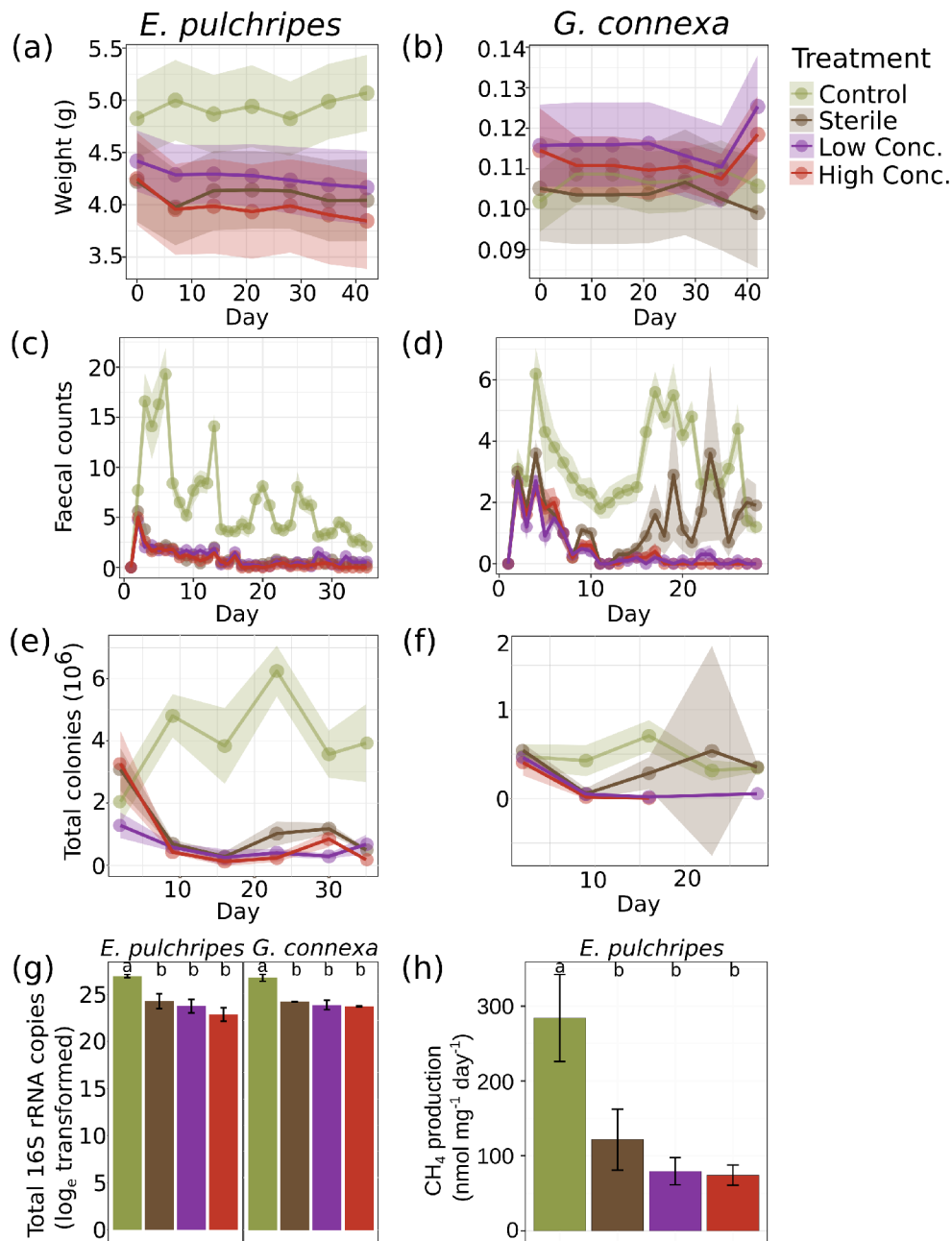


Fig. 1: Effect of antibiotic treatment on *E. pulchripes* and *G. connexa*. Time series of mean weight loss (mean  $\pm$  SE ribbon) in (a) *E. pulchripes* and (b) *G. connexa*; faecal counts in (c) *E. pulchripes* and (d) *G. connexa*; total colony forming units in (e) *E. pulchripes* and (f) *G. connexa*; (g) 16S rRNA gene copy numbers in the faeces; and (h) CH<sub>4</sub> production rate after 35 days of antibiotics treatment in *E. pulchripes*. 'High Conc.' and 'Low Conc.' refer to the concentration of applied antibiotics (see Materials and Methods for more details). Different lower case letters in panels g and h denote statistical significance. See Results for a detailed description of the statistical tests performed on the time series (panels a-f).

## **Prokaryotic community compositions after treatment**

We sequenced 48 samples of *E. pulchripes* and *G. connexa*, consisting of 12 hindguts and 12 faecal samples for each species. The average sequencing depth stood at ca. 40K reads per sample, post-processing of reads and decontamination (Table S7 and S8). The two millipede species differed remarkably in their microbial composition, with the phylum *Bacteroidota* dominating the hindgut of *E. pulchripes* and *Pseudomonadota* that of *G. connexa*. In each case, these phyla comprised over 50% of the abundance regardless of treatment (Fig. 2a and b; Table S9).

*Pseudomonadota* dominated both species' faecal pellets, and *Actinobacteriota*, although rare in the gut, were prominent. (Fig. 2c and d). On the genus level, *E. pulchripes*' hindgut and faecal samples were primarily dominated by *Citrobacter*, *Bacteroides*, and *Pseudomonas* (Fig. 2e-h; Table S9). In contrast, *G. connexa* showed differences between hindgut and faecal sample compositions, with faecal samples appearing more diverse (Fig. 2h).

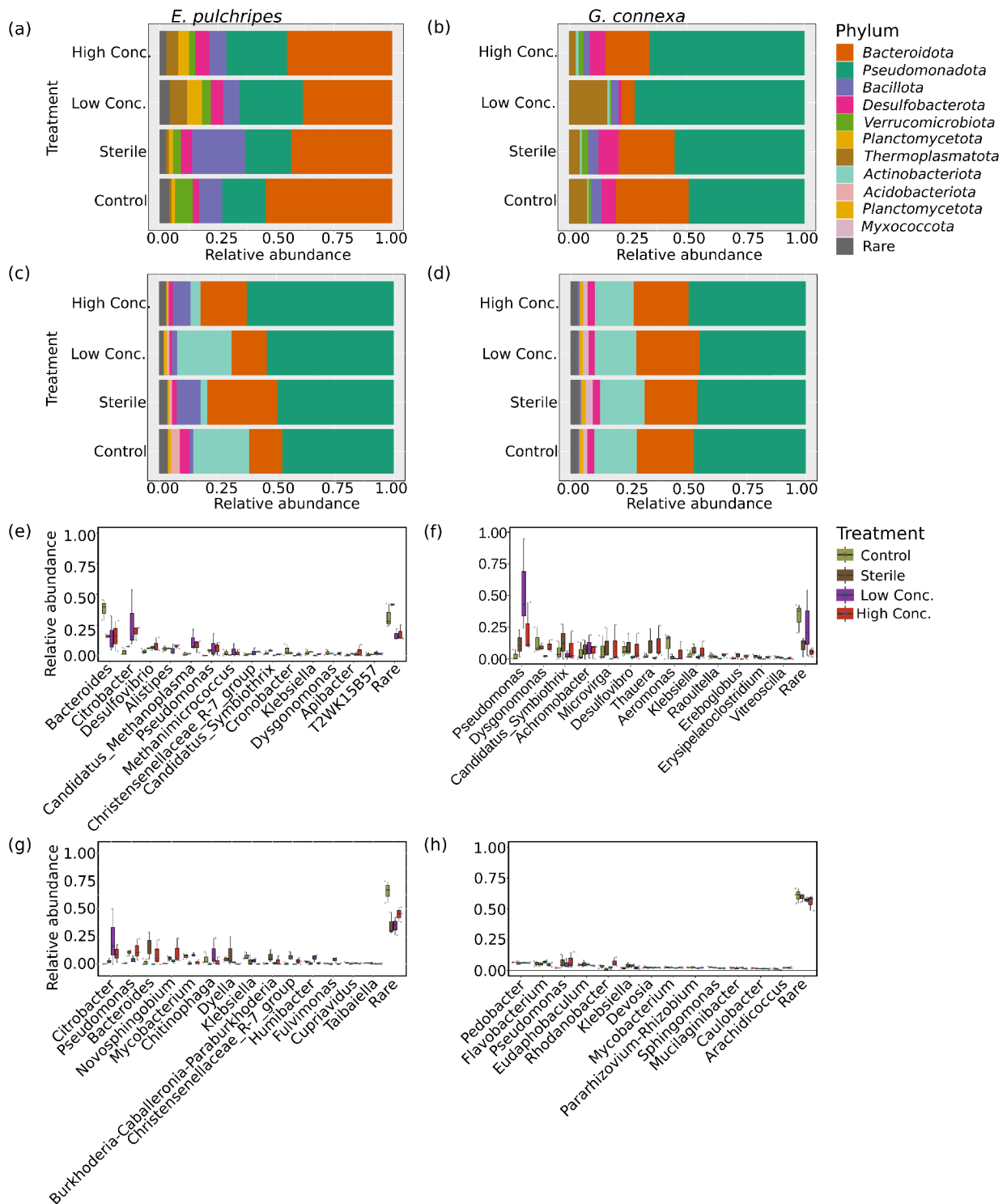


Fig. 2. Effect of antibiotic treatment on the taxonomic composition of prokaryotes in *E. Pulchripes* (left) and *G. Connexa* (right) following treatment. Phylum distribution in the hindguts (a and b) and the faeces (c and d). Distribution at genus level in the hindguts (e and f) and faeces (g and h). 'High Conc.' and 'Low Conc.' refer to the conc. of antibiotics applied (see Materials and Methods for more details).

## **Impact of treatment on prokaryotic community structures**

Overall, no significant differences were found in alpha diversity within or between treatment groups in the hindguts (Fig. 3a & b; Table S10) or faeces (Fig. 3c & d; Table S10) of *E. pulchripes* and *G. connexa*. *E. pulchripes*' hindgut groups displayed greater bacterial diversity and richness than *G. connexa*. In comparison, *G. connexa*'s faecal samples showed higher diversity and richness compared to *E. pulchripes* (also see Fig. S2).

Constrained analysis of principal coordinates (CAP) revealed significant differences in microbial community composition among sterile feeding or antibiotics treatments in both hindguts and faeces of both species (Fig. 3e, f, g & h). ANCOM-BC2 analysis identified only a handful of microbial genera with differential abundance between treatments (Table S11; Fig. S3), indicating that the antibiotic treatment worked relatively non-selective. The few taxa with a decrease in the mean absolute abundance (e.g. *Streptomycetaceae* and *Mucilaginibacter* from the *E. pulchripes*' faeces) are known to often possess antibiotic resistance genes.

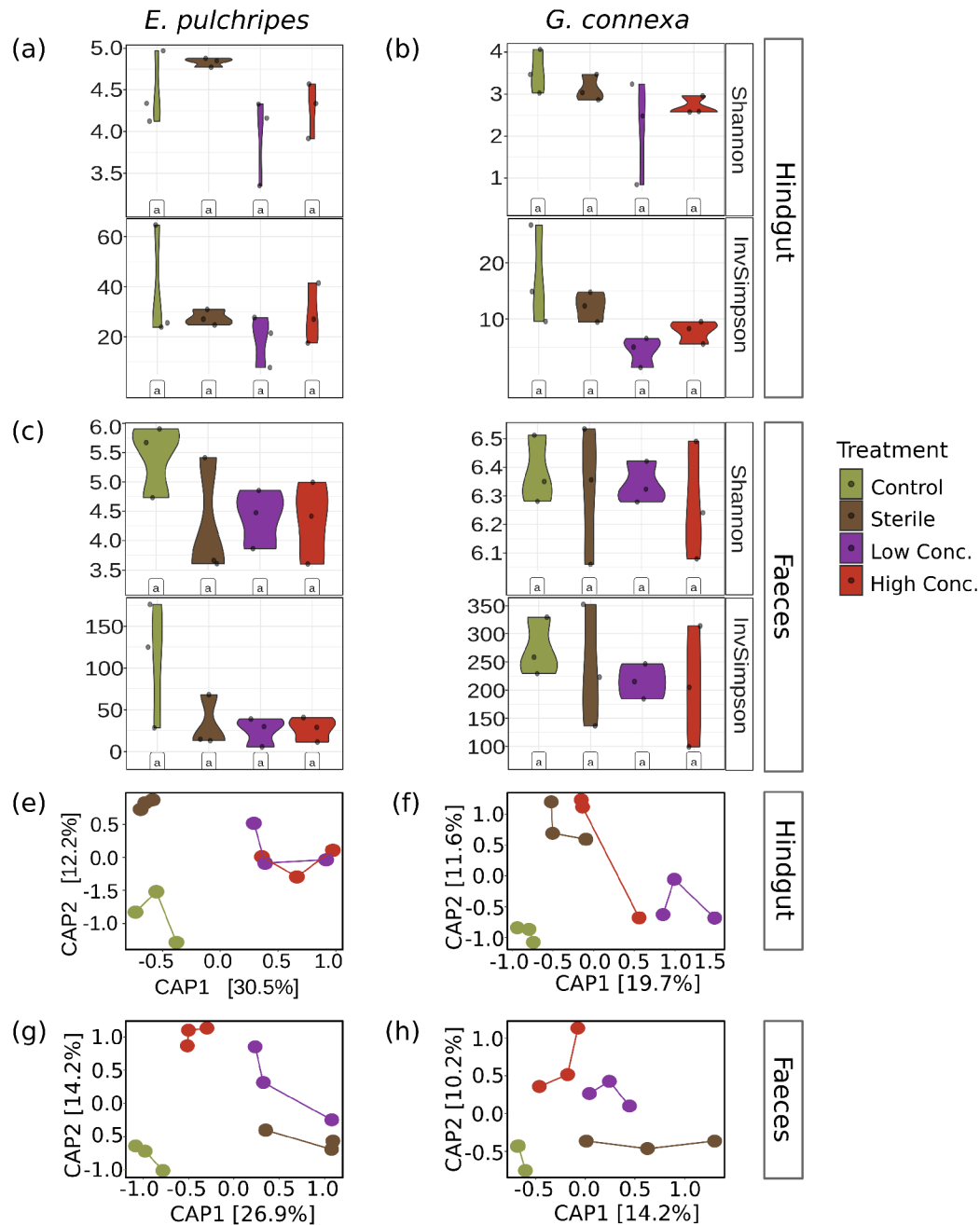


Fig. 3. Effect of antibiotic treatment on the alpha and beta diversity indices of the microbial communities in the hindgut and faeces in *E. pulchripes* (left) and *G. connexa* (right). Alpha diversity values for each species, stratified by treatment groups for hindgut (a and b) and faeces samples (c and d) from *E. pulchripes* and *G. connexa*. The statistical test was based on Kruskal–Wallis (identical letters denote  $p > 0.05$ ). Dissimilarity between hindgut (e and f) and faeces (g and h) microbial communities in the different treatments using constrained principal coordinates analysis (PcoA) with the model  $Dist.Mat \sim Treatment$  for each species and sample type separately.

## **Influence of BES inhibition on methanogenesis in *E. pulchripes***

A diet of Na-BES-treated litter was provided to investigate the importance of methanogenesis in CH<sub>4</sub>-emitting *E. pulchripes*. Methane emissions showed no significant differences on days 0 (P = 0.19) and day 7 (P = 0.08; Fig. 4A; Table S12). However, by day 14, CH<sub>4</sub> production was nearly fully inhibited and remained so for an additional 21 days, with significant inhibition on day 14 (P = 2.7 x 10<sup>-4</sup>) and day 21 (P = 2.2 x 10<sup>-5</sup>). Upon switching to untreated litter on day 35, methane emissions began recovering after 14 days (on day 49). Despite some average weight increase in treated groups, no significant difference was detected at any time (Fig. 4b).

After inhibiting methane production for 21 days, a suspension made from fresh faeces was examined under a bright-field microscope, revealing various protists, nematodes, and rotifers ranging from 12 to 100 µm in size (Fig. S4). The ciliate abundance averaged 3 × 10<sup>5</sup> ml<sup>-1</sup>, regardless of treatment (Fig. 4c; Table S13). Quantification of the *mcrA* gene, pivotal in methane production (Hedderich & Whitman 2006), showed a significant reduction in the two Na-BES-treated groups compared to the control (P = 0.02; Fig. 4d). CARD-FISH was used to detect the presence of free-living (Fig. S5) and symbiotic archaea (Fig. S6), primarily methanogens, in protists from faecal samples. The amplicon sequencing data indicated that members of the *Methanomassciillicoccales* and *Methanobacteriales* were the dominant methanogens in *E. pulchripes*, and these orders were accordingly targeted. Although *mcrA* copy numbers declined, positive hybridisation signals for these methanogen orders were observed in both Na-BES treatments. Methanogens were detected on the 0.2 µm filter (Fig. S5) and associated with protists as endosymbionts (Fig. 4e; Fig. S6), with no significant changes in its count per ciliate (Fig. 4f).



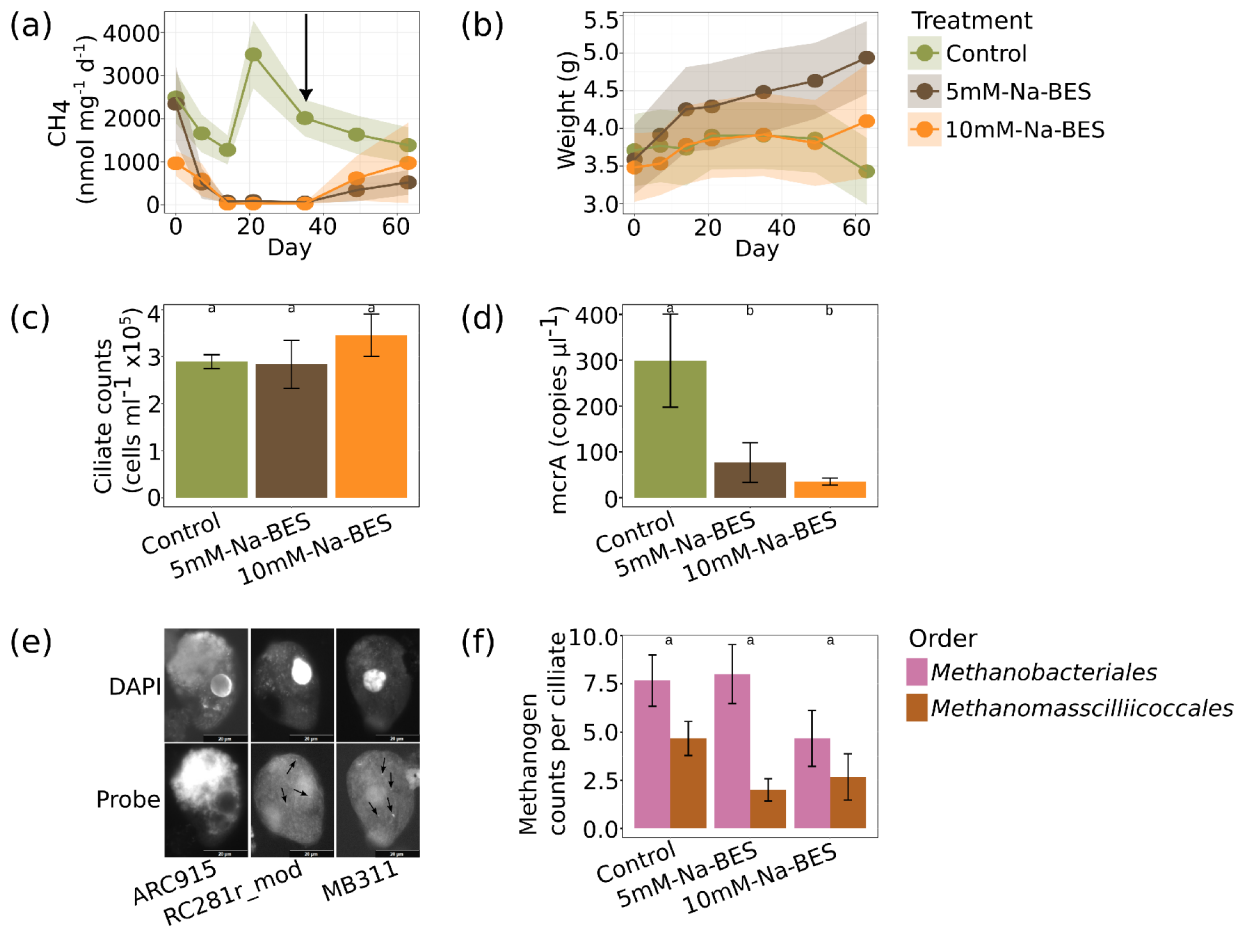


Fig. 4. Effect of BES treatment on CH<sub>4</sub> emissions from *E. pulchripes*, animal weight, ciliates and ciliate-associated methanogens. (a) Emission rates of CH<sub>4</sub> over time followed by recommence of methane production after the switch to untreated litters (indicated by the arrow). (b) Change in the weight of *E. pulchripes* over time. (c) Enumeration of symbiotic ciliates found in the faeces following BES treatment. (d) *mcrA* gene copy numbers in the faecal samples following BES treatment. (e) Fluorescence microscopy images of ciliates and the two most-abundant endosymbiotic methanogens in faecal samples of *E. pulchripes* using DAPI and CARD-FISH probes. ARC915: general archaea, RC281r\_mod: Methanomasscilliicoccales, and MB311: Methanobacteriales in the 10mM-Na-BES-treated group. (f) Enumeration of the methanogens associated with ciliates using FISH signals.

## Detection of active microbiota with <sup>13</sup>C-RNA-SIP

RNA-SIP was used to identify the active microorganisms in the millipedes' gut on a temporal scale (Table S14). The shift in peak of 16S rRNA towards the denser gradient fractions, indicating label incorporation, was evident by day 3 and more prominently by day 7 for *E. pulchripes* and day 14 for *G. connexa* (Fig. 5). Nevertheless, Despite feeding on fully-labelled litter for 21 days, a significant portion of RNA remained unlabelled. Surprisingly, the labelling of the fungal biomass, represented by the 18S rRNA peak, shifted earlier towards denser gradient fractions compared to 16S rRNA in both millipede species (Fig. S7). However, the lack of pronounced peak deviation compared to the control in some replicates and days does not necessarily imply unsuccessful labelling since the labelled fraction of the community might still be too small. Indeed, there was a noticeable and significant change in community composition in the heavy fractions of labelled gradients compared to unlabelled ones already by day 3 (Fig. S8; Table S15).

For comparing heavy fractions in labelled versus unlabelled gradients of 16S RNA, an average of  $1305 \pm 59$  and  $579 \pm 41$  ASVs were used for *E. pulchripes* and *G. connexa* per time point after filtering (Table S16). Surprisingly, the model identified only around 22% of ASVs in *E. pulchripes* and 24% in *G. connexa*, on average, as labelled. Moreover, this proportion of labelled ASVs remained consistent over time in both species. Therefore, the shift in copy-number peaks towards denser fractions, as observed in Fig. 5, was due to increased labelling in already labelled ASVs rather than a change in the proportion of labelled ASVs.

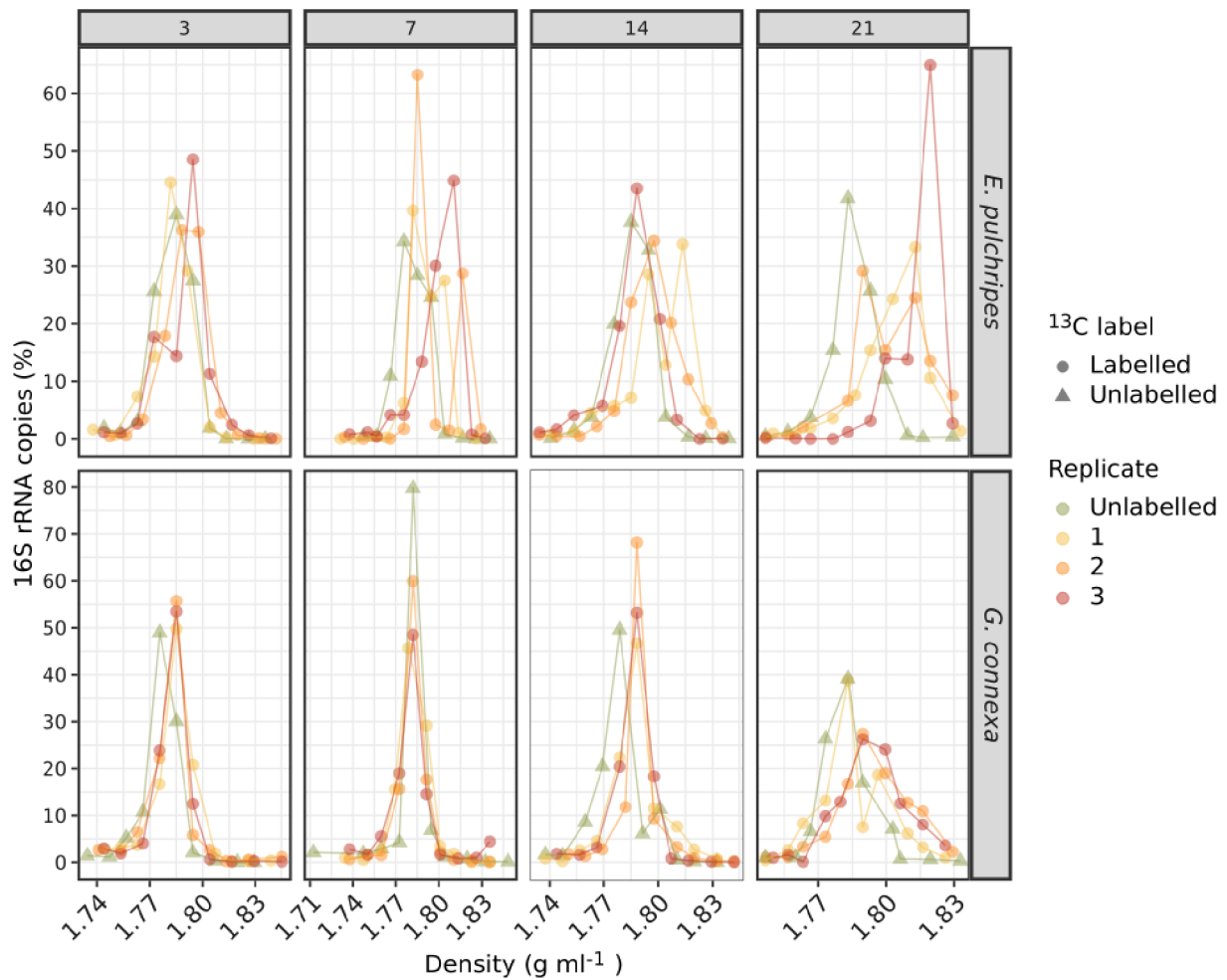


Fig. 5. Bacterial 16S rRNA copies recovered from each fraction in the SIP gradients. rRNA copies relative to the total number of rRNA copies obtained from the entire gradient against the buoyant density of each fraction. Labelled RNA is expected to be found in fractions with density  $>1.795 \text{ g ml}^{-1}$ .

## Diversity of active microbiota in a heavy fraction of $^{13}\text{C}$ -RNA-SIP

In agreement with the general bacterial diversity in the gut, the major phyla whose members were flagged as labelled were *Actinobacteriota*, *Bacillota*, *Bacteroidota*, and *Pseudomonadota* (Fig. 6; Table S16). In *E. pulchripes*, *Bacillota* comprised 35 to 55.3%, *Bacteroidota* 13.1 to 15.1% and *Pseudomonadota* from 13.8 to 23% of the total labelled ASVs. In *G. connexa*,

*Bacillota* comprised 20.4 to 45.9% of total significant ASVs, *Pseudomonadota* ranged from 20 to 51.6%, *Actinobacteriota* from 15.1% to 22.6%, and *Bacteroidota* from 3.2 to 10.8%. Fig. S9-15 show the phylogenetic distribution of the labelled ASVs across the samples in each of the major bacterial classes. Despite our expectation for a gradual labelling of the microorganisms with time, we see the same ASVs consistently labelled throughout the incubation.

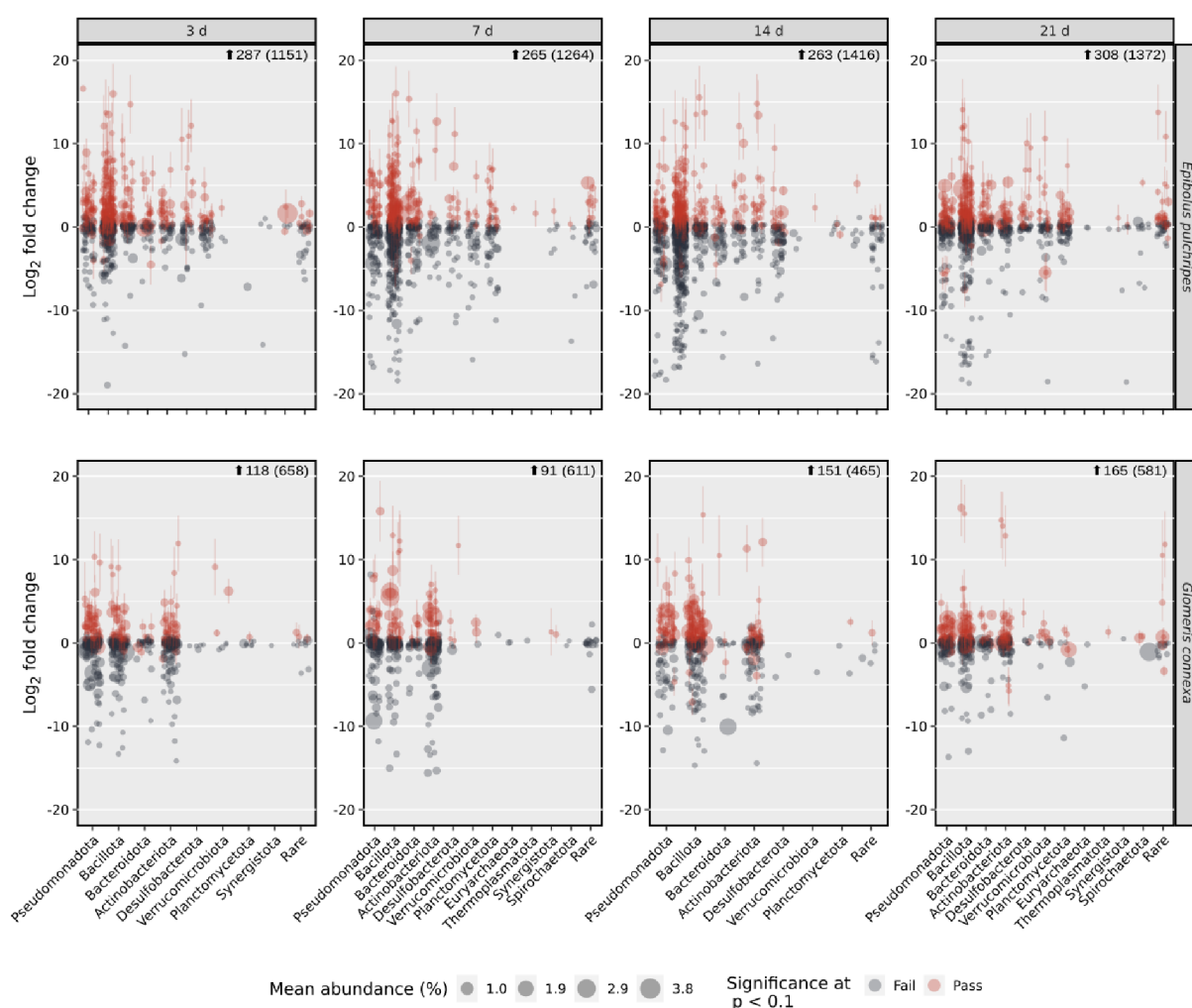


Fig. 6. Differentially abundant ASVs between the labelled and unlabelled gradients of the SIP experiments. Comparison of the relative abundance of each ASV from *E. pulchripes* and *G. connexa*. Each subfigure represents a triplicate. The plot shows the most abundant phyla in the dataset in decreasing abundance. The differential abundance of any particular ASV is given in Log<sub>2</sub> fold change. “Rare” indicates phyla with mean relative abundance below 0.45%.

*E. pulchripes* displayed consistent order across multiple time points. *Enterobacterales* and *Bacteroidales* remained prominent, with base mean values of almost 5 on day 3 and 7, 4 on day 14, and 5.1 and 6.1 on day 21. Other prevalent orders included *Burkholderiales*, *Aeromonadales*, and *Oscillospirales*. In *G. connexa*, *Burkholderiales* and *Lachnospirales* were present at all time points with a base mean value ranging from 2.2 to 6.2 and 1.1 to 4.8. Additionally, the base mean value for *Enterobacterales* rose from 0.9 on day 3 to a range of 5.0 to 8.9 from day 7 to day 21.

## Discussion

The gut microbiota plays a vital role in the ecophysiology of many animals, specifically arthropods. This is particularly true for detritivores because they rely on food rich in recalcitrant plant polymers and poor in nitrogen. In consistency with earlier reviewed culture-based research (Dhivya & Alagesan 2017) and recent molecular studies (Nweze *et al.* 2024; Sardar *et al.* 2022a, b), the results highlight a generally stable and species-specific millipede gut microbiota, even in response to inhibitors. The difference in microbiota between millipede species has been shown before (Nweze *et al.* 2024). Variances among closely related arthropods may stem from gut conditions like pH and oxygen availability (Engel & Moran 2013) and gut topography (Nardi *et al.* 2016). Specifically for millipedes, it was hypothesised that much of the discrepancy in gut microbiota stems from the volume of the hindgut, which in turn directly affects its redox potential, favouring fermentation and methanogenesis in larger species, like *E. pulchripes*, but not in smaller ones, like *G. connexa* (Nweze *et al.* 2024; Šustr *et al.* 2014b).

Curing or sterilising arthropods to evaluate the degree of dependence on their gut microbiota has been performed on several arthropods. Not surprisingly, if

wood-feeding termites are exposed to high oxygen levels, their flagellates disappear, and they die of starvation (Brune 2014; Ebert & Brune 1997). This is because wood-feeding termites rely on short-chain fatty acids, which are the products of cellulose fermentation, for their nutrition. Other studies on cured arthropods showed a more moderate response like a decrease in feeding and altered microbiota, such as in two desert millipedes (Taylor 1982), members of the Carabidae (Lundgren & Lehman 2010) or egg-hatching cockroaches (Tegtmeier *et al.* 2016). In contrast, larval Lepidoptera, which feed on fresh leaves exclusively and likely rely on simple sugars, showed no physiological response to antibiotic curing (Hammer *et al.* 2017). Both millipede species studied here maintained a stable weight throughout the experiment, demonstrating that they might not require fermentation products for their nutrition. However, the marked decrease in faecal production and the fact that, with some exceptions, the taxonomic composition remained intact indicated that the microbiota might nevertheless have an important role. However, we note some shift in abundance towards bacterial strains known for harbouring antibiotic resistance, such as *Citrobacter*, *Bacteroides* (Jabeen *et al.* 2023; Rasmussen *et al.* 1993) in the case of *E. pulchripes* and *Pseudomonas* and *Achromobacter* in *G. connexa* (Abbott & Peleg 2015; Pang *et al.* 2019).

This study confirmed CH<sub>4</sub> release in *E. pulchripes* but not in *G. connexa*, aligning with earlier findings (Horváthová *et al.* 2021; Šustr *et al.* 2014a). Antibiotics reduced CH<sub>4</sub> emission, probably by disrupting bacterial fermentation. A similar observation was made in cockroaches targeting bacteria and flagellates (Gijzen 1991). Not surprisingly, the application of BES, an effective and specific methanogenesis inhibitor (Zhou *et al.* 2011), reduced CH<sub>4</sub> production to below the detection limit. However, this suppression had no apparent effect on the fitness of *E. pulchripes*. Since CH<sub>4</sub> production is needed

as a sink for hydrogen in anaerobic systems to drive syntrophic fermentation processes (Pereira *et al.* 2022), this serves as an additional indication that gut fermentation is not essential for the millipede's nutrition. The dominant methanogens in our millipedes, namely members of the *Methanobacteriales* and *Methanomassiliicoccales*, are known inhabitants of millipede guts (Horváthová *et al.* 2021). Surprisingly, though, while CH<sub>4</sub> production was suppressed and *mcrA* gene copy numbers dropped about 10-fold, this did not affect the observed density of methanogens in the gut. Since the gut is a dynamic system where members must continue to proliferate to avoid being flushed out, this indicates that much like in termites, methanogens live as symbionts of protists and likely directly benefit from their fermentation products (Husseneder 2010; Messer & Lee 1989).

Labelling of the RNA in the SIP experiment was slow and gradual, leaving a significant part of the RNA unlabelled even after a full 21 days. This indicates that the millipede gut system is inefficient in degrading leaf litter and assimilating carbon. In contrast, fungal biomass was labelled quicker, producing higher labelling (esp. in *G. connexa*). Previous research on soil litter decomposition indicates that fungi thrive first on recalcitrant and nutrient-poor litter, with bacteria flourishing later on nutrient-rich litter where carbon is readily available (Schneider *et al.* 2012; Tláskal *et al.* 2016). In the soil, *Ascomycota* prevails at early decomposition stages (Schneider *et al.* 2012) and is later replaced by *Basidiomycota* (Purahong *et al.* 2016; Voříšková & Baldrian 2013). These phyla dominate the hindgut of both millipede species (Nweze *et al.* 2024; Sardar *et al.* 2022b). Although detritivores like millipedes may not exclusively depend on microbial symbiont enzymes for nutrition, salivary gland studies indicate the presence of enzymes that hydrolyse various polysaccharides, lipids, and proteins (Nunez & Crawford 1976), complemented



by resident microbes (Geib *et al.* 2008). Despite indications of cellulose-rich plant material consumption and anoxic conditions in the digestive tract (Horváthová *et al.* 2021) and methanogenesis (Hackstein & Stumm 1994; Šustr *et al.* 2014a), current quantitative data fall short of establishing the significance of cellulose digestion in millipede metabolism, as studies on millipedes fed pure cellulose showed low metabolic rates, suggesting an inability to maintain a positive energy balance (Šustr *et al.* 2020).

The labelled microbiota, mainly Bacillota, Bacteroidota, and Pseudomonadota, exhibit distinct patterns in *E. pulchripes* and *G. connexa*, suggesting their involvement in polysaccharide degradation, aligning with recent studies in millipedes (Sardar *et al.*, 2022a; Nweze *et al.*, 2023b). Members of these phyla were labelled in a similar study in scarab beetles using <sup>13</sup>cellulose (Alonso-Pernas *et al.* 2017). Moreover, while many of the labelled taxa (e.g. *Bacteroidales*, *Burkholderiales* and *Enterobacterales*) are known to be involved in (ligno)cellulose fermentation in millipedes (Nweze *et al.* 2024; Sardar *et al.* 2022a, b), many others (e.g. members of the *Desulfovibrionales* and the two archaeal orders) are hindgut microorganisms involved different processes such as sulfate reduction and methanogenesis, and are likely not involved in this particular fermentation process. Accordingly, we conclude that while cellulolytic fermentation certainly occurs in the millipede gut, it likely makes little to no contribution to the host's diet.

If fermentation products are not a (significant) nutritional source for the millipede, what is it then? Classical studies using <sup>14</sup>C-labelling suggested that the assimilation of bacteria into the millipede's biomass exceeded that of the plant but included lab-grown strains only and failed to include fungi (Bignell 1989). However, the preference for fungi- or bacteria-colonised leaf tissues over natural fresh litter has been demonstrated for woodlice, which is also a

common detritivore (Ihnen & Zimmer 2008). Evidence for the capacity of the millipede gut microbiome to digest plant biomass effectively comes from a recent genomic and transcriptomic screening of the millipede species studied here (Nweze *et al.* 2024). In this work, glycoside hydrolases (GH) capable of degrading chitin and peptidoglycan were as, or even more abundant than cellulose-degrading GHs. The significant decrease in ergosterol levels in the faeces of some millipedes post-digestion (Maraun & Scheu 1996) further supports the notion that substantial fungal digestion is occurring within the millipede gut. While millipedes typically feed on litter and bark, some researchers observed a preference for fungal fruiting bodies, algae, and lichen films among certain species (Semenyuk & Tiunov 2019). In addition to digesting fungi, the millipede midgut fluid has also been shown to kill bacteria effectively in a species-specific manner (Byzov *et al.* 1998). The importance of coprophagy in millipedes has long been debated (Farfan 2010). In light of these results, it may be hypothesised that coprophagy allows millipedes to access fresh microbial and fungal biomass that proliferated thanks to the partial mechanical and chemical breakdown of the recalcitrant plant material (Joly *et al.* 2020). In addition to consuming fungal and microbial biomass, millipedes also produce a variety of endogenous GHs in their salivary glands and midgut that can help them digest non-structural plant material (Nunez & Crawford 1976; Sardar *et al.* 2022a, b). Recently, fluid feeding was described in millipedes of the clade Colobognatha, which enables feeding on fresh plant material (Moritz *et al.* 2022). Naturally, these findings do not exclude other beneficial roles of the millipede gut microbiota, such as detoxification of plant toxins (Hammer & Bowers, 2015), protection against pathogens (Nweze *et al.* 2023, 2024) and even as a source for acquiring new genes through horizontal transfer (So *et al.* 2022).

## Conclusions

Millipedes are keystone detritivores that harbour species-unique and stable microbiota. This work demonstrates that cellulose fermentation likely plays a minor role, at best, in the millipede's nutrition. Further work is needed to decipher their exact trophic function in nature and the potential role their microbiota plays in their survival and modulating greenhouse gas emissions.

## Acknowledgements

We are grateful for the support of Lucie Faktorová and Eva Petrová in collecting *G. connexa* samples, Lucie Faktorová in maintaining the *E. pulchripes* colony and assisting with millipede dissection, and Eva Petrová for her guidance and assistance in DNA and RNA extractions and quantification. We are thankful to Radka Malá for her assistance in the filtration and fixation of CARD-FISH samples. Special thanks to Meador Travis Blake, Jabinski Stanislav, and Ljubov Poláková for their contributions to stable isotope detection and quantification in millipede faeces. RA and JEN were supported by a Junior Grant from the Czech Science Foundation (GA ČR), grant number 19-24309Y.

## Author Contributions

The approach for this study was conceptualised RA and VS, experiments were carried out by SG, JN, MS and TH, and the data analysis was designed by RA and JN. The bioinformatics analyses were carried out by JN and RA. The manuscript was written by JN, SG and RA. All authors have thoroughly reviewed and approved the final version of the manuscript.

## Availability of data and analysis scripts

The short-read amplicon sequencing data have been deposited under the NCBI BioProject PRJNA948469 with BioSample SUB13838396 for antibiotics treatment and SUB13843680 for RNA-SIP. For reproducibility, reusability, and transparency, the scripts used in this study are available on GitHub (<https://github.com/ISBB-anaerobic/Active-microbial-community-pre-and-post-inhibition.git>).

## References

- Abbott, I.J. & Peleg, A.Y. (2015). *Stenotrophomonas*, *Achromobacter*, and nonmelioid *Burkholderia* species: antimicrobial resistance and therapeutic strategies. In: *Seminars in respiratory and critical care medicine*. Thieme Medical Publishers, pp. 099–110.
- Alagesan, P. (2003). Isolation and characterization of gut bacteria of millipede, *Xenobolus carnifex* (Fabricius). *Indian J. Microbiol.*, 43, 111–113.
- Alonso-Pernas, P., Bartram, S., Arias-Cordero, E.M., Novoselov, A.L., Halty-deLeon, L., Shao, Y., et al. (2017). *in vivo* isotopic labeling of symbiotic bacteria involved in cellulose degradation and nitrogen recycling within the gut of the forest cockchafer (*Melolontha hippocastani*). *Front Microbiol*, 8.
- Anderson, M.J. (2001). A new method for non-parametric multivariate analysis of variance. *Austral Ecol.*, 26, 32–46.
- Anderson, M.J. & Willis, T.J. (2003). Canonical analysis of principal coordinates: a useful method of constrained ordination for ecology. *Ecol.*, 84, 511–525.
- Angel, R. (2019). Stable isotope probing techniques and methodological considerations using <sup>15</sup>N. In: *Methods in Molecular Biology: Stable Isotope Probing*. Springer, pp. 175–187.
- Angel, R., Matthies, D. & Conrad, R. (2011). Activation of methanogenesis in arid biological soil crusts despite the presence of oxygen. *PLoS One*, 6, e20453.
- Bates, D., Mächler, M., Bolker, B. & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.*, 67, 1–48.
- Benjamini, Y. & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc., Ser. B, Methodol.*, 57, 289–300.
- Bignell, D. (1989). Relative assimilations of <sup>14</sup>C-labelled microbial tissues and <sup>14</sup>C-plant fibre ingested with leaf litter by the millipede *Glomeris marginata* under experimental conditions. *Soil Biol. Biochem.*, 21, 819–827.
- Bordenstein, S.R. & Theis, K.R. (2015). Host biology in light of the microbiome: ten principles of holobionts and hologenomes. *PLoS Biol.*, 13, e1002226.
- Bouchon, D., Zimmer, M. & Dittmer, J. (2016). The terrestrial isopod microbiome: An all-in-one toolbox for animal–microbe interactions of ecological relevance. *Front. Microbiol.*, 7, 1472.
- Brune, A. (2014). Symbiotic digestion of lignocellulose in termite guts. *Nat Rev Microbiol*, 12, 168–180.
- Byzov, B.A. (2006). Intestinal microbiota of millipedes. In: *Intestinal microorganisms of*

- termites and other invertebrates* (eds. König, H. & Varma, A.). Springer-Verlag, Berlin/Heidelberg, pp. 89–114.
- Byzov, B.A., Thanh, V.N., Bab'Eva, I.P., Tretyakova, E.B., Dyvak, I.A. & Rabinovich, Y.M. (1998). Killing and hydrolytic activities of the gut fluid of the millipede *Pachyiulus flavipes* C.L. Koch on yeast cells. *Soil Biol. Biochem.*, 30, 1137–1145.
- Callahan, B.J., Sankaran, K., Fukuyama, J.A., McMurdie, P.J. & Holmes, S.P. (2016). Bioconductor Workflow for Microbiome Data Analysis: from raw reads to community analyses. *F1000Res*, 5, 1492.
- Crawford, C.S. (1992). Millipedes as model detritivores. *Ber. nat.-med. Verein Innsbruck*, 12.
- Crocetti, G., Murto, M. & Björnsson, L. (2006). An update and optimisation of oligonucleotide probes targeting methanogenic Archaea for use in fluorescence *in situ* hybridisation (FISH). *J Microbiol Methods*, 65, 194–201.
- David, J.-F. (2014). The role of litter-feeding macroarthropods in decomposition processes: A reappraisal of common views. *Soil Biol. Biochem.*, 76, 109–118.
- Davis, N.M., Proctor, D.M., Holmes, S.P., Relman, D.A. & Callahan, B.J. (2018). Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome*, 6, 226.
- Dhivya, A. & Alagesan, P. (2017). Millipedes as Host for Microbes - A Review. *Int. J. Microbiol. Res.*, 8, 19–24.
- Dinno, A. & Dinno, M.A. (2017). Package 'dunn.test.' *CRAN Repos*, 10, 1–7.
- Dixon, P. (2003). VEGAN, a package of R functions for community ecology. *J. Veg. Sci.*, 14, 927–930.
- Douglas, A.E. (2009). The microbial dimension in insect nutritional ecology. *Funct. Ecol.*, 23, 38–47.
- Ebert, A. & Brune, A. (1997). Hydrogen Concentration Profiles at the Oxic-Anoxic Interface: a Microsensor Study of the Hindgut of the Wood-Feeding Lower Termite *Reticulitermes flavipes* (Kollar). *Appl Environ Microbiol*, 63, 4039–4046.
- Engel, P. & Moran, N.A. (2013). The gut microbiota of insects – diversity in structure and function. *FEMS Microbiol. Rev.*, 37, 699–735.
- Enghoff, H. (2011). East African giant millipedes of the tribe *Pachybolini* (Diplopoda, Spirobolida, Pachybolidae). *Zootaxa*, 2753, 1–41.
- Farfan, M.A. (2010). Some Aspects of the Ecology of Millipedes (Diplopoda). Masters Degree. Ohio State University, Ohio, United States.
- Geib, S.M., Filley, T.R., Hatcher, P.G., Hoover, K., Carlson, J.E., Jimenez-Gasco, M. del M., *et al.* (2008). Lignin degradation in wood-feeding insects. *Proc. Natl. Acad. Sci.*, 105, 12932–12937.
- Gerstaecker, A. (1873). *Die gliederthier-fauna des Sansibar-gebietes*. CF Winter.
- Gijzen, H.J. (1991). Methanogenic bacteria as endosymbionts of the ciliate *Nyctotherus ovalis* in the Cockroach Hindgut. *Appl. Environ. Microbiol.*, 57, 5.
- Girden, E.R. (1992). *ANOVA: Repeated measures*. sage.
- Goel, M.K., Khanna, P. & Kishore, J. (2010). Understanding survival analysis: Kaplan-Meier estimate. *Int J Ayurveda Res*, 1, 274–278.
- Hackstein, J.H. & Stumm, C.K. (1994). Methane production in terrestrial arthropods. *Proc. Natl. Acad. Sci.*, 91, 5441–5445.
- Hammer, T.J., Janzen, D.H., Hallwachs, W., Jaffe, S.P. & Fierer, N. (2017). Caterpillars lack a resident gut microbiome. *Proc. Natl. Acad. Sci.*, 114, 9641–9646.
- Hammer, T.J., Sanders, J.G. & Fierer, N. (2019). Not all animals need a microbiome. *FEMS Microbiology Letters*, 366, fnz117.
- Hedderich, R. & Whitman, W.B. (2006). Physiology and biochemistry of the methane-

