Adaptive strategies of *Enterococcus mundtii* to different living conditions in the gut microbiome of *Spodoptera littoralis* larvae

Dissertation

To Fulfill the
Requirements for the Degree of „doctor rerum naturalium“
(Dr. rer. nat.)

Submitted to the Council of the Faculty
of Biological Sciences
of the Friedrich Schiller University Jena

By
Tilottama Mazumdar,
Masters in Biotechnology,
born on 17.04.1992 in
India

Reviewers

1. Prof. Dr. Wilhelm Boland,
   Department of Bioorganic Chemistry, Max Planck Institute for Chemical Ecology
2. Prof. Dr. Erika Kothe, Institute of Microbiology, Friedrich Schiller University
3. Prof. Dr. David Heckel, Department of Entomology, Max Planck Institute for Chemical Ecology
4. Dr. Mark S Gresnigt, Leibniz Institute for Natural Product Research and Infection Biology – Hans Knöll Institute (HKI)
5. Prof. Dr. Dirk Hoffmeister, Leibniz Institute for Natural Product Research and Infection Biology – Hans Knöll Institute (HKI)
6. Prof. Dr. Dino McMahon, Institute of Biology – Zoology, Freie Universität Berlin

Date of Defense- 19th November, 2020

"We are all of us walking communities of bacteria. The world shimmers, a pointillist landscape made of tiny living beings"

--- Lynn Marguilis, Microcosmos: Four Billion Years of Microbial Evolution, 1986

"We can allow satellites, planets, suns, universe, nay whole systems of universe, to be governed by laws, but the smallest insect, we wish to be created at once by special act."

--- Charles Darwin, Darwin's religious odessey, 2002

"Science cannot solve the ultimate mystery of nature. And that is because, in the last analysis, we ourselves are a part of the mystery that we are trying to solve."

--- Max Planck, Where is Science going? , 1981

Contents

1 Introduction

Arthropods constitute the largest and the most diverse phylum in the animal kingdom. They have all evolved from a single ancestor and exhibit certain common characteristics: Bilateral symmetry, segmented body, a hard exoskeleton, jointed appendages and many pairs of limbs. They have gained an evolutionary success owing to their diversity and ability to survive in a wide range of habitats. What adds up to their success is their small size, ability to conserve water by excreting pellets of uric acid instead of urea, less food requirement, ability to fly and evade danger, short generation times, and capability of withstanding hardships with their tough exoskeletons. Humans have categorized insects into beneficial ones and pests. Only 1% of all insects in the world are pests [1]. Pests have a record of causing several diseases like chagas disease (*Triatoma* sp), dengue (*Aedes aegypti*), malaria (*Anopheles* sp), filariasis (*Anopheles* sp, *Aedes aegypti*) or typus (*Pediculus humanus humanus*). Pests are known to cause crop and forest loses. For example, the European corn borer, a lepidopteran causes great crop losses in the United States. What is pest from a human perspective can be beneficial in a different circumstance. A housefly can be a detritivorous. Termites, although are pests at home have ecological roles in forests. Most flowering plants require pollinators, and among 200000 pollinators, vast majority are insects [2]. Apart from a few very specific pollinator-flower relationships, most of the insect pollinators are generalists. Generalist lepidopteran pollinators are more common than specialists like Yucca moth and Yucca flowers [3].

Insects are not lone players. They have a community of microorganisms inhabiting their guts. Several of the microorganisms play with the fitness and well-being of the host. It could be very much possible that the success of arthropods is dependent on their resident gut microbiome. Hence, the importance of investigating deeper into this topic is necessary. Since years scientists are trying to study as many insect-bacterial pairs as possible that have resulted in volumes of research on this topic.

This thesis is a contribution to such a pool of insect-symbiont research. In the first part it focuses on the survival strategies of the dominant gut symbiont *Enterococcus mundtii* in the gut of *Spodoptera littoralis* larvae. The second part investigates how the host genetics is shaping the bacterial population in the gut of the same.

1.1 Symbiosis: How it all started

The complex interaction of the community of microorganisms inhabiting within higher animals has been a subject of immense interest among researchers. This interaction precisely defines the phenomenon of symbiosis which is nothing but the coexistence of two or more living beings at close proximity, or within each other, bringing about a beneficial, harmful or neutral consequence for the partners. Coined by a German Biologist, Anton de Bary in the year 1879, "Symbiosