

# **Bacterial adaptation to host association**

## **Dissertation**

in fulfilment of the requirements for the degree  
*Doctor rerum naturalium*  
of the Faculty of Mathematics and Natural Sciences at the  
University of Kiel

Submitted by Nancy Obeng  
Department of Evolutionary Ecology and Genetics  
Zoological Institute, Kiel University  
Kiel, 2020



## **Declaration**

I, Nancy Obeng, declare that:

Apart from my supervisor's guidance the content and design of the thesis is all my own work;

Specific aspects of my thesis were supported by colleagues; their contribution is specified in detail in the following section "Contribution of authors";

The thesis has not already been submitted neither partially nor wholly as part of a doctoral degree to another examining body. Apart from the included published papers no other part of the thesis has been published nor submitted for publishing;

The thesis has been prepared subject to the Rules of Good Scientific Practice of the German Research Foundation (DFG).

Signature: \_\_\_\_\_



First referee: Prof. Dr. Hinrich Schulenburg

Second referee: Prof. Dr. Arne Traulsen

Date of oral examination: December 3<sup>rd</sup>, 2020

Signature: \_\_\_\_\_



## Acknowledgements

It has been a privilege and joy to engage in this thesis project in the last years and it would neither have been possible nor nearly as much fun without a large group of people involved.

It felt much like being part of a microbiota community characterized by mutualisms.

Most importantly, I would like to thank you, Hinrich Schulenburg, for the guidance, support, and freedom you provided as a supervisor. You always pushed me to go a step further and continue to inspire me. I also cannot thank you enough for all opportunities you have provided me with to develop.

Dear Arne Traulsen, Ute Hentschel Humeida and Michael Sieber, thank you for always lending an open and critical ear, a quick mind and calm support as my thesis advisory committee.

Many thanks also to Florence Bansept, Johannes Zimmermann, Georgios Marinos, Christoph Kaleta for fantastic, fruitful, and very fun collaborations. You have always challenged me to think differently and expand my horizons. Similarly, I would like to thank Joanna Summers, Dave Rogers, and Paul Rainey for stimulating discussions and excellent advice on wrinkly *Pseudomonads* and their genetic manipulation. Further, I would like to thank Janina Fuß and Sören Franzenburg at the CCGA, Kiel, for guiding and supporting my sequencing efforts.

I would also like to thank all former and current members of the Schulenburg group, which have provided a stimulating scientific atmosphere, great advice, and discussions, and more importantly have made the lab a lovely place to be. Aditi Batra, Agnes Piecyk, Andrei Papkou, Anke Kloock, Alejandra Zárata-Potes, Ashley Gedon, Barbara Pees, Camilo Barbosa, Carola Petersen, Christiana Anagnostou, Christina Griebner, Florian Henkies, Jack Aidley, João Botelho, Julia Johnke, Katja Dierking, Kohar Annie Kissoyan, Leif Tüffers, Lena Peters, Meike Friedrichsen, Niels Mahrt, Philipp Dirksen, Roderich Roemhild, Sabrina Butze, Sabrina Köhler, Silvia Dähn and Yang Wentao – thank you in so many ways, I feel extremely lucky to be part of this group. Melinda Kemlein and Anna Czerwinski, working together and supervising you has been truly rewarding, and I am especially proud to see you progress in troublesome Covid times. Thank you also Thekla Schultheiss, Adina-Malin Tietje, Román Zapién-Campos, Annika Tiesler and Shindhuja Joel for joining in and helping along the way.

For your interest and exciting discussions along the way, I would like to thank Brendan Bohannon, Angela Douglas and Bärbel Stecher. Thank you also Jordi van Gestel, for being a role model and for graphical inspiration on my thesis layout and Sander van Doorn, Dick van Elsas and Franjo Weissing for preparing and supporting me to embark on the journey of a PhD thesis. Venturing further, I would like thank Ute Jülly for much reflection and support.

For funding, infrastructure and the fora of people and discussions, I thank the CRC1182 and the IMPRS for Evolutionary Biology. Thank you, Cleo Pietschke, Anika Hintz, and Kerstin Mehnert for facilitating this.

Thank you also Danielle Harris, Julia Neelsen and Esther Schmidt for being wonderful scientific advisors, curious minds and more importantly friends that have closely followed the journey of this thesis.

Finally, I would like to thank my family, my parents and my sister, for always supporting me and being role models of following your passions and pursuing them diligently and fiercely. And Rudolf, for emotional support and motivation, and for being an immensely patient, yet always a demanding advisor pushing me to put in my very best in this work.





## Contents

<b>Contribution of authors .....</b>	<b>11</b>
<b>Summary.....</b>	<b>15</b>
<b>Zusammenfassung .....</b>	<b>17</b>
<b>Introduction</b>	
Bacterial life in host association.....	19
<b>Chapter 1</b>	
The functional repertoire encoded within the native microbiota of the model nematode <i>Caenorhabditis elegans</i> .....	33
<b>Chapter 2</b>	
Evolution of microbiota-host associations: the microbe's perspective .....	65
<b>Chapter 3</b>	
Selection gradients on microbial life history traits in the context of host association.....	77
<b>Chapter 4</b>	
Wrinkly formation underlies bacterial adaption to a host-associated lifestyle .....	97
<b>Chapter 5</b>	
Distinct bacterial life history strategies within the natural microbiota of <i>C. elegans</i> .....	135
<b>Epilogue</b>	
Observing bacterial strategies and manipulating host association.....	161
<b>Bibliography.....</b>	<b>173</b>
<b>Curriculum vitae.....</b>	<b>199</b>



## Contribution of authors

### Chapter 1

#### **The functional repertoire contained within the native microbiota of the model nematode *Caenorhabditis elegans*.**

*ISME Journal* (2019).

*Johannes Zimmermann\**, *Nancy Obeng\**, *Wentao Yang*, *Barbara Pees*, *Carola Petersen*, *Silvio Waschina*, *Kohar A. B. Kissoyan*, *Jack Aidley*, *Marc P. Hoepfner*, *Boyke Bunk*, *Cathrin Spröer*, *Matthias Leippe*, *Katja Dierking*, *Christoph Kaleta\**, *Hinrich Schulenburg\**

\* Shared first or senior authorship

HS, CK, JZ and NO conceived the study and wrote the original draft. HS and CK supervised the work and provided funding. JZ, WY, SW, MPH analyzed the genomic data. NO, BP, CP, KABK, and JA designed, performed, and analyzed experiments. BB, CS, MLKDCK and HS gave intellectual input. All authors discussed the data, read, and approved the final manuscript.

### Chapter 2

#### **Evolution of microbiota-host associations: the microbe's perspective.**

*Manuscript ready for submission.*

*Nancy Obeng*, *Florence Bansept*, *Michael Sieber*, *Arne Traulsen*, *Hinrich Schulenburg*

NO and HS conceived the study, reviewed the relevant literature, and wrote the original draft. FB and MS developed mathematical models. HS and AT supervised the work and provided funding. All authors provided intellectual input, read, and approved the final manuscript.

### Chapter 3

#### **Selection gradients on microbial life history traits in the context of host association.**

*Manuscript in preparation.*

Florence Bansept, Nancy Obeng, Hinrich Schulenburg, Arne Traulsen

All authors provided intellectual input. FB developed and analyzed mathematical models, performed simulations, and wrote the original draft. FB, AT and NO revised the manuscript. HS and AT supervised the work and provided funding.

### Chapter 4

#### **Wrinkly formation underlies bacterial adaption to a host-associated lifestyle.**

*Manuscript in preparation.*

*Nancy Obeng*, *Melinda Kemlein*, *Anna Czerwinski*, *Thekla Schultheiss*, *Janina Fuß*, *Hinrich Schulenburg*

NO and HS conceived the study. NO designed and performed the evolution experiment, analyzed, and interpreted data and wrote the manuscript. NO, AC, TS and MK performed phenotypic analyses of evolved bacteria. JF and the CCGA, Kiel provided material and sequencing services. HS supervised the project and provided funding.

## Chapter 5

**Distinct bacterial life history strategies within the natural microbiota of *C. elegans***  
*Manuscript in preparation.*

*Nancy Obeng, Hinrich Schulenburg*

NO and HS conceived the study. NO designed and performed the evolution experiment and phenotypic analyses of evolved bacteria, analyzed and interpreted data, wrote image analysis macro and wrote the manuscript. HS supervised the project and provided funding.

As supervisor I confirm the above stated contributions

Signature: \_\_\_\_\_





## Summary

Microorganisms are commonly found living in multicellular organisms. Taxonomically, individual symbionts and communities of microbes, microbiotas, of a variety of animals and plants have been well-described and it has become clear that they can have profound effects on their hosts' biology. The bacterial side of the association, however, has received comparatively little attention. In particular how microbes adjust to life with a host, and what drives the emergence and maintenance of microbe-host associations remains understudied.

In this thesis, my main objective is therefore to improve our understanding of the traits and life history strategies that facilitate microbe-host associations. To that end I combine experimental study of bacterial isolates from the natural microbiota of the nematode *Caenorhabditis elegans* as a model system with predictions from mathematical models.

In the first part of the thesis I focus on the functional ecology of the natural *C. elegans* microbiota. Integrating whole genome sequencing of bacterial isolates, metabolic modeling, and experiments, we found that the microbiota can synthesize all amino acids and vitamins essential to the worm. Further, we demonstrate that their predicted traits and metabolic repertoires lead to distinct life history strategies, shape interactions within the microbiota as well as worm colonization and worm population growth.

The second part of the thesis addresses the evolution of bacteria in host-association. The first step is the development of the concept of a biphasic life cycle as an evolutionary intermediate between a free-living and a host-associated lifestyle. At this transition, matrix population models predict that bacteria can optimize fitness either by increasing replication rates or by modulating migration rates between host and environment. Experimentally evolving the microbiota isolate *Pseudomonas lurida* MYb11 in a biphasic life cycle with *C. elegans*, we find that a key step to associating is to simply stick to the host. Specifically, we observe the evolution of wrinkly colony types that show improved persistence in worms and a greater ability to form biofilms than the ancestral MYb11. This goes hand in hand with genetic mutations in regulators of the universal second messenger cyclic di-GMP, previously linked to the life history transition from motile to sessile and stress-tolerant. These findings highlight that bacteria can adapt to a host that they encounter periodically during a biphasic life cycle. Finally, I extend the focus to a two-member microbiota of MYb11 and *Ochrobactrum vermis* MYb71 and study interactions between these species and differences in their life history strategies when co-associating in *C. elegans*. This confirms that interactions within the microbiota are context-dependent and shift along the stages of the biphasic life cycle. Further, it demonstrates that different life history strategies co-exist in the microbiota, with the one centered on stress-tolerance being the most advantageous.

Overall, this thesis emphasizes the importance of bacterial traits and life history to shaping life in association with a host. It thereby provides fundamental insights into the driving forces and possibly origin of microbe-host associations.





## Zusammenfassung

Mikroorganismen leben oft in Assoziation mit mehrzelligen Organismen. Einzelne Symbionten und komplexere Mikrobengemeinschaften (Mikrobiotas), die einen starken Einfluss auf das Leben ihres Wirtes haben können, sind für viele Tiere und Pflanzen gut in ihrer Artenzusammensetzung beschrieben. Im Gegensatz dazu ist vergleichsweise wenig dazu bekannt welchen Einfluss das Leben im Wirt auf die Mikroben selbst hat. Insbesondere die Frage, wie sich Mikroben an das Leben mit einem Wirt anpassen und was die Entstehung und Aufrechterhaltung von Mikroben-Wirt-Assoziationen ermöglicht, ist weitgehend unerforscht. In dieser Doktorarbeit besteht mein Hauptziel darin, unser Verständnis der mikrobiellen Merkmale und lebensgeschichtlichen Strategien zu verbessern, die Mikroben-Wirt-Assoziationen ermöglichen. Zu diesem Zweck kombiniere ich die Experimente mithilfe des Modellsystems aus dem Nematoden *Caenorhabditis elegans* und Bakterienisolaten aus seiner natürlichen Mikrobiota mit Vorhersagen aus mathematischen Modellen.

Der erste Teil dieser Arbeit beschäftigt sich mit der funktionellen Ökologie der natürlichen Mikrobiota von *C. elegans*. Mithilfe einer Kombination aus genomischer Sequenzierung einer Großzahl an Bakterienisolaten aus der Wurmmikrobiota, metabolischer Modellierung und Laborexperimenten konnten wir zeigen, dass die assoziierten Bakterien gemeinsam alle für den Wurm essenziellen Aminosäuren und Vitamine synthetisieren können. Darüber hinaus führen die unterschiedlichen, im Genom angelegten Merkmale und metabolischen Profile zu unterschiedlichen Lebensstrategien und Bakterieninteraktionen innerhalb der Mikrobiota und beeinflussen die Besiedlung und das Populationswachstum des Wurmwirtes. Im zweiten Teil der Arbeit liegt der Fokus auf der Evolution von Bakterien in Wirtassoziation. Zunächst stelle ich das Konzept eines biphasischen Lebenszyklus vor, der ein evolutionäres Zwischenstadium zwischen einem freilebenden und einem wirtassoziierten Lebensstil darstellt. Ein darauf abgestimmtes Matrixpopulationsmodell sagt vorher, dass Bakterienpopulationen innerhalb eines solchen Lebenszyklus ihre Fitness entweder durch erhöhte Replikationsraten oder eine Veränderung der Migrationsraten zwischen Wirt und Umwelt optimieren können. Gleichzeitig konnten wir im Rahmen eines Evolutionsexperimentes mit dem Mikrobiotabakterium *Pseudomonas lurida* MYb11 beobachten, dass das Verbleiben im Wirt ein Schlüsselschritt in der Wirtsanpassung ist. Nachdem wir MYb11 im Labor an einen biphasischen Lebenszyklus mit *C. elegans* als Wirt haben anpassen lassen, entstanden aufgefaltete Kolonietypen, die besser im Wurm verbleiben und stärker Biofilme bilden als das ursprüngliche Bakterium. Diese neuen Kolonietypen, die mit genetischen Mutationen in Regulatoren des sekundären Botenstoffs cyclic-di-GMP einhergehen, scheinen daher eine Anpassung an einen sessilen, stress-toleranteren Lebensstil im Wurm zu sein. Dies zeigt, dass schon ein biphasischer Lebenszyklus, in dem sich Bakterium und Wirt nur periodisch treffen, für die bakterielle Anpassung an den Wirt ausreicht. In Experimenten mit einer simplen Mikrobiotagemeinschaft aus MYb11 und einem weiteren Mikrobiotaisolat, *Ochrobactrum vermis* MYb71, untersuche ich weiterhin die Kontextabhängigkeit von Bakterieninteraktionen und die Diversität von Lebensstrategien in der Wurmmikrobiota. Dies zeigt, dass Bakterien mit unterschiedliche Lebensstrategien zwar koexistieren können, sich eine stresstolerante Strategie jedoch am vorteilhaftesten ist.

In ihrer Gesamtheit unterstreicht diese Doktorarbeit die Bedeutung bakterieller Merkmale und Lebensstrategien in der Anpassung an einen Wirt. Sie liefert damit grundlegende Erkenntnisse über die treibenden Kräfte und den möglichen Ursprung von Mikroben-Wirt-Assoziationen.



# **INTRODUCTION**

---

## **Bacterial life in host association**



Invisible to the naked eye exists a strikingly complex world of microbes living with multicellular organisms. Modern sequencing approaches have revealed that bacteria, a large group of microorganisms, can be found associating with animals and plants across biospheres, including the gut of tiny nematode worms living in backyard composts. Often communities of bacteria from different taxa, so-called microbiotas, reside within and on hosts. Yet, while we know a lot about the bacterial taxa found in different hosts, and appreciate that they can shape how animals and plants develop, grow and even reproduce, much less is known about the life and evolution of bacteria in host association. In this dissertation, I therefore take the bacterial perspective and investigate bacterial function and adaptation to life inside a host.

This first chapter introduces microbe-host associations, highlighting what we have yet to learn about bacterial adaptation and strategies to cope with life in a host. It further discusses why the natural microbiota of the nematode *Caenorhabditis elegans* is a useful model system to investigate bacterial life history. I then conclude with a brief summary of aims for each thesis chapter and spanning from microbiota function (**chapter 1**) to theoretical and experimental exploration (**chapters 2-5**) of microbiota evolution.

## **Living together – an introduction to microbe-host associations**

The diversity of life on Earth is paralleled with a myriad of interactions between organisms of different kinds. Organisms sharing time and space can shape each other's lives greatly, as seen in clownfish inhabiting anemones, spiders spinning webs on plants or ants tending aphids (Litsios et al., 2012; Stadler and Dixon, 2005; Styrsky, 2014). Such living together of organisms with different names has been called symbiosis since de Bary popularized this term describing lichens, a union of fungi and cyanobacteria or algae, in the late 1870s (De Bary, 1879; Oulhen et al., 2016). It usually consists of a host and a hosted organism; the latter being called the symbiont. Notably, this ancient phenomenon even precedes the origin of animals and plants, as eukaryotic cells are assumed to be the outcome of a bacterium living inside an archaeon (Sagan, 1967). Today, we are uncovering novel associations as we proceed to study life invisible to the naked eye. Microbes, including bacteria, which I focus on in this thesis, are cosmopolitan and thus found in association with organisms across all the world. This includes mussels near deep-sea hydrothermal vents (Picazo et al., 2019), the roots of plants in soil below our feet (Berendsen et al., 2012), or the inside of our guts (The Human Microbiome Project Consortium et al., 2012). With technologies to visualize organisms and sequence their genetic material, it has become clear that microbes inhabit most animals and plants, where they can be found on the surface, inside hosts or intracellularly (McFall-Ngai et al., 2013). Some organisms have intimate relationships with primarily single bacterial species, but often many different bacterial taxa can be found in a host. Following the early study of mostly one-to-one symbioses, high throughput sequencing technologies have allowed detecting such microbiotas in the 21<sup>st</sup> century. From this we have learned that host organisms may host specific microbiotas that vary in diversity and are usually distinct from the environmental microbial communities (e.g. (Johnke et al., 2020; Thomas et al., 2016)). Notably, with the technological advances, an important semantic distinction was made between the microorganisms associated with a host, the microbiota, and the microbiome, being their genetic content (Berg et al., 2020; Egan et al., 2020). Previously used for both and still favored by many, the term microbiome may refer to the organismal community, too (Lederberg and McCray, 2001). In this thesis, I will stick to the microbiome-microbiota distinction, and will further treat one-to-one and many-to-one microbe-host associations as a continuum rather than distinguishing symbiosis and microbiota research per se.

### *Interactions between microbes and their hosts*

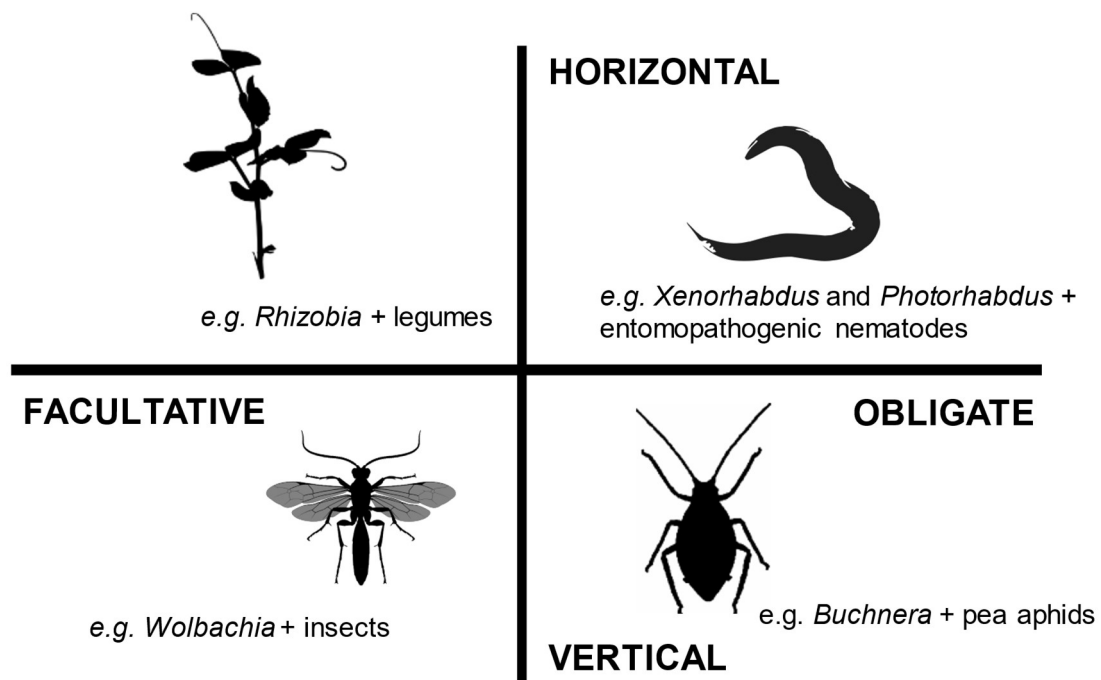
Microbes may vary in their impact on the host and are of key importance to host biology by providing a variety of services. Different types of interaction may be classified based on the net impact a symbiont has on its host. When the symbiont provides a benefit to the host, it is considered a mutualist. In case of no detectable difference to the host it is a commensal, and when there is a cost to the host, the symbiont can be a parasite or pathogen.

In beneficial associations, of main interest in this thesis, there are numerous ways in which bacteria may provide services to their hosts. Firstly, bacteria may supply and exchange nutrients with the host. Bacteria may act as food themselves, as providers of (essential) nutrients and break down substrates their hosts could otherwise not use (Consuegra et al., 2020; Flint et al., 2012; Reis et al., 2020; Salem et al., 2017; Shukla et al., 2018a; Zhu et al., 2011). Plant-sap feeding insects and their symbionts illustrate that the exchange can go both ways, as here bacteria trade sugars with amino acids required by the host (Douglas, 2018). Secondly, associated microbes can protect their hosts against pathogens. In the simplest case via colonization resistance, i.e. by occupying a niche an incoming pathogen would otherwise have invaded (Stecher and Hardt, 2008). Further, symbionts may provide direct protection against harmful organisms via toxin production, e.g. antibacterial compounds, or other mechanisms of bacterial competition (e.g. (Ford et al., 2016; Kissoyan et al., 2019)). In addition to protecting against pathogenic organisms, associated microbes may neutralize toxins, and degrade or modify incoming chemicals of potential harm (Wang et al., 2020). Another, quite flamboyant, service is the provision of luminescence by bacterial symbionts. Most famously described in counter-illumination provided by *Vibrio fischerii* to the Hawaiian bobtail squid (Nyholm and McFall-Ngai, 1998), but also produced by *Cryptopsaras couesii* in deep-sea angler fishes (Baker et al., 2019), bacterial metabolism may output light that in turn provides a possible selective advantage to the respective host. Finally, host biology from immunity, to development and reproduction, can be fundamentally impacted by the associated microbes (see (Douglas, 2018) for a synthesis). Therefore bacteria might to a certain extent be considered to extend the host's phenotype, linking to Richard Dawkins' idea of the extended phenotype (Hunter, 2018). A similar line of thought has led to the idea that host and associated bacteria should be considered as one unit, the holobiont or metaorganism, and possibly even under selection as such (Bosch and Miller, 2016; Rosenberg and Zilber-Rosenberg, 2016; Zilber-Rosenberg and Rosenberg, 2008). Next to debates about the realism and value of considering hosts and microbes as a functional unit (Bordenstein and Theis, 2015; Douglas and Werren, 2016), these concepts highlight an anthropocentric focus on the host perspective that manifests in a bias of research about microbial influence on hosts unparalleled by the inverse. Although generally helpful to distinguish interaction types, it becomes difficult to infer the impact of individual microbiota members on a host in many-to-one associations. This is due to possible changes in the influence of bacteria in the presence of other species, as well as possible emergent effects the microbiota community may have on the host (Mushegian and Ebert, 2016). A bottom-up approach to understanding the interactions between host and microbiota, is therefore to study individual symbiont taxa in host association followed by investigation of successively more complex communities.

From the bacterial perspective, the host can be considered a habitat defined by a certain set of abiotic and biotic characteristics. It importantly provides spatial niches, which bacteria may colonize and in which they proliferate (33). Although bacteria can be found on most organismal epithelia, the gut microbiota is often the focus of research, as this is where the majority of associated bacteria reside in humans and where the most influential interactions with the host appear to take place (34). In the gut, abiotic challenges include gradients of oxygen supply and pH, mechanic digestion and peristalsis (33,35). Further, bacteria are exposed to digestive enzymes and host immune effectors. In this habitat, spatial structure, shaped by host crypts and microvilli for example, and temporal variation, introduced by peristalsis or circadian rhythms, present higher order abiotic challenges. In the digestive tract, on other host surfaces, or within host cells, bacteria may proliferate on

nutrients acquired through host feeding, or exchange nutrients with the host and other symbionts (36). Additionally, hosts may also serve as a vector for bacterial dispersal otherwise inaccessible to microbes of limited size and motility (Chase et al., 2019; Diaz and Restif, 2014; Moeller et al., 2018).

Classic symbiosis research has provided a general framework to classify the interactions between host and symbiont including the strength of association between the partners, the transmission or inheritance of symbionts, as well as the impact of the partners on each other.



**Figure 1** *Microbe-host associations along gradients of association strength and inheritance.* Examples microbe-host associations include: rhizobia and legumes (Beringer et al., 1979), *Xenorhabdus* and *Photorhabdus* associated with entomopathogenic nematodes (Ciche et al., 2006), *Wolbachia* and insects (Werren et al., 2008) and *Buchnera* and pea aphids (Douglas, 2018). This figure was inspired by (Fisher et al., 2017). Images of hosts were collected from phylopic.org.

The strength of association between host and symbiont(s) ranges from sporadic encounters to complete reliance on each other (Figure 1). The closest form of living together are obligate symbioses, in which neither partner can live without the other. Classic examples of this are bacterial *Buchnera* symbionts in pea aphids (Russell et al., 2013) as well as *Sulcia* and *Nasuia*, obligately associated with the leafhopper *Macrostelus quadrilineatus* (Mao et al., 2018), all endosymbionts living inside host cells. On the other end of the spectrum lie facultative symbioses. In these symbionts and hosts can live both in association as well as independently, as is the case for leguminous plants and rhizobia bacteria (Beringer et al., 1979).

Further, microbe-host associations differ in their mode of establishment and inheritance. They may be acquired from an environmental source anew at every host generation, i.e. horizontally acquired, or transmitted directly from host to host, and thus vertically acquired. Commonly, facultative symbionts transmit horizontally, while obligate symbioses are stabilized via vertical transmission. This appears to be the case for bacteria found in rotting plant material that facultatively colonize the gastro-intestinal tract of the model nematode *Caenorhabditis elegans* (Dirksen et al., 2016), and the insect endosymbiont *Buchnera* (Douglas, 2018), respectively. Burying beetles that use vertebrate carcasses as a

habitat and food source for offspring have a mixed transmission strategy (Shukla et al., 2018b; Wang and Rozen, 2017). During parental care, oral and anal secretions are transferred to larvae and the carcass, thereby being transmitted vertically. Additionally, larvae can acquire bacteria from the carcass horizontally. In these cases, it is especially important to consider the free-living environment as an important alternative habitat. For this, the concept of nested ecosystems has been proposed, which places the microbiota inside a host, yet inside the external environment (Douglas, 2018; McFall-Ngai et al., 2013). Importantly, the strength of the association and the mechanism of inheritance do not always align (Fisher et al., 2017). Notably, the obligate symbioses between *Photorhabdus* or *Xenorhabdus* and entomopathogenic nematodes (Ciche et al., 2006), or between sulfur-oxidizing bacteria and the giant tube worm *Riftia pachyptila* are established horizontally (Dubilier et al., 2008; Fisher et al., 2017), while *Wolbachia* symbionts of a variety of different arthropod and even nematode hosts may be vertically inherited, yet associates only facultatively (Werren et al., 2008). Notably, the position of a given microbe along gradients of association strength and routes of transmission can shape the traits, life history and evolutionary trajectories taken.

### *Bacteria-bacteria interactions*

In addition to abiotic challenges and interaction with the host, bacteria usually encounter other microbes in the host. Within these consortia of associated microbes, bacteria can interact with each other in different ways. During cross-feeding for example, one bacterium may take up and grow on the metabolic side-products of another bacterium, as in pectin degradation by *Bacteroides* from the human colon or symbionts of reed beetles (Luis et al., 2018; Reis et al., 2020). Sharing similar substrates, bacteria may also compete for resources in the host, directly or indirectly (Coyte and Rakoff-Nahoum, 2019). Again, these biotic interactions may be classified according to their impact on the participating species (Douglas, 2018). A mutualism occurs when both bacteria benefit from the interaction and antagonism or competition when neither does. In case one partner is gaining a benefit at the cost of the other, we speak of parasitism. Further, commensalism and amensalism describe the cases, in which one partner is unaffected, while the other receives a benefit or has a cost from the interaction. To infer these interactions, co-occurrence networks can be inferred from 16S rRNA or metagenomic sampling of microbiomes (Faust et al., 2012; Johnke et al., 2020). Repeated co-existence would be interpreted as positive, and complementary absences as negative interactions. Yet, what determines these patterns often remains unanswered from co-existence networks. A better understanding of bacterial functional repertoires should improve their biological interpretation and prediction, as shared substrates for growth for example, might provide a mechanistic explanation for co-existence.

### **Responding to the life in the host**

To proliferate, or at least survive, bacteria may respond to the challenges faced inside the host phenotypically or evolutionarily. One of the fastest responses bacteria can mount is metabolic. Without a change in gene expression, cells may adjust their physiology to the nutrient environment. Depending on the enzymes and pathways at their disposal, bacteria may thus shift their physiological state inside a host to use the available substrates. Whole genome or metagenomic sequencing can provide insights into the possible metabolic functions bacteria in a given microbiota encode. Such an approach has indicated that *Poribacteria*, symbiotic to sponges, should be able to digest their host's extracellular matrix (Kamke et al., 2013). Furthermore, symbionts of reed beetles were predicted to provide amino acids and riboflavin (vitamin B2) to sap feeding larvae, and aid adult beetles feeding on specific plants in the digestion of pectin (Reis et al., 2020). Predictions exceeding the catabolism of single substrates are possible with system-level reconstructions of metabolic



networks (Bordbar et al., 2014). These can be used to predict bacterial metabolism in a given environment taking into account their overall metabolic constraints. Using flux balance analysis, a constraint-based modeling technique, it is further possible to predict nutrient exchanges and growth of a cell given its metabolic network, including bacterial metabolism in symbiosis (Ankrah et al., 2017, 2020; Basile et al., 2020; Orth et al., 2010).

For other cellular responses tailored to the experienced environment, bacteria need to sense the conditions first to then act accordingly. Two component systems (TCS) allow exactly that. A sensor histidine kinase usually detects an environmental stimulus and relays this input to a specific cytoplasmic response regulator via phosphorylation (Capra and Laub, 2012; Hoch, 2000). In turn, signaling cascades, often affecting multiple downstream targets, are kicked off, allowing the bacterium to respond to the stimulus. Such TCSs are both common and diverse within and across bacteria (Galperin, 2005). Notably, they are also involved in mediating a life history transition from motile to sessile in bacteria (Bantinaki et al., 2007; Hengge, 2020; Prüß, 2017). As I will discuss in **chapters 2** and **4** in more detail, this is a key transition allowing bacteria to stay associated with a host in contrast to a free-living lifestyle. In the long term, bacteria may adapt to a host-associated lifestyle evolutionarily. Growing clonally by cell division, one of the major modes of genetic change is via *de novo* genomic mutations (incl. point mutations, insertions, deletions, frameshift mutations, as well as duplications and genomic rearrangements). Additionally, bacteria may exchange genetic material with other bacteria via horizontal gene transfer (e.g. via plasmids, phages, or other mobile genetic elements; see e.g. (Botelho and Schulenburg, 2020; Frazão et al., 2019; Jahn et al., 2019)). Next to changes in the genome, bacteria may inherit biochemical gradients or protein concentrations that determine their physiological state, and thus response function (e.g. (Ackermann, 2015; Kordes et al., 2019)). Such changes experienced by the individual cell can improve offspring to better match life inside a host. In the context of host adaptation, most research has been focusing on understanding the evolutionary fit between pathogens and their hosts (Chaguza, 2020; Otto, 2014; Sheppard et al., 2018). More recently, however, beneficial microbe-host associations have gained interest. Comparative as well as release-and-resequence approaches have started uncovering genes more commonly found in host-associated bacteria than those isolated from other sources, including genes involved in adhesion (e.g. type IV pili), virulence (e.g. type 6 secretion systems), immune evasion (e.g. capsular polysaccharides) and metabolic pathways (e.g. carbohydrate digestion, nitrate respiration) (Burghardt et al., 2018; Kwong et al., 2014; Martino et al., 2016; Newell et al., 2014; Winans et al., 2017; Zhang et al., 2020). Further, experimental evolution, a controlled approach of understanding bacterial change under defined conditions, has proven useful to understand bacterial responses to diverse selection pressures in the lab (Ford et al., 2016; Jansen et al., 2015; Lebov et al., 2020; Martino et al., 2018a; Remigi et al., 2019; Robinson et al., 2018; Soto et al., 2014). Few studies have leveraged this approach to study bacteria in host association (Hoang et al., 2016). Thus far, serial passaging experiments of *Shewanella oneidensis* and *Aeromonas veronii* in zebrafish (Lebov et al., 2020; Robinson et al., 2018), *Lactobacillus plantarum* in fruit fly hosts (Martino et al., 2018a), and *Vibrio fischerii* the light organs of bobtail squid (Soto et al., 2012, 2014), have shown that generally, bacteria improve host colonization over time, going hand in hand with changes in the motile-sessile continuum depending on the specific host, or changes in substrate use. Unfortunately, the insights gained are limited to few bacterial clones from evolving populations of bacteria and thus may miss important variation at the true scale of complexity of the evolved microbiota. From experiments passaging *Escherichia coli* in mice, it has become evident, that within host-evolution is indeed accompanied by clonal diversification (Frazão et al., 2019; Sousa et al., 2017), but to what extent this has consequences for overall strategies bacteria may use to associate with the host, remains diffuse. Evolution experiments tracking microbiota communities in *Arabidopsis thaliana* and *Brassica rapa*, have shown that whole communities may also adjust to living with the plant host (Panke-Buisse et al., 2015). Changes in taxonomic composition accompanied a functional shift in the microbiota, but it remains unclear if and how individual microbiota members adapt over time. Thus, we need to both consider the

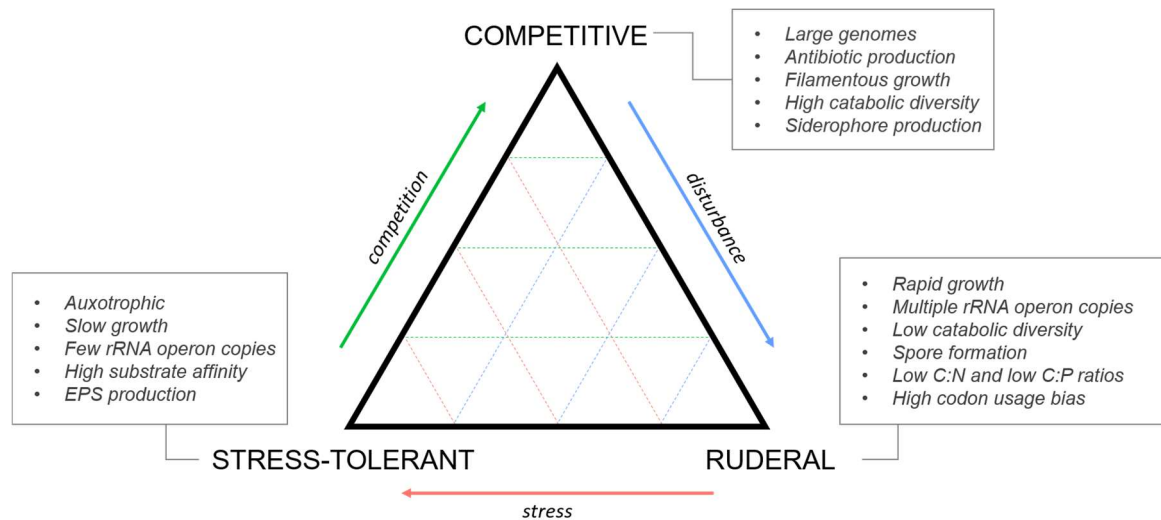
relevant scale(s) of the evolving microbiota (Gorter et al., 2020), be it bacterial populations or communities, to understand how individual members respond in the complexity of the microbiota. Moreover, these studies usually focus on the host only, and thus neglected that bacteria in most symbioses are not only nested in but also exposed to an alternative, free-living environment. Thus, to provide a more holistic understanding of host association given the complete set of environments, or life cycle, experienced, I develop the concept of a biphasic symbiosis life cycle in this thesis (**chapter 2**). I then go on and show predictions (**chapter 3**) and tests of its impact on bacterial responses to life with a host both at an ecological as well as evolutionary time scale (**chapters 4 + 5**). For this life history theory provides a helpful framework to understand how microbe-host associations emerge and are maintained.

## Understanding bacterial life histories

On the way to a more holistic understanding of host-associated microbes, we have to move from viewing microbiotas as communities of different taxa towards groups of organisms with shared life histories (Fierer, 2017). To that end life history theory provides helpful concepts that may be applied to understand bacterial evolutionary ecology better. Life history theory is concerned with ways in which organisms balance reproduction and survival, two key components of fitness, from birth to death. This includes the strategies organisms may use to optimize the trade-offs between the two in a given environment and the trait correlated with this. Over time, a variety of concepts have been proposed to describe and distinguish different life history strategies.

The r/K selection theory is a classic and well-known example, which assumes that organisms fall on a continuum between two extreme reproductive strategies (MacArthur and Wilson, 1967; Pianka, 1970). Ranging from r-selected organisms that invest into producing many offspring yet have a short life expectancy, to K-selected organisms with few offspring but a high life expectancy. This theory on the trade-off between reproduction and survival has been widely applied to animals and plants. In unicellular organism reproducing clonally, such as bacteria, this trade-off manifests between growth in terms of cell divisions and survival, maintenance and resilience of cells (Hengge, 2020). Similarly, this juxtaposes quantitative and qualitative strategies.

The Universal Adaptive Strategy Theory (UAST) assumes that environments are characterized by competition, stress and disturbance and therefore, partitions adaptive strategies into competitive, stress-tolerant and ruderal (Grime, 1974, 1977). Organisms with a competitive strategy are characterized by their ability to outcompete other organisms for resources or space, stress-tolerators can survive and persist under adverse abiotic conditions or low resources, while ruderals are apt at taking over a novel niche or uncolonized habitat in which they grow rapidly (Grime, 1974, 1977). Initially proposed for the study of plants, and in that context called CSR framework (Grime, 1977), it has more recently proven valuable in describing microbial strategies in response to climate change (Evans and Wallenstein, 2014; Fierer, 2017; Malik et al., 2020; Wallenstein and Hall, 2012). In this thesis, I move forward and compare bacterial strategies in host association in light of the UAST framework.



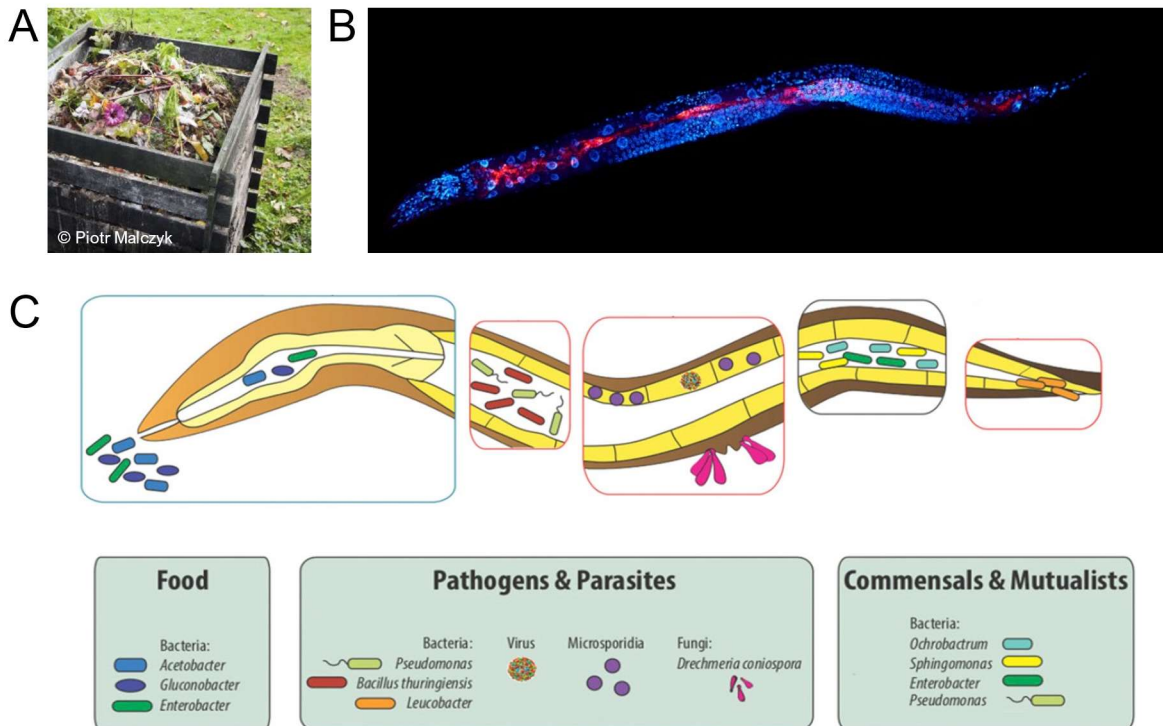
**Figure 2 Bacterial life history strategies according to the Universal Adaptive Strategy Theory.** Following Grime, organisms may follow three adaptive strategies, namely competitive, stress-tolerant and ruderal depending on the environment experienced. Here gradients of competition, stress and disturbances shape the corresponding strategy or mix of strategies. Their contributions are also highlighted by contours within the triangular model of bacterial life history strategies (stress = red, competition = green, disturbance = blue). Traits contributing to the different strategies that can be inferred from genomic data are highlighted in boxes. Adapted from (Fierer, 2017; Grime, 1974).

Genomic and metagenomic sequence data has been used to infer growth rates (e.g. from codon usage patterns), stress tolerance (e.g. from genes for sporulation and EPS production), and nutrient acquisition (e.g. from enzyme genes) ((Fierer, 2017); Figure 2). Attractive in its promise, it remains open how well these genomically predicted strategies map onto realized bacterial strategies. For this, a direct comparison between predicted traits and traits measured *in situ* is required. This should happen in traits co-varying with the strategies along the CSR axes of variation, as has been done in plants with plant specific leaf area, plant height and seed mass (Westoby, 1998).

For a trait-based understanding of host-associated microbes, bacterial attributes should be both inferred from genomic sequences and validated empirically. Put together, this will provide a window into key traits and strategies required for successful host association. Thus, life history theory may advance our current understanding of microbe-host associations and indicate what allows their emergence and maintenance.

## Modeling microbe-host associations using *C. elegans* and its naturally associated bacteria

To study function and strategies of bacteria in host association, as well as its emergence and maintenance, the nematode *Caenorhabditis elegans* and members of its naturally microbiota present an ideal model system. Importantly, this system has all attributes required to experimentally test hypotheses on the ecology and evolution of bacteria inside a host.



**Figure 3** *Discovering the natural microbiota of C. elegans.* (A) In nature, the nematode *Caenorhabditis elegans* lives in rotting plant material, including compost heaps. (B) Bacteria from the natural microbiota consortium CeMbio colonize the gastrointestinal tract of N2 *C. elegans*. In this confocal images, bacterial (in red) are stained using fluorescence in situ hybridization with the eubacterial probe EUB338 and worm nuclei (in blue) using DAPI. Image taken from (Dirksen et al., 2020). (C) In association with the worm, bacteria can act as food, pathogens, or parasites as well as commensals and mutualists. Adapted from (Schulenburg and Félix, 2017).

Like most animals and plants, the nematode harbors a microbiota in its natural habitat, although its role has been neglected and its physiology and evolution remain largely unclear. In nature, *C. elegans* lives in microbe-rich habitats such as compost heaps, rotting fruits and decaying plant matter (Félix and Braendle, 2010; Schulenburg and Félix, 2017). Yet most of the worm's biology has been studied in the context of bacteria not found in these habitats. Classically a model organism in biological disciplines including genetics, development, or neurobiology, the worm is traditionally maintained on *Escherichia coli* OP50 in the lab. This is a Gram-negative bacterium, which does not colonize worms until old age, as it is easily digested by worms crushing cells in their grinder (Avery and You, 2012). Furthermore, this bacterium grows only thinly on nematode growth agar, therefore making experimental handling and microscopy easy. *E. coli* OP50 is, however, not found in worms or their substrate in natural habitats (Dirksen et al., 2016; Samuel et al., 2016; Zhang et al., 2017). In a natural settings, *C. elegans* have been shown to host a diverse and species-rich microbiota. Crawling through composts and rotting plant material, worms pick up bacteria from an environmental source pool, thus acquiring a microbiota horizontally from the substrate. Importantly, microbiota bacteria can survive digestion to different extents and thus colonize the worm (Dirksen et al., 2016). Three studies published in 2016 have described the taxonomic composition of the microbiota independently, all taking slightly different, yet complementary approaches. Bacteria were either sampled from habitats *C. elegans* was found in (Samuel et al., 2016), wild worms captured in rotting plant material (Dirksen et al., 2016) or worms exposed to compost microcosms in the lab (Berg et al., 2016). A meta-analysis found strong overlap in bacterial communities across these studies indicating a specific assembly of bacterial taxa in the worm (Zhang et al., 2017). On the family level, *Xanthomonadaceae*,

*Pseudomonadaceae*, *Enterobacteriaceae*, *Comamonadaceae*, *Sphingomonadaceae*, *Sphingobacteriaceae*, *Weksellaceae* and *Flavobacteriaceae* can be found in worms from all study approaches (Zhang et al., 2017). These findings are also supported by additional experiments with worms in lab microcosms (Berg et al., 2019), as well as sampling of natural worms in the Kiel Botanical Garden over successive seasons (Johnke et al., 2020). On the species and strain level, however, great variation exists between samples, indicating that there is no stringent co-evolution of specific isolates with this worm species. This raises the question what other factors may shape the fit between microbe and host, which will be addressed in **chapter 1** using functional inference from bacterial genomes and using colonization experiments and experimental evolution in **chapters 4 and 5**.

In the experiments presented in this thesis, I used bacterial isolates from wild *C. elegans* captured in Northern Germany paired with their natural worm host (Dirksen et al., 2016). Specifically, a culture collection of 77 isolates of the microbiota community was used for functional profiling, characterization of colonization and impact on host fitness (**chapter 1**). To fundamentally assess the evolutionary fate of bacteria following a life history with *C. elegans* as a host as compared to their evolution without the worm (**chapters 4 and 5**), two isolates from bacterial genera of particular interest were selected as model organisms. These are *Pseudomonas lurida* (MYb11) and *Ochrobactrum vermis* (MYb71). Both MYb11 and MYb71 are enriched in worms compared to their substrate, and can be cultured in the lab (Dirksen et al., 2016; Sieber et al., 2019), making experiments both in worms and on agar feasible. More importantly, *P. lurida* MYb11 is both member of one of the core families of the worm microbiota as well one of the best researched bacterial genera, thus a biologically relevant microbiota member whose evolutionary change may be well interpreted with the help of comparative annotation. Although not relevant in the scope of this thesis, MYb11 has further raised interest for protecting *C. elegans* from infection by fungal and bacterial pathogens (Dirksen et al., 2016; Kissoyan et al., 2019). The second isolate, MYb71 presents a model for a host-adapted bacterium, as it is very abundant when sampled in worm microbial communities (Sieber et al., 2019), and has been shown to modulate *C. elegans* metabolism and reproduction, visible also in changes in the transcriptome and proteome as compared to *E. coli* OP50 (Cassidy et al., 2018; Yang et al., 2019). Further, *Ochrobactrum* is a genus also reported in symbiosis with entomopathogenic nematodes (Babic et al., 2000). It is thus of interest to learn about functions relevant and maintained in the interaction with *C. elegans* or lost in the absence of the host. Given their co-isolation from a single worm, these two isolates further allow investigating bacteria-bacteria interactions within the microbiota and their impact on adaptation to the worm in a small two-member model community (**chapter 5**). Both MYb11 and MYb71 are also part of a more recently established microbiota consortium, the CeMbio, including 12 bacteria representing the core microbiota of *C. elegans* (Dirksen et al., 2020), thus making the work presented here relatable to future research. Taken together, the natural microbiota provides bacterial isolates that for my purposes function as models to investigate fundamental questions of microbiota ecology and evolution.

In addition to association with a variety of bacterial taxa, it is known that *C. elegans* can interact with bacteria in different ways. Being a bacterivore, *C. elegans* uses bacteria as a food source. For this, the worm uses the grinder in the pharynx, low pH, and digestive enzymes to kill bacteria and access them as nutrients. Here *C. elegans* can use different bacterial species as food of varying quality to support development (Avery and You, 2012; Fang-Yen et al., 2009; Grewal, 1991; Shtonda and Avery, 2006). Notably, also naturally colonizing bacteria serve as food for growth and development, and thus generally serve a dual role as symbiont and food. Briefly addressed in **chapter 2**, an association between bacteria and worms can still be maintained in spite of digestion, as only a part of the bacteria taken up are digested, while the remaining live cells may interact with worms and can even be released back into the environment where they may grow again, a process sometimes even interpreted as farming (Hoang et al., 2019; Thutupalli et al., 2017). Next to serving as food, live bacteria may also influence worm physiology and fitness as mutualists. Bacteria may synthesize and provide essential nutrients and cofactors required for worm development, survival and

reproduction (e.g. (Chaudhari et al., 2016; MacNeil et al., 2013a; Na et al., 2018; Virk et al., 2016; Warnhoff and Ruvkun, 2019)). Furthermore, bacteria, including members of the natural microbiota, can protect the worms from pathogen infection (Ford et al., 2016; King et al., 2016; Kissoyan et al., 2019). Finally, some bacteria are also pathogens to *C. elegans*. This includes classic human pathogens such as *Pseudomonas aeruginosa* or *Salmonella enterica*, but also bacteria from natural habitats of the worm such as *Bacillus thuringiensis*, *Serratia marcescens* or *Leucobacter* species (Samuel et al., 2016; Schulenburg and Félix, 2017). Although intriguing in their effect on the worm, pathogens are not a focus of this thesis. Most research on bacterial impact on the worm, particularly a metabolic perspective, has been obtained with bacteria not found to associate with *C. elegans* naturally. So, it remains unclear what the natural microbiota can do and how relevant these interactions would be in a natural context. In this thesis I therefore build on the taxonomic description of the natural microbiota and investigate functional repertoires of its members as well as their impact and response to host adaptation (see **chapter 1**).

*C. elegans* and its natural microbiota are a great model system to study the early stages of microbe-host association. This is because specific taxa reliably associate with the worm yet colonize it anew every generation. It thus presents an intermediate between well-established obligate symbioses, and naïve pairs of microbes and hosts. This makes selection towards a more stringent host association, as well as the loss of host adaptive traits in its absence, possible. Furthermore, this experimental system allows us to study both defined association of naturally associating bacteria with axenic hosts and bacterial culture *in vitro* under controlled lab conditions (Berg et al., 2016; Dirksen et al., 2016, 2020; Samuel et al., 2016; Zhang et al., 2017). Well-established techniques for synchronizing worms in their development and rendering them sterile before exposing them to bacteria make this an easy to handle system to assess microbe associations under defined conditions (Douglas, 2019; Stiernagle, 2006; Zhang et al., 2017). At the same time, bacteria can be subjected to experimental evolution scaffolded along the complete host life cycle of only a few days. The worm's short generation time makes it possible to expose bacteria to worms across developmental stages and has proven valuable in studying bacterial within-host evolution (in the context of pathogens) as well as co-evolution of host and pathogen, or host and symbionts (Hoang et al., 2016; Jansen et al., 2015; King et al., 2016; Masri et al., 2015; Papkou et al., 2018; Rafaluk-Mohr et al., 2018; Schulte et al., 2010; Zhang et al., 2017). Finally, being one of many fascinating microbe-host associations, this system further benefits from *C. elegans* being a model organism in biology and having been studied extensively, thus providing a framework in which to interpret bacterial ecology and evolution (Petersen et al., 2015). At the same time, however, we currently have a limited understanding of what the bacteria in the worm can do or how they adjust to life in the worm.

## Objectives and overview of this thesis

In this thesis, I aim to explore the bacterial perspective to association with *C. elegans* as a host, to more generally learn about bacterial life history strategies in host association. Structured in two parts, I first present work on the functional ecology of host-associated bacteria and then focus on bacterial evolution in this context. In a general discussion at the end, the insights gained will be synthesized in light of bacterial life histories in the context of host association. Here I briefly introduce all thesis chapters and their aims.

### Part I. Functional ecology of microbiotas

**Chapter 1** is aimed at enhancing our understanding of the functional repertoire of bacteria naturally associated with the nematode *C. elegans*. Taking a bottom up approach, the metabolic competencies of 77 bacteria from the natural microbiota of *C. elegans* were

predicted from whole genome sequences and experimentally validated. This showed that together the community of bacteria may synthesize all nutrients essential to the worm host, but cholesterol. Notably, the interactions between bacteria in the community are context-dependent and carbon sources may be used to specifically modulate them. Finally, predicted bacterial life history strategies and traits are related to worm colonization and worm fitness. Here bacteria that invest mainly in stress-tolerance and competition appeared most abundant in worms, while the capacity to degrade hydroxyproline linked to increased host fitness.

## Part II. Evolution of bacteria in host association

In **chapter 2**, I present an opinion piece with the aim of providing a novel framework for the emergence and maintenance of microbe-host associations from the microbial perspective. In it, we argue that microbes in many extant microbiotas, but also at the origin of host association, follow a biphasic life cycle switching between hosts and free-living life. Connecting bacterial function to life history strategy, the traits required along the life cycle and their contributions to microbial fitness are examined and provide a basis for novel approaches to model microbial side of symbiosis. It is then discussed how bottlenecks along the life cycle and consequences for the host may shape microbial adaptation in an evolutionary time frame.

In the following **chapter 3**, I present a theoretical study that aims to identify which life history traits optimize microbial fitness in a biphasic life cycle. For this, a deterministic model describing a population of microbes transitioning between host association and free-living life in the environment is described. Using population growth rate as a proxy for fitness shows that greater fitness can be attained either by investment in increased replication or increased migration between host and the environment. Particularly when population size in the host, and thus growth inside it, are limited, microbes should invest in migration towards the host rather than replication.

In **chapter 4**, I present a study that aims to understand how the bacterium *Pseudomonas lurida*, a member of the natural *C. elegans* microbiota, adapts to the worm upon experimentally imposing a biphasic life cycle. Here, biphasically evolved populations showed greater colonization and persistence in worms than ancestral, while the absence of a host bacteria improved dispersal autonomy in terms of colony expansion. Further, a wrinkly colony type specialized on host association emerged and dominated biphasically evolved populations. Via genetic changes in the regulation of the second messenger c-di-GMP, wrinkly types improve biofilm formation and persistence in worms. This illustrates sticking to a host as a key step in the emergence and establishment of microbe-host-association.

**Chapter 5** is aimed at understanding how two microbiota members, *Pseudomonas lurida* and *Ochrobactrum vermis* co-colonize, sub-specialize and evolve together in *C. elegans* as a host. Both part of the natural microbiota of the *C. elegans*, they can colonize the worm alone and together. Here the bacteria can be distinguished in their ability to compete in the agar plate environment, quickly enter as well as persist inside the host. From this we infer distinct life history strategies of *P. lurida* and *O. vermis* in association with the worm. We further consider how shifts in the composition of the two-member community feed back on host development and fitness.

Finally, I provide a general discussion of bacterial responses to life inside a host and present perspectives for future research in the **epilogue**. Specifically, I review the role of biofilms in beneficial microbe-host associations, discuss what we have learned about bacterial life history strategies in host association, and provide a perspective on how to apply the insights gained to manipulate microbe-host associations.





# 1

## The functional repertoire encoded within the native microbiota of the model nematode *Caenorhabditis elegans*

Johannes Zimmermann\*

Nancy Obeng\*

Wentao Yang

Barbara Pees

Carola Petersen

Silvio Waschina

Kohar Annie Kissoyan

Jack Aidley

Marc P. Hoepfner

Boyke Bunk

Cathrin Spröer

Matthias Leippe

Katja Dierking

Christoph Kaleta#

Hinrich Schulenburg#

\* joint first author

# joint last author

## Abstract

The microbiota is generally assumed to have a substantial influence on the biology of multicellular organisms. The exact functional contributions of the microbes are often unclear and cannot be inferred easily from 16S rRNA genotyping, which is commonly used for taxonomic characterization of bacterial associates. In order to bridge this knowledge gap, we here analyzed the metabolic competences of the native microbiota of the model nematode *Caenorhabditis elegans*. We integrated whole genome sequences of 77 bacterial microbiota members with metabolic modeling and experimental characterization of bacterial physiology. We found that, as a community, the microbiota can synthesize all essential nutrients for *C. elegans*. Both metabolic models and experimental analyses revealed that nutrient context can influence how bacteria interact within the microbiota. We identified key bacterial traits that are likely to influence the microbe's ability to colonize *C. elegans* (i.e., the ability of bacteria for pyruvate fermentation to acetoin) and affect nematode fitness (i.e., bacterial competence for hydroxyproline degradation). Considering that the microbiota is usually neglected in *C. elegans* research, the resource presented here will help our understanding of this nematode's biology in a more natural context. Our integrative approach moreover provides a novel, general framework to characterize microbiota-mediated functions.

## Introduction

Multicellular organisms are continuously associated with microbial communities. The ongoing interactions have likely influenced evolution of the involved microbes and hosts, affecting bacterial growth characteristics or host development, metabolism, immunity, and even behavior (McFall-Ngai et al., 2013). Host organisms and their associated microorganisms (i.e., the microbiota) are thus widely assumed to form a functional unit, the metaorganism, where microbial traits expand host biology (Bosch and Miller, 2016). To date, most microbiota studies focus on describing bacterial taxonomic composition, using 16S rRNA amplicon sequencing (Pascoe et al., 2017). These studies revealed that specific taxa reliably associate with certain hosts, for example Bacteroidetes and Firmicutes with humans, *Snodgrassella* and *Gilliamella* with honeybees, or *Lactobacillus* and *Acetobacter* with *Drosophila* (Moran et al., 2012; The Human Microbiome Project Consortium et al., 2012; Wong et al., 2011). 16S profiling, however, is insufficient to identify bacterial functions relevant for the interaction (Louca et al., 2016). More insights can be obtained from bacterial genome sequences. For example, genomic analysis of bee microbiota members revealed complementary functions in carbohydrate metabolism, suggesting syntrophic interactions among bacteria (Kwong et al., 2014). Further, the systems biology approach of constraint-based modeling permits inference of genome-scale metabolic models and prediction of microbial phenotypes (Bordbar et al., 2014), as demonstrated for whiteflies and their endosymbionts (Ankrah et al., 2017; Luan et al., 2015) and also hosts with complex microbiotas (Bauer et al., 2015; Magnúsdóttir et al., 2017).

The nematode *Caenorhabditis elegans* is an important model organism in biomedical research. Yet, almost all *C. elegans* research has been without microbiota. In fact, the nematode's microbiota was only characterized recently, consisting mostly of Gammaproteobacteria (*Enterobacteriaceae*, *Pseudomonaceae*, *Xanthomonadaceae*) and Bacteroidetes (*Sphingobacteriaceae*, *Weeksellaceae*, *Flavobacteriaceae*) (Berg et al., 2016; Dirksen et al., 2016; Samuel et al., 2016; Zhang et al., 2017), some of which persist in the worm intestine (Berg et al., 2019; Dirksen et al., 2016; Kissoyan et al., 2019). The microbiota composition is influenced by both host genotype and environment, and appears similar across geographic regions ((Berg et al., 2016; Dirksen et al., 2016); see meta-analysis in (Zhang et al., 2017)). The few studies on microbiota functions highlight an influence on *C. elegans* fitness, stress resistance, and pathogen protection (Dirksen et al., 2016). Previous studies also combined *C. elegans* with soil bacteria, revealing that these can provide specific nutrients (Chaudhari et al., 2016; MacNeil et al., 2013b; Shapira, 2017; Virk et al., 2016; Watson et al., 2014). Bacterial metabolism can also affect the worm's response to drugs against cancer and diabetes (Cabreiro et al., 2013; García-González et al., 2017; Norvaisas and Cabreiro, 2018; Scott et al., 2017). To date, the functions of the native microbiota have not been systematically explored.

Our aim was to establish the natural *C. elegans* microbiota as a model for studying microbiota functions. We extended previous 16S rRNA data (Dirksen et al., 2016) by sequencing whole genomes for 77 bacteria, which are associated with *C. elegans* in nature, most likely as part of the intestinal microbiota, and also *Escherichia coli* OP50, the nematode's standard laboratory food. We reconstructed metabolic networks from the genomes to explore the bacteria's metabolic competences and microbe-microbe interactions. We additionally characterized bacterial physiology and assessed which bacterial traits shape colonization ability and influence *C. elegans* fitness.

## Materials and methods

### *Materials*

Microbiota strains were previously isolated from natural *C. elegans* isolates or corresponding substrates in Northern Germany ((Dirksen et al., 2016); Supplementary Table S1). Briefly, bacteria from worms were obtained after nematodes were thoroughly washed and broken up by vortexing with Zirconium beads. Most bacteria are likely from the intestines, yet an association with the nematode cuticle cannot be excluded (Dirksen et al., 2016). A representative set of 77 strains was chosen for genome sequencing, covering 79.5% of the diversity and abundance of the top ten genera and still 54.6% of that of the top 20 genera from the native *C. elegans* microbiota (some common microbiota members such as *Flavobacteria* could not yet be isolated (Dirksen et al., 2016)). For physiological analysis, bacteria were cultured in tryptic soy broth (TSB) at 28 °C. We performed experiments with the main *C. elegans* laboratory strain N2 (see below), thus allowing us to use previous literature and concurrent experiments with the standard food *E. coli* OP50 as a reference. For these experiments, bacterial TSB cultures (500 µl at OD600 = 10) were spread onto peptone-free medium (PFM). Maintenance and bleaching, to obtain gnotobiotic, agesynchronized worms, followed standard methods (Stiernagle, 2006).

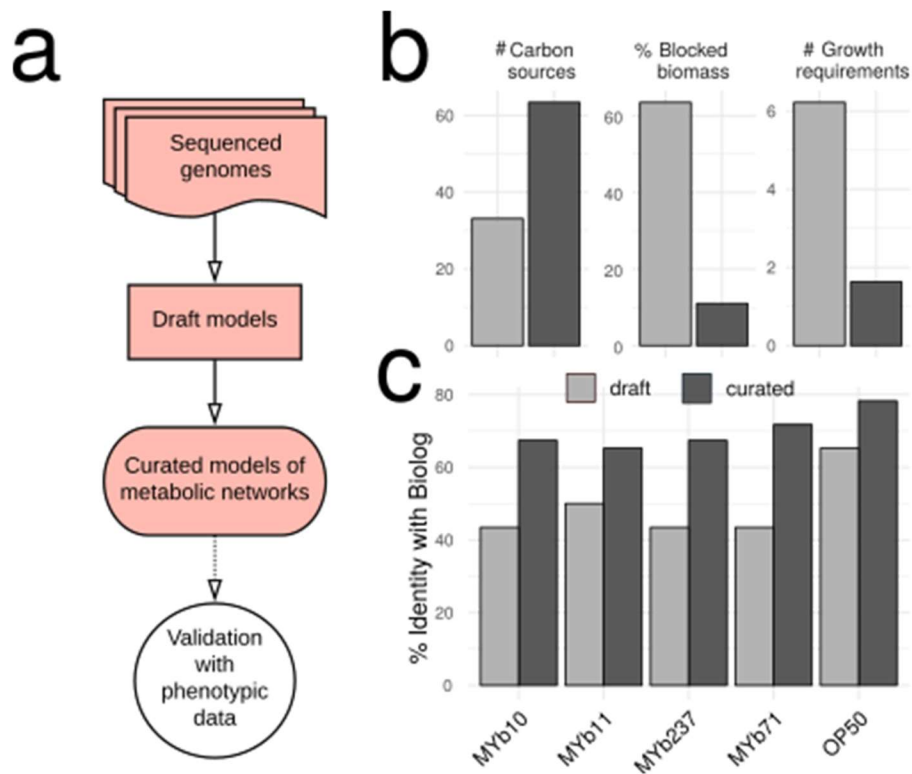
### *Genome sequencing*

Total bacterial DNA was isolated using a cetyl-trimethylammonium-bromid (CTAB) approach (Schulenburg et al., 2001). Sequencing was based on Illumina HiSeq and in a subset of nine strains additionally PacBio (Supplementary Table S1). For PacBio, SMRTbell™ template library was prepared following manufacturer's instructions (Pacific Biosciences, US; Protocol for Greater Than 10 kb Template Preparation). SMRT sequencing was performed on the PacBio RSII (Pacific Biosciences, US), applying 240-min movie lengths. PacBio data was assembled using the RS\_HGAP\_Assembly.3 protocol (SMRT Portal version 2.3.0). Chromosomes and chromids were circularized, unusual redundancies at contig ends and artificial contigs were removed. Error correction was performed by Illumina reads mapping onto genomes using BWA (Li and Durbin, 2009) with subsequent variant calling using VarScan (Koboldt et al., 2012). QV60 consensus concordances were confirmed for all genomes. Annotations were obtained with the NCBI Prokaryotic Genome Annotation Pipeline (PGAP). For samples with only Illumina data, low-quality reads and adaptors were trimmed with Trimmomatic v0.36 (Bolger et al., 2014). Genomes were assembled using SPAdes v3.8.0 (Bankevich et al., 2012) and contigs greater than 1000 bp annotated with PGAP and Prokka v1.11 (Seemann, 2014). Genomes were compared with BRIG (Alikhan et al., 2011). BUSCO (Simão et al., 2015) analysis revealed high-genome completeness, irrespective of sequencing technology (mean completeness of 96.81%; Supplementary Figs. S12 and S13). As assembly quality of the available OP50 genome was low (NCBI project PRJNA41499; >2900 contigs, 73% completeness), we sequenced and assembled it again (218 contigs, 98.7% completeness). All sequences are available from NCBI Genbank, Bioproject PRJNA400855.

### *Reconstruction of metabolic networks*

Metabolic networks were reconstructed following two steps (Fig. 1a). First, genomes were used to create draft metabolic models, using ModelSEED version 2.0 (Henry et al., 2010) and associated SEED reaction database. Second, we corrected errors and extended drafts by (i) finding futile cycles, (ii) allowing growth with the isolation medium (TSB), (iii) improving biosynthesis of biomass components, (iv) extending capacities to use different carbon sources, and (v) checking for additional fermentation products. Curation was based on combining topological- and sequenced-based gap-filling using gapseq (version 0.9;

<https://github.com/jotech/gapseq>), pathway definitions of MetaCyc release 22 (Caspi et al., 2018), and UniProt (Consortium, 2014). The presence of enzymatic reactions was inferred by BLAST with bitscore of  $\geq 50$  ( $\geq 150$  for more conservative estimation), and 75% minimum query coverage. Moreover, reactions were assumed to be present if overall pathway completeness was  $>75\%$  or if it was  $>66\%$  and key pathway enzymes were present. Host-microbe interaction genes were identified with the virulence factor database (Chen et al., 2016). The resulting curated models (Supplementary data S1) were used for further metabolic network analysis. Genome incompleteness did not have a large effect on pathway reconstruction (Supplementary Fig. S14). Computations were done with GNU parallel (Tange, 2011).



**Fig. 1 Genomes of bacterial isolates, reconstruction and validation of metabolic networks.** *a* Pipeline for metabolic network reconstruction. Sequenced genomes were used to create draft metabolic models. Draft models were curated using topological- and sequenced based gap-filling. The resulting models were validated with physiological data (BIOLÓG GN2; see Fig. 3); these models represent the metabolic networks of microbiome isolates and were used for functional inference. *b* Model improvements by curation, leading to an increase in accurate prediction of uptake of carbon sources, and decreases in the prediction of non-producible biomass components and the number of components needed for growth. *c* Model curation improved agreement with experimental data, as for example the BIOLÓG results.

#### Phylogenetic correlation and clustering of metabolic pathways

We assessed the correlation between metabolic pathway similarity and phylogenetic relationships, using pairwise comparisons of bacteria. We specifically focused on 16S rRNA sequences to calculate phylogenetic similarities, in order to enable comparisons with the standard microbiota analysis approach, based on 16S amplicon sequencing. 16S similarity was scored as percent identity with biostrings (Pagès et al., 2018), using data from the SILVA database (Quast et al., 2013) based on best hits of extracted genomic 16S rRNA using RNAmmer (Lagesen et al., 2007). To determine overall metabolic distances between isolates,

metabolic networks were treated as vectors, clustered horizontally, followed by computation of Euclidean distances between vectors. Cluster similarity was estimated by average linkage and assessed via multi-scale bootstrapping (10,000 replications) using pvclust (Suzuki and Shimodaira, 2006).

### *BIOLOG experiments*

We used BIOLOG GN2 plates to assess metabolic competences of selected bacteria, including MYb10, MYb11, MYb71, MYb237, and OP50. Bacterial cultures were washed thrice using phosphate-buffered saline (PBS) and density adjusted to OD600 = 1.150  $\mu$ l bacterial suspension per well of BIOLOG plate was incubated at 28 °C for 46 h. Tetrazolium dye absorption (OD595) was measured every 30 min (three replicates per strain). Substrate reduction was inferred from fold-change in tetrazolium absorbance:

$$foldchange = \frac{OD_{t46} - OD_{t0}}{OD_{t0}} - OD_{control}$$

Fold-changes in water were subtracted as background. Hierarchical clustering of strains was based on average foldchange profiles (Ward's clustering; Euclidean distance) and bootstrapping (n = 100). We analyzed metabolic specialization by k-means clustering of substrates (k = 7, n = 103; (Hartigan and Wong, 1979)) (Supplementary Fig. S1). Statistical analyses were performed in R version 3.3.1 (R Core Team, 2013) and ggplot2 (Wikham, 2016).

### *Bacterial growth experiments*

To validate BIOLOG results, we assessed growth of MYb11, MYb71, and their co-culture in defined media with either alpha-D-glucose or D-(+)-sucrose as carbon sources. We focused on these two isolates, because they are members of two common taxa of the native microbiota of *C. elegans* (Dirksen et al., 2016) and because detailed information is available on the interaction of these two isolates with *C. elegans*, including their ability to colonize the nematode gut, persist under stressful conditions, influence nematode population growth, and provide protection against pathogens (Dirksen et al., 2016; Kissoyan et al., 2019). Our defined medium is related to S medium (Stiernagle, 2006): 0.3% NaCl, 1 mM MgSO<sub>4</sub>, 1mM CaCl<sub>2</sub>, 25mM KPO<sub>4</sub>, 0.1% NH<sub>4</sub>NO<sub>3</sub>, 0.05 mM EDTA, 0.025 mM FeSO<sub>4</sub>, 0.01mM MnCl<sub>2</sub>, 0.01mM ZnSO<sub>4</sub>, 0.01 mM CuSO<sub>4</sub>, and 1% carbon source. Medium without carbon source served as negative and TSB as positive control. Overnight cultures were washed and adjusted to 3.94  $\times$  10<sup>7</sup> CFUs for growth experiments. Microtiter plates were incubated as BIOLOG plates above. OD600 was measured every 30 min, and cultures plated after 48 h. Selective plating of MYb71 using kanamycin (10  $\mu$ g/ml) allowed to quantify MYb11/MYb71 proportions in co-culture. Three independent runs with technical replicates were assessed with Mann-Whitney U-tests and Pvalue adjustment by false discovery rate (fdr, Benjamini Yoav et al. (Benjamini and Hochberg, 1995)).

### *Simulation of bacterial in silico growth*

We used the curated models to simulate growth of MYb11 and MYb71 with sucrose as carbon source using BacArena (Bauer et al., 2017). Sucrose invertases were identified with gapseq (<https://github.com/jotech/gapseq>) and secreted peptides with SignalP 4.1 (Petersen et al., 2011). The MYb71 extracellular sucrose invertase was modeled as independent species with a single sucrose invertase reaction and exchange reactions for sucrose, glucose, and fructose. Carbon source utilization and metabolic by-products were predicted using flux balance and variability analysis in R with sybil (Gelius-Dietrich et al., 2013). Flux balance analysis is a constrained-based method to estimate intra-cellular reaction activities by linear optimization

(Orth et al., 2010), permitting inference of bacterial growth. A carbon source was assumed utilizable if the minimal solution of the corresponding exchange was negative and a byproduct producible if the maximal solution of exchange positive.

### *Simulation of ecological interactions*

We assessed possible interactions among bacteria using joined models, assuming a common compartment for metabolite exchange between microbes. Activity of individual reactions (i.e., fluxes) was linearly coupled to biomass production to prevent unrealistic exchange fluxes, such as those that solely benefit the partner but not the producer (Heinken et al., 2013). The objective function was set to maximize the sum of fluxes through both biomass reactions. Two growth media were used for simulations, TSB (Supplementary Table S2) and a glucose minimal medium with thiamine and traces (0.001 mM) of sucrose and methionine to allow initial bacterial growth (Supplementary Table S3). Joined growth rates ( $j_1, j_2$ ) were compared to single growth rates ( $s_1, s_2$ ). Mutualism was defined as  $j_1 > s_1$  and  $j_2 > s_2$ , competition as  $j_1 < s_1$  and simultaneously  $j_2 < s_2$ , parasitism as  $j_1 < s_1$  and simultaneously  $j_2 > s_2$  (or  $j_2 < s_2$  and  $j_1 > s_1$ ), and commensalism as  $j_1 = s_1$  and  $j_2 > s_2$  (or  $j_2 = s_2$  and  $j_1 > s_1$ ).

### *Experimental analysis of bacterial colonization and bacterial effects on *C. elegans* population growth*

We examined bacterial colonization of *C. elegans* (i.e., bacteria attached to worms after the washing protocol, thus mainly consisting of intestinal bacteria) by quantifying CFUs extracted from young adults exposed to bacteria for 24 h. In detail, L4 larvae were raised on OP50 lawns and placed on each of the considered bacteria (500  $\mu$ l, OD600 = 10; only one bacterium present). After 24 h, they were washed in a series of buffers (2  $\times$  M9 buffer with 25mM tetramisole, 2  $\times$  M9 with 25mM tetramisole and 100  $\mu$ g/ml gentamicin, 1  $\times$  PBS with 0.025% Triton-X) to remove bacteria from nematode surfaces, and homogenized in the GenoGrinder 2000 using 1 mm zirconia beads (1200 strokes/min, 3 min). Worm homogenate and supernatant control were plated onto tryptic soy agar for quantification. We further measured worm population growth as a proxy for worm fitness. Three L4, raised on OP50 lawns, were transferred onto lawns with the considered bacteria and total worm numbers counted after five days at 20 °C.

### *Regression models*

We analyzed the association between phenotypic measurements (i.e., bacterial colonization, worm fitness) and metabolic or virulence characteristics using Spearman rank correlation and random forest regression analysis. Significance of correlations was assessed with permutation tests (100 randomly generated features, FDR-adjusted Pvalues). For random forest regression, the R package VSURF served to select features via permutation-based score of importance (Genuer et al., 2015) and otherwise default settings (ntree = 2000, ntry = p/3).

### *Adaptive strategies*

According to the universal adaptive strategy theory (UAST) (Fierer, 2017; Grime, 1977), heterotrophic bacteria follow one of three strategies: (i) rapid growth and thus good competitor, (ii) high resistance and thus stress-tolerator, or (iii) fast niche occupation and thus ruderal. We categorized bacterial isolates using published UAST criteria for genomic data (Fierer, 2017), which are based on three scores, inferred from genome sequences and metabolic models. In detail, the components of a competitive strategy were large genome size, antibiotics production (presence of pathways belonging to “Antibiotic-Biosynthesis” category in MetaCyc), high-catabolic diversity (Metacyc: “Energy-Metabolism”), and siderophore biosynthesis (Metacyc: “Siderophores-Biosynthesis”). The criteria for stress-tolerators were

auxotrophies, slow growth rates in TSB, few rRNA copies, and exopolysaccharides production (MetaCyc pathways: PWY-6773, PWY-6655, PWY-6658, PWY-1001, PWY-6068, PWY-6082, PWY-6073). Those for a ruderal strategy were fast growth in TSB, multiple rRNA copies, and low catabolic diversity (Metacyc: "Energy-Metabolism"). The characteristics of each isolate were related to the other bacteria, yielding a relative score, thereby assuming that different strategies are present in the microbial community as a whole. For each isolate, we assessed whether the inferred value belonged to the lower or upper quantile of this criterium (in case of growth rates we used the mean instead). The total adaptive score per strategy was scaled by the number of features considered per strategy. An isolate was assumed to follow the strategy, for which it produced the highest score. If two strategies had the same score, then we assumed a mixed strategy. The UAST classification remained stable with the same qualitative order, irrespective of genome completeness or sequencing technology (Supplementary Figs. S15 and S16).



## Results

### *Genomes of bacterial isolates, reconstruction and validation of metabolic networks*

We obtained whole genome sequences for 77 bacterial isolates of the *C. elegans* microbiota (Table 1). Of these, nine were sequenced with PacBio technology, allowing their full assembly, yielding either a single-circular chromosome (four strains) or three chromosomes/chromids in case of the five isolates of the genus *Ochrobactrum*, which is known to have more than one chromosome (Chain et al., 2011) (Supplementary Table S1). The remaining isolates were sequenced with Illumina only, resulting in assemblies with 11 up to 243 contigs. For four genera (*Ochrobactrum*, *Pseudomonas*, *Arthrobacter*, *Microbacterium*), we included more than five strains and identified substantial intragenomic genome variation (Supplementary Fig. S2).

**Table 1 Overview of bacterial isolates from the natural microbiota of *C. elegans* included in this study**

Phylum	Order	Genus	Isolate
Proteobacteria	Xanthomonadales	<i>Stenotrophomonas</i>	MYb238, <b>MYb57</b>
Proteobacteria	Pseudomonadales	<i>Pseudomonas</i>	MYb1, MYb114, MYb115, MYb117, MYb12, MYb13, MYb16, MYb17, MYb184, MYb185, MYb2, MYb22, MYb3, MYb60, MYb75, <b>MYb11</b> , MYb187, <b>MYb193</b>
Proteobacteria	Pseudomonadales	<i>Acinetobacter</i>	MYb10
Proteobacteria	Enterobacteriales	<i>Erwinia</i>	MYb121
Proteobacteria	Enterobacteriales	<i>Escherichia</i>	MYb137, MYb5, OP50
Terrabacteria group	Actinobacteria	<i>Micrococcaceae</i>	MYb211, MYb213, MYb214, MYb216, MYb221, MYb222, MYb224, MYb227, MYb229, MYb23, MYb51
Terrabacteria group	Actinobacteria	<i>Microbacteriaceae</i>	MYb24, MYb32, MYb40, MYb43, MYb45, MYb50, MYb54, MYb62, MYb64, MYb66, MYb72
FCB group	Bacteroidetes	<i>Flavobacteriales</i>	MYb25, MYb44, MYb7
Proteobacteria	Caulobacterales	<i>Brevundimonas</i>	MYb31, MYb33, MYb46, MYb52
Terrabacteria group	Bacilli	<i>Paenibacillaceae</i>	MYb63
Proteobacteria	Rhizobiales	<i>Ochrobactrum</i>	<b>MYb6</b> , MYb14, <b>MYb15</b> , MYb18, MYb19, MYb29, <b>MYb49</b> , <b>MYb58</b> , MYb68, <b>MYb71</b> , MYb237
Proteobacteria	Burkholderiales	<i>Achromobacter</i>	MYb9, <b>MYb73</b>
Terrabacteria group	Bacilli	<i>Bacillaceae</i>	MYb48, MYb56, MYb67, MYb78, MYb209, MYb212, MYb220
Bacteroidetes	Sphingobacteriales	<i>Sphingobacterium</i>	MYb181
Actinobacteria	Actinomycetales	<i>Rhodococcus</i>	MYb53

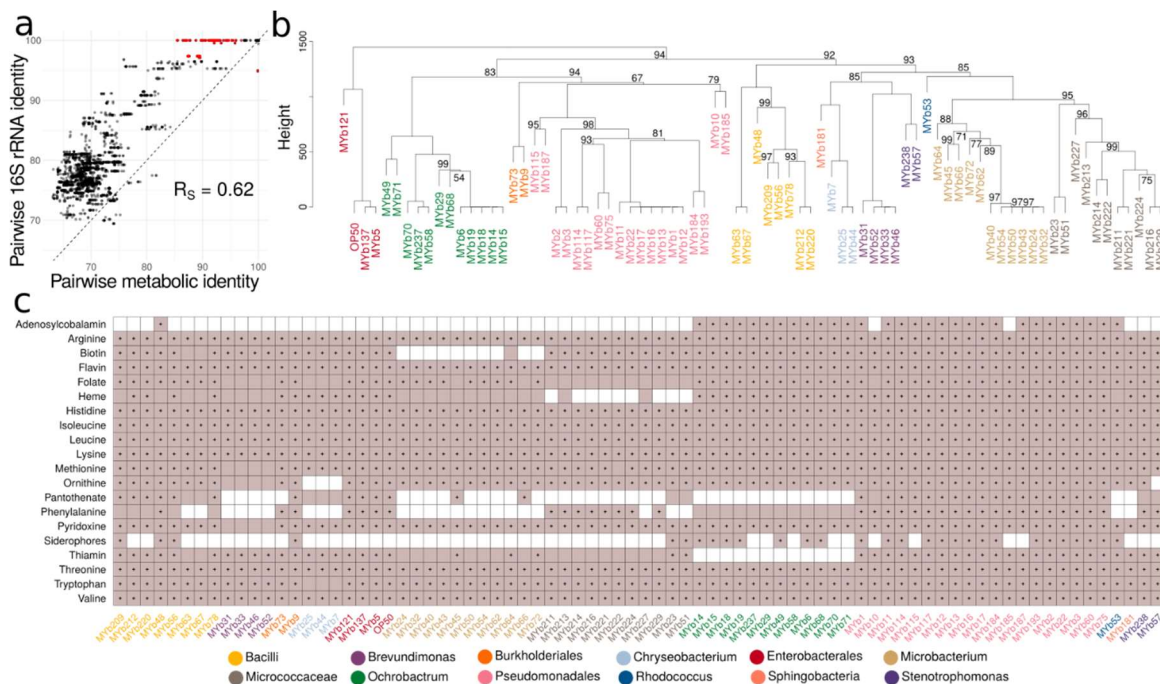
Strains with PacBio sequencing data are given in bold

To study the microbiota's functional repertoire, we reconstructed genome-scale metabolic models (Fig. 1a and Supplementary Data S1). The initial metabolic models were curated by screening for transporter proteins and filling of missing reactions (gap-filling). Curation increased model quality, including doubling of utilizable carbon sources, reduced absence of essential biosynthesis pathways (e.g., for nucleotides or amino acids) from 60% to below 10%, and reduction in the required additional compounds for growth on defined media from on average six to one (Fig. 1b). To validate our metabolic models, we experimentally

quantified the ability of five selected bacteria to utilize 46 carbon sources using the BIOLOG approach. The results produced 49.6% overlap with the initial and 70% overlap with the curated models (Fig. 1c and Supplementary Fig. S9). A 70% overlap is generally consistent with previous studies with model organisms like *Salmonella enterica*, *E. coli*, *Bacillus subtilis*, or *Pseudomonas putida* (AbuOun et al., 2009; Kumar and Maranas, 2009; Oh et al., 2007). Notably, the models in these studies were manually reconstructed, highlighting the quality of our automated reconstructions.

### Metabolic diversity within the microbiome of *C. elegans*

We used the curated metabolic networks to assess the relationship between metabolic and phylogenetic similarities and the bacteria's metabolic potential. For phylogenetic relationships, we specifically focused on 16S rRNA sequences, as they are most commonly used to characterize microbiota communities (Pascoe et al., 2017). We found that pairwise 16S phylogenetic relationships are generally indicative of metabolic network similarities (Fig. 2a; Spearman rank correlation,  $R_s = 0.6199$ ,  $P < 0.0001$ ). Phylogenetic similarities appeared to be larger than metabolic similarities, suggesting some variation in metabolic competences within taxa. Such variation even occurred among isolates with 16S identity above 97%, often used as a species cut-off. This was confirmed through hierarchical clustering of inferred metabolic networks (Fig. 2b), for example for the genus *Pseudomonas* with three clearly separated clusters (see similar patterns for *Enterobacter*, *Ochrobactrum*, or *Microbacterium*). We conclude that variation in metabolic competences is generally related to bacterial phylogeny albeit some variation being present within genera.



**Fig. 2 Metabolic network clustering and distribution of important pathways.** **a** Correlation between pairwise similarities in 16S rRNA sequences and metabolic networks is shown. Red indicates pairs with a 16S rRNA identity above 97% and metabolic identity below 97% and vice versa. **b** Hierarchical clustering of metabolic networks based on pathway prediction. P-values were calculated via multi-scale bootstrap resampling. In case of full support (i.e.,  $P = 100$ ), P-values are not shown (For a complete list of different unbiased P-values and bootstrap values see Supplementary Fig. S11). **c** Prediction of bacterial capacity to produce metabolites favoring *C. elegans* growth. Filled squares in light purple indicate that the metabolic networks predict the presence of the biosynthetic pathway required to produce essential amino acids and co-factors. Black dots within the filled squares indicate that pathway presence is supported by more conservative parameters (BLAST bitscore  $\geq 150$ ). Different bacterial genera in b and c are indicated by different colors of the strain names (Table 1).

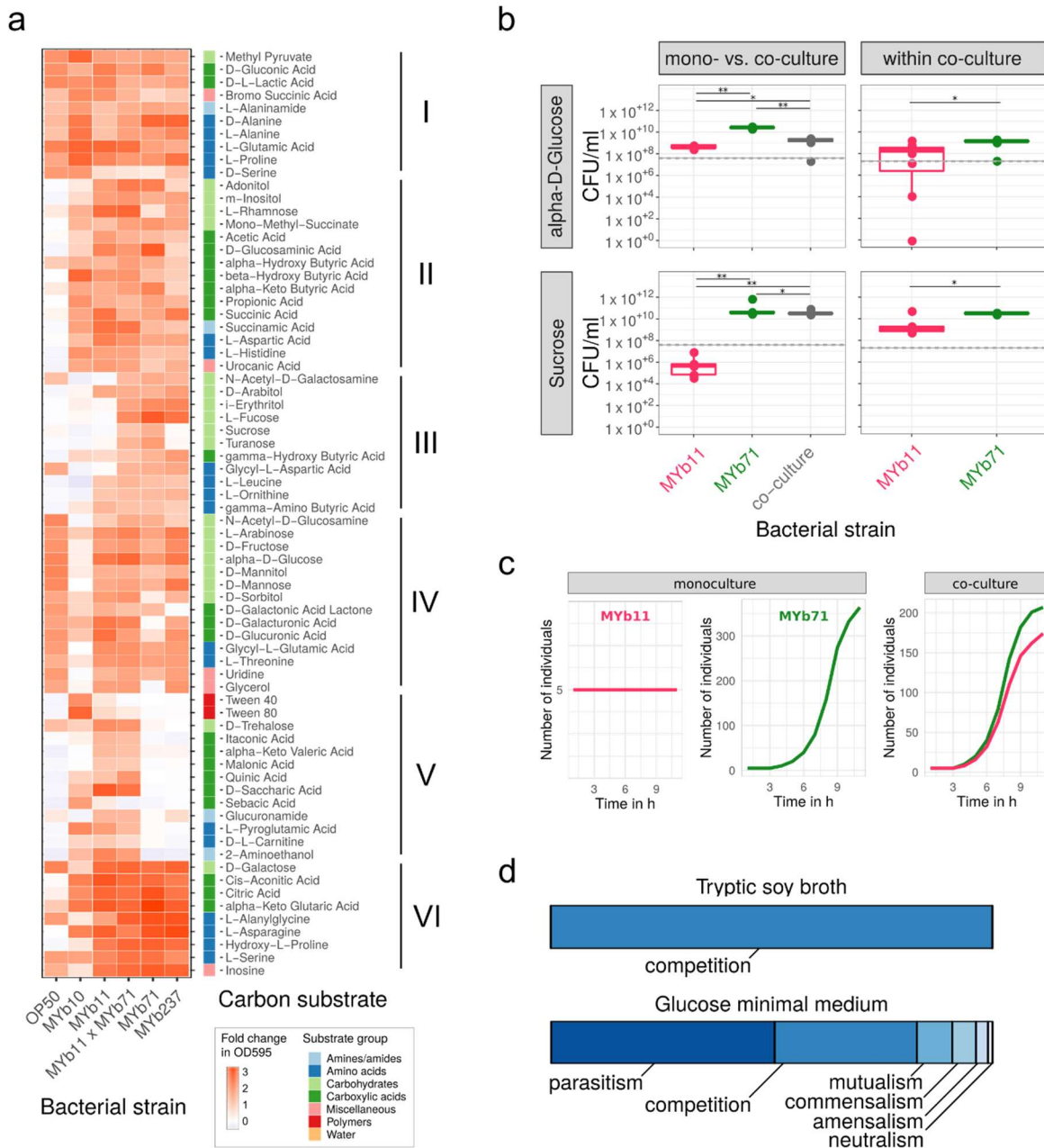
We next assessed the bacteria's metabolic competences (Supplementary Table S4). In general, the inferred metabolic competences are consistent with the aerobic and heterotrophic lifestyle of the *C. elegans* host. The glycolysis, at least the partial pentose phosphate pathway, the tricarboxylic acid cycle, and enzymes enabling oxidative phosphorylation (cytochrome oxidases) were present in all genomes. Almost all isolates possessed enzymes enabling tolerance to microaerobic conditions (e.g., cytochrome bd oxidase). Some *Bacilli*, *Pseudomonas*, and *Ochrobactrum* showed competences for chemolithotrophic lifestyle (nitrite and formate oxidation) and anaerobic respiration (nitrate, arsenate reduction). Pathways related to CO<sub>2</sub> fixation (reductive TCA or anaplerosis) were found in several *Pseudomonas*, *Microbacterium*, or *Bacilli*. Two *Bacillales* strains showed capacity to degrade polysaccharides, such as starch, cellulose, mannan, rhamnogalacturonan (e.g., *Paenibacillus* MYb63, *Bacillus* MYb67). The microbiota members are able to produce all essential substances required for *C. elegans* growth, which the nematode cannot synthesize on its own (i.e., all essential amino acids and vitamins; Fig. 2c). Most variation among isolates was observed in the biosynthetic pathways of B12, pantothenate, phenylalanine, and siderophores (Fig. 2c). Simulation of *in silico* growth (Supplementary Fig. S9) suggests that all bacteria can use simple sugars, such as glucose, ribose, or arabinose, while only some can degrade lactose, maltodextrin, or sucrose. Short-chain fatty acids can be generated by all bacteria (Supplementary Fig. S9), while they vary in succinate, cysteine, and valine production. Several microbes possess potential virulence genes, especially *Pseudomonas* and *Escherichia* isolates (Supplementary Table S5).

We subsequently focused on *Ochrobactrum* and *Pseudomonas* isolates. These two genera are enriched in the native microbiota of *C. elegans*, comprising 10–20% of the associated bacteria, they are particularly well able to colonize the nematode gut (Dirksen et al., 2016), and some isolates can protect *C. elegans* from pathogens (Dirksen et al., 2016; Kissoyan et al., 2019). Most *Pseudomonas* isolates can provide all required substances for nematode growth. *Ochrobactrum* isolates can produce vitamin B12, like *Pseudomonas* isolates, but unlike most other microbiota members (Fig. 2c). Moreover, the *Ochrobactrum* isolates vary from other microbiota members in degradation pathways, energy metabolism, vitamin biosynthesis, and potential virulence factors (Supplementary Table S6). They apparently lack thiamine and pantothenate vitamin biosynthetic pathways, essential for *C. elegans*. They possess a unique *Brucella*-like putatively immune-modulating LPS (Supplementary Table S5).

In summary, *C. elegans* harbors a microbial community with diverse metabolic competences, which can supply all essential nutrients for *C. elegans* and includes several *Ochrobactrum* and *Pseudomonas* isolates capable of producing important vitamins such as B12.

#### *Nutrient context influences ecological interactions within the microbiota*

To study how metabolic repertoires affect bacterial growth and interactions within the microbiota, we characterized carbon source utilization of selected isolates and tested growth in different nutrient environments *in vitro* and *in silico*. Using the BILOG approach, we focused on prominent *C. elegans* microbiota members that colonize worms and affect host fitness, including MYb71, MYb237 (both *Ochrobactrum*), MYb10 (*Acinetobacter*), MYb11 (*Pseudomonas lurida*), and *E. coli* OP50 as control (Supplementary Fig. S3; ref. (Dirksen et al., 2016)). For a first insight into bacterial interactions, we additionally included a MYb11-MYb71 mixture (two strains that can co-exist in *C. elegans* (Dirksen et al., 2016)).



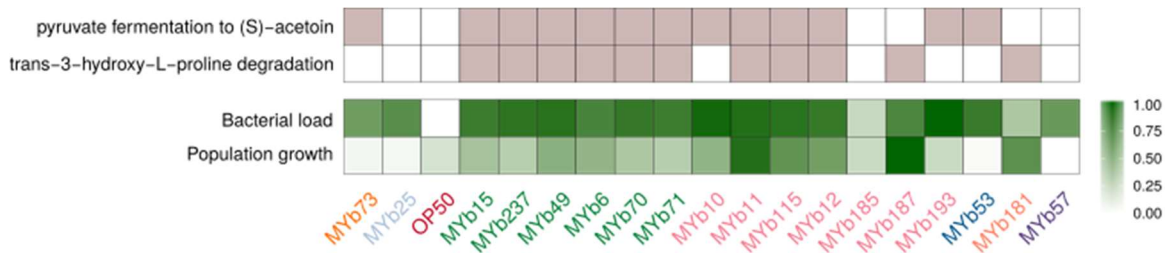
**Fig. 3 Realized carbon metabolism and growth.** **a** Profiles of carbon substrate use of *Acinetobacter* sp. (MYb10), *Pseudomonas lurida* (MYb11), *Ochrobactrum* sp. (MYb71), *Ochrobactrum* sp. (MYb237), and *E. coli* OP50 in BIOLOG GN2 plates over 46 h. The fold-change in indicator dye absorption from 0 to 46 h indicates that the particular compound is metabolized. *k*-means clustering ( $k = 7$ ) of substrates by fold-change highlights metabolic differences between strains. See Supplementary Fig. S5 for cluster VII with substrates used poorly across most strains. **b** Colony-forming units per ml (CFU/ml) of MYb11 and MYb71 in mono- and co-culture at 48 h in alpha-D-glucose and sucrose-containing minimal media. The horizontal and dashed lines indicate mean and SD of CFU/ml at inoculation. Statistical differences were determined using Mann-Whitney U-tests and corrected for multiple testing using *fdr*, where appropriate. Significant differences are indicated by stars (\*\* for  $P < 0.01$ ; \* for  $P < 0.05$ ). Data from three independent experiments is shown. **c** In silico growth of MYb11 and MYb71 in mono- and co-culture in sucrose-thiamine medium using BacArena with an arena of  $20 \times 20$  and five initial cells per species. **d** Bacterial interaction types observed during in silico cocultures of all combinations of the 77 microbiota isolates and OP50.

Metabolic repertoires of the strains differ and the four microbiota isolates deviate from OP50 in carboxylic and amino acid metabolism (Fig. 3a, cluster II; and Supplementary Fig. S4). MYb10 was least versatile at using carboxylic acids and sugar alcohols (Fig. 3a, cluster IV), while MYb11 and both *Ochrobactrum* could additionally metabolize unique sets of carboxylic acids and sugar alcohols, respectively (Fig. 3a, cluster V and III). Notably, the disaccharides sucrose and turanose were only metabolized by MYb71 (Fig. 3a, cluster III), although sucrose invertases were present in the genomes of both MYb71 and MYb11 (cf. pathway: sucrose degradation I, Supplementary Table S4). In co-culture, the metabolic repertoires of MYb11 and MYb71 appeared additive.

We next assessed whether the differences in MYb11 and MYb71 metabolic competences shape bacterial interactions in growth media with only a single carbon source. We focused on these two isolates as a model and proof-of-principle, because their interaction with *C. elegans* has been characterized in detail, including efficient colonization of nematodes, persistence under stressful conditions, an effect on nematode fitness, and protection against pathogens (Dirksen et al., 2016; Kissoyan et al., 2019). We considered growth in the presence of two sugars, which are characteristic for the *C. elegans* natural habitat (e.g., rotting fruits and plant matter). We did not observe any growth in a control medium without carbon source, and thus the tested bacteria are not chemoautotrophic (Supplementary Fig. S6). In minimal medium with alpha-D-glucose, both MYb11 and MYb71 grew, yet exhibited distinct growth dynamics (Fig. 3b and Supplementary Fig. S6). MYb71 produced more CFUs than MYb11 in co-culture (Fig. 3b), suggesting that MYb71 has a growth advantage and/or interferes with MYb11 in some other way. In agreement with the BIOLOG results, a medium including only sucrose supported growth of MYb71 but not MYb11 in monoculture (Fig. 3b and Supplementary Fig. S6). Surprisingly, MYb11 grew in co-culture, indicating parasitic growth (Fig. 3b). Thus, the presence of different carbon sources can change the interaction type between two isolates.

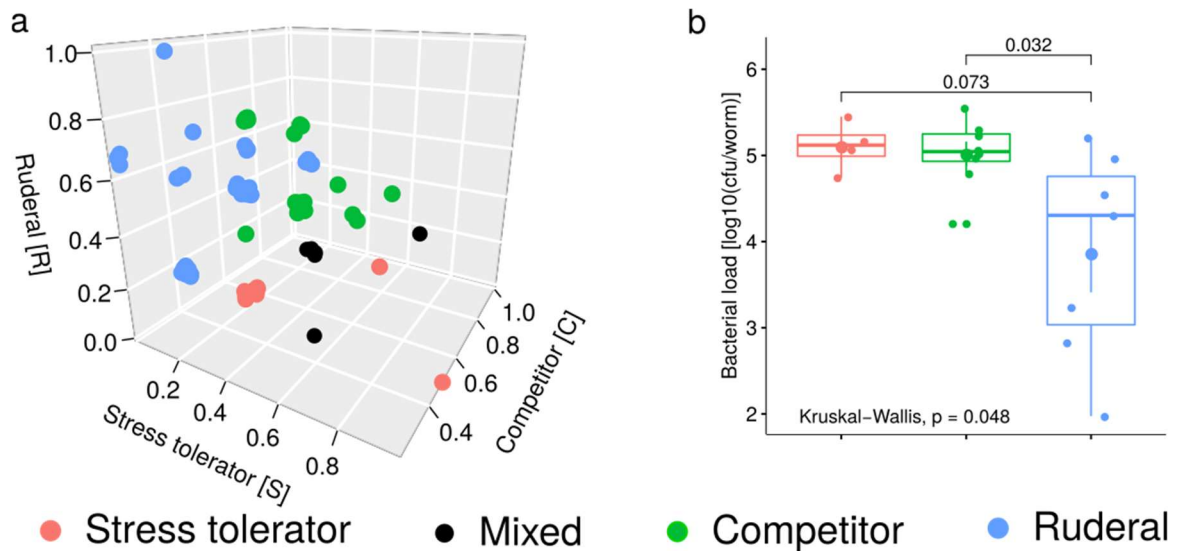
We subsequently assessed the basis for co-growth of MYb11 and MYb71 in sucrose medium, using genome sequence information and *in silico* growth simulations. Interestingly, we found a secreted sucrose invertase in the genome of MYb71 but not MYb11 (Supplementary Fig. S10). *In silico* simulations demonstrated that MYb71 can grow in sucrose medium, MYb11 alone does not, while both grow in co-culture (Fig. 3c), confirming our *in vitro* results. Genome sequence information strongly suggests that growth of both in co-culture is mediated by a secreted enzyme from MYb71. Taking a more global perspective, we next investigated *in silico* the potential ecological interactions among bacteria. We compared bacterial growth characteristics in monoculture and co-culture in different nutrient environments. In rich medium (TSB), the exclusive interaction type was competition, indicated by lower growth rates in co- vs. monoculture (Fig. 3d). This changed completely in glucose minimal medium: 50% interactions were parasitic (i.e., the growth rate for one isolate was higher in co-culture than in monoculture, while this pattern was opposite for the other isolate of a pair), 30% were competitive, and 8% mutualistic (i.e., growth rates for both isolates higher in co-culture than the monocultures; Fig. 3d). Under these minimal medium conditions, the most frequently exchanged metabolites across bacteria were glyceraldehyde, acetate, and ethanol (Supplementary Fig. S7). We conclude that the nutrient context modulates bacterial growth, consistently identified *in silico* and *in vitro*, and thereby shapes bacteria-bacteria interactions.

*Specific metabolic competences predict bacterial colonization ability and bacterial effects on nematode fitness*



**Fig. 4 Relationship of bacterial metabolic competences with their colonization ability and their effects on nematode fitness.** Presence of metabolic traits (light purple color), which were found to be associated with the bacteria’s ability to colonize *C. elegans* or affect nematode population growth as a proxy for worm fitness (green color). Regression models suggested that the pathway of pyruvate fermentation to acetoin influences bacterial load while the presence of hydroxyproline degradation is associated with *C. elegans* population growth. Colonization and population growth data was normalized; darker colors indicate increased capacities. Different bacterial genera are indicated by the different colors of the strain names (Table 1).

We next characterized traits involved in the interaction between bacteria and *C. elegans*, especially the bacteria’s colonization ability and their effects on worm fitness. We focused on 18 microbiota isolates based on (i) their abundance in the *C. elegans* microbiota, (ii) enrichment in worms, and (iii) effects on worm population growth [15, 63]. OP50 was included as control. The bacteria varied substantially in their ability to colonize *C. elegans* and their effects on nematode fitness (Fig. 4; Supplementary Fig. S3; and Supplementary Table S7). Importantly, these two microbiota characteristics were significantly related with certain metabolic competences. Pyruvate fermentation to (S)-acetoin was significantly associated with bacterial load and the degradation of trans-3-hydroxyproline with nematode population growth (Fig. 4 and Supplementary Table S8).



**Fig. 5 Different adaptive strategies within the microbiota and their relationship to worm colonization.** We applied the universal adaptive strategy theory proposed for soil bacteria [58] to categorize the bacterial isolates. **a** Based on genomic and metabolic features, each isolate obtained a score for the competitive (C), stress-tolerating (S), and ruderal (R) strategy, which is represented in the 3D-coordinate system. **b** Bacterial colonization behavior in comparison to adaptive strategies. Isolates that were categorized as ruderal produced the lowest bacterial load, whereas stress-tolerator and competitors had the highest values. The difference in bacterial load between ruderal and other strategies was significant (Wilcoxon rank-sum test,  $P = 0.01$ )

To further explore the potential behavior of all microbiota isolates in an ecological context, we interpreted their genomic and metabolic traits in light of the universal adaptive strategy theory [57, 58]. Twenty-six isolates were associated with a competitive, nine with a stress-tolerating, and 37 with a ruderal (fast niche occupiers) strategy (Fig. 5a and Supplementary Table S9). The remaining six isolates showed a mixed strategy (same score for competition and stress-tolerance). Interestingly, isolates with different adaptive strategies also varied in their colonization ability (Fig. 5b): bacteria with competitive or stress-tolerance strategies showed higher bacterial load than those with ruderal strategy (Wilcoxon rank-sum test,  $P = 0.01$ ). Moreover, for the competitive and stress-tolerance isolates, bacterial load was significantly correlated with the inferred score (Spearman,  $RS = 0.37$ ,  $P = 0.1$ ; Supplementary Fig. S8). Taken together, the competitive and stress-tolerating strategies are most prevalent within the *C. elegans* microbiota and relate to bacterial colonization capacity.

## Discussion

We here present the first overview of the functional repertoire contained within the native microbiota of *C. elegans* and provide a metabolic framework for functional analysis of host-associated microbial communities. Whole-genome sequences were used to reconstruct the metabolic network of 77 microbiota members. We found that bacterial metabolic competences vary and that the community as a whole can produce nutrients essential for *C. elegans* growth. We identified a significant correlation between metabolic similarities and phylogenetic relationships inferred from 16S rRNA sequences, which are commonly used for bacterial classification. Metabolic variation was larger than evident from 16S data alone, suggesting that metabolic competences can be derived to only limited extent from 16S sequences and should ideally be reconstructed from whole genome information. For selected bacteria, we validated the model predictions using physiological analyses. Moreover, both in

vitro and in silico approaches demonstrated that the nutrient environment can modulate bacterial interactions, for example, from competition to mutualism. We further identified specific metabolic modules that appear to shape the interaction with the host. Finally, we considered a combination of genomic, metabolic, and cellular traits to infer bacterial life history strategies according to the universal adaptive strategy theory (Fierer, 2017; Grime, 1977), finding that bacterial colonization ability is associated with a competitive or stress-tolerant strategy. In the following, we will discuss in more detail (i) the diversity of metabolic competences in the microbiota and possible implications for *C. elegans* biology, (ii) how metabolic networks shape bacteria–bacteria interactions, and (iii) how bacterial traits affect colonization and *C. elegans* fitness.

Our analysis revealed that the microbiota members are jointly able to synthesize all essential nutrients required by *C. elegans*. The considered isolates varied in their capacity to produce vitamins essential to *C. elegans*, such as folate, thiamine, and vitamin B12, which are known to affect nematode physiology and life history (Bito et al., 2013; Cabreiro et al., 2013; Chaudhari et al., 2016; MacNeil et al., 2013b; Virk et al., 2016; Watson et al., 2016). For example, vitamin B12 influences propionate breakdown, it can accelerate development, and reduce fertility (MacNeil et al., 2013b; Watson et al., 2016). Of the characterized bacteria, only *Pseudomonas* and *Ochrobactrum* strains had the pathways to produce vitamin B12. Their enrichment in the microbiota should therefore affect the metabolic state and fitness of *C. elegans*.

Our study demonstrated that the nutrient environment can change bacterial interactions. In our simulations, competitive interactions dominated in rich medium, while parasitic and mutualistic interactions in minimal medium. Interactions between *Pseudomonas lurida* MYb11 and *Ochrobactrum* MYb71 shifted from parasitic to competitive in a sucrose- vs. glucose-supplemented medium. We detected a secreted sucrose invertase in the MYb71 genome, which otherwise lacks sucrose transporters. Thus, we propose that MYb71 breaks down sucrose extracellularly, and the monosaccharides glucose and fructose become exploitable by MYb11. While a similar phenomenon was described for yeast with engineered auxotrophies (Hoek et al., 2016; Zomorodi and Segre, 2017), it was here observed for naturally coexisting host-associated bacteria. This emphasizes the relevance of nutrient context in host microbiota interactions. Importantly, no single growth medium might reliably predict all possible interaction types. It is therefore essential to consider the nutrient context to fully understand bacterial interactions within the microbiota (e.g., ref. (Coyte et al., 2015)).

Our analysis further identified two bacterial traits that appear to influence the interaction with the host. Colonization ability was associated with pyruvate fermentation to (S)-acetoin. This fermentation pathway includes the ketone diacetyl as an intermediate, whose buttery odor attracts *C. elegans* and promotes feeding behavior (Ryan et al., 2014). In detail, diacetyl binds the transmembrane odor receptor odr-10 and affects odortaxis (Ryan et al., 2014; Sengupta et al., 1996; Zhang et al., 1997). As a result, worms are more attracted to bacterial lawns with this particular smell (Ryan et al., 2014). Indeed, lactic acid bacteria in rotting citrus fruits were more attractive to worms when releasing diacetyl (Choi et al., 2016). Similarly, entomopathogenic *Steinernema* nematodes were more attracted to insect cadavers infected with the diacetyl-producing bacterial symbionts of the nematode (Baiocchi et al., 2017). Thus, if worms are attracted to diacetyl-producing bacteria, they should spend more time in their presence. This alone could increase bacterial uptake and colonization.

We also found that trans-3-hydroxyproline degradation in bacteria is associated with increased nematode fitness. In worms, hydroxyproline is present in collagen type IV, a major component of the extracellular matrix in the pharynx, intestine, and cuticle (Graham et al., 1997; Hutter et al., 2000; Page and Johnstone, 2007). The breakdown of hydroxyproline can generate reactive oxygen species (Cooper et al., 2008). These may act as signaling molecules, which could affect cellular proliferation (Halliwell, 2007) and *C. elegans* reproduction (De Henau et al., 2015). Whether ROS in the gut increases brood size is unknown. Alternatively, bacteria with the degradation pathway may utilize the amino acid as a carbon source,



consistent with the “microbiome on the leash” hypothesis, characterized by host-selection of beneficial bacterial traits (Foster et al., 2017).

In conclusion, our study provides a resource of naturally associated bacteria, their whole-genome sequences, and reconstructed metabolic competences that can be exploited to study and understand *C. elegans* in an ecologically meaningful context. This resource may help to further establish *C. elegans* as a model for studying host–microbe interactions.

## Acknowledgements

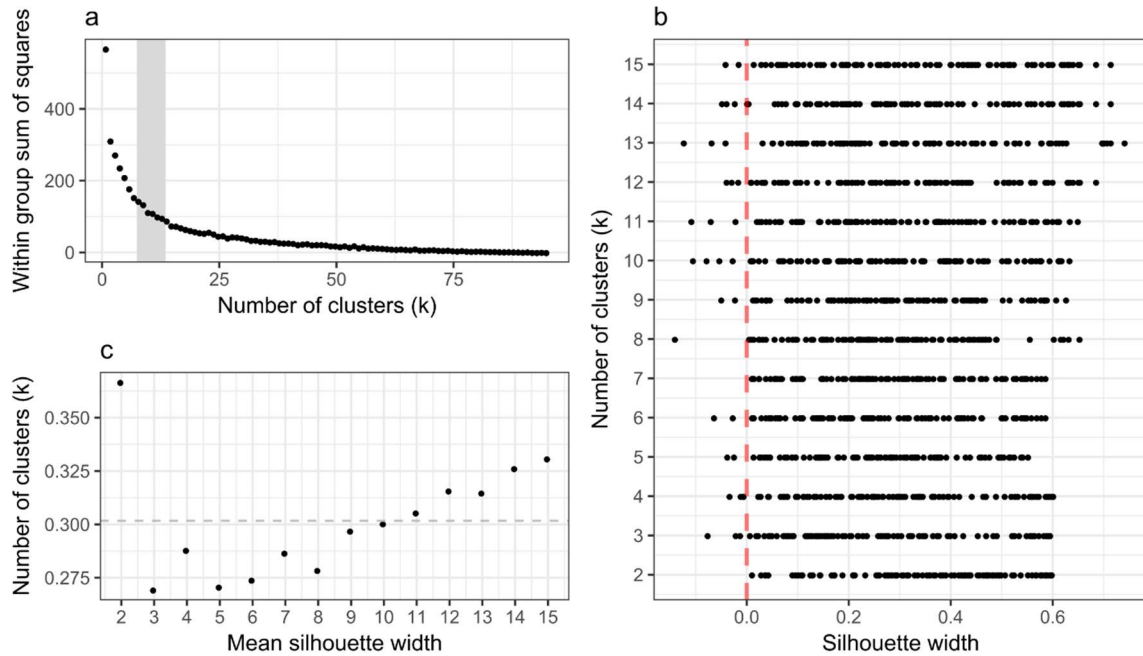
We thank Simone Severitt and Nicole Heyer for technical assistance regarding PacBio genome sequencing, Jolantha Swiderski for long-read genome assemblies, Peter Deines for advice on BIOLOG assays, the Kiel BiMo/LMB for access to their core facilities, and the CRC 1182 and the Schulenburg group for general advice.

## Funding

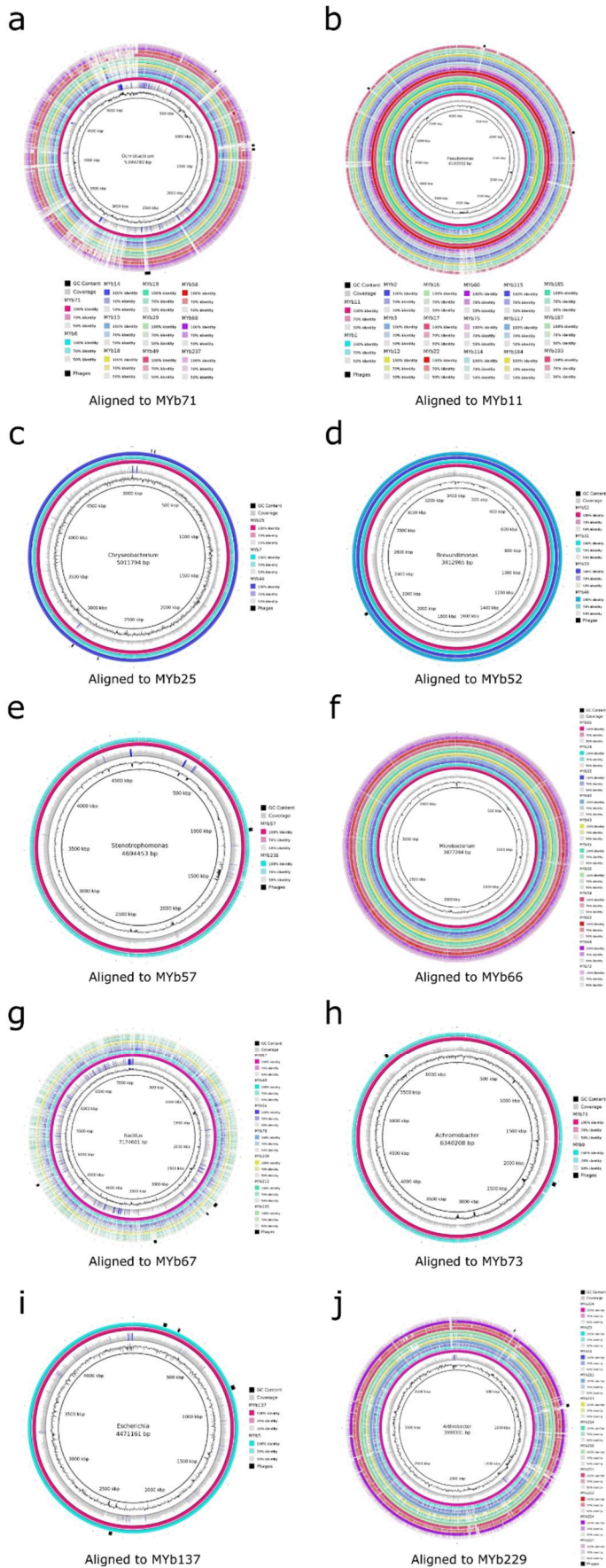
German Science Foundation within the Collaborative Research Center CRC 1182 on Origin and Function of Metaorganisms, projects A1 (KD, ML, HS), A4 (HS), and INF (MPH, CK) and under Germany’s Excellence Strategy – EXC 22167-390884018 (Precision Medicine in Chronic Inflammation; CK, HS); the Competence Center for Genome Analysis Kiel (CCGA Kiel; CK, HS); the Max-Planck Society (Fellowship to HS); and the International Max-Planck Research School for Evolutionary Biology (NO).

## Supplement

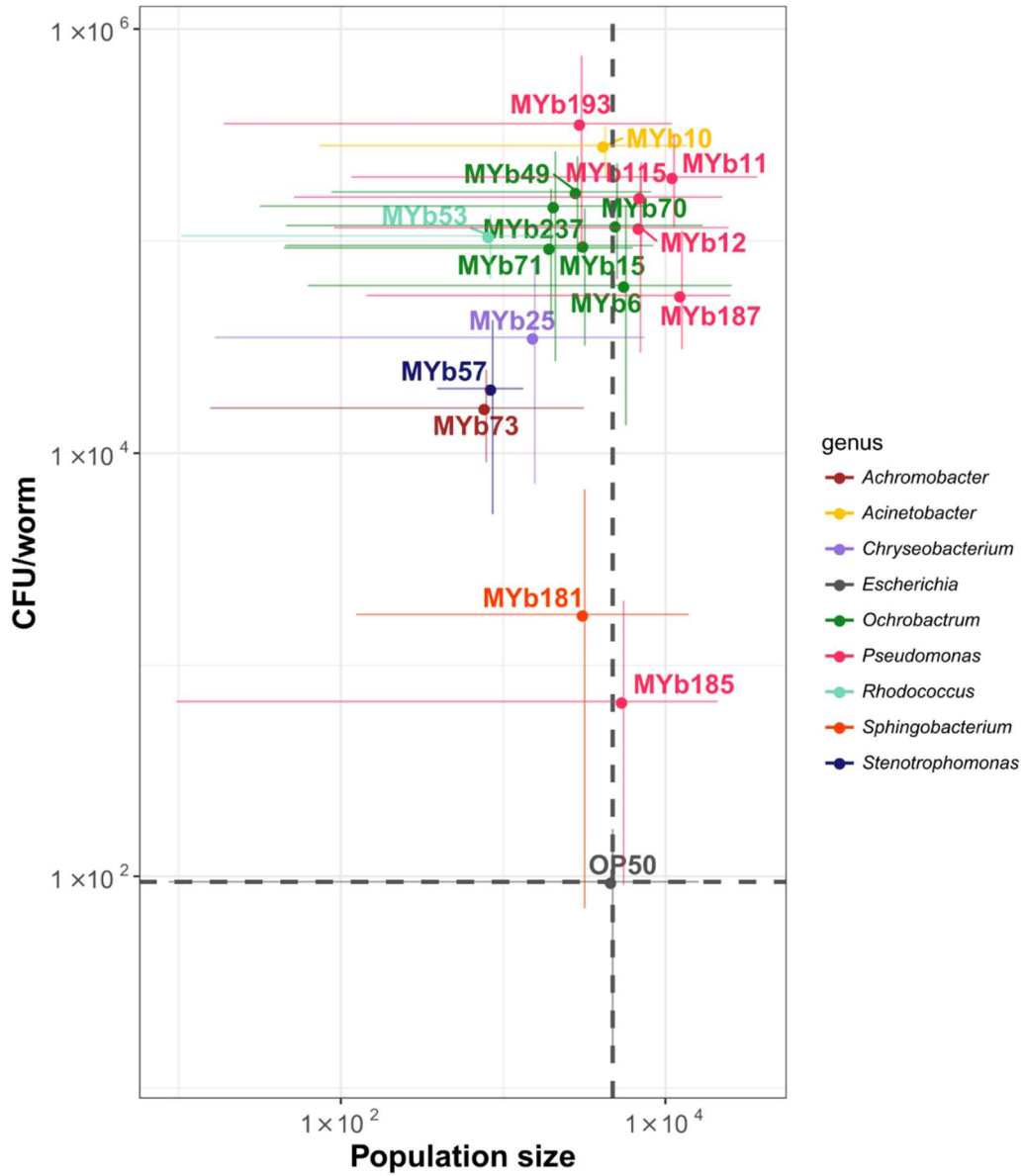
### Supplementary figures



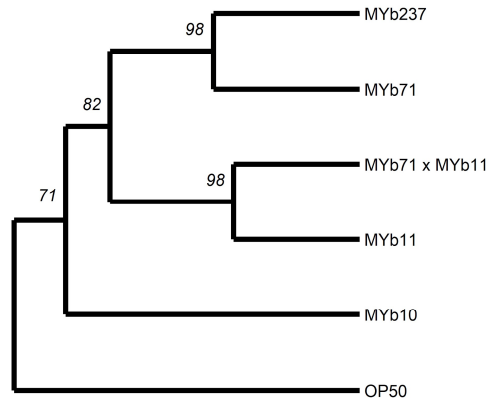
**Supplementary Fig. S1. Diagnostic plots of k-means clustering.** (a) Within group sum of squares (SS) across clustering with different  $k$ . Grey shaded region highlights number of clusters where within group SS starts decreasing less rapidly with an increase in  $k$ . (b) Silhouette width of clustering with different  $k$ . A silhouette width = 0 indicates points of several clusters overlapping, while a silhouette width = 1 implies a point being exclusive to one cluster. (c) Mean silhouette width across clustering with different  $k$  100 times. Dashed grey line indicates overall mean silhouette width observed.



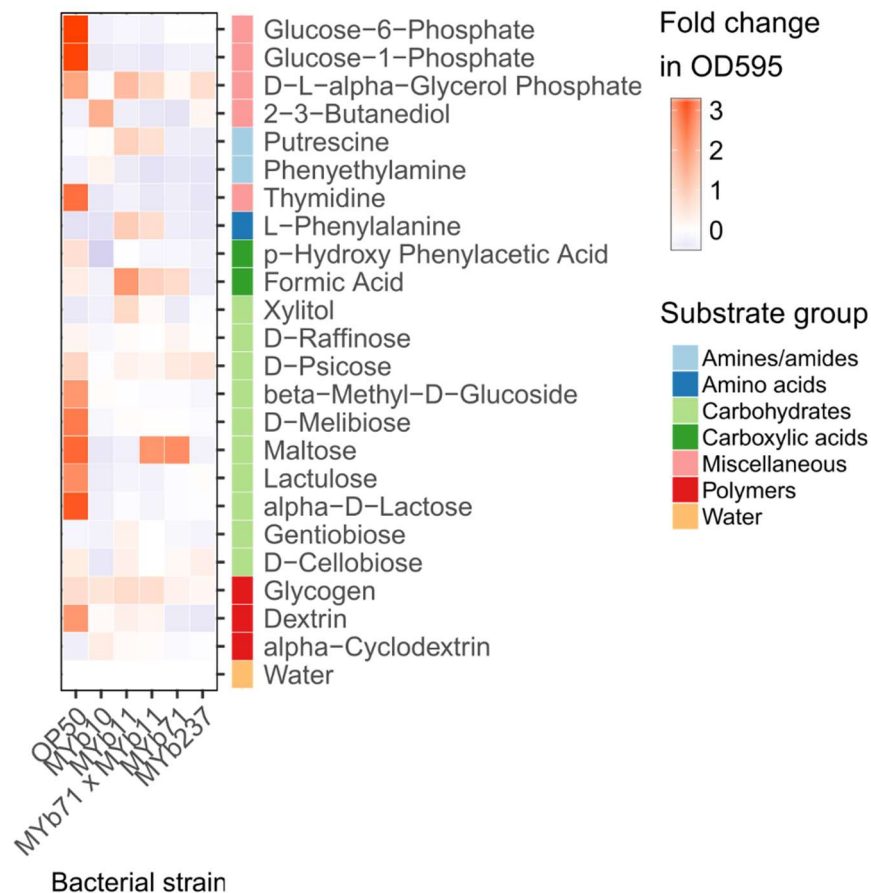
**Supplementary Fig. S2. Genome comparison of selected bacterial taxa using circular plots.** Selected taxa include: (a) *Ochrobactrum*, (b) *Pseudomonas*, (c) *Chryseobacterium*, (d) *Brevundimonas*, (e) *Stenotrophomonas*, (f) *Microbacterium*, (g) *Bacillus*, (h) *Achromobacter*, (i) *Escherichia*, and (j) *Arthrobacter*. For each taxon, the genome sequence with highest quality (i.e., genome sequence with the smallest number of contigs) was identified and then used as a reference for alignment of the remaining genomes. GC content, coverage, and predicted bacterial phage information are given for the reference genome of each taxon. The color intensity in each ring indicates BLAST match identity. The genomes within a particular taxon are usually highly similar, except in the case of *Bacillus*, which produces higher diversity across the included isolates. Predicted phage regions tend to be highly diverse across isolates within a taxon, except for *Chryseobacterium*, *Brevundimonas*, *Stenotrophomonas*, *Achromobacter*, and *Escherichia*, which may be due to the small number of species in these taxa.



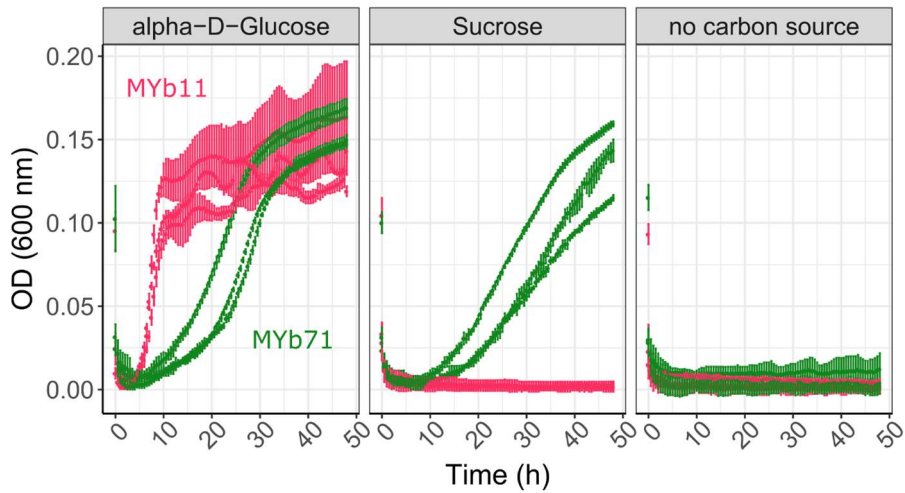
**Supplementary Fig S3. Variation in population growth and colonization levels of *C. elegans* in mono-association with natural microbiome isolates.** Three L4 larvae were exposed to bacterial lawns on PFM plates for five days, and F2 population sizes quantified ( $n = 3-6$ ). To count colony forming units per worm, L4 larvae were transferred from NGM plates with OP50, exposed to microbiota lawns for 24 h, washed and the associated bacteria extracted ( $n = 5$ ). Dashed lines show the mean population size and bacterial load of the canonical food bacteria *E. coli* OP50.



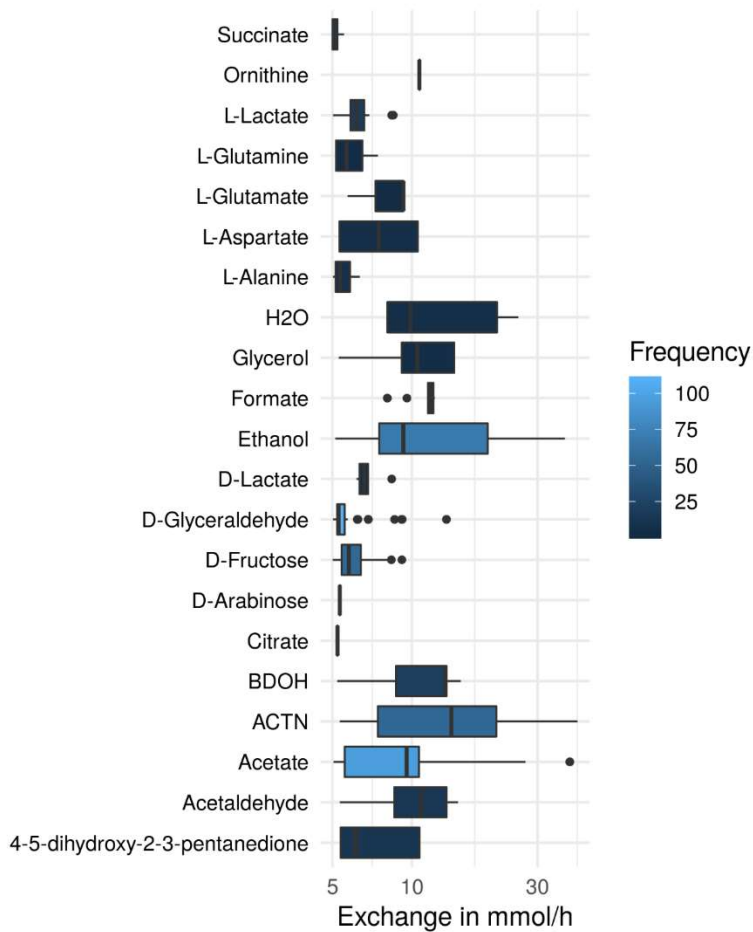
**Supplementary Fig. S4. Hierarchical clustering of strains based on BIOLOG profiles.** Clustering is based on Ward's algorithm and Euclidean distance measures. Bootstrap support ( $n_{boot} = 1000$ ) is shown on nodes.



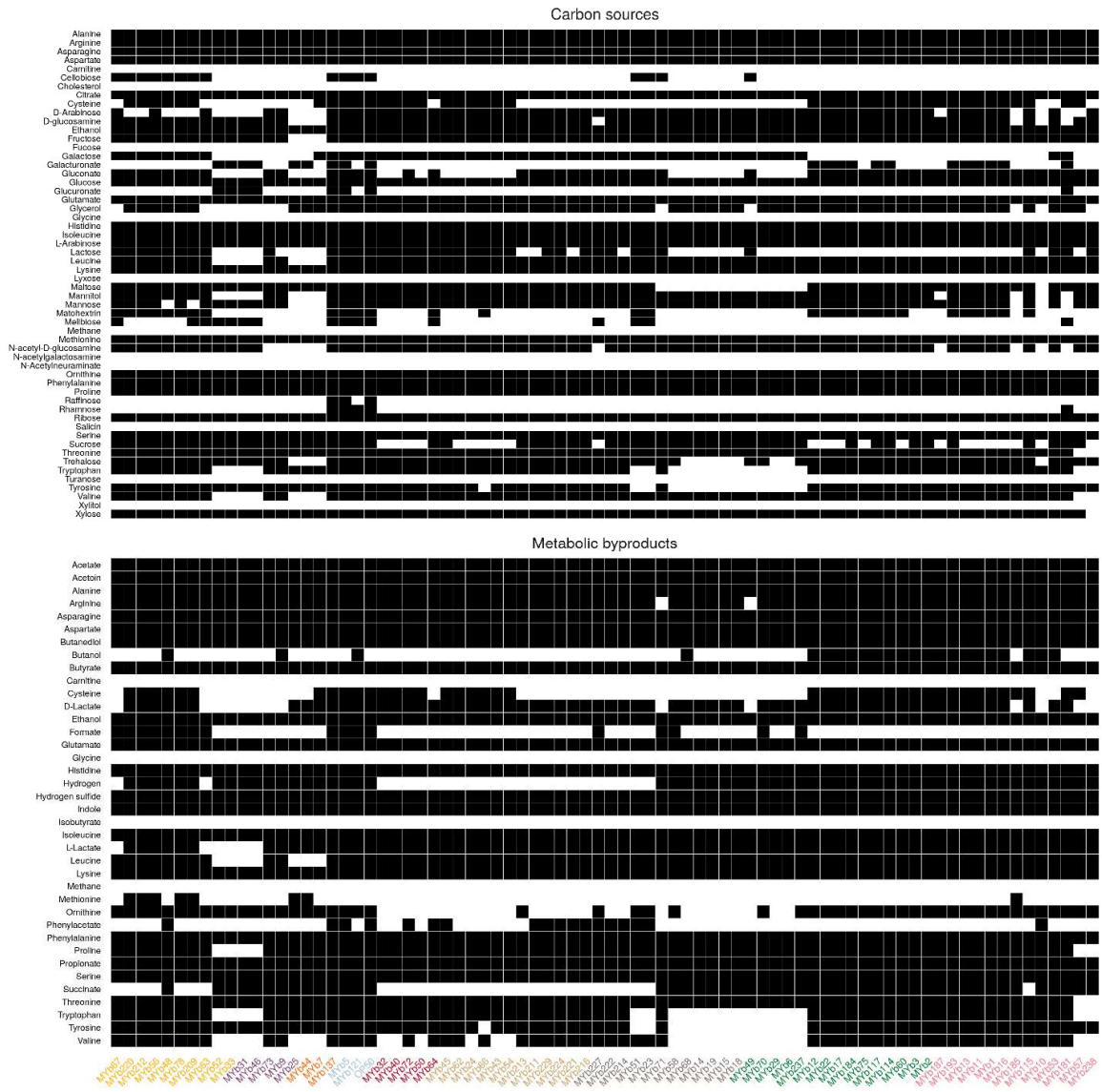
**Supplementary Fig. S5. Cluster 7 of BIOLOG profiling.** Profiles of carbon substrate use of *Acinetobacter* sp. (MYb10), *Pseudomonas lurida* (MYb11), *Ochrobactrum* sp. (MYb71), *Ochrobactrum* sp. (MYb237), and *E. coli* OP50 in BIOLOG GN2 plates over 46 h. The fold-change in indicator dye absorption from 0 to 46 h indicates that the indicated compound is metabolized. K-means clustering ( $k = 7$ ) of substrates by fold-change highlights metabolic differences between strains. Clusters I - VI are shown in figure 3 in the main text.



**Supplementary Fig. S6. Culture of MYb11 and MYb71 in defined media with single carbon substrates.** Growth of MYb11 and MYb71 in chemically defined media with a single carbon source (i.e., alpha-D-glucose or sucrose), or no-carbon control (over 46 h in 96-well plates).



**Supplementary Fig. S7. Metabolites exchanged in in silico pairwise interactions on a minimal medium.** Substances which were exchanged between bacterial isolates in simulation of ecological interactions. In a minimal medium, metabolic byproducts could influence the growth rates of organisms. The figure shows metabolites, which were predicted to be exchanged most frequently and in highest quantity.



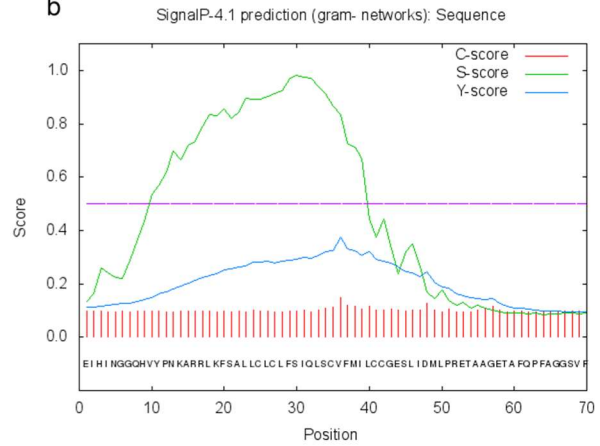
**Supplementary Fig. S9. Prediction of carbon sources and metabolic byproducts.** In simulation of bacteria metabolism, we predicted substances which could be used as carbon sources (left panel) and substances that may be secreted as metabolic byproducts during for example fermentative processes (right panel). Microbiome isolates are given along the x-axis and compounds along the y-axis.

a

**>Sucrose invertase in MYb71 genome**

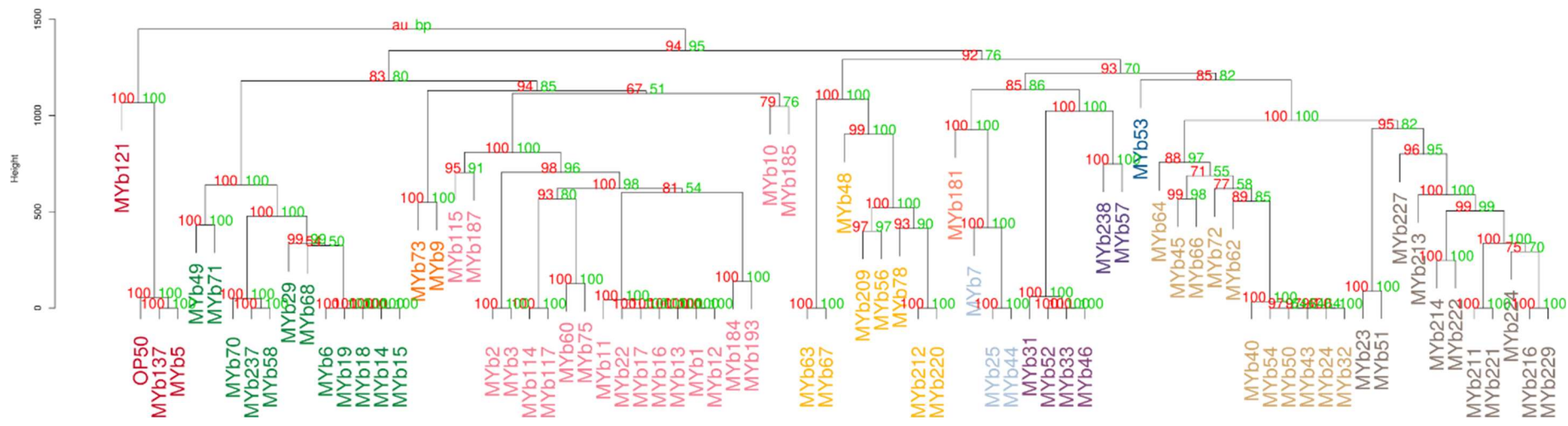
```
EIHINGGQHV*----YPNKARRLKFSALLCLCLF*SI*QLSCV
FMILCCGESLIDMLPRETAAG--ETAFQPFAGGSVFNTAIA
LGRLDVPTGFFSGISSDFFGEVLRDNLARSNVDYSFAAIS
DRPTT-LAFVRL-VDGQARYAFYDENTAGRMLTESDMPY-
VDDAIDAMLFGCISLISEPCGSVYEALMT-REAPRRVMFL
DPNIRAGFITDREKHLHRMKRMIALADIVKLSDEDLAWF
GEKGSHEIAAEWLKLGPKLVITKGAGADAYTAKATV
RVPGVKVDVVDTVGAGDTVNAGILASLHNQGLLDKDAL
VELTEDQIHSAVALGVRAAAVTVSRAGANPPW
```

b



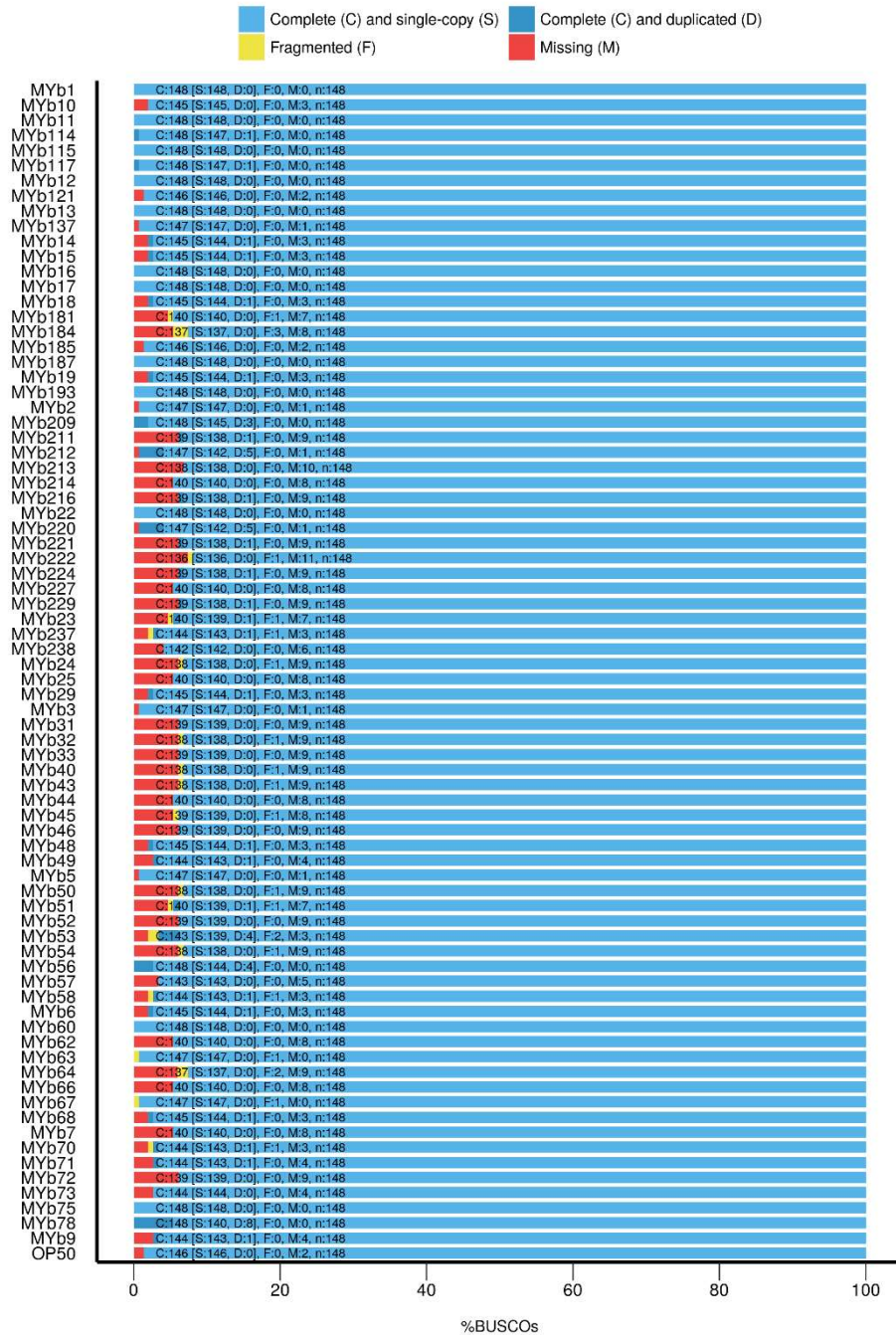
**Supplementary Fig. S10. External sucrose invertase in MYb71.** (a) Amino acid sequence of the sucrose invertase found in the MYb71 genome. (b) We used SignalP to check for secretory signatures. We found that a discrimination score of  $D = 0.518$  (i.e. weighted average of the mean  $S$  and the max.  $Y$  scores) which indicates a signal peptide (signalP: Name=Sequence SP='YES' Cleavage site between pos. 35 and 36: LSCVF  $D=0.518$   $D$ -cutoff=0.420 Networks=SignalP-noTM).



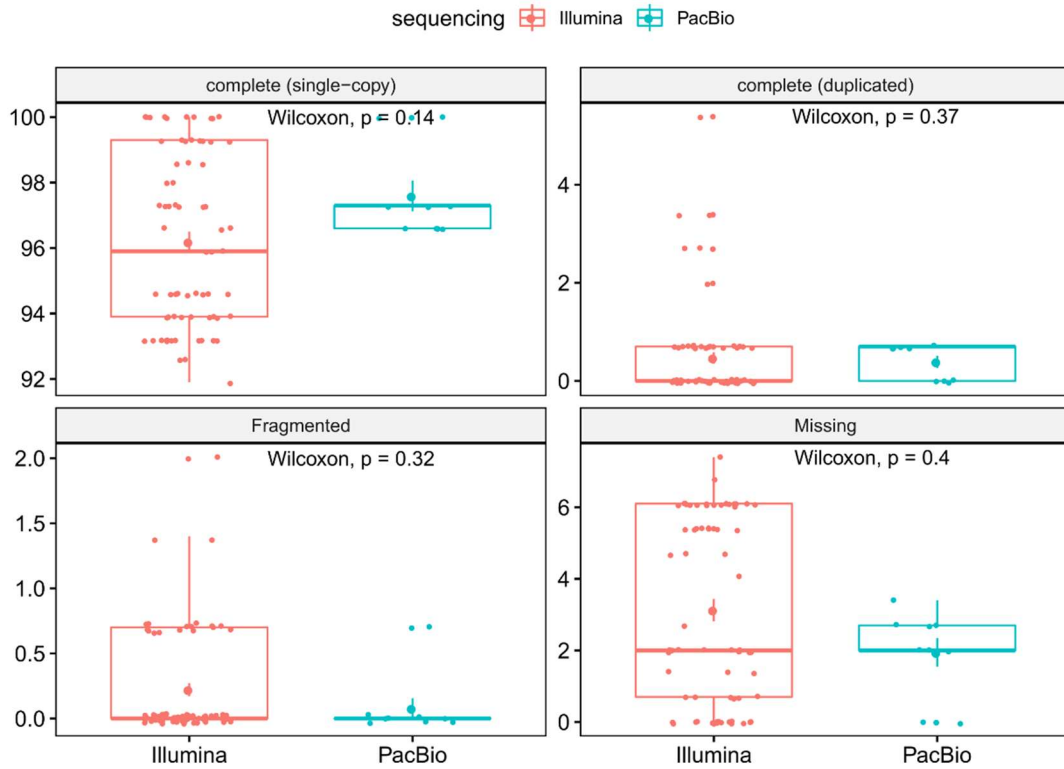


**Supplementary Fig. S11. Hierarchical clustering of metabolic networks based on pathway prediction.** Metabolic networks were clustered according to their pathway completeness score by Euclidean distances and similarity of clusters was estimated by average linkage. The quality of the clustering was tested by multiscale bootstrap resampling. Green values next to branches indicate the bootstrap probability (number of 90 means e.g. that the cluster exists in 90 of 100 runs). In addition to this, the approximately unbiased *p*-value from multiscale bootstrap is shown in red (see (1) for details).

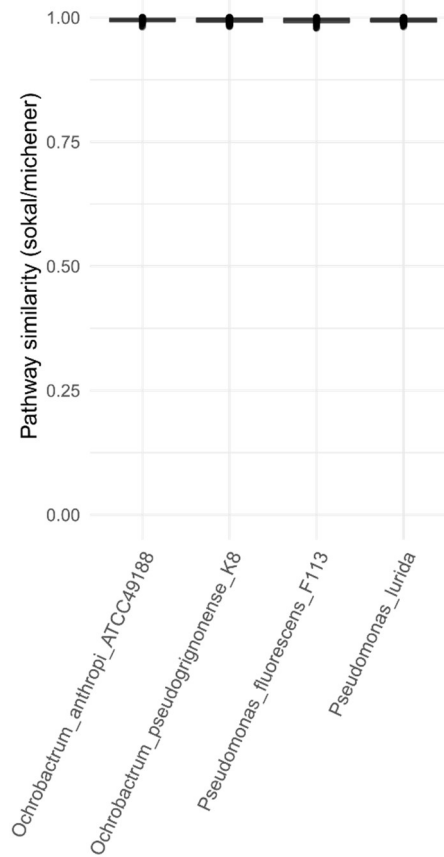
## BUSCO Assessment Results



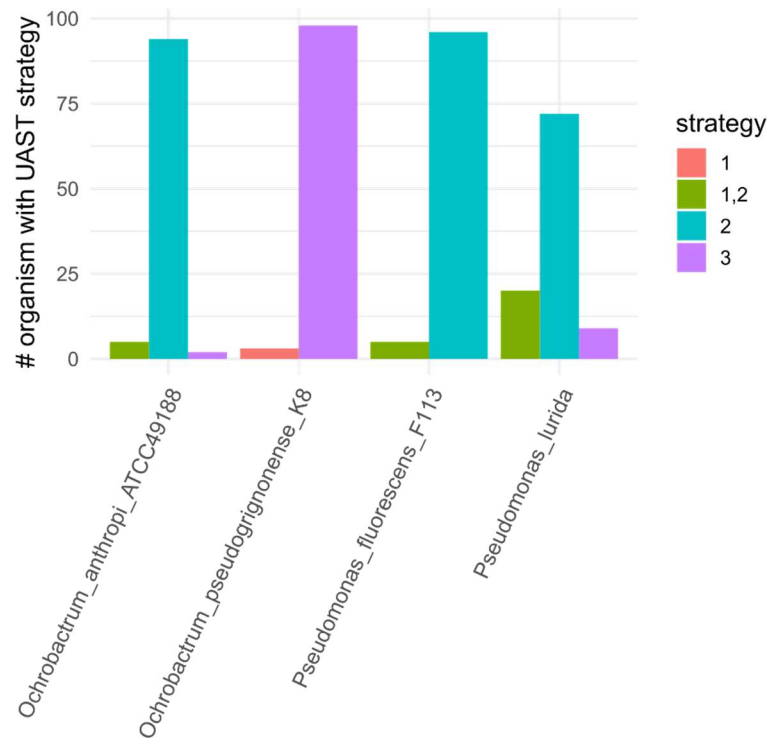
**Supplementary Fig. S12. Genomic completeness as assessed with BUSCO.** Overview of complete, fragmented and missing matches that were found for all sequenced genomes. The following categories are considered by BUSCO (2): i) complete and single copy: high scoring (i.e.: 90% of the minimum bitscore from an HMM search) and large alignment length matches with one copy; ii) complete and duplicated: high scoring and large alignment length matches with several copies; iii) fragmented: high scoring but shorter alignment length matches; and iv) missing: low scoring matches. The mean completeness level was 96.81% with a standard deviation of 2.65%. For further details, see the BUSCO user guide ([http://gitlab.com/ezlab/busco/raw/master/BUSCO\\_v3\\_userguide.pdf](http://gitlab.com/ezlab/busco/raw/master/BUSCO_v3_userguide.pdf)).



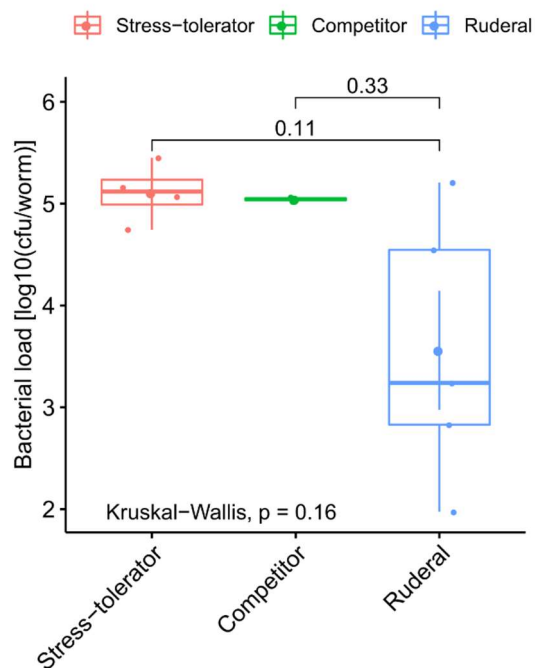
**Supplementary Fig. S13. Comparison of genome completeness for different sequencing techniques.** Complete, fragmented and missing matches found by BUSCO were compared between genomes reconstructed with either Illumina or PacBio technology, in order to assess to what extent sequencing technique influenced draft genome quality. The four BUSCO categories are the same as those explained in the legend to Supplementary Figure S12. The difference in inferred values between Illumina- and PacBio-sequenced genomes was tested with a Wilcoxon rank sum test. None of the comparisons was significant.



**Supplementary Fig. S14. Pathway similarity of incomplete genomes.** To assess the influence of incomplete genome coverage on inferred metabolic pathways, we downloaded four high quality genomes of our two focal genera (*Pseudomonas* and *Ochrobactrum*) from NCBI. We used the following, published, high quality genomes for this analysis: i) *Ochrobactrum anthropi* ATCC49188 (RefSeq assembly accession numbers: GCF\_001652485.1\_ASM165248v1); ii) *Ochrobactrum pseudogrignonense* K8 (GCF\_000017405.1\_ASM1740v1); iii) *Pseudomonas fluorescens* F113 (GCF\_000237065.1\_ASM23706v1); and iv) *Pseudomonas lurida* (GCF\_002966835.1\_ASM296683v1). We used these genomes to randomly remove sequence chunks. For this, we simulated similar levels of genome completeness as in our data by drawing from a normally distributed sample parameterized by the results of our BUSCO genome analysis (mean = 96.81%, sd = 2.65%; Supplementary Figure 12). The presence of metabolic pathways for randomly shortened genomes was compared to the prediction for the original reference genomes. The overlap ranged from 97.8 to 100%. The overall mean pathway similarity between 400 randomly shortened genomes and their complete counterparts was 99.4%.



**Supplementary Fig. S15. Stability of adaptive strategy predictions.** The robustness of UAST classification was tested with randomly shortened representative genomes. High-quality reference genomes were randomly shortened to the same extent as found by BUSCO for the genomes of this study. See legend to Supplementary Figure S14 for more details. For each taxon, 100 shortened genomes were generated and then reclassified according to the universal adaptive strategies (UAST) as done for all other genomes of this study. The bars show how many times a certain strategy was predicted for the incomplete genomes of a particular taxon. The analysis identified the same strategy in at least 94% of shortened genomes for *Ochrobactrum anthropi* ATCC49188, *Ochrobactrum pseudogrignonense* K8, and also *Pseudomonas fluorescens* F113. For *Pseudomonas lurida*, 72% of shortened genomes were consistent with strategy 2 and 20% still with both strategies 1 and 2. A description of the strategies and their definitions is given in (Fierer, 2017).



**Supplementary Fig. S16. Bacterial colonization behavior in comparison to adaptive strategies for Illumina genomes.** Since sequencing technique could impact rRNA copy number detection, which is used for adapted strategies classification, we repeated the analysis regarding the colonization potential for Illumina genomes only (69 isolates). Here we found the same qualitative results that competitive and stress-tolerating strategies were associated with higher colonization phenotypic data.

## Supplementary tables

All supplementary tables can be found at <https://www.nature.com/articles/s41396-019-0504-y> and only title legends are reprinted here for reference.

**Table S1: Genome characteristics (Samples\_all.xlsx).** Genomic overview of all bacterial isolates. Based on sequence identity, the closest related species/strains are shown in column B. Genome and assembly statistics (contigs, N50, GC, ...) can be found in columns G:M. The corresponding bioproject to access data via NCBI is shown in column N.

**Table S2: In silico TSB-based medium (In silico TSB medium.csv).** Compounds and maximal uptake rates for Tryptic soy broth (TSB) medium used to simulate bacterial growth.

**Table S3: In silico glucose minimal medium with thiamine (in silico glucose minimal medium with thiamine.csv).** Compounds and maximal uptake rates for glucose minimal medium with thiamine used to simulate bacterial growth.

**Table S4: Predicted pathways (predicted\_pathways.xlsx).** For each bacterial isolate (column A) the predicted presence (column D and column E for with more conservative bitscore cutoff) for all considered metabolic pathways (column C) is given. Here, a number of 1 means present and 0 means not available. In column G the hierarchy (i.e. subsystem) of the pathway is shown.

**Table S5: Predicted virulence (Predicted\_virulence.xlsx).** In this table the presence of virulence traits based on homology with the virulence factor database is shown. In column A the isolates are given and columns B:AV show the presence of virulence factors (1 for true and 0 for false).

**Table S6: Traits with differences in Ochrobactrum (ochrobactrum.xls).** Table consists of metabolic pathways and virulence factors which showed to be significantly changed in isolates belonging to the Ochrobactrum genus (based on a Wilcoxon signed rank test). For each trait (column B) a FDR corrected P-value (column C) and the mean for Ochrobactrum (column D) and all other isolates (column E) is given.

**Table S7: Experimental phenotypic data: bacterial load and C. elegans population growth.** (TableS\_phenotypes.xlsx) Sheet 1 shows the number of colony forming units per worm (column D) shown across microbiome isolates (column A) and experimental repetitions of the analysis (i.e., runs; column C). Sheet 2 shows the mean number of worms counted (3 samples counted per population) per population of worms (column C) and standard deviation (column D) on the respective microbiome isolates (column A).

**Table S8: Regression analysis to infer metabolic competences associated with bacterial colonization and host fitness.** (colonization\_regression.xlsx) Two regression approaches were used to find traits which were associated with experimental data of bacterial load in C. elegans and the fitness of C. elegans when the worm was grown together with the bacterial isolates. In this table, the significant/important traits of both approaches are listed.

**Table S9: Traits and scores used to categorize isolates according to adaptive strategies.** (adaptive\_strategies.xlsx) Table of metabolic traits and model features associated with stress-tolerating, competitive, or ruderal strategies. For each isolate the scores and classification for each strategy are listed.

Supplementary data

Supplementary data can be found at <https://www.nature.com/articles/s41396-019-0504-y> and only file names are reprinted here for reference.

**Supplementary data S1. Genome-scale metabolic models of the microbiome of *C. elegans*.** Zip archive of metabolic models for each isolate in Systems biology markup language (SBML) format. In addition to this, an R file with a list of all models is provided.





# 2

## Evolution of microbiota-host associations: the microbe's perspective

Nancy Obeng  
Florence Bansept  
Michael Sieber  
Arne Traulsen  
Hinrich Schulenburg

Manuscript ready for submission

## **Abstract**

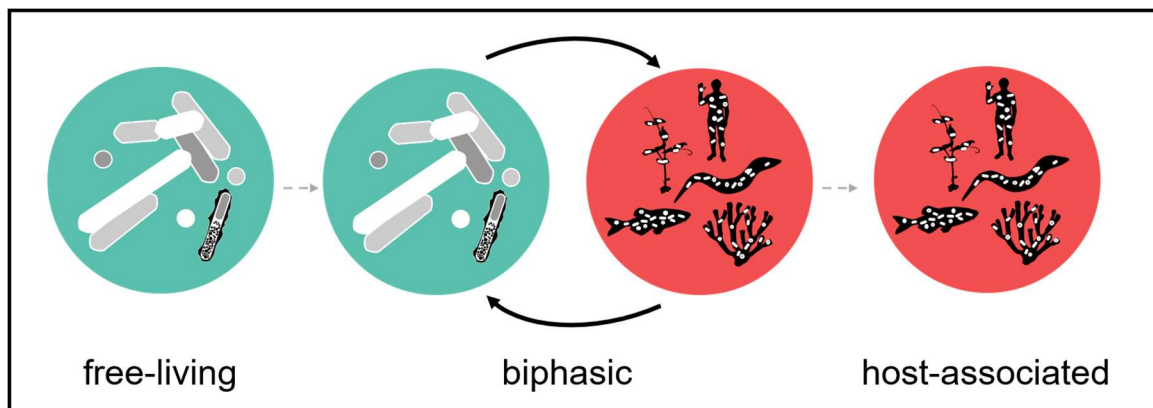
Microbiota-host associations are ubiquitous in nature. They are often studied using a host-centered view. We here put the spotlight on the microbes and how they optimize their fitness in association with the host. We propose that on the route to exclusive host association, free-living bacteria follow a biphasic life cycle, alternating between hosts and free-living habitats, as still observed for many extant microbiota-host interactions. We develop a new integrative measure of microbial fitness across the distinct stages of the life cycle and explain how microbes are subjected to selection at different levels of biological organization. Together, our approach provides a new life-history framework to assess emergence and maintenance of microbiota-host associations with a particular focus on microbial fitness.

## Highlights

- A microbial perspective is needed for full understanding of microbiota-host symbiosis
- A biphasic life cycle likely shapes the origin of a microbe's association with host
- Microbes must thus succeed in the host, the environment and reciprocal transmission
- Symbiosis is driven by microbial fitness integrated across the entire life cycle
- A new formalized integrative fitness measure can guide novel research on symbiosis

## On the origin of microbiota-host associations

Microbiota-host associations are widespread in nature and thus essential for our understanding of life (Browne et al., 2017; Dodds et al.; McFall-Ngai et al., 2013). Host and microbial partners can benefit from the association in different ways. The microbes may act as food, provide essential nutrients, catabolize otherwise indigestible food (Consuegra et al., 2020; Zhu et al., 2011; Zimmermann et al., 2019), protect the host from pathogens and toxic substances (Kissoyan et al., 2019; Wang et al., 2020), or regulate host physiology, including immunity, development and even host social behavior (e.g. (Douglas, 2018; Sarkar et al.)). Microbes, too, can profit from the association, although this has been much less studied. For them, the host can be a habitat (Donaldson et al., 2016), or a vector for their transmission between habitats otherwise inaccessible (Diaz and Restif, 2014; Moeller et al., 2018). But how did these associations originate and which selective processes favored their emergence and maintenance? These questions are usually addressed by taking a host perspective and pointing to the various, above listed benefits provided by the microbes (e.g. (Berg et al., 2019; Franzenburg et al., 2013; Kaltenpoth et al., 2014)). Here, we argue that a microbial perspective is similarly if not more important for a full understanding of the evolution these associations. The main reason is that microbes usually show a much higher evolutionary potential than their hosts because of their comparatively shorter generation times, larger populations, and/or often haploid genomes. Thus, they should be able to adapt much faster to a host environment and also avoid hosts, if there is no benefit. Moreover, if we assume that a host-free lifestyle in some environmental habitat is ancestral, then a loose temporal association of the microbes with a more complex organism is likely the first step in symbiosis (Figure 1). Consequently, a characterization of the selective benefits across such a biphasic life cycle will likely help our understanding of the evolutionary origins of these associations. Below, we first highlight the distribution of biphasic life cycles in symbiosis. We then explore how microbial fitness is shaped across the different stages of the biphasic cycle and in box 1 we use a mathematical approach to develop a corresponding, integrative measure of microbial fitness that is analogous to the parasite fitness parameter  $R_0$ . We further discuss at which levels of biological organization selection acts upon the microbes and point to promising future research avenues. To enhance clarity, we on purpose focus on the interaction between a single microbial taxon and its host, as this allows us to specifically highlight the selective constraints central for the association with the host. We thus neglect ecological dynamics shaped by microbe-microbe interactions, although these can be added in future extensions of our framework. For simplicity, we mainly use bacterial symbionts as examples; yet most arguments also apply to other microbial associates.



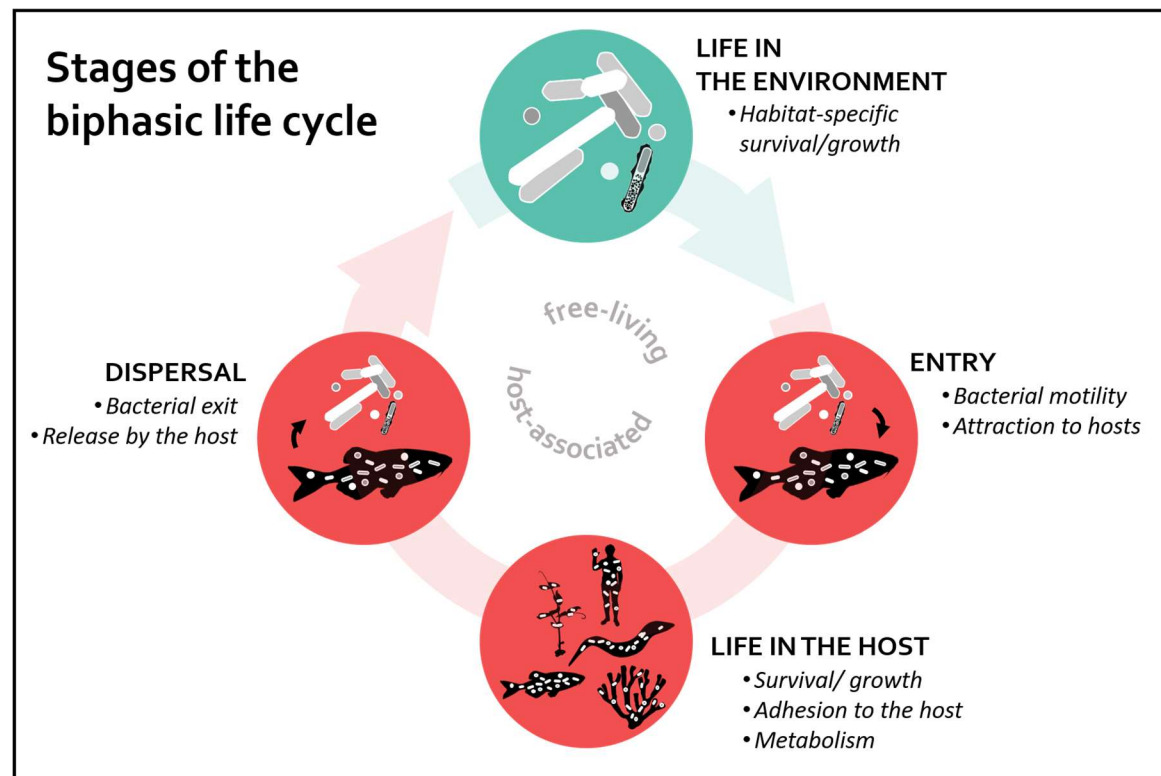
**Figure 1 Evolution from free-living to host-associated life via a biphasic life cycle.** Bacteria may follow a purely free-living (left), exclusively host-associated (right) or biphasic life history (middle), transitioning between host and free-living habitat. The biphasic cycle presents both a steppingstone for free-living bacteria towards host association, but also a symbiosis strategy on its own. Images of hosts were collected from [phylopic.org](http://phylopic.org).

## Transitioning between host and environment is common for many microbiota-host associations

A biphasic life cycle is widespread among extant symbioses. In fact, some form of a biphasic cycle occurs for all but strictly vertically transmitted symbionts, which account for about a third of all symbioses (Russell, 2019), and even here bacteria may transition across different within-host habitats. For example, the obligate association of entomopathogenic nematodes (EPNs) with their bacterial symbionts relies on symbionts shifting between colonization of nematode larvae during dispersal to new insect prey and proliferation and production of virulence factors in the prey (Cao and Goodrich-Blair, 2017; Ciche et al., 2006). Similarly, *Vibrio fischeri*, the luminescence-providing symbiont of the bobtail squid *Euprymna*, is recruited from sea water early during host development and the majority of symbionts growing at excess is expelled into the water again daily (Nyholm and McFall-Ngai, 1998). Presumably a biphasic life cycle also exists for members of the human microbiota traversing the built environment, therefore resulting in co-housed humans being more similar in microbiota composition (Browne et al., 2017).

A bi- or multiphasic life cycle likely applies also to many cases of so-called phylosymbiosis. Here, microbiotas of closely related host species resemble each other more than those of distantly related hosts (e.g. (Berg et al., 2016; Brooks et al., 2016; Moeller et al., 2019)), leading to congruent host and microbiota phylogenies (Brooks et al., 2016; Theis et al., 2016). One interpretation of phylosymbiosis is past coevolution of microbes and hosts. Empirical support for this interpretation, however, has been mixed, and an overlap in ecological factors including diet, habitat, or geography appears to be rather involved in shaping the observed patterns (Kohl, 2020). Moreover, strict coevolution, due to reciprocal adaptations between host and microbes, should manifest at the strain or at least species rather than the family or genus level. While this might be true for intracellular symbionts (Douglas and Werren, 2016), evidence for strict coevolution is lacking for most other associations. From honey bees, deep-sea mussels, and *C. elegans* to humans, host individuals vary substantially in associated bacterial species and strains, thus shifting taxonomic conservation across microbiotas above the genus level (Ansorge et al., 2019; Ellegaard and Engel, 2019; Schloissnig et al., 2013; Zhang et al., 2017). This suggests that host individuals are colonized by a set of non-coevolved microbes from the environment at some or several time-points during their life. The pattern of phylosymbiosis is produced, because microbes

evolved to exploit a certain type of host habitat and related hosts simply provide similar habitat types (Douglas, 2018). Alternatively, related hosts have evolved to select similar (i.e., related) microbes from the environment, in order to acquire specific functions from the microbes (Douglas, 2018). Irrespective of the explanation, microbes always transition between a host and an environmental habitat in these cases.



**Figure 2** Microbial fitness is defined by four stages across the biphasic life cycle. Assuming that a free-living lifestyle is ancestral, then three of the four stages are driven by adaptations to the host habitat (in red), including (i) microbial entry into the host (transmission from environment to host), (ii) life in the host, and (iii) dispersal to the environment (transmission from host to environment). The free-living phase (turquoise) is characterized by persistence and replication in the specific environment. Some of the traits required at the individual stages are indicated in bullet points.

Along the biphasic life cycle, microbial fitness is influenced by persistence and proliferation in either of the two habitats and successful transmission between the two. This framework is related to the concept of nested ecosystems, describing microbes as nested within hosts nested in the environment, which already accounts for interrelated levels of organization and the importance of the environment (Douglas, 2018; McFall-Ngai et al., 2013). Our framework specifically emphasizes the importance of four distinct stages across the lifecycle that jointly define microbial fitness and permit us to formalize an integrative fitness measure (Box 1). We assume that a free-living life-style is ancestral and thus propose that three of the four stages are mainly driven by adaptation to the host habitat, including entry into the host, maintenance and replication within the host, and release from the host (Figure 2). Different trait functions are likely key for bacterial success during each stage.

Firstly, microbes need to enter a host. Here motility and attraction of a host can play a role. By actively moving, microbes may influence their chance of meeting and entering a host. This includes entry from the environment and within-host migration (Raina et al., 2019). The mechanisms used by bacteria (e.g. flagellar-propelled motion, gliding, twitching, or surfing (Raina et al., 2019)) are shaped by host and environmental structure, and bacterial identity (Lebov et al., 2020; Robinson et al., 2018; Soto et al., 2014). Instead of moving

## Glossary

**Microbiota** = sum of microorganisms associated with a host

**Symbiosis** = association, or “living together” (De Bary, 1879), of at least two different organisms

**Biphasic life cycle** = microbial life history including a host-associated and a free-living phase connected by transmission between them

**Phylosymbiosis** = congruence in phylogeny of host and associated microbiota

**Bet hedging** = concept of fitness optimization: under fluctuating conditions, individuals vary in phenotype stochastically to reduce variation in fitness over time and thus increase their geometric mean fitness

**Phenotypic plasticity** = ability to express distinct phenotypes in response to the environmental conditions

**SIR model** = Mathematical framework to describe the epidemiology of infectious disease by dividing the host population into three groups, consisting of susceptible (hosts susceptible to the disease, S), infectious (infected individuals which can infect others, I), and recovered hosts (hosts which are now immune, R), and dissecting the rate of change among these groups dependent on parasite transmission

towards hosts, microbes may also attract them and thereby indirectly stimulate their uptake. This can involve settlement in areas attractive to hosts, secretion of odors, or even modulation of host preferences (O'Donnell et al., 2020; Zimmermann et al., 2019).

After entering a host, microbes need to survive and persist in the host, where metabolism may fuel growth. This usually includes evading a variety of abiotic stressors, host digestion and immunity. Here bacteria, which have experienced similar conditions in a previous habitat, are more likely to succeed. One strategy is formation of a protective barrier using biofilms (Flemming et al., 2016). Biofilms can further help to stick with the host, ideally in a location fulfilling bacterial resource needs (Brooks et al., 2014; Donelli et al., 2012; McLoughlin et al., 2016; Powell et al., 2016). To improve fitness, microbes may grow by metabolizing resources from the host, its diet, or other host-associated microbes (Douglas, 2020; Fisher et al., 2017; Ford et al., 2016; Mazumdar et al., 2020). Notably, symbioses can be stable even when hosts digest microbial partners, so long as part of the symbiont population survives, as for example known for EPN symbioses (Cao and Goodrich-Blair, 2017; Ciche et al., 2006; Hoang et al., 2019).

To complete the life cycle, microbes need to exit the host and enter the environment again. While dispersal has been extensively studied for pathogens, less is known for microbiota taxa. Generally, microbial release may be caused by the host, via peristalsis (Dimov and Maduro, 2019) or regurgitation (Ciche et al., 2006), possibly even induced by the microbes. Alternatively, microbes actively exit the host, for example via increased motility (McLoughlin et al., 2016), aggregate formation (Bansept et al., 2019), or direct association with gut content exiting the host.

Closing the circle, microbes also have to survive and proliferate in the environmental habitat. This is likely dependent on the bacteria's ability to metabolize nutritious substances and activate defense responses (in case of abiotic stressors). The exact favorable life history strategy depends on the particular environmental conditions and may thus rely on phenotypically plastic responses.

Overall, a detailed and integrative analysis of bacterial fitness across the life cycle will be of central importance for our understanding of microbiota-host associations, including both their initial emergence and their long-term maintenance. Such an analysis may benefit from host models that are amenable to experimental manipulation, thus allowing control of the relevant abiotic and biotic parameters (Bosch et al., 2019; Douglas, 2019). As a complement, the involved parameters may also be explored with the help of mathematical models, for example those based on a re-formalization of epidemiological models, which focused on parasite fitness in terms of the basic reproductive number,  $R_0$ , and which were crucial in

understanding the emergence, spread and distribution of infectious disease (see Box 1).

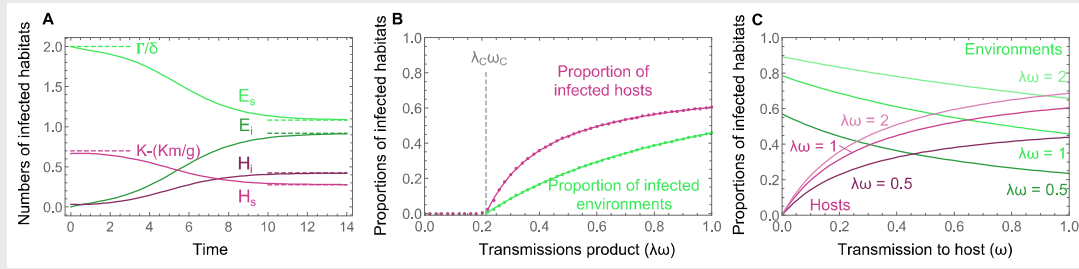
**Box 1: A mathematical framework for an integrative microbial fitness measure**

The basic reproduction number  $R_0$  used in epidemiological modeling recapitulates the chances of success for a parasite or pathogen to start an epidemic based on the key parameters of its life cycle. We derive a similar quantity for the chances of microbial spread in a model which captures the essential components of a biphasic life cycle. The model resembles a coupled SI model with two types of habitats for microbial populations: hosts ( $H$ ) and environments ( $E$ ), both of which can either be microbe-free ( $H_S$  and  $E_S$ ) or colonized by microbes ( $H_I$  and  $E_I$ ). Environments are colonized upon encounter with an infected host with transmission rate  $\lambda$ , and susceptible hosts are colonized with transmission rate  $\omega$  when they enter an infected environment. We assume a constant renewal rate  $\Gamma$  of microbe-free environments. Hosts grow logistically with equal intrinsic growth rate for colonized and microbe-free hosts ( $g$ ), and there is no vertical transmission of microbes, i.e. all hosts are initially microbe-free. All habitats also decay at constant rates -  $m$  for the hosts and  $\delta$  for the environments - which leads to the following equations:

$$\left\{ \begin{array}{l} \frac{\partial E_S}{\partial t} = \Gamma - \lambda E_S H_I - \delta E_S \\ \frac{\partial E_I}{\partial t} = \lambda E_S H_I - \delta E_I \\ \frac{\partial H_I}{\partial t} = \omega H_S E_I - m H_I \\ \frac{\partial H_S}{\partial t} = g(H_I + H_S) \left(1 - \frac{H_I + H_S}{K}\right) - \omega H_S E_I - m H_S. \end{array} \right.$$

Figure 3A shows a numerical example for these dynamics. Analyzing the stability of the microbe-free state  $(E_S = \frac{\Gamma}{\delta}, E_I = 0, H_I = 0, H_S = K - \frac{K m}{g})$  allows us to determine when the microbes can spread and be maintained across the two habitat types. This stability is determined by the product of the transmission rates  $\lambda\omega$ , which works as analogous to  $R_0$  in our model. Below the critical value  $\lambda_c \omega_c = \frac{\delta^2}{\Gamma K (1/m - 1/g)}$ , the microbial population does not spread among the habitats. Above the critical value, the microbial population can spread (Figure 3B), in which case an equilibrium with colonized and microbe-free hosts and environments is reached (Figure 3A). That the chance of successful spread of the microbes depends on the product of the transmission rates suggests that microbes can trade off the two transmission routes, for example by increasing their ability to colonize hosts while decreasing their release rate into the environment. Figure 3C shows how a bias in favor of one or the other of the transmission rates shifts the distribution of infected habitats towards hosts or environments.

Box 1 (continued)



**Figure 3 Analysis of microbial fitness in the context of a biphasic life cycle.** **A.** Example of habitat dynamics with microbial spread: a new equilibrium with microbes present in both host ( $H_i$ ) and environment ( $E_i$ ) is reached over time. Parameter values:  $g = 2, m = 0.6, \delta = 0.5, K = 1, \Gamma = 1, \lambda = \omega = 1$ . Initial conditions chosen close to the microbe-free equilibrium (with an introduction of 5% of infected hosts). **B.** Proportion of infected habitats in equilibrium as a function of the product  $\lambda\omega$  (points: numerical solution at time  $t=100$ ; lines: analytical stationary solution). Colonization of the habitats is only observed above the critical value  $\lambda_c\omega_c$ . Other parameters as in A. **C.** Proportion of infected environments and hosts in the new equilibrium with microbes (derived analytically), in function of  $\omega$ , for fixed values of  $\lambda\omega$ . Other parameters as in A. As expected, if transmissions are biased towards the host, it increases the proportion of infected hosts while it decreases the proportion of infected environments.

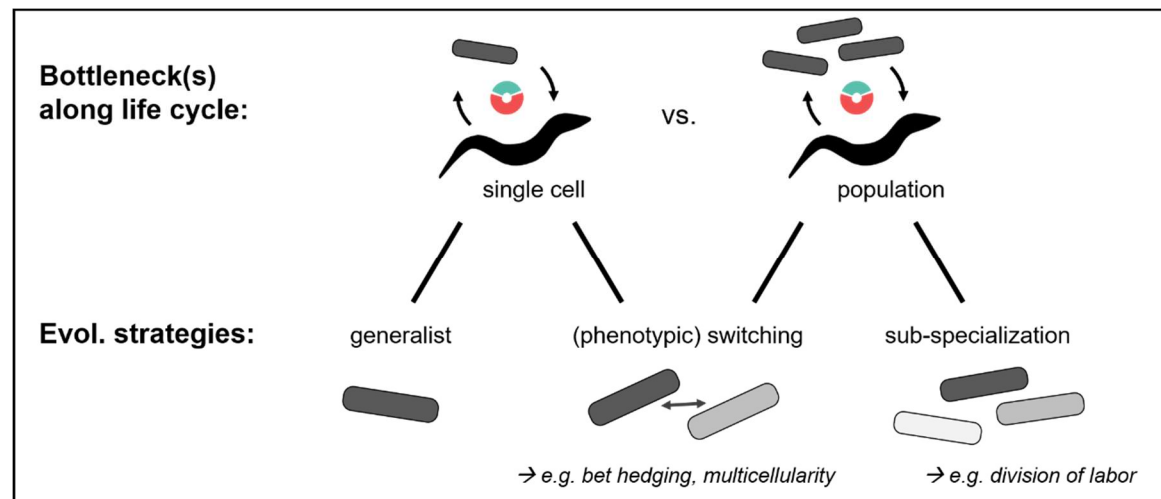
## Evolving within the biphasic life cycle

Free-living microbes can adapt to the outlined biphasic life cycle over evolutionary time scales. In addition to the encountered selective constraints, variation in population genetic characteristics are likely to determine the evolutionary dynamics. Of particular importance are bottlenecks at transmission stages, especially during host entry, because these affect the level of genetic diversity (Wein and Dagan, 2019) and thereby the process of adaptation (Figure 4). Because of the recurrent bottlenecks and thus increased levels of drift, selection across the life cycle may in the long run favor robust expression of one of the following non-exclusive strategies. Firstly, a generalist strategy, where bacteria are robustly adapted to the specifics of both host and environmental habitats. Secondly, bacteria may switch between host and environment by alternating physiological states, for example via phase variation (a genetically based switch to regulate protein expression), as observed for *Photorhabdus* and *Xenorhabdus* associating with EPNs (Cao and Goodrich-Blair, 2017). Thirdly, memory effects, possibly due to inherited protein concentrations, can further help to tune gene expression to the respective habitat, as described for *P. aeruginosa* colonizing *Galleria mellonella* moths (Kordes et al., 2019). Fourthly, if transitions between life cycle stages are unpredictable, bacteria within a population may vary between the phenotypic states stochastically, thus following a bet hedging strategy. Finally, a reliable progression of events may favor the expression of phenotypic plasticity, as for example in environmentally induced spore-formation.

A completely different strategy relies on specialization within the microbial population (Figure 4). Here, distinct sub-populations exist that are specifically adapted to different stages of the life cycle. Their abundances would be expected to fluctuate according to cycle progression. This strategy could be considered a division of labor similar to that in



biofilm formation (Gestel and Nowak, 2016). Importantly, multiple cells must survive at each stage of the cycle for the collective to be stable. A single type that gives rise to specialized offspring via development, however, can generate diversity under single cell bottlenecks. Here a genetic switch could evolve allowing a single cell to respond to changes along the life cycle, making this an early form of multicellularity (Gao et al., 2019; Hammerschmidt et al., 2014). In natural microbiota communities, these routes of diversification may contribute to the commonly observed strain-level diversity (e.g. (Ansorge et al., 2019; Ellegaard et al., 2020; Picazo et al., 2019)).



**Figure 4 Possible evolutionary strategies in response to different transmission bottlenecks during the biphasic life cycle.** When only single cells are transmitted from environment to host or vice versa, selection may favor generalists in the long run or the ability of microbes to adapt phenotypically (e.g., phenotype switching). The latter may represent a bet hedging strategy or help establishment in the new habitat through a form of a multicellular lifestyle (e.g., biofilm formation). When multiple cells, or populations, are transmitted, they can diversify via (phenotypic) switching or sub-specialize, which may lead to division of labor.

Taken together, we predict that the evolutionary path of a bacterium entering a biphasic life cycle will strongly depend on the bottlenecks experienced and the diversity of selective constraints imposed across the life cycle. Moreover, long-term persistence of the initial association will be important, as it may permit selection to act on the geometric mean of fitness across the life cycle and thus to favor a life history strategy that optimizes fitness along all four stages (in analogy to selection favoring a bet hedging strategy in fluctuating environments (Beaumont et al., 2009)).

## Perspectives

We propose that an enhanced conceptual understanding of microbial fitness across the biphasic life cycle will generally advance our knowledge of microbial life histories and the origin and function of microbiota-host symbiosis. A key strategy to help specify, test, and generalize hypotheses lies in the combination of theory and experiment. Here, advances have been made using mechanistic models of bacterial metabolism to infer interactions with the host, while classic ecological theories helped explain microbiota composition and dispersal (Miller et al., 2018; Sieber et al., 2019). What is lacking, is the consideration of the full life cycle and models of the evolutionary transitions towards host association. A first step would be models depicting the complete life cycle (Box 1) and testing what traits bacteria should optimize to succeed. Further, conceptual models may focus on conditions under which

bacteria might invest in host association. Here different aspects of bacterial biology under selection can help predict how bacteria should balance investment in host vs. environment, respectively. For example, the recent currency-exchange-inheritance model (Wein et al., 2019) proposes that, if bacteria are able to trade a currency (e.g. metabolites) with a novel host, they should invest even more into efficient or stable exchange, thereby biasing the life cycle towards host association. Alternatively, bacteria might use hosts as a vector for transmission, in which case greater chances of increased bacterial uptake should be selectively favored, which can be achieved via growth in the host, but also via an environmental reservoir (Wein et al., 2019).

Theoretical expectations should be compared with empirical observations in nature and the lab. Comparative studies of natural populations, release-and-resequence approaches, and mutagenesis screens, can help to identify traits and the underlying genetic basis required by bacteria in or from natural contexts. To study the origin and first steps in a biphasic life cycle, experimental evolution (EE), a tool previously used to study different aspects of bacterial evolution (e.g. (Ford et al., 2016; Jansen et al., 2015; Lebov et al., 2020; Martino et al., 2018a; Robinson et al., 2018; Soto et al., 2014)), can be employed. Combined with functional genetics, this can be powerful approach to identify relevant evolutionary paths and their heritable underpinnings. Although EE has provided first insights into host association, these studies usually neglected bacterial fitness in the environment or the transitory aspect of bacterial life histories. A further, important extension is to consider microbial adaptation in the presence of a diverse microbial community, shaped by distinct, often context-dependent microbe-microbe interactions (Chatzidaki-Livanis et al., 2016; Rakoff-Nahoum et al., 2016; Stephens et al., 2015). These interactions and the resulting ecological dynamics are likely to have an additional influence on individual bacterial fitness at each of the four stages of the life cycle. For example, adaptation to host association may slow down if negative interactions (i.e., competition or exploitation) predominate in the community. In this case, the selective advantages obtained while in association with a host and also the ability to transmit may become particularly important to establish long-term symbiosis. At the same time, genetic exchange within the microbial community can also fuel the adaptive potential of individual members, as observed for the evolution of virulence or antibiotic resistance in bacteria (Botelho and Schulenburg, 2020; Frazão et al., 2019). To experimentally explore these ideas, microbiota-host associations in the lab should not rely on single microbial isolates (Lebov et al., 2020; Martino et al., 2018a; Robinson et al., 2018), but rather use more complex communities (Brugiroux et al., 2016; Dirksen et al., 2020).

## Outstanding questions

- Which specific traits determine microbial fitness in each of the different stages of a biphasic life cycle in different microbiota-host systems?
- Which trait combinations are key for an integrative measure of microbial fitness across the entire life cycle?
- How important are the population genetic characteristics for microbial adaptation to host association, especially the possible bottlenecks experienced during transmission?
- Which specific life history strategies provide a selective advantage to the microbes across the biphasic life cycle? How important are bet hedging strategies or the simultaneous presence of distinct phenotypes within the microbial population?
- Which type of microbiota-host models are needed for a critical, experimental analysis of a biphasic life cycle and the evolution of symbiosis?
- How can we integrate a more complex microbial community into the theoretical and empirical analysis of biphasic life cycles and individual microbial fitness? Which community interactions are most important at which stage of the life cycle?

## **Acknowledgements**

We thank the Schulenburg and Traulsen groups for discussion. For funding we thank the German Science Foundation within the Collaborative Research Center CRC 1182 on Origin and Function of Metaorganisms, projects A4 (HS, AT, MS, FB); the Max-Planck Society (AT, MS, Fellowship to HS); and the International Max-Planck Research School for Evolutionary Biology (NO).



# 3

## **Selection gradients on microbial life history traits in the context of host association**

Florence Bansept  
Nancy Obeng  
Hinrich Schulenburg  
Arne Traulsen

## Abstract

While in the context of microbiology, the notion of fitness is often assimilated to a single fitness component, like the replication rate in a given habitat, in reality, each step of the microbial life cycle can contribute to reproductive success. This is of particular importance for microbes experiencing life cycles that involve a phase of host association, which translates to a selection pressure not only on the replication rates, but also on the phenotypic traits associated to migrating from the external environment to the host and vice-versa. Here, we investigate a simple model of a microbial population living, replicating, migrating and competing in and between two compartments: a host and an environment. We perform an elasticity analysis on the global growth rate to determine the selection gradient experienced by the microbial population, and focus on the leading direction of the selection gradient at each point of the traits space, defining an optimal strategy for the microbial population to maximize its fitness. We show that depending on the initial values of the traits, the initial distribution of the population within the compartments, the intensity of competition, and the time scales involved in the life cycle versus the time scale of adaptation (which determines the adequate probing time to measure fitness), it may be more optimal for the microbial population to act either upon its replication rates or its migration rates.

## Introduction

Fitness is a central notion in evolutionary biology, and more precisely, in the theory of natural selection. Fitness measures how well a phenotype performs in terms of reproductive success, i.e. in terms of its ability to survive and reproduce. Natural selection is then nothing else than the "survival of the fittest" phenotype, allowed by reproduction and inheritance of the phenotypic traits. However, despite the ubiquity of this notion, when it comes to the study of particular experimental or modeling systems, the question of how to measure fitness is often subtle and may vary. For example, fitness may be quantified as a proportion of habitats successfully colonized or as a replication rate measured over a limited period of time in fixed laboratory conditions. But none of these fitness components taken alone provides an inclusive view of what fitness is under natural conditions. Indeed, in nature individuals undergo complex life cycles to produce new offspring, which makes fitness a multi-dimensional notion resulting from all the life history traits characterizing that organism's life cycle. This may include, in addition to offspring production, the ability of that offspring to migrate or disperse to the appropriate environments, or the ability to find mates in the case of sexual reproduction.

The importance of taking into account life cycles complexity to characterize fitness has been highlighted in different, rather practical contexts. Historically, it has been studied in the context of human demography (Caswell, 2001). In the context of species conservation or, at the other end of the spectrum, pest management, it has proved useful to find the "Achilles heel" of a species life cycle and design efficient strategies to act upon it, in order to shape and preserve biodiversity. This idea has further been developed theoretically, within the conceptual framework of metapopulation dynamics (Andow et al., 1990; Hanski, 1998). Finally, it is also a concept central to the study of the onset of multicellularity, to understand why and how group replication can be selected for (Pichugin et al., 2017, 2019).

Microbial communities that associate with hosts – microbiotas – are of particular interest in the study of how life cycles components contribute to fitness. Nature abounds with examples of intricate life cycles, where microbes can for example use hosts as vectors between different habitats (Ciche et al., 2006; Goodrich-Blair and Clarke, 2007). Having a living host as a habitat adds one layer of complexity to the assessment of fitness, given that the presence of the microbes impacts host fitness and vice-versa. It is in fact the whole life cycle of host-associating microbes that is intertwined with that of their host. Research has often been biased towards the host perspective, and has focused on how microbes can contribute to host fitness by extending the host functional repertoire, performing digestive or immune tasks (Hrček et al., 2018; Wiles et al., 2016). An exception is epidemiology, that has also demonstrated the impact of the host fitness on the pathogen, in the form of trade-offs between transmission and within-host virulence. In this context, cross-scale modeling has proved particularly useful (Park et al., 2013; Schreiber et al., 2018). But what about commensal relationships? In this context, what are the factors that determine a microbial population fitness?

Here we propose an analysis framework to assess the gradient of selection acting upon the life history traits of a microbial population which life cycle includes host association. The gradient of selection gives the direction in the phenotypic space that evolution should follow to maximize fitness. Our general aim is to provide a tool to compare the relative importance of the different life history traits of a microbial population, starting only from the equations that describe the population dynamics experienced throughout the life cycle. We explore a simple two-compartment model that allows microbes to migrate from the environment to the host and back. We use a continuous time version of the elasticity analysis method (Caswell, 2001; Grant and Benton, 2000; Tienderen, 2000) to infer the strength of the population growth rate dependence on the traits we are considering. In the baseline version of the model, we consider unconstrained growth, before extending our framework numerically to non-linear dynamics including competition terms. We define the local leading

direction of the selection gradient as the optimal strategy for a microbial population to adapt to its life cycle, starting from the local values of the traits. We show the existence of defined regions of different optimal strategies in the phenotypic space, that are modified by the modeling assumptions such as competition, and the probing time chosen to measure fitness.

## Model

We focus on a single microbial type and ask how the growth rate of the population is affected by its life history traits. We study the population in two compartments corresponding to communicating habitats: the host and the environment. Note that in more complex models, the number of compartments could be increased, to represent for example several hosts, several anatomical compartments in a host or several extra-host environments. Let us write  $n_H$  for the number of host-associated microbes and  $n_E$  for the number of environmental ones. We define the life history traits of the microbial population as the rates at which individuals of the compartmental populations reproduce and die, compete and migrate from one compartment to another (Figure 1A).  $r_E$  and  $r_H$  are the net replication rates in the environment and within the host, respectively. They thus could encompass both offspring production and death.  $m_E$  and  $m_H$  are the migration rates from the host to the environment and from the environment to the host. We start with exponentially growing populations and later introduce competition of intensity  $k_{ij}$  felt by the microbes of compartment  $i$  due to the abundance of microbes in the compartment  $j$ . We assume that the number of microbes is large enough to be described by differential equations and assume that all rates introduced above are constant. This leads to the general equations:

$$\text{Eq. 1} \quad \begin{cases} \frac{\partial n_H}{\partial t} = r_H n_H + m_H n_E - m_E n_H - k_{HE} n_H n_E - k_{HH} n_H^2 \\ \frac{\partial n_E}{\partial t} = r_E n_E + m_E n_H - m_H n_E - k_{EH} n_E n_H - k_{EE} n_E^2 \end{cases}$$

In the following, we will first consider unconstrained growth, where there is no competition ( $k_{EE} = k_{HH} = k_{EH} = k_{HE} = 0$ ), before adding global competition ( $k_{EE} = k_{HH} = k_{EH} = k_{HE} = k$ ), competition limited to one of the compartments ( $k_{EH} = k_{HE} = 0$  and  $k_{EE}$  or  $k_{HH} \neq 0$ ), and finally, equal competition in each of the compartments ( $k_{EH} = k_{HE} = 0$  and  $k_{EE} = k_{HH} = k$ ). While in nature it is likely that none of the  $k_{ij}$  vanishes and that a wide range of values are possible, the study of these limit cases gives a powerful insight into what is to be expected in a wide range of situations.

## Results

### 1. Baseline model: no competition ( $k_{EE} = k_{HH} = k_{EH} = k_{HE} = 0$ )

We start by neglecting competition and consider unconstrained growth in each of the two compartments. In this case, the equations describing our model become linear and can be rewritten in matrix form as follows:

$$\text{Eq. 2} \quad \begin{pmatrix} \frac{\partial n_H}{\partial t} \\ \frac{\partial n_E}{\partial t} \end{pmatrix} = \underbrace{\begin{pmatrix} r_H - m_E & m_H \\ m_E & r_E - m_H \end{pmatrix}}_{\text{projection matrix}} \begin{pmatrix} n_H \\ n_E \end{pmatrix}$$



The dominant eigenvalue  $\lambda$  of the above defined projection matrix gives the asymptotic growth rate of the whole population, thus representing a good proxy for fitness. The dominant right eigenvector represents the stable distribution in the two compartments of the exponentially growing population.  $\lambda$  can be evaluated at each point of the trait space defined by the ranges of possible values that could be taken by the life history traits  $\{r_E, r_H, m_E, m_H\}$ . The dependence of  $\lambda$  on these traits tells us at which points of the trait space fitness is maximized (locally or globally).

From the projection matrix, we derive the dominant eigenvalue, which is

$$\lambda = \frac{1}{2} \left( \sqrt{(-r_E - r_H + m_E + m_H)^2 - 4(r_E r_H - r_E m_E - r_H m_H)} + r_E + r_H - m_E - m_H \right)$$

Note that if microbes replicate at the same rate in the host and in the environment, i.e. if  $r_E = r_H = r$ ,  $\lambda$  simplifies to  $r$ , regardless of the migration rates  $m_H$  and  $m_E$ . When there is an asymmetry between the two replication rates however, which is very likely to be the case in nature, we see that also migration rates affect the population growth rate. In the following sections, we study this effect as compared with the effect of replication rates, by arbitrarily setting  $r_H < r_E$ . In biological terms, this corresponds to the situation where the microbial population is initially more adapted to the environment than to the host. But mathematically, in this model host and environment are symmetrical, i.e. they only differ by the rates defined above. Thus, the chosen direction of this inequality does not carry any strong meaning, and there is no loss of generality in making this choice.

### 1.1. Symmetric migration rates

Let us first study the case where the migration rates from and towards the environment are equal, i.e.  $m_E = m_H = m > 0$ . Let us denote  $\rho = \frac{r_H}{r_E} < 1$  the ratio of the replication rates. Then, setting  $r_E = 1$  to scale time (and thus, measuring all further rates in units of the replication rate of the microbe in the environment),  $\lambda$  rewrites as

$$\lambda_{sym} = \frac{1}{2} \left( 1 + \rho - 2m + \sqrt{(1 - \rho)^2 + 4m^2} \right)$$

For any fixed positive value of  $m$ ,  $\lambda_{sym}$  is a strictly increasing function of  $\rho$ , ranging from  $\frac{1}{2}(1 - 2m + \sqrt{1 + 4m^2})$  (where  $\rho = 0$ , i.e. microbes do not replicate inside the host) to 1 (where  $\rho = 1$ , i.e. host and environment are identical). For any fixed positive value of  $\rho$ ,  $\lambda_{sym}$  is a strictly decreasing function of  $m$ , ranging from 1 ( $m = 0$ , no transfer) to  $\frac{1}{2}(1 + \rho)$  ( $m \rightarrow \infty$ ). Figure 1B shows the value of  $\lambda$  on the reduced trait space  $\{\rho, m\}$ . Thus, the maximum possible value for  $\lambda$  is 1 (in units of  $r_E$ ). This value is achieved either by increasing the ratio of replication rates between host and environment, so that both microbial populations grow nearly at the same rate (strategy 1), or by reducing migration between host and environment (strategy 2). This second strategy allows microbes to spend a longer time in the environment on average. The distribution of the time between two emigration events is exponential of mean  $1/m_H = 1/m$ . Note however, that this strategy is limited, since setting  $m$  to zero decouples the two compartments completely, in which case the expression of  $\lambda$  loses its meaning.

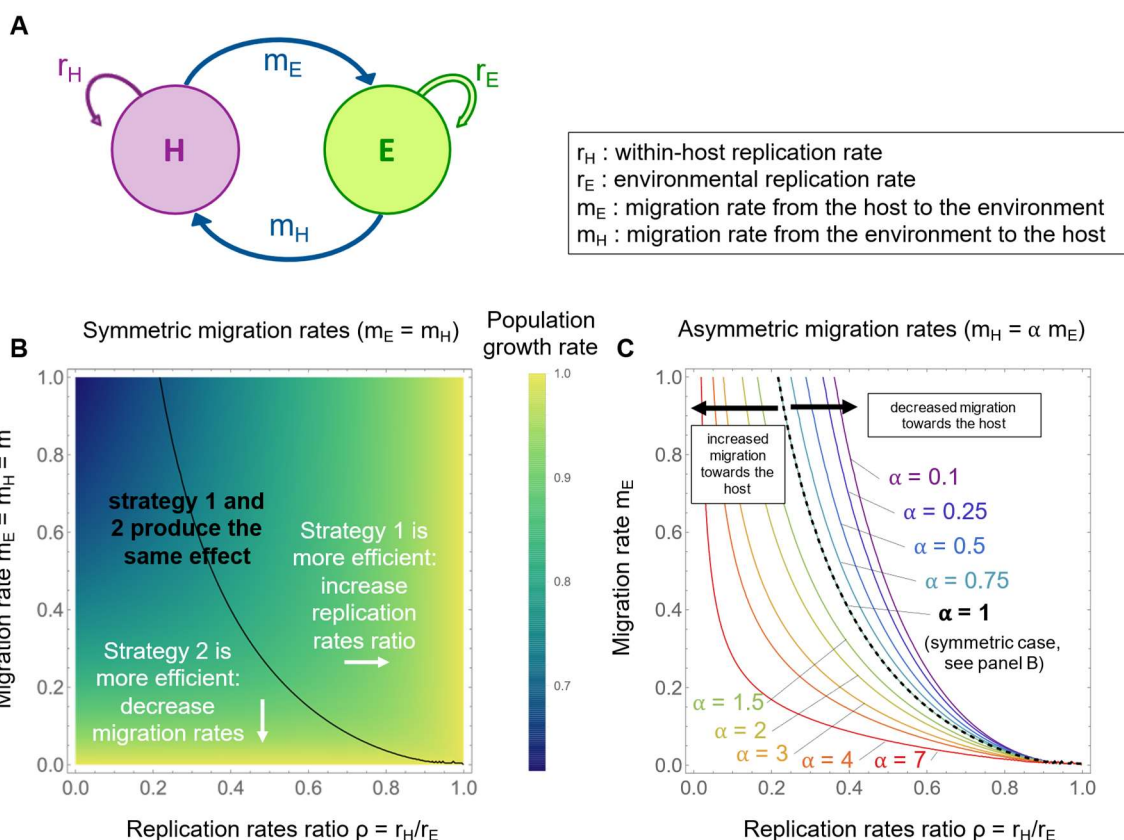
How strong is the selection on these traits? This question can be approached by inferring the strength of the population growth rate dependence on the traits we are considering. One standard approach to measure this is elasticity (Caswell, 2001; Grant and

Benton, 2000; Tienderen, 2000). One defines the elasticity of the population growth rate  $\lambda$  to the life history trait  $x_i$  as

$$\epsilon_i(\mathbf{x}) = \frac{x_i(\mathbf{x})}{\lambda(\mathbf{x})} \left. \frac{\partial \lambda}{\partial x_i} \right|_{\mathbf{x}},$$

which gives the proportional change in the value of  $\lambda$  that results from a proportional increment of the trait  $x_i$ . As elasticities sum to one (Caswell, 2001), they can also be interpreted as the relative contribution of each of the traits to the fitness, measured here as the global growth rate. This quantity is dimensionless and is a local property that can be calculated for each point  $\mathbf{x}$  of the trait space. The vector of the elasticities at point  $\mathbf{x}$  gives the direction of the selection gradient on the fitness landscape. In other words, to phenotypically adapt, the population should move in the trait space following the direction of this gradient.

If the population can invest in phenotypic adaptation only by tuning one of its life history traits at a time, then it should act upon the trait that has the largest elasticity at the current position of the population in the trait space. This reasoning allows to divide the trait space into regions of distinct optimal strategies, as shown in Figure 1B. In the limit of elevated migration rates (i.e. when the switch between the compartments is rapid, so that the population is almost experiencing the average environment), strategy 1 (increasing  $\rho$ ) becomes almost always optimal, except for small replication ratios, so when there is almost no replication in the host. In summary, migration rates are important when replication in the host is slow compared to the environment, and when migration itself is slow.



**Figure 1 Optimal strategies in the baseline model (with no competition).** (A) Schematic diagram and definition of the rates in the 2-compartments baseline model of a microbial population migrating between a host and its environment and replicating in each compartment. (B) Population growth rate  $\lambda$  (in color scale) on the trait space  $\{\rho, m\}$ . The population growth rate  $\lambda$  is maximized for small  $m$  or for large  $\rho$ . The contour line shows the line of the traits space that equalizes the absolute values of the two

(Figure 1 continued) elasticities,  $|\epsilon_m| = |\epsilon_\rho|$ , delimiting the regions of optimality of the two different strategies. Note that we take the absolute values of the elasticities, because in the baseline model the elasticity of  $\lambda$  to changes in  $m$  is always negative, while it is always positive to changes in  $\rho$ . When  $\left|\frac{\epsilon_m}{\epsilon_\rho}\right| < 1$ , the optimal strategy is to increase the replication rates ratio (strategy 1). When  $\left|\frac{\epsilon_m}{\epsilon_\rho}\right| > 1$ , the optimal strategy it is to decrease the migration rate (strategy 2). **(C)** Change in the contour line delimiting the regions of optimality of the two strategies defined in panel B with  $\alpha$ , the ratio of the migration rates. The dashed line shows  $\alpha = 1$ , i.e. the symmetric case identical to the contour shown in panel B. Increasing the migration towards the host is unfavorable for the considered population. With high values of  $\alpha$ , one needs to decrease  $m$  more to achieve similar results on  $\lambda$ , thus decreasing the area of optimality of strategy 2.

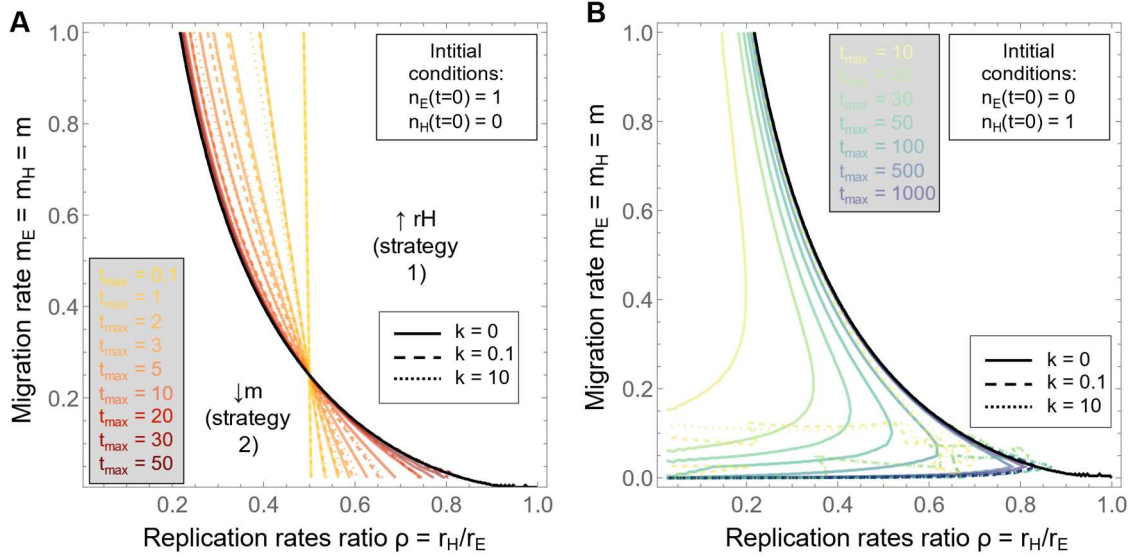
### 1.2 Asymmetric migration rates

Let us now add an asymmetry between the migration rates. Setting  $m_E = m$  and  $m_H = \alpha m_E$ , the expression of the population growth rate becomes

$$\lambda_\alpha = \frac{1}{2} \left( 1 + \rho - (1 + \alpha)m + \sqrt{(m(1 + \alpha) - 1 - \rho)^2 - 4(\rho - m(1 + \alpha\rho))} \right)$$

The resulting shape of the fitness landscape is qualitatively unchanged:  $\lambda$  is still maximized for low values of the migration rates and for high values of the replication rates ratio, and the maximum value for the population growth rate  $\lambda$  is still 1 (in units of  $r_E$ ). However, the depth of the landscape changes with  $\alpha$ : the larger  $\alpha$  (i.e. the more migration towards the host is favored compared to migration into the environment), the larger the amplitude of the landscape, and the steeper the slope to climb to reach the maxima. How does it translate in terms of optimal strategy? Figure 1C shows how the contour delimiting the optimality of the two strategies is modified. When  $\alpha$  is increased the range of optimality of strategy 2 is narrower. This is expected, as the most important factor to maximize  $\lambda$  is to stay in the environment - where replication is faster - as long as possible, and thus to reduce  $m_H$ . In the unfavorable condition of a large  $\alpha$ , one needs to reduce  $m$  more to achieve the same effect on  $m_H$ . Thus, strategy 2 to decrease migration rates becomes less efficient.

2. Model with limited overall population size ( $k_{HH} = k_{EE} = k_{EH} = k_{HE} = k$ )



**Figure 2 Optimal strategies in the model with a limited overall population size. (A)** Change in the contour line delimiting the regions of optimality of the two strategies (strategy 1: increasing the replication rates ratio; strategy 2: decreasing migration) with  $t_{max}$ , the time chosen to measure the final population size. Initially all the microbes are in the environment. With  $t_{max}$  long enough, one recovers the contour line of the baseline model calculated analytically (black line). The continuous lines show the case with  $k = 0$ , i.e. no competition, while the dashed lines show increasing values of  $k$  the competition rate. In the case of global competition, the results are barely different from the linear case. **(B)** Extension of A with all the microbes in the host initially. In this case the convergence to the linear case appears to be slightly slower, while increasing  $k$  values seems to accelerate this convergence.

In the baseline model there are no constraints on population growth. In nature, however, microbial populations do face limits to their growth. Since the equations above are linear and can only give rise to exponential growth or exponential decay, they can only describe the dynamics of a population over a limited period of time. In order to account for saturation and competition during growth, we keep a continuous time description and extend elasticity analysis numerically. This has previously been proposed to treat density dependence in discrete times (Grant, 1997; Grant and Benton, 2000).

In this section we study the case of a microbial population limited in size by global competition occurring at rate  $k$ . This situation could correspond to a microbiota living in direct contact with an external environment, like on the skin. Alternatively, what we call the "environment" in our model could represent another host compartment in direct contact with the other, like the gut lumen and the gut epithelium. In that case, microbes living in association with the host are in direct contact with those in the environment and can mutually impact each other's growth. It is of particular relevance if both microbial subpopulations rely on and are limited by the same nutrients for growth. The equations we consider are

$$\begin{cases} \frac{\partial n_H}{\partial t} = r_H n_H + m_H n_E - m_E n_H - k n_H (n_E + n_H) \\ \frac{\partial n_E}{\partial t} = r_E n_E + m_E n_H - m_H n_E - k n_E (n_E + n_H) \end{cases}$$

From the microbial abundances in the different compartments obtained by numerically solving the equations, one can build a proxy for the population growth rate. To remain coherent with the previous section, we define

$$\text{Eq. 3} \quad \Lambda(x) = \frac{1}{t_{max}} \log \left( \frac{n_E(t_{max}) + n_H(t_{max})}{n_E(0) + n_H(0)} \right)$$

i.e. the effective exponential growth rate of the whole microbial population over a defined period of time  $[0, t_{max}]$ . There are several fundamental differences between  $\Lambda$  in a non-linear system and  $\lambda$  in a linear one, the dominant eigenvalue of the projection matrix as defined in the baseline model. First, while  $\lambda$  was the asymptotic growth rate of both subpopulations in the baseline model,  $\Lambda$  only provides a measure of growth for the whole population. Second, it does not correspond to an asymptotic growth rate either: in the case of global saturation, the asymptotic growth rate for the whole population is zero, because replication stops when the carrying capacity is reached. Therefore, the choice of the probing time  $t_{ma}$  has an impact on  $\Lambda$ , which we will see in more detail below. Third, the choice of the exact form of  $\Lambda$  now implies biological assumptions on the selection pressure felt by the population: choosing the effective exponential growth rate over the whole population as we do implies that selection is acting on the whole population evenly. There may be some situations, for example experiments in which the population of one of the compartments is artificially selected for, where it would make more sense to define  $\Lambda$  as the effective exponential growth rate over just this subpopulation. One must thus adapt  $\Lambda$  to the specifics of the modeled system.

Using the Euler method, we now define the non-linear elasticity in the direction of the trait  $x_i$  as

$$\text{Eq. 4} \quad e_i = \frac{\Lambda(x_1, x_2, \dots, x_{i-1}, x_i + \delta x_i, x_{i+1}, \dots, x_N) - \Lambda(x_1, x_2, \dots, x_N)}{\delta \Lambda(x_1, x_2, \dots, x_N)}$$

with  $\delta$  the discretization parameter, and  $N$  the number of traits defining the phenotypic space  $\{x_1, \dots, x_N\}$ . Note that contrary to the previous section, the elasticities do not necessarily sum to one anymore. One advantage of using elasticities that is particularly useful for non-linear systems, is that it allows to compare traits of different dimensions, and different orders of magnitude, since it characterizes proportional changes.

For numerical solving, additional choices need to be made. First, the trait space needs to be discretized. Then, to calculate Eq. 3, one needs to choose a set of initial conditions and a final time at which to measure the population sizes, as exposed in details for the linear case in (Stott et al., 2011). Finally, we need to choose the discretization parameter  $\delta$ . In the following, we always choose  $\delta$  sufficiently small for convergence, i.e. so that it does not significantly impact the numerical values of the elasticities, and focus on the choices of the other parameters and the influence of the competition intensity  $k$ . One strategy to explore the possible sets of initial conditions is to use "stage biased vectors" (Stott et al., 2011), i.e. extreme distributions of the population. This correspond to initial conditions where microbes either exist only in the host or only in the environment.

In Figure 2, we show how the contour lines delimiting the two optimal strategies change with the final time  $t_{max}$  chosen to measure the population growth rate and with the intensity of competition  $k$ , for these two extreme cases ( $\{n_{E,0} = 0, n_{H,0} = 1\}$  and  $\{n_{E,0} = 1, n_{H,0} = 0\}$ ). In all cases, with sufficiently long  $t_{max}$ , the contours converge to the contour plot of the baseline model in the previous section. This is expected, since competition here affects all the microbes in the same way, so that the equilibrium distribution is the same as the asymptotic distribution of the baseline model (given by the dominant eigenvector). In the case where all the microbes are initially in the host (Figure 2A), the convergence to the linear case requires higher values of  $t_{max}$  than in the case where all the microbes are initially in the environment (Figure 2B). Intuitively, this corresponds to the fact that convergence to the baseline distribution requires population growth, and growth is slower if all the microbes are initially in the host compartment - where replication is slower. There is thus a time delay between these two situations, corresponding to the time it takes for migration to carry microbes into the environment - where replication is faster. Additionally, with the "all in the

host" initial condition, we note the persistence of a lower horizontal branch of the numerical contour that does not exist in the baseline model. This is an artifact introduced by the discretization of the parameter space. When  $m = 0$ , no microbe ever migrates to the environment, which therefore remains empty. Thus, when  $m = 0$ , a proportional increase of  $m$  has no effect on  $\Lambda$ , the elasticity of migration is zero, and the strategy 1 of increasing the ratio of replication rates is necessarily optimal.

Finally, we observe that the intensity of competition has only a small effect on the contours when all the microbes are initially in the environment but a larger effect with the opposite initial conditions. In both cases, adding a  $k$  appears to accelerate convergence to the baseline contour. This is because during the transient dynamics, the distribution has to balance to the expected asymptotic distribution from the initial conditions. In most cases, this equilibrium distribution is a mixed state, where a part of the population lives in the environment and another in the host. To reach this balance quickly from a pure biased state necessitates immigration to and replication in the initially empty compartment, while ideally immigration to and replication in the other compartment remain slow in comparison. Competition limits the growth in the compartment that is not initially empty, and thus helps this balancing process. This effect is even stronger if the initially empty compartment happens to be the environment, where microbes replicate faster. This explains why the effect of  $k$  is stronger in this case (Figure 2B).

### 3. Model with limited population size in one of the compartments only

( $k_{EH} = k_{HE} = 0$  and  $k_{EE}$  or  $k_{HH} \neq 0$ )

In this section we consider competition happening inside one of the compartments only (i.e.  $k_{EH} = k_{HE} = 0$  and  $k_{EE}$  or  $k_{HH} \neq 0$ ). We will start by considering competition in the host only, as it seems likely, from the biological point of view, that resources may be more limited in the host than in the environment. However, we also later consider competition in the environment only. Although this situation is biologically less plausible, given the symmetry of our model definition, a simple inversion of all the H and E indices allows to also cover the situation where the host is the fast-replicating compartment, e.g. when the microbial population is better adapted to the host than to the environment, and competition is limited to the host.

In the case where competition is limited to only one of the compartments, we do not expect an equilibrium population size to be reached, nor an asymptotic growth rate like in the baseline model. What we know should happen, is that the population in the unconstrained compartment keeps growing exponentially, while the subpopulation in the compartment with competition reaches an equilibrium size. The contribution of this subpopulation to the global population thus rapidly becomes negligible and we expect the disappearance of one of the two strategies at sufficiently long times. However, and although the theoretical perspective often focuses on long term behaviors, it can be of high practical relevance to study the transient optimal strategies, as it is not always the case that experimental times correspond to long-term behaviors.

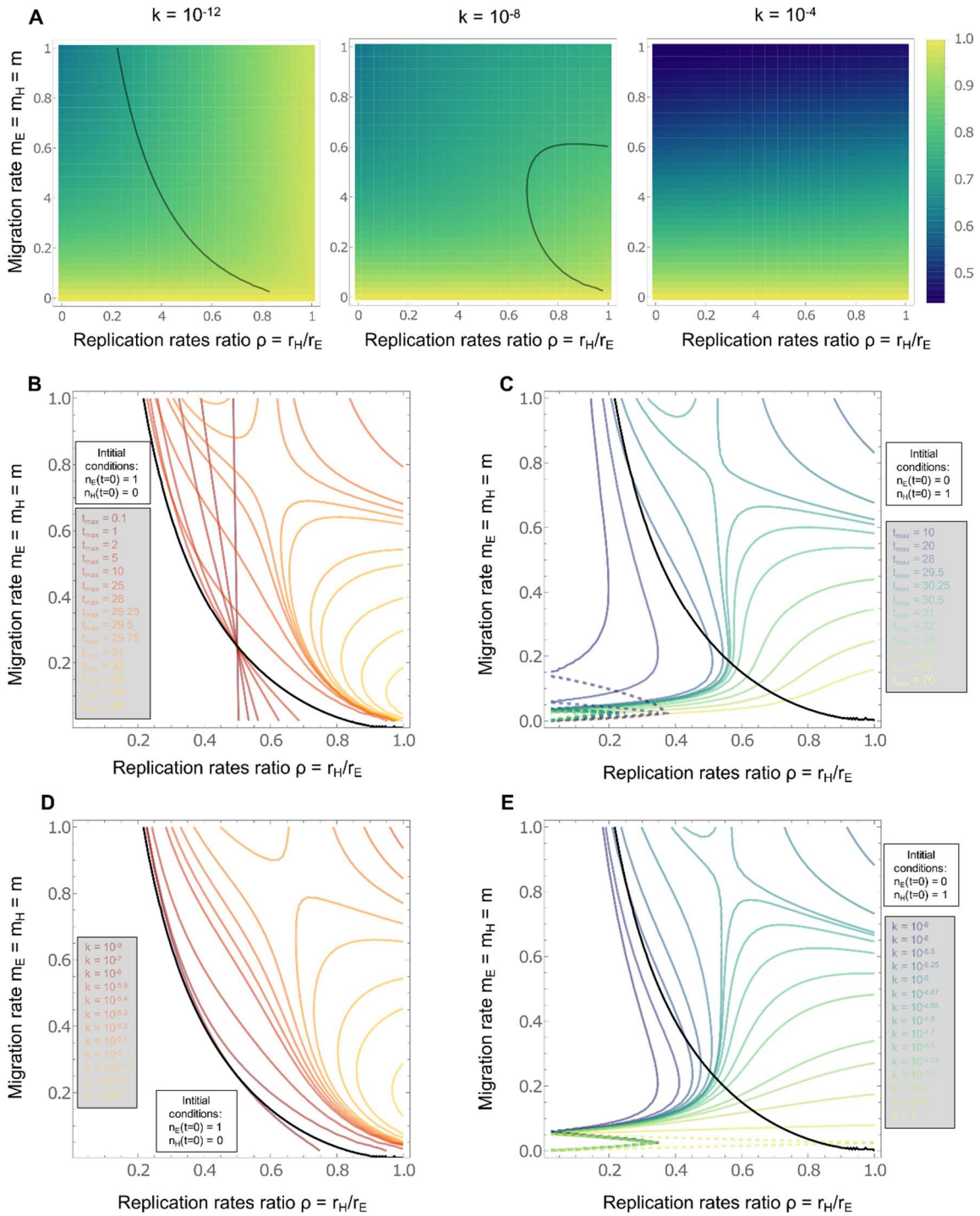
#### 3.1. Competition in the host (slow-replicating compartment)

When there is competition in the host only, then growth inside the host becomes irrelevant, so that the elasticity of parameter  $\rho$  should tend to zero, and only the migration rates should continue to matter. We thus expect the contour lines to disappear, whatever be the other parameters (initial conditions, intensity of competition) with  $t_{max}$  large enough.

Figure 3 verifies this verbal argument: as expected, for a fixed  $t_{max}$ , we recover the shape of the fitness landscape of the baseline model for small values of  $k$ . Strategy 1 of increasing the replication rates ratio, however, is disappearing with increasing values of  $k$  (Figure 3A). The values of  $\lambda$  also become smaller overall: growth is slower due to competition.

As expected, for a fixed value of  $k$ , the contours delimiting the two optimal strategies disappear with large values of  $t_{max}$  (Figure 3B-C). As previously, there is one exception: when the initial population is located in the host and  $m = 0$  there is no growth in the environmental population at all. Here the only remaining strategy has to be strategy 1, i.e. increasing  $\rho$ . Additionally, the disappearance of strategy 1 is slightly faster when initially the microbes are in the environment, the fast-replicating compartment, because this initial distribution is closer to the final one, where the host subpopulation is of negligible size. On the other hand, when the initial distribution is completely opposite to the equilibrium relative distribution (Figure 3C), we can observe the transient appearance of a third optimal strategy: increasing the migration rate. This strategy becomes optimal intermediately for small values of  $\rho$  and small values of  $m$ : in these extremely unfavorable conditions, increasing the microbial flux towards the environment becomes more important than limiting the flux of microbes leaving it.

The impact of increasing  $k$  at fixed  $t_{max}$  on the contour lines delimiting the two optimal strategies is clear in Figure 3D and E. We see that increasing  $k$  and increasing  $t_{max}$  have very similar effects: with small values of  $k$ , the baseline case is recovered, as expected. When increasing  $k$  sufficiently, the contour line disappears with strategy 1 of increasing the replication rates ratio. This is because if growth is limited in the host, increasing growth in the host can have only limited effect, which makes strategy 2 comparatively more efficient.

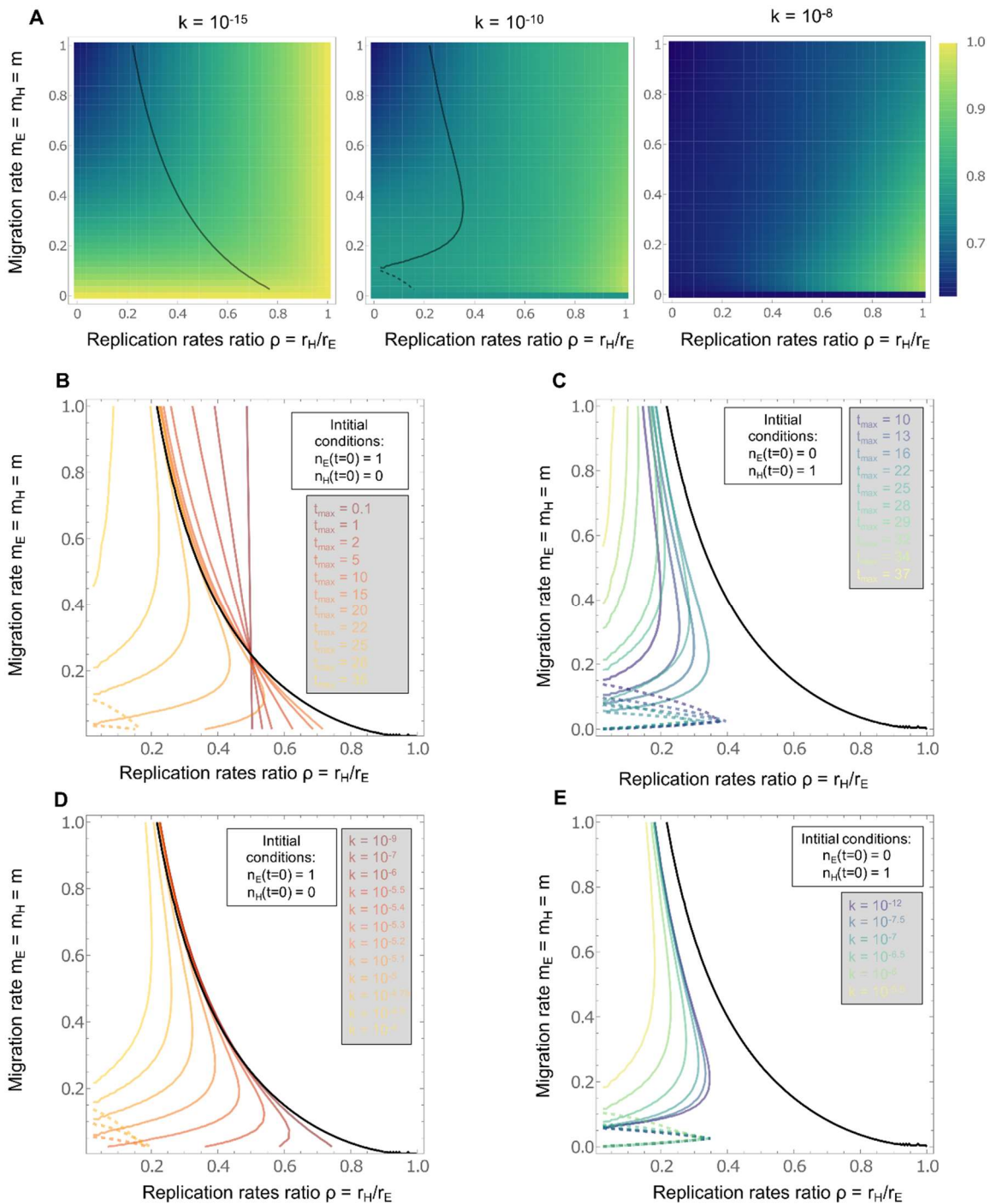


**Figure 3 Optimal strategies in the model with limited growth in the host only.** (A) Change in the fitness landscape with  $k$ , the intensity of competition within the host compartment. Other parameters values are  $t_{max} = 30$ ,  $n_E(t = 0) = 1$ ,  $n_H(t = 0) = 0$ . Black line: contour line of equal elasticities. (B-E) Change in the contour lines delimiting the regions of optimality of the strategies. Solid line: between the two strategies 1. increasing the replication rates ratio and 2. decreasing migration, and dashed line: between the strategies 1. increasing the replication rates ratio and 3. increasing migration with (B-C)  $t_{max}$  (the probing time chosen to measure the final population size) (and  $k = 10^{-8}$ ) and (D-E)  $k$  (the intensity of within-host competition) (and  $t_{max} = 20$ ), in the case of initial conditions where all the microbes are in the environment (B and D,  $n_E(t = 0) = 1$ ,  $n_H(t = 0) = 0$ ) or in the case where all the microbes are initially in the host (C and E,  $n_E(t = 0) = 0$ ,  $n_H(t = 0) = 1$ ). In the case of competition limited to the host, whatever be the value of the initial populations and the intensity of competition, at sufficiently large times, the strategy 1 (increasing the replication rates ratio) tends to disappear.



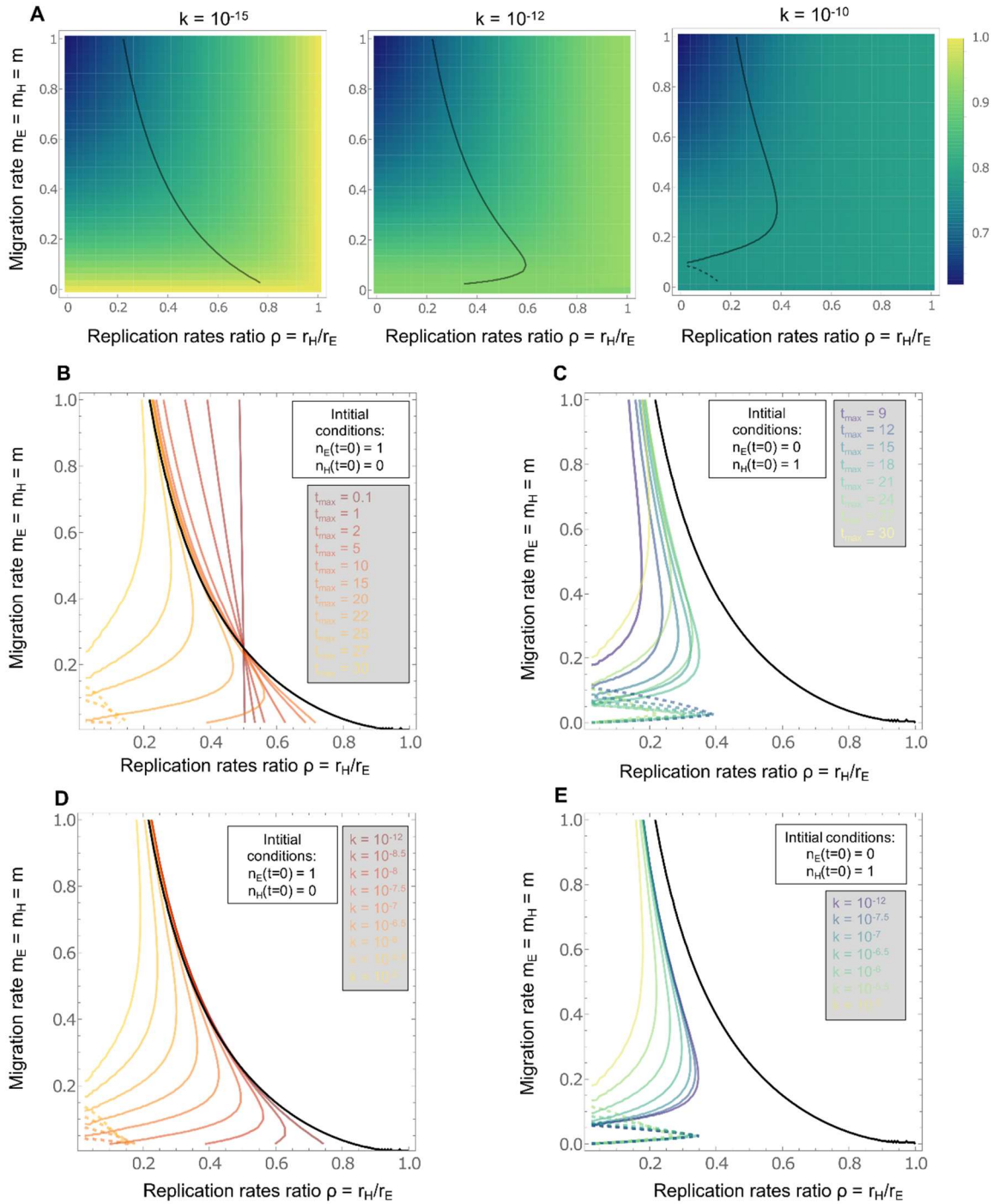
### 3.2. Competition in the environment (fast-replicating compartment)

When there is competition in the environment only, the size of the environmental population becomes negligible compared to that of the host-associated after some time. As a consequence, strategy 1 that increases the replication rate within the host becomes more important, so that we should see its area of optimality extend. Figure 4 shows that this is indeed what is happening. For a fixed  $t_{max}$ , with a small value of  $k$  we recover the shape of the fitness landscape from the baseline model with no competition, but increasing  $k$  shifts the contour line to the left until the strategy 2 of decreasing migration disappears. Like in the previous section, we also observe the transient appearance and disappearance of a third optimal strategy, increasing migration, in the most unfavorable case of a low migration and a low replication rates ratio. For a fixed value of  $k$ , with increasing  $t_{max}$  the contour line starts by getting closer to the baseline model contour, before diverging from this limit and disappearing to the left of the panel. This finally leaves strategy 1 as the only optimal strategy at sufficiently long times (Figure 4B, C). Increasing  $k$  for a fixed value of  $t_{max}$  (Figure 4D, E) has a very similar effect on the contour, except for the initial dynamics towards the baseline model: strategy 2 disappears.



**Figure 4 Optimal strategies in the model with limited growth in the environment only. (A)** Change in the fitness landscape with  $k$ , the intensity of competition within the environment. Other parameters values are  $t_{max} = 30$ ,  $n_E(t=0) = 1$ ,  $n_H(t=0) = 0$ . Black line: contour line of equal elasticities. **(B-E)**. Change in the contour lines delimiting the regions of optimality of the two strategies: increasing the replication rates ratio (strategy 1) and decreasing migration (strategy 2) with **(B-C)**  $t_{max}$  (the time chosen to measure the final population size) (and  $k = 10^{-8}$ ) and **(D-E)**  $k$  (the intensity of within-host competition) (and  $t_{max} = 20$ ), in the case of initial conditions where all the microbes are in the environment **(B and D)**,  $n_E(t=0) = 1$ ,  $n_H(t=0) = 0$  or in the case where all the microbes are initially in the host **(C and E)**,  $n_E(t=0) = 0$ ,  $n_H(t=0) = 1$ . In the case of competition limited to the environment, whatever be the value of the initial populations and the intensity of competition, at sufficiently large times, the strategy "2. decreasing migration" tends to disappear.

#### 4. Competition of equal intensity in each compartment



**Figure 5 Optimal strategies in the model with limited growth in the host and the environment. (A)** Change in the fitness landscape with  $k$ , the intensity of competition within the host and within the environment. Other parameters values are  $t_{max} = 30$ ,  $n_E(t=0) = 1$ ,  $n_H(t=0) = 0$ . Black line: contour line of equal elasticities. **(B-E)** Change in the contour lines delimiting the regions of optimality of the two strategies "1. increasing the replication rates ratio" and "2. decreasing migration" with **(B-C)**  $t_{max}$  (the time chosen to measure the final population size) (and  $k = 10^{-8}$ ) and **(D-E)**  $k$  (the intensity of within-host competition) (and  $t_{max} = 20$ ), in the case of initial conditions where all the microbes are in the environment (B and D,  $n_E(t=0) = 1$ ,  $n_H(t=0) = 0$ ) or in the case where all the microbes are initially in the host (C and E,  $n_E(t=0) = 0$ ,  $n_H(t=0) = 1$ ). In the case of competition in the host and in the environment, the effect of the competition in the environment (the fast-replicating compartment) seems to dominate, so that at sufficiently large times, the strategy 2 (decreasing migration) tends to disappear.

When there is competition of equal intensity in the host and the environment (i.e.  $k_{EH} = k_{HE} = 0$  and  $k_{EE} = k_{HH} = k$ ), we observe very similar results as in the previous section, with competition in the environment only (see Figure 5): increasing  $k$  or increasing  $t_{max}$  lead to the disappearance, at long times, of strategy 2, decreasing migration. This suggests that the effect of competition in the fast-replicating compartment has a dominating effect on the global population growth rate.

## Discussion

While in the context of microbiology, the notion of fitness is often limited to a single fitness component, like a replication rate measured in a single habitat, we emphasize that in reality, each step of the microbial life cycle can contribute to reproductive success. This is particularly true for microbes experiencing life cycles that involve a phase of host association, which translates as a selection pressure on phenotypic traits associated to migrating from external environment to the host and vice-versa. A framework to study fitness inclusive of all its components is needed. Here, we investigate a model of a microbial population living, replicating, migrating, and competing in and between two compartments: a host - assumed to be, throughout the paper, the slow-replicating compartment, and an environment. To determine the selection gradient felt by the microbial population experiencing this biphasic life cycle - with phases in the environment and phases in the host - we perform elasticity analysis. We focus on the leading direction of the selection gradient at each point of the trait space, defining an optimal strategy for the microbial population to maximize its fitness. We show that in the case of unconstrained exponential growth in both of the compartments, there are two optimal strategies: increasing the ratio of replication rates in the host versus the environment (strategy 1), and decreasing the migration rates (strategy 2) so as to maximize the time spent in the fast-replicating compartment. The first is optimal at initially high ratio of replication rates and high migration rates, while the latter is optimal at initially small migration rates and small ratio of replication rates. Next, we extend the model to a scenario where microbial growth is limited by competition. We start with global competition, a case which could describe competition for a resource homogeneously distributed across the host and the environment. Biologically, this corresponds to communities of microbes that are associated with hosts, microbiotas, but have extensive contact with the environment, as the skin or other epithelial microbiotas for example (Chen et al., 2018; Fraune and Bosch, 2007). In this case, we show that apart from a transient effect, the optimality of the strategies is conserved from the case without competition. With competition in the host only (the slow-replicating compartment), at longer probing times or at higher competition intensities, the strategy 1 (increasing the ratio of replication rates) disappears, to leave strategy 2 (decreasing the migration rates) as the sole optimal strategy. Inversely, with competition in the environment only (the fast-replicating compartment), or with competition of same intensity within the host and within the environment, the strategy 2 disappears, leaving strategy 1 as the only optimal strategy. Unsurprisingly, this suggests that competition within the fast-replicating compartment dominates the effect on the selection gradient.

While this analysis provides crucial information on the selection gradient that shapes microbial adaptation to life cycles involving host association, it does not take into account the evolvability of the traits themselves. Although the selection gradient is a good indicator of the expected evolutionary path in the phenotypic space, the underlying genotype/phenotype mapping does not always allow for this path to be taken (Houle, 1992; Lande, 1982; Orr, 2005; Tienderen, 2000). The discrete nature, the non-additivity and non-linearity of genetic information, as well as the existence of trade-offs and costs may prevent the predicted continuous change on the phenotypic trait. In addition, using elasticities is built on the assumption that adaptation generates multiplicative changes in life history traits. Although this is the most common assumption because it allows for a fair comparison of traits of

different dimensions and different orders of magnitude, different assumptions are sometimes made. For example, additive changes of the traits are assumed in sensitivity analysis (Caswell, 2001). These fundamental assumptions can sometimes result in different inferred selection gradients (S. Giaimo, personal communication).

Stepping back, we can evaluate the predictions of our model in the light of experimental work. Evolution experiments where microbial populations are sequentially passaged through a host and an environment are of particular interest here, to assess the response to selection resulting from biphasic life cycles. The key role of immigration in early-evolved adaptation in zebrafish has for example been highlighted in Robinson et al., (2018). In *Drosophila* and in *C. elegans* (Jansen et al., 2015), host selection for more subtle microbial life history traits that directly impact host fitness is shown. In the first case, there is evolution towards by-product mutualism, and in the second, which concerns an initially pathogenic population, evolution towards less virulence and an increased carrying capacity. Conceptually, our elasticity analysis approach could provide a complementary insight to the patch invasion approaches developed to analyze such evolution experiments, for example in (Hurford et al., 2010; Miller and Bohannan, 2019; Miller et al., 2018). While patch invasion relies on assessing the chances of successful invasion of an established population at equilibrium by a new mutant strain of defined traits values, elasticity analysis provides a more systematic framework to exhaustively explore the possible traits space. In addition, it is a framework that can be applied to out-of-equilibrium systems. Both frameworks rely on different proxies to assess a fitness inclusive of its different components - in one case, the frequency of patches where the microbe is present, and in the other, the microbial population growth. Both frameworks converge on the key role of migration between compartments.

In future work, we could extend our framework in different directions to capture additional characteristics of host-associating microbial populations life cycles. The first extension could concern the number of compartments: while the question of fluctuating environments had been studied before, in discrete times or in a different context (Grant and Benton, 2000; Pichugin et al., 2019), in our context it may be profitable to consider and include host population dynamics. This would notably allow to include microbial traits that affect the host fitness in our analysis. A second direction could be to include stochasticity and non-homogeneities. Indeed, our deterministic description is valid only if the size of the microbial population is sufficiently large at all times and can only describe the average selection gradient experienced by the population. Introducing stochasticity would allow to study differentiation, which may play a role in the response to complex life cycles. Differentiation, in the form of speciation, phenotypic plasticity, or bet-hedging, is indeed observed in microbial evolution experiments and in natural microbial populations (Hammerschmidt et al., 2014; Medina and Langmore, 2015; Rainey and Travisano, 1998; Xue and Leibler, 2016; Zhang and Rainey, 2010) and is thus expected during evolution of life cycles that include a host association phase. Finally, a key aspect that we have so far excluded is spatiality. Effects of spatiality on the selection pressure are known for example in a simple Petri dish system, where the existence of an optimal expansion speed for a given habitat size has been shown (Liu et al., 2019; Mattingly and Emonet, 2019). Generally, hosts are highly structured habitats with variation in nutrients and chemical and physical gradients shaping for example the gut (Donaldson et al., 2016; Schlomann, 2018; Schlomann et al., 2018). The introduction of several compartments or sub-compartments within the hosts could represent a first stepping stone in this direction.

Thus, our model provides the key ingredients to study the onset of host association, and meets the need to conceptualize fitness inclusively of all the aspects of microbial life cycles. With the development of this framework, we hope to better understand the mutual benefits that microbes and hosts can retrieve from such associations.

## **Acknowledgements**

For funding we acknowledge the German Science Foundation and the Collaborative Research Center CRC 1182 on Origin and Function of Metaorganisms, project A4 (FB, NO, HS, AT), the Max-Planck Society (Fellowship to HS); and the International Max-Planck Research School for Evolutionary Biology (NO). Further, we thank the Traulsen and Schulenburg groups for discussion and general advice.





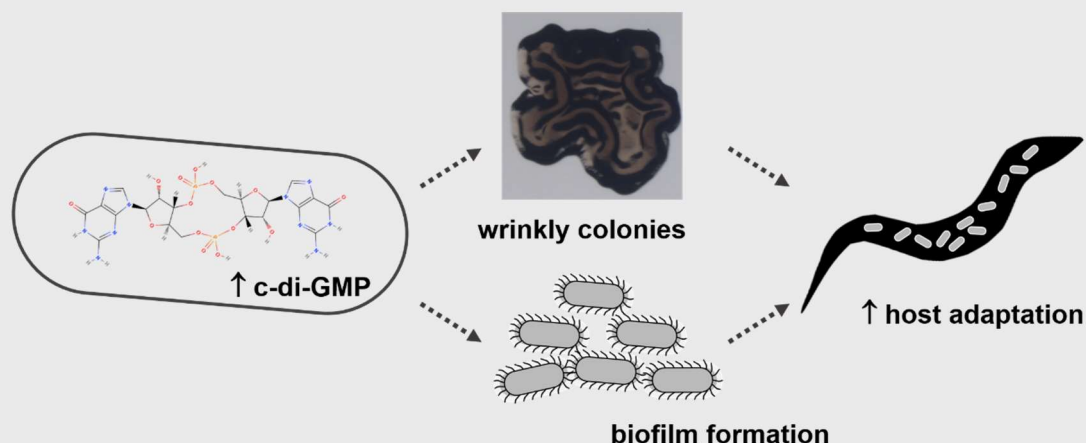


# 4

## Wrinkly formation underlies bacterial adaption to a host-associated lifestyle

Nancy Obeng  
Melinda Kemlein  
Anna Czerwinski  
Thekla Schultheiss  
Janina Fuß  
Hinrich Schulenburg

## Graphical summary



**Upregulation of the second messenger *c*-di-GMP as a strategy for biofilm-mediated host-adaptation.** Increased cellular levels of *c*-di-GMP induce biofilm formation and wrinkly colony morphology in *Pseudomonas lurida* MYb11. This appears to lead to improved host adaptation via greater association and persistence. The structural formula of *c*-di-GMP was generated using [molview.org/?cid=135440063](https://molview.org/?cid=135440063).

## Abstract

Bacteria associate with animals and plants as symbionts or microbiota communities and fundamentally shape their host's biology. It is unclear, however, which bacterial traits enable the transition from a free-living to a host-associated lifestyle. At this transition, microbes follow a biphasic life cycle and switch between hosts and the environment. This lifestyle is also widespread among microbe-host associations in nature. We here dissect the specific adaptations and molecular processes defining the biphasic symbiosis life cycle combining evolution experiments, whole genome sequencing and functional genetics of the bacterium *Pseudomonas lurida* MYb11 and its nematode host *Caenorhabditis elegans* as a study system. We find that biphasically evolved MYb11 increased colonization of and persistence of *C. elegans*, whereas absence of the host favored dispersal autonomy (colony expansion) and growth. MYb11 populations diversified, and unexpectedly, wrinkly morphotypes uniquely emerged and spread in the biphasic, host-associated populations. Wrinkly types increased biofilm formation, and persistence in worms, thereby improving transition across the life cycle. We causally link the emergence of the wrinkly type to underlying molecular changes in regulation of the second messenger cyclic-di-GMP. This study integrates the environmental context into microbe-host interactions and provides an eco-evolutionary perspective on the first steps towards symbiosis.

## Introduction

Microbes accompany hosts across the tree of life. From single bacterial species colonizing the roots of legumes as symbionts to complex communities of microbes found in humans as microbiotas, microbes can have a profound effect on their hosts' biology. They can provide essential nutrients (Consuegra et al., 2020; Flint et al., 2012; Zhu et al., 2011; Zimmermann et al., 2019), guide host development and physiology (Douglas, 2018; Sarkar et al.), and protect against abiotic stressors and pathogens (Kissoyan et al., 2019; Stecher and Hardt, 2008; Wang et al., 2020). For microbes, hosts present a habitat, that too can offer nutrients (Fischer et al., 2017), or serve as a vector for dispersal (Chase et al., 2019; Diaz and Restif, 2014; Moeller et al., 2018). Despite the well-described prevalence and importance of microbe-host associations, little is known about the drivers allowing their emergence and maintenance. Specifically, the microbial perspective in the association has been largely neglected. To understand the early steps involved in bacteria becoming symbionts, we here study bacterial traits required to initiate the association and provide the basis for its maintenance.

In the early stages of symbiosis, free-living bacteria shifted to a host-associated lifestyle. A first, loose association with a host arises via a biphasic life cycle, in which bacteria alternate between host and the free-living environment (**chapter 2**). Moreover, it appears that this biphasic life cycle can arise easily or at least be maintained efficiently, as it is still widespread among microbe-host associations. Famously, the bacterial symbiont providing bobtail squid with ventral illumination, *Vibrio fischeri*, colonizes every host anew from sea water, as do microbiota members of corals (van Oppen and Blackall, 2019) and angler fish (Baker et al., 2019). In terrestrial symbioses also, as that of *Photorhabdus* and *Xenorhabdus* bacteria colonizing entomopathogenic nematodes or the microbiota of burying beetles (Ciche et al., 2006; Shukla et al., 2018b), bacteria need to survive and/or proliferate in both the host and an alternative free-living habitat. Thus, understanding what allows bacteria to follow a biphasic life cycle will provide fundamental insights microbe-host associations.

Here, we aim to understand how bacteria transition from a free-living to a host-associated lifestyle via the biphasic symbiotic life cycle. To study the traits involved in following such a life cycle, we use the bacterium *Pseudomonas lurida* (MYb11) and its natural host, the nematode *Caenorhabditis elegans*. As a culturable isolate of the worm's microbiota (Dirksen et al., 2016, 2020), MYb11 can be specifically associated with sterile worms in the lab to follow the biphasic life cycle. This system thus presents an intermediate between well-established obligate symbioses and naïve pairs of microbes and hosts, thus allowing to select for a more stringent host association, as well as the loss of host adaptive traits. Aiming to pinpoint key traits required for MYb11 to associate with *C. elegans*, we experimentally allowed bacterial populations to evolve in a biphasic life cycle alternating between the worm as a host and an agar plate environment. As a free-living contrast, replicate populations were kept monophasically on agar only. After ten cycles mapping to ten worm generations, we analyzed phenotypes relevant across the life cycle and whole genomes of the evolved bacterial populations and isolates thereof. To understand molecular changes, we used functional genetics and established casual relationships between mutated genes and the observed phenotypic change. Together, we provide insights into the fundamental properties of bacteria allowing to stick to and interact with a host.

## Materials and methods

### *Bacterial and worm strains*

*Pseudomonas lurida* (MYb11), member of the natural microbiota of *Caenorhabditis elegans* (Dirksen et al., 2016), was used as a model bacterium in this study. In assays including worms, *Escherichia coli* OP50, the standard lab food of *C. elegans*, served as non-colonizing food bacterium that does not form biofilms (Arata et al., 2020; Portal-Celhay et al., 2012; Stiernagle, 2006). The natural *C. elegans* isolate MY316, recovered from a rotting apple in the Kiel Botanical Gardens and original host of MYb11 (Dirksen et al., 2016), was used as a host organism in the evolution experiment and all subsequent assays described in this study.

A hermaphrodite-enriched population of worms was stored at -80°C and aliquots thawed and raised on nematode growth medium (NGM) and *E. coli* OP50 lawns (0.5ml, OD<sub>600</sub>=3) at 20°C. A standard bleaching protocol was used to collect sterile eggs, and synchronize hatched L1 larvae in M9 buffer overnight (Stiernagle, 2006).

### *Evolution experiment*

Bacterial populations were serially passaged in biphasic life cycle, i.e. maintained on agar plates with exposure to *C. elegans* in regular intervals, or a monophasic life cycle, i.e. maintained on agar plates only (Fig. 1A). Specifically, MYb11 was added to nematode growth medium agar (NGM: 2.5g/l peptone from meat, 51.7mM NaCl, 1 mM MgSO<sub>4</sub>, 1mM CaCl<sub>2</sub>, 25mM KPO<sub>4</sub>, 0.0005% cholesterol) as a bacterial lawn and allowed to grow for 3.5 days. In the host-associated phase, ten *C. elegans* L4 larvae (raised until this stage on *E. coli* OP50) were added to NGM plates allowing bacteria to associate with worms for until the worms' F1 generation. In the monophasic treatment, bacteria were maintained on NGM agar only for a full cycle. At the end of every cycle, a bottleneck of 10% of the bacterial population associated with the worm hosts was imposed, i.e. 10% of host-associated bacteria were transferred to the next cycle. To impose a comparable bottleneck on free-living populations, the mean number of colony forming units (CFUs) extracted from host populations in the biphasic treatment of the previous cycle was used as a bottleneck for free-living populations. To generate a fossil record, samples of bacterial populations were collected at every bottleneck in 10% DMSO and frozen at -80°C. To allow bacteria to evolve, replicate populations of MYb11 were passaged for ten cycles, when a notably greater abundance of bacteria was found to associate with worm populations (Fig. S1).

### *Population recovery and evolution experiment re-run*

Following the evolution experiment, evolved populations from cycle ten were recovered from frozen stocks and a re-run of one cycle of the evolution experiment was done as during the evolution experiment to minimize any potential effects of freezing populations. In preparation for following assays (host-association: colonization of, persistence in and release from *C. elegans* hosts; environmental performance: growth assay, colony expansion and swarming on NGM agar; DNA isolation for whole genome sequencing), populations were collected from either worm populations or NGM agar plates, adjusted to a common optical density (as a proxy for biomass) and grown on NGM agar for two days as a common garden treatment.

### *Isolation of morphotypes*

After observing distinct colony morphologies emerging in the evolution experiment, they were isolated from cycle ten evolved populations. For this, evolved populations were recovered from frozen fossil record, serially diluted, and plated on TSA. Colonies that appeared to have a unique combination of surface texture, margin, opacity, and variation in

colony size were re-streaked on TSA (48h, 20°C) and stored at -80°C in glycerol. This was repeated for six biphasically (B1-B6) and six monophasically evolved populations (M1-M6; Table S1). To assess true breeding, colony morphotypes were thawed, and re-streaked once.

#### *Culturing of isolated morphotypes in preparation for phenotypic assays*

Ancestral MYb11 and evolved morphotypes were stored at -80°C in glycerol stocks. For phenotypic assays, 10 ml tryptic soy broth (TSB; Roth) in 50 ml Falcon tubes were inoculated with single colonies grown on tryptic soy agar (TSA; Roth) and incubated at 28°C and 150rpm overnight. Cultures were centrifuged (10 min, 4000 rpm; Eppendorf large centrifuge), washed in M9 buffer and the optical density adjusted using M9 buffer.

#### *Biofilm assay*

To quantify the morphotypes' ability in forming biofilms, a biofilm assay was conducted in microtiter plates (based on (O'Toole, 2011)). For this, bacterial overnight cultures were adjusted to an  $OD_{600} = 1$  in M9 buffer and 1:10 dilutions in TSB were pipetted into 96-well plates in a randomized layout. These were and incubated at 20°C for 48h with orbital shaking (180 rpm) in an Epoche2 BioTek plate reader. Subsequently, non-adherent cells and medium were washed off with deionized water. The remaining biofilms were stained with 0.1% crystal violet, rinsed four times with 300  $\mu$ l water each and dissolved in 200  $\mu$ l 30% acetic acid for one hour. Absorption of dyed biofilm solutions was measured at 550 nm.

#### *Host-association: colonization, persistence, and release assays*

The different stages during host association (Fig. 1D), namely colonization of L4 stage worms and adults, short-term persistence in L4 stage larvae, long-term persistence until adulthood in the presence of no food source, as well as release into the free-living environment were assayed as described below.

For all assays, bacterial lawns from 125  $\mu$ l bacterial suspension at  $OD_{600}=2$  were prepared on NGM agar (6cm plates) from ancestral MYb11 and evolved populations (post common garden) or morphotypes (post overnight culture). Depending on the assay, 30-40 synchronized L1 stage *C. elegans* larvae (assays: L4 and adult colonization, all persistence and release assays) or 10-15 L4 stage larvae raised on lawn of *E. coli* OP50 (assay: early colonization), were added to the lawn. Worms could associate with bacterial populations while plates were incubated for 48h at 20°C. To quantify bacterial load per worm, worms were washed off after incubation with M9 buffer containing 0.025% Triton-100 and 25mM tetramisole, a paralyzing antihelminthic. Worms were then washed three times in buffer using a custom-made filter tip washing system(Papkou et al., 2018). To assay long-term persistence, L4 worms were then transferred to fresh and empty NGM plates and collected for bacterial sampling after an additional 24h. Across assays, washed worms were collected in M9 buffer with Triton-100 (release: 300 $\mu$ l, all other assays: 200 $\mu$ l) into sterile 2ml Eppendorf tubes and the number of worms per sample counted under the stereo microscope. Next, 100  $\mu$ l worm-free supernatant was collected as a background sample. To extract bacteria, worm, and supernatant samples (treatment control) were ruptured using bead beating with 1mm zirconia beads in the GenoGrinder. Bacterial suspensions were serially diluted in M9 buffer, samples plated on tryptic soy agar and colonies counted after one day incubation at 20°C. The CFUs/worm during colonization were calculated as the difference in CFUs between worm and supernatant samples. To quantify bacterial release, supernatant sample of 100 $\mu$ l was taken to quantify CFUs remaining in the buffer after washing (supernatant 1). After 1 h, 100  $\mu$ l worm-free supernatant containing released bacteria (supernatant 2) was collected. Subsequently CFU quantification was done as described for the other assays. The CFUs released per worm were calculated as the difference in CFUs in supernatant 2 and supernatant 1.

To highlight the evolutionary response of biphasically and monophasically evolved populations, the ratio CFU/worm or CFU/released of the evolved vs. ancestral populations is plotted. A relative value > 1 thus indicates a greater number of CFUs reached by the evolved populations, whereas a value < 1 indicates a lower number reached by evolved populations.

#### *Environmental performance: Growth assay, colony expansion and swarming*

To measure bacterial growth, bacterial populations from the common garden treatment or washed over night cultures of the isolated morphotypes were adjusted to  $OD_{600}=0.1$  and 50  $\mu$ l spotted centrally to 6cm NGM agar plates. After 24h and 3 days at 20°C respectively, bacterial lawns were washed off in 1 ml M9 buffer. A sterile spreader was used to help scrape off tightly adhering lawns, which were then homogenized by repeated pipetting in M9 buffer. Bacterial suspensions were serially diluted in M9 buffer, samples plated on tryptic soy agar and colonies counted after one day incubation at 20°C. Swarming was assayed NGM with 0.5% agar, while colony expansion was tested on NGM with 3.4% agar, analogous to the evolution experiment the morphotypes emerged in. In both cases, 0.5 $\mu$ l of cell suspension ( $OD_{600nm} = 1$ ) were spotted centrally onto 6 cm, surface-dried (30 min in clean bench at RT) agar plates. After 24h, 3 days and 7 days at 20°C, plates were photographed (iPad mini, HR modus), and maximum colony expansion measured using Fiji/ImageJ2 (Rueden et al., 2017; Schindelin et al., 2012). All morphotypes were assayed in technical triplicate.

#### *Competition assays*

To compare fitness of ancestral MYb11 and evolved morphotype isolates, short-term persistence and growth on agar were measured in competition. For this, representative morphotype isolates were selected from population 3 (MT11 – MT14) and competed against ancestral, red-fluorescently labeled MYb11 (dTomato). Short-term persistence and growth assays were performed as described above. However, in preparation, bacterial isolates were grown in overnight cultures (TSB, 28°C, 150 rpm) and mixed in even proportions based on optical density at 600nm before inoculation.

#### *Fluorescent labeling of bacteria*

*E. coli* OP50 and the evolved, wrinkly morphotype MT12 were labeled with red fluorescent dTomato. Fluorescent labels were introduced using Tn7 transposon-based chromosomal insertion as in (Wiles et al., 2018). Briefly, bacterial strains of interest were mixed with the *E. coli* SM10 containing pTNS2 (helper strain with transposase) and the *E. coli* SM10 containing pTn7xKS dTomato (donor strain with dTomato gene and kill-switch) and incubated for a triparental mating. Subsequently, modified clones were selected based on fluorescent signals and growth on gentamicin-selective plates (10 $\mu$ g/ml). Correct insertion of *dTomato* at the attTn7 site was confirmed via PCR.

#### *Microscopy*

Colonization of *E. coli* OP50::dTomato, MYb11::dTomato and MT12::dTomato in *C. elegans* MY316 was visualized using confocal laser scanning microscopy. For this, synchronized L1 larvae were raised on the respective bacterial lawns until L4 stage as described in the colonization assays yet on PFM agar inoculated with 750 $\mu$ l bacterial suspension at  $OD_{600}=10$ . Worms were collected in 5ml of M9 buffer and washed 3 times by exchanging the buffer. For imaging, worms were anaesthetized in M9 with 25mM tetramisole and 0.025% Triton (~2h) and placed on agar patches on microscopy slides directly before imaging. Bacterial colonization was visualized using a Zeiss CLSM 700 using a 20X objective. Images were false colored and scale bars added using the Leica Zen 3.2 software.

## Genome sequencing and analysis

Total DNA was isolated using a cetyl-trimethylammonium-bromid-based (CTAB) protocol (Schulenburg et al., 2001). For Illumina HiSeq400 (paired-end, 300bp) sequencing, libraries were prepared using the Nextera DNA Flex kit using manufacturer's instructions. Genomic analyses were performed on the NEC HPC-Linux-Cluster. Python scripts were written and computed using Python version 3.7.4 (<https://www.python.org/>). Read quality of individual raw and trimmed sequences (fastq.gz files) was inspected using FastQC v0.11.8 (Andrews et al., 2018) and an overview of quality across samples was synthesized using MultiQC (Ewels et al., 2016). Reads were trimmed using Trimmomatic v0.3.9 (Bolger et al., 2014) to remove adapters and low quality bases/ sequences. Paired reads were aligned to the MYb11 reference genome (RefSeq: GCF\_002966835.1) using Bowtie2 version 2.3.3 (Langmead and Salzberg, 2012). Duplicate regions were removed using MarkDuplicates in Picardtools (version 2.22.2)(Broad Institute, 2019). Conversion from .sam- to .bam-file was done using samtools v1.5(Li et al., 2009). Genome-wide coverage was calculated using bedtools v2.29.2 (Quinlan and Hall, 2010). Average genome coverage was computed using R v3.6.1 in the RStudio environment(R Core Team, 2016; RStudio Team, 2015) as follows:

$$\frac{\text{average coverage}}{\text{genome}} = \frac{\sum_i(\text{coverage depth}_i * \text{number bases}_i)}{\text{genome length}}$$

To call variants using BCFtools version 1.10.2 (Li, 2011), zlib version 1.2.11 (Gailly and Adler, 2017), bzip2 version v2-1.0.6 (Seward, 2010) and xz Utils version 5.2.4 (Collin, 2018) were installed. Variants were called using VarScan v2.3.9 (Koboldt et al., 2012) Curl.h from curl version 7.68.0 (curlProject, 2020) was used to support the installation of BCFtools. Called variants were annotated using snpEff (Cingolani et al., 2012a, 2012b). Synonymous variants and those present in the ancestral MYb11 were filtered in R and excluded from the analysis. Gene ontology was based on GO entries of genes in the Pseudomonas database (Winsor et al., 2016). When no GO annotation was present, GO of homologous genes, as defined as top blast hits (NCBI blastn) in *Pseudomonas fluorescens* SBW25, was noted. Amino acid sequence similarity of *wspE* from MYb11 to SBW25 was compared using EMBLs MUSCLE (Madeira et al., 2019) alignment using ClustalW and standard settings. InterPro (Mitchell et al., 2019) using standard settings was used to predict domains of the MYb11 *wspE*.

### Allelic exchange

To confirm the phenotypic effect of mutations identified from genome sequencing, we used the two-step allelic replacement method as previously described (protocol updated by David Rogers and Joanna Summers; Bantinaki et al., 2007; Hmelo et al., 2015; Rainey, 1999). Here we introduced the mutant alleles into the WT background and reverted mutations by introducing WT alleles in the mutant background. Briefly, we used PCR to amplify ~700 bp surrounding each mutation, cloned products into pUIsacB plasmids, and confirmed via Sanger sequencing. Assembled pUIsacB plasmids were transformed into competent *E. coli* cells and transferred to MYb11 isolates via conjugative mating on LB agar. MYb11 transconjugants were selected on LB agar containing 25 µg/mL tetracycline and 100 µg/mL nitrofurantoin. Subsequently, counterselection for cells that lost the pUIsacB plasmid and thus tetracycline resistance was performed using sucrose in TYS10 plates (Huang and Wilks, 2017). Loss of tetracycline resistance was confirmed using selective plating. Single tetracycline resistant colonies were isolated, and the region containing the mutation in selected clones was amplified and sequenced to confirm successful allelic replacement.

## Statistical analyses

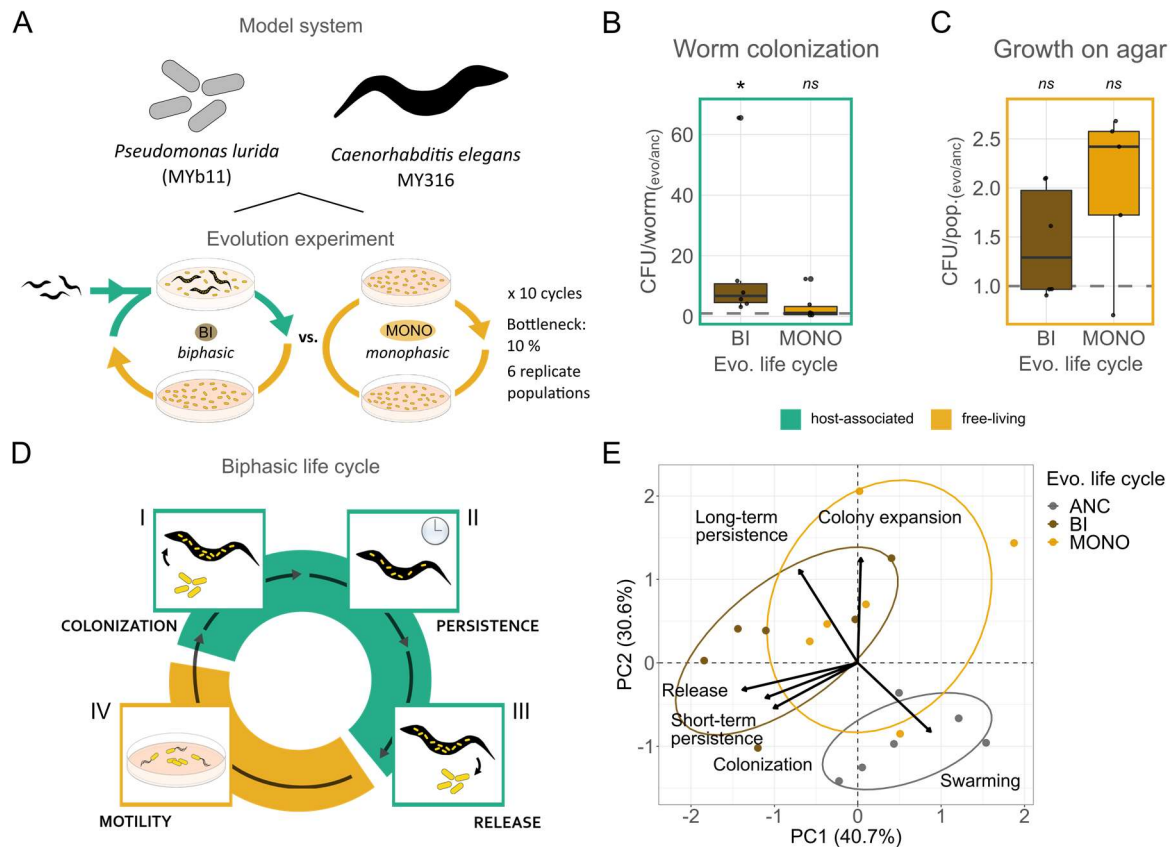
All analyses were done using R (v.3.6.1)(R Core Team, 2016) in the R Studio environment(RStudio Team, 2015). Graphs were plotted using ggplot2(Wickham) and compiled for publication in InkScape (version 0.92.2). The assumptions of parametric models (normality and homogeneity of variances) were checked by inspecting box plots and qqplots of the data and by using Shapiro-Wilk and Levene tests. When these were not met, non-parametric tests were applied. Within bacterial treatments, two-sided t-tests with false-discovery rate correction were used to test for differences of phenotypes of evolved populations vs. ancestral bacteria. Changes in morphotype proportions over time were tested for using beta-regressions in the package gamlss(Rigby and Stasinopoulos, 2005). Differences between morphotype phenotypes were detected using ANOVAs and Generalized linear models followed by Tukey post-hoc comparisons using the package lme4(Bates et al., 2015), and functional specializations of morphotypes were detected using principal component analyses using the packages ggbiplot(Vu, 2020) and missMDA(2020) and plotted using ggbiplot. Principal component analysis was performed to distinguish whether treatment groups of evolutionary life cycle or isolate morphologies showed different sets of phenotypes. In addition to plotting confidence ellipses (one standard deviation, i.e. 68%) surrounding the treatment groups, we performed permutational analysis of variances (package vegan(Oksanen et al., 2019): adonis and pairwise.adonis (Arbizu, 2020)) followed by pairwise comparisons of groups (fdr-corrected(Benjamini and Hochberg, 1995)) to test for differences in phenotype sets of ancestral and evolved groups (with 1000 permutations each).

## Results

### *Bacterial fitness in host and environment respond given the evolved life cycle*

To understand how bacteria transition to symbiosis, we experimentally evolved a biphasic life cycle alternating between host and free-living environment. Specifically, we allowed *Pseudomonas lurida* (MYb11), naturally associated with the model nematode *Caenorhabditis elegans*, to adapt to either a biphasic life history, with worms and on agar plates, or a monophasic, free-living only life history (Fig. 1A). As a member of the worm's microbiota, MYb11 provides insight into both the adaptation to defined host association in the biphasic life cycle, as well as the potential loss of traits relevant for interaction with *C. elegans* during the monophasic life cycle. As illustrated in Fig. 1A, initially clonal populations of MYb11 were exposed to *C. elegans* L4 larvae, which developed and proliferated to the F1 generation for half a cycle (3.5 days). Samples of 10% of the bacterial population associated with washed worms, were then transferred to nematode growth medium agar (NGM) and allowed to proliferate for the isochrone free-living phase (Fig. 1A). Bottlenecks imposed in the monophasic treatment had the same magnitude as those experienced by biphasic populations in the previous cycle. After passaging bacterial populations for 10 cycles, when bacterial load in worm populations had notably increased (Fig. S1), we examined the fitness consequences the evolved life cycles had on bacteria in hosts and the environment.





**Fig. 1 Evolving a biphasic life cycle with *C. elegans* as a host improves host association, while a monophasic free-living life supports bacterial growth in the environment.** (A) *Pseudomonas lurida* MYb11 as a symbiont and *C. elegans* MY316 as a host were used as a model. In an evolution experiment, initially clonal bacteria were passaged for 10 cycles in either a biphasic or monophasic life cycle (6 replicates each). This included in the former case the alternation between a host-associated phase with a free-living phase on growth agar, or in the latter case an exclusive free-living phase under otherwise identical conditions (see methods). At the end of the evolution experiment, bacterial populations were frozen and later used for phenotypic and genomic analysis. Bacterial performance across the life cycle was assessed for (B) the ability to colonize worm hosts (average number of colony forming units, CFUs, per host) and (C) the ability to proliferate on nematode growth agar for 24h hours (shown as average CFUs). To compare the direction of evolutionary change, the inferred number of CFUs in evolved populations was divided by those in ancestral populations, which were always characterized in parallel. The dashed line indicates no difference between ancestral MYb11 and evolved populations. Evolutionary shifts from the ancestral phenotype were detected using one-sample t-tests (fdr-corrected; \*\*\* ( $p < 0.001$ ), \*\* ( $p < 0.01$ ), \* ( $p < 0.05$ )). Green and yellow frames around graphs indicate whether data was collected in a host or free-living habitat. (D) The impact of experimental evolution on bacterial phenotypes along the biphasic life cycle was measured in a series of additional experiments with worm hosts of defined, synchronized developmental stage (L4 larvae and adults). The stages include I) entry and colonization of the worm, II) persistence in association with the worm (1h and 24h), III) release into the free-living environment and IV) motility in the environment. (E) Principal component analysis on phenotypes along the life cycle of ancestral, biphasic, and monophasic evolved populations. Individual data points refer to replicate populations colored according to evolved life cycle.

Biphasically evolved populations show a significantly greater association with populations of *C. elegans* than ancestral MYb11, as we observe on average more colony forming units per worm (Fig. 1B;  $t(5)=4.887$ ,  $p=0.009$ ). In contrast, monophasic populations did not shift from the ancestral level of association (Fig. 1B;  $t(5)=0.749$ ,  $p=0.488$ ). In the free-living environment, we find that bacterial population size after 24h of growth is overall comparable to ancestral levels, although the majority of monophasic populations exceeds ancestral population sizes (Fig. 1C;  $t_{BI-ANC}(5)=0.127$ ,  $p=0.064$ ;  $t_{MONO-ANC}(5)=2.398$ ,  $p=0.127$ ). At later time points, namely about a half or full cycle (3 and 7 days respectively), population densities reach

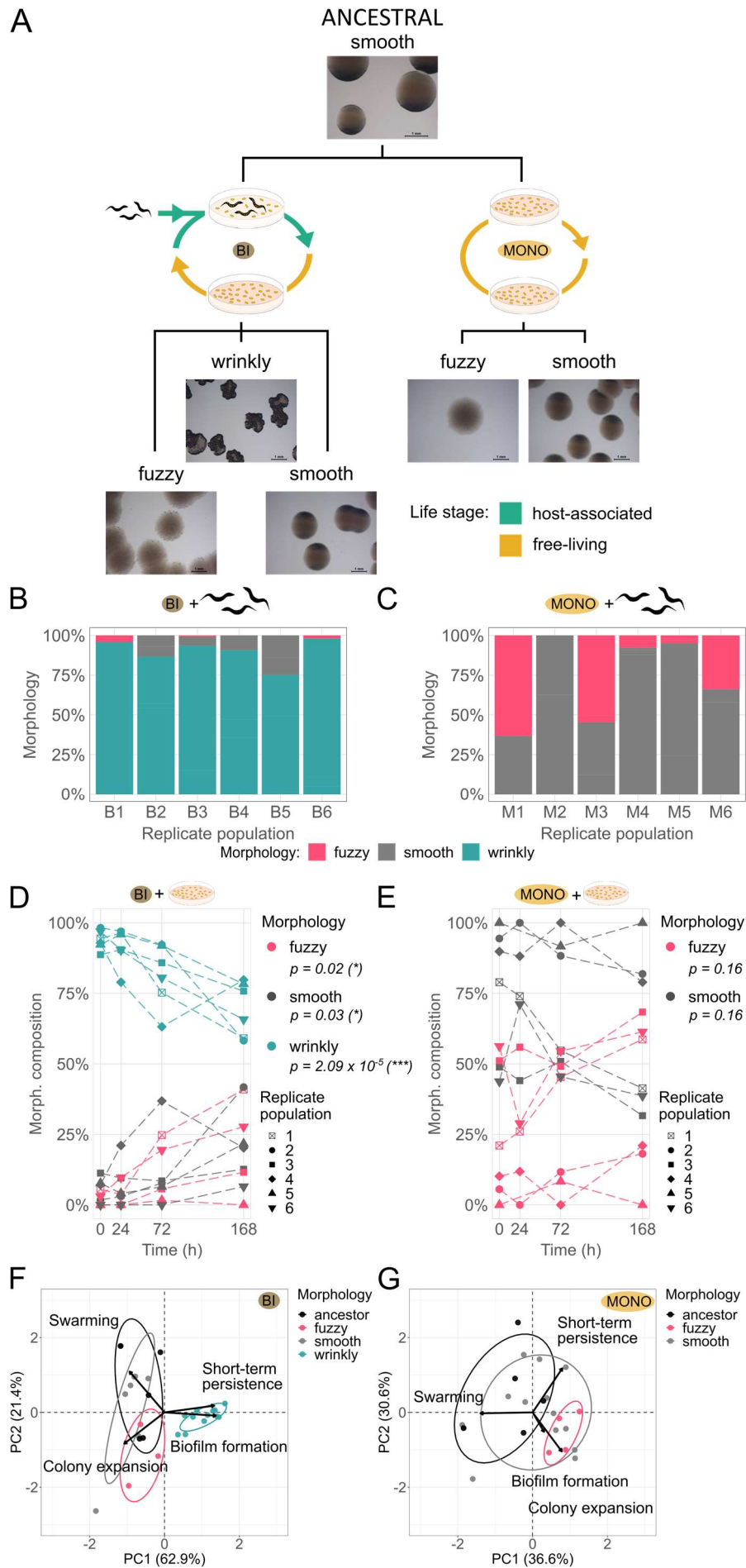
a plateau (Fig. S3D) and evolved and ancestral population sizes cannot be distinguished (Fig. S3B-C; half cycle:  $t_{\text{BI-ANC}}(5)=0.634$ ,  $p=0.554$ ;  $t_{\text{MONO-ANC}}(5)=1.329$ ,  $p=0.482$ ; full cycle:  $t_{\text{BI-ANC}}(5)=-1.314$ ,  $p=0.309$ ;  $t_{\text{MONO-ANC}}(5)=-1.132$ ,  $p=0.309$ ). Thus, bacterial fitness is generally maintained, or improves with respect to the habitat the evolving populations experienced at the time of selection.

*Along the stages of the life cycle, host association is improved via persistence in worms and plate life promotes colony expansion*

Observing a shift in fitness in worms and on plates, we investigated which stages of the life cycle were most affected by bacterial life history evolution. Here, we considered i) bacterial ability to enter, ii) colonize and iii) persist in worms, as well as iv) to be released into the free-living environment and v) disperse (Fig. 1D). In the host associated phase, we generally observe that biphasically evolved populations do not always exceed ancestral bacterial loads per worm, and monophasic populations do not always show lower abundances in worm association than ancestral populations (Fig. S4B-H). Here, we specifically consider association with L4 stage *C. elegans*, as they are both the founding development stage in the evolution experiment and present in F1 worm populations at the time of bacterial selection and can be easily identified. The colonization of worms (L4 stage) by evolved and ancestral populations of MYb11, so the number of bacteria found in worms exposed to the different bacteria since hatching, cannot be distinguished statistically (Fig. S4B;  $t_{\text{BI}}(5)=1.36$ ,  $p=0.464$ ;  $t_{\text{MONO}}(5)=0.13$ ,  $p=0.901$ ). Still, in biphasic populations we note an almost 10x greater median colonization than ancestrally (Fig. S4B). Focusing on sustained persistence of bacteria in worms in the absence of additional uptake, we find that biphasic populations most commonly maintain higher bacterial loads compared to ancestral and monophasic populations (Fig. S4C,D). While there are no statistical differences in the short-term persistence in worms within one hour (Fig. S4C;  $t_{\text{BI}}(5)=1.04$ ,  $p=0.722$ ;  $t_{\text{MONO}}(5)=-0.592$ ,  $p=0.722$ ), we find that evolved biphasic populations can maintain greater densities in worms kept on empty agar plates for 24h (Fig. S4D;  $t_{\text{BI}}(5)=6.77$ ,  $p=0.002$ ;  $t_{\text{MONO}}(5)=2.29$ ,  $p=0.071$ ). Notably, some biphasically evolved populations can maintain more than 100x greater bacterial load than the ancestral MYb11 (Fig. S4D). This suggests that bacterial populations, in particular biphasic ones, improve host association via greater persistence in worms. Finally, we quantified bacterial release from worms, thereby connecting the host and the environment. Here, we again observe great variation between replicate populations and no directional change from the ancestor across evolved populations (Fig. S4E;  $t_{\text{BI}}(5)=0.899$ ,  $p=0.722$ ;  $t_{\text{MONO}}(5)=-0.074$ ,  $p=0.944$ ). In the free-living environment, we see clear consequences for the dispersal strategies of the evolved populations. Both mono- and biphasic populations showed significantly decreased swarming ability, with the strongest decrease seen in biphasic populations (Fig. S4F;  $t_{\text{BI}}(5)=-6.07$ ,  $p=0.004$ ;  $t_{\text{MONO}}(5)=-4.3$ ,  $p=0.008$ ). On the contrary, monophasic colonies expanded more strongly on the dense agar experienced during the evolution experiment (Fig. S4G;  $t_{\text{BI}}(5)=1.79$ ,  $p=0.134$ ;  $t_{\text{MONO}}(5)=4.77$ ,  $p=0.01$ ). Here, monophasic populations used the plate surface better and showed about 1.4 times greater colony diameter than the ancestral MYb11 (Fig. S4G).

Considering general phenotypic shifts across the stages of the biphasic life cycle, a principal component analysis shows that biphasic populations have significantly diverged in phenotype from the ancestral MYb11 (Fig. S4K, Table S3, Table S2, PERMANOVA (fdr-corrected),  $F(1)=8.856$ ,  $p=0.012$ ). Similar to monophasic populations, they differ in their swarming ability, but show distinct behavior in colonization, short-term persistence in and release from worms (Fig. S4K, Table S2). Mapping to PC1, monophasic populations show the greatest colony expansion (Fig. S4K, Table S2), but cannot be distinguished from ancestral MYb11 (Fig. S4K, Table S3, Table S2, PERMANOVA (fdr-corrected),  $F(1)=2.970$ ,  $p=0.093$ ). Thus, making worm presence a key driver of phenotypic evolution. We therefore conclude that experimental evolution of a biphasic life cycle led to greater association with and

persistence in the worm, while experimental evolution of a monophasic life cycle promoted greater dispersal autonomy of bacteria.



(see left) **Fig. 2 Distinct bacterial colony morphotypes emerge during experimental evolution and enhance performance across the imposed life cycle.** (A) Distinct colony morphologies observed in ancestral and bi- and monophasically evolved populations of MYb11. Images were taken after 48 h at 20°C on tryptic soy agar. (B, C) Morphological composition of MYb11 populations from the end of the evolution experiment under either (B) a biphasic or (C) a monophasic life cycle, and subsequently passaged through *C. elegans* hosts for half a life cycle (i.e. 3.5 days as in the evolution experiment). (D, E) Morphological composition of the same populations from either (D) biphasic and (E) monophasic evolved MYb11 populations, which were grown on NGM agar for 24h, ~half a cycle (3 days), and a full cycle (7 days) of. Beta-regressions were used to test for a change in proportions over time (fdr-corrected p-values: \*\*\* ( $p < 0.001$ ), \*\* ( $p < 0.01$ ), \* ( $p < 0.05$ )). (F, G) Functional differentiation of morphologies analyzed using principal component analyses (PCA) of the phenotypes (biofilm formation, short-term persistence in worms, colony expansion and swarming). Morphotypes isolated from (F) biphasically evolved and (G) monophasically evolved populations are shown as individual data points, colored according to morphology.

#### *Populations of Pseudomonas lurida MYb11 diversify in a life cycle-specific manner*

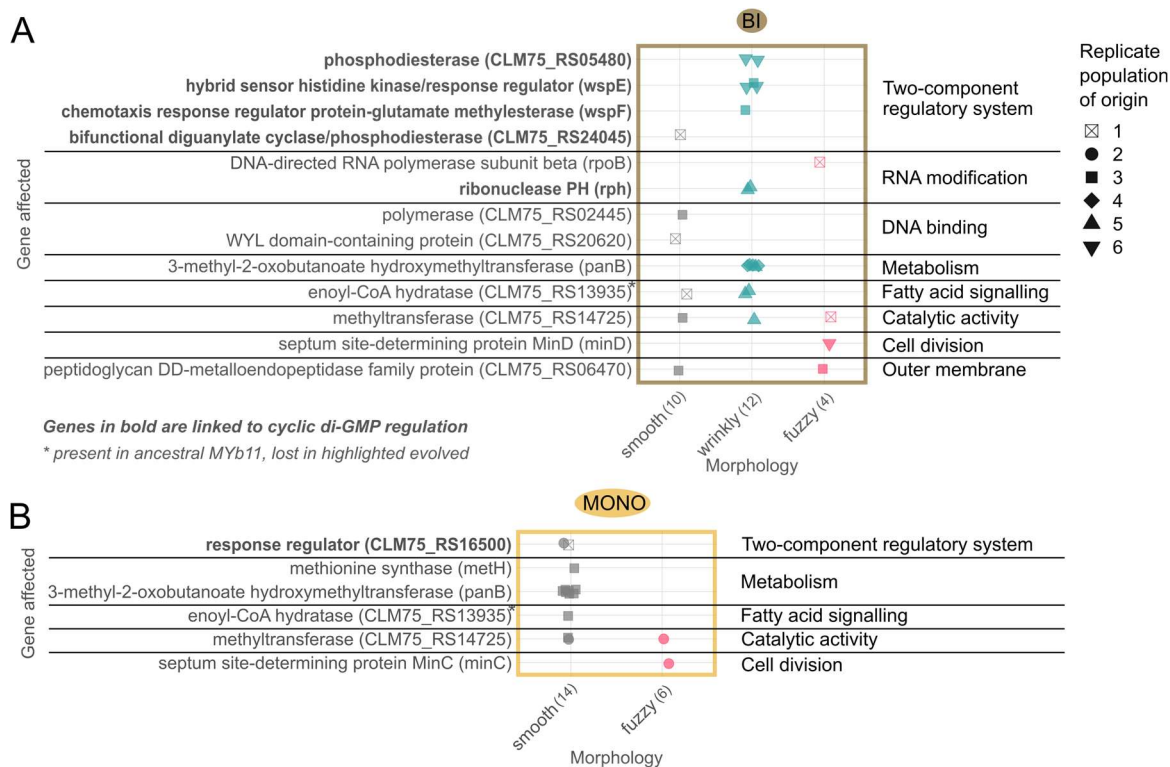
During experimental evolution of the bacterial life cycles, diverse colony morphologies emerged within passaged populations. While colonies of the ancestral MYb11 have smooth edges, are round, turbid and slightly yellow, markedly different morphotypes emerged in both bi- and monophasic populations (Fig. 2A). In the biphasic treatment, MYb11 we observed the strongest morphological change with wrinkly colonies, which appear rather dark caramel under the dissecting scope appearing across replicate populations (Fig. 2A). These wrinkly colonies only emerged in the biphasic life history, which uniquely allows bacteria to associate with worm hosts. Under both bi- and monophasic treatments, we further noted a colony type with fuzzy edges and grained colony interior (Fig. 2A), as well as a variety of colony morphologies with smooth edges and yellow appearance, much like the ancestral MYb11 (Fig. 2A). Differentiation in colony morphology was observed as early as cycle 2 of the evolution experiment and consolidated in populations collected after cycle 10 (Fig. S5).

Given the overlap of smooth and fuzzy emergence across evolutionary treatments, yet uniquely biphasic emergence of wrinkly types in biphasically evolving populations, we hypothesized that morphologies map to habitats of the biphasic life cycle, and thus checked their relative proportions in worms and on plates. Within worm populations, wrinkly morphotypes dominated all biphasically evolved populations, while the relative contributions of fuzzy and smooth types varied across replicates (Fig. 2B). Within monophasic populations, the proportions of smooth and fuzzy types in worms varied across replicate populations between roughly even proportions and smooth dominance greater than 90% of the populations (Fig. 2C). Thus, wrinkly types appear most adapted to worms, showing a consistent competitive advantage, whereas no clear difference in relative abundances of smooth and fuzzy types appears. In the free-living habitat of the NGM agar plate, we tracked the populations' morphological composition over time (Fig. 2D, E). In biphasic populations, which experience at least half of their life cycle on agar, wrinkly morphotypes consistently dominate over a period of seven days (Fig. 2D). At inoculation, which was preceded by a re-run of a biphasic life cycle and a two day common garden on NGM agar, more than 88 % of wrinkly types were detected in all replicate populations (Fig. 2D). Over time, however, the proportion of wrinkly types significantly drops (Fig. 2D, Fig. S6A,B; beta-regression:  $\beta = -0.011$ ,  $t(5) = -6.121$ ,  $p = 2.09 \times 10^{-5}$ ). Concomitantly, the proportion of both fuzzy and smooth types in the population increases significantly (Fig. 2D, Fig. S6A,B; beta-regressions:  $\beta_{\text{fuzzy}} = 0.010$ ,  $t(3) = 2.872$ ,  $p = 0.023$ ;  $\beta_{\text{smooth}} = 0.006$ ,  $t(4) = 2.370$ ,  $p = 0.032$ ). The morphological composition of monophasic populations does not vary over time (Fig. 2E). As within worms, the replicate populations appear to differ in their proportions of the two morphologies, but no directional change can be detected over time (Fig. 2E, Fig. S6C,D; beta-regression:  $\beta_{\text{smooth}} = -0.004$ ,  $t(5) = 0.003$ ,  $p = 0.163$ ). These quasi-competition experiments thus show that wrinkly types are strong competitors within worms, yet have a slight disadvantage on agar plates, manifesting after half a cycle. In contrast, fuzzy and smooth types are both

stronger competitors on agar but remain at stable proportions within monophasic populations within one cycle.

To understand whether the morphological diversification has functional consequences, we isolated morphologically distinct colony types, and screened them for phenotypes relevant in the biphasic life cycle. Notably, these isolates generally breed true to type, except for MT21, MT22 and MT43 (all pop. B5), and MT5, in which reversal from wrinkly or fuzzy to ancestral morphology was noted after 3 days of experimentation, and which are not further characterized here. In defined competition experiments with the ancestor, representative isolates of the three morphologies from population 3 recapitulate these observations (Czerwinski MSc thesis, unpublished). Under identical starting conditions, wrinkly morphotypes (MT12 and MT14) do outcompete ancestral MYb11 as well as co-evolved smooth and fuzzy isolates in worms, but not on agar plates (Czerwinski MSc thesis, unpublished).

Isolated morphotypes were further assayed for phenotypes relevant in the host-associated phase of the life cycle, namely biofilm formation and short-term persistence inside *C. elegans* MY316, as well as swarming and colony expansion on agar plates. Generally, we find significant phenotypic variation within populations, found across all phenotypes tested (Fig. 2F,G; Fig. S7-Fig. S11). Thus, in addition to morphological differentiation, we also observe functional differentiation. Across populations, principal component analyses reveal general differentiation between the morphologies, especially in the biphasically evolved isolates (Fig. 2F,G; Table S4). In a principal component analysis on biphasic isolates, the first principal component (explaining 62.9% of the variance) shows that morphotypes either show great short-term persistence in the worms and biofilm formation, or the ability to disperse (Fig. 2F). Notably, wrinkly morphotypes form a distinct cluster, with both the highest levels of biofilms and persistence, and thus the prime morphology, which can be statistically distinguished from the ancestral MYb11 (PERMANOVA (fdr-corrected),  $F(1)=28.098$ ,  $p=0.003$ ). Fuzzy types, too, appear distinct from the ancestral MYb11 (PERMANOVA (fdr-corrected),  $F(1)=5.165$ ,  $p=0.039$ ). On PC2, smooth and ancestral morphotypes have a tendency for better swarming, while fuzzy types show greater investment into colony expansion. In the monophasic treatment, evolved morphotypes largely overlap (Fig. 3G). Still, the fuzzy isolates appear on the far right of PC1, again showing the greatest colony expansion, but cannot statistically distinguished from the ancestor (PERMANOVA (fdr-corrected),  $F(1)=4.189$ ,  $p=0.130$ ). In contrast to the biphasic treatment, biofilm formation and short-term persistence are not aligned for the monophasic treatment but are almost orthogonal to each other. Thus, biphasic populations show a greater phenotypic divergence than monophasic morphologies, with wrinkly types standing out in their ability to form biofilms and to persist in worms and fuzzy types in their colony expansion. Thus, diverging morphologies indeed mirror functional differentiation with wrinkly types being most derived. Yet, smooth morphotypes resembling the ancestral MYb11 can, but do not always overlap in phenotypic traits with the ancestor.



**Fig. 3 Adaptive diversification is underpinned by specific genomic changes. (A)** Genomic variants detected in evolved (A) biphasic and (B) monophasic isolates of MYb11. Each data point represents a variant detected, color indicates morphology and shape replicate population the isolate came from. Number of isolates per morphology is shown in brackets next to morphology, and genes linking to c-di-GMP regulation are printed bold.

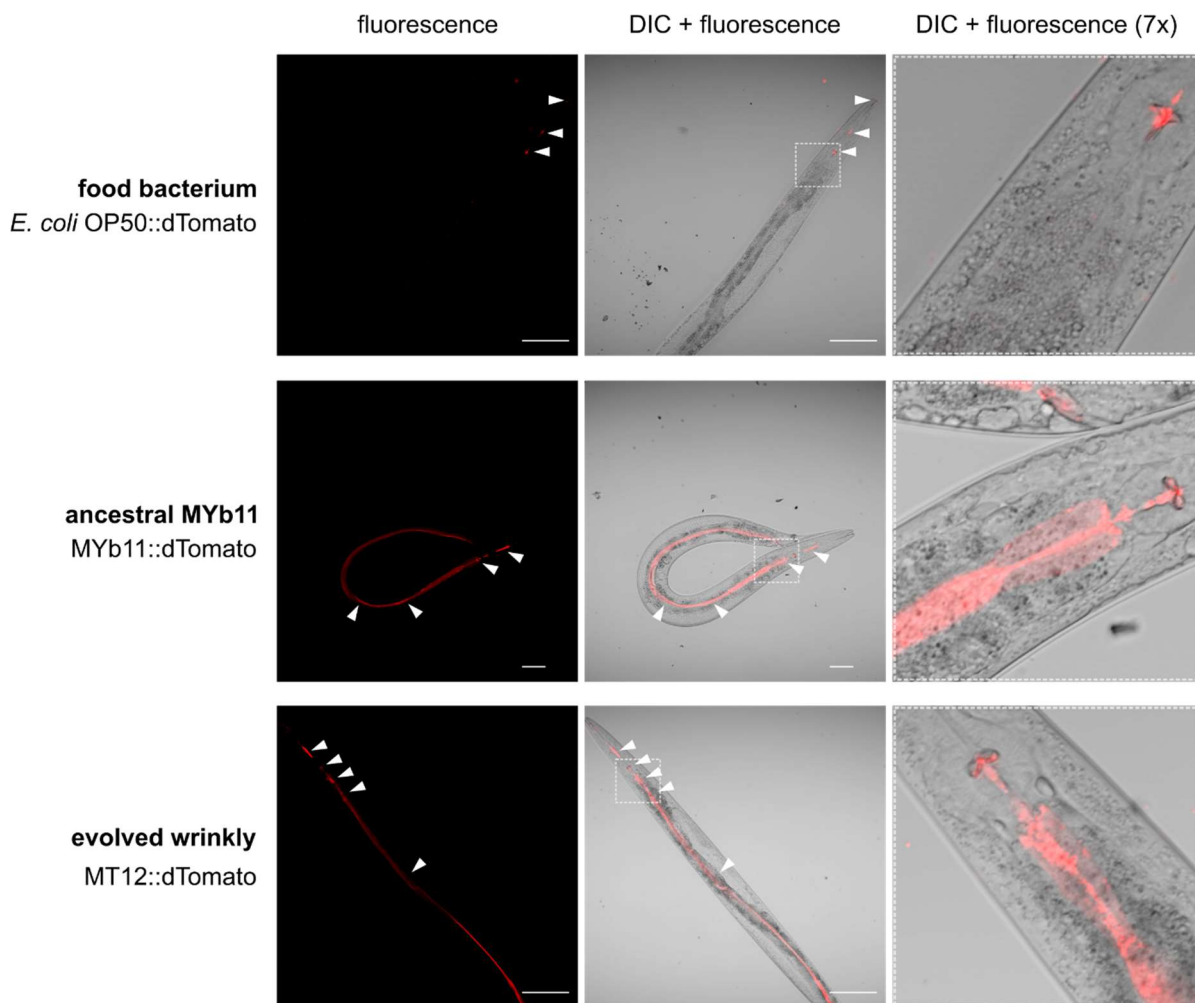
*Life cycle evolution results in genomic changes, including morphotype-specific genetic variants*

The genetic basis of life cycle adaptation was investigated using whole genome sequencing. Generally, we find novel genomic variants in all evolved populations (Fig. 3, Fig. S15), showing that there is a genetic component to bacterial life cycle evolution. These include both intragenic missense variants, frameshifts and deletions, as well as intergenic variants including gained stop codons and upstream gene variants (Fig. S12, Fig. S14). In isolates from biphasic populations, these affect various cellular functions and components including two-component regulatory systems, RNA modification, DNA binding, metabolism, fatty acid signaling, catalytic activities, cell division and the outer membrane (Fig. 3). Similarly, two-component regulatory systems, metabolism, fatty acid signaling, catalytic activity and cell division are affected in isolates from monophasic populations (Fig. 3).

Comparing the isolates with different morphologies, we detect genomic variants unique to each morphology. Most notably, only wrinkly morphotypes from biphasic populations B3 and B6 show variants in a histidine sensor kinase and a chemotaxis response regulator protein-glutamate methyltransferase (CLM75\_RS06090 and CLM75\_RS06095), homologous to *wspE* and *wspF* in *Pseudomonas fluorescens* SBW25 (Fig. 3A). Using an alignment of CLM75\_RS06090 to *wspE* from *P. fluorescens* SBW25 shows 96.82% identity, and allows assigning the frameshift mutation (p.Glu458fs) to the cheW-like domain of the protein (Fig. S16). Wrinkly types in population 6 further show changes upstream of a phosphodiesterase, while those from population 5 show missense mutations in the ribonuclease PH. In comparison to wrinkly types, neither fuzzy nor smooth isolates show great parallelism across replicate populations or evolved life cycles. Most notably, we detected mutations in the genes *minC* and *D*, both involved in cell division, in fuzzy isolates from both life cycle treatments (Fig. 3A, B). Smooth isolates, although resembling the ancestral MYb11 most closely, show diversity of mutations which belong to different gene

ontological families (Fig. 3 A,B). Further, we note that while there is great overlap in the variants detected in isolates and their respective populations of origin, there are a few differences, highlighting that sampling based on distinct morphologies did not exhaustively capture the present genetic variation (Fig. 3, Fig. S15). Taken together, we note diverse genetic differences between morphotypes evolved under the two life history regimes, with wrinkly isolates showing the greatest parallelism.

Focusing on the basis for host-adaptation in wrinkly isolates, we used functional genetics to causally link the detected mutations to the phenotypic evolution. By means of allelic exchange, we both introduced specific mutations detected using whole genome sequencing of evolved morphotypes and reverted mutants by re-introducing wild type alleles. We confirm that mutations observed in *wspF* (c.231\_236delTATCGT; in MT12) and *rph* (c.398G>A; in MT22 and MT21) indeed lead to morphological shifts from smooth to wrinkly (Czerwinski MSc thesis, unpublished). Thus, wrinkly morphotypes that evolved when experimentally evolving in the biphasic life cycle have a genomic basis, that can be recapitulated by genetic modification.



**Fig. 4 Ancestral MYb11 and wrinkly morphotype MT12 colonize the *C. elegans* gut.** Confocal laser scanning images of *C. elegans* MY316 maintained on *E. coli* OP50::dTomato, ancestral MYb11::dTomato and the evolved, wrinkly MT12::dTomato (*wspF* inframe deletion mutant), respectively. White arrows point to localization of single cells in the pharynx, pre- and post-grinder, and the intestine where applicable. Insets show a 7x zoom of the grinder and beginning of the intestine. Length of the scale bar: 100µm. Microscopic images were collected by Melinda Kemlein (MD thesis, unpublished).



Finally, we checked whether wrinkly formation has visible consequences for colonization. For this we exemplarily fluorescently labeled and visualized the wrinkly *wspF* mutant MT12 associated with worms and compared this to the ancestral MYb11 and the food bacterium *E. coli* OP50. We confirm, that the food bacterium *E. coli* OP50::dTomato does not colonize the intestine of worms as no fluorescent signal was reliably detected post grinders (Fig. 4, top panel). In contrast, in worms exposed to either the ancestral MYb11::dTomato and the wrinkly morphotype MT12::dTomato, single cells and diffuse fluorescent signal were detected in the pharynx, post-grinder, in the anterior part of the gut (defined by the first intestinal ring) and along the remainder of the intestine (Fig. 4, middle and bottom panels). Surface colonization of worms could be excluded by manually checking sections, thus limiting bacterial colonization to the worm gut. No differences in localization patterns of single cells could be found when comparing MYb11::dTomato and MT12::dTomato colonization across worms (Kemlein, unpublished MD thesis). However, the intestines of worms on MT12 appeared to be more distended than those on the ancestral MYb11 (Kemlein, unpublished MD thesis). Together, this emphasizes the role of MYb11 and evolved morphotypes as gut-associated symbionts of the worm.

## Discussion

We here provide novel insights into the early transition of bacteria towards host association, using experimental evolution of *Pseudomonas lurida* MYb11 in its host, the model nematode *C. elegans*. Passaging bacterial populations in a biphasic life cycle shared with the host, we find that persistence within worms is key to establishing and maintaining an association with the host. On the contrary, bacteria maintained free-living on agar plates improved their dispersal ability in this environment, thus rather investing in dispersal autonomy. Across evolving populations, we observe diversity on the level of colony morphology, which manifests functionally as well as genetically. Notably, wrinkly colony types unique to the biphasic life cycle with host association outcompete ancestral and co-evolved types within worms and appear best adapted to the host. We link this to improved biofilm formation, which is enabled by mutations in regulators of the intracellular second messenger c-di-GMP. Based on these insights, we now discuss (i) the evolutionary emergence of host-specialists in populations evolving in a biphasic life cycle, (ii) molecular paths to c-di-GMP-mediated biofilm formation and (iii) biofilm formation as a life history adaptation to host association.

### *Evolutionary emergence of host-specialists within biphasically evolving populations*

On the path to host association, evolving populations of *P. lurida* MYb11 included multiple functionally distinct morphotypes in the biphasic life cycle. While wrinkly types are unique to the biphasic life cycle and best adapted to worms, they did not outcompete smooth or fuzzy types in the agar environment. After ten worm generations, we thus observe multiple morphological types with heritable bases rather than the emergence and complete invasion of a single specialist strain that matches both host and environment optimally. The coexistence of distinct morphotypes during adaptation can have the following two possible explanations. On the one hand, our evolution experiment did not last long enough to produce the single genotype that can outcompete all others across all stages of the life cycle. Our results might thus be snapshot from the middle of an adaptive process, where distinct genotypes emerged, but are not yet competitively superior across the entire life cycle. On the other hand, adaptation to the biphasic life cycle as a whole may proceed on the level of the population and coexistence of different genotypes is the product. This might be considered an adaptive radiation, as wrinkly types appear to specialize on worms, while smooth and

fuzzy types have a slight advantage on agar over time. The latter explanation is consistent with the common observation of high strain-level diversity in microbiota communities, for example in *Bathymodiolus* mussel, honey bee, human and *C. elegans* hosts (Ansorge et al., 2019; Ellegaard and Engel, 2019; Schloissnig et al., 2013; Zhang et al., 2017). In the latter case, adaptation to a bacterial life history as complex as the biphasic life cycle in symbiosis cannot be achieved by a single genotype, but rather prompts diversification and niche specialization and therefore underlies a component of microbiota complexity.

#### *Molecular paths to c-di-GMP-mediated biofilm formation*

In our evolution experiments, we observed wrinkly colonies well-adapted to association with *C. elegans* hosts. Despite phenotypic parallelism in wrinkly formation, we observed a certain degree of molecular diversity in achieving it. Multiple wrinkly isolates showed mutations in *wspF*, a methylesterase encoded in the wrinkly spreader phenotype (*wsp*) operon, which regulates c-di-GMP levels via a negative feedback loop (Bantinaki et al., 2007; Lind et al., 2015; McDonald et al., 2009). Moreover, we observed wrinkled *wspE* mutants, not previously implicated in wrinkly formation (Bantinaki et al., 2007; Lind et al., 2015). We hypothesize that the observed frameshifts in *wspE* could either disrupt the downstream *wspF*, or alternatively lose their capacity to interact with WspF as a negative regulator of c-di-GMP while keeping its stimulatory effect on WspR, which in turn upregulates c-di-GMP production. The latter case is supported by the frameshift occurring in the cheW domain of WspE. This would present a novel genetic route to wrinkly formation via the *wsp* operon not previously described in *P. fluorescens* SBW25.

Next to mutations in *wspE* and *wspF*, we detected lesions in the ribonuclease PH in wrinkly (Fig. 3). There are two potential ways in which ribonuclease may affect c-di-GMP levels. The ribonucleases could bind c-di-GMP, and thus act as riboswitches, which promote or inhibit translation of specific biofilm-associated mRNAs (Jenal et al., 2017). Alternatively, ribonucleases could be involved in c-di-GMP degradation (Cohen et al., 2015; Orr et al., 2015, 2018), and their mutation would thus allow increased c-di-GMP levels then inducing biofilm genes. In cases of wrinkly formation without detectable genetic changes, more transient “memory effects” could lead to similar host adaptation. In *P. aeruginosa* passaged in *Galleria melonella* larvae, a “cellular memory” was invoked, that increased with the number of host passages and decreased with environmental passages (Kordes et al., 2019). Thus, in addition to genetic mutations, transcriptomic changes, changes in protein concentrations or other cellular asymmetries developing over time could support the host association. This should be addressed in future research. Thus, we see diverse mutations converge on increased c-di-GMP levels resulting in biofilm formation and greater association with worm hosts.

#### *Biofilm formation as an adaptation to association to the worm host*

The emergence of wrinkly morphotypes that effectively form biofilms highlights the importance of sticking to the host, which enables persistent and recurring contact to the host and is thus fundamental to symbiosis initiation. Wrinkly bacterial types sticking to each other or surfaces has been described for a variety of bacteria. In static liquid culture, *Pseudomonas fluorescens* SBW25, for example diversifies into wrinkly, smooth and fuzzy types, with wrinkly types forming a biofilm pellicle at the air liquid interface (Bantinaki et al., 2007; Rainey and Travisano, 1998). Introducing spatial structure by adding beads or increasing medium viscosity led to similar observations in *Burkholderia cenopacia* and *Pseudomonas aeruginosa* (Schick and Kassen, 2018; Traverse et al., 2013). Furthermore, wrinkly, rugose or red-dry-and-rough (radr) morphologies have been described in pathogens especially during chronic infections, including *Pseudomonas aeruginosa* (Starkey et al., 2009), *Salmonella enterica* (Anriany et al., 2001), *E. coli* (Bokranz et al., 2005), *Vibrio cholerae* O1 El Tor (Yildiz and Schoolnik, 1999) and *Mycobacterium tuberculosis* (Koch, 1884). In line with our observations here, this morphological variant is often triggered by increased intracellular levels of the

second messenger c-di-GMP. These, in turn, lead to biofilm formation and other structural changes linked to cellular attachment, including production of exopolysaccharide and proteinaceous structural components of biofilms, type IV pili, required for adhesion as well as stress resistance (reviewed in (Hengge, 2020; Jenal et al., 2017)).

Although better characterized in the context of pathogenic bacterial infections, biofilm formation has also been noted for the human intestinal microbiota (de Vos, 2015), symbionts of entomopathogenic nematodes (An and Grewal, 2011), and *P. fluorescens* and *Burkholderia* species colonizing plant and insect host (Kaltenpoth and Flórez, 2020; Liu et al., 2018; Romero-Gutiérrez et al., 2020) as well as *V. fischeri* passaged in bobtail squid light organs (Pankey et al., 2017). Further, mutations in two-component systems linking to c-di-GMP in *Helicobacter pylori* isolates in human stomachs (Ailloud et al., 2019) and *Lactobacillus plantarum* experimentally evolved in *Drosophila* (Martino et al., 2018a) hint that a biofilm strategy is more prevalent across microbe-host associations than assumed thus far.

More fundamentally, wrinkly formation may be an indicator for a shift in life history strategy towards stress-tolerance and survival in the host. C-di-GMP-mediated biofilm formation may shape bacterial fitness by providing protection or homeostasis to limit death rates (Hengge, 2020). Inside the worm host, bacteria are exposed to mechanical and chemical stress of the *C. elegans* digestive tract, where they have to surpass the grinder to enter the gut, are exposed to waves of low pH during peristalsis (Bender et al., 2013), and an arsenal of digestive enzymes (Dimov and Maduro, 2019). Such self-protection may also be viewed in light of the universal adaptive strategy theory (Fierer, 2017; Grime, 1977), in which bacteria may be categorized into stress-tolerating, competitive or ruderal (i.e. fast niche occupiers) types based on their functional repertoires. Most interestingly, MYb11 had previously been classified as competitive (Zimmermann et al., 2019). Upon experimental adaptation to the host-associated biphasic life cycle, this strain has now evolved a stress-tolerating strategy, dominated by the wrinkly types which produce biofilms and show slow colony expansion on agar. Such a stress-tolerating strategy has previously linked to greater worm host colonization for other members of the *C. elegans* microbiota (Zimmermann et al., 2019). The consequence of being able to survive and persist within the host should be especially beneficial when resources or competition limit growth in the environment, and hosts present an alternative habitat or vector for dispersal. In combination with the loss in autonomous motility of wrinkly types, host association then increases reliance on the host, and potentially collectivization of host and bacterium (Militello, 2019).

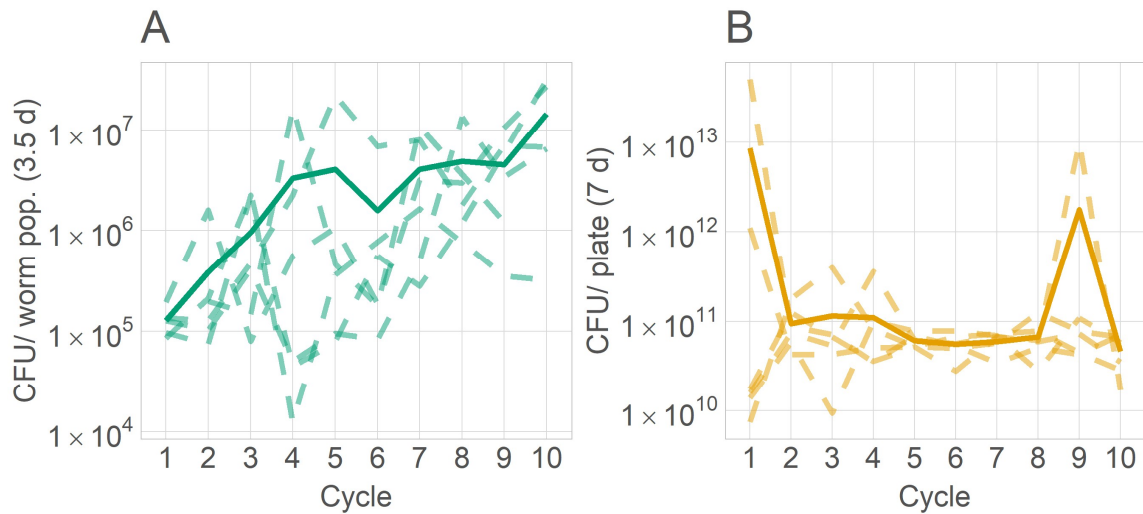
Taken together, our study shows that bacteria can improve host association via a biphasic life cycle. They do so by shifting from a motile to a sessile, stress-tolerating life history using the universal second messenger c-di-GMP. This life history transition presents a key adaptation to initiate and maintain contact with the host. A bacterial pre-adaptation to settle can therefore explain the early steps toward symbiosis and microbiota emergence.

## Acknowledgements

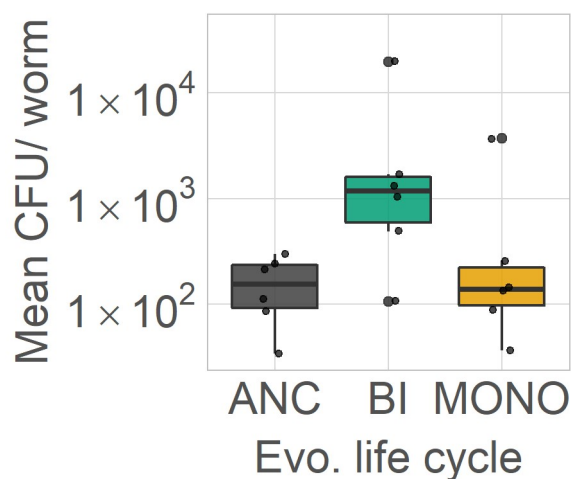
We thank Christopher Giez for preparing MYb11::dTomato, Shindhuja Joel for help during worm sampling, Paul Rainey, Joana Summers and David Rogers for support for allelic exchange mutagenesis and general discussion, the staff of the microbiome sequencing unit at the CCGA, Kiel, Kiel BiMo/LMB for access to their core facilities, and the CRC1182 and the Schulenburg group for general advice. For funding, we thank the German Science Foundation within the Collaborative Research Center CRC 1182 on Origin and Function of Metaorganisms, projects A1 (HS), A4 (HS, NO, AC, MK), the Competence Center for Genome Analysis Kiel (CCGA Kiel; HS); the Max-Planck Society (Fellowship to HS); and the International Max-Planck Research School for Evolutionary Biology (NO).

## Supplement

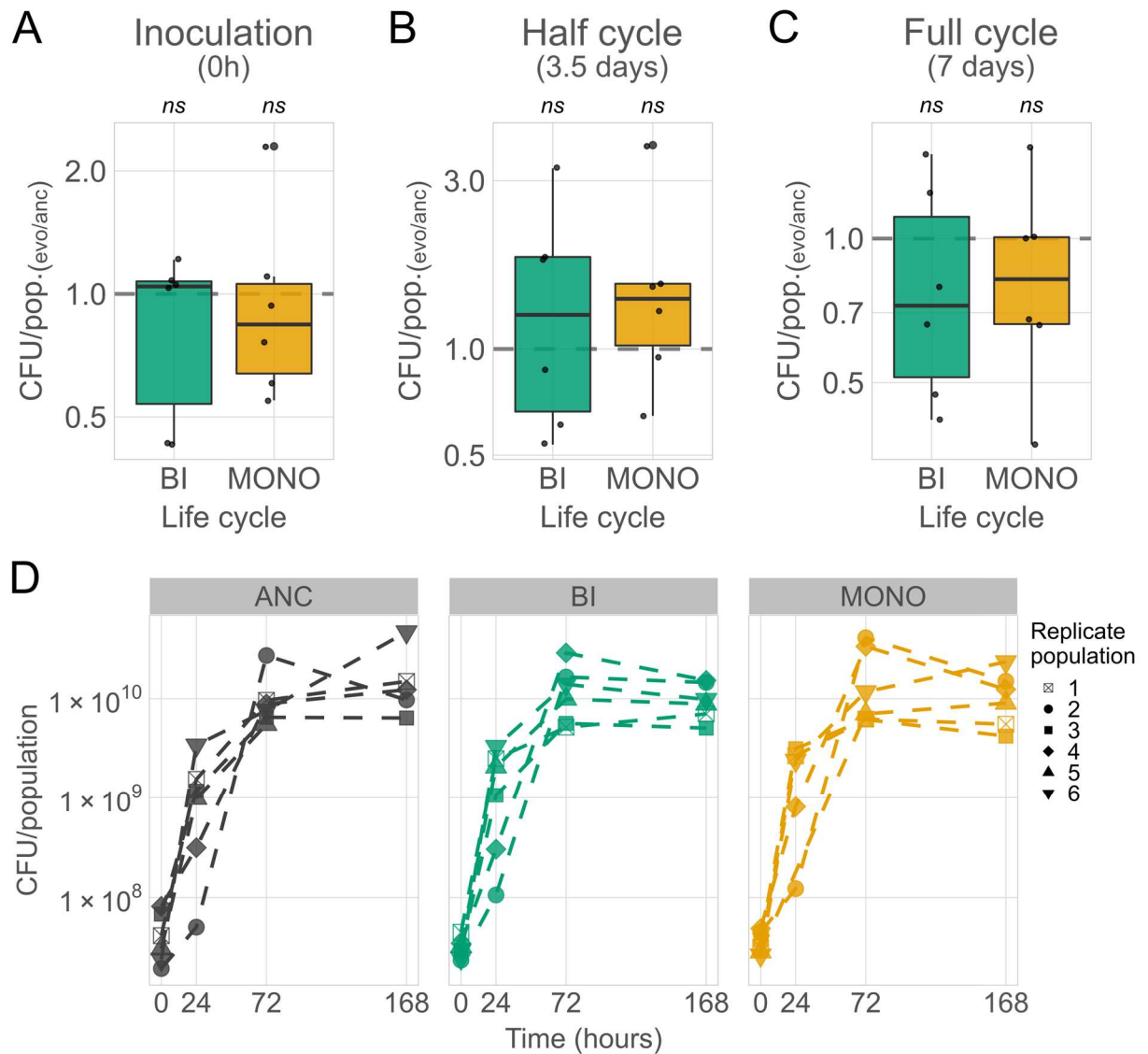
### Supplementary figures



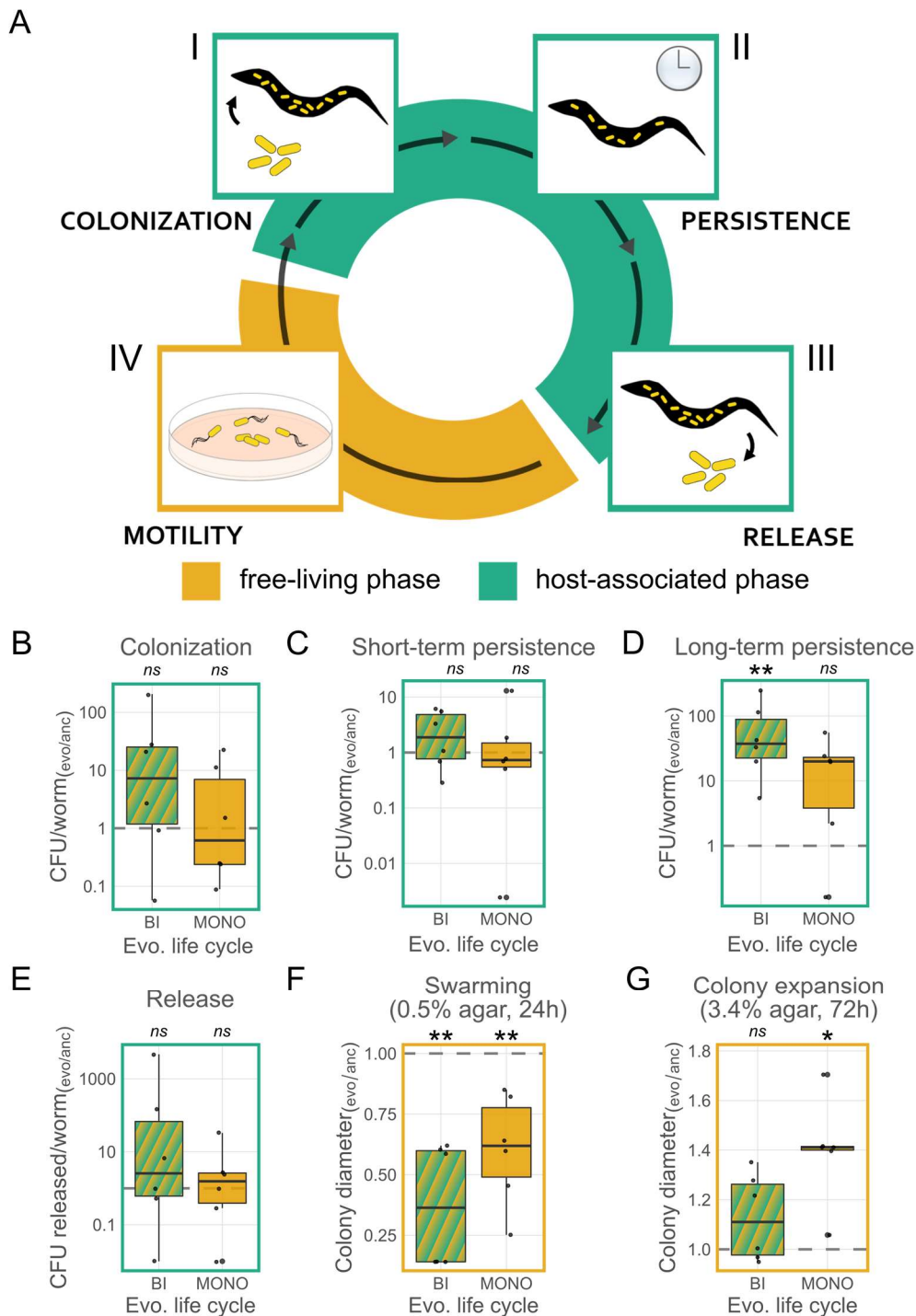
**Fig. S1 Bacterial fitness during the evolution experiment. (A)** Bacterial fitness in worms, as number of colony forming units associated with a worm population founded by 10 L4 larvae over 3.5 days. **(B)** Bacterial fitness on nematode growth agar, as CFUs collected per population after growth for one monophasic cycle of 7 days. Replicate populations are shown as separate lines, thick line shows mean and standard deviation across replicates. Data was collected during the evolution experiment, always at the bottleneck of the noted cycle.



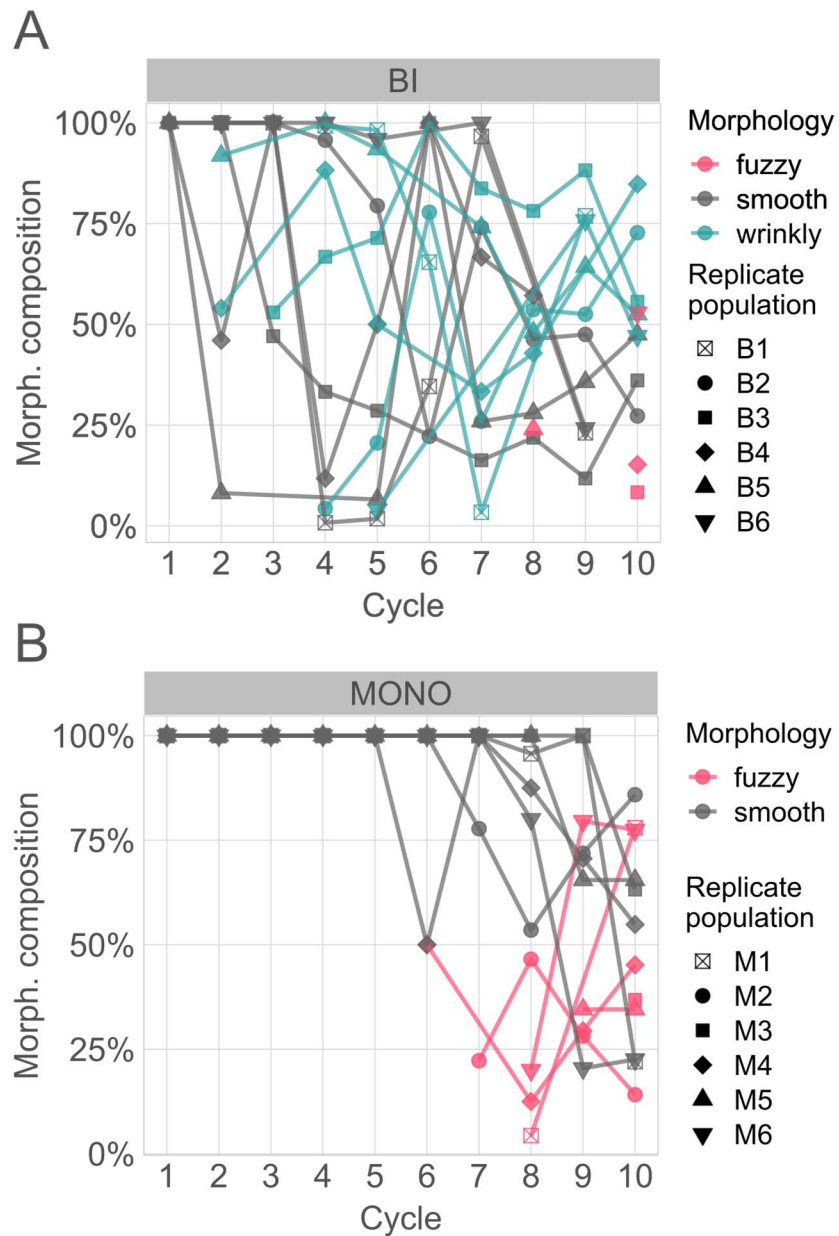
**Fig. S2 Mean number of colony forming units (CFU) per worm in a worm population.** Five L4 *C. elegans* larvae proliferated on evolved and ancestral bacterial lawns for 3.5 days (reaching F1 generation) and CFUs were extracted from the whole worm population. CFUs per population were divided by the estimated number of worms in the population. Replicate populations are shown as individual data points.



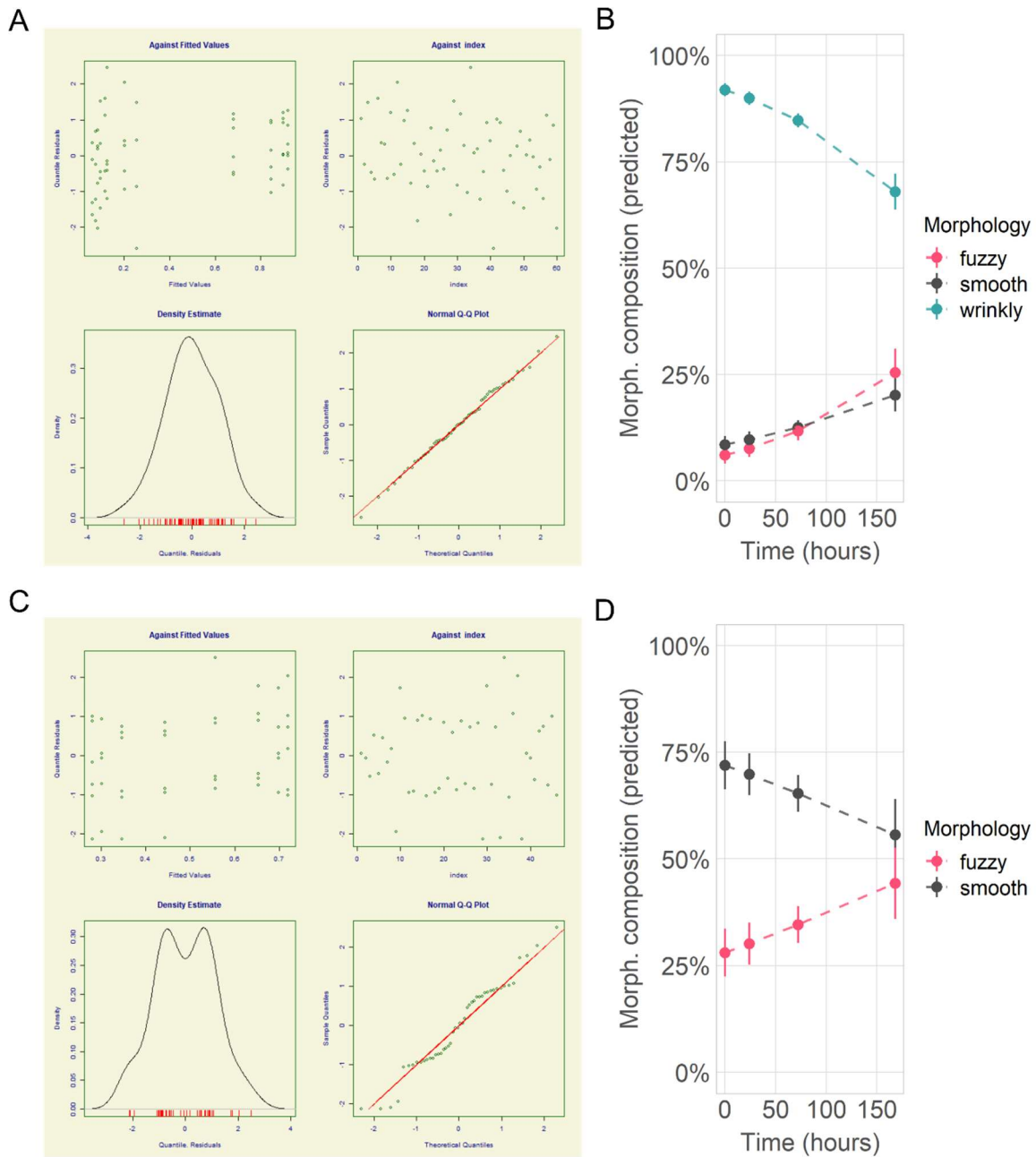
**Fig. S3 Bacterial growth on nematode growth agar. (A-C)** Evolved shift in bacterial population growth on nematode growth agar at **(A)** inoculation, **(B)** after ~half a cycle (3 days), and **(C)** after a full cycle (7 days). Quotients of evolved over ancestral MYb11 replicate populations are shown as individual data points. The dashed line indicates no difference between ancestral MYb11 and evolved populations. Evolutionary shifts from the ancestral phenotype were detected using one-sided one-sample *t*-tests (*fdr*-corrected; \*\*\* ( $p < 0.001$ ), \*\* ( $p < 0.01$ ), \* ( $p < 0.05$ )). **(D)** Absolute CFUs detected per population over time. Replicate populations are shown as individual data points in population-specific shapes.



**Fig. S4 Evolutionary life history shapes adaptation to the stages of the biphasic life cycle.** (A) For bacteria, the stages of the biphasic life cycle imposed include I) colonization of the worm, II) persistence in association with the worm, III) release into the free-living environment and IV) motility in here. Shifts in phenotypes from the ancestral MYb11 to the evolved populations associated with worms of defined developmental stage were analyzed. This includes change in (B) colonization of L4 larvae (exposed to bacteria since L1 stage), (C) short-term persistence in L4 larvae kept in M9 buffer for 1h (raised on bacteria), (C) persistence in worms kept on empty NGM agar for 24h post-exposure, and (D) release from L4 larvae into buffer within 1h (previously raised on bacteria from L1 to L4), and (E) swarming distance on 0.5% agar within 24, as well as (F) colony expansion on 3.4% agar within 72h. Quotients of evolved over ancestral MYb11 replicate populations are shown as individual data points. The dashed line indicates no difference between ancestral MYb11 and evolved populations. Evolutionary shifts from the ancestral phenotype were detected using *t*-tests (*fdr*-corrected; \*\*\* ( $p < 0.001$ ), \*\* ( $p < 0.01$ ), \* ( $p < 0.05$ )). Green and yellow frames around graphs indicate whether data was collected in a host or free-living habitat.

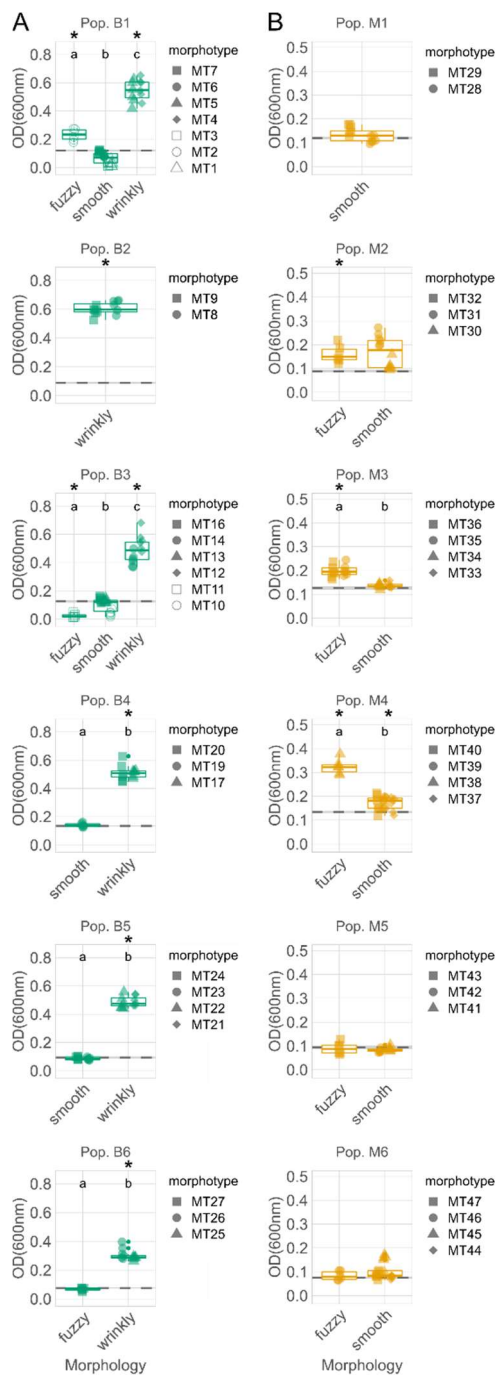


**Fig. S5 Emergence of novel colony morphologies across experimental evolution.** Proportions of morphologies (smooth, wrinkly, fuzzy) scored during colony counting after the regular bottlenecks of the evolution experiment for biphasic (BI) or monophasic (MONO) populations. First appearance in this graph indicates the cycle when the morphotype was first observed during the evolution experiment. Replicate populations are differentiated by shape and colors indicate morphology.

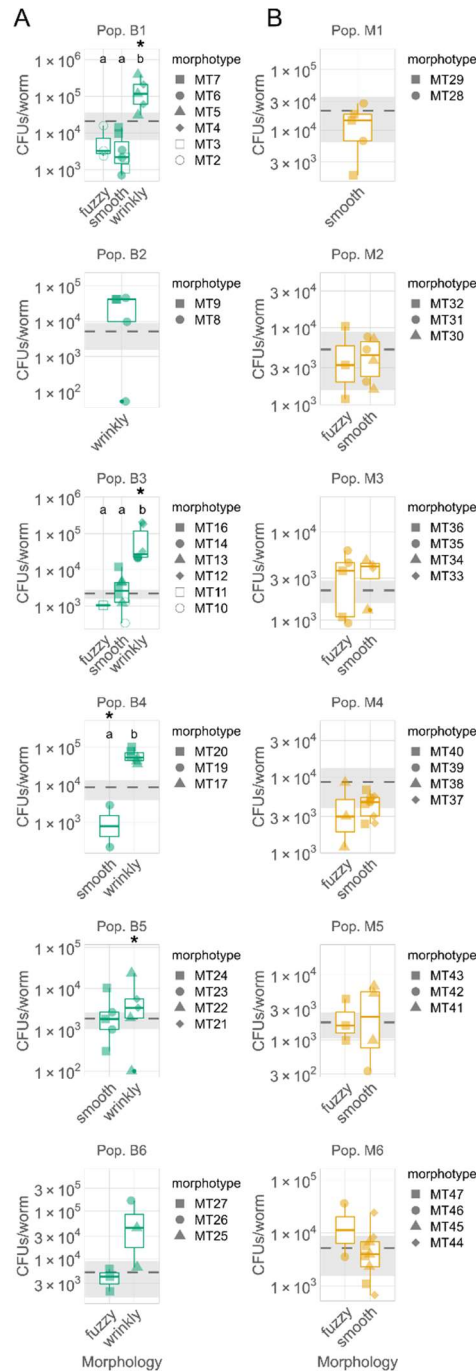


**Fig. S6** Fit and predictions of beta regression on morphological composition of evolved populations over time on agar. (A-B) Model fit and predictions for evolved biphasic and (C-D) monophasic populations.

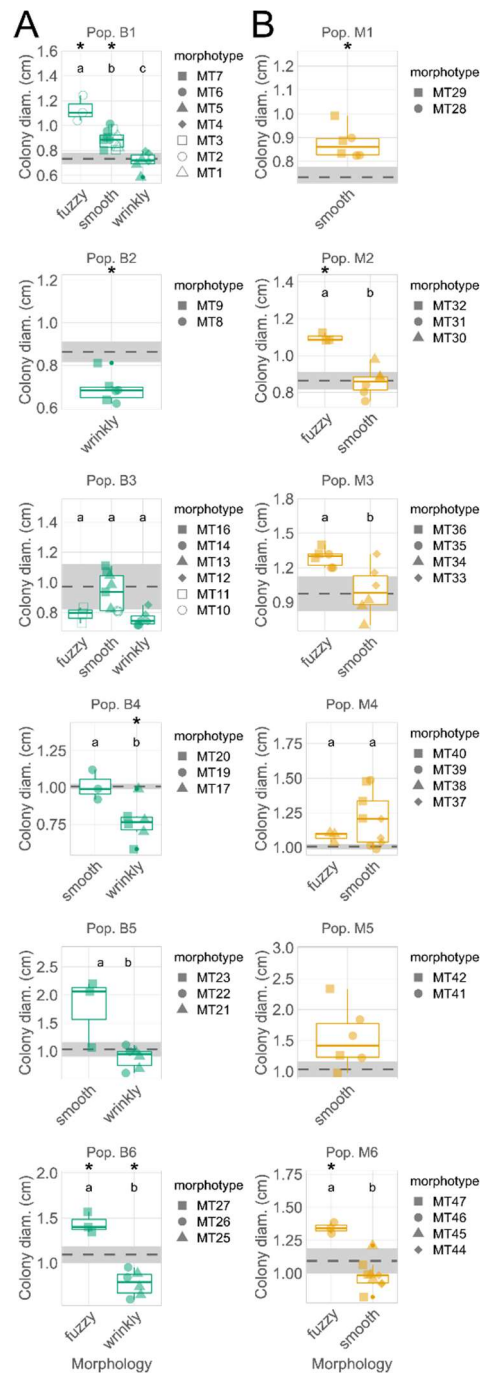




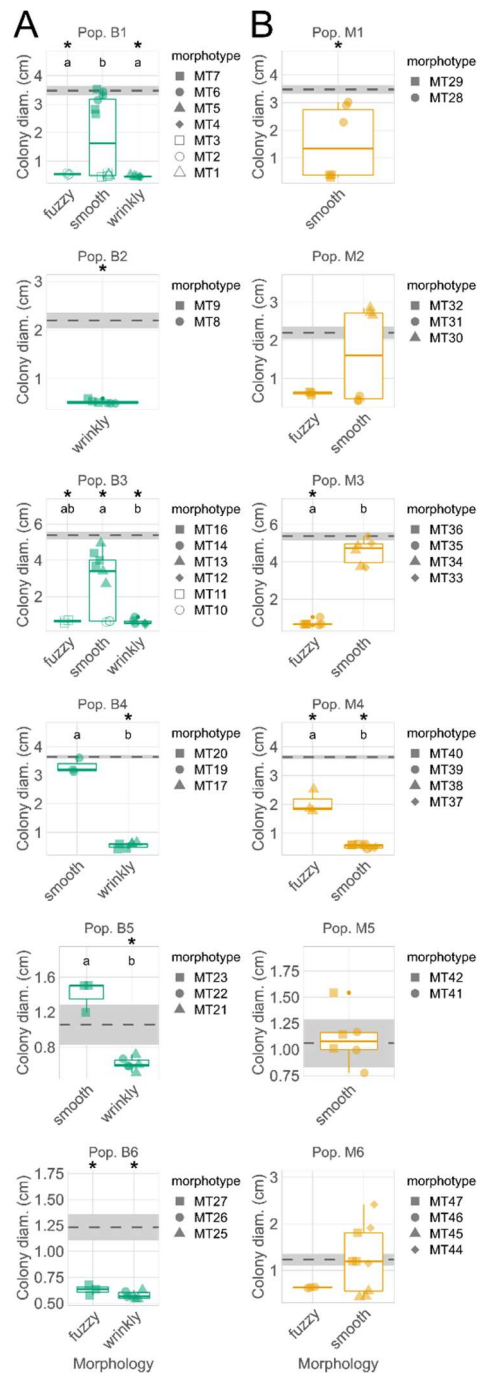
**Fig. S7 Biofilm formation of evolved morphotypes from the different replicate populations.** Biofilm formation of isolates from (A) biphasic (in green), and (B) monophasic populations (yellow) was measured after 2 days as absorption of stained biofilm previously attached to wells of a multititer plate. Different isolates are numbered consecutively (MT1, MT2, etc). For each population, they are shown in different shapes, while technical replicates ( $n=5$ ) are given as individual data points. Mean and standard error of ancestral MYb11 are shown as dashed line and shaded grey area, respectively. Differences between morphologies and ancestral MYb11 were detected using ANOVA and Tukey post-hoc tests. Morphologies with significantly different phenotypes are labeled with different letters. Additionally, asterisks above an isolate highlight significant deviation from the ancestral phenotype ( $p < 0.05$ ).



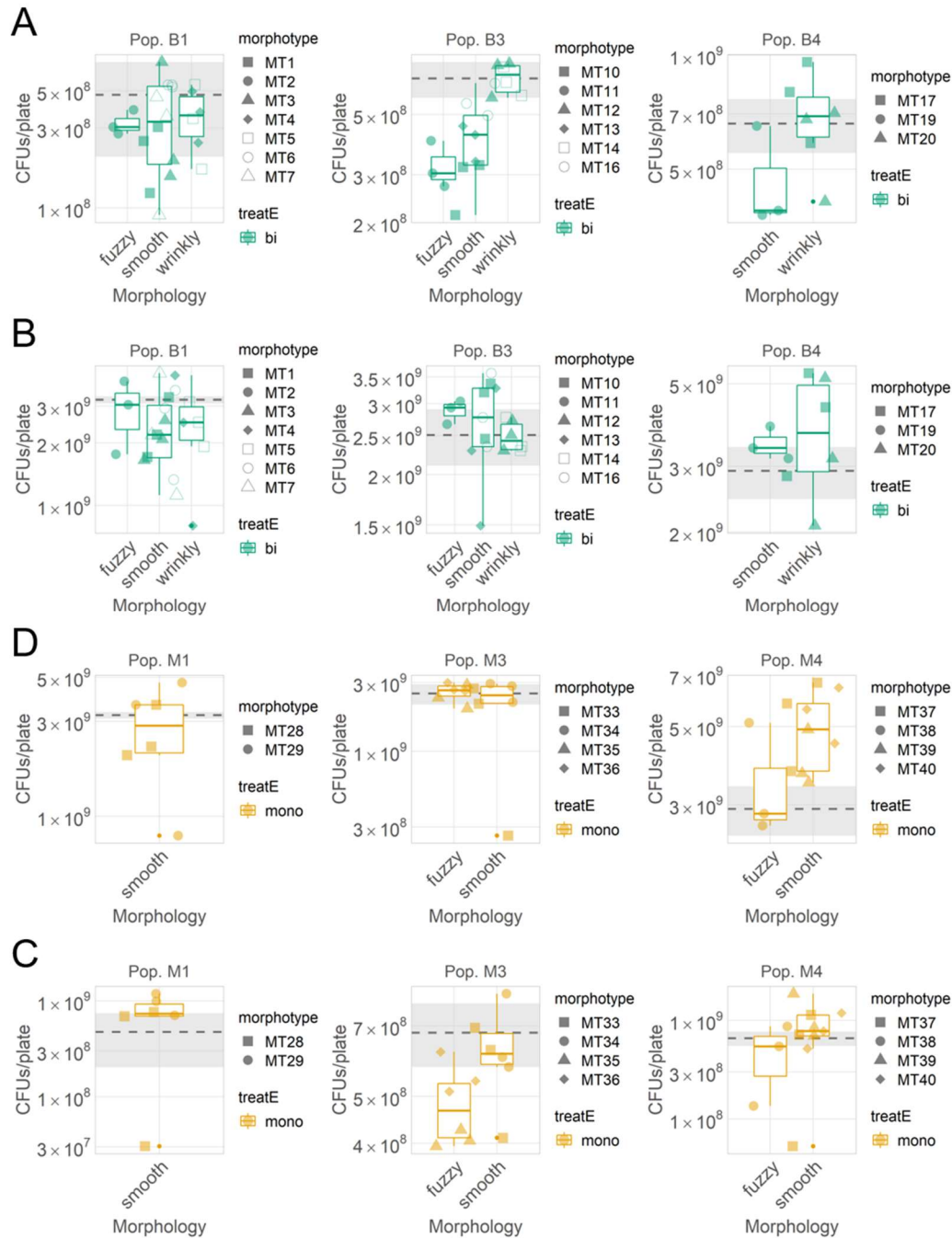
**Fig. S8 Short-term persistence of evolved morphotypes from the different replicate populations.** (A) Isolates from biphasic (green), and (B) monophasic populations (yellow). Short-term persistence in *C. elegans* hosts previously raised on the respective morphotypes were measured as colony forming units (CFUs) remaining in L4 stage worms after 1 hour in bacteria-free buffer. Different isolates are numbered consecutively. For each population, they are shown in different shapes and technical replicates ( $n=3$ ) are given as individual data points. Mean and standard error of ancestral MYb11 are shown as dashed line and shaded grey area, respectively. Differences in  $\log_{10}(\text{CFU}/\text{worm})$  between morphologies and ancestral MYb11 were detected using Generalized linear models (family=Gamma) and Tukey post-hoc tests. Morphologies with significantly different phenotypes are labeled with different letters. Additionally, asterisks above an isolate highlight significant deviation from the ancestral phenotype ( $p < 0.05$ ).



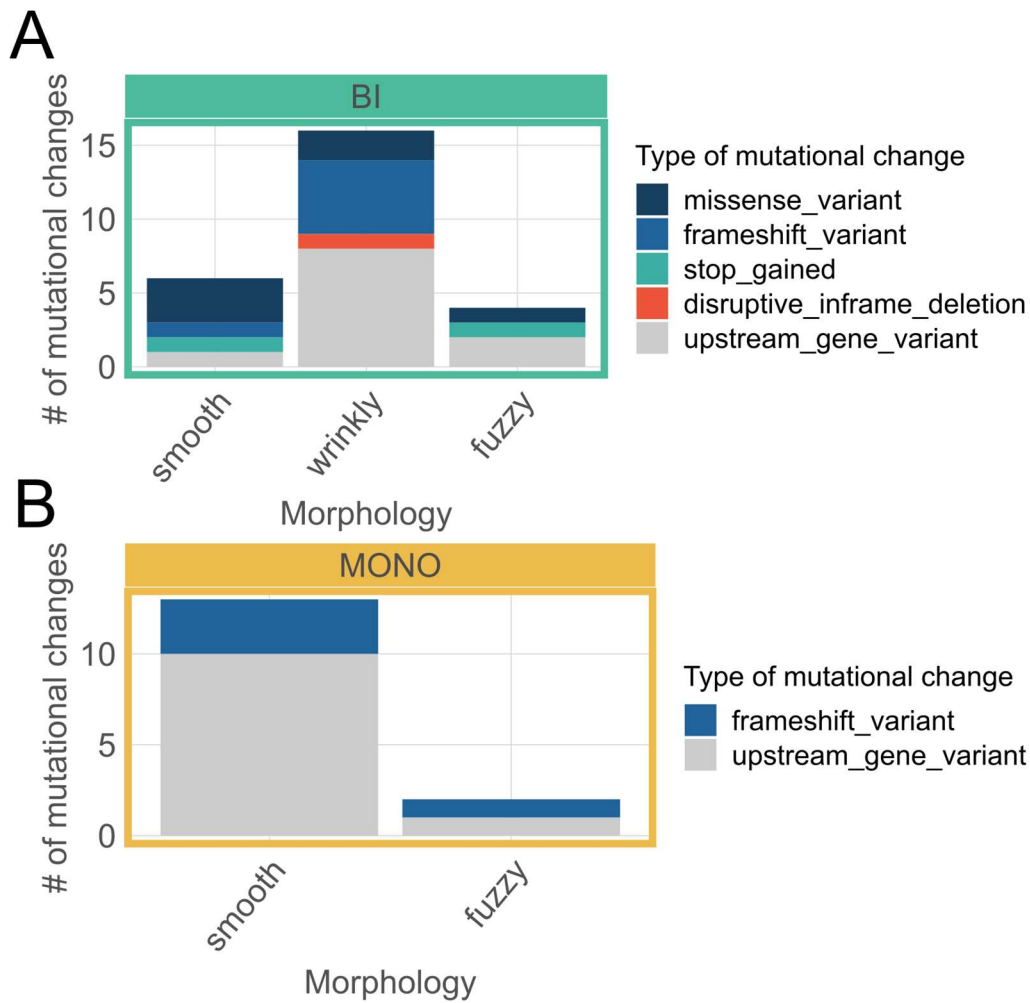
**Fig. S9 Bacterial colony expansion of evolved morphotypes from the different replicate populations.** Expansion was measured on NGM agar (3.4% agar) after 72h. **(A)** Isolates from biphasic, and **(B)** monophasic populations are shown. Different isolates are consecutively numbered. For each population, they are shown in different shapes and technical replicates ( $n=3$ ) are given as individual data points. Mean and standard error of ancestral MYb11 are shown as dashed line and shaded grey area, respectively. Differences between morphologies and ancestral MYb11 were detected using ANOVA and Tukey post-hoc tests. Morphologies with significantly different phenotypes are labeled with different letters. Additionally, asterisks above an isolate highlight significant deviation from the ancestral phenotype ( $p < 0.05$ ).



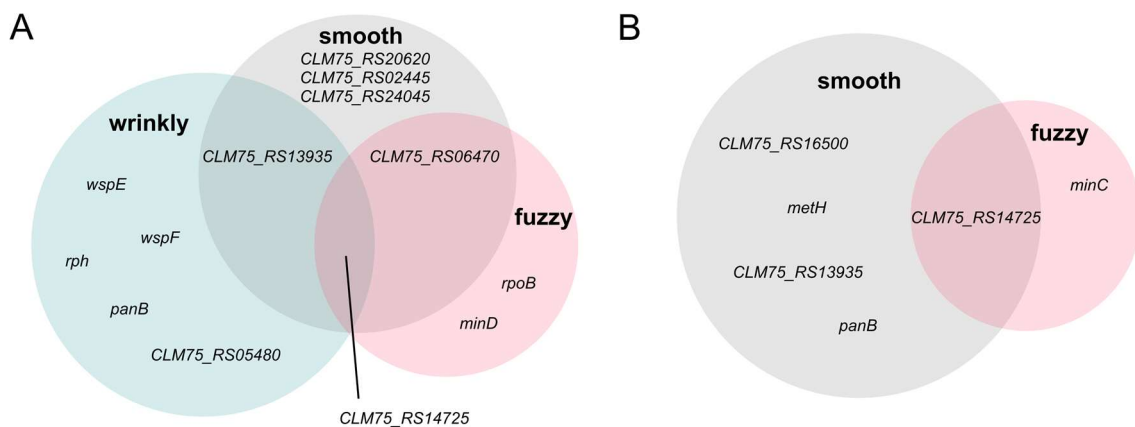
**Fig. S10 Bacterial swarming of evolved morphotypes from the different replicate populations.** Swarming was measured as maximum spot diameter after 24h on 0.5% agar. **(A)** Isolates from biphasic (in green), and **(B)** monophasic populations (yellow). Different isolates are consecutively numbered. For each replicate population, they are shown in different shapes and technical replicates ( $n=3$ ) are given as individual data points. Mean and standard error of ancestral MYb11 are shown as dashed line and shaded grey area, respectively. Differences between morphologies and ancestral MYb11 were detected using ANOVA and Tukey post-hoc tests. Morphologies with significantly different phenotypes are labeled with different letters. Additionally, asterisks above an isolate highlight significant deviation from the ancestral phenotype ( $p < 0.05$ ).



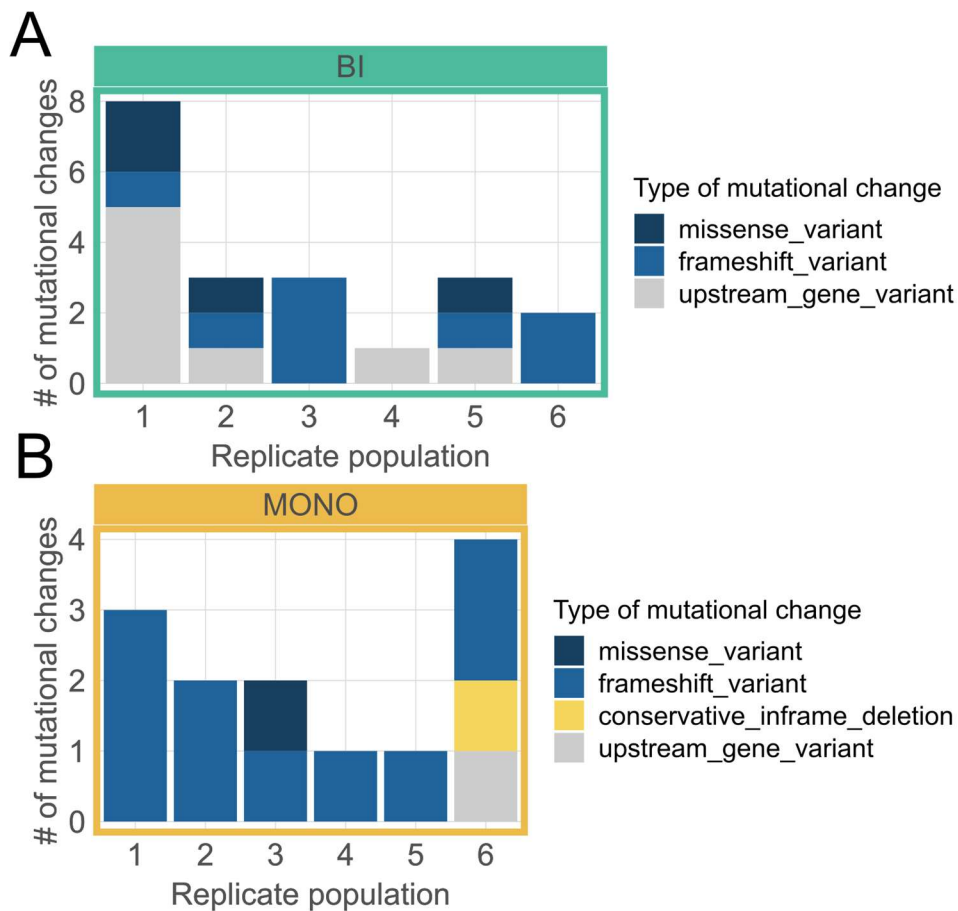
**Fig. S11 Growth of morphotypes from the different evolved replicate populations on nematode growth agar (3.4%).** Number of colony forming units collected for the (A, B) biphasic isolates (in green) after 24h and 3 days (~half a cycle), as well as (C, D) monophasic isolates (in yellow). The isolated morphotypes are numbered consecutively (MT1, MT2, etc). The different isolates per replicate population are shown in different shapes and technical replicates ( $n=3$ ) are given as individual data points. Mean and standard error of ancestral MYb11 are shown as dashed line and shaded grey area, respectively. Differences between morphologies and ancestral MYb11 were assessed using ANOVA and Tukey post-hoc tests. Morphologies with significantly different phenotypes are labeled with different letters. Additionally, asterisks above an isolate highlight significant deviation from the ancestral phenotype ( $p < 0.05$ ).



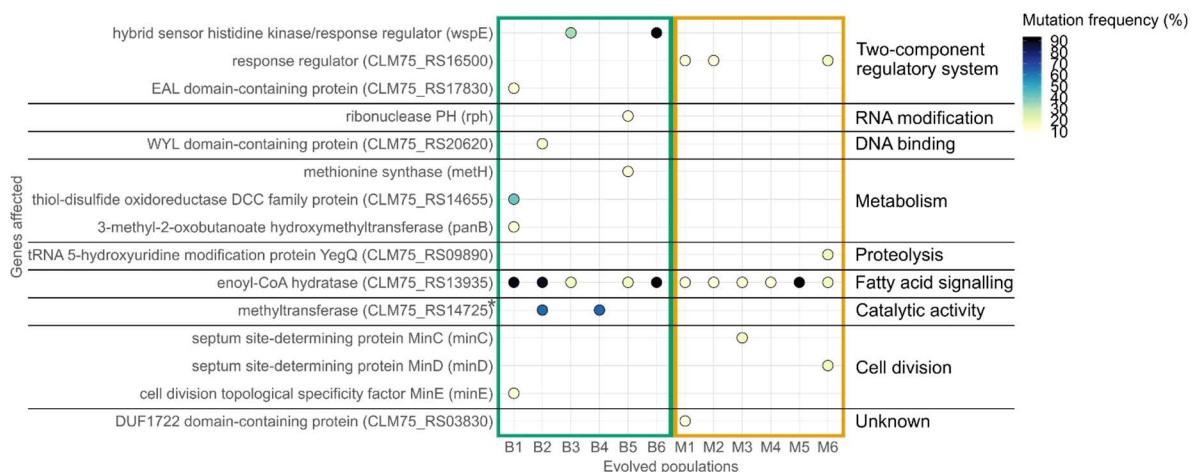
**Fig. S12 Genomics changes in isolated morphotypes from evolved populations.** Average number of mutations detected per isolate from (A) biphasic and (B) monophasic populations shown by evolutionary life cycle treatment and morphology. Biphasic isolates include 10 smooth, 12 wrinkly and 4 fuzzy isolates, while monophasic isolates include 14 smooth and 6 fuzzy isolates. Types of mutational changes are highlighted by color in the stacked bar plot.



**Fig. S13 Venn diagrams of genes with evolved variants detected in wrinkly, fuzzy, and smooth morphotypes considering isolates from (B) biphasic and (C) monophasic populations.**



**Fig. S14 Genomics changes in evolved populations.** Average number of mutations detected per isolate shown in (A) biphasic and (B) monophasic populations. Types of mutational changes are highlighted by color in the stacked bar plot.



**Fig. S15 Genomic changes in evolving populations of MYb11.** (A) Genomic variants detected in evolved biphasic and monophasic populations. Datapoints indicate genomic variants in the respective genes that differ from the ancestral MYb11. Color shows the mutation frequency (i.e. percentage of population carrying the novel genomic variant) with low frequencies in yellow ranging to high in dark blue and black. Green and yellow frames around graphs indicate whether bacterial populations evolved in association with a host (biphasic life cycle) or without (monophasic life cycle).

**A**

AM181176.4:1356352-1358619 NZ_CP023272.1:1337209-1339485	MTDQMRDASLLELFSLEADAQTQVLSAGLLALERNPTQADQL EACMRAHSLKGAARIV MTDQMRDASLLELFSLEADAQTQVLSAGLLALERNPTQADQL EACMRAHSLKGAARIV *****
AM181176.4:1356352-1358619 NZ_CP023272.1:1337209-1339485	GVDAGVSVSHMEDCLVSAQENRILYQPEHIDALLQGTDLNRITATPGNDVGPADVEAYV GVDAGVSVSHMEDCLVSAQENRILYQPEHIDALLQGTDLNRITATPGNDVGPADIEAYV *****
AM181176.4:1356352-1358619 NZ_CP023272.1:1337209-1339485	ALMERLLDPSQAPVNVAP--SPEPAPVVEELPPEPEPAPVISEPPRQKRMTEGGERV ALMERLLDPSQAPVNVAP--SPEPAPVVEELPPEPEPAPVISEPPRQKRMTEGGERV *****
AM181176.4:1356352-1358619 NZ_CP023272.1:1337209-1339485	LRVTAERLNSLDLSSKSLVETQRLKPYLALQRLKRIQSQGTRALDTLDGQLKTQVLSL LRVTAERLNSLDLSSKSLVETQRLKPYLALQRLKRIQSQGTRALDTLDGHLKTQHLNL *****
AM181176.4:1356352-1358619 NZ_CP023272.1:1337209-1339485	EAQEALADTRRLLSEAQALLAEKHAELDEFQWQAGQRAQVLYDTALACRMRPFADVLAGQ EAQEALADTRRLLSEAQALLAEKHAELDEFQWQAGQRAQVLYDTALACRMRPFADVLAGQ *****
AM181176.4:1356352-1358619 NZ_CP023272.1:1337209-1339485	VRMVRDLGRSLGKQVRLIEIEGEKTVDRDVLKLEAPLTHLLRNAVDHGIEMPEQRLLAG VRMVRDLGRSLGKQVRLIEIEGEKTVDRDVLKLEAPLTHLLRNAVDHGIEMPEQRLLAG *****
AM181176.4:1356352-1358619 NZ_CP023272.1:1337209-1339485	KPAEGLIRLRASHQAGLLVLELSDDGNGVDLERLGRITVDRHLSPVETALRSEEEELTF KPAEGLIRLRASHQAGLLVLELSDDGNGVDLERLGRITVDRHLSPAETALRSEEEELTF *****
AM181176.4:1356352-1358619 NZ_CP023272.1:1337209-1339485	LFLPGFSLRDTVTEVSGRQVGLDAVQHMVRLRGAVLEQTAGQGRFHLVPLTSLVWR LFLPGFSLRDKVTEVSGRQVGLDAVQHMVRLRGAVLEQTAGRGRFHLVPLTSLVWR *****
AM181176.4:1356352-1358619 NZ_CP023272.1:1337209-1339485	SLVVEGEEAYAFPLAHIERMCDLAPDDIVQL EGRQHFHWEGRHVLVAASQLLQRPAGQ SLVVEGEEAYAFPLAHIERMCDLAPDDIVQL EGRQHFHWEGRHVLVAASQLLQRPAGQ *****
AM181176.4:1356352-1358619 NZ_CP023272.1:1337209-1339485	SPSETLKVWIRERDVTYVIGIAVERF IGERTLVVLPDRLGKVVQDISAGALLDDGQSVLI NQSETLKVWIRERDAVYVIGIAVERF IGERTLVVLPDRLGKVVQDISAGALLDDGQSVLI *****
AM181176.4:1356352-1358619 NZ_CP023272.1:1337209-1339485	VDVEMLRSDVKLLNTGRLEIARRSQOTTEAPRKRVLVDDSLTVRELQRKLLNLRGYE VDVEMLRSDVKLLNTGRLEIARRSQOTTEAPRKRVLVDDSLTVRELQRKLLNLRGYE *****
AM181176.4:1356352-1358619 NZ_CP023272.1:1337209-1339485	VAVAVDGMGNALRSEDFDLITDIDMPMDGIELVTLRRDSRLQSLPMMVSYKDRE VAVAVDGMGNALRSEDFDLITDIDMPMDGIELVTLRRDSRLQSLPMMVSYKDRE *****
AM181176.4:1356352-1358619 NZ_CP023272.1:1337209-1339485	EDRRGLDAGADYYLAKASFHDDALLDAVVELGGARA EDRRGLDAGADYYLAKASFHDDALLDAVVELGGARA *****

**B**

Accession	Name	Source Database	Type	Integrated Into	Integrated Signatures	GO Terms	Protein Accession	Protein Length	Matches
IPR008207	Signal transduction histidine kinase, phosphotransfer (Hpt) domain	interpro	domain		SM00073,PF01627,PS50894,cd00088	GO:0000160	c3ke14	755	4..108
IPR005467	Histidine kinase domain	interpro	domain		PS50109		c3ke14	755	225..473
IPR003594	Histidine kinase/HSP90-like ATPase	interpro	domain		SM00387,PF02518,PF13581		c3ke14	755	330..473
IPR004358	Signal transduction histidine kinase-related protein, C-terminal	interpro	domain		PR00344	GO:0016772,GO:0016310	c3ke14	755	375..389,433..451,457..470
IPR002545	CheW-like domain	interpro	domain		SM00260,PF01584,PS50851	GO:0006935,GO:0007165	c3ke14	755	465..609
IPR001789	Signal transduction response regulator, receiver domain	interpro	domain		SM00448,PF00072,PS50110,cd00156	GO:0000160	c3ke14	755	632..750
IPR036641	HPT domain superfamily	interpro	homologous_superfamily		G3DSA:1.20.120.160,SSF47226	GO:0000160	c3ke14	755	5..140
IPR036890	Histidine kinase/HSP90-like ATPase superfamily	interpro	homologous_superfamily		G3DSA:3.30.565.10,SSF55874		c3ke14	755	284..472
IPR036061	CheW-like domain superfamily	interpro	homologous_superfamily		SSF50341	GO:0006935,GO:0007165	c3ke14	755	472..604
IPR011006	CheY-like superfamily	interpro	homologous_superfamily		SSF52172		c3ke14	755	626..752

**Fig. S16 Interpretation of sensor histidine kinase (CLM75\_RS06090).** (A) Amino acid alignment of CLM75\_RS06090 of ancestral MYb11 (RefSeq: NZ\_CP023272.1) with *wspE* from *Pseudomonas fluorescens* SBW 25 (Genbank: AM181176.4) using ClustalW. (B) Interpro protein domain annotation of *wspE*.



Supplementary tables

**Table S1 Isolated morphotypes emerging in life cycle evolution experiment.** Morphotypes were numbered consecutively given their origin in biphasic, then monophasic evolved populations. MT15 and MT18 were excluded from analyses, as sequencing revealed them to be food *E. coli* OP50, which had hitchhiked along the passaging regime and are thus not of interest in this study.

Morphotype ID	Replicate population of origin	Evo. life cycle	Morphology	Description
MT48	-	ancestral	anc	ancestral
MT1	1	biphasic	smooth	ancestral-like
MT2	1	biphasic	fuzzy	fuzzy
MT3	1	biphasic	smooth	ancestral-like, more wet colony
MT4	1	biphasic	wrinkly	crumbly
MT5	1	biphasic	wrinkly	crumbly
MT6	1	biphasic	smooth	ancestral-like, colony size variation
MT7	1	biphasic	smooth	ancestral-like
MT28	1	monophasic	smooth	very wet and flat colony
MT29	1	monophasic	smooth	ancestral-like, more wet colony
MT8	2	biphasic	wrinkly	crumbly
MT9	2	biphasic	wrinkly	crumbly
MT30	2	monophasic	smooth	ancestral-like, colony size variation
MT31	2	monophasic	smooth	very wet and flat colony
MT32	2	monophasic	fuzzy	fuzzy
MT10	3	biphasic	smooth	ancestral-like, colony size variation
MT11	3	biphasic	fuzzy	ancestral-like, slightly fuzzy
MT12	3	biphasic	wrinkly	crumbly
MT13	3	biphasic	smooth	ancestral-like, colony size variation
MT14	3	biphasic	wrinkly	slightly crumbly
MT15	3	biphasic	fuzzy	fuzzy
MT16	3	biphasic	smooth	ancestral-like, more wet colony
MT33	3	monophasic	smooth	ancestral-like, colony size variation
MT34	3	monophasic	smooth	ancestral-like, colony size variation

MT35	3	monophasic	fuzzy	fuzzy
MT36	3	monophasic	fuzzy	fuzzy
MT17	4	biphasic	wrinkly	crumbly
MT18	4	biphasic	fuz.ecol	resembles E. coli OP50 (fuzzy)
MT19	4	biphasic	smooth	ancestral-like, colony size variation
MT20	4	biphasic	wrinkly	crumbly
MT37	4	monophasic	smooth	very wet and flat colony
MT38	4	monophasic	fuzzy	fuzzy
MT39	4	monophasic	smooth	very wet and flat colony
MT40	4	monophasic	smooth	ancestral-like, more wet colony
MT21	5	biphasic	wrinkly	crumbly
MT22	5	biphasic	wrinkly	crumbly
MT23	5	biphasic	smooth	ancestral-like, more wet colony
MT24	5	biphasic	smooth	ancestral-like, more wet colony
MT41	5	monophasic	smooth	ancestral-like, colony size variation
MT42	5	monophasic	smooth	very wet and flat colony
MT43	5	monophasic	fuzzy	fuzzy
MT25	6	biphasic	wrinkly	crumbly
MT26	6	biphasic	wrinkly	crumbly
MT27	6	biphasic	fuzzy	fuzzy
MT44	6	monophasic	smooth	ancestral-like, more wet colony
MT45	6	monophasic	smooth	ancestral-like, more wet colony
MT46	6	monophasic	fuzzy	fuzzy
MT47	6	monophasic	smooth	ancestral-like, more wet colony

**Table S2 Principal component analysis (PCA) of the phenotypes of ancestral MYb11 and evolved biphasic and monophasic populations along the stages of the biphasic life cycle. This includes early colonization, colonization, short-term persistence (1h) and long-term persistence (24h) in worms without access to additional bacteria, as well as release from *C. elegans* and swarming and colony expansion on NGM agar.**

	PC1	PC2	PC3	PC4	PC5	PC6
Standard deviation	1.5621	1.3558	0.9912	0.6874	0.4264	0.2914
Proportion of Variance	0.4067	0.3064	0.1637	0.0788	0.0303	0.0142
Cumulative Proportion	0.4067	0.7131	0.8768	0.9556	0.9859	1.0000
Colonization	-0.4335	-0.2690	0.5708	-0.2582	0.5325	0.2524
Short-term persistence	-0.4754	-0.2076	-0.5819	-0.0360	-0.1707	0.6015
Release	-0.5944	-0.1599	-0.1283	-0.2249	-0.1957	-0.7181
Long-term persistence	-0.3018	0.5514	-0.2680	0.4254	0.5817	-0.1169
Swarming	0.3761	-0.4088	-0.4928	-0.3249	0.5566	-0.1821
Colony expansion	0.0158	0.6227	-0.0667	-0.7713	-0.0269	0.1091

**Table S3 Pairwise comparison of evolved vs. ancestral phenotype sets following permutational analysis of variance. P-values were adjusted using *fdr*.**

pairs	Df	Sum of Squares	F-value	R <sup>2</sup>	p-value	Adj. p-value
anc vs bi	1	0.1056	8.8564	0.4697	0.0040	0.0120
anc vs mono	1	0.0445	2.9695	0.2290	0.0310	0.0929

**Table S4 Principal component analysis (PCA) of the phenotypes (biofilm formation, short-term persistence in worms, colony expansion and swarming) of morphotypes isolated from biphasically evolved and monophasically evolved populations.**

Biphasic morphotypes					
		PC1	PC2	PC3	PC4
	Standard deviation	1.5868	0.9246	0.6383	0.4688
	Proportion of Variance	0.6295	0.2137	0.1019	0.0549
	Cumulative Proportion	0.6295	0.8432	0.9451	1.0000
	Biofilm formation	-0.3918	0.7854	0.4364	0.1979
	Swarming	-0.4518	-0.6020	0.6568	0.0462
	Colony expansion	0.5573	0.1295	0.5452	-0.6127
	Short-term persistence	-0.3918	0.7854	0.4364	0.1979
Monophasic morphotypes					
		PC1	PC2	PC3	PC4
	Standard deviation	1.2095	1.1068	0.9949	0.5678
	Proportion of Variance	0.3657	0.3062	0.2475	0.0806
	Cumulative Proportion	0.3657	0.6720	0.9194	1.0000
	Biofilm formation	0.1688	-0.3133	0.9131	-0.1989
	Swarming	-0.7707	-0.0147	-0.0014	-0.6370
	Colony expansion	0.4302	-0.6288	-0.4054	-0.5052
	Short-term persistence	0.4387	0.7115	0.0439	-0.5472

**Table S5** Pairwise comparison of morphologies within biphasic and monophasic isolate phenotype sets following permutational analysis of variance. *P*-values were adjusted using *fdr*.

Biphasic morphotypes							
	pairs	Df	Sum of Squares	F-value	R <sup>2</sup>	p-valze	Adj. p-value
	smooth vs ancestor	1	0.0075	0.3502	0.0309	0.7303	0.7303
	fuzzy vs ancestor	1	0.0959	5.1650	0.4246	0.0260	0.0390
	wrinkly vs ancestor	1	0.2619	28.0978	0.6372	0.0010	0.0030
Monophasic morphotypes							
	pairs	Df	Sum of Squares	F-value	R <sup>2</sup>	p-valze	Adj. p-value
	smooth vs ancestor	1	0.0475	2.3348	0.1208	0.1349	0.1349
	fuzzy vs ancestor	1	0.0680	4.1882	0.3436	0.0649	0.1299



# 5

## **Distinct bacterial life history strategies within the natural microbiota of *C. elegans***

Nancy Obeng  
Hinrich Schulenburg

Manuscript in preparation

## Abstract

Commonly microbiotas are composed of bacteria from different taxa that may differ in their functional repertoires. To associate with the host, they must enter and persist in a host, and often grow or survive in environmental reservoirs, thus following a biphasic life cycle. To what extent species co-inhabiting a host differ in their life history and fitness strategies, however, is unknown. We here use a two-member model microbiota based on two key members of the natural microbiota of the nematode *Caenorhabditis elegans* to infer bacterial strategies of host association. Specifically, we compare *Pseudomonas lurida* strain MYb11 and *Ochrobactrum vermis* strain MYb71 in mono- and co-culture in worms and during growth on nematode growth agar. While on agar the two species show antagonistic interactions. These are attenuated during co-colonization of their native host strain *C. elegans* MY316. Tracking bacterial fitness along a biphasic life cycle with a host-associated and a free-living phase in worms and on agar, we find the species use distinct fitness strategies. While MYb11 specializes on rapid growth in the free-living phase, quick entry into and long-term persistence within worms, MYb71 invests in competition on agar and accumulation in worms over time. Experimentally co-evolving in such a biphasic life cycle emphasizes these interspecies differences and leads to MYb11 dominating the community on agar and increasing its proportion during worm association. Finally, we examine the effect of microbiota composition on host development and fitness. Here MYb11 and MYb71 have largely similar effects, though worms grown on MYb71 alone are slenderer than on the co-culture. Together, this shows how intra-microbiota dynamics are context-dependent and highlights a diversity in life history strategies across bacterial species from the same microbiota. It therefore presents an important building block in understanding the complexity and fate of interspecies interactions in microbiotas.



## Introduction

Bacteria from different taxa live together and interact with each other in host-associated communities, so-called microbiotas. Yet, while the taxonomic composition of microbiotas from wild wheat (Hassani et al., 2020) to gutless tube worms (Dubilier et al., 2001), and different body sites of humans (The Human Microbiome Project Consortium et al., 2012) has been described, we are only starting to understand the reasons for microbial adaptation to their host. Moving from sequencing of 16S rRNA genes, which allows identifying operational taxonomic units or amplified sequence variants, towards metagenomics and whole genome sequencing approaches has helped in predicting bacterial functions and relate them to life with a host. From this we learned for example, that bacteria can synthesize and exchange nutrients with the host (Kwong et al., 2014; Martino et al., 2018b; Zimmermann et al., 2019). Such functional catalogues, however, only provide a basis to imagine the ecological reality of bacteria in the microbiota. What aspects of host association certain bacterial taxa specialize in, how interactions within the microbiota affect the success of individual species, and how this related to bacterial fitness in a symbiotic life history cannot easily be inferred from such sequence data.

Naturally, most microbiota bacteria follow a biphasic life cycle, in which they must both grow and survive in a free-living environment as well as associate with a host (see **chapter 2**). In such a biphasic life cycle, bacteria transition through different stages of association from initiation and establishment of host association to release back into the environment, and possibly dispersal to a novel host. Such a life cycle has been described in many facultative symbioses, including the famous association of *Vibrio fischeri* and the bobtail squid (Nyholm and McFall-Ngai, 1998), and is also experienced by horizontally acquired microbiota bacteria (Martino et al., 2016; Shukla et al., 2018b). Experimentally quantifying bacterial abundances along the life cycle can allow us to infer key stages of the life cycle that contribute to fitness. Here bacteria must balance resource allocation to survival and reproduction, two key fitness components, along the stages of the life cycle (Fierer, 2017). How they balance this trade-off along their life then indicates bacterial life history strategies in host association and may reveal variation in strategies between microbiota members.

In this study, we aim to understand how different species adjust to association with the same host and how their interaction shapes host association. We address this in a simplified community of two species co-isolated from the natural microbiota of the model nematode *Caenorhabditis elegans* (Dirksen et al., 2016). Here we chose the culturable isolates *Pseudomonas lurida* MYb11 and *Ochrobactrum vermis* MYb71 capable of colonizing and persisting in worms and shown to interact with each other in a context-dependent fashion *in vitro* (Dirksen et al., 2016; Zimmermann et al., 2019). Notably, both belong to core taxa of the *C. elegans* microbiota (Zhang et al., 2017) and are of key relevance for *C. elegans*, as both, especially MYb11, can protect worm hosts from pathogen infection (Dirksen et al., 2016; Kissoyan et al., 2019), and MYb71 was shown to induce physiological changes in the worm spanning from development to metabolism (Cassidy et al., 2018; Yang et al., 2019). For the current study, we associate MYb11 and MYb71 with worms, alone and in combination, and compare their strategies of host association. Further, we consider how these species interact along the symbiotic life cycle composed of a host-associated and a free-living phase and how community composition is affected by the life cycle. For this, we allow the two isolates (alone and in co-culture) to experimentally evolve in a biphasic life cycle with a free-living and a host-associated phase vs. a monophasic free-living lifestyle along ten worm generations and study evolved microbiota composition along the life cycle. Finally, we compare the effect of the MYb11 and MYb71 on *C. elegans* to understand how microbiota composition feeds back on host development and fitness. Together, this provides an eco-evolutionary perspective on bacterial adaptation to host association.

## Materials and methods

### *Worm and bacterial strains*

This study includes two key members of the *C. elegans* microbiome, *Pseudomonas lurida* MYb11 and *Ochrobactrum vermis* MYb71, isolated from the natural *C. elegans* strain MY316 (Dirksen et al., 2016). Both bacteria can colonize worms and be maintained on nematode growth medium (NGM) agar. Based on antibiotic resistance profiles of the strains, selective plating of co-cultures using kanamycin (10 µg/ml) was used to determine proportions of MYb11 (kanamycin-sensitive) and MYb71 (kanamycin-resistant). Further, *Escherichia coli* OP50 was used as a food bacterium to maintain *C. elegans* in preparation for microbiota experiments (Stiernagle, 2006). As *E. coli* OP50 does not colonize *C. elegans* larvae and young adults, and does not form biofilms (Arata et al., 2020; Portal-Celhay et al., 2012), it does not affect microbiota establishment. The natural *C. elegans* isolate MY316, recovered from a rotting apple in the Kiel Botanical Gardens (Dirksen et al., 2016), is used as a host organism.

### *Growth on agar*

Bacterial growth of MYb11, MYb71 or a 1:1 co-culture (inoculation based on optical density) was quantified as colony forming units grown on nematode growth agar (NGM) at 24h, 72h and 168h as described in chapter 4 of this thesis.

### *Motility*

Bacterial motility, in terms of colony expansion on 3.4% NGM and swarming at 0.5% agar, was assayed as described in chapter 4 of this thesis.

### *Worm association or colonization assays*

Bacterial colonization of *C. elegans* was measured along the stages of association. This includes early colonization of *C. elegans* MY316 L4 larvae (after 1.5h exposure), established colonization of L4 larvae (exposed to bacteria since L1 stage), short-term persistence in L4 larvae kept in M9 buffer for 1h (raised on bacteria), long-term persistence in worms kept on empty NGM agar for 24h post-development on co-culture from L1 to L4, and release from L4 larvae into buffer within 1h (previously raised on bacteria from L1 to L4). For experimental details, please refer to chapter 4 of this thesis, in which the same protocols were used for MYb11 in mono-association with *C. elegans*.

### *Evolution experiment*

Bacterial populations were serially passaged in biphasic life cycle, i.e. maintained on agar plates with exposure to *C. elegans* in regular intervals, or a monophasic life cycle, i.e. maintained on agar plates only (Figure 3A).

Specifically, MYb11 was added to nematode growth medium agar (NGM) as a bacterial lawn and allowed to grow for 3.5 days. In the host-associated phase, ten *C. elegans* L4 larvae (raised until this stage on the non-colonizing food bacterium *Escherichia coli* OP50) were added to NGM plates allowing bacteria to associate with worms until the worms' F1 generation. In the monophasic treatment, bacteria were maintained on NGM agar only for a full cycle. At the end of every cycle, a bottleneck of 10% of the bacterial population associated with the worm hosts was imposed, i.e. 10% of host-associated bacteria were transferred to the next cycle. To impose a comparable bottleneck on free-living populations, the mean number of colony forming units (CFUs) extracted from host populations in the biphasic treatment of the previous cycle was used as a bottleneck for free-living populations. To generate a fossil record, samples of bacterial populations were collected at every bottleneck in 10% DMSO and frozen at -80°C. To allow

bacteria to evolve, replicate populations of *Pseudomonas lurida* (MYb11) were passaged for 10 cycles.

#### *Analysis of evolved bacteria*

Following the evolution experiment, evolved populations from cycle 10 were recovered from frozen stocks and adaptation to the host-associated phase of the biphasic life cycle was measured in terms of colonization of and release from *C. elegans* hosts (see Figure 2). A re-run of one cycle of the evolution experiment was done to minimize any potential effects of freezing populations. Subsequently, they were adjusted to a common optical density (as a proxy for biomass) and grown on NGM agar for two days as a common garden treatment. Subsequently, growth on agar and different stages of colonization were measured as described above.

#### *Automated measurement of worm size*

Sizes of individual worms were measured from photographs of washed, frozen, and thawed worm populations. For this 5 L4 stage MY316 worms were placed onto bacterial lawns and incubated at 20°C. After 3.5 days, worms were washed off as described for the colonization assays, and washed on filter tips. Post-washing, a sample of worms in 150µl M9 buffer were frozen at -20°C in a 24-well plate. Photos were taken with the Leica fluorescence dissecting scope. To measure worm sizes, images were imported into ImageJ2 (Rueden et al., 2017) and detected as particles using a custom-written Macro (Figure S5).

Briefly, detected particles were approximated as ellipses and ellipse major and minor axis length taken as proxies for worm length and width. Worm area was calculated as the product of minor and major axis length, while worm breadth is the quotient of minor and major axis length. Particles with dimensions observed for *C. elegans* (major axis between 0.18-1.3 mm, minor axis  $\leq$  0.1 mm; as in (Mörck and Pilon, 2006)) were counted and measured. As a measure of detection quality, numbers of worms counted automatically were correlated by counts of two independent humans (Figure S5).

#### *Measuring eggs laid and worm population size*

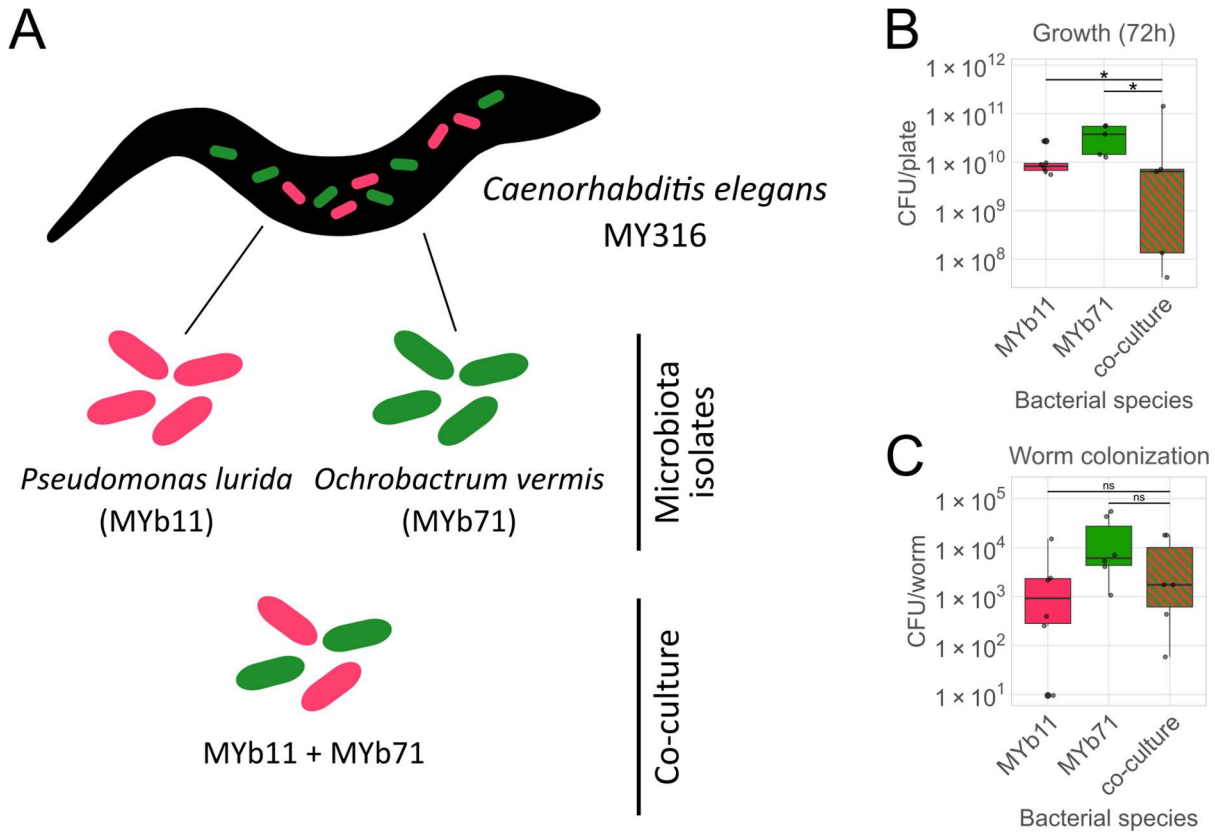
As measures of reproduction, eggs laid by worms within 24h and populations sizes reached after 3.5 days (half a life cycle) of exposure to bacterial lawns were quantified. For this 5 L4 stage MY316 worms were placed onto bacterial lawns and eggs counted under the stereomicroscope after 24h. Worm population sizes at 3.5 days were based on the number of worms detected by the image analysis algorithm above (automated measurement of worm size).

#### *Statistical analysis*

All statistical analyses were performed using R 3.6.1 in the RStudio environment (R Core Team, 2016; RStudio Team, 2015). All plots were generated using the ggplot2 package (Wickham). The assumptions of parametric models (normality and homogeneity of variances) were checked by inspecting box plots of the data. When these were not met, non-parametric tests were applied.

## Results

As a simplified model microbiota, we used two isolates from the natural microbiota of *C. elegans*, *Pseudomonas lurida* MYb11 and *Ochrobactrum vermis* MYb71 (Figure 1A). In the following we compare their abundances in worms and on agar plates as a proxy for bacterial fitness in mono- and co-culture.



**Figure 1 Interaction of two-member microbiota during growth on agar and colonization of *C. elegans* as a host.** (A) Two bacterial isolates from the natural microbiota of *Caenorhabditis elegans* MY316, *Pseudomonas lurida* MYb11 and *Ochrobactrum vermis* MYb71, serve as a model for microbe-host associations. The two bacterial species were studied in monoculture and as a two-member community in co-culture. (B) Bacterial growth was measured as colony forming units (CFUs) collected from nematode growth agar after 72h of incubation at 20°C. (C) Worm colonization was measured as CFUs collected per washed L4 *C. elegans* larva raised on the respective bacterial lawns since L1 stage. Replicates are shown as individual data points. Differences between monocultures of MYb11 and MYb71 and the co-culture detected using Dunnett tests (\*\*\* ( $p < 0.001$ ), \*\* ( $p < 0.01$ ), \* ( $p < 0.05$ )).

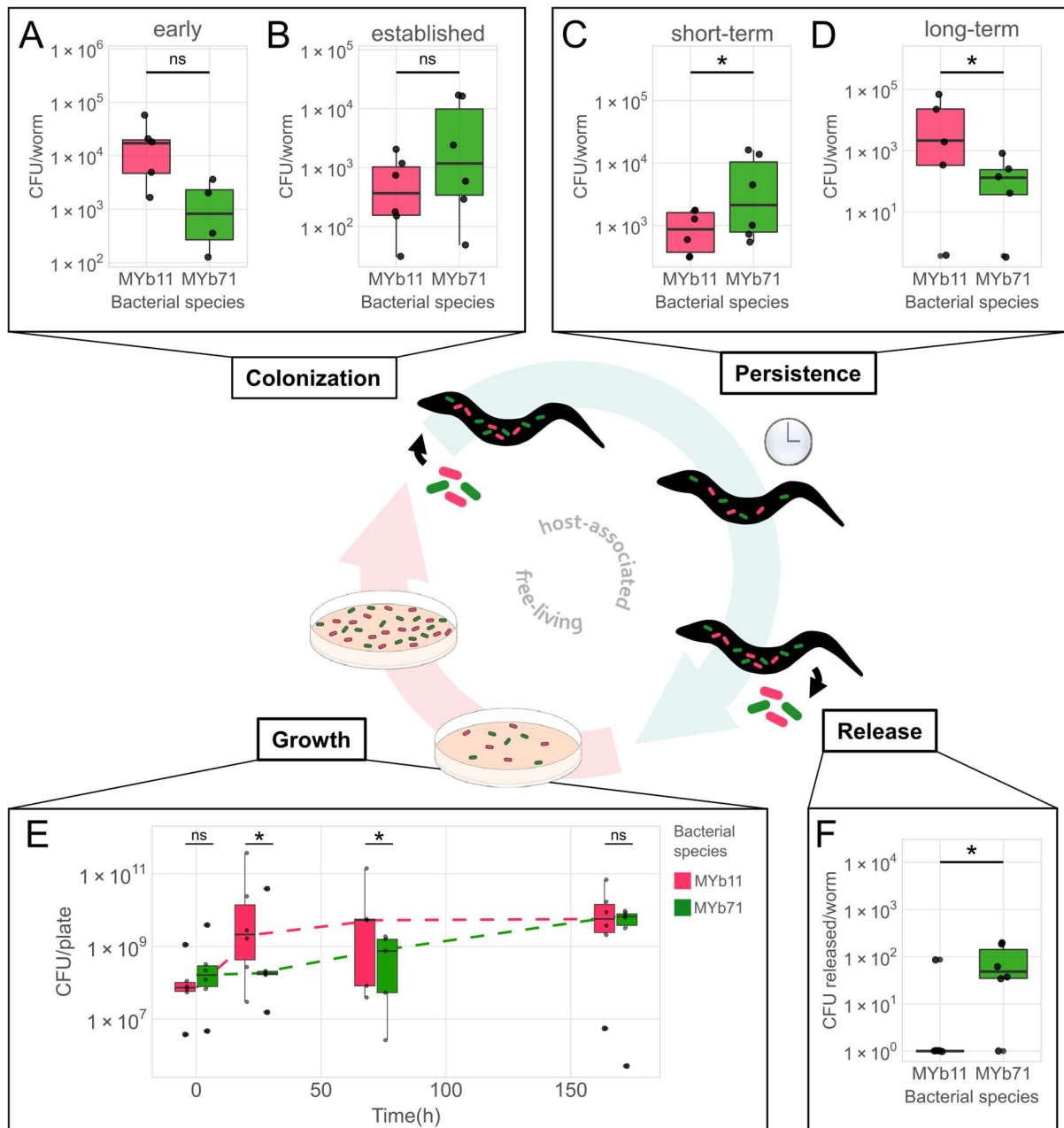
### Interspecies interactions are context-dependent

Both microbiota members can be cultivated on nematode growth agar where they acted antagonistically in co-culture (Figure 1B). In detail, after 72 hours of growth, where both species reached median population sizes around  $10^{10}$  cells per bacterial lawn (on plates of 6 cm in diameter), both *Pseudomonas* MYb11 (Dunnett,  $t(4)=3.323$ ,  $p=0.014$ ; Table S1) as well as *Ochrobactrum* MYb71 (Dunnett,  $t(4)=2.716$ ,  $p=0.039$ ; Table S1) showed significantly greater population sizes when growing alone than together in co-culture. Thus, they both suffered a growth cost indicating a competitive interaction between the two. On nematode growth agar used to maintain *C. elegans*, both bacterial species grew over a period of seven days, at which they exhibited a growth plateau (Figure S1). As expected based on their co-isolation from natural *C. elegans*, MYb11 and MYb71 could colonize worms alone and in combination (Figure 1C). In worms raised on bacterial lawns of MYb11, MYb71 or both, until L4 stage, around 1000 bacteria were

found per worm (Figure 1C). Here no significant differences between MYb11 and MYb71 mono-associations and the abundance of the two-member community in worms could be detected (MYb11: Dunnett,  $t(5)=0.023$ ,  $p=1$ ; MYb71: Dunnett,  $t(5)=0.439$ ,  $p=0.871$ ; Table S2). Thus, colonization of L4 larvae by MYb11 and MYb71 appeared additive. Based on these observations, we conclude that the interaction between the two species changes between habitats.

*The two bacteria express distinct strategies across the biphasic symbiosis life cycle*

To understand the change in interaction and bacterial fitness in more detail, we characterized species interactions along a biphasic symbiotic life cycle with *C. elegans* as a host and a free-living phase on nematode growth medium (NGM) agar. In this life cycle, bacterial communities progress from early to established colonization, are challenged to persist in worms in the short and long term (24h) without influx of additional cells, release back into the agar environment, and may grow there (Figure 2). Using selective plating, we quantified species abundances within the co-cultures and thus assessed variation in bacterial performance along the life cycle.

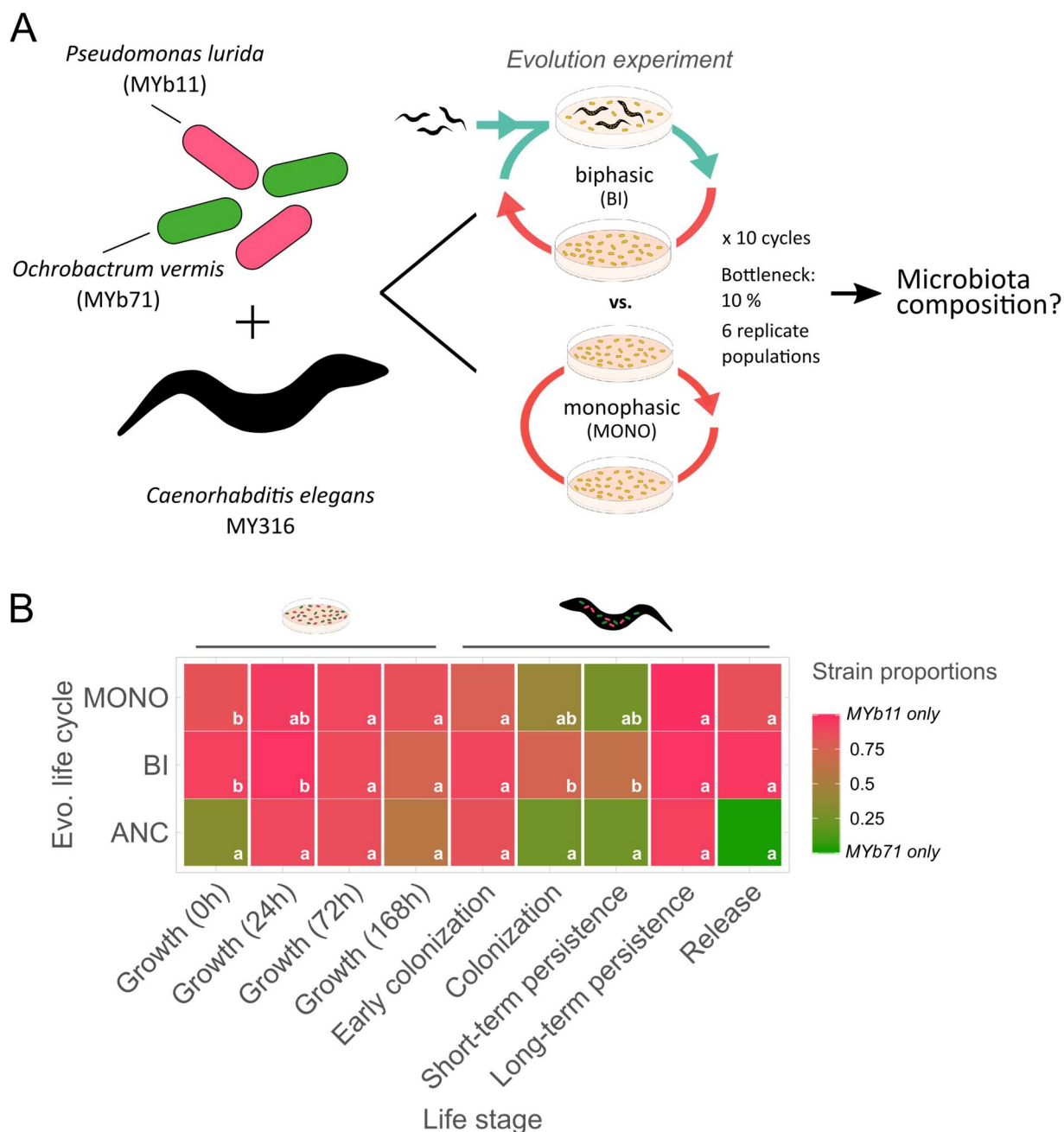


**Figure 2** *Pseudomonas lurida* MYb11 and *Ochrobactrum vermis* MYb71. have distinct fitness strategies along the symbiotic life cycle. Bacterial fitness, in terms of colony forming units per worm or agar plate, was measured in competition along the symbiotic life cycle. The stages of the life cycle include early colonization of *C. elegans* MY316 L4 larvae (after 1.5h exposure), established colonization of L4 larvae (exposed to bacteria since L1 stage), short-term persistence in L4 larvae kept in M9 buffer for 1h (raised on bacteria since L1), long-term persistence in worms kept on empty NGM agar for 24h post-development on co-culture from L1 to L4, release from L4 larvae into buffer within 1h (previously raised on bacteria from L1 to L4) and growth on nematode growth agar over a period of seven days (sampled at 0h, 24h, 72h and 168h). Paired t-tests were used to compare species abundances (\*\*\* ( $p < 0.001$ ), \*\* ( $p < 0.01$ ), \* ( $p < 0.05$ )).

In the early phase of colonization, i.e. after 1.5 hours of worm exposure, similar numbers of MYb11 and MYb71 were found per L4 stage worm, although often MYb11 dominated (Figure 2A; paired t-test,  $t(4) = 2.295$ ,  $p = 0.083$ ). Similarly, established bacterial communities in L4 larvae raised on a mixed lawn for 48 hours showed on average roughly even species composition (Figure 2B; paired t-test,  $t(5) = -2.315$ ,  $p = 0.068$ ), with MYb71 dominating in more samples than did MYb11. Persisting in worms in the stage of the life cycle, when no additional bacteria are taken up, we observed clear differences between the species, almost reversing the ratios seen during colonization. While *Ochrobactrum* MYb71 was more abundant in worms in the short term (i.e. one hour in buffer;

Figure 2C; paired t-test,  $t(5)=-3.596$ ,  $p=0.016$ ), *Pseudomonas* MYb11 could persist at significantly greater levels in the long term (i.e. 24 hours on empty agar plates; Figure 2D; paired t-test,  $t(4)=3.354$ ,  $p=0.028$ ). Here MYb11 appeared particularly good at maintaining the level of colonization around 1000 cells per worm, while MYb71 abundance dropped by about one order of magnitude (Figure 2D). This relates to the number of live bacteria that were released from worms into the environment. Here more MYb71 than MYb11, or commonly only MYb71 cells were detected after keeping worms in buffer for one hour (Figure 2F; paired t-test,  $t(5)=-3.487$ ,  $p=0.018$ ), thus indicating lowered bacterial load to a certain extent due to emigration of cells.

Bacteria further showed distinct growth strategies in the environment of the nematode growth agar. Starting growth in the co-culture at roughly even CFU densities (adjusted based on optical density of inocula; Figure 2E; paired t-test,  $t(5)=-2.139$ ,  $p=0.085$ ), *Pseudomonas* MYb11 outgrew *Ochrobactrum* MYb71 at 24 hours (paired t-test,  $t(5)=3.403$ ,  $p=0.019$ ) as well as after three days of incubation (paired t-test,  $t(4)=3.023$ ,  $p=0.040$ ). This indicates, that *Ochrobactrum* suffered the greatest growth cost in the co-culture on plate already observed in the sum of cells grown in the co-culture after three days (Figure 1). After 7 days on plate, however, both strains showed even abundances again (Figure 2E; paired t-test,  $t(5)=1.304$ ,  $p=0.249$ ), indicating an earlier exponential growth phase of MYb11 compared to MYb71 but a similar, or shared carrying capacity. In addition to growth rates, *Pseudomonas* MYb11 and *Ochrobactrum* MYb71 also differed in their spatial growth strategies. In terms of spatial expansion of colonies on nematode growth agar (3.4% agar), as well as swarming on agar plates with lower density (0.5% agar), MYb11 showed significantly greater motility than MYb71 (Figure S2, Table S3). A co-culture of the two appeared most similar to MYb11, where motility was likely exhibited mostly by this species (Figure S2, Table S3). Taken together, we observed clear differences in fitness between the species along the stages of a biphasic symbiosis life cycle.



**Figure 3 Microbiota composition responds to a biphasic symbiosis life cycle evolutionarily. (A)** In an evolution experiment, a community of *Pseudomonas lurida* MYb11 and *Ochrobactrum vermis* MYb71 was passaged either in a biphasic life cycle (with a host-associated phase with *C. elegans* MY316 and a free-living phase on nematode growth agar), or an only free-living, monophasic life cycle on NGM. At each transition between the ten cycles of evolution (here: ten worm generations), a bottleneck equivalent to 10% of the bacterial population associated with worms was imposed. Evolved microbiota communities from the end of the evolution experiment were recovered from frozen fossil record and fitness along the life cycle was assayed in comparison to ancestral co-cultures of MYb11 and MYb71. **(B)** Proportions of MYb11 and MYb71 in ancestral, evolved biphasic and evolved monophasic microbiota communities along the life stages of the biphasic symbiosis cycle. Median proportions from five to six replicates are shown using a heatmap. Differences in species proportions across co-cultures from different life cycles (ancestral (ANC), biphasically (BI) and monophasically (MONO) evolved) were detected using Generalized Linear Models with a binomial distribution. Proportions differing between evo. life cycles (vertical axis) at each of the considered life stages (horizontal axis) are indicated by different letters based on Tukey post-hoc tests.



## *Microbiota composition changed in response to the experimentally evolved bacterial life cycle*

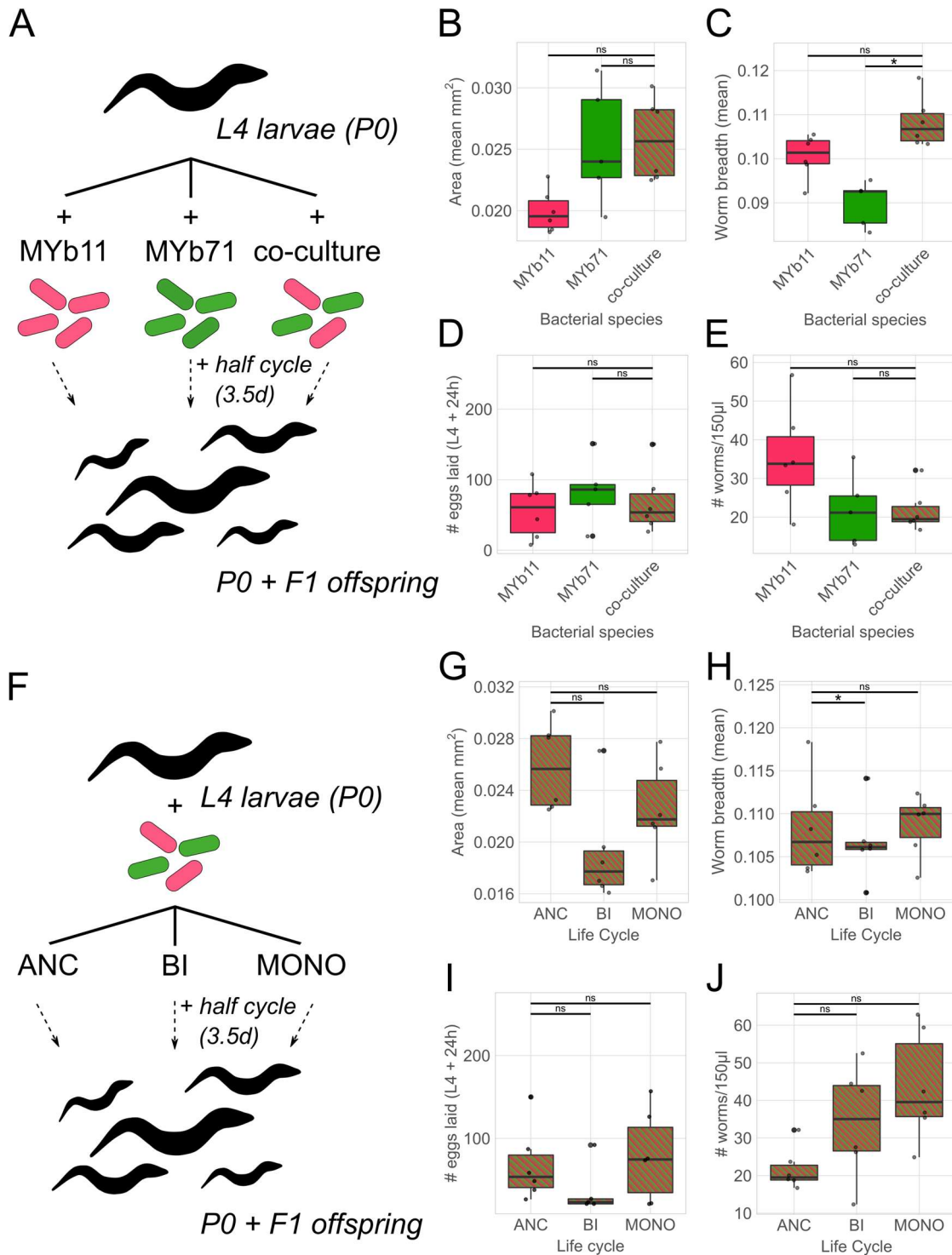
Expanding the scope from ecological interactions of the two species to the impact of evolved life histories, we imposed either a biphasic life cycle with worm hosts or a purely free-living life cycle on agar only. Here we aimed to understand whether microbiota composition responds to the evolved life history. Especially in the biphasic life cycle shared with *C. elegans* as a host, species proportions might further reveal whether the distinct strengths of the two species might stabilize the community or lead to dominance of either species. After the ten cycles of experimental evolution (scaffolded along ten worm generations (Figure 3A), The evolved co-cultures were compared in species composition along the biphasic symbiotic life cycle.

Across the stages within the life cycle, the proportions within the two species microbiota varied with MYb11 clearly establishing dominance on agar and sometimes even in hosts, depending on the evolution treatment. Starting with an even species ratio, co-cultures on agar were dominated by MYb11 within the first 24 hours of growth in the ancestral co-culture (Figure 3B). Across evolved co-cultures, MYb11 dominated co-cultures on agar across sampled time points, including the time of inoculation (Figure 3B). This suggests that the dominance of this species at the end of the first evolutionary cycle was maintained and carried over to the next cycles of the evolution experiment.

In worms on the other hand, microbiota composition only shifted in certain stages of the association with the worm. Species proportions in worms during establish colonization and short-term persistence differed depending on the evolved life cycle. Here, the ancestral communities were dominated by MYb71, while biphasically evolved populations became dominated by MYb11 (Figure 3B, Table S4). Monophasically evolved co-cultures showed similarity to both ancestral and biphasic co-cultures, indicating a more subtle increase in MYb11 proportions (Figure 3B, Table S4). Similar trends were observed for species proportions during release, however, they could not be confirmed statistically due to limited sample sizes (see Figure S3). During early colonization of *C. elegans*, as well as long-term persistence in worms, MYb11 dominated co-cultures across evolutionary treatments (Figure 3B, Table S4). Here the habitat presented a strong filter favoring MYb11, as the population could not be distinguished from the ancestral state (Figure 3B, Table S4). Thus, microbiota composition on agar as well as inside worms can respond to the evolved life cycle in a habitat and life stage sensitive-manner.

### *Effects of species sorting on the host*

Finally, we aimed to understand the influence MYb11 and MYb71 on the development and fitness of *C. elegans*, and to what extent their co-existence or life history affects this. Comparing worm populations grown on mono-lawns of MYb11 and MYb71 or a co-culture until reaching the F1 generation after 3.5 days, we generally found the effect of the co-culture to be similar to that of the component species (Figure 4A-E). We did not observe statistical differences in worm size in terms of worm area (Figure 4B; Dunnett<sub>MYb11:co-culture</sub>,  $t(5)=-0.693$ ,  $p=0.720$ , Dunnett<sub>MYb71:co-culture</sub>,  $t(4)=1.282$ ,  $p=0.368$ ; Table S5), the number of eggs laid at 24 hours (Figure 4D; Dunnett<sub>MYb11:co-culture</sub>,  $t(5)=-0.143$ ,  $p=0.986$ , Dunnett<sub>MYb71:co-culture</sub>,  $t(4)=1.121$ ,  $p=0.459$ ; Table S5) or the population size of reached after 3.5 days (Figure 4E; Dunnett<sub>MYb11:co-culture</sub>,  $t(5)=0.728$ ,  $p=0.695$ , Dunnett<sub>MYb71:co-culture</sub>,  $t(4)=-0.233$ ,  $p=0.961$ ; Table S5) between worm developing in mono-associations or the co-cultures. Still, worm area and population size reached on the co-culture resembled that of MYb71 more than that of MYb11 (Figure 4A, E). The breadth of worms, however, was lower in worms grown on MYb71 than on the co-culture (Figure 4C; Dunnett,  $t(3)=-3.340$ ,  $p=0.012$ ; Table S5). This suggests that worms on MYb71 were slenderer than those in co-culture, and that MYb11 thus had a dominant effect over MYb71 in shaping worm breadth. From this, we take that the two species have largely similar effects on *C. elegans* MY316, with the exception of worm breadth mostly shaped by MYb11.



**Figure 4** Indirect effects of microbiota composition and bacterial life cycle on *C. elegans* hosts. **(A)** In a worm population growth assay 5 L4 stage *C. elegans* MY316 were grown monocultures of *Pseudomonas lurida* MYb11, *Ochrobactrum vermis*. MYb71, or a co-culture of both. After 24h, the number of eggs laid were counted and after 3.5 days worm populations sizes (P0 + F1 offspring) and the dimensions of worms therein were quantified. As a measure of worm development, **(B)** worm area (major\*minor axis length) and **(C)** worm breadth (minor/major axis length) were calculated, while **(D)** eggs laid at 24 h and **(E)** worm population sizes after 3.5 days are indicative of host fitness. **(F)** To compare the impact of bacterial life histories, a worm population growth assay was performed with ancestral and biphasically and monophasically evolved co-cultures of MYb11 and MYb71. Again, **(G)** worm area, **(H)** worm breadth, **(I)** eggs laid at 24h post exposure and **(J)** worm population size were measured. Replicate worm populations are shown as individual data points. Differences between treatments (species or evolved life cycles) were tested using ANOVAs with Dunnett post hoc comparisons of mono-associations with the co-culture or ancestral with evolved co-cultures respectively (\*\*\*) ( $p < 0.001$ ), \*\* ( $p < 0.01$ ), \* ( $p < 0.05$ )).

Next, we tested whether evolution of a biphasic life history by the microbiota would change the impact on *C. elegans* hosts (Figure 4F). Knowing that co-cultures passaged in a biphasic symbiosis life cycle or monophasically on agar differed in their species composition from ancestral co-cultures (Figure 3), we expected worm phenotypes to respond accordingly. In line with the similarity in worm area, eggs laid at 24h and worm population sizes on MYb11 and MYb71, these phenotypes did not show significant responses to the evolved life cycle (Figure 4G, I, J; Table S6). Careful inspection of the data, however, indicated a slight tendency of worms on biphasically evolved co-cultures to be smaller (Figure 4G) and to have laid fewer eggs than on ancestral co-cultures (Figure 4I), despite reaching slightly greater population sizes on evolved bacteria (Figure 4J). The majority of monophasically evolved replicate co-cultures, too, yielded greater worm population sizes (Figure 4J) with worms more similar to those on ancestral co-cultures (Figure 4G-I). Worm broadness, however, was significantly decreased on biphasically evolved bacteria in comparison to ancestral co-cultures (Dunnett,  $t(5)=2.313$ ,  $p=0.048$ ; Table S6). Given the increase in MYb11 in the microbiota in the biphasic life cycle (Figure 3B), which was the species prompting less broad worms (Figure 4C), this cannot be due to species sorting only, but must, at least partially, have an evolutionary basis. Taken together, we find that the life history background of the microbiota does, very mildly, affect worm development and possibly fitness.

## Discussion

Biotic interactions are key to the importance and functioning of host-associated microbial communities. We here studied a simplified two-member community of *Pseudomonas lurida* MYb11 and *Ochrobactrum vermis* MYb71 in their natural host *C. elegans*. Tracking bacterial fitness, in terms of abundance, along the symbiotic life cycle, from initiation and establishment of host colonization after growth in the environment, persistence and release back into the environment provides insight into co-existing life history strategies within microbiotas. We found that the two species have distinct life history strategies while living in the same host. In the following, we discuss the shift in interspecies interactions of MYb11 and MYb71 along the phases of the symbiotic life cycle and contextualize the impact of the two-member community on *C. elegans* hosts.

It has been proposed that, in theory, stable microbiotas are characterized by antagonistic (here: competitive) interactions (Coyte et al., 2015). The reasoning here being that during a disturbance cooperating species that rely on each other will be lost together, while competing species will iteratively benefit from each others decline and thereby be maintained together. Too strong competition can also lead to species displacement by competitive exclusion (Gause, 1934). In this study, we observed two species interacting antagonistically on agar plates, as previously reported during growth *in vitro* (Zimmermann et al., 2019), while antagonism is reduced inside worm hosts. Similar to our observations here, the two main colonizers of the fresh water polyp *Hydra*, the bacteria *Duganella* and *Curvibacter* only stably co-exist on hosts, while *Duganella* outcompetes *Curvibacter in vitro* (Deines et al., 2020). For this, three alternative explanations emerge: either the theoretical prediction is simply incorrect, MYb11 and MYb71 are not members of a stable microbiota, or these species' competition is context-dependent and niches in the host are sufficiently different to attenuate the antagonism observed during growth *in vitro*. Firstly, the data here is insufficient to reject the theory as such, as we are looking at a two-member community only, and do not simulate disturbances. Further, we can approximate that MYb11 and MYb71 do stably co-associate with worms, as they (i) were co-isolated from a worm (Dirksen et al., 2016), (ii) shown to co-exist in different microbiota configurations in *C. elegans* experimentally (Dirksen et al., 2016, 2020), and (iii) both belong to core taxa found to associate with *C. elegans* naturally (Zhang et al., 2017). There is evidence, however, that different characteristics of MYb11 and MYb71 during the phases of host association may reduce competition by providing different temporal and possibly spatial niches, which may be considered niche differentiation. *Pseudomonas lurida* MYb11 shows comparatively rapid growth and spatial expansion on agar, possibly explaining the high proportion of MYb11 in worms during early colonization. Further,

this bacterium is capable of persisting in worms over extended periods of time, either by evading host digestion and immunity and thereby maintaining bacterial load, or by growing inside worms. These observations link to a the life history history strategy previously attributed to MYb11 based on genomic traits interpreted in the Universal Adaptive Strategy Theory (UAST) framework (Fierer, 2017; Grime, 1977; Zimmermann et al., 2019). Predictions suggest that MYb11 is mainly following a competitive strategy, but also has attributes relating to a ruderal and to a lesser extent a stress-tolerant strategy (Zimmermann et al., 2019). Our observations confirm this. Specifically, the advantage on agar and the capacity to dominate in worms after biphasic evolution illustrate its competitive ability. Further, the rapid initial growth and early colonization match a ruderal strategy successful at colonizing a novel habitat early on in successions. Persistence in worms, especially in the long term highlight attributes of a stress-tolerant strategy. *Ochrobactrum vermis* MYb71, on the contrary, was predicted to have a mixed ruderal and competitive strategy (Zimmermann et al., 2019). Its competitive ability is apparent from both growth and colonization experiments, as this species increases its proportion over time in the co-culture despite MYb11 having a growth advantage with an earlier exponential phase. Whether the accumulation of MYb71 over time, also observed in the domesticated *C. elegans* strain N2 (Dirksen et al., 2020), is due to fewer cells entering and colonizing a worm at any given moment or growth within worms remains open. To experimentally distinguish these scenarios, future investigations may use pulse chase assays, in which MYb11 or MYb71 labeled with different fluorescent proteins can be sequentially exposed to the worms and growth or replacement dynamics in the worm monitored based on fluorescent signals. Ruderal attributes can be confirmed, too, as during early colonization abundances of around  $10^3$  MYb71 CFUs per worm are comparable to those of established MYb11 abundances. In contrast to genomic predictions, our experimental data further indicates that MYb71 can also persist in worms, although to a lesser extent than MYb11, also previously shown in N2 strain *C. elegans* hosts (Dirksen et al., 2016). For a better comparison of genomically predicted and experimentally inferred strategies, the scoring of strategies should thus be more quantitative. For this, comparing a panel of bacterial species with a range of phenotypes should help. Already now, however, the mapping of predicted strategies and observed strategies largely overlaps.

While these different strategies of MYb11 and MYb71 may reduce their temporal overlap and thus the strength of competition, it cannot completely explain the lost signal of competition inside the worm. Alternatively, it is possible that the realized functions inside the worm are sufficiently different from those in the environment, that the two species no longer compete here. Additionally, spatial sorting within the host may attenuate competition (Amarasekare, 2003). Previous work suggests that during mono-colonizations, MYb11 predominantly colonizes and persists in the pharynx, while MYb71 can be found in the intestine (Dirksen et al., 2016). Such specialization on specific within host or even within gastrointestinal tract habitats has also been described in other microbiotas, for example honey bees, *Drosophila* flies or humans (Kwong and Moran, 2016; Pais et al., 2018; The Human Microbiome Project Consortium et al., 2012). Differential patterns of colonization could not only explain the lack of competition when species do not interact directly, but possibly link to the different strategies of colonization matching different within-host challenges. As MYb11 has previously been described to colonize the pharyngeal part of the worm gut, it might be more adapted to the lower pH here and the specific immune environment of the first intestinal segment in *C. elegans* than MYb71, which has rather been localized to the anterior and posterior intestine (Bender et al., 2013; Dierking et al., 2016; Dirksen et al., 2016). This emphasizes the importance of *in vivo* studies of microbiota communities and highlights hosts as potential reservoirs for bacterial diversity.

Microbiota bacteria can differentially shape host development and reproductive behaviour via nutrient provisioning and immune system communication (Grenier and Leulier, 2020; Jose et al., 2019; MacNeil et al., 2013b; Metcalf et al., 2019). Here, we observed largely overlapping impact of MYb11 and MYb71 on *C. elegans* development and reproduction. Both sustain growth, developmental progression and generation of offspring. In contrast to previous studies (Dirksen et al., 2016, 2020), we did not observe faster development or greater population sizes on MYb11 than on MYb71. However, the measures here compound development with the onset of egg laying and study population growth within one, but not across generations as

previously shown (Dirksen et al., 2016, 2020). While the observed differences may stem from differences in experimental design, they could also relate to the genetic background of the *C. elegans* host. Here we pair MYb11 and MYb71 with the worm isolate MY316 from compost (Dirksen et al., 2016), which is thus their natural host, and not the domesticated lab strain *C. elegans* N2. If there are indeed differences in the influence of these two host genetic backgrounds, it would be of great interest to compare their basis. For this, whole genome sequencing of MY316 and inclusion of this strain in panels for genome-wide association studies, such as the CeNDR resource (Cook et al., 2017), could allow identifying host genes involved in responding to these microbiota members. Similarly, the molecular response to of MY316 could be assessed via transcriptomic or proteomic approaches, which have shown N2 respond to MYb71 in metabolism, development and reproduction, as well as immune function (Cassidy et al., 2018; Yang et al., 2019).

Furthermore, we here observed worms growing on MYb71 to be slenderer than those on the co-culture, while those on MYb11 grew equally broad to worms on the two-member community. This might simply be due to different patterns of colonization, or possibly virulent behaviours, leading to specific patterns of gut lumen distension, observed in aging *C. elegans* abundantly colonized by *E. coli* OP50 or under infection with the pathogen *Pseudomonas aeruginosa* (Garigan et al., 2002; Singh and Aballay, 2019). Alternatively, worm breadth could relate to fat storage and therefore indirectly to differential nutrient provisioning or sharing by the two bacterial species (Brooks et al., 2009). Based on their genomes, they should both be capable of synthesizing essential folate, vitamin B12 and tryptophan, for example (Zimmermann et al., 2019). Yet, only MYb11 has the complete pathways to produce thiamine (vitamin B1) and pantothenate (vitamin B5) while MYb71 does not (Zimmermann et al., 2019). These vitamins are taken up from the bacterial diet, and might therefore play a role in determining worm breadth. Instead MYb71 might use some of the nutrients taken up by the worm, thus compete for resources, and thereby lead to slenderer worms. In addition to worm breadth, we observed the subtle tendency to have smaller, but more numerous offspring in worms growing on biphasically evolved, and thus presumably host-adapted co-cultures. This might hint at a bacterially induced shift in a trade-off between offspring quality (size) and quantity in worms (see (Metcalf et al., 2019)). For this, extending the population growth assay to the F2 generation might magnify potential differences between the co-cultures of different life history backgrounds. Lastly, the general similarity in the impact the individual species and the co-culture have on the worm, suggests that bacterial influence on the host is additive. Whether this holds true across different species, even when they have more distinct influences on the worm, and to what extent this is scalable to communities with more than two species, remains open. This fundamental aspect is also of applied value, as it could inform the design of bacterial communities in microbiota transplants, for example when aiming to keep or enrich a protective microbe in the community (Dahan et al., 2020). More fundamentally, the phenotypic plasticity bacteria may provide to or provoke in their hosts may also shape host evolutionary trajectories (Kolodny and Schulenburg, 2020).

In conclusion, this study provides a bottom-up perspective to interspecies interaction within the natural microbiota of *C. elegans* and its implications for worm hosts. Using co-culture experiments of two key species, we show that bacterial strategies can be inferred from host entry, colonization and persistence and differ between co-existing microbiota members. It therefore presents a first initiative and important contribution to unravel the importance of bacterial life history strategies during host association and the implications for the stability of microbe-host systems.

## **Acknowledgements**

We thank Melinda Kemlein, Anna Czerwinski and Shindhuja Joel for experimental support, Kiel BiMo/LMB for access to their core facilities, and the CRC1182 and the Schulenburg group and Johannes Zimmerman for critical discussions and general advice. For funding, we thank the CRC1182 (projects A1.1 and A4.3; NO and HS) and the Max Planck Society (fellowship to HS and IMPRS membership of NO).

## Supplement

**Table S1 Comparing MYb11 and MYb71 growth in monocultures with growth in co-culture using ANOVA followed by Dunnett post-hoc test.**

ANOVA					
	Df	Sum Sq	Mean Sq	F-value	P-value
Strain	2	3.348	1.674	3.495	0.071
Replicate population (pop)	1	0.333	0.333	0.696	0.424
Strain:Pop	2	3.745	1.872	3.908	0.056
Residuals	10	4.791	0.479	NA	NA

### Dunnett post-hoc

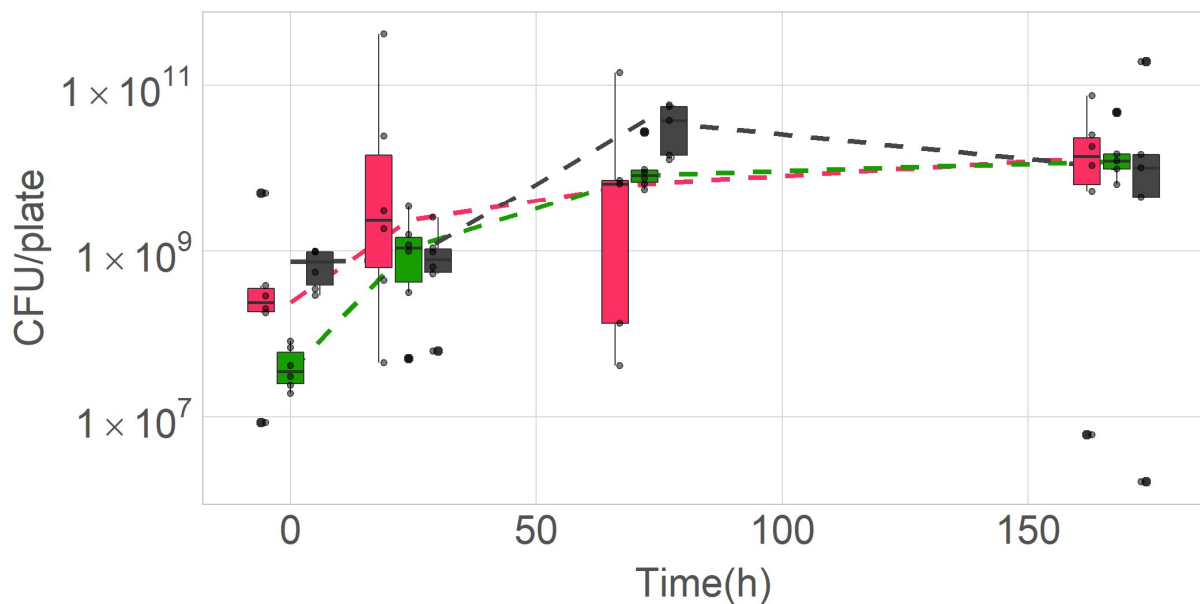
Hypothesis	Estimate	Std. Error	t-value	P-value
MYb11 - co-culture == 0	2.493	0.918	2.716	0.039
MYb71 - co-culture == 0	3.307	0.995	3.323	0.014

**Table S2 Comparing MYb11 and MYb71 colonization of L4 *C. elegans* larvae in monoculture and in co-culture using ANOVA followed by Dunnett post-hoc test.**

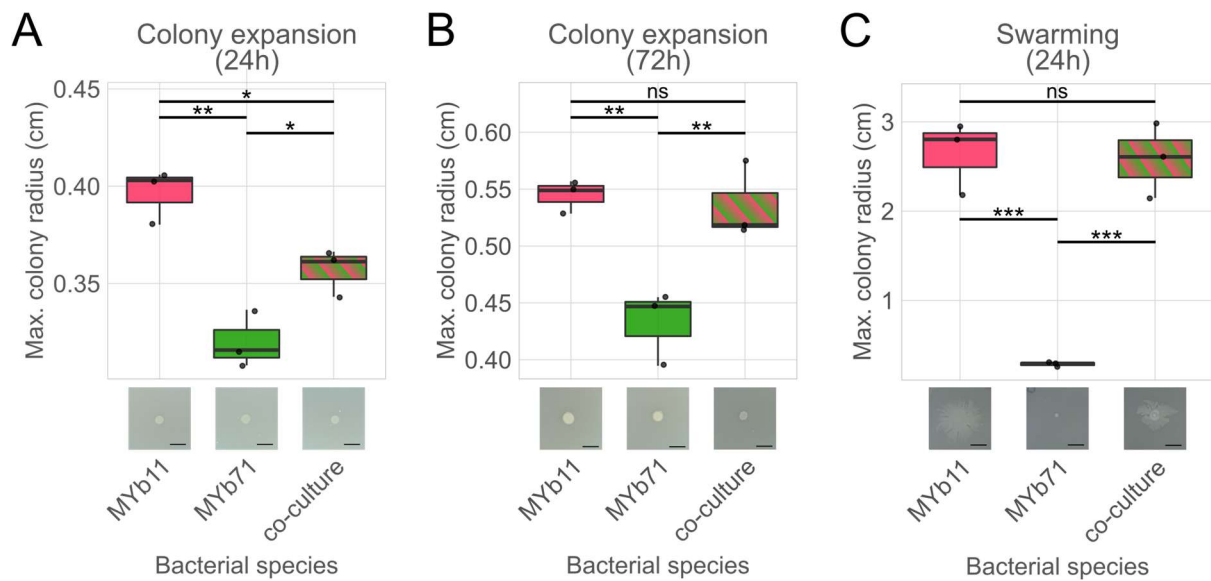
ANOVA					
	Df	Sum Sq	Mean Sq	F-value	P-value
Strain	2	3.843	1.921	1.924	0.189
Replicate population (pop)	1	0.516	0.516	0.517	0.486
Strain:Pop	2	0.261	0.130	0.131	0.879
Residuals	12	11.986	0.999	NA	NA

### Dunnett post-hoc

Hypothesis	Estimate	Std. Error	t-value	P-value
MYb11 - co-culture == 0	0.030	1.316	0.023	1.000
MYb71 - co-culture == 0	0.578	1.316	0.439	0.871



**Figure S1 Growth of *Pseudomonas lurida* (MYb11; pink), *Ochrobactrum vermis* (MYb71; green) and a co-culture of the two (grey) on nematode growth agar over time. Colony forming units (CFUs) collected from bacterial lawns are shown, with replicate populations or co-cultures shown as individual data points. The dashed line connects of median CFUs per species treatment, collected from independent plates, over time.**



**Figure S2 Motility of *Pseudomonas lurida* MYb11, *Ochrobactrum vermis* MYb71 and a co-culture of the two on nematode growth agar. (A) Colony expansion on 3.4% agar was measured as maximum colony radius after 24h of spotting and (B) 72 hours after spotting. (C) Swarming motility was assayed on 0.5% agar and max. colony diameter measured after 24h. Technical replicates are shown as individual data points. Species were compared using ANOVAs followed by Tukey post-hoc tests (\*\*\*) ( $p < 0.001$ ), \*\* ( $p < 0.01$ ), \* ( $p < 0.05$ ). Images of representative replicates are shown on the x-axis.**



**Table S3 Comparing MYb11 and MYb71 motility (colony expansion and swarming) in monocultures and in co-culture using ANOVA followed by Tukey post-hoc tests.**

**Colony expansion (24h)**

<b>ANOVA</b>					
	<b>Df</b>	<b>Sum Sq</b>	<b>Mean Sq</b>	<b>F-value</b>	<b>P-value</b>
Strain	2	0.009	0.004	23.228	0.001
Residuals	6	0.001	0.000	NA	NA

<b>Tukey post-hoc</b>				
<b>Hypothesis</b>	<b>Estimate</b>	<b>Std. Error</b>	<b>t-value</b>	<b>P-value</b>
MYb11 - co-culture == 0	0.039	0.011	3.526	0.029
MYb71 - co-culture == 0	-0.037	0.011	-3.288	0.038
MYb71 - MYb11 == 0	-0.076	0.011	-6.814	0.001

**Colony expansion (72h)**

<b>ANOVA</b>					
	<b>Df</b>	<b>Sum Sq</b>	<b>Mean Sq</b>	<b>F-value</b>	<b>P-value</b>
Strain	2	0.024	0.012	14.643	0.005
Residuals	6	0.005	0.001	NA	NA

<b>Tukey post-hoc</b>				
<b>Hypothesis</b>	<b>Estimate</b>	<b>Std. Error</b>	<b>t-value</b>	<b>P-value</b>
MYb11 - co-culture == 0	0.009	0.023	0.378	0.925
MYb71 - co-culture == 0	-0.104	0.023	-4.486	0.010
MYb71 - MYb11 == 0	-0.113	0.023	-4.864	0.007

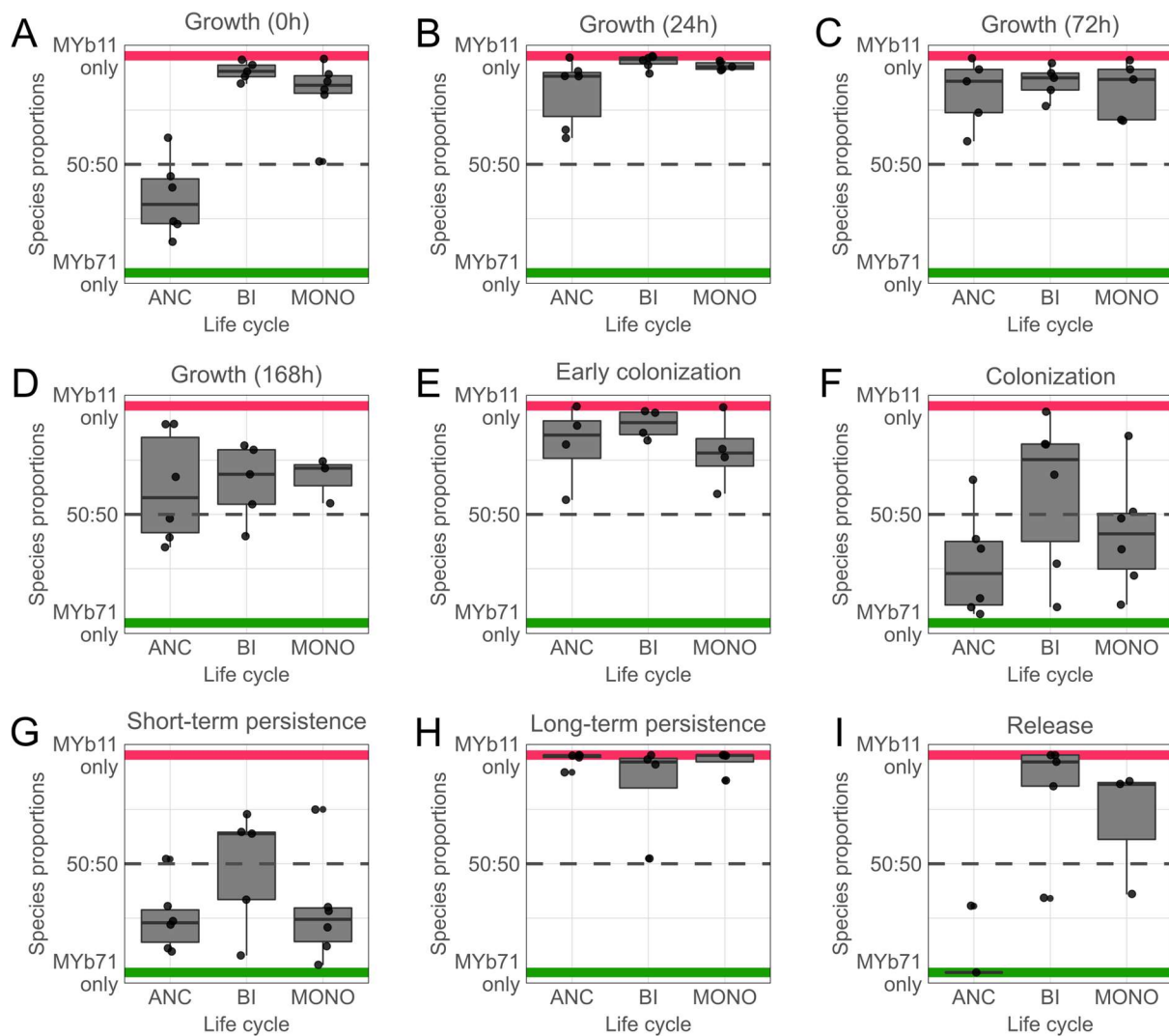
**Swarming (24h)**

<b>ANOVA</b>					
	<b>Df</b>	<b>Sum Sq</b>	<b>Mean Sq</b>	<b>F-value</b>	<b>P-value</b>
Strain	2	10.819	5.409	47.566	0.000
Residuals	6	0.682	0.114	NA	NA

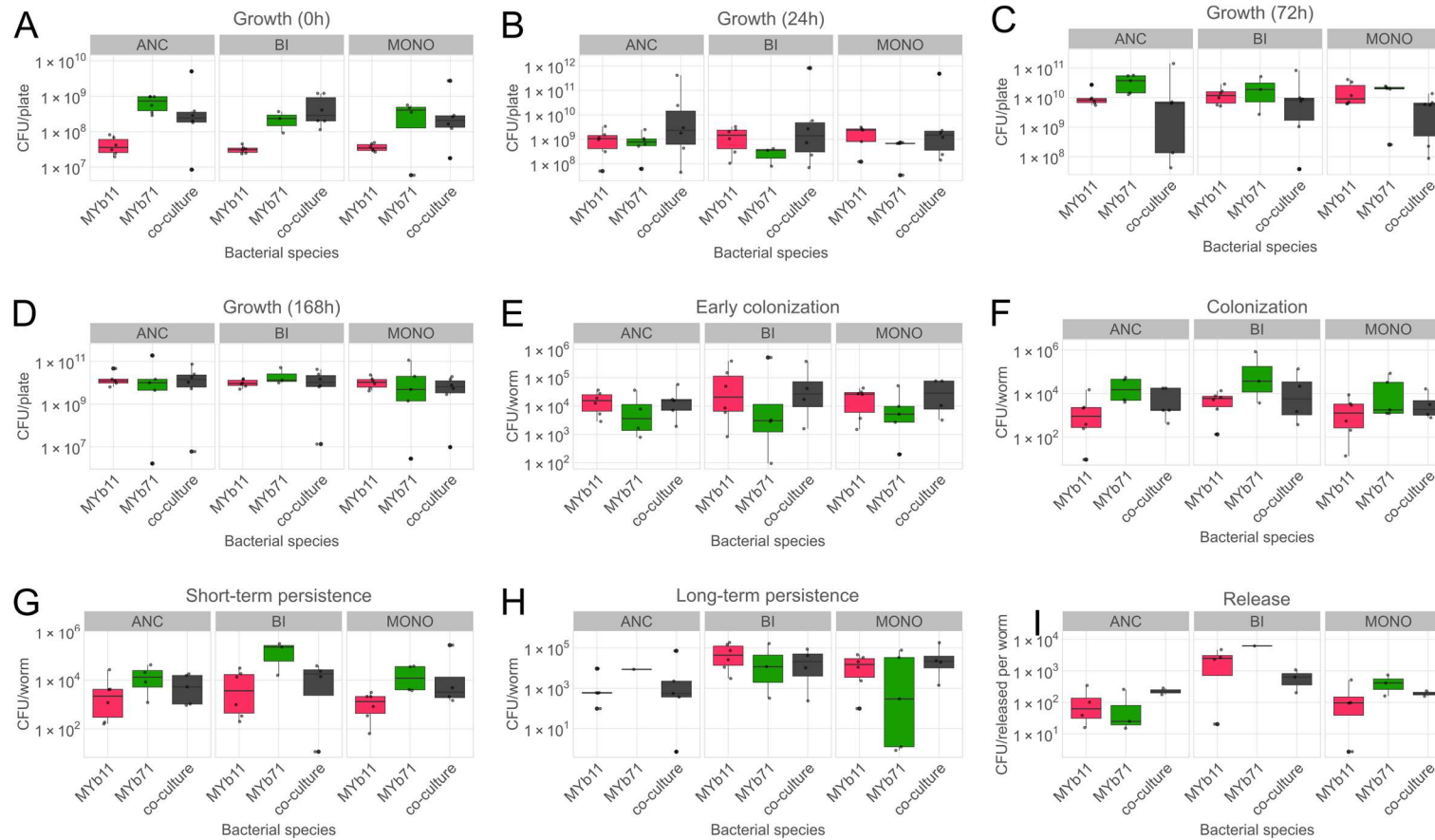
<b>Tukey post-hoc</b>				
<b>Hypothesis</b>	<b>Estimate</b>	<b>Std. Error</b>	<b>t-value</b>	<b>P-value</b>
MYb11 - co-culture == 0	0.063	0.275	0.230	0.971
MYb71 - co-culture == 0	-2.293	0.275	-8.329	<0.001
MYb71 - MYb11 == 0	-2.357	0.275	-8.560	<0.001

**Table S4 Tukey post hoc contrasts of generalized linear model (binomial) comparing the proportions of MYb11 and MYb71 in ancestral and evolved (bi- and monophasically) co-cultures.**

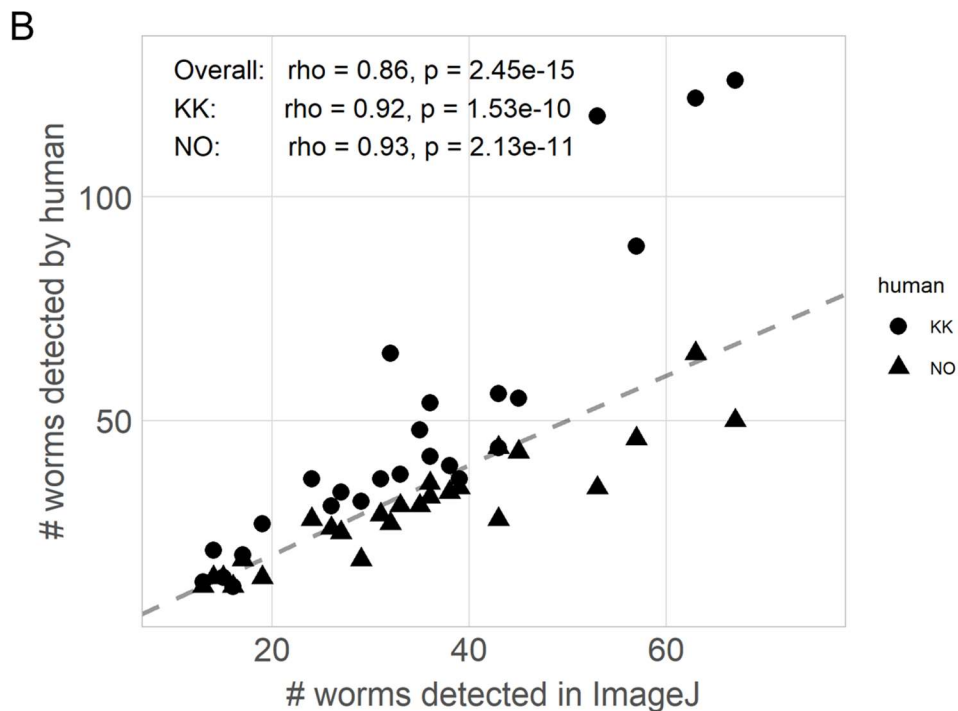
<b>Growth (t0h)</b>	<b>Hypothesis</b>	<b>Estimate</b>	<b>Std. Error</b>	<b>z-value</b>	<b>P-value</b>
	bi - anc	3.193	0.672	4.752	0.000
	mono - anc	2.232	0.482	4.630	0.000
	mono - bi	-0.961	0.710	-1.354	0.360
<b>Growth (24h)</b>	<b>Hypothesis</b>	<b>Estimate</b>	<b>Std. Error</b>	<b>z-value</b>	<b>P-value</b>
	bi - anc	1.964	0.787	2.496	0.033
	mono - anc	1.312	0.689	1.904	0.134
	mono - bi	-0.652	0.947	-0.689	0.766
<b>Growth (72h)</b>	<b>Hypothesis</b>	<b>Estimate</b>	<b>Std. Error</b>	<b>z-value</b>	<b>P-value</b>
	bi - anc	0.379	0.638	0.594	0.823
	mono - anc	0.015	0.615	0.024	1.000
	mono - bi	-0.364	0.655	-0.555	0.844
<b>Growth (168h)</b>	<b>Hypothesis</b>	<b>Estimate</b>	<b>Std. Error</b>	<b>z-value</b>	<b>P-value</b>
	bi - anc	0.106	0.519	0.204	0.977
	mono - anc	0.173	0.611	0.284	0.956
	mono - bi	0.067	0.634	0.106	0.994
<b>Early colonization</b>	<b>Hypothesis</b>	<b>Estimate</b>	<b>Std. Error</b>	<b>z-value</b>	<b>P-value</b>
	bi - anc	2.192	1.742	1.259	0.406
	mono - anc	-0.051	0.902	-0.057	0.998
	mono - bi	-2.244	1.738	-1.291	0.388
<b>Colonization (L4s / established)</b>	<b>Hypothesis</b>	<b>Estimate</b>	<b>Std. Error</b>	<b>z-value</b>	<b>P-value</b>
	bi - anc	1.451	0.556	2.610	0.025
	mono - anc	0.773	0.551	1.403	0.339
	mono - bi	-0.678	0.524	-1.294	0.398
<b>Short-term persistence</b>	<b>Hypothesis</b>	<b>Estimate</b>	<b>Std. Error</b>	<b>z-value</b>	<b>P-value</b>
	bi - anc	1.533	0.504	3.040	0.007
	mono - anc	1.028	0.499	2.059	0.098
	mono - bi	-0.505	0.467	-1.080	0.526
<b>Long-term persistence</b>	<b>Hypothesis</b>	<b>Estimate</b>	<b>Std. Error</b>	<b>z-value</b>	<b>P-value</b>
	bi - anc	-3.139	1.850	-1.697	0.196
	mono - anc	-1.990	2.003	-0.994	0.569
	mono - bi	1.149	1.060	1.083	0.512
<b>Release</b>	<b>Hypothesis</b>	<b>Estimate</b>	<b>Std. Error</b>	<b>z-value</b>	<b>P-value</b>
	bi - anc	3.382	1.843	1.835	0.154
	mono - anc	3.765	1.822	2.066	0.094
	mono - bi	0.383	1.196	0.320	0.944



**Figure S3 Proportions of MYb11 and MYb71 in ancestral and evolved co-culture across the stages of the biphasic life cycle.** Species proportions were quantified at (A) inoculation of bacterial lawns at  $t=0h$ , (B) after growth on nematode growth agar (3.4% agar) after 24h, (C) after 72h, (D) after 168h, (E) during early colonization of *C. elegans* MY316 L4 larvae (after 1.5h exposure), (F) during colonization of L4 larvae (exposed to bacteria since L1 stage), (G) during short-term persistence in L4 larvae kept in M9 buffer for 1h (raised on bacteria), (H) long-term persistence in worms kept on empty NGM agar for 24h post-development on co-culture from L1 to L4, and (I) released from L4 larvae into buffer within 1h (previously raised on bacteria from L1 to L4). Replicate co-cultures are depicted as individual data points.



**Figure S4 Comparison of mono- and co-cultures of MYb11 and MYb71 along the stages of the biphasic life cycle.** Colony forming units quantified at **(A)** inoculation of bacterial lawns at  $t=0h$ , **(B)** after growth on nematode growth agar (3.4% agar) after 24h, **(C)** after 72h, **(D)** after 168h, **(E)** during early colonization of *C. elegans* MY316 L4 larvae (after 1.5h exposure), **(F)** during colonization of L4 larvae (exposed to bacteria since L1 stage), **(G)** during short-term persistence in L4 larvae kept in M9 buffer for 1h (raised on bacteria), **(H)** long-term persistence in worms kept on empty NGM agar for 24h post-development on co-culture from L1 to L4, and **(I)** released from L4 larvae into buffer within 1h (previously raised on bacteria from L1 to L4). Replicate co-cultures are depicted as individual data points.



**Figure S5 Workflow to count and measure *C. elegans* populations from photographs. (A)** 1) Samples of worm populations were collected in 48-well plates, washed, frozen and thawed to straighten. Photographs of worm samples were taken using a Leica fluorescence dissecting scope. A custom ImageJ macro was used to 2) adjust images to black and white (B/W) and deduct background noise, 3) set a pixel threshold to emphasize high contrast particles (incl. worms) and 4) detect particles and define by ellipses in an automated manner. Ellipse dimensions were taken as a proxy for particle dimensions and filtered given expected worm dimensions. **(B)** Correlation of automatically detected worms using ImageJ and human counts. These counts were correlated with those from two humans (KK and NO) who independently counted the number of worms present in a photograph. Counts from 25 different images were correlated using Spearman's rho.

**Table S5 Comparing worm development and fitness on monocultures of MYb11 and MYb71 with a co-culture of the two species using ANOVA followed by Dunnett post-hoc tests.**

**Worm area**

<b>ANOVA</b>	<b>Df</b>	<b>Sum Sq</b>	<b>Mean Sq</b>	<b>F-value</b>	<b>P-value</b>
Strain	2	0.000	0.000	5.076	0.027
Replicate population (pop)	1	0.000	0.000	0.078	0.785
Strain:Pop	2	0.000	0.000	1.221	0.332
Residuals	11	0.000	0.000	NA	NA

**Tukey post-hoc**

<b>Hypothesis</b>	<b>Estimate</b>	<b>Std. Error</b>	<b>t-value</b>	<b>P-value</b>
MYb11 - co-culture == 0	-0.003	0.005	-0.693	0.720
MYb71 - co-culture == 0	0.006	0.005	1.282	0.368
MYb71 - MYb11 == 0	-0.003	0.005	-0.693	0.720

**Worm breadth**

<b>ANOVA</b>	<b>Df</b>	<b>Sum Sq</b>	<b>Mean Sq</b>	<b>F-value</b>	<b>P-value</b>
Strain	0.001	0.000	15.432	0.001	0.001
Replicate population (pop)	0.000	0.000	0.700	0.421	0.000
Strain:Pop	0.000	0.000	0.663	0.535	0.000
Residuals	0.000	0.000	NA	NA	0.000

**Tukey post-hoc**

<b>Hypothesis</b>	<b>Estimate</b>	<b>Std. Error</b>	<b>t-value</b>	<b>P-value</b>
MYb11 - co-culture == 0	-0.014	0.007	-2.003	0.123
MYb71 - co-culture == 0	-0.025	0.007	-3.340	0.012
MYb71 - MYb11 == 0	-0.014	0.007	-2.003	0.123

**Eggs laid at 24h**

<b>ANOVA</b>	<b>Df</b>	<b>Sum Sq</b>	<b>Mean Sq</b>	<b>F-value</b>	<b>P-value</b>
Strain	2	1959.927	979.963	0.458	0.644
Replicate population (pop)	1	957.653	957.653	0.448	0.517
Strain:Pop	2	2389.811	1194.905	0.559	0.587
Residuals	11	23524.874	2138.625	NA	NA

**Tukey post-hoc**

<b>Hypothesis</b>	<b>Estimate</b>	<b>Std. Error</b>	<b>t-value</b>	<b>P-value</b>
MYb11 - co-culture == 0	-9.810	68.440	-0.143	0.986
MYb71 - co-culture == 0	69.930	62.390	1.121	0.459
MYb71 - MYb11 == 0	-9.810	68.440	-0.143	0.986

**Worm population size (L4 + 3.5 d)**

<b>ANOVA</b>	<b>Df</b>	<b>Sum Sq</b>	<b>Mean Sq</b>	<b>F-value</b>	<b>P-value</b>
Strain	2	716.262	358.131	2.906	0.097
Replicate population (pop)	1	28.041	28.041	0.227	0.643
Strain:Pop	2	10.563	5.281	0.043	0.958
Residuals	11	1355.843	123.258	NA	NA

**Tukey post-hoc**

<b>Hypothesis</b>	<b>Estimate</b>	<b>Std. Error</b>	<b>t-value</b>	<b>P-value</b>
MYb11 - co-culture == 0	10.639	14.617	0.728	0.698
MYb71 - co-culture == 0	-3.496	14.978	-0.233	0.961
MYb71 - MYb11 == 0	10.639	14.617	0.728	0.698

**Table S6 Comparing worm development and fitness on co-cultures of MYb11 and MYb71 with different life histories (ancestral, biphasically or monophasically evolved) using ANOVA followed by Dunnett post-hoc tests.**

**Worm area**

<b>ANOVA</b>	<b>Df</b>	<b>Sum Sq</b>	<b>Mean Sq</b>	<b>F-value</b>	<b>P-value</b>
Evo. life cycle	2	0.000	0.000	6.282	0.004
Replicate population (pop)	1	0.000	0.000	0.441	0.510
Strain:Pop	2	0.000	0.000	0.086	0.918
Residuals	43	0.001	0.000	NA	NA

**Tukey post-hoc**

<b>Hypothesis</b>	<b>Estimate</b>	<b>Std. Error</b>	<b>t-value</b>	<b>P-value</b>
mono - anc == 0	-0.002	0.003	-0.570	0.797
bi - anc == 0	-0.004	0.003	-1.334	0.320

**Worm breadth**

<b>ANOVA</b>	<b>Df</b>	<b>Sum Sq</b>	<b>Mean Sq</b>	<b>F-value</b>	<b>P-value</b>
Evo. life cycle	2	0.000	0.000	3.013	0.060
Replicate population (pop)	1	0.000	0.000	0.029	0.866
Strain:Pop	2	0.000	0.000	1.971	0.152
Residuals	43	0.003	0.000	NA	NA

**Tukey post-hoc**

<b>Hypothesis</b>	<b>Estimate</b>	<b>Std. Error</b>	<b>t-value</b>	<b>P-value</b>
mono - anc == 0	0.002	0.007	0.242	0.959
bi - anc == 0	0.016	0.007	2.313	0.048

**Eggs laid at 24h**

<b>ANOVA</b>	<b>Df</b>	<b>Sum Sq</b>	<b>Mean Sq</b>	<b>F-value</b>	<b>P-value</b>
Evo. life cycle	2	5551.930	2775.965	0.930	0.403
Replicate population (pop)	1	1731.556	1731.556	0.580	0.451
Strain:Pop	2	1366.628	683.314	0.229	0.796
Residuals	41	122372.655	2984.699	NA	NA

**Tukey post-hoc**

<b>Hypothesis</b>	<b>Estimate</b>	<b>Std. Error</b>	<b>t-value</b>	<b>P-value</b>
mono - anc == 0	21.250	44.540	0.477	0.851
bi - anc == 0	-35.250	47.010	-0.750	0.679

**Worm population size (L4 + 3.5 d)**

<b>ANOVA</b>	<b>Df</b>	<b>Sum Sq</b>	<b>Mean Sq</b>	<b>F-value</b>	<b>P-value</b>
Evo. life cycle	2	1411.862	705.931	4.005	0.025
Replicate population (pop)	1	155.273	155.273	0.881	0.353
Strain:Pop	2	69.647	34.824	0.198	0.821
Residuals	43	7578.654	176.248	NA	NA

**Tukey post-hoc**

<b>Hypothesis</b>	<b>Estimate</b>	<b>Std. Error</b>	<b>t-value</b>	<b>P-value</b>
mono - anc == 0	16.757	10.411	1.609	0.201
bi - anc == 0	7.731	11.002	0.703	0.711





**Observing bacterial strategies and  
manipulating host association**

Nancy Obeng



Many multicellular organisms provide a habitat and interaction partner for microbes. The associated microbes, the so-called microbiota, may consist of single taxa or harbor complex microbial communities (Fisher et al., 2017; McFall-Ngai et al., 2013). These microbe-host associations are an ancient phenomenon common to many eukaryotic hosts (McFall-Ngai et al., 2013; Pascoe et al., 2017; The Human Microbiome Project Consortium et al., 2012; Zhang et al., 2017). Yet what drives the emergence and maintenance of these associations and how living with a host shapes bacterial traits and life history, remains underexplored (Koch and McFall-Ngai, 2018). In this thesis I have studied bacteria naturally associated with the nematode *Caenorhabditis elegans* to understand the functional repertoires of host-associated bacteria and the evolutionary forces shaping their establishment and maintenance in the host. To this end empirical insights from a combination of phenotypic characterization, genomic analysis, and evolution experiments were compared with predictions from mechanistic constraint-based metabolic models as well as more phenomenological matrix population models. In summary, the main findings emerging are:

- (i) The natural microbiota of *C. elegans* includes bacteria with distinct functional repertoires that together may synthesize all essential vitamins and amino acids for the worm. Their interaction depends on the context, including the nutrient context (**chapter 1**) and the habitat, i.e. the worm or the free-living agar environment (**chapter 5**).
- (ii) Bacterial isolates from the natural microbiota of *C. elegans* vary in their life history strategies predicted in the context of the Universal Adaptive Strategy Theory (UAST) framework (**chapter 1**). Exemplified by the two microbiota isolates *Pseudomonas lurida* MYb11 and *Ochrobactrum vermis* MYb71, the stress-tolerant fares best in the host (**chapter 1**).
- (iii) In theory, evolving host association in a biphasic life cycle, bacteria should ultimately invest either into greater replication or migration rates (**chapter 3**).
- (iv) Rather than generally increasing its replication, the microbiota isolate MYb11 increases its stress-tolerance via improved biofilm formation when experimentally evolving with *C. elegans* as a host in a biphasic life cycle, (**chapter 4+5**). This provides a competitive advantage in co-culture with MYb71 in the host and the free-living environment (**chapter 5**).
- (v) Attachment to the host via biofilm formation presents a key mechanism allowing for extended contact between symbiont and host in the early stages of association. In *C. elegans*, the microbiota isolate *Pseudomonas lurida* MYb11 achieves this via regulation with the second messenger cyclic di-GMP (**chapter 4**).
- (vi) Composition of a two-member microbiota composed of MYb11 and MYb71 shifts under experimental evolution of a biphasic life cycle. These changes indirectly affect the host phenotype, as they have a mild impact on worm development.
- (vii) Periodic interaction with the host within a biphasic life cycle is thus sufficient to select for improved host association and can thus be considered an evolutionary route for free-living bacteria towards symbiosis (**chapters 2-4**).

From this I conclude that key aspects of bacterial association with *C. elegans* as a host are firstly species-specific, resulting in distinct life history strategies co-existing in the natural microbiota, and secondly include simply sticking to the host to establish and maintain contact.

Overall, this project provides a bacterial perspective on the evolution of microbe-host interactions, taking into account a biphasic symbiosis life cycle. It uses the natural microbiota of *C. elegans* as a model, thereby allowing us to specifically associate bacteria and study the consequences for both the symbiont and host. Conceptually, it advances microbiota research by adding the relevant dimension of the free-living environment in the biphasic symbiosis life cycle. Further, it provides an example of how to infer and compare life history strategies both from genomic data and metabolic reconstructions, but also (co-)culture experiments with the host. Methodologically, the combination of modeling and empirical approaches, has yielded proximate mechanisms of bacterial colonization and persistence and ultimate insights into life history strategies for host association. Together, this thesis therefore contributes to eco-evolutionary study of microbe-host associations.

In the following sections, I will step back and frame these insights considering the current literature and outline emerging research directions. I will focus on three aspects: (i) biofilm formation in beneficial microbe-host associations, (ii) bacterial life history strategies in the microbiota and (iii) the transition from prediction to observation and finally manipulation of bacteria in host association.

## Getting stuck – biofilms in beneficial microbe-host associations

To establish contact with the host and allow for persistent colonization, bacteria need to stay inside the host. Within this thesis, I showed that attachment to the host via biofilm formation can be a first step to engage in symbiosis. Studying *Pseudomonas lurida* MYb11 adapt to *C. elegans* in a biphasic life cycle experimentally, I found that wrinkly bacterial types, with improved ability to form biofilms and greater persistence in the worms, evolved in host association (**chapter 4**).

Biofilms are bacterial aggregates embedded in an extracellular matrix that allows cells to adhere to each other (Geesey et al., 1977; Vert et al., 2012). Commonly, it is a sessile form of bacterial life that may take a variety of shapes. The matrix surrounding the bacterial cells is made up of different extracellular polysaccharides, structural proteins, and extracellular DNA (Flemming et al., 2016). These not only allow bacteria to attach to each other, but also to different substrates, such as host surfaces (Flemming and Wuertz, 2019). Furthermore, the biofilm matrix presents a protective structure buffering for example changes in pH or to keep homeostasis as a protective barrier against abiotic, digestive, and immune attack (Flemming et al., 2016). The close contact between cells also facilitates interaction and communication, and thereby a multicellular form of bacterial life (Flemming and Wuertz, 2019).

In the context of host association, biofilms have mostly been studied in pathogenic bacteria and have received less attention in beneficial associations. Pathogens may use biofilms to adhere to hosts, especially in chronic infections. This has been well-documented in Gram-negative pathogens such as *Salmonella enterica* (Anriany et al., 2001), *Pseudomonas aeruginosa* (Starkey et al., 2009) or different *Escherichia coli* isolates (Bokranz et al., 2005), in which also wrinkly colony types have been documented. These can form biofilms in different within-host habitats from burn wounds of the skin, to lung tissue and the gastrointestinal tract (Hall-Stoodley et al., 2004). Genetically encoded biofilm formation can be upregulated during extended contact to the host and can be linked to pathogen virulence (Hall-Stoodley et al., 2004; Kordes et al., 2019; Moreillon et al., 1995; Sullam et al., 1996). In *C. elegans*, biofilms of *P. aeruginosa* have also been shown to entrap worms (Chan et al., 2020) or those of *Yersinia pestis* block the worms' mouth (Darby et al., 2002), thereby exerting their pathogenic effect. For beneficial associations much less is known about biofilms. So far, insights have come from genomic predictions and experimental observations in only a few host organisms. This includes the identification of genes for biofilm formation in *Lactobacillus* and *Acetobacter* associated with *Drosophila* (Martino et al., 2016; Newell et al., 2014; Winans

et al., 2017), *Gilliamella apicola* and *Snodgrassella alvi* in honey bees (Kwong et al., 2014), *Burkholderia* associated with the orchid rhizosphere and *Vibrio fischerii*, which aggregate on their way to the squid light organ (Bongrand et al., 2020; Visick, 2009). Inside hosts, biofilms of *Lactobacillus* have been observed in the anterior parts of murine and avian guts (Lebeer et al., 2011) and shown to improve colonization of *Photorhabdus* in the entomopathogenic nematode *Heterhabditis* (An and Grewal, 2011) as well as *Burkholderia* colonization of the bean bug midgut (Kim et al., 2014). In *C. elegans*, biofilms of *Bacillus subtilis* have been implicated in protection against pathogens and accumulation of alpha-synuclein aggregates (Goya et al., 2020; Smolentseva et al., 2017). Two main sampling biases could explain the lack of studies on beneficial biofilms. Firstly, biofilms are likely not localizing to non-mucosal surfaces, which have gained a lot of attention in microbiota research. It has been argued that the rapid turnover of mucous in the human colon exceeds that of biofilm formation, and thus makes biofilm formation here unlikely (de Vos, 2015). This matches the observation of biofilms localizing in the anterior sections of gastrointestinal tracts in bees, flies, chickens and mice (Kwong and Moran, 2016; Lebeer et al., 2011; Pais et al., 2018). The *C. elegans* gut with its low pH and limited space for mucous might therefore be closer to such a habitat, and therefore select for biofilm formation in *P. lurida* MYb11 (**chapter 4**; (Bender et al., 2013; Dimov and Maduro, 2019)). Secondly, the use of fecal samples to study microbes in host association prevents understanding the persistent fractions of bacteria, which remain as biofilms in the host (discussed in **chapter 2**).

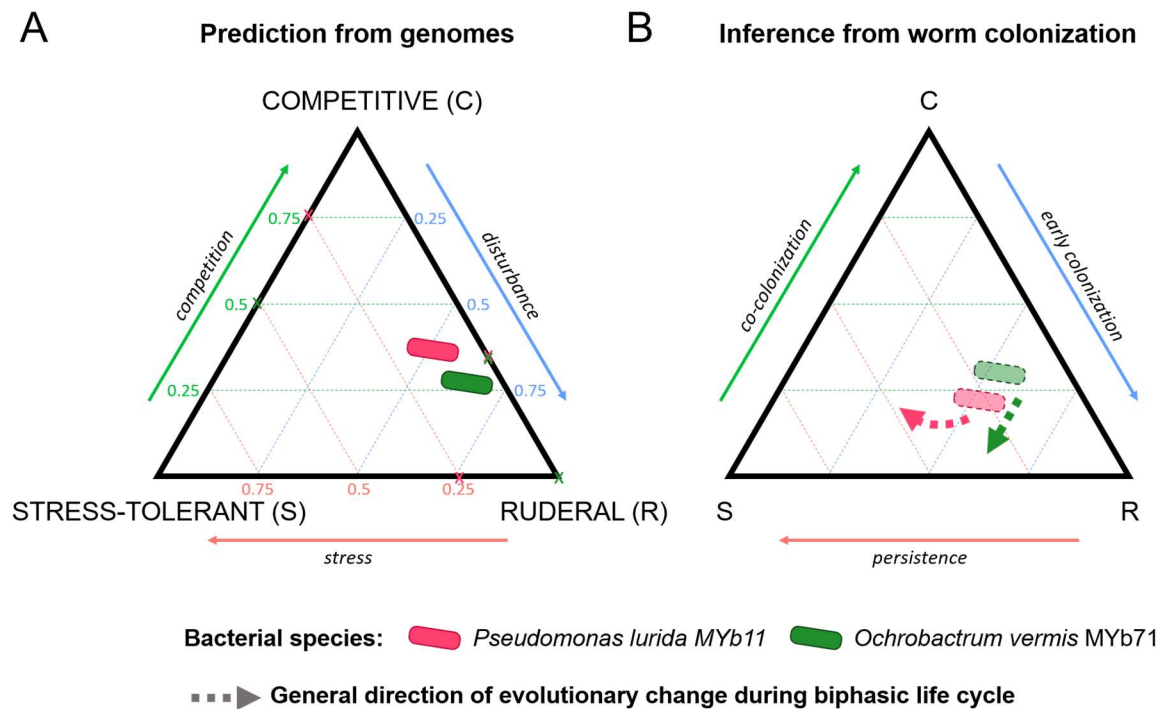
Importantly, this leaves open the following questions: where (in the gastrointestinal tract) do biofilms localize in (*C. elegans*) hosts? What are the causes and consequences of bacterial aggregation inside the host? And what is the origin of biofilm formation in beneficial microbe-host associations? The last question might link to the evolutionary routes towards host association. Both the evolution of mutualists from pathogens (Sachs et al., 2011; Wein et al., 2019), as well as the transition of environmental bacteria, which may need to form biofilms to persist in alternative habitats such as streams (Geesey et al., 1977; Sachs et al., 2011), would explain a pre-adaptations to biofilm formation. The questions on the localization of and potential mechanisms for biofilm formation can be addressed in *C. elegans* as a model host in future research. The transparency of the worm is ideal to track both the spatial localization and assembly of biofilms and their components, as well as their regulation using *in vivo* fluorescence microscopy. Such *in vivo* studies can then be used as a general tool to understand the progression of symbiont establishment. Future research should further consider biofilms across host systems and aim to understand the functional consequences for both partners in the microbe-host association.

## Strategic adaptation to life with the host

In a life cycle with a host, bacteria may follow or evolve life history strategies suitable for the association. Using the UAST framework, bacterial life history strategies in the microbiota can be classified into competitive, stress-tolerant and ruderal (Fierer, 2017; Grime, 1977). As shown in **chapter 1**, the trait sets inferred from the genomes of 77 bacterial isolates from the natural microbiota of *C. elegans* yielded a variety of different adaptive strategies. Similar to Grime's original application in plants, some isolates showed specialization in a single of these three primary strategies, while most appeared to follow intermediate strategies (**chapter 1**; (Grime, 1974)). Such a mix of strategies may well be attributed to the complexity of the host and the biphasic life cycle experienced by the microbiota. As each stage in the life cycle contributes to survival and reproductive, a bacterial strategy should thus integrate fitness along the life cycle and its respective habitats (as argued in **chapter 3**).

Certainly, the basis for the different predicted life history strategies are the distinct yet overlapping functional repertoires of bacteria in the natural microbiota of *C. elegans*. This confirms two attributes commonly observed for microbiotas and organismal communities in

general: functional redundancy and niche differentiation (Moya and Ferrer, 2016). The functional redundancy in part stems from basic cell functions but extends to the synthesis of amino acids and vitamins (**chapter 1**). In the context of the microbiota, this results in an insurance effect allowing worms to maintain functions despite disturbances that affect species abundances or composition (Moya and Ferrer, 2016). Similarly, functional interactions, such as cross-feeding of MYb11 and MYb71 on sucrose (chapter 1), may be maintained in the community if multiple species share the same pathways. Niche differentiation, manifesting from the trait sets and transferring to the adaptive strategies, likely facilitates the co-existence of taxa in the microbiota. This is in part due to the different metabolic profiles, which provide sufficient information to classify distinct niche spaces (see also (Fahimipour and Gross, 2020)). These should allow for co-existence via sub-specialization along the stages of the biphasic life cycle, as well as the different habitats passed therein. Comparable sub-specialization has been noted in two *Photorhabdus* species that differ in the production of secondary metabolites during the worm associated and free-living (inside the insect cadaver) phase of their symbiotic life cycle with EPNs (Maher et al., 2020). Observing co-existence along the life cycle with varying species proportions between MYb11 and MYb71 (in **chapter 5**), confirms that distinct adaptive strategies can help maintain species diversity in the microbiota.



**Figure 1** Inferred bacterial life history strategies in the microbiota of *C. elegans*. The life history strategies of *Pseudomonas lurida* MYb11 and *Ochrobactrum vermis* MYb71 were **(A)** predicted from genomically inferred traits (chapter 1) and **(B)** roughly approximated from experimental observations of worm colonization in a relative approach (chapters 4 and 5). The experimental inferences were based on the genomic predictions as a starting point. They were shifted along the axes where the relative strength in colonization of the species differed from the genomic predictions. Species positioning are shown in transparent colors to highlight the non-quantitative nature of inference from the experimental observations. The contributions of competition/co-colonization, stress-tolerance/persistence in worms and disturbance/early colonization in strategy space are shown as contours. The position of the species was determined using approximate triangular ordination. The general direction of evolutionary change based on phenotypic observations (chapters 4 and 5) is indicated by arrows colored by species.

Practically, the co-culture experiments and measures of bacterial entry, co-colonization, and persistence, provide an empirical approach to infer bacterial life history strategies in the microbiota. Using this small set of traits, the strategies predicted for MYb11

and MYb71 based on genomic data (**chapter 1**) could be roughly confirmed (**chapter 5**) (Figure 1). For this, the ability of the bacterial species to successfully colonize, compete and persist in the worms was inferred from bacterial counts in worms and compared between the two species. While the placement of bacteria into strategies based on these empirical observations is very rough (see Figure 1), this suggests that these three components of association with *C. elegans* can be used analogously to the LHS (specific leaf area, plant height, and seed mass) scheme used to infer strategies in plants (Westoby, 1998). Still, important differences between strategies predicted from genomes and those inferred from empirical observations emerged. While genomic traits suggested that MYb11 and MYb71 should be equally good at dealing with post-disturbance habitats for example, MYb11 colonized empty worms more quickly (**chapters 1 and 5**). This might be due to a number of reasons. Firstly, the set of two bacterial species is too limited to infer adaptive strategies from experimental data in a relative approach. Secondly, the specific bacterial habitat and context likely matter, and traits or their relative importance assumed in *in silico* predictions might not be of key relevance during worm colonization. Thirdly, predicted traits and those realized in association with the worm might differ. Therefore, to validate this approach, entry co-colonization and persistence should be measured for all 77 isolates from the natural microbiota of *C. elegans* for which strategic predictions exist, tested and calibrated. The empirical approach might then serve as a practical tool to quantify bacterial life history strategies in *C. elegans*.

The different life history strategies predicted are expected to lead to varying levels of colonization, and thus fitness, in the host. For this, an important assumption must be met: the bacterial side matters. In the most simple scenario, bacterial colonization is completely stochastic and bacterial identity, traits or strategy do not affect colonization. In this case, all bacteria should colonize worms to the same or a completely random amount in mono-colonization experiments. Further, under such neutral conditions microbiota communities across worms should start to resemble each other over time and even out in species proportions (Sieber et al., 2019; Vega and Gore, 2017). As we can observe species-specific differences in colonization (**chapter 1 and 5**; e.g. (Dirksen et al., 2016, 2020)) and deviations from evenness and neutrality have been described (e.g. Berg et al., 2016; Dirksen et al., 2016, 2020; Sieber et al., 2019), bacterial identity and therefore traits and strategy at least partially affect the success of host association. In **chapter 1** of this thesis, we found that bacterial isolates from the worm microbiota with stress-tolerating and competitive strategies showed greater abundance in worms than those with a ruderal strategy in mono-colonization experiments. As has been argued for the soil microbiota (Malik et al., 2020), a ruderal strategy might be less relevant in worm hosts as bacteria will rarely colonize empty hosts and thus be less adapted to the worm. However, *Ochrobactrum vermis* MYb71 whose primary strategy (based on genomic predictions) is ruderal, is known for its ability to both reach high levels of colonization alone and in a microbiota community (**chapters 1**; (Dirksen et al., 2016, 2020; Sieber et al., 2019)). Limiting the correlation of life history strategies and colonization levels to the primary strategies might be too simplified. The success of a ruderal strategy might also depend on the environmental conditions and the level of stochasticity in worm colonization. When worms are exposed to low densities of bacteria, the rate of colonization events is low compared to that of bacterial growth inside worms, and ruderal strategists that benefit from the empty worm are likely able to establish and dominate (Vega and Gore, 2017). On the contrary, high bacterial densities in the environment and a greater rate of colonization reduce this priority effect and species proportions should be balanced in the absence of non-neutral processes (Vega and Gore, 2017). This might explain why MYb11 started to outcompete MYb71 after multiple cycles of experimental evolution in a lab context (**chapter 5**). Furthermore, not only the in-host replication rate, but also the migration rate towards the host determine the success of colonization (**chapter 3**). Thus, a combination of quick growth and spatial exploration (not included in the ruderal strategy so far) of a new habitat should be traits required for colonization success. In line with a niche-based, deterministic account of host association, a stress-tolerant strategy is best suited for life in the worm. In a context

with possibly low resource levels and predation pressure, stress tolerance may favor maintenance over cellular growth (Hengge, 2020). Under these circumstances, investment into a qualitative rather than quantitative strategy should lead to success. This was confirmed by MYb11 strengthening its stress-tolerance by evolving improved biofilm formation when experimentally evolving in the biphasic life cycle (**chapter 4**; Figure 1B). Further, MYb11 increased in proportion to MYb71 within the biphasically evolved co-culture (**chapter 5**; Figure 1B), suggesting that this strategy can be dominant in this life cycle. This in turn links back to *C. elegans* as a host environment, in which bacteria are subjected to low and variable pH, immune attack and digestion (Bender et al., 2013; Dierking et al., 2016; Dimov and Maduro, 2019). Finally, the lack of competition between the two species in co-colonization as opposed to the antagonism observed on the agar plate (**chapter 5**), indicates that the competitive ability might not be selected for during co-colonization. Thus, highlighting the importance of context for both bacteria-bacteria interactions and life history strategies.

While one-to-one symbioses in a simple model host provide tractable mechanistic insights, we need to scale up to more natural complexity. Firstly, this means comparing life history strategies of a variety of bacteria in a *C. elegans* host. Strategies could be predicted for bacteria non-native to the worm to test the value of stress-tolerance across microbial taxa and origins. Secondly, one should check to what extent adaptive strategies influence co-existence in microbiotas containing more than two species and whether this is equally difficult to predict as community composition based on pairwise interactions (Lopez et al., 2019). Finally, it should be tested to what extent the stress-tolerating strategy is superior in other hosts, too. As a first step, *Drosophila* is a good choice, as here too, bacteria can readily be collected from well-studied host organisms along the stages of host association (Douglas, 2019). Together, these insights can inform our fundamental understanding of bacterial adaptation to hosts and advance microbiota transplants and other manipulative approaches.

## Perspective: From observation to prediction to manipulation

Depending on their functional repertoires, bacteria can influence the physiology of their hosts. Predicting and leveraging such differences can be used to selectively modulate host health and fitness. In the scope of this thesis, I collected evidence that members of the natural microbiota of *C. elegans* vary in their functional repertoires and strategies of host association. This includes differential production of essential vitamins and amino acids required by the worm (**chapter 1**), and slight differences in the impact on worm development and population growth (**chapters 1 + 5**). Furthermore, they differ in their ability to protect worms against bacterial or worm pathogens, with *Pseudomonas lurida* MYb11 protecting worms from infection both directly and indirectly (Dirksen et al., 2016; Kissoyan et al., 2019). In addition, I could show that supplementation of the growth medium with specific carbon sources can be used to manipulate the relative abundances of *Ochrobactrum vermis* MYb71 and *Pseudomonas lurida* MYb11 in co-culture in a directed manner (**chapter 1**). Together, this could be leveraged to specifically enrich beneficial microbes in the host based on the known metabolic profiles.

For the success of targeted dietary modulation of the microbiota, the beneficial effect of a focal bacterium needs to be dose-dependent and stable in the context of the microbiota community. A recent study investigating the protective effects of *Enterococcus faecalis* against *Staphylococcus aureus* infection in *C. elegans* showed that in this case protection could be maintained in the context of a complex microbiota (Dahan et al., 2020). However, the specifics likely depend on the beneficial contribution in questions. In the case of MYb11, protection against the pathogen *Bacillus thuringiensis* occurs via production of the antimicrobial peptide massetolide E (Kissoyan et al., 2019). To enrich a protective effect, massetolide E should be produced constitutively and proportional to MYb11 abundance when in the worm. The same would need to be true for the synthesis of certain vitamins, as for example vitamin B12, which



may be used to modulate worm developmental speed or reproduction (Watson et al., 2014). An appropriate experimental test would be to check for a protective, or developmental effect in worms exposed to microbiota communities with varying proportions of MYb11. For this, the CeMbio, a 12-member microbiota consortium derived from the natural microbiota of *C. elegans*, presents a tractable and well-characterized model system (Dirksen et al., 2020). Already established metabolic models of these bacterial isolates can be used to predict nutrients that can serve as a prebiotic, i.e. nutrients favoring the growth of selected microbiota members. This is work currently ongoing in collaboration of the Schulenburg group with Georgios Marinos and Christoph Kaleta, who are leveraging metabolic models as presented in **chapter 1** to predict candidate prebiotics. This approach can be extended from native microbiota members to the introduction of engineered bacteria to modulate hosts in a specific way (Sun and Gao, 2020). An example would be the enrichment of *Snodgrassella alvi* optimized to protect honey bees against infecting mites (Leonard et al., 2020). Iterative cycles of prediction and empirical validation of dietary modulation will not only provide insight into bacterial functioning in the microbiota but allow moving towards manipulation.

In the long term, a prebiotic can also support sustained or repeated co-existence of host and microbe, and thus allow for symbiosis and possibly coevolution. Depending on whether the probiotic can be found only in the host or also the free-living habitat, this should determine the selection on bacteria to invest rather in migration toward the host or in increased replication rates, respectively (**chapter 3**). If greater association with the host is favored, other bacterial traits will further determine the success of the association. Emerging from **chapters 2 to 5**, the capacities of bacteria to firstly associate with, possibly enter, and certainly persist within the host, are key. This highlights the value of combining three different theoretical frameworks, constraint-based metabolic modeling, matrix population models and UAST framework to investigate key aspects of host association from different perspectives. Proximately, metabolic models reveal specific mechanisms used by individual bacteria and possible interactions between them. They can further be used to fine-tune manipulation of the microbiota using prebiotics (**chapter 1**). Elasticity analysis of matrix population models tracking host-associated and free-living bacteria, provided insight into ultimate strategies leading to greater bacterial fitness in the biphasic life cycle (**chapter 3**). Similarly, inferring life history strategies using the UAST framework gives insights into ultimate strategies for host association, with stress-tolerance emerging as being most relevant in the worm host (**chapters 1, 4 and 5**).

These insights can be applied to introduce, maintain, or re-establish microbiotas. Introducing novel microbiota members or communities can be of value in a medical, agricultural, or industrial context to provide novel functions to a host, e.g. protection against a pathogen or specific catabolism of toxins and other pharmaceuticals in the host (Kissoyan et al., 2019; Leonard et al., 2020; Norvaisas and Cabreiro, 2018; Scott et al., 2017; Wang et al., 2020). Prebiotic supplementation and knowledge on traits and strategies used by the introduced and resident microbes can help maintain new additions to the microbiota stably. This knowledge could be developed further to aid conserving microbial diversity that appears to be lost in the human microbiota along a gradient of industrialization (Sonnenburg and Sonnenburg, 2019). Furthermore, it can help in re-establishing native microbiotas, or transplanting those with desired impact on the host, for better treatment of disturbed microbiotas after by antibiotic treatment for example (Pamer, 2016). Therefore, studying the fundamental aspects of bacterial adaption to host association has great potential to inform applied microbiota modulation.

## Concluding remarks

Taken together, this thesis has investigated bacterial function and evolution in the microbiota of *C. elegans*. Venturing beyond lists of species, the survey of metabolic repertoires has

provided a window into the activities and interactions of bacteria to be found in this tiny worm. What these bacteria can do, informs us also about life history strategies they follow to live and associate with the worm host. Experimentally evolving *Pseudomonas lurida* (MYb11) with the worm, it became obvious that simply sticking to the host is a key first step to ensure extended contact to the worm. In the community of bacteria naturally associated with *C. elegans*, we found bacteria to follow a variety of different strategies to survive and compete in the host, thereby likely allowing them to co-exist. Combining experiment and theory revealed traits and strategies involved in the emergence and maintenance of microbe-host association. Reading my thesis, I thus hope you are convinced that taking the bacterial perspective can enrich our fundamental understanding of microbe-host association and its origin.





## Bibliography

### A

AbuOun, M., Suthers, P.F., Jones, G.I., Carter, B.R., Saunders, M.P., Maranas, C.D., Woodward, M.J., and Anjum, M.F. (2009). Genome Scale Reconstruction of a Salmonella Metabolic Model comparison of similarity and differences with a commensal Escherichia coli strain. *J. Biol. Chem.* *284*, 29480–29488.

Ackermann, M. (2015). A functional perspective on phenotypic heterogeneity in microorganisms. *Nat. Rev. Microbiol.* *13*, 497–508.

Ailloud, F., Didelot, X., Woltemate, S., Pfaffinger, G., Overmann, J., Bader, R.C., Schulz, C., Malfertheiner, P., and Suerbaum, S. (2019). Within-host evolution of Helicobacter pylori shaped by niche-specific adaptation, intragastric migrations and selective sweeps. *Nat. Commun.* *10*, 2273.

Alikhan, N.-F., Petty, N.K., Ben Zakour, N.L., and Beatson, S.A. (2011). BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics* *12*, 402.

Amarasekare, P. (2003). Competitive coexistence in spatially structured environments: a synthesis. *Ecol. Lett.* *6*, 1109–1122.

An, R., and Grewal, P.S. (2011). purL gene expression affects biofilm formation and symbiotic persistence of Photorhabdus temperata in the nematode Heterorhabditis bacteriophora. *Microbiology*, *157*, 2595–2603.

Andow, D.A., Kareiva, P.M., Levin, S.A., and Okubo, A. (1990). Spread of invading organisms. *Landsc. Ecol.* *4*, 177–188.

Andrews, S., Krueger, F., Segonds-Pichon, A., Laura, B., Krueger, C., and Wingett, S. (2018). (Babraham Institute).

Ankrah, N.Y.D., Luan, J., and Douglas, A.E. (2017). Cooperative Metabolism in a Three-Partner Insect-Bacterial Symbiosis Revealed by Metabolic Modeling. *J. Bacteriol.* *199*, e00872-16.

Ankrah, N.Y.D., Wilkes, R.A., Zhang, F.Q., Zhu, D., Kaweesi, T., Aristilde, L., and Douglas, A.E. (2020). Syntrophic splitting of central carbon metabolism in host cells bearing functionally different symbiotic bacteria. *ISME J.* 1–12.

Anriany, Y.A., Weiner, R.M., Johnson, J.A., Rezende, C.E.D., and Joseph, S.W. (2001). Salmonella enterica Serovar Typhimurium DT104 Displays a Rugose Phenotype. *Appl. Environ. Microbiol.* *67*, 4048–4056.

Ansorge, R., Romano, S., Sayavedra, L., Porras, M.Á.G., Kupczok, A., Tegetmeyer, H.E., Dubilier, N., and Petersen, J. (2019). Functional diversity enables multiple symbiont strains to coexist in deep-sea mussels. *Nat. Microbiol.* *4*, 2487–2497.

Arata, Y., Oshima, T., Ikeda, Y., Kimura, H., and Sako, Y. (2020). OP50, a bacterial strain conventionally used as food for laboratory maintenance of C. elegans, is a biofilm formation defective mutant. *MicroPublication Biol.* *2020*.

Arbizu, P.M. (2020). pmartinezarbizu/pairwiseAdonis.

Avery, L., and You, Y.-J. (2012). *C. elegans* feeding. In *The Worm Book*, p.

## B

Babic, I., Fischer-Le Saux, M., Giraud, E., and Boemare, N. (2000). Occurrence of natural dixenic associations between the symbiont *Photobacterium luminescens* and bacteria related to *Ochrobactrum* spp. in tropical entomopathogenic *Heterorhabditis* spp. (Nematoda, Rhabditida). *Microbiology* 146, 709–718.

Baiocchi, T., Lee, G., Choe, D.-H., and Dillman, A.R. (2017). Host seeking parasitic nematodes use specific odors to assess host resources. *Sci. Rep.* 7, 6270.

Baker, L.J., Freed, L.L., Easson, C.G., Lopez, J.V., Fenolio, D., Sutton, T.T., Nyholm, S.V., and Hendry, T.A. (2019). Diverse deep-sea anglerfishes share a genetically reduced luminous symbiont that is acquired from the environment. *ELife* 8, e47606.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., et al. (2012). SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *J. Comput. Biol.* 19, 455–477.

Bansept, F., Schumann-Moor, K., Diard, M., Hardt, W.-D., Slack, E., and Loverdo, C. (2019). Enchained growth and cluster dislocation: A possible mechanism for microbiota homeostasis. *PLOS Comput. Biol.* 15, e1006986.

Bantinaki, E., Kassen, R., Knight, C.G., Robinson, Z., Spiers, A.J., and Rainey, P.B. (2007). Adaptive Divergence in Experimental Populations of *Pseudomonas fluorescens*. III. Mutational Origins of Wrinkly Spreader Diversity. *Genetics* 176, 441–453.

Basile, A., Campanaro, S., Kovalovszki, A., Zampieri, G., Rossi, A., Angelidaki, I., Valle, G., and Treu, L. (2020). Revealing metabolic mechanisms of interaction in the anaerobic digestion microbiome by flux balance analysis. *Metab. Eng.* 62, 138–149.

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *J. Stat. Softw.* 67, 1–48.

Bauer, E., Laczny, C.C., Magnusdottir, S., Wilmes, P., and Thiele, I. (2015). Phenotypic differentiation of gastrointestinal microbes is reflected in their encoded metabolic repertoires. *Microbiome* 3, 55.

Bauer, E., Zimmermann, J., Baldini, F., Thiele, I., and Kaleta, C. (2017). BacArena: Individual-based metabolic modeling of heterogeneous microbes in complex communities. *PLOS Comput. Biol.* 13, e1005544.

Beaumont, H.J.E., Gallie, J., Kost, C., Ferguson, G.C., and Rainey, P.B. (2009). Experimental evolution of bet hedging. *Nature* 462, 90–93.

Bender, A., Woydziak, Z.R., Fu, L., Branden, M., Zhou, Z., Ackley, B.D., and Peterson, B.R. (2013). Novel Acid-Activated Fluorophores Reveal a Dynamic Wave of Protons in the Intestine of *Caenorhabditis elegans*. *ACS Chem. Biol.* 8, 636–642.

Benjamini, Y., and 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.

- Berendsen, R.L., Pieterse, C.M.J., and Bakker, P.A.H.M. (2012). The rhizosphere microbiome and plant health. *Trends Plant Sci.* *17*, 478–486.
- Berg, G., Rybakova, D., Fischer, D., Cernava, T., Vergès, M.-C.C., Charles, T., Chen, X., Cocolin, L., Eversole, K., Corral, G.H., et al. (2020). Microbiome definition re-visited: old concepts and new challenges. *Microbiome* *8*, 103.
- Berg, M., Stenuit, B., Ho, J., Wang, A., Parke, C., Knight, M., Alvarez-Cohen, L., and Shapira, M. (2016). Assembly of the *Caenorhabditis elegans* gut microbiota from diverse soil microbial environments. *ISME J.* *10*, 1998–2009.
- Berg, M., Monnin, D., Cho, J., Nelson, L., Crits-Christoph, A., and Shapira, M. (2019). TGF $\beta$ /BMP immune signaling affects abundance and function of *C. elegans* gut commensals. *Nat. Commun.* *10*, 604.
- Beringer, J.E., Brewin, N., Johnston, A.W.B., Schulman, H.M., and Hopwood, D.A. (1979). The Rhizobium-Legume Symbiosis. *Proc. R. Soc. Lond. B Biol. Sci.* *204*, 219–233.
- Bito, T., Matsunaga, Y., Yabuta, Y., Kawano, T., and Watanabe, F. (2013). Vitamin B12 deficiency in *Caenorhabditis elegans* results in loss of fertility, extended life cycle, and reduced lifespan. *FEBS Open Bio* *3*, 112–117.
- Bokranz, W., Wang, X., Tschäpe, H., and Römling, U. (2005). Expression of cellulose and curli fimbriae by *Escherichia coli* isolated from the gastrointestinal tract. *J. Med. Microbiol.* *54*, 1171–1182.
- Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* *30*, 2114–2120.
- Bongrand, C., Moriano-Gutierrez, S., Arevalo, P., McFall-Ngai, M., Visick, K.L., Polz, M., and Ruby, E.G. (2020). Using Colonization Assays and Comparative Genomics To Discover Symbiosis Behaviors and Factors in *Vibrio fischeri*. *MBio* *11*.
- Bordbar, A., Monk, J.M., King, Z.A., and Palsson, B.O. (2014). Constraint-based models predict metabolic and associated cellular functions. *Nat. Rev. Genet.* *15*, 107–120.
- Bordenstein, S.R., and Theis, K.R. (2015). Host Biology in Light of the Microbiome: Ten Principles of Holobionts and Hologenomes. *PLOS Biol.* *13*, e1002226.
- Bosch, T.C.G., and Miller, D.J. (2016). *The Holobiont Imperative: Perspectives from Early Emerging Animals* (Springer).
- Bosch, T.C.G., Guillemin, K., and McFall-Ngai, M. (2019). Evolutionary “Experiments” in Symbiosis: The Study of Model Animals Provides Insights into the Mechanisms Underlying the Diversity of Host–Microbe Interactions. *BioEssays* *41*, 1800256.
- Botelho, J., and Schulenburg, H. (2020). The Role of Integrative and Conjugative Elements in Antibiotic Resistance Evolution. *Trends Microbiol.*
- Broad Institute (2019). Picard toolkit.
- Brooks, A.W., Kohl, K.D., Brucker, R.M., Opstal, E.J. van, and Bordenstein, S.R. (2016). Phylosymbiosis: Relationships and Functional Effects of Microbial Communities across Host Evolutionary History. *PLOS Biol.* *14*, e2000225.

Brooks, J.F., Gyllborg, M.C., Cronin, D.C., Quillin, S.J., Mallama, C.A., Foxall, R., Whistler, C., Goodman, A.L., and Mandel, M.J. (2014). Global discovery of colonization determinants in the squid symbiont *Vibrio fischeri*. *Proc. Natl. Acad. Sci.* *111*, 17284–17289.

Brooks, K.K., Liang, B., and Watts, J.L. (2009). The Influence of Bacterial Diet on Fat Storage in *C. elegans*. *PLOS ONE* *4*, e7545.

Browne, H.P., Neville, B.A., Forster, S.C., and Lawley, T.D. (2017). Transmission of the gut microbiota: spreading of health. *Nat. Rev. Microbiol.* *15*, 531–543.

Brugiroux, S., Beutler, M., Pfann, C., Garzetti, D., Ruscheweyh, H.-J., Ring, D., Diehl, M., Herp, S., Lötscher, Y., Hussain, S., et al. (2016). Genome-guided design of a defined mouse microbiota that confers colonization resistance against *Salmonella enterica* serovar Typhimurium. *Nat. Microbiol.* *2*, 1–12.

Burghardt, L.T., Epstein, B., Guhlin, J., Nelson, M.S., Taylor, M.R., Young, N.D., Sadowsky, M.J., and Tiffin, P. (2018). Select and resequence reveals relative fitness of bacteria in symbiotic and free-living environments. *Proc. Natl. Acad. Sci.* 201714246.

## C

Cabreiro, F., Au, C., Leung, K.-Y., Vergara-Irigaray, N., Cochemé, H.M., Noori, T., Weinkove, D., Schuster, E., Greene, N.D.E., and Gems, D. (2013). Metformin Retards Aging in *C. elegans* by Altering Microbial Folate and Methionine Metabolism. *Cell* *153*, 228–239.

Cao, M., and Goodrich-Blair, H. (2017). Ready or Not: Microbial Adaptive Responses in Dynamic Symbiosis Environments. *J. Bacteriol.* JB.00883-16.

Capra, E.J., and Laub, M.T. (2012). Evolution of Two-Component Signal Transduction Systems. *Annu. Rev. Microbiol.* *66*, 325–347.

Caspi, R., Billington, R., Fulcher, C.A., Keseler, I.M., Kothari, A., Krummenacker, M., Latendresse, M., Midford, P.E., Ong, Q., Ong, W.K., et al. (2018). The MetaCyc database of metabolic pathways and enzymes. *Nucleic Acids Res.* *46*, D633–D639.

Cassidy, L., Petersen, C., Treitz, C., Dierking, K., Schulenburg, H., Leippe, M., and Tholey, A. (2018). The *Caenorhabditis elegans* Proteome Response to Naturally Associated Microbiome Members of the Genus *Ochrobactrum*. *PROTEOMICS* *18*, 1700426.

Caswell, H. (2001). *Matrix population models* (Sinauer Associates, Sunderland MA).

Chaguza, C. (2020). Bacterial survival: evolve and adapt or perish. *Nat. Rev. Microbiol.* *18*, 5–5.

Chain, P.S.G., Lang, D.M., Comerci, D.J., Malfatti, S.A., Vergez, L.M., Shin, M., Ugalde, R.A., Garcia, E., and Tolmasky, M.E. (2011). Genome of *Ochrobactrum anthropi* ATCC 49188 T, a versatile opportunistic pathogen and symbiont of several eukaryotic hosts. *J. Bacteriol.* *193*, 4274–4275.

Chan, S.Y., Liu, S.Y., Seng, Z., and Chua, S.L. (2020). Biofilm matrix disrupts nematode motility and predatory behavior. *ISME J.* 1–10.



- Chase, A.B., Arevalo, P., Brodie, E.L., Polz, M.F., Karaoz, U., and Martiny, J.B.H. (2019). Maintenance of Sympatric and Allopatric Populations in Free-Living Terrestrial Bacteria. *MBio* 10.
- Chatzidaki-Livanis, M., Geva-Zatorsky, N., and Comstock, L.E. (2016). *Bacteroides fragilis* type VI secretion systems use novel effector and immunity proteins to antagonize human gut Bacteroidales species. *Proc. Natl. Acad. Sci.* 113, 3627–3632.
- Chaudhari, S.N., Mukherjee, M., Vagasi, A.S., Bi, G., Rahman, M.M., Nguyen, C.Q., Paul, L., Selhub, J., and Kipreos, E.T. (2016). Bacterial Folates Provide an Exogenous Signal for *C. elegans* Germline Stem Cell Proliferation. *Dev. Cell* 38, 33–46.
- Chen, L., Zheng, D., Liu, B., Yang, J., and Jin, Q. (2016). VFDB 2016: hierarchical and refined dataset for big data analysis—10 years on. *Nucleic Acids Res.* 44, D694–D697.
- Chen, Y.E., Fischbach, M.A., and Belkaid, Y. (2018). Skin microbiota–host interactions. *Nature* 553, 427–436.
- Choi, J.I., Yoon, K., Subbammal Kalichamy, S., Yoon, S.-S., and Il Lee, J. (2016). A natural odor attraction between lactic acid bacteria and the nematode *Caenorhabditis elegans*. *ISME J.* 10, 558–567.
- Ciche, T.A., Darby, C., Ehlers, R.-U., Forst, S., and Goodrich-Blair, H. (2006). Dangerous liaisons: The symbiosis of entomopathogenic nematodes and bacteria. *Biol. Control* 38, 22–46.
- Cingolani, P., Platts, A., Wang, L.L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X., and Ruden, D.M. (2012a). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly (Austin)* 6, 80–92.
- Cingolani, P., Patel, V.M., Coon, M., Nguyen, T., Land, S.J., Ruden, D.M., and Lu, X. (2012b). Using *Drosophila melanogaster* as a Model for Genotoxic Chemical Mutational Studies with a New Program, SnpSift. *Front. Genet.* 3.
- Cohen, D., Mechold, U., Nevenzal, H., Yarmiyhu, Y., Randall, T.E., Bay, D.C., Rich, J.D., Parsek, M.R., Kaefer, V., Harrison, J.J., et al. (2015). Oligoribonuclease is a central feature of cyclic diguanylate signaling in *Pseudomonas aeruginosa*. *Proc. Natl. Acad. Sci.* 112, 11359–11364.
- Collin, L. (2018). XZ Utils.
- Consortium, T.U. (2014). Activities at the Universal Protein Resource (UniProt). *Nucleic Acids Res.* 42, D191–D198.
- Consuegra, J., Grenier, T., Baa-Puyoulet, P., Rahioui, I., Akherraz, H., Gervais, H., Parisot, N., Silva, P. da, Charles, H., Calevro, F., et al. (2020). *Drosophila*-associated bacteria differentially shape the nutritional requirements of their host during juvenile growth. *PLoS Biol.* 18, e3000681.
- Cook, D.E., Zdraljevic, S., Roberts, J.P., and Andersen, E.C. (2017). CeNDR, the *Caenorhabditis elegans* natural diversity resource. *Nucleic Acids Res.* 45, D650–D657.
- Cooper, S.K., Pandhare, J., Donald, S.P., and Phang, J.M. (2008). A Novel Function for Hydroxyproline Oxidase in Apoptosis through Generation of Reactive Oxygen Species. *J. Biol. Chem.* 283, 10485–10492.

Coyte, K.Z., and Rakoff-Nahoum, S. (2019). Understanding Competition and Cooperation within the Mammalian Gut Microbiome. *Curr. Biol.* 29, R538–R544.

Coyte, K.Z., Schluter, J., and Foster, K.R. (2015). The ecology of the microbiome: Networks, competition, and stability. *Science* 350, 663–666.

curlProject (2020). curl.

## D

Dahan, D., Preston, G.M., Sealey, J., and King, K.C. (2020). Impacts of a novel defensive symbiosis on the nematode host microbiome. *BMC Microbiol.* 20.

Darby, C., Hsu, J.W., Ghori, N., and Falkow, S. (2002). Plague bacteria biofilm blocks food intake. *Nature* 417, 243–244.

De Bary, A. (1879). *Die Erscheinung der Symbiose* (Strasbourg: Verlag J. Trubner).

De Henau, S., Tilleman, L., Vangheel, M., Luyckx, E., Trashin, S., Pauwels, M., Germani, F., Vlaeminck, C., Vanfleteren, J.R., Bert, W., et al. (2015). A redox signalling globin is essential for reproduction in *Caenorhabditis elegans*. *Nat. Commun.* 6, 8782.

Deines, P., Hammerschmidt, K., and Bosch, T.C.G. (2020). Microbial Species Coexistence Depends on the Host Environment. *MBio* 11.

Diaz, S.A., and Restif, O. (2014). Spread and Transmission of Bacterial Pathogens in Experimental Populations of the Nematode *Caenorhabditis elegans*. *Appl. Environ. Microbiol.* 80, 5411–5418.

Dierking, K., Yang, W., and Schulenburg, H. (2016). Antimicrobial effectors in the nematode *Caenorhabditis elegans*: an outgroup to the Arthropoda. *Phil Trans R Soc B* 371, 20150299.

Dimov, I., and Maduro, M.F. (2019). The *C. elegans* intestine: organogenesis, digestion, and physiology. *Cell Tissue Res.* 377, 383–396.

Dirksen, P., Marsh, S.A., Braker, I., Heitland, N., Wagner, S., Nakad, R., Mader, S., Petersen, C., Kowallik, V., Rosenstiel, P., et al. (2016). The native microbiome of the nematode *Caenorhabditis elegans*: gateway to a new host-microbiome model. *BMC Biol.* 14, 38.

Dirksen, P., Assié, A., Zimmermann, J., Zhang, F., Tietje, A.-M., Marsh, S.A., Félix, M.-A., Shapira, M., Kaleta, C., Schulenburg, H., et al. (2020). CeMbio - The *Caenorhabditis elegans* Microbiome Resource. *G3 Genes Genomes Genet.*

Dodds, W.K., Zeglin, L.H., Ramos, R.J., Platt, T.G., Pandey, A., Michaels, T., Masigol, M., Klompen, A.M.L., Kelly, M.C., Jumpponen, A., et al. Connections and Feedback: Aquatic, Plant, and Soil Microbiomes in Heterogeneous and Changing Environments. *BioScience*.

Donaldson, G.P., Lee, S.M., and Mazmanian, S.K. (2016). Gut biogeography of the bacterial microbiota. *Nat. Rev. Microbiol.* 14, 20–32.

Donelli, G., Vuotto, C., Cardines, R., and Mastrantonio, P. (2012). Biofilm-growing intestinal anaerobic bacteria. *FEMS Immunol. Med. Microbiol.* 65, 318–325.

Douglas, A.E. (2018). *Fundamentals of Microbiome Science: How Microbes Shape Animal Biology* (Princeton University Press).

Douglas, A.E. (2019). Simple animal models for microbiome research. *Nat. Rev. Microbiol.* *17*, 764–775.

Douglas, A.E. (2020). The microbial exometabolome: ecological resource and architect of microbial communities. *Philos. Trans. R. Soc. B Biol. Sci.* *375*, 20190250.

Douglas, A.E., and Werren, J.H. (2016). Holes in the Hologenome: Why Host-Microbe Symbioses Are Not Holobionts. *MBio* *7*, e02099-15.

Dubilier, N., Mülders, C., Ferdelman, T., de Beer, D., Pernthaler, A., Klein, M., Wagner, M., Erséus, C., Thiermann, F., Krieger, J., et al. (2001). Endosymbiotic sulphate-reducing and sulphide-oxidizing bacteria in an oligochaete worm. *Nature* *411*, 298–302.

Dubilier, N., Bergin, C., and Lott, C. (2008). Symbiotic diversity in marine animals: the art of harnessing chemosynthesis. *Nat. Rev. Microbiol.* *6*, 725–740.

## E

Egan, S., Fukatsu, T., and Francino, M.P. (2020). Opportunities and Challenges to Microbial Symbiosis Research in the Microbiome Era. *Front. Microbiol.* *11*.

Ellegaard, K.M., and Engel, P. (2019). Genomic diversity landscape of the honey bee gut microbiota. *Nat. Commun.* *10*, 446.

Ellegaard, K.M., Suenami, S., Miyazaki, R., and Engel, P. (2020). Vast Differences in Strain-Level Diversity in the Gut Microbiota of Two Closely Related Honey Bee Species. *Curr. Biol.* *30*, 2520-2531.e7.

Evans, S.E., and Wallenstein, M.D. (2014). Climate change alters ecological strategies of soil bacteria. *Ecol. Lett.* *17*, 155–164.

Ewels, P., Magnusson, M., Lundin, S., and Käller, M. (2016). MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* *32*, 3047–3048.

## F

Fahimipour, A.K., and Gross, T. (2020). Mapping the bacterial metabolic niche space. *Nat. Commun.* *11*, 4887.

Fang-Yen, C., Avery, L., and Samuel, A.D.T. (2009). Two size-selective mechanisms specifically trap bacteria-sized food particles in *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci.* *106*, 20093–20096.

Faust, K., Sathirapongsasuti, J.F., Izard, J., Segata, N., Gevers, D., Raes, J., and Huttenhower, C. (2012). Microbial Co-occurrence Relationships in the Human Microbiome. *PLOS Comput. Biol.* *8*, e1002606.

Félix, M.-A., and Braendle, C. (2010). The natural history of *Caenorhabditis elegans*. *Curr. Biol.* *20*, R965–R969.

- Fierer, N. (2017). Embracing the unknown: disentangling the complexities of the soil microbiome. *Nat. Rev. Microbiol.* *15*, 579–590.
- Fischer, C., Trautman, E.P., Crawford, J.M., Stabb, E.V., Handelsman, J., and Broderick, N.A. (2017). Metabolite exchange between microbiome members produces compounds that influence *Drosophila* behavior. *ELife* *6*, e18855.
- Fisher, R.M., Henry, L.M., Cornwallis, C.K., Kiers, E.T., and West, S.A. (2017). The evolution of host-symbiont dependence. *Nat. Commun.* *8*, ncomms15973.
- Flemming, H.-C., and Wuertz, S. (2019). Bacteria and archaea on Earth and their abundance in biofilms. *Nat. Rev. Microbiol.* *17*, 247–260.
- Flemming, H.-C., Wingender, J., Szewzyk, U., Steinberg, P., Rice, S.A., and Kjelleberg, S. (2016). Biofilms: an emergent form of bacterial life. *Nat. Rev. Microbiol.* *14*, 563–575.
- Flint, H.J., Scott, K.P., Duncan, S.H., Louis, P., and Forano, E. (2012). Microbial degradation of complex carbohydrates in the gut. *Gut Microbes* *3*, 289–306.
- Ford, S.A., Kao, D., Williams, D., and King, K.C. (2016). Microbe-mediated host defence drives the evolution of reduced pathogen virulence. *Nat. Commun.* *7*, 13430.
- Foster, K.R., Schluter, J., Coyte, K.Z., and Rakoff-Nahoum, S. (2017). The evolution of the host microbiome as an ecosystem on a leash. *Nature* *548*, 43–51.
- Franzenburg, S., Walter, J., Künzel, S., Wang, J., Baines, J.F., Bosch, T.C.G., and Fraune, S. (2013). Distinct antimicrobial peptide expression determines host species-specific bacterial associations. *Proc. Natl. Acad. Sci.* *110*, E3730–E3738.
- Fraune, S., and Bosch, T.C.G. (2007). Long-term maintenance of species-specific bacterial microbiota in the basal metazoan *Hydra*. *Proc. Natl. Acad. Sci.* *104*, 13146–13151.
- Frazão, N., Sousa, A., Lässig, M., and Gordo, I. (2019). Horizontal gene transfer overrides mutation in *Escherichia coli* colonizing the mammalian gut. *Proc. Natl. Acad. Sci.* *116*, 17906–17915.

## G

- Gailly, J.-L., and Adler, M. (2017). zlib.
- Galperin, M.Y. (2005). A census of membrane-bound and intracellular signal transduction proteins in bacteria: bacterial IQ, extroverts and introverts. *BMC Microbiol.* *5*, 35.
- Gao, Y., Traulsen, A., and Pichugin, Y. (2019). Interacting cells driving the evolution of multicellular life cycles. *PLOS Comput. Biol.* *15*, e1006987.
- García-González, A.P., Ritter, A.D., Shrestha, S., Andersen, E.C., Yilmaz, L.S., and Walhout, A.J.M. (2017). Bacterial Metabolism Affects the *C. elegans* Response to Cancer Chemotherapeutics. *Cell* *169*, 431–441.e8.
- Garigan, D., Hsu, A.-L., Fraser, A.G., Kamath, R.S., Ahringer, J., and Kenyon, C. (2002). Genetic Analysis of Tissue Aging in *Caenorhabditis elegans*: A Role for Heat-Shock Factor and Bacterial Proliferation. *Genetics* *161*, 1101–1112.

Gause, G.F. (1934). Experimental Analysis of Vito Volterra's Mathematical Theory of the Struggle for Existence. *Science* 79, 16–17.

Geesey, G.G., Richardson, W.T., Yeomans, H.G., Irvin, R.T., and Costerton, J.W. (1977). Microscopic examination of natural sessile bacterial populations from an alpine stream. *Can. J. Microbiol.* 23, 1733–1736.

Gelius-Dietrich, G., Desouki, A.A., Fritzeimer, C.J., and Lercher, M.J. (2013). sybil – Efficient constraint-based modelling in R. *BMC Syst. Biol.* 7, 125.

Genuer, R., Poggi, J.-M., and Tuleau-Malot, C. (2015). VSURF: An R Package for Variable Selection Using Random Forests. *R J.* 7, 19.

Gestel, J. van, and Nowak, M.A. (2016). Phenotypic Heterogeneity and the Evolution of Bacterial Life Cycles. *PLOS Comput. Biol.* 12, e1004764.

Goodrich-Blair, H., and Clarke, D.J. (2007). Mutualism and pathogenesis in *Xenorhabdus* and *Photorhabdus*: two roads to the same destination. *Mol. Microbiol.* 64, 260–268.

Gorter, F.A., Manhart, M., and Ackermann, M. (2020). Understanding the evolution of interspecies interactions in microbial communities. *Philos. Trans. R. Soc. B Biol. Sci.* 375, 20190256.

Goya, M.E., Xue, F., Sampedro-Torres-Quevedo, C., Arnaouteli, S., Riquelme-Dominguez, L., Romanowski, A., Brydon, J., Ball, K.L., Stanley-Wall, N.R., and Doitsidou, M. (2020). Probiotic *Bacillus subtilis* Protects against  $\alpha$ -Synuclein Aggregation in *C. elegans*. *Cell Rep.* 30, 367-380.e7.

Graham, P.L., Johnson, J.J., Wang, S., Sibley, M.H., Gupta, M.C., and Kramer, J.M. (1997). Type IV collagen is detectable in most, but not all, basement membranes of *Caenorhabditis elegans* and assembles on tissues that do not express it. *J. Cell Biol.* 137, 1171–1183.

Grant, A. (1997). Selection pressures on vital rates in density-dependent populations. *Proc Biol Sci* 264, 303–306.

Grant, A., and Benton, T.G. (2000). Elasticity analysis for density-dependent populations in stochastic environments. *Ecology* 81, 680–693.

Grenier, T., and Leulier, F. (2020). How commensal microbes shape the physiology of *Drosophila melanogaster*. *Curr. Opin. Insect Sci.*

Grewal, P.S. (1991). Influence of Bacteria and Temperature On the Reproduction of *Caenorhabditis Elegans* (Nematoda: Rhabditidae) Infesting Mushrooms (*Agaricus Bispor Us*). *Nematologica* 37, 72–82.

Grime, J.P. (1974). Vegetation classification by reference to strategies. *Nature* 250, 26–31.

Grime, J.P. (1977). Evidence for the Existence of Three Primary Strategies in Plants and Its Relevance to Ecological and Evolutionary Theory. *Am. Nat.* 111, 1169–1194.

## H

Halliwell, B. (2007). Biochemistry of oxidative stress. *Biochem. Soc. Trans.* 35, 1147–1150.

- Hall-Stoodley, L., Costerton, J.W., and Stoodley, P. (2004). Bacterial biofilms: from the Natural environment to infectious diseases. *Nat. Rev. Microbiol.* 2, 95–108.
- Hammerschmidt, K., Rose, C.J., Kerr, B., and Rainey, P.B. (2014). Life cycles, fitness decoupling and the evolution of multicellularity. *Nature* 515, 75–79.
- Hanski, I. (1998). Metapopulation dynamics. *Nature* 396, 41–49.
- Hartigan, J.A., and Wong, M.A. (1979). Algorithm AS 136: A K-Means Clustering Algorithm. *J. R. Stat. Soc. Ser. C Appl. Stat.* 28, 100–108.
- Hassani, M.A., Özkurt, E., Franzenburg, S., and Stukenbrock, E.H. (2020). Ecological Assembly Processes of the Bacterial and Fungal Microbiota of Wild and Domesticated Wheat Species. *Phytobiomes J.* 4, 217–224.
- Heinken, A., Sahoo, S., Fleming, R.M.T., and Thiele, I. (2013). Systems-level characterization of a host-microbe metabolic symbiosis in the mammalian gut. *Gut Microbes* 4, 28–40.
- Hengge, R. (2020). Linking bacterial growth, survival, and multicellularity – small signaling molecules as triggers and drivers. *Curr. Opin. Microbiol.* 55, 57–66.
- Henry, C.S., DeJongh, M., Best, A.A., Frybarger, P.M., Linsay, B., and Stevens, R.L. (2010). High-throughput generation, optimization and analysis of genome-scale metabolic models. *Nat. Biotechnol.* 28, 977–982.
- Hmelo, L.R., Borlee, B.R., Almlblad, H., Love, M.E., Randall, T.E., Tseng, B.S., Lin, C., Irie, Y., Storek, K.M., Yang, J.J., et al. (2015). Precision-engineering the *Pseudomonas aeruginosa* genome with two-step allelic exchange. *Nat. Protoc.* 10, 1820–1841.
- Hoang, K.L., Morran, L.T., and Gerardo, N.M. (2016). Experimental Evolution as an Underutilized Tool for Studying Beneficial Animal–Microbe Interactions. *Front. Microbiol.* 7.
- Hoang, K.L., Morran, L.T., and Gerardo, N.M. (2019). Can a Symbiont (Also) Be Food? *Front. Microbiol.* 10.
- Hoch, J.A. (2000). Two-component and phosphorelay signal transduction. *Curr. Opin. Microbiol.* 3, 165–170.
- Hoek, T.A., Axelrod, K., Biancalani, T., Yurtsev, E.A., Liu, J., and Gore, J. (2016). Resource Availability Modulates the Cooperative and Competitive Nature of a Microbial Cross-Feeding Mutualism. *PLOS Biol* 14, e1002540.
- Houle, D. (1992). Comparing evolvability and variability of quantitative traits. *Genetics* 130, 195–204.
- Hrček, J., Parker, B.J., McLean, A.H.C., Simon, J.-C., Mann, C.M., and Godfray, H.C.J. (2018). Hosts do not simply outsource pathogen resistance to protective symbionts. *Evolution* 72, 1488–1499.
- Huang, W., and Wilks, A. (2017). A rapid seamless method for gene knockout in *Pseudomonas aeruginosa*. *BMC Microbiol.* 17.
- Hunter, P. (2018). The revival of the extended phenotype. *EMBO Rep.* 19.

Hurford, A., Cownden, D., and Day, T. (2010). Next-generation tools for evolutionary invasion analyses. *J. R. Soc. Interface* 7, 561–571.

Hutter, H., Vogel, B.E., Plenefisch, J.D., Norris, C.R., Proenca, R.B., Spieth, J., Guo, C., Mastwal, S., Zhu, X., Scheel, J., et al. (2000). Conservation and novelty in the evolution of cell adhesion and extracellular matrix genes. *Science* 287, 989–994.

## J

Jahn, M.T., Arkhipova, K., Markert, S.M., Stigloher, C., Lachnit, T., Pita, L., Kupczok, A., Ribes, M., Stengel, S.T., Rosenstiel, P., et al. (2019). A Phage Protein Aids Bacterial Symbionts in Eukaryote Immune Evasion. *Cell Host Microbe* 26, 542–550.e5.

Jansen, G., Crummenerl, L.L., Gilbert, F., Mohr, T., Pfefferkorn, R., Thänert, R., Rosenstiel, P., and Schulenburg, H. (2015). Evolutionary Transition from Pathogenicity to Commensalism: Global Regulator Mutations Mediate Fitness Gains through Virulence Attenuation. *Mol. Biol. Evol.* msv160.

Jenal, U., Reinders, A., and Lori, C. (2017). Cyclic di-GMP: second messenger extraordinaire. *Nat. Rev. Microbiol.* 15, 271–284.

Johnke, J., Dirksen, P., and Schulenburg, H. (2020). Community assembly of the native *C. elegans* microbiome is influenced by time, substrate and individual bacterial taxa. *Environ. Microbiol.* 22, 1265–1279.

Jose, P.A., Ben-Yosef, M., Jurkevitch, E., and Yuval, B. (2019). Symbiotic bacteria affect oviposition behavior in the olive fruit fly *Bactrocera oleae*. *J. Insect Physiol.* 117, 103917.

## K

Kaltenpoth, M., and Flórez, L.V. (2020). Versatile and Dynamic Symbioses Between Insects and Burkholderia Bacteria. *Annu. Rev. Entomol.* 65, 145–170.

Kaltenpoth, M., Roeser-Mueller, K., Koehler, S., Peterson, A., Nechitaylo, T.Y., Stubblefield, J.W., Herzner, G., Seger, J., and Strohm, E. (2014). Partner choice and fidelity stabilize coevolution in a Cretaceous-age defensive symbiosis. *Proc. Natl. Acad. Sci.* 111, 6359–6364.

Kamke, J., Sczyrba, A., Ivanova, N., Schwientek, P., Rinke, C., Mavromatis, K., Woyke, T., and Hentschel, U. (2013). Single-cell genomics reveals complex carbohydrate degradation patterns in poribacterial symbionts of marine sponges. *ISME J.* 7, 2287–2300.

Kim, J.K., Kwon, J.Y., Kim, S.K., Han, S.H., Won, Y.J., Lee, J.H., Kim, C.-H., Fukatsu, T., and Lee, B.L. (2014). Purine Biosynthesis, Biofilm Formation, and Persistence of an Insect-Microbe Gut Symbiosis. *Appl. Environ. Microbiol.* 80, 4374–4382.

King, K.C., Brockhurst, M.A., Vasieva, O., Paterson, S., Betts, A., Ford, S.A., Frost, C.L., Horsburgh, M.J., Haldenby, S., and Hurst, G.D. (2016). Rapid evolution of microbe-mediated protection against pathogens in a worm host. *ISME J.* 10, 1915–1924.

Kissoyan, K.A.B., Drechsler, M., Stange, E.-L., Zimmermann, J., Kaleta, C., Bode, H.B., and Dierking, K. (2019). Natural *C. elegans* Microbiota Protects against Infection via Production of a Cyclic Lipopeptide of the Viscosin Group. *Curr. Biol.* 29, 1030-1037.e5.

Koboldt, D.C., Zhang, Q., Larson, D.E., Shen, D., McLellan, M.D., Lin, L., Miller, C.A., Mardis, E.R., Ding, L., and Wilson, R.K. (2012). VarScan 2: Somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res.* 22, 568–576.

Koch, R. (1884). Die Ätiologie der Tuberkulose. *Mitteilungen Aus Dem Kaiserl Gesundheitsamte II.*

Koch, E.J., and McFall-Ngai, M. (2018). Model systems for the study of how symbiotic associations between animals and extracellular bacterial partners are established and maintained. *Drug Discov. Today Dis. Models* 28, 3–12.

Kohl, K.D. (2020). Ecological and evolutionary mechanisms underlying patterns of phyllosymbiosis in host-associated microbial communities. *Philos. Trans. R. Soc. B Biol. Sci.* 375, 20190251.

Kolodny, O., and Schulenburg, H. (2020). Microbiome-mediated plasticity directs host evolution along several distinct time scales. *Philos. Trans. R. Soc. B Biol. Sci.* 375, 20190589.

Kordes, A., Grahl, N., Koska, M., Preusse, M., Arce-Rodriguez, A., Abraham, W.-R., Kaever, V., and Häussler, S. (2019). Establishment of an induced memory response in *Pseudomonas aeruginosa* during infection of a eukaryotic host. *ISME J.* 13, 2018.

Kumar, V.S., and Maranas, C.D. (2009). GrowMatch: An Automated Method for Reconciling In Silico/In Vivo Growth Predictions. *PLOS Comput. Biol.* 5, e1000308.

Kwong, W.K., and Moran, N.A. (2016). Gut microbial communities of social bees. *Nat. Rev. Microbiol.* 14, 374–384.

Kwong, W.K., Engel, P., Koch, H., and Moran, N.A. (2014). Genomics and host specialization of honey bee and bumble bee gut symbionts. *Proc. Natl. Acad. Sci.* 111, 11509–11514.

## L

Lagesen, K., Hallin, P., Rødland, E.A., Staerfeldt, H.-H., Rognes, T., and Ussery, D.W. (2007). RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.* 35, 3100–3108.

Lande, R. (1982). A Quantitative Genetic Theory of Life History Evolution. *Ecology* 63, 607–615.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359.

Lebeer, S., Verhoeven, T.L.A., Claes, I.J.J., Hertogh, G.D., Vermeire, S., Buyse, J., Immerseel, F.V., Vanderleyden, J., and Keersmaecker, S.C.J.D. (2011). FISH analysis of *Lactobacillus* biofilms in the gastrointestinal tract of different hosts. *Lett. Appl. Microbiol.* 52, 220–226.



Lebov, J.F., Schlomann, B.H., Robinson, C.D., and Bohannan, B.J.M. (2020). Phenotypic parallelism during experimental adaptation of a free-living bacterium to the zebrafish gut (Evolutionary Biology).

Lederberg, J., and McCray, A.T. (2001). 'Ome Sweet 'Omics - A Genealogical Treasury of Words | National Library of Medicine. *Scientist* 15, 8.

Leonard, S.P., Powell, J.E., Perutka, J., Geng, P., Heckmann, L.C., Horak, R.D., Davies, B.W., Ellington, A.D., Barrick, J.E., and Moran, N.A. (2020). Engineered symbionts activate honey bee immunity and limit pathogens. *Science* 367, 573–576.

Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* 27, 2987–2993.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 25, 1754–1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078–2079.

Lind, P.A., Farr, A.D., and Rainey, P.B. (2015). Experimental evolution reveals hidden diversity in evolutionary pathways. *ELife* 4, e07074.

Litsios, G., Sims, C.A., Wüest, R.O., Pearman, P.B., Zimmermann, N.E., and Salamin, N. (2012). Mutualism with sea anemones triggered the adaptive radiation of clownfishes. *BMC Evol. Biol.* 12, 212.

Liu, W., Cremer, J., Li, D., Hwa, T., and Liu, C. (2019). An evolutionarily stable strategy to colonize spatially extended habitats. *Nature* 575, 664–668.

Liu, Z., Beskrovnaya, P., Melnyk, R.A., Hossain, S.S., Khorasani, S., O'Sullivan, L.R., Wiesmann, C.L., Bush, J., Richard, J.D., and Haney, C.H. (2018). A Genome-Wide Screen Identifies Genes in Rhizosphere-Associated *Pseudomonas* Required to Evade Plant Defenses. *MBio* 9.

Lopez, A.O., Vega, N.M., and Gore, J. (2019). Interspecies bacterial competition determines community assembly in the *C. elegans* intestine. *BioRxiv* 535633.

Louca, S., Parfrey, L.W., and Doebeli, M. (2016). Decoupling function and taxonomy in the global ocean microbiome. *Science* 353, 1272–1277.

Luan, J.-B., Chen, W., Hasegawa, D.K., Simmons, A.M., Wintermantel, W.M., Ling, K.-S., Fei, Z., Liu, S.-S., and Douglas, A.E. (2015). Metabolic Coevolution in the Bacterial Symbiosis of Whiteflies and Related Plant Sap-Feeding Insects. *Genome Biol. Evol.* 7, 2635–2647.

Luis, A.S., Briggs, J., Zhang, X., Farnell, B., Ndeh, D., Labourel, A., Baslé, A., Cartmell, A., Terrapon, N., Stott, K., et al. (2018). Dietary pectic glycans are degraded by coordinated enzyme pathways in human colonic *Bacteroides*. *Nat. Microbiol.* 3, 210–219.

## M

- MacArthur, R.H., and Wilson, E.D. (1967). *The theory of island biogeography* (N.J.: Princeton University Press).
- MacNeil, L., Watson, E., Arda, H.E., Zhu, L.J., and Walhout, A.J.M. (2013a). Diet-Induced Developmental Acceleration Independent of TOR and Insulin in *C. elegans*. *Cell* 153.
- MacNeil, L.T., Watson, E., Arda, H.E., Zhu, L.J., and Walhout, A.J.M. (2013b). Diet-Induced Developmental Acceleration Independent of TOR and Insulin in *C. elegans*. *Cell* 153, 240–252.
- Madeira, F., Park, Y.M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., Basutkar, P., Tivey, A.R.N., Potter, S.C., Finn, R.D., et al. (2019). The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* 47, W636–W641.
- Magnúsdóttir, S., Heinken, A., Kutt, L., Ravcheev, D.A., Bauer, E., Noronha, A., Greenhalgh, K., Jäger, C., Baginska, J., Wilmes, P., et al. (2017). Generation of genome-scale metabolic reconstructions for 773 members of the human gut microbiota. *Nat. Biotechnol.* 35, 81–89.
- Maher, A.M.D., Asaiyah, M., Quinn, S., Burke, R., Wolff, H., Bode, H.B., and Griffin, C.T. (2020). Competition and Co-existence of Two *Photobacterium* Symbionts with a Nematode Host. *Microb. Ecol.*
- Malik, A.A., Martiny, J.B.H., Brodie, E.L., Martiny, A.C., Treseder, K.K., and Allison, S.D. (2020). Defining trait-based microbial strategies with consequences for soil carbon cycling under climate change. *ISME J.* 14, 1–9.
- Mao, M., Yang, X., and Bennett, G.M. (2018). Evolution of host support for two ancient bacterial symbionts with differentially degraded genomes in a leafhopper host. *Proc. Natl. Acad. Sci.* 115, E11691–E11700.
- Martino, M.E., Bayjanov, J.R., Caffrey, B.E., Wels, M., Joncour, P., Hughes, S., Gillet, B., Kleerebezem, M., van Hijum, S.A.F.T., and Leulier, F. (2016). Nomadic lifestyle of *Lactobacillus plantarum* revealed by comparative genomics of 54 strains isolated from different habitats. *Environ. Microbiol.* 18, 4974–4989.
- Martino, M.E., Joncour, P., Leenay, R., Gervais, H., Shah, M., Hughes, S., Gillet, B., Beisel, C., and Leulier, F. (2018a). Bacterial Adaptation to the Host's Diet Is a Key Evolutionary Force Shaping *Drosophila*-*Lactobacillus* Symbiosis. *Cell Host Microbe*.
- Martino, M.E., Joncour, P., Leenay, R., Gervais, H., Shah, M., Hughes, S., Gillet, B., Beisel, C., and Leulier, F. (2018b). Bacterial adaptation to diet is a key evolutionary force shaping *Drosophila*-*Lactobacillus* symbiosis. *BioRxiv* 222364.
- Masri, L., Branca, A., Sheppard, A.E., Papkou, A., Laehnemann, D., Guenther, P.S., Pahl, S., Saebelfeld, M., Hollensteiner, J., Liesegang, H., et al. (2015). Host-Pathogen Coevolution: The Selective Advantage of *Bacillus thuringiensis* Virulence and Its Cry Toxin Genes. *PLoS Biol.* 13.
- Mattingly, H., and Emonet, T. (2019). A rule from bacteria to balance growth and expansion. *Nature* 575, 602–603.

- Mazumdar, T., Teh, B.S., Murali, A., Schmidt-Heck, W., Schlenker, Y., Vogel, H., and Boland, W. (2020). Survival strategies of *Enterococcus mundtii* in the gut of *Spodoptera littoralis*: a live report. *BioRxiv* 2020.02.03.932053.
- McDonald, M.J., Gehrig, S.M., Meintjes, P.L., Zhang, X.-X., and Rainey, P.B. (2009). Adaptive Divergence in Experimental Populations of *Pseudomonas fluorescens*. IV. Genetic Constraints Guide Evolutionary Trajectories in a Parallel Adaptive Radiation. *Genetics* 183, 1041–1053.
- McFall-Ngai, M., Hadfield, M.G., Bosch, T.C.G., Carey, H.V., Domazet-Lošo, T., Douglas, A.E., Dübilier, N., Eberl, G., Fukami, T., Gilbert, S.F., et al. (2013). Animals in a bacterial world, a new imperative for the life sciences. *Proc. Natl. Acad. Sci.* 110, 3229–3236.
- McLoughlin, K., Schluter, J., Rakoff-Nahoum, S., Smith, A.L., and Foster, K.R. (2016). Host Selection of Microbiota via Differential Adhesion. *Cell Host Microbe* 19, 550–559.
- Medina, I., and Langmore, N.E. (2015). Coevolution is linked with phenotypic diversification but not speciation in avian brood parasites. *Proc. R. Soc. B Biol. Sci.* 282, 20152056.
- Metcalf, C.J.E., Henry, L.P., Rebolledo-Gómez, M., and Koskella, B. (2019). Why Evolve Reliance on the Microbiome for Timing of Ontogeny? *MBio* 10.
- Militello, G. (2019). Motility Control of Symbionts and Organelles by the Eukaryotic Cell: The Handling of the Motile Capacity of Individual Parts Forges a Collective Biological Identity. *Front. Psychol.* 10.
- Miller, E.T., and Bohannan, B.J.M. (2019). Life Between Patches: Incorporating Microbiome Biology Alters the Predictions of Metacommunity Models. *Front. Ecol. Evol.* 7.
- Miller, E.T., Svanbäck, R., and Bohannan, B.J.M. (2018). Microbiomes as Metacommunities: Understanding Host-Associated Microbes through Metacommunity Ecology. *Trends Ecol. Evol.* 33, 926–935.
- Mitchell, A.L., Attwood, T.K., Babbitt, P.C., Blum, M., Bork, P., Bridge, A., Brown, S.D., Chang, H.-Y., El-Gebali, S., Fraser, M.I., et al. (2019). InterPro in 2019: improving coverage, classification and access to protein sequence annotations. *Nucleic Acids Res.* 47, D351–D360.
- Moeller, A.H., Suzuki, T.A., Phifer-Rixey, M., and Nachman, M.W. (2018). Transmission modes of the mammalian gut microbiota. *Science* 362, 453–457.
- Moeller, A.H., Gomes-Neto, J.C., Mantz, S., Kittana, H., Segura Munoz, R.R., Schmaltz, R.J., Ramer-Tait, A.E., and Nachman, M.W. (2019). Experimental Evidence for Adaptation to Species-Specific Gut Microbiota in House Mice. *MSphere* 4.
- Moran, N.A., Hansen, A.K., Powell, J.E., and Sabree, Z.L. (2012). Distinctive Gut Microbiota of Honey Bees Assessed Using Deep Sampling from Individual Worker Bees. *PLoS ONE* 7.
- Mörck, C., and Pilon, M. (2006). *C. elegans* feeding defective mutants have shorter body lengths and increased autophagy. *BMC Dev. Biol.* 6, 39.
- Moreillon, P., Entenza, J.M., Francioli, P., McDevitt, D., Foster, T.J., François, P., and Vaudaux, P. (1995). Role of *Staphylococcus aureus* coagulase and clumping factor in pathogenesis of experimental endocarditis. *Infect. Immun.* 63, 4738–4743.

Moya, A., and Ferrer, M. (2016). Functional Redundancy-Induced Stability of Gut Microbiota Subjected to Disturbance. *Trends Microbiol.* *24*, 402–413.

Mushegian, A.A., and Ebert, D. (2016). Rethinking “mutualism” in diverse host-symbiont communities. *BioEssays* *38*, 100–108.

## N

Na, H., Ponomarova, O., Giese, G.E., and Walhout, A.J.M. (2018). *C. elegans* MRP-5 Exports Vitamin B12 from Mother to Offspring to Support Embryonic Development. *Cell Rep.* *22*, 3126–3133.

Newell, P.D., Chaston, J.M., Wang, Y., Winans, N.J., Sannino, D.R., Wong, A.C.N., Dobson, A.J., Kagle, J., and Douglas, A.E. (2014). In vivo function and comparative genomic analyses of the *Drosophila* gut microbiota identify candidate symbiosis factors. *Front. Microbiol.* *5*.

Norvaisas, P., and Cabreiro, F. (2018). Pharmacology in the age of the holobiont. *Curr. Opin. Syst. Biol.* *10*, 34–42.

Nyholm, S.V., and McFall-Ngai, M.J. (1998). Sampling the Light-Organ Microenvironment of *Euprymna scolopes*: Description of a Population of Host Cells in Association With the Bacterial Symbiont *Vibrio fischeri*. *Biol. Bull.* *195*, 89–97.

## O

O’Donnell, M.P., Fox, B.W., Chao, P.-H., Schroeder, F.C., and Sengupta, P. (2020). A neurotransmitter produced by gut bacteria modulates host sensory behaviour. *Nature* 1–6.

Oh, Y.-K., Palsson, B.O., Park, S.M., Schilling, C.H., and Mahadevan, R. (2007). Genome-scale Reconstruction of Metabolic Network in *Bacillus subtilis* Based on High-throughput Phenotyping and Gene Essentiality Data. *J. Biol. Chem.* *282*, 28791–28799.

Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P.R., O’Hara, R.B., Simpson, G.L., Solymos, P., et al. (2019). *vegan*: Community Ecology Package. R package version 2.5-6.

van Oppen, M.J.H., and Blackall, L.L. (2019). Coral microbiome dynamics, functions and design in a changing world. *Nat. Rev. Microbiol.* *17*, 557–567.

Orr, H.A. (2005). The genetic theory of adaptation: a brief history. *Nat. Rev. Genet.* *6*, 119–127.

Orr, M.W., Donaldson, G.P., Severin, G.B., Wang, J., Sintim, H.O., Waters, C.M., and Lee, V.T. (2015). Oligoribonuclease is the primary degradative enzyme for pGpG in *Pseudomonas aeruginosa* that is required for cyclic-di-GMP turnover. *Proc. Natl. Acad. Sci.* *112*, E5048–E5057.

Orr, M.W., Weiss, C.A., Severin, G.B., Turdiev, H., Kim, S.-K., Turdiev, A., Liu, K., Tu, B.P., Waters, C.M., Winkler, W.C., et al. (2018). A Subset of Exoribonucleases Serve as Degradative Enzymes for pGpG in c-di-GMP Signaling. *J. Bacteriol.* *200*.

Orth, J.D., Thiele, I., and Palsson, B.Ø. (2010). What is flux balance analysis? *Nat. Biotechnol.* *28*, 245–248.

O'Toole, G.A. (2011). Microtiter Dish Biofilm Formation Assay. *J. Vis. Exp. JoVE*.

Otto, M. (2014). Physical stress and bacterial colonization. *FEMS Microbiol. Rev.* *38*, 1250–1270.

Oulhen, N., Schulz, B.J., and Carrier, T.J. (2016). English translation of Heinrich Anton de Bary's 1878 speech, 'Die Erscheinung der Symbiose' ('De la symbiose'). *Symbiosis* *69*, 131–139.

## P

Page, A.P., and Johnstone, I.L. (2007). The cuticle (WormBook).

Pagès, H., Aboyoun, P., Gentleman, R., and DebRoy, S. (2018). Biostrings: efficient manipulation of biological strings. R package version 2.50.1.

Pais, I.S., Valente, R.S., Sporniak, M., and Teixeira, L. (2018). *Drosophila melanogaster* establishes a species-specific mutualistic interaction with stable gut-colonizing bacteria. *PLOS Biol.* *16*, e2005710.

Pamer, E.G. (2016). Resurrecting the intestinal microbiota to combat antibiotic-resistant pathogens. *Science* *352*, 535–538.

Panke-Buisse, K., Poole, A.C., Goodrich, J.K., Ley, R.E., and Kao-Kniffin, J. (2015). Selection on soil microbiomes reveals reproducible impacts on plant function. *ISME J.* *9*, 980–989.

Papkou, A., Guzella, T., Yang, W., Koepper, S., Pees, B., Schalkowski, R., Barg, M.-C., Rosenstiel, P.C., Teotónio, H., and Schulenburg, H. (2018). The genomic basis of Red Queen dynamics during rapid reciprocal host–pathogen coevolution. *Proc. Natl. Acad. Sci.* 201810402.

Park, M., Loverdo, C., Schreiber, S.J., and Lloyd-Smith, J.O. (2013). Multiple scales of selection influence the evolutionary emergence of novel pathogens. *Philos. Trans. R. Soc. B Biol. Sci.* *368*.

Pascoe, E.L., Hauffe, H.C., Marchesi, J.R., and Perkins, S.E. (2017). Network analysis of gut microbiota literature: an overview of the research landscape in non-human animal studies. *ISME J.* *11*, 2644–2651.

Petersen, C., Dirksen, P., and Schulenburg, H. (2015). Why we need more ecology for genetic models such as *C. elegans*. *Trends Genet.* *31*, 120–127.

Petersen, T.N., Brunak, S., von Heijne, G., and Nielsen, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods* *8*, 785–786.

Pianka, E.R. (1970). On r- and K-Selection. *The American Naturalist* *104*, 592–597.

Picazo, D.R., Dagan, T., Ansorge, R., Petersen, J.M., Dubilier, N., and Kupczok, A. (2019). Horizontally transmitted symbiont populations in deep-sea mussels are genetically isolated. *ISME J.* *13*, 2954–2968.

Pichugin, Y., Peña, J., Rainey, P.B., and Traulsen, A. (2017). Fragmentation modes and the evolution of life cycles. *PLOS Comput. Biol.* *13*, e1005860.

Pichugin, Y., Park, H.J., and Traulsen, A. (2019). Evolution of simple multicellular life cycles in dynamic environments. *J. R. Soc. Interface* 16, 20190054.

Portal-Celhay, C., Bradley, E.R., and Blaser, M.J. (2012). Control of intestinal bacterial proliferation in regulation of lifespan in *Caenorhabditis elegans*. *BMC Microbiol.* 12, 49.

Powell, J.E., Leonard, S.P., Kwong, W.K., Engel, P., and Moran, N.A. (2016). Genome-wide screen identifies host colonization determinants in a bacterial gut symbiont. *Proc. Natl. Acad. Sci.* 201610856.

Prüß, B.M. (2017). Involvement of Two-Component Signaling on Bacterial Motility and Biofilm Development. *J. Bacteriol.* 199.

## Q

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., and Glöckner, F.O. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* 41, D590–D596.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841–842.

## R

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing.

R Core Team (2016). R: A Language and Environment for Statistical Computing.

Rafaluk-Mohr, C., Ashby, B., Dahan, D.A., and King, K.C. (2018). Mutual fitness benefits arise during coevolution in a nematode-defensive microbe model. *Evol. Lett.* 2, 246–256.

Raina, J.-B., Fernandez, V., Lambert, B., Stocker, R., and Seymour, J.R. (2019). The role of microbial motility and chemotaxis in symbiosis. *Nat. Rev. Microbiol.* 17, 284–294.

Rainey, P.B. (1999). Adaptation of *Pseudomonas fluorescens* to the plant rhizosphere. *Environ. Microbiol.* 1, 243–257.

Rainey, P.B., and Travisano, M. (1998). Adaptive radiation in a heterogeneous environment. *Nature* 394, 69–72.

Rakoff-Nahoum, S., Foster, K.R., and Comstock, L.E. (2016). The evolution of cooperation within the gut microbiota. *Nature* 533, 255–259.

Reis, F., Kirsch, R., Pauchet, Y., Bauer, E., Bilz, L.C., Fukumori, K., Fukatsu, T., Kölsch, G., and Kaltenpoth, M. (2020). Bacterial symbionts support larval sap feeding and adult folivory in (semi-)aquatic reed beetles. *Nat. Commun.* 11, 2964.

Remigi, P., Masson-Boivin, C., and Rocha, E.P.C. (2019). Experimental Evolution as a Tool to Investigate Natural Processes and Molecular Functions. *Trends Microbiol.* 27, 623–634.

Rigby, R.A., and Stasinopoulos, D.M. (2005). Generalized additive models for location, scale and shape. *J. R. Stat. Soc. Ser. C Appl. Stat.* *54*, 507–554.

Robinson, C.D., Klein, H.S., Murphy, K.D., Parthasarathy, R., Guillemin, K., and Bohannan, B.J.M. (2018). Experimental bacterial adaptation to the zebrafish gut reveals a primary role for immigration. *PLOS Biol.* *16*, e2006893.

Romero-Gutiérrez, K.J., Dourado, M.N., Garrido, L.M., Olchanheski, L.R., Mano, E.T., Dini-Andreote, F., Valvano, M.A., and Araújo, W.L. (2020). Phenotypic traits of Burkholderia spp. associated with ecological adaptation and plant-host interaction. *Microbiol. Res.* *236*, 126451.

Rosenberg, E., and Zilber-Rosenberg, I. (2016). Microbes Drive Evolution of Animals and Plants: the Hologenome Concept. *MBio* *7*, e01395-15.

RStudio Team (2015). RStudio: Integrated Development for R.

Rueden, C.T., Schindelin, J., Hiner, M.C., DeZonia, B.E., Walter, A.E., Arena, E.T., and Eliceiri, K.W. (2017). ImageJ2: ImageJ for the next generation of scientific image data. *BMC Bioinformatics* *18*, 529.

Russell, S.L. (2019). Transmission mode is associated with environment type and taxa across bacteria-eukaryote symbioses: a systematic review and meta-analysis. *FEMS Microbiol. Lett.* *366*.

Russell, C.W., Bouvaine, S., Newell, P.D., and Douglas, A.E. (2013). Shared Metabolic Pathways in a Coevolved Insect-Bacterial Symbiosis. *Appl. Environ. Microbiol.* *79*, 6117–6123.

Ryan, D.A., Miller, R.M., Lee, K., Neal, S.J., Fagan, K.A., Sengupta, P., and Portman, D.S. (2014). Sex, Age, and Hunger Regulate Behavioral Prioritization through Dynamic Modulation of Chemoreceptor Expression. *Curr. Biol.* *24*, 2509–2517.

## S

Sabrina Pankey, M., Foxall, R.L., Ster, I.M., Perry, L.A., Schuster, B.M., Donner, R.A., Coyle, M., Cooper, V.S., and Whistler, C.A. (2017). Host-selected mutations converging on a global regulator drive an adaptive leap towards symbiosis in bacteria. *ELife* *6*, e24414.

Sachs, J.L., Skophammer, R.G., and Regus, J.U. (2011). Evolutionary transitions in bacterial symbiosis. *Proc. Natl. Acad. Sci.* *108*, 10800–10807.

Sagan, L. (1967). On the origin of mitosing cells. *J. Theor. Biol.* *14*, 225-IN6.

Salem, H., Bauer, E., Kirsch, R., Berasategui, A., Cripps, M., Weiss, B., Koga, R., Fukumori, K., Vogel, H., Fukatsu, T., et al. (2017). Drastic Genome Reduction in an Herbivore's Pectinolytic Symbiont. *Cell* *171*, 1520-1531.e13.

Samuel, B.S., Rowedder, H., Braendle, C., Félix, M.-A., and Ruvkun, G. (2016). *Caenorhabditis elegans* responses to bacteria from its natural habitats. *Proc. Natl. Acad. Sci.* *113*, E3941–E3949.

Sarkar, A., Harty, S., Johnson, K.V.-A., Moeller, A.H., Carmody, R.N., Lehto, S.M., Erdman, S.E., Dunbar, R.I.M., and Burnet, P.W.J. The role of the microbiome in the neurobiology of social behaviour. *Biol. Rev.* *n/a*.

Schick, A., and Kassen, R. (2018). Rapid diversification of *Pseudomonas aeruginosa* in cystic fibrosis lung-like conditions. *Proc. Natl. Acad. Sci.* *115*, 10714–10719.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. *Nat. Methods* *9*, 676–682.

Schloissnig, S., Arumugam, M., Sunagawa, S., Mitreva, M., Tap, J., Zhu, A., Waller, A., Mende, D.R., Kultima, J.R., Martin, J., et al. (2013). Genomic variation landscape of the human gut microbiome. *Nature* *493*, 45–50.

Schlomann, B.H. (2018). Stationary moments, diffusion limits, and extinction times for logistic growth with random catastrophes. *J. Theor. Biol.* *454*, 154–163.

Schlomann, B.H., Wiles, T.J., Wall, E.S., Guillemain, K., and Parthasarathy, R. (2018). Bacterial Cohesion Predicts Spatial Distribution in the Larval Zebrafish Intestine. *Biophys. J.* *115*, 2271–2277.

Schreiber, S.J., Ke, R., Loverdo, C., Park, M., Ahsan, P., and Lloyd-Smith, J.O. (2018). Cross-scale dynamics and the evolutionary emergence of infectious diseases. *BioRxiv* 066688.

Schulenburg, H., and Félix, M.-A. (2017). The Natural Biotic Environment of *Caenorhabditis elegans*. *Genetics* *206*, 55–86.

Schulenburg, V.D., Graf, J.H., Hancock, J.M., Pagnamenta, A., Sloggett, J.J., Majerus, M.E.N., and Hurst, G.D.D. (2001). Extreme Length and Length Variation in the First Ribosomal Internal Transcribed Spacer of Ladybird Beetles (Coleoptera: Coccinellidae). *Mol. Biol. Evol.* *18*, 648–660.

Schulte, R.D., Makus, C., Hasert, B., Michiels, N.K., and Schulenburg, H. (2010). Multiple reciprocal adaptations and rapid genetic change upon experimental coevolution of an animal host and its microbial parasite. *Proc. Natl. Acad. Sci.* *107*, 7359–7364.

Scott, T.A., Quintaneiro, L.M., Norvaisas, P., Lui, P.P., Wilson, M.P., Leung, K.-Y., Herrera-Dominguez, L., Sudiwala, S., Pessia, A., Clayton, P.T., et al. (2017). Host-Microbe Co-metabolism Dictates Cancer Drug Efficacy in *C. elegans*. *Cell* *169*, 442–456.e18.

Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics* *30*, 2068–2069.

Sengupta, P., Chou, J.H., and Bargmann, C.I. (1996). odr-10 Encodes a Seven Transmembrane Domain Olfactory Receptor Required for Responses to the Odorant Diacetyl. *Cell* *84*, 899–909.

Seward, J. (2010). bzip2.

Shapira, M. (2017). Host-microbiota interactions in *Caenorhabditis elegans* and their significance. *Curr. Opin. Microbiol.* *38*, 142–147.

Sheppard, S.K., Guttman, D.S., and Fitzgerald, J.R. (2018). Population genomics of bacterial host adaptation. *Nat. Rev. Genet.* *1*.

Shtonda, B.B., and Avery, L. (2006). Dietary choice behavior in *Caenorhabditis elegans*. *J. Exp. Biol.* *209*, 89–102.



Shukla, S.P., Plata, C., Reichelt, M., Steiger, S., Heckel, D.G., Kaltenpoth, M., Vilcinskas, A., and Vogel, H. (2018a). Microbiome-assisted carrion preservation aids larval development in a burying beetle. *Proc. Natl. Acad. Sci.* *115*, 11274–11279.

Shukla, S.P., Vogel, H., Heckel, D.G., Vilcinskas, A., and Kaltenpoth, M. (2018b). Burying beetles regulate the microbiome of carcasses and use it to transmit a core microbiota to their offspring. *Mol. Ecol.* *27*, 1980–1991.

Sieber, M., Pita, L., Weiland-Bräuer, N., Dirksen, P., Wang, J., Mortzfeld, B., Franzenburg, S., Schmitz, R.A., Baines, J.F., Fraune, S., et al. (2019). Neutrality in the Metaorganism. *PLOS Biol.* *17*, e3000298.

Simão, F.A., Waterhouse, R.M., Ioannidis, P., Kriventseva, E.V., and Zdobnov, E.M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* *31*, 3210–3212.

Singh, J., and Aballay, A. (2019). Intestinal infection regulates behavior and learning via neuroendocrine signaling. *ELife* *8*, e50033.

Smolentseva, O., Gusarov, I., Gautier, L., Shamovsky, I., DeFrancesco, A.S., Losick, R., and Nudler, E. (2017). Mechanism of biofilm-mediated stress resistance and lifespan extension in *C. elegans*. *Sci. Rep.* *7*.

Sonnenburg, J.L., and Sonnenburg, E.D. (2019). Vulnerability of the industrialized microbiota. *Science* *366*.

Soto, W., Punke, E.B., and Nishiguchi, M.K. (2012). Evolutionary perspectives in mutualism of sepiolid squid and bioluminescent bacteria: combined usage of microbial experimental evolution and temporal population genetics. *Evol. Int. J. Org. Evol.* *66*, 1308–1321.

Soto, W., Rivera, F.M., and Nishiguchi, M.K. (2014). Ecological Diversification of *Vibrio fischeri* Serially Passaged for 500 Generations in Novel Squid Host *Euprymna tasmanica*. *Microb. Ecol.* *67*, 700–721.

Sousa, A., Frazão, N., Ramiro, R.S., and Gordo, I. (2017). Evolution of commensal bacteria in the intestinal tract of mice. *Curr. Opin. Microbiol.* *38*, 114–121.

Stadler, B., and Dixon, A.F.G. (2005). Ecology and Evolution of Aphid-Ant Interactions. *Annu. Rev. Ecol. Evol. Syst.* *36*, 345–372.

Starkey, M., Hickman, J.H., Ma, L., Zhang, N., Long, S.D., Hinz, A., Palacios, S., Manoil, C., Kirisits, M.J., Starner, T.D., et al. (2009). *Pseudomonas aeruginosa* Rugose Small-Colony Variants Have Adaptations That Likely Promote Persistence in the Cystic Fibrosis Lung. *J. Bacteriol.* *191*, 3492–3503.

Stecher, B., and Hardt, W.-D. (2008). The role of microbiota in infectious disease. *Trends Microbiol.* *16*, 107–114.

Stephens, W.Z., Wiles, T.J., Martinez, E.S., Jemielita, M., Burns, A.R., Parthasarathy, R., Bohannon, B.J.M., and Guillemin, K. (2015). Identification of Population Bottlenecks and Colonization Factors during Assembly of Bacterial Communities within the Zebrafish Intestine. *MBio* *6*.

- Stiernagle, T. (2006). Maintenance of *C. elegans*. WormBook.
- Stott, I., Townley, S., and Hodgson, D.J. (2011). A framework for studying transient dynamics of population projection matrix models. *Ecol. Lett.* *14*, 959–970.
- Styrsky, J.D. (2014). An orb-weaver spider exploits an ant–acacia mutualism for enemy-free space. *Ecol. Evol.* *4*, 276–283.
- Sullam, P.M., Bayer, A.S., Foss, W.M., and Cheung, A.L. (1996). Diminished platelet binding in vitro by *Staphylococcus aureus* is associated with reduced virulence in a rabbit model of infective endocarditis. *Infect. Immun.* *64*, 4915–4921.
- Sun, Q., and Gao, B. (2020). Programming animal physiology and behaviors through engineered bacteria. *BioRxiv* 2020.08.15.232637.
- Suzuki, R., and Shimodaira, H. (2006). Pvcust: an R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics* *22*, 1540–1542.

## T

- Tange, O. (2011). GNU parallel—the command-line power tool. *USENIX Magazine* 42–47.
- The Human Microbiome Project Consortium, Huttenhower, C., Gevers, D., Knight, R., Abubucker, S., Badger, J.H., Chinwalla, A.T., Creasy, H.H., Earl, A.M., FitzGerald, M.G., et al. (2012). Structure, function and diversity of the healthy human microbiome. *Nature* *486*, 207–214.
- Theis, K.R., Dheilly, N.M., Klassen, J.L., Brucker, R.M., Baines, J.F., Bosch, T.C.G., Cryan, J.F., Gilbert, S.F., Goodnight, C.J., Lloyd, E.A., et al. (2016). Getting the Hologenome Concept Right: an Eco-Evolutionary Framework for Hosts and Their Microbiomes. *MSystems* *1*.
- Thomas, T., Moitinho-Silva, L., Lurgi, M., Björk, J.R., Easson, C., Astudillo-García, C., Olson, J.B., Erwin, P.M., López-Legentil, S., Luter, H., et al. (2016). Diversity, structure and convergent evolution of the global sponge microbiome. *Nat. Commun.* *7*, 11870.
- Thutupalli, S., Uppaluri, S., Constable, G.W.A., Levin, S.A., Stone, H.A., Tarnita, C.E., and Brangwynne, C.P. (2017). Farming and public goods production in *Caenorhabditis elegans* populations. *Proc. Natl. Acad. Sci.* 201608961.
- Tienderen, P.H. van (2000). Elasticities and the link between demographic and evolutionary dynamics. *Ecology* *81*, 666–679.
- Traverse, C.C., Mayo-Smith, L.M., Poltak, S.R., and Cooper, V.S. (2013). Tangled bank of experimentally evolved *Burkholderia* biofilms reflects selection during chronic infections. *Proc. Natl. Acad. Sci.* *110*, E250–E259.

## V

- Vega, N.M., and Gore, J. (2017). Stochastic assembly produces heterogeneous communities in the *Caenorhabditis elegans* intestine. *PLOS Biol.* *15*, e2000633.

Vert, M., Doi, Y., Hellwich, K.-H., Hess, M., Hodge, P., Kubisa, P., Rinaudo, M., and Schué, F. (2012). Terminology for biorelated polymers and applications (IUPAC Recommendations 2012). *Pure Appl. Chem.* *84*, 377–410.

Virk, B., Jia, J., Maynard, C.A., Raimundo, A., Lefebvre, J., Richards, S.A., Chetina, N., Liang, Y., Helliwell, N., Cipinska, M., et al. (2016). Folate Acts in *E. coli* to Accelerate *C. elegans* Aging Independently of Bacterial Biosynthesis. *Cell Rep.* *14*, 1611–1620.

Visick, K.L. (2009). An intricate network of regulators controls biofilm formation and colonization by *Vibrio fischeri*. *Mol. Microbiol.* *74*, 782–789.

de Vos, W.M. (2015). Microbial biofilms and the human intestinal microbiome. *Npj Biofilms Microbiomes* *1*, 1–3.

Vu, V. (2020). *vqv/ggbiplot*.

## W

Wallenstein, M.D., and Hall, E.K. (2012). A trait-based framework for predicting when and where microbial adaptation to climate change will affect ecosystem functioning. *Biogeochemistry* *109*, 35–47.

Wang, Y., and Rozen, D.E. (2017). Gut Microbiota Colonization and Transmission in the Burying Beetle *Nicrophorus vespilloides* throughout Development. *Appl. Environ. Microbiol.* *83*.

Wang, G.-H., Berdy, B.M., Velasquez, O., Jovanovic, N., Alkhalifa, S., Minbiole, K.P.C., and Brucker, R.M. (2020). Changes in Microbiome Confer Multigenerational Host Resistance after Sub-toxic Pesticide Exposure. *Cell Host Microbe* *0*.

Warnhoff, K., and Ruvkun, G. (2019). Molybdenum cofactor transfer from bacteria to nematode mediates sulfite detoxification. *Nat. Chem. Biol.* *15*, 480–488.

Watson, E., MacNeil, L.T., Ritter, A.D., Yilmaz, L.S., Rosebrock, A.P., Caudy, A.A., and Walhout, A.J.M. (2014). Interspecies Systems Biology Uncovers Metabolites Affecting *C. elegans* Gene Expression and Life History Traits. *Cell* *156*, 759–770.

Watson, E., Olin-Sandoval, V., Hoy, M.J., Li, C.-H., Louise, T., Yao, V., Mori, A., Holdorf, A.D., Troyanskaya, O.G., Ralser, M., et al. (2016). Metabolic network rewiring of propionate flux compensates vitamin B12 deficiency in *C. elegans*. *ELife* *5*, e17670.

Wein, T., and Dagan, T. (2019). The Effect of Population Bottleneck Size and Selective Regime on Genetic Diversity and Evolvability in Bacteria. *Genome Biol. Evol.* *11*, 3283–3290.

Wein, T., Picazo, D.R., Blow, F., Woehle, C., Jami, E., Reusch, T.B.H., Martin, W.F., and Dagan, T. (2019). Currency, Exchange, and Inheritance in the Evolution of Symbiosis. *Trends Microbiol.* *27*, 836–849.

Werren, J.H., Baldo, L., and Clark, M.E. (2008). Wolbachia: master manipulators of invertebrate biology. *Nat. Rev. Microbiol.* *6*, 741–751.

Westoby, M. (1998). A leaf-height-seed (LHS) plant ecology strategy scheme. *Plant Soil* *199*, 213–227.

- Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*. 2009.
- Wickham, H. (2016). *ggplot: elegant graphics for data analysis*. Springer- Verlag.
- Wiles, T.J., Jemielita, M., Baker, R.P., Schlomann, B.H., Logan, S.L., Ganz, J., Melancon, E., Eisen, J.S., Guillemin, K., and Parthasarathy, R. (2016). Host Gut Motility Promotes Competitive Exclusion within a Model Intestinal Microbiota. *PLOS Biol.* *14*, e1002517.
- Wiles, T.J., Wall, E.S., Schlomann, B.H., Hay, E.A., Parthasarathy, R., and Guillemin, K. (2018). Modernized Tools for Streamlined Genetic Manipulation and Comparative Study of Wild and Diverse Proteobacterial Lineages. *MBio* *9*.
- Winans, N.J., Walter, A., Chouaia, B., Chaston, J.M., Douglas, A.E., and Newell, P.D. (2017). A genomic investigation of ecological differentiation between free-living and *Drosophila*-associated bacteria. *Mol. Ecol.* *26*, 4536–4550.
- Winsor, G.L., Griffiths, E.J., Lo, R., Dhillon, B.K., Shay, J.A., and Brinkman, F.S.L. (2016). Enhanced annotations and features for comparing thousands of *Pseudomonas* genomes in the *Pseudomonas* genome database. *Nucleic Acids Res.* *44*, D646–D653.
- Wong, C.N.A., Ng, P., and Douglas, A.E. (2011). Low-diversity bacterial community in the gut of the fruitfly *Drosophila melanogaster*. *Environ. Microbiol.* *13*, 1889–1900.

## X

- Xue, B., and Leibler, S. (2016). Evolutionary learning of adaptation to varying environments through a transgenerational feedback. *Proc. Natl. Acad. Sci.* *113*, 11266–11271.

## Y

- Yang, W., Petersen, C., Pees, B., Zimmermann, J., Waschina, S., Dirksen, P., Rosenstiel, P., Tholey, A., Leippe, M., Dierking, K., et al. (2019). The Inducible Response of the Nematode *Caenorhabditis elegans* to Members of Its Natural Microbiota Across Development and Adult Life. *Front. Microbiol.* *10*.
- Yildiz, F.H., and Schoolnik, G.K. (1999). *Vibrio cholerae* O1 El Tor: Identification of a gene cluster required for the rugose colony type, exopolysaccharide production, chlorine resistance, and biofilm formation. *Proc. Natl. Acad. Sci.* *96*, 4028–4033.

## Z

- Zhang, X.-X., and Rainey, P.B. (2010). Bet hedging in the underworld. *Genome Biol.* *11*, 137.
- Zhang, F., Berg, M., Dierking, K., Félix, M.-A., Shapira, M., Samuel, B.S., and Schulenburg, H. (2017). *Caenorhabditis elegans* as a Model for Microbiome Research. *Front. Microbiol.* *8*.
- Zhang, W., Zhang, X., Su, Q., Tang, M., Zheng, H., and Zhou, X. (2020). Genomic features underlying the evolutionary transitions of *Apibacter* to honey bee gut symbionts. *BioRxiv* 2020.09.30.321786.

Zhang, Y., Chou, J.H., Bradley, J., Bargmann, C.I., and Zinn, K. (1997). The *Caenorhabditis elegans* seven-transmembrane protein ODR-10 functions as an odorant receptor in mammalian cells. *Proc. Natl. Acad. Sci.* *94*, 12162–12167.

Zhu, L., Wu, Q., Dai, J., Zhang, S., and Wei, F. (2011). Evidence of cellulose metabolism by the giant panda gut microbiome. *Proc. Natl. Acad. Sci.* *108*, 17714–17719.

Zilber-Rosenberg, I., and Rosenberg, E. (2008). Role of microorganisms in the evolution of animals and plants: the hologenome theory of evolution. *FEMS Microbiol. Rev.* *32*, 723–735.

Zimmermann, J., Obeng, N., Yang, W., Pees, B., Petersen, C., Waschina, S., Kissoyan, K.A., Aidley, J., Hoepfner, M.P., Bunk, B., et al. (2019). The functional repertoire contained within the native microbiota of the model nematode *Caenorhabditis elegans*. *ISME J.* 1–13.

Zomorodi, A.R., and Segre, D. (2017). Intracellular metabolic circuits shape inter-species microbial interactions. *BioRxiv* 127332.

(2020). *cran/missMDA (cran)*.



## Curriculum vitae

### Personal information

Date of birth: Dec. 16th, 1991  
Place of birth: Hamburg, Germany  
Nationality: German



### Education

10/2016 – 12/2020	<b>Christian-Albrechts-Universität zu Kiel</b> <b>Int'l Max Planck Research School for Evolutionary Biology</b> Phd studies	Kiel, DE
09/2014 – 09/2016	<b>Rijksuniversiteit Groningen</b> MSc Ecology and Evolution	Groningen, NL
04/2016 – 08/2016	<b>Ludwig-Maximilians-Universität</b> MSc thesis; ERASMUS research project	Munich, DE
09/2013 – 05/2014	<b>University of Toronto</b> BSc study abroad	Toronto, CA
04/2011 – 08/2014	<b>Freie Universität Berlin</b> BSc in Biology	Berlin, DE

### Publications

Spriewald S., Stadler E., Hense B. A., Münch P. C., McHardy A. C., Weiss A. S., **Obeng N.**, Müller, J., Stecher B. (2020). Evolutionary Stabilization of Cooperative Toxin Production through a Bacterium-Plasmid-Phage Interplay. *mBio*. 11 (4) e00912-20

Zimmermann, J.\*, **Obeng, N.\***, Yang, W., Pees, B., Petersen, C., Waschina, S., Kissoyan, K. A., Aidley, J., Hoepfner, M. P., Bunk, B., Spröer, C., Leippe, M., Dierking, K., Kaleta, C., Schulenburg, H. The functional repertoire contained within the native microbiota of the model nematode *Caenorhabditis elegans*. *ISME J.* 1–13 (2019). doi:10.1038/s41396-019-0504- \*shared first authorship; chapter 1 of this thesis

Engerer, P., Suzuki, S.C., Yoshimatsu, T., Chapouton, P., **Obeng, N.**, Odermatt, B., Williams, P. R., Misgeld, T., Godinho, L. (2017). Uncoupling of neurogenesis and differentiation during retinal development. *EMBO J.* 36, 1134–1146

**Obeng, N.**, Pratama, A. A., van Elsas, J. D. (2016). The Significance of Mutualistic Phages for Bacterial Ecology and Evolution. *Trends in Microbiology*. 10.1016/j.tim.2015.12.009