- producing Archaea. In: *In: Rosenberg, E., DeLong, E.F., Lory, S., Stackebrandt, E., Thompson, F. (eds) The Prokaryotes*. Springer, Berlin, Heidelberg, pp. 1050–1079.
- Hoess, R. & Scholl, A. (2001). Allozyme and Literature Study of *Glomeris guttata* Risso, 1826, and *G. connexa* Koch, 1847, a Case of Taxonomic Confusion (Diplopoda: Glomeridae). *Zool. Anz.*, 240, 15–33.
- Horváthová, T., Šustr, V., Chroňáková, A., Semanová, S., Lang, K., Dietrich, C., *et al.* (2021). Methanogenesis in the Digestive Tracts of the Tropical Millipedes *Archispirostreptus gigas* (Diplopoda, Spirostreptidae) and *Epibolus pulchripes* (Diplopoda, Pachybolidae). *Appl. Environ. Microbiol.*, 87, e00614-21.
- Hurst, G.D.D. & Darby, A.C. (2009). The inherited microbiota of arthropods, and their importance in understanding resistance and immunity. In: *Insect Infection and Immunity: Evolution, Ecology, and Mechanisms* (eds. Rolff, J. & Reynolds, S.). Oxford University Press, pp. 119–136.
- Husseneder, C. (2010). Symbiosis in Subterranean Termites: A Review of Insights From Molecular Studies. *Environmental Entomology*, 39, 378–388.
- Ihnen, K. & Zimmer, M. (2008). Selective consumption and digestion of litter microbes by *Porcellio scaber* (Isopoda: Oniscidea). *Pedobiologia*, 51, 335–342.
- Iino, T., Tamaki, H., Tamazawa, S., Ueno, Y., Ohkuma, M., Suzuki, K., *et al.* (2013). *Candidatus Methanogranum caenicola*: a novel methanogen from the anaerobic digested sludge, and proposal of *Methanomassiliicoccaceae* fam. nov. and *Methanomassiliicoccales* ord. nov., for a methanogenic lineage of the class *Thermoplasmata*. *Microbes Environ.*, 28, 244–250.
- Jabeen, I., Islam, S., Hassan, A.K.M.I., Tasnim, Z. & Shuvo, S.R. (2023). A brief insight into *Citrobacter* species - a growing threat to public health. *Frontiers in Antibiotics*, 2.
- Joly, F.-X., Coq, S., Coulis, M., David, J.-F., Hättenschwiler, S., Mueller, C.W., *et al.* (2020). Detritivore conversion of litter into faeces accelerates organic matter turnover. *Communications Biology*, 3, 660.
- Katoh, K., Misawa, K., Kuma, K. & Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.*, 30, 3059–3066.
- Keselman, H.J. & Rogan, J.C. (1977). The Tukey multiple comparison test: 1953–1976. *Psychol. Bull.*, 84, 1050.
- Kime, R.D. & Golovatch, S.I. (2000). Trends in the ecological strategies and evolution of millipedes (Diplopoda). *Biological Journal of the Linnean Society*, 69, 333–349.
- Kocourek, P., Tajovský, K. & Dolejš, P. (2017). New species of millipedes occurring in the Czech Republic: species discovered in the period 2003–2017, 5.
- Koubová, A., Lorenc, F., Horváthová, T., Chroňáková, A. & Šustr, V. (2023). Millipede gut-derived microbes as a potential source of cellulolytic enzymes. *World J. Microbiol. Biotechnol.*, 39, 169.
- Lin, H. & Peddada, S.D. (2020). Analysis of compositions of microbiomes with bias correction. *Nat. Commun.*, 11, 3514.
- Liu, C.M., Kachur, S., Dwan, M.G., Abraham, A.G., Aziz, M., Hsueh, P.-R., *et al.* (2012). FungiQuant: a broad-coverage fungal quantitative real-time PCR assay. *BMC Microbiol.*, 12, 1–11.
- Love, M.I., Huber, W. & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.*, 15, 550.
- Lundgren, J.G. & Lehman, R.M. (2010). Bacterial gut symbionts contribute to seed digestion in an omnivorous beetle. *PLoS One*, 5, e10831.
- Mangiafico, S. & Mangiafico, M.S. (2017). Package ‘rcompanion.’ *Cran Repos*, 20, 1–71.

- Maraun, M. & Scheu, S. (1996). Changes in microbial biomass, respiration and nutrient status of beech (*Fagus sylvatica*) leaf litter processed by millipedes (*Glomeris marginata*). *Oecologia*, 107, 131–140.
- Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*, 17, 10–12.
- Martinez Arbizu, P. (2020). pairwiseAdonis: Pairwise multilevel comparison using adonis. *R package version 0.4*, 1.
- Mattila, J.M., Zimmer, M., Vesakoski, O. & Jormalainen, V. (2014). Habitat-specific gut microbiota of the marine herbivore *Idotea balthica* (Isopoda). *J. Exp. Mar. Biol. Ecol.*, 455, 22–28.
- McKight, P.E. & Najab, J. (2010). Kruskal-Wallis Test. In: *The Corsini Encyclopedia of Psychology*. John Wiley & Sons, Ltd, pp. 1–1.
- McMurdie, P.J. & Holmes, S. (2013). phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLOS One*, 8, e61217.
- Messer, A.C. & Lee, M.J. (1989). Effect of chemical treatments on methane emission by the hindgut microbiota in the termite *Zootermopsis angusticollis*. *Microbial Ecology*, 18, 275–284.
- Mikaelyan, A., Thompson, C.L., Hofer, M.J. & Brune, A. (2016). Deterministic assembly of complex bacterial communities in guts of germ-free cockroaches. *Appl. Environ. Microbiol.*, 82, 1256–1263.
- Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., Von Haeseler, A., et al. (2020). IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.*, 37, 1530–1534.
- Moran, N.A., Ochman, H. & Hammer, T.J. (2019). Evolutionary and ecological consequences of gut microbial communities. *Annu. Rev. Ecol. Evol. Syst.*, 50, 451–475.
- Moritz, L., Borisova, E., Hammel, J.U., Blanke, A. & Wesener, T. (2022). A previously unknown feeding mode in millipedes and the convergence of fluid feeding across arthropods. *Sci. Adv.*, 8, eabm0577.
- Naqib, A., Poggi, S. & Green, S.J. (2019). Deconstructing the polymerase chain reaction II: an improved workflow and effects on artifact formation and primer degeneracy. *PeerJ*, 7, e7121.
- Nardi, J.B., Bee, C.M. & Taylor, S.J. (2016). Compartmentalization of microbial communities that inhabit the hindguts of millipedes. *Arthropod Struct. Dev.*, 45, 462–474.
- Nunez, F.S. & Crawford, C.S. (1976). Digestive enzymes of the desert millipede *Orthoporus ornatus* (Girard) (Diplopoda: Spirostreptidae). *Comp. Biochem. Physiol. Part A Mol. Integr. Physiol.*, 55, 141–145.
- Nweze, J., Schweichhart, J. & Angel, R. (2023). Viral Communities in Millipede Guts: Insights into Diversity and the Potential Role in Modulating the Microbiome. *Preprint (Version 1)*.
- Nweze, J.E., Šustr, V., Brune, A. & Angel, R. (2024). Functional similarity, despite taxonomical divergence in the millipede gut microbiota, points to a common trophic strategy. *Microbiome*, 12, 16.
- Pang, Z., Raudonis, R., Glick, B.R., Lin, T.-J. & Cheng, Z. (2019). Antibiotic resistance in *Pseudomonas aeruginosa*: mechanisms and alternative therapeutic strategies. *Biotechnol. Adv.*, 37, 177–192.
- Pereira, A.M., de Lurdes Nunes Enes Dapkevicius, M. & Borba, A.E.S. (2022). Alternative pathways for hydrogen sink originated from the ruminal fermentation of carbohydrates: Which microorganisms are involved in lowering methane emission? *Anim. Microbiome*, 4, 5.



- Perlmutter, J.I. & Bordenstein, S.R. (2020). Microorganisms in the reproductive tissues of arthropods. *Nat Rev Microbiol*, 18, 97–111.
- Petersen, J.M. & Osvatic, J. (2018). Microbiomes in nature: Importance of invertebrates in understanding the natural variety of animal-microbe interactions. *mSystems*, 3, e00179-17.
- Piwosz, K., Mukherjee, I., Salcher, M.M., Grujčić, V. & Šimek, K. (2021). CARD-FISH in the sequencing era: Opening a new universe of protistan ecology. *Front. Microbiol.*, 12.
- Purahong, W., Wubet, T., Lentendu, G., Schloter, M., Pecyna, M.J., Kapturska, D., *et al.* (2016). Life in leaf litter: novel insights into community dynamics of bacteria and fungi during litter decomposition. *Mol. Ecol.*, 25, 4059–4074.
- Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., *et al.* (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.*, 41, D590-596.
- Ramanathan, B. & Alagesan, P. (2012). Isolation, characterization and role of gut bacteria of three different millipede species. *Indian J. Sci. Res.*, 3, 55–61.
- Rasmussen, B.A., Bush, K. & Tally, F.P. (1993). Antimicrobial Resistance in Bacteroides. *Clin. Infect. Dis.*, 16, S390–S400.
- RCore, T. (2016). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Russell, J.A., Dubilier, N. & Rudgers, J.A. (2014). Nature’s microbiome: introduction. *Mol. Ecol.*, 23, 1225–1237.
- Sardar, P., Šustr, V., Chroňáková, A. & Lorenc, F. (2022a). Metatranscriptomic holobiont analysis of carbohydrate-active enzymes in the millipede *Telodeinopus aoutii* (Diplopoda, Spirostreptida). *Front. Ecol. Evol.*, 10.
- Sardar, P., Šustr, V., Chroňáková, A., Lorenc, F. & Faktorová, L. (2022b). *De novo* metatranscriptomic exploration of gene function in the millipede holobiont. *Sci. Rep.*, 12, 16173.
- Schmidt, K. & Engel, P. (2021). Mechanisms underlying gut microbiota–host interactions in insects. *J. Exp. Biol.*, 224, jeb207696.
- Schneider, T., Keiblinger, K.M., Schmid, E., Sterflinger-Gleixner, K., Ellersdorfer, G., Roschitzki, B., *et al.* (2012). Who is who in litter decomposition? Metaproteomics reveals major microbial players and their biogeochemical functions. *ISME J.*, 6, 1749–1762.
- Semenyuk, I.I. & Tiunov, A.V. (2019). Foraging behaviour as a mechanism for trophic niche separation in a millipede community of southern Vietnam. *Eur. J. Soil Biol.*, 90, 36–43.
- So, W.L., Nong, W., Xie, Y., Baril, T., Ma, H., Qu, Z., *et al.* (2022). Myriapod genomes reveal ancestral horizontal gene transfer and hormonal gene loss in millipedes. *Nat. Commun.*, 13, 3010.
- Stahl, D.A. (1991). Development and application of nucleic acid probes in bacterial systematics. In: *Nucleic Acid Techniques in Bacterial Systematics* (eds. Stackebrandt, E. & Goodfellow, M.). John Wiley & Sons Ltd., Chichester, UK, pp. 205–248.
- Šustr, V., Chroňáková, A., Semanová, S., Tajovský, K. & Šimek, M. (2014a). Methane production and methanogenic archaea in the digestive tracts of millipedes (Diplopoda). *PLoS one*, 9.
- Šustr, V., Šimek, M., Faktorová, L., Macková, J. & Tajovský, K. (2020). Release of greenhouse gases from millipedes as related to food, body size, and other factors. *Soil Biol. Biochem.*, 144, 107765.
- Šustr, V., Stingl, U. & Brune, A. (2014b). Microprofiles of oxygen, redox potential, and pH,

- and microbial fermentation products in the highly alkaline gut of the saprophagous larva of *Penthetria holosericea* (Diptera: Bibionidae). *J. Insect Physiol.*, 67, 64–69.
- Taylor, E.C. (1982). Role of aerobic microbial populations in cellulose digestion by desert millipedes. *Appl. Environ. Microbiol.*, 11.
- Tegtmeier, D., Thompson, C.L., Schauer, C. & Brune, A. (2016). Oxygen Affects Gut Bacterial Colonization and Metabolic Activities in a Gnotobiotic Cockroach Model. *Appl Environ Microbiol*, 82, 1080–1089.
- Tinker, K.A. & Ottesen, E.A. (2016). The core gut microbiome of the American cockroach, *Periplaneta americana*, is stable and resilient to dietary shifts. *Appl. Environ. Microbiol.*, 82, 6603–6610.
- Tláskal, V., Voříšková, J. & Baldrian, P. (2016). Bacterial succession on decomposing leaf litter exhibits a specific occurrence pattern of cellulolytic taxa and potential decomposers of fungal mycelia. *FEMS Microbiol. Ecol.*, 92, fiw177.
- Vavre, F. & Kremer, N. (2014). Microbial impacts on insect evolutionary diversification: from patterns to mechanisms. *Curr. Opin. Insect. Sci.*, 4, 29–34.
- Voříšková, J. & Baldrian, P. (2013). Fungal community on decomposing leaf litter undergoes rapid successional changes. *ISME J*, 7, 477–486.
- Wallner, G., Amann, R. & Beisker, W. (1993). Optimizing fluorescent *in situ* hybridization with rRNA-targeted oligonucleotide probes for flow cytometric identification of microorganisms. *Cytom., : j. Int. Soc. Anal. Cytol.*, 14, 136–143.
- Walters, W., Hyde, E.R., Berg-Lyons, D., Ackermann, G., Humphrey, G., Parada, A., *et al.* (2016). Improved bacterial 16S rRNA gene (V4 and V4-5) and fungal internal transcribed spacer marker gene primers for microbial community surveys. *mSystems*, 1, e00009-15.
- Watanabe, H. & Tokuda, G. (2001). Animal cellulases. *Cell. Mol. Life Sci.*, 58, 1167–1178.
- Yu, Y., Lee, C., Kim, J. & Hwang, S. (2005). Group-specific primer and probe sets to detect methanogenic communities using quantitative real-time polymerase chain reaction. *Biotechnol. Bioeng.*, 89, 670–679.
- Zhou, Z., Meng, Q. & Yu, Z. (2011). Effects of methanogenic inhibitors on methane production and abundances of methanogens and cellulolytic bacteria in *in vitro* ruminal cultures. *Appl Environ Microbiol*, 77, 2634–2639.
- Zhu, A., Ibrahim, J.G. & Love, M.I. (2019). Heavy-tailed prior distributions for sequence count data: removing the noise and preserving large differences. *Bioinform.*, 35, 2084–2092.
- Zilber-Rosenberg, I. & Rosenberg, E. (2008). Role of microorganisms in the evolution of animals and plants: the hologenome theory of evolution. *FEMS Microbiol. Rev.*, 32, 723–735.
- Zimmer, M. & Bartholmé, S. (2003). Bacterial endosymbionts in *Asellus aquaticus* (Isopoda) and *Gammarus pulex* (Amphipoda) and their contribution to digestion. *Limnol. Oceanogr.*, 48, 2208–2213.
- Zou, K.H., Tuncali, K. & Silverman, S.G. (2003). Correlation and simple linear regression. *Radiology*, 227, 617–628.

# **Cellulose fermentation by the gut microbiome is likely not essential for the nutrition of millipedes**

**Julius Eyiuche Nweze<sup>1,2</sup>, Shruti Gupta<sup>1</sup>, Michaela Salcher<sup>3</sup>, Vladimír Šustr<sup>1</sup>, Terézia Horváthová<sup>1</sup>, Roey Angel<sup>1,2\*</sup>**

<sup>1</sup>Institute of Soil Biology and Biogeochemistry, Biology Centre CAS, České Budějovice, Czechia

<sup>2</sup>Faculty of Science, University of South Bohemia in České Budějovice, Czechia

<sup>3</sup>Institute of Hydrobiology, Biology Centre CAS, České Budějovice, Czechia

\* Correspondence: roey.angel@bc.cas.cz

## **Supplementary material**

### **Supplementary Tables**

See file: Nweze\_et\_al\_Supplementary\_Table.xlsx

## **Supplementary methods**

### **Identification and enumeration of protists and symbiotic methanogens**

The polycarbonate filters were embedded in warm, 0.2% low-melting agarose (Thermo Fisher Scientific) and dried at 37 °C in an oven. Cells on the filters were permeabilised with a lysozyme solution (10 mg ml<sup>-1</sup>; Sigma-Aldrich) at 37 °C for 45 minutes, then incubated in achromopeptidase solution (2 µl of 30 KU achromopeptidase in 1 ml of NaCl Tris buffer; Sigma-Aldrich Co) for 15 minutes at 37 °C. For hybridisation, filters were cut, labelled, and hybridised using 300 µl hybridisation buffer and 2 µl probe for 2 hours at 35 °C, followed by 20-30 min wash in a 37 °C washing buffer and incubation in phosphate-buffered saline with Tween-20 for 45 minutes at 37 °C. The signal from the probe was amplified by incubating blot-paper-dabbed filters with fluorescently labelled tyramide (Sigma-Aldrich Co) and 0.15% H<sub>2</sub>O<sub>2</sub> for 30 min in the dark. Subsequently, fluorescein isothiocyanate (FITC; Sigma-Aldrich) was used as a fluorochrome, and 4',6-diamidino-2-phenylindole (DAPI; Sigma-Aldrich) as a sample counterstain.

### **Nucleic acid extraction and quantification**

Nucleic acids were extracted as follows: the samples were subjected to three consecutive bead beating rounds (Lysing Matrix E tubes; MP Biomedicals™) in a FastPrep-24™ 5G (MP Biomedicals™) in the presence of CTAB, phosphate buffer (pH 8.0) and E-saturated phenol, followed by phenol-chloroform-isoamyl alcohol (25:24:1; all from Thermo Fisher Scientific) purification, precipitation using Invitrogen™ UltraPure™ Glycogen (Thermo Fisher Scientific) and then purification with the OneStep™ PCR Inhibitor Removal Kit (Zymo Research). The resulting DNA was then quantified using the Quant-it™ PicoGreen DNA Assay Kit (Thermo Fisher Scientific). The RNA was purified from the DNA/RNA hindgut extract using TURBO™ DNase and GeneJET RNA Cleanup and Concentration Micro Kit (Thermo Fisher Scientific) for the SIP experiment. The quantity and quality of the RNA were determined using

Quant-it™ RiboGreen RNA Assay Kit (Thermo Fisher Scientific) and the Agilent 2100 bioanalyzer (Agilent Technologies).

### **Isopycnic ultra-centrifugation of <sup>13</sup>C labelled RNA**

Each gradient was prepared with 4.848 ml of CsFTA solution (GE Healthcare), 1.083 ml of gradient buffer (0.1 M Tris-HCl, 0.1 M KCl, 1 mM EDTA; Sigma Aldrich) and 211 µl of Hi-Di Formamide (3.56 % v/v; Thermo Fisher Scientific). The density was confirmed using an AR200 Automatic Digital Refractometer (Reichert) against a calibration curve, and occasionally, adjustments were made using CsTFA or gradient buffer until a final density of 1.79 g ml<sup>-1</sup> was reached. Finally, each 6 ml Ultracrimp PA centrifugation tube (Thermo Fisher Scientific) contained approximately 5.8 ml of the density gradient solution and ca. 500 ng of RNA. Additionally, a control tube containing no RNA was used to exclude the presence of DNA or RNA contamination. Tubes were centrifuged in a TV-1665 vertical rotor in a Sorvall WX Ultra 100 Ultracentrifuge (Thermo Fisher Scientific) at 20 °C and 130,000 ×g for 72 h. Twelve fractions of ca. 500 µl were collected into 2.0 ml low-binding collection tubes (Eppendorf) using a NE-300 Just Infusion™ Syringe Pump (NEW ERA PumpSystem Inc.), and the buoyant density (BD) of each fraction from the control tube was determined using a refractometer. RNA was then precipitated from the gradient fractions using 2 µl of GlycoBlue (Thermo Fisher Scientific), 47 µl of 3M Na-Acetate (pH 5.5) (Thermo Fisher Scientific) and 1.175 ml ethanol (absolute), and dissolved in 10 µl of RNA Storage solution (Thermo Fisher Scientific). Lastly, RNA from the gradient fractions 2–11 was converted into cDNA in 20 µl reactions using SuperScript IV RT (Thermo Fisher Scientific). The resulting cDNA was then stored at -20 °C until further processing.

## Supplementary figures

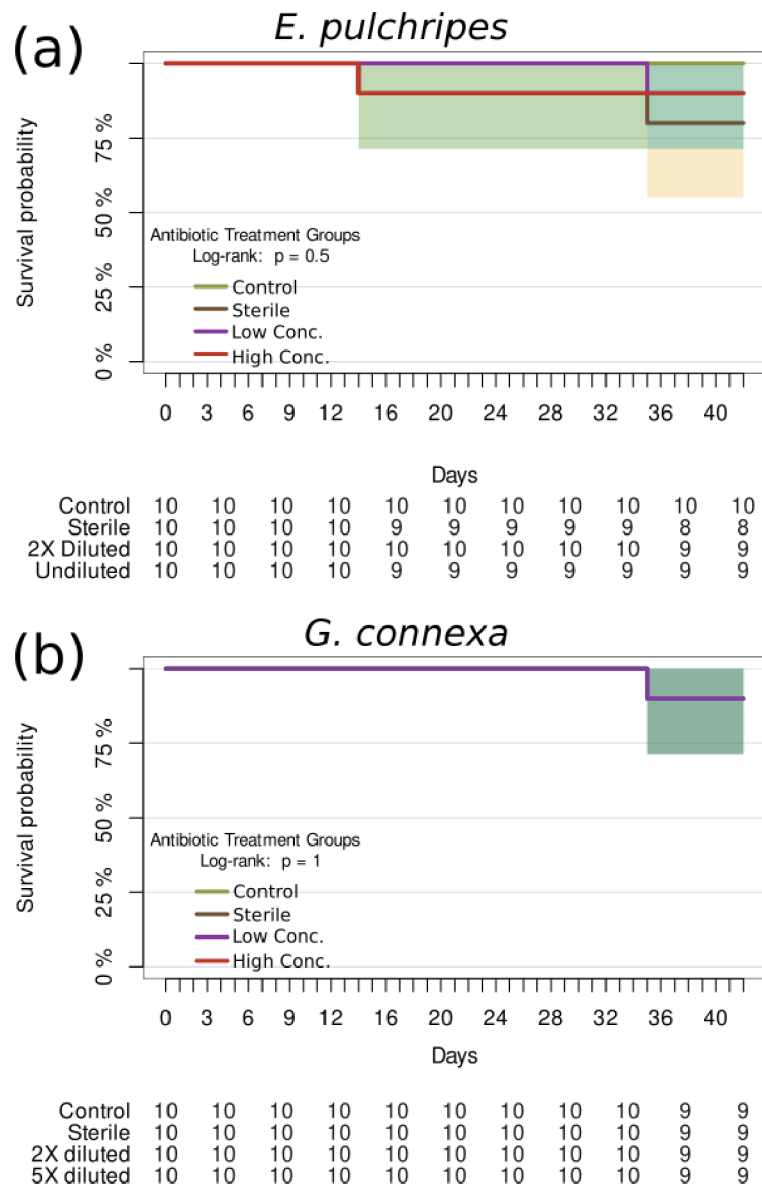


Fig. S1 Kaplan-Meier survival curves based on treatment for *E. pulchripes* and *G. connexa*. The study involved a total of 40 individuals, evenly distributed across four groups. The term 'High Conc.' indicates the group fed litter-treated antibiotics diluted at a ratio of 1:2 (2X-diluted), whereas 'Low Conc.' pertains to antibiotics diluted at a ratio of 1:5 (2X-diluted). The term 'Undiluted' denotes the original antibiotics solution containing penicillin G: 10,000 units  $ml^{-1}$ , streptomycin sulfate: 10 mg  $ml^{-1}$ , and amphotericin B: 25  $\mu g \mu l^{-1}$ .

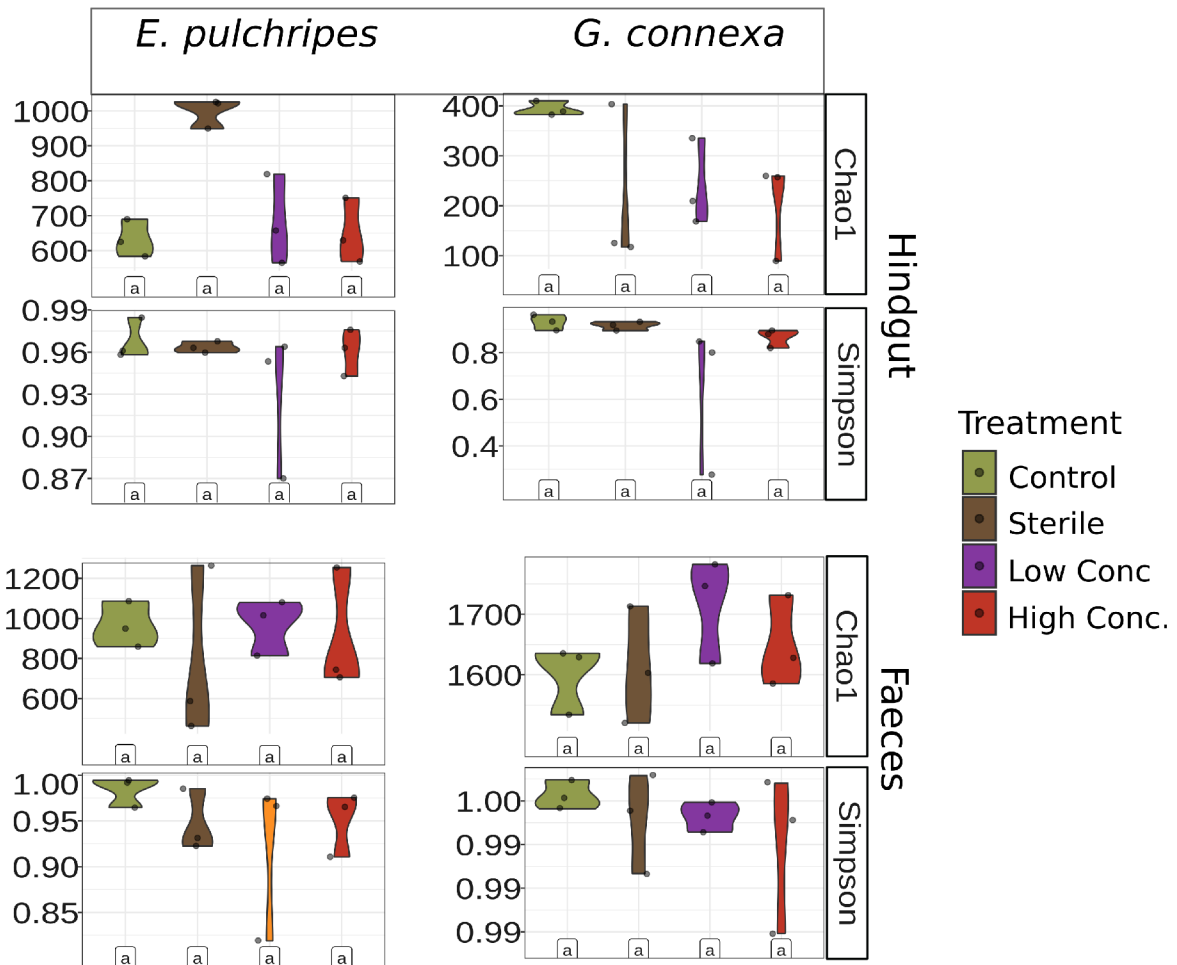


Fig. S2 Alpha diversity indices of the microbial communities in the hindgut and faeces after antibiotics treatment in *E. pulchripes* and *G. connexa*. Additional alpha diversity values for each species to those shown in Fig. 3, stratified by treatment groups of hindguts and faeces from *E. pulchripes* and *G. connexa*. The statistical test was based on Kruskal–Wallis (identical letters indicate  $p > 0.05$ ). Refer to Fig. S1 for details regarding the treated groups.



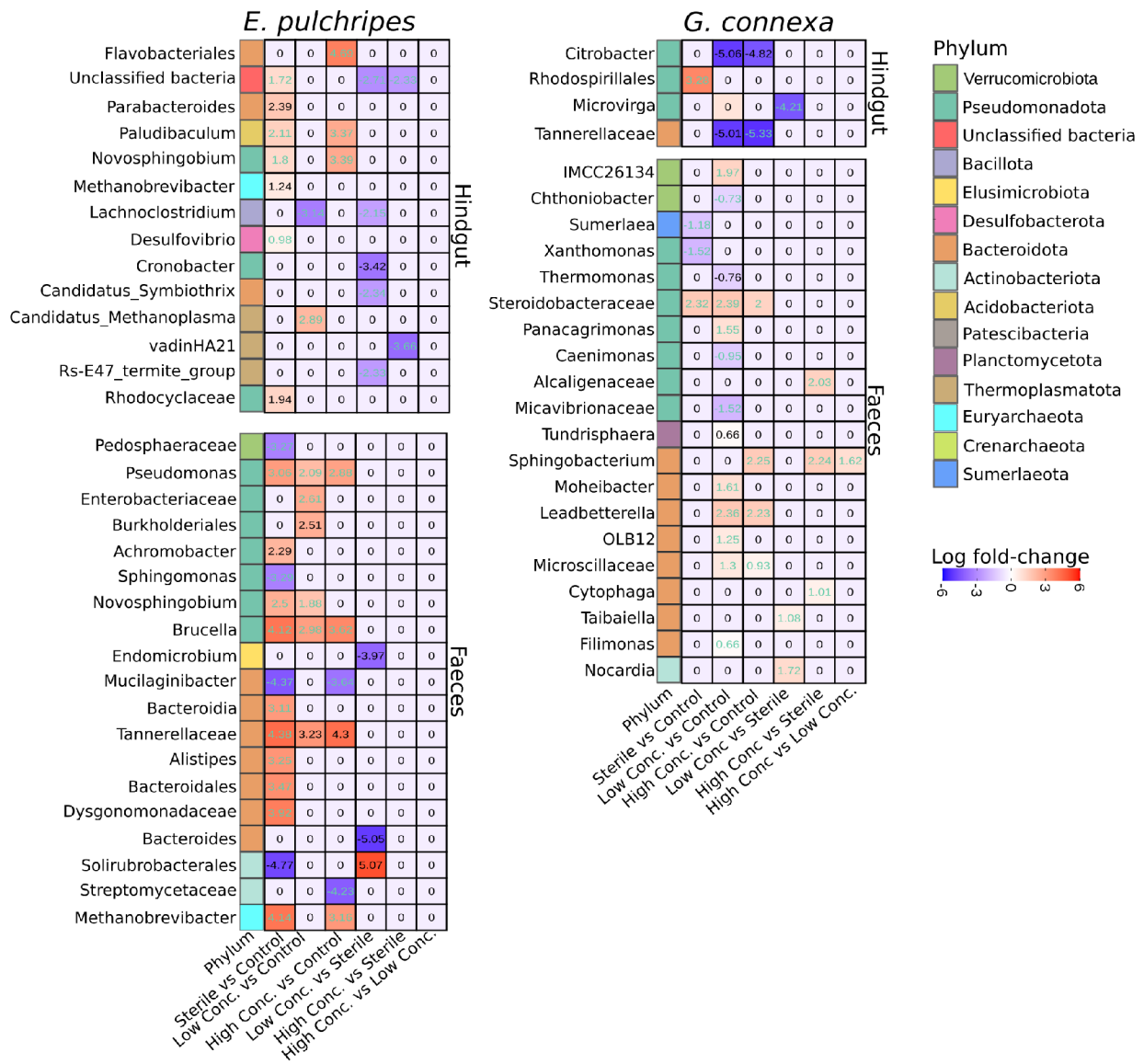


Fig. S3 Heatmaps of the ANCOM-BC2 pairwise analysis for the effect of antibiotics or sterile feeding on the microbial relative abundances in the hindgut and faecal samples from *E. pulchripes* and *G. connexa*. The heatmaps show multiple pairwise comparisons on the genus level among the four groups: control, sterile-fed, low-conc. antibiotics and high-conc. antibiotics. The X-axis represents treatment comparison, while the Y-axis displays significant genera and their phylum, identified by ANCOM-BC2. Each cell is colour-coded with blue representing reduced abundance and red representing increased abundance in response to the treatment. The numbers in each cell indicate the log fold-change. The Benjamini-Hochberg method was used to correct for multiple testing, and taxa with log fold-change values marked in green have successfully passed the sensitivity analysis for pseudo-count addition. Refer to Fig. S1 for details regarding the treated groups.

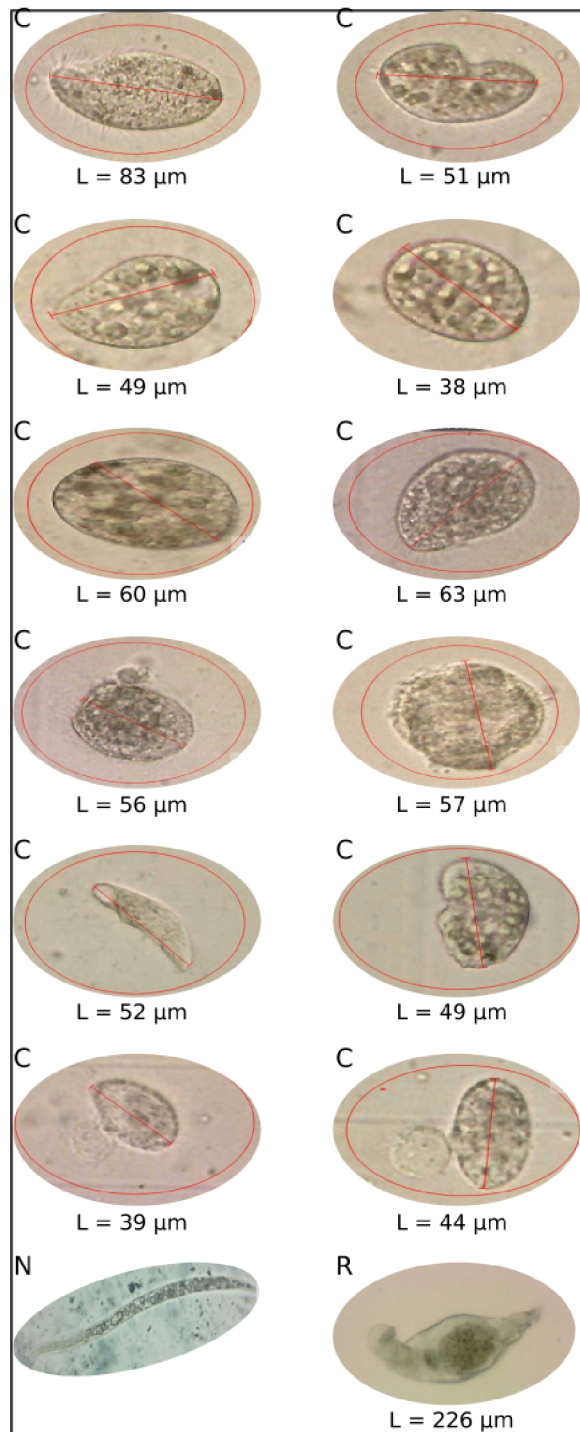


Fig. S4 Light-microscopy crop images of symbiotic ciliates, rotifer and nematode found in the faeces of *E. pulchripes*. “L” represents the length of the organisms as measured from the image.

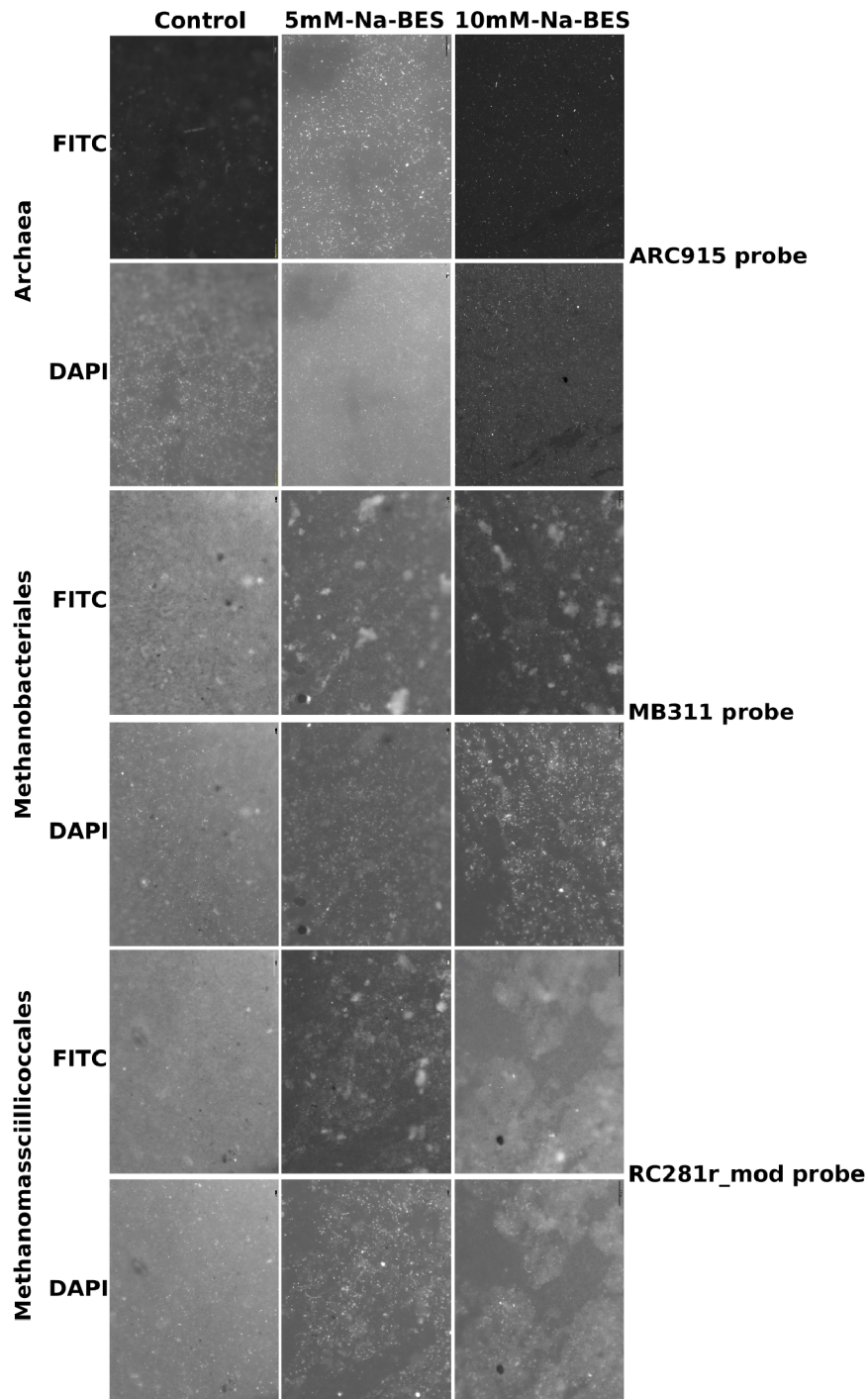


Fig. S5 CARD-FISH images of filtered cells from faecal samples of Na-BES-fed *E. pulchripes*. The free-living or detached methanogens were captured from faecal samples using a 0.2  $\mu\text{m}$  filter and labelled with the following probes: ARC915 for general archaea, MB311 for Methanobacteriales, and RC281r\_mod for Methanomassciillicoccales. Each FITC-probe image has a parallel DAPI-stained image. The treatments were represented by control (group fed with untreated litters); 10mM-Na-BES-treated litters (group fed with 10mM-Na-BES-treated litters); and 5mM-Na-BES-treated litters (group fed with 5mM-Na-BES-treated litters).

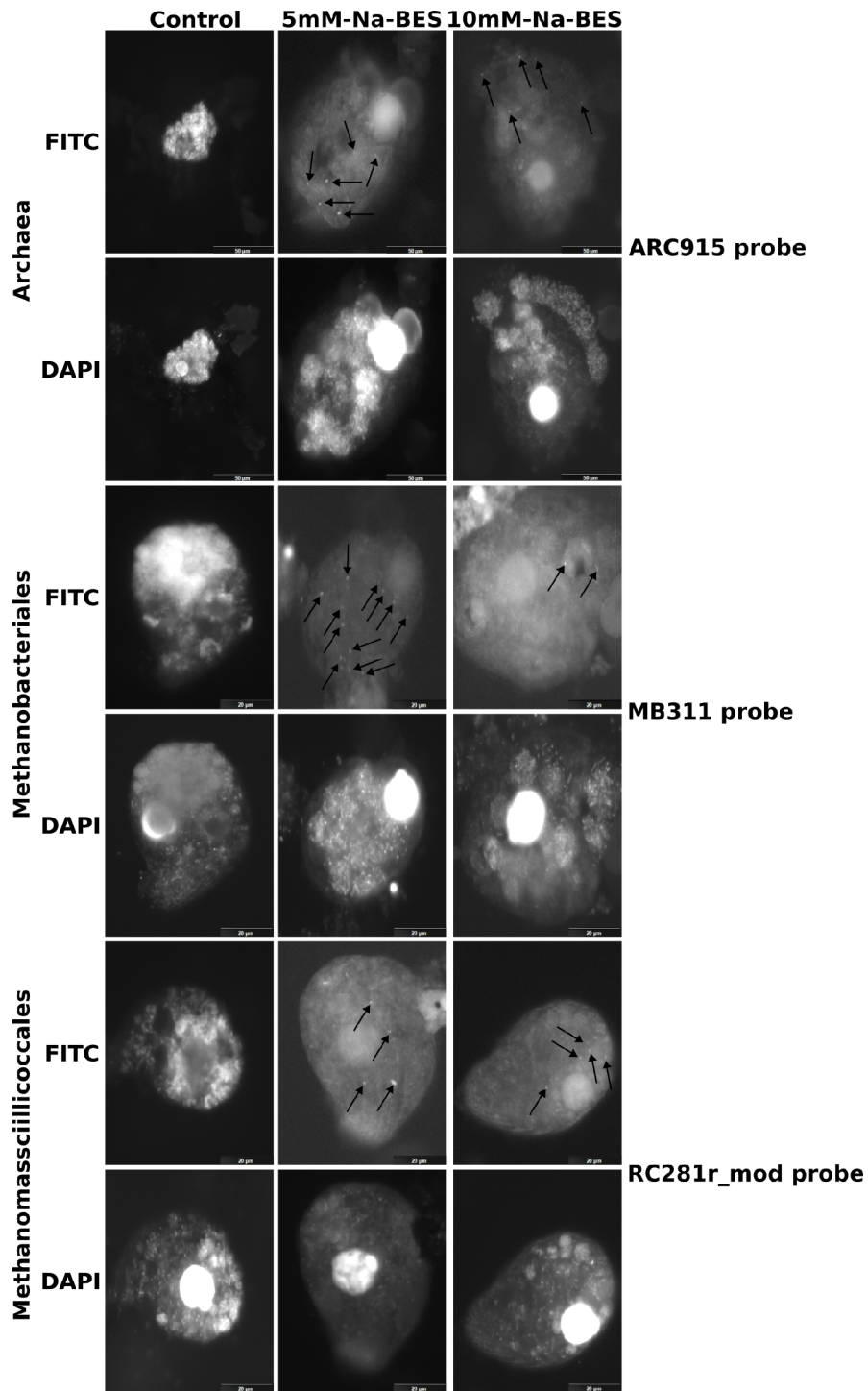


Fig. S6. CARD-FISH images of ciliate-associated cells from faecal samples of Na-BES-fed *E. pulchripes*. Ciliates were captured from faecal samples using a 10  $\mu$ m filter and labelled with the following probes: ARC915 for general archaea, MB311 for Methanobacteriales, and RC281r\_mod for Methanomasscillicoocales. Each FITC-probe image has a parallel DAPI-stained image.

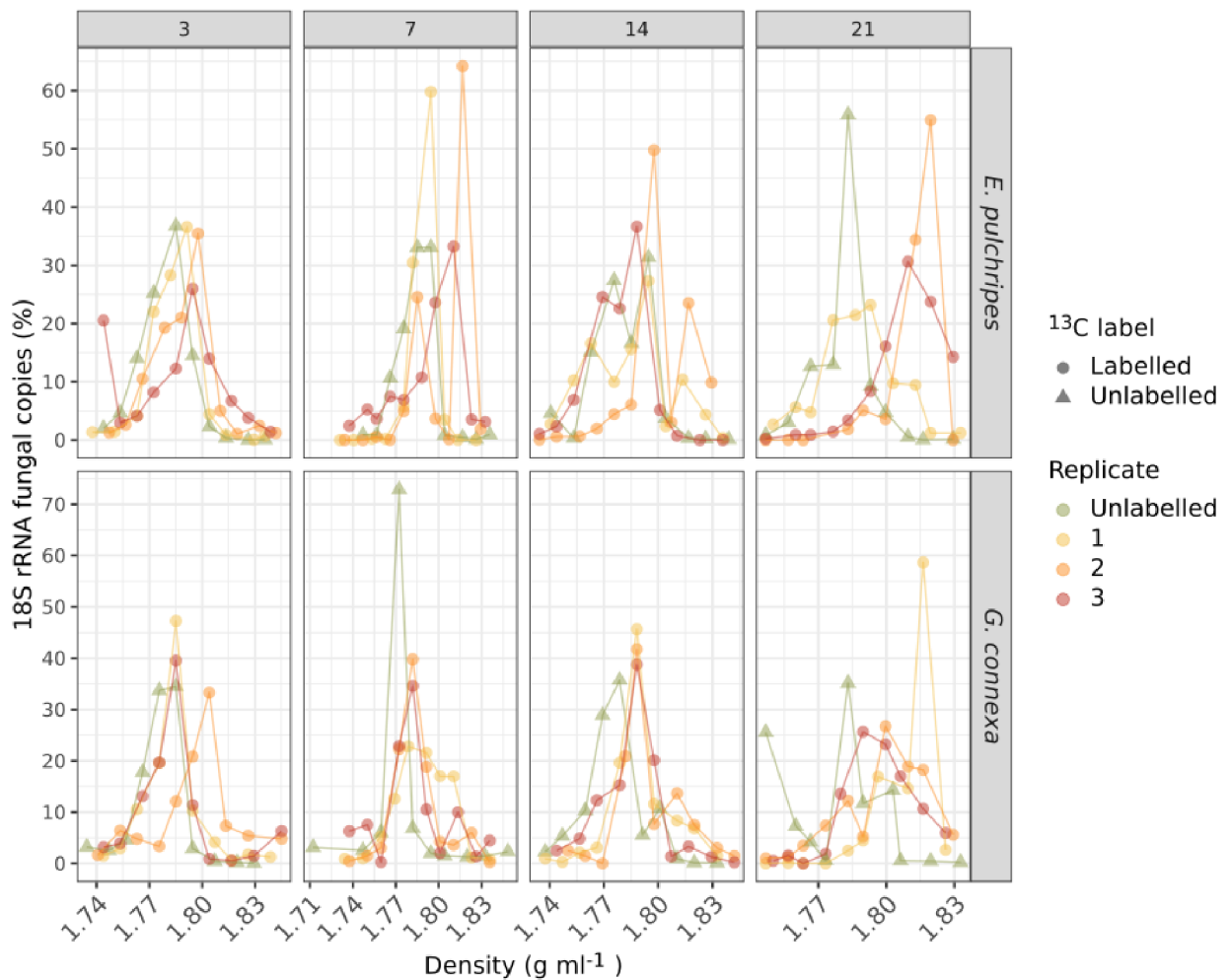


Fig. S7. Fungal 18S rRNA copies recovered from each fraction in the SIP gradients. Values on the y-axis are the rRNA copies relative to the total number of rRNA copies obtained from the entire gradient in %. Values on the x-axis show the buoyant density of each fraction. Labelled RNA is expected to be found in fractions with density  $>1.795 \text{ g ml}^{-1}$ .



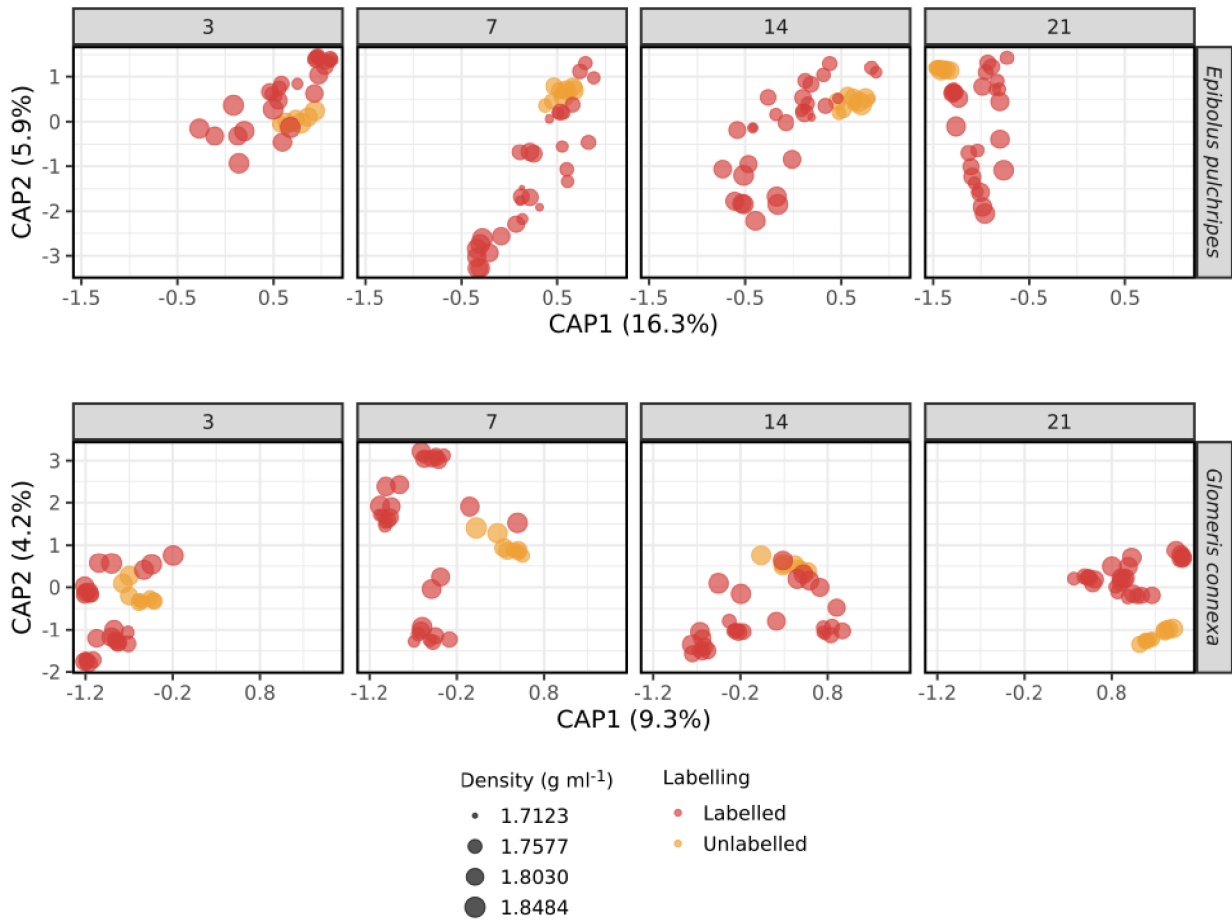


Fig. S8. Constrained principal coordinates analysis (PCoA) of Morisita-Horn dissimilarities in community composition of rRNA sequences from the SIP fractions. An ordination model using the formula:  $Dist.Mat \sim Day + Density.zone$  was calculated for each millipede species separately.

*Bacteroidia*

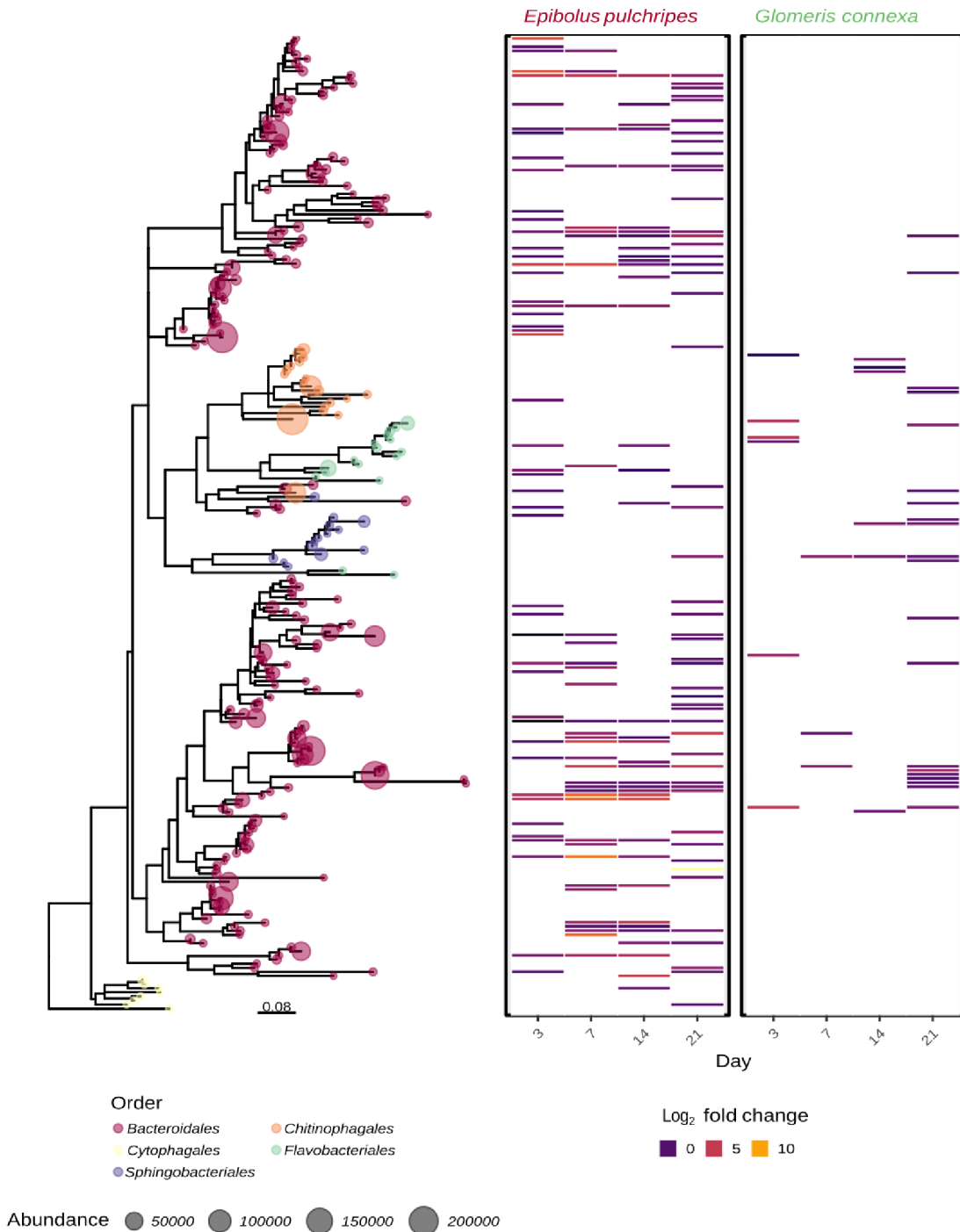


Fig. S9. A heatmap and a phylogenetic tree of ASVs from the class *Bacteroidia* (phylum: *Bacteroidota*). Each tip in the tree represents an ASV, and its circle size is proportional to the combined abundance. The tips are also colour-coded according to the order to which they are classified. Each heatmap column represents a time point, and cells of labelled ASVs are filled to represent their Log<sub>2</sub>-fold change in abundance compared to the unlabelled controls.



Bacillota

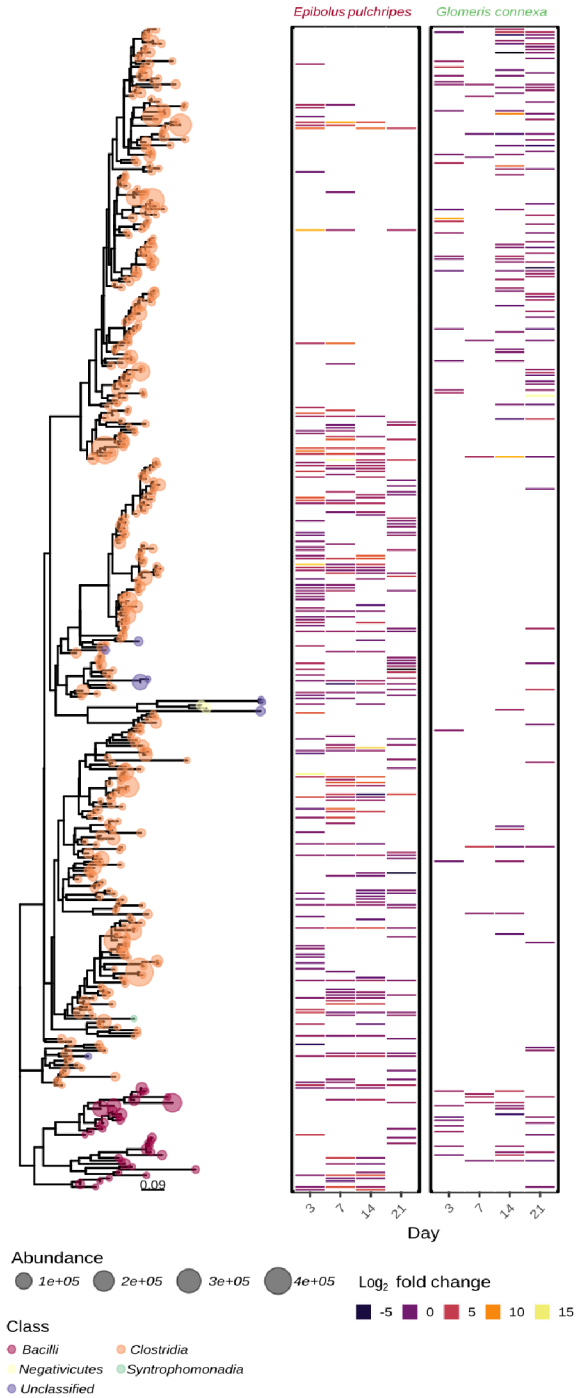


Fig. S10. A heatmap and a phylogenetic tree of ASVs from the phylum Bacillota. Each tip in the tree represents an ASV, and its circle size is proportional to the combined abundance. The tips are also colour-coded according to the class to which they are classified. Each heatmap column represents a time point, and cells of labelled ASVs are filled to represent their  $\text{Log}_2$ -fold change in abundance compared to the unlabelled controls.

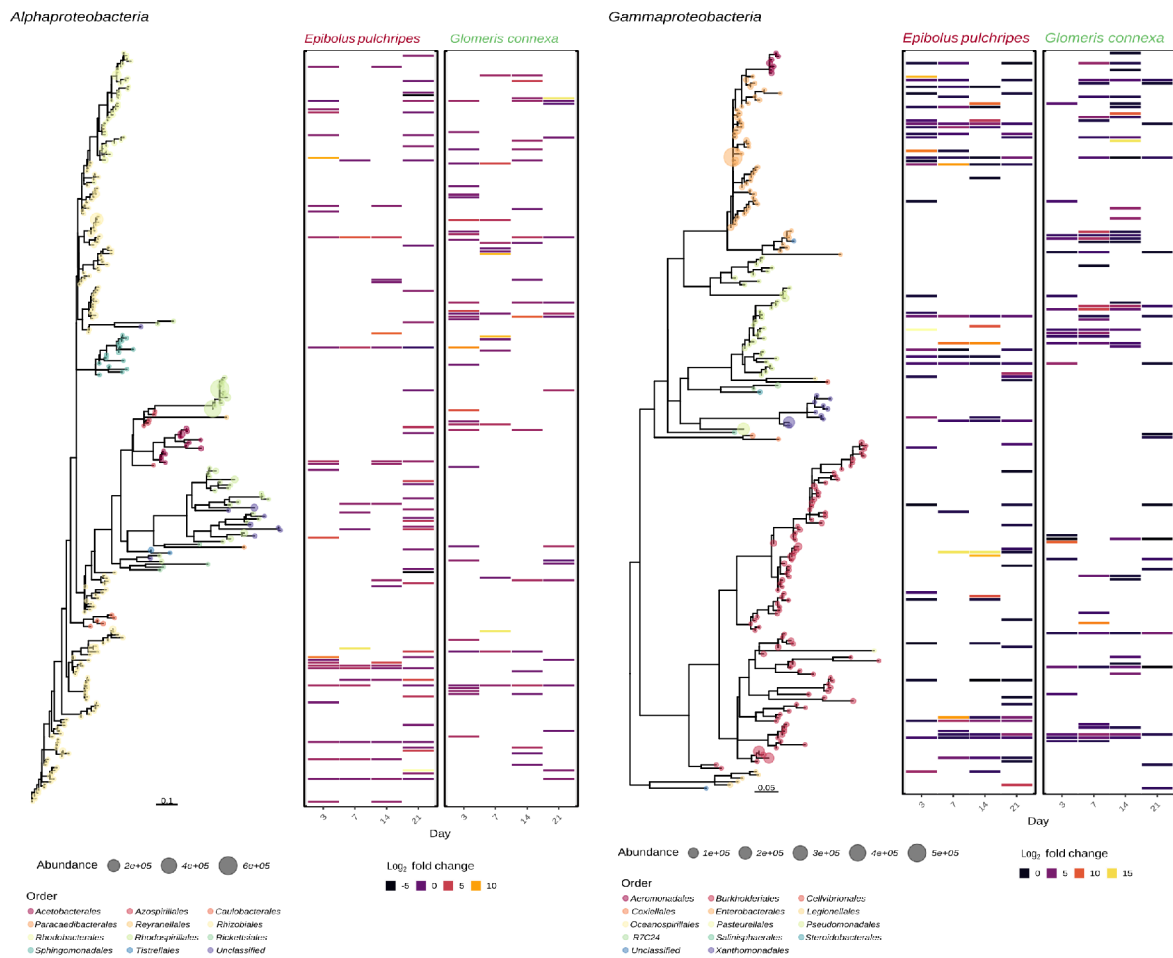


Fig. S11. A heatmap and a phylogenetic tree of ASVs from the classes Alphaproteobacteria and Gammaproteobacteria (phylum: Pseudomonadota). Each tip in the tree represents an ASV, and its circle size is proportional to the combined abundance. The tips are also colour-coded according to the order to which they are classified. Each heatmap column represents a time point, and cells of labelled ASVs are filled to represent their Log<sub>2</sub>-fold change in abundance compared to the unlabelled controls.

Actinobacteria

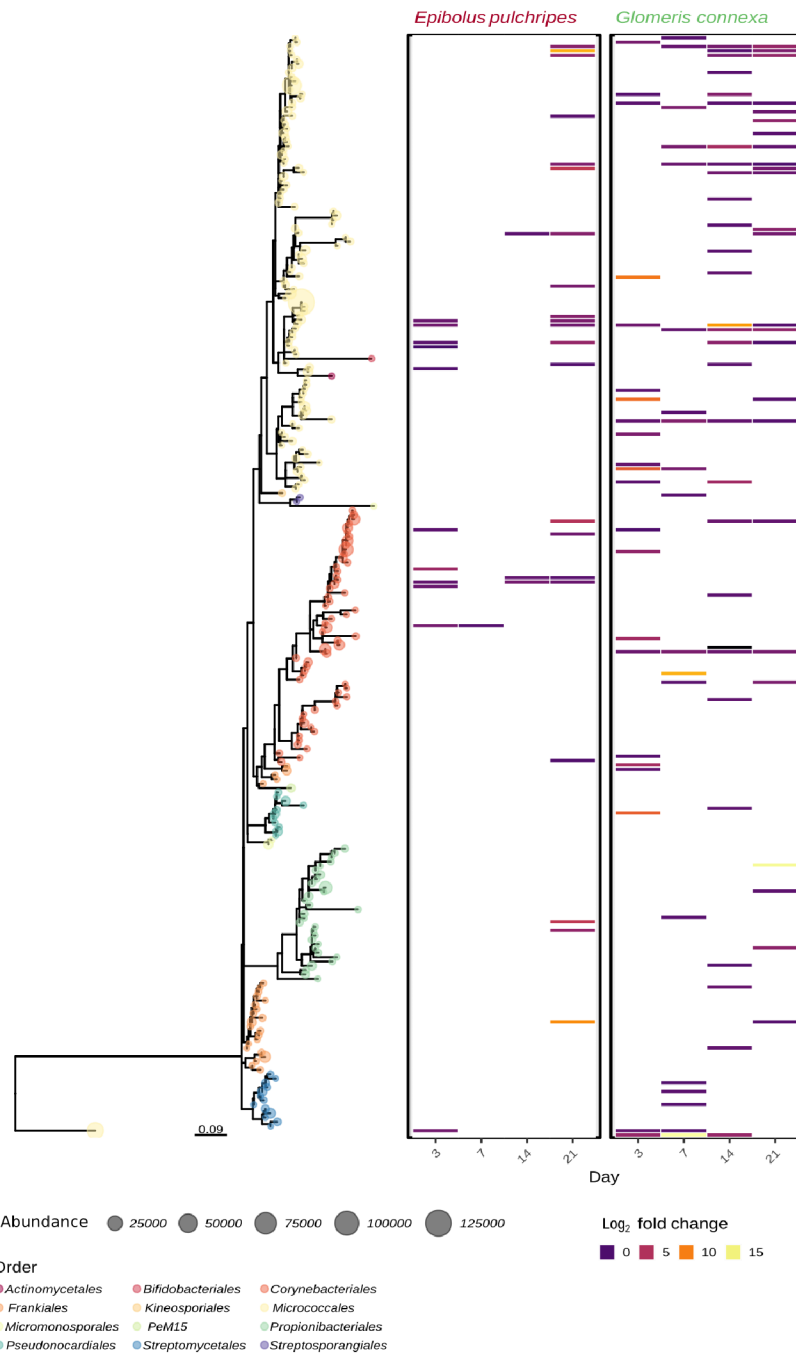


Fig. S12. A heatmap and a phylogenetic tree of ASVs from the class Actinobacteria (Actinomycetia; phylum Actinomycetota). Each tip in the tree represents an ASV, and its circle size is proportional to the combined abundance. The tips are also colour-coded according to the order to which they are classified. Each heatmap column represents a time point, and cells of labelled ASVs are filled to represent their Log<sub>2</sub>-fold change in abundance compared to the unlabelled controls.

Desulfobacterota

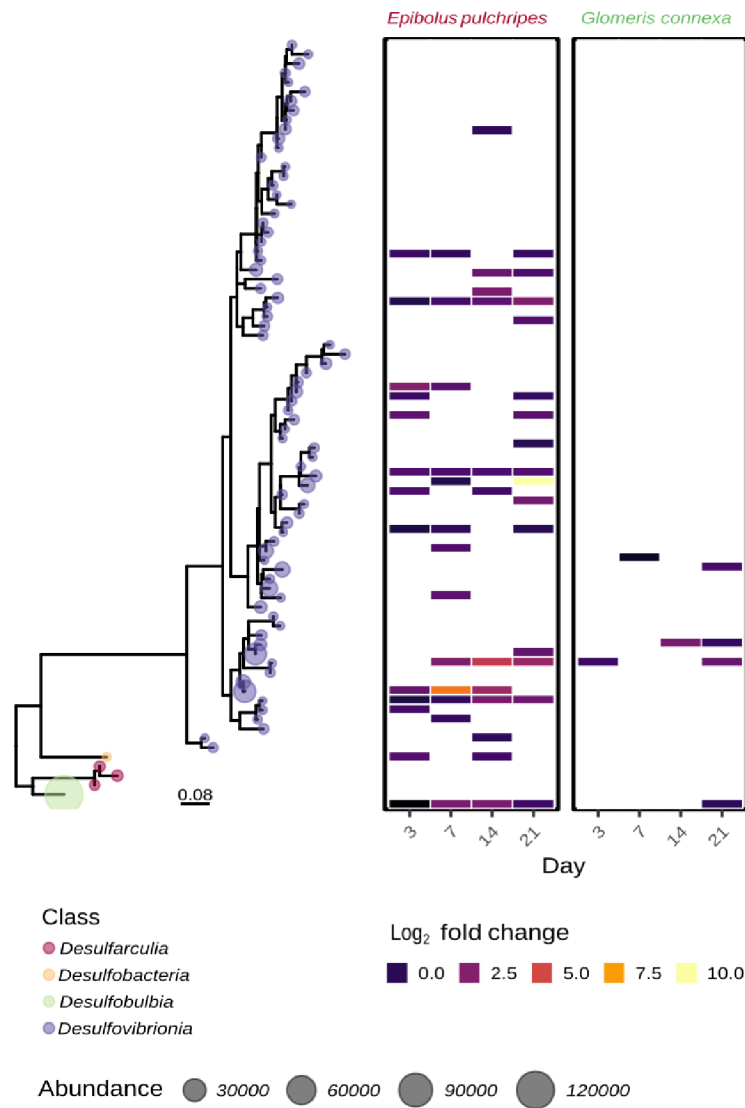


Fig. S13. A heatmap and a phylogenetic tree of ASVs from the phylum Desulfobacterota. Each tip in the tree represents an ASV, and its circle size is proportional to the combined abundance. The tips are also colour-coded according to the class to which they are classified. Each heatmap column represents a time point, and cells of labelled ASVs are filled to represent their Log<sub>2</sub>-fold change in abundance compared to the unlabelled controls.

Planctomycetota

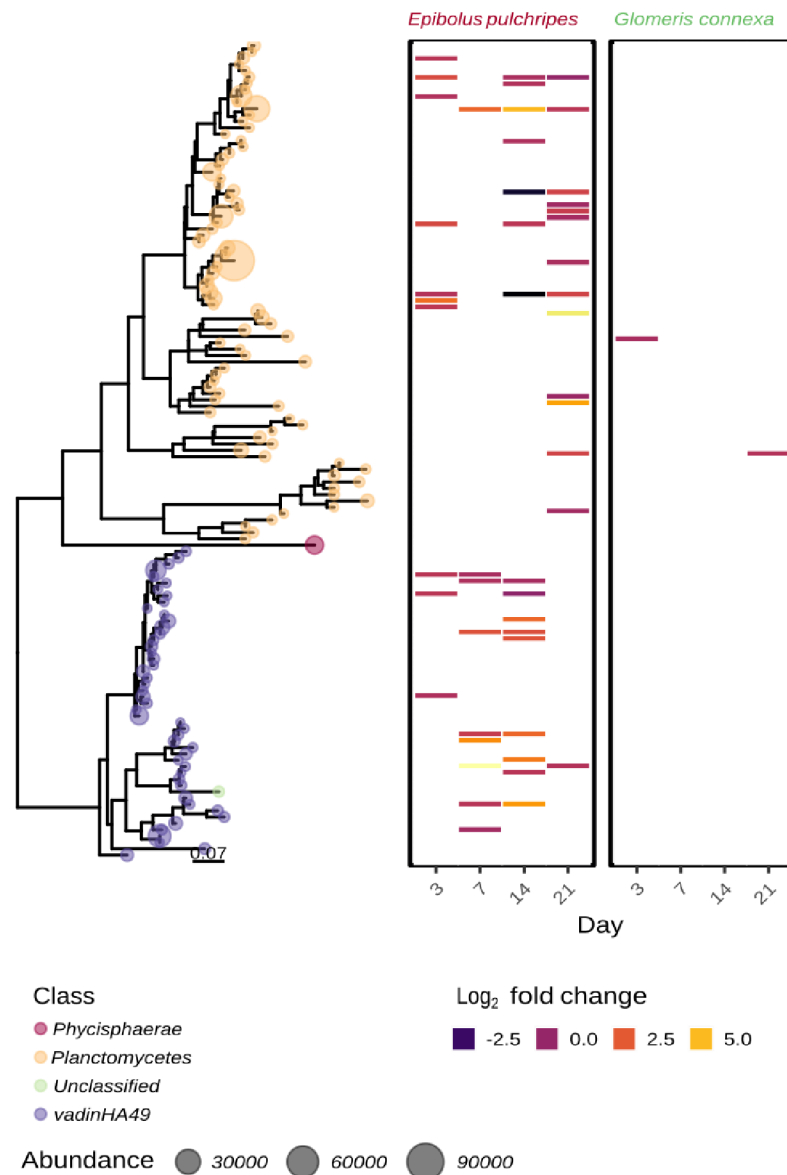


Fig. S14. A heatmap and a phylogenetic tree of ASVs from the phylum Planctomycetota. Each tip in the tree represents an ASV, and its circle size is proportional to the combined abundance. The tips are also colour-coded according to the class to which they are classified. Each heatmap column represents a time point, and cells of labelled ASVs are filled to represent their Log<sub>2</sub>-fold change in abundance compared to the unlabelled controls.

Verrucomicrobiota

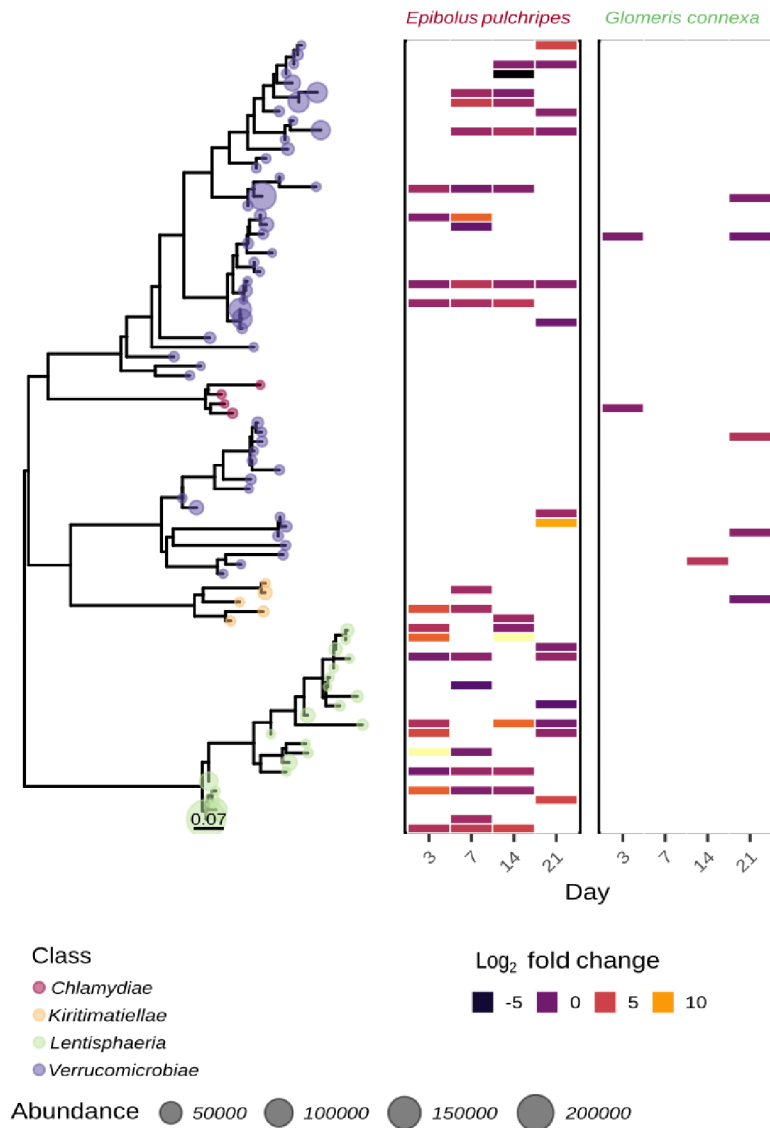


Fig. S15. A heatmap and a phylogenetic tree of ASVs from the phylum Verrucomicrobiota. Each tip in the tree represents an ASV, and its circle size is proportional to the combined abundance. The tips are also colour-coded according to the class to which they are classified. Each heatmap column represents a time point, and cells of labelled ASVs are filled to represent their Log<sub>2</sub>-fold change in abundance compared to the unlabelled controls.

## 6 Curriculum vitae



## Personal Information

**Name:** Julius Eyiuche Nweze

**Residential address:** Budivojova 1513/21 370 04, Ceske Budejovice, Czech Republic

**Permanent address:** No. 5 Ikpa Road, Olido 413105, Enugu Ezike, Enugu State, Nigeria

**Date of birth:** 13th March, 1986

**Phone Number:** +420777326352

**E-mail:** julius.nweze@bc.cas.cz, julipeale2001@gmail.com

**Researchgate:** <https://scholar.google.com/citations?user=dzEvwSIAAAAJ&hl=en>

**Google scholar:** <https://www.researchgate.net/profile/Eyiuche-Nweze>

**ORCID:** <https://orcid.org/0000-0002-2093-0438>

## Educational Background

**01/07/2019 – Present:** PhD student at the Institute of Soil Biology and Biogeochemistry, Biology Centre, Czech Academy of Sciences

**01/07/2019 – Present:** PhD student at the Department of Ecosystem Biology, Faculty of Science, University of South Bohemia in Ceske Budejovice, Czech Republic

**01/2022 -04/2022:** Guest Scientist at MPI of Terrestrial Microbiology in Marburg, Germany

**03/2016 – 02/2017:** Exchange masters' student at University of Tsukuba, Ibaraki, Japan

**2014 – 2016:** Master's Degree (M. Sc.) in Industrial Microbiology, University of Nigeria, Nsukka

**2012 – 2013:** National Youth Service Corps Certificate, Bayelsa State, Nigeria

**2007 – 2011:** Bachelor Degree (B. Sc.) in Microbiology, University of Nigeria, Nsukka, Nigeria

**2006 – 2007:** Pre-Science Programme at Enugu State University of Science and Technology, Nigeria

**2005 – 2006:** Pre-Degree Programme at Ebonyi State University, Abakaliki, Ebonyi State, Nigeria

**1998 – 2004:** Special Science School, Ukehe, Igbo-Etiti L.G.A., Enugu State, Nigeria

**1992 – 1998:** Community Primary School, Olido, Igbo Eze North L.G.A., Enugu State, Nigeria

## Publications (Research articles)

Yoshinori Murata, Charles O. Nwuche, **Julius E. Nweze**, Ifeanyi A. Ndubuisi, James C. Ogbonna (2021). Potentials of multi-stress tolerant yeasts, *Saccharomyces cerevisiae* and *Pichia kudriavzevii* for fuel ethanol production from industrial cassava wastes, *Process Biochemistry*, Volume 111(2): 305-314.

**Julius E. Nweze**, Ifeanyi Ndubisi, Nnaemeka J. Onoyima et al. (2019). Isolation and evaluation of thermotolerant-xylose fermenting yeasts for bioethanol production. *Biofuels*, 961-970.

Charles O. Nwuche, Yoshinori Murata, **Julius E. Nweze**, Ifeanyi A. Ndubuisi, Hide Ohmae, Masayoshi Saito, James C. Ogbonna (2018). Bioethanol production under multiple stress condition by a new acid and temperature tolerant *Saccharomyces cerevisiae* strain LC 269108 isolated from rotten fruits. *Process Biochemistry*.

Ifeanyi Ndubisi, **Julius E. Nweze**, Nnaemeka J. Onoyima et al. (2018). Ethanol production from cassava pulp by a newly isolated thermotolerant *Pichia kudriavzevii* LC375240. *Energy and Power Engineering*. 10(10); 457-474.

Justus Amuche Nweze, J. I. Okafor, Emeka I. Nweze and **Julius Eyiuche Nweze** (2017). Evaluation of physicochemical and antioxidant properties of two stingless bee honeys: a

comparison with *Apis mellifera* honey from Nsukka, Nigeria. BMC Res Notes, 10: 566. doi: 10.1186/s13104-017-2884-2

**Nweze Julius Eyiuche**, Shiho Asakawa, Takahiro Yamashita, Atsuo Ikeguchi, Yutaka Kitamura and Hiroshi Yokoyama (2017). Community analysis of biofilms on flame-oxidized stainless steel anodes in microbial fuel cells fed with different substrates. *BMC Microbiology*, 17:145.

Justus Amuche Nweze, Josephine I. Okafor, Emeka Innocent Nweze and **Julius Eyiuche Nweze** (2016). Comparison of antimicrobial potential of honey samples from *Apis mellifera* and two stingless bees from Nsukka, Nigeria. *J Pharmacogn Nat Prod* 2: 124.

## **Books (Published book chapters)**

Charles O. Nwuche, Shruti Gupta, Joseph Akor, **Julius Eyiuche Nweze**, Justus Amuche Nweze, and Victor U. Unah (2023). Biogas from Manure: The Future of Renewable Natural Gas and Its Implications. In: Arshad, M. (eds) *Climate Changes Mitigation and Sustainable Bioenergy Harvest Through Animal Waste*. Springer, Cham. doi.org/10.1007/978-3-031-26224-1\_7

Justus Amuche Nweze, Shruti Gupta, Joseph Akor, Charles O. Nwuche, **Julius Eyiuche Nweze**, and Victor U. Unah (2023). Animal Waste: An Environmentally Sustainable Management Approach. In: Arshad, M. (eds) *Climate Changes Mitigation and Sustainable Bioenergy Harvest Through Animal Waste*. Springer, Cham. doi.org/10.1007/978-3-031-26224-1\_1

**Julius Eyiuche Nweze**, Justus Amuche Nweze, Uche Oyeagu, Shruti Gupta, Joseph Akor, Charles Nwuche (2023). Chapter 11 - Bioremediation of pharmaceutical wastewater.: In *Development in Wastewater Treatment Research and Processes Innovative Trends in Removal of Refractory Pollutants from Pharmaceutical Wastewater*. Elsevier 1st Edition, ISBN: 9780323992787

Joseph Akor, **Julius Eyiuche Nweze**, Justus Amuche Nweze, Charles Nwuche (2023). Chapter 4 - Application of biosensors/biological assays for the analysis of contaminants present in pharmaceutical waste water: In *Development in Wastewater Treatment Research and Processes- Innovative Trends in Removal of Refractory Pollutants from Pharmaceutical Wastewater*. Elsevier 1st Edition, ISBN: 9780323992787

Justus Amuche Nweze, **Julius Eyiuche Nweze**, Somtochukwu Cecilia Agbo, Emeka Innocent Nweze and Emmanuel A. Eze (2023). Chapter 10. Some Rare and Endemic Fungal Species: The Cause of Difficult-To-Diagnose Community-Acquired Mycoses: In *The Book of Fungal Pathogens*. Nova Science Publishers. ISBN: 979-8-88697-454-6

Aparna Gunjal, Shruti Gupta, **Julius Eyiuche Nweze**, Justus Amuche Nweze (2023). Chapter 4 - Metagenomics in bioremediation: Recent advances, challenges, and perspectives: In *Developments in Applied Microbiology and Biotechnology, Metagenomics to Bioremediation*, Academic Press, 81-102, ISBN 9780323961134

**Julius Eyiuche Nweze**, Justus Amuche Nweze, Shruti Gupta (2022). Application of extremophiles in sustainable agriculture: Physiology, genomics, and biotechnological applications of extremophiles. Academic Press. ISBN13: 97817991444. Published Date: 15th January 2021

Shruti Gupta, **Julius Eyiuche Nweze**, Sharad D Subugade (2022). Marine microbial hotspots - especially related to corals: In *Microbial diversity and ecology in hotspots*. Academic Press. ISBN: 9780323901482. Published Date: 15th November 2021

## **Awards/Grants**

**01/2022 – 04/2022:** ERASMUS+ Mobility Grant (CZ CESKE01)

**2022:** Grant from Incorporation of the Biology Centre of the CAS into the European Research Area (IBERA)

**11/2016:** Outstanding poster award for a masters' student for a poster presented at the Ag-ESD International Symposium held at Tsukuba University, Japan

**11/2016:** Certificate of participation at Ag-ESD International Symposium held at Tsukuba University, Japan

**03/2016 - 02/2017:** Japan Student Services Organization (JASSO) scholarship

## **International Conference/Symposium**

**14/08 – 19/08/2022:** Bacteroidota and Proteobacteria are the major complex polysaccharide degraders in the hindguts of tropical and temperate millipedes. A poster presentation at the International Society for Microbial Ecology (ISME) held in Lausanne, Switzerland

**19/06 – 23/06/2022:** Tropical and temperate millipedes share their diet but host a very different gut microbiome. A short-talk and poster presentation at the Ecology of Soil Microorganisms held in Prague Czech Republic.

**23/05 – 27/05/2022:** Do methane and non-methane releasing millipedes depend on their gut microbiome to digest leaf litter? A short-talk at European Geosciences Union (EGU) conference held at the Austria Center Vienna (ACV) in Vienna, Austria.

**11/2016:** Isolation and evaluation of thermotolerant-xylose fermenting yeasts for bioethanol production. Poster presentation at Ag-ESD International Symposium held at Tsukuba University, Japan

## **Special Skills**

**Molecular Biology:** Total nucleic acid extraction and purification, PCR, nucleic acid quantification (qPCR and ddPCR), gel electrophoresis, primer design, FISH, CARD-FISH, DNA/RNA-SIP

**Microbial techniques:** Fermentation and extraction, anaerobic cultivation, batch/continuous cultivation, optimisation

**Spectroscopy:** UV and visible spectroscopy

**Programming/coding:** Unix (Linux), R, QIIME2, Python, Perl, HTML

**Bioinformatics:** Amplicon analysis, metagenomic and metatranscriptomic analyses

**Data science:** JSON

**Database management:** MS Access

**Graphics and design:** Inkscape, Photoshop, Acrobat, Corel Draw, Microsoft Publisher, Pix

**Maintenance:** Configuration, administration, tech support, installations, Linux/Unix

**Networking:** Configuration, security, automation, cloud management, WAN/LAN, Router, DNS, DHCP, debugging, troubleshooting

**Operating Systems:** Microsoft Windows, Linux and MacOS

**Languages:** English (fluent), Igbo (fluent), Japanese (elementary)

**Personal:** Hard working, disciplined, learning attitude, ability to work in teams as well as independently, ability to manage stress, sincere and self-confident, possess positive attitude, strong communication and interpersonal skills.



© for non-published parts Julius Eyiuche Nweze

julius.nweze@bc.cas.cz

With a little help from my friends: “understanding the roles and importance of the millipede gut microbiome”

Ph.D. Thesis, 2024

All rights reserved

For non-commercial use only

Printed in the Czech Republic by REPRO CO., v.o.s.

Edition of 5 copies

University of South Bohemia in České Budějovice

Faculty of Science

Branišovská 1760

CZ-37005 České Budějovice, Czech Republic

Phone: +420 387 776 201

[www.prf.jcu.cz](http://www.prf.jcu.cz), e-mail: [sekret-fpr@prf.jcu.cz](mailto:sekret-fpr@prf.jcu.cz)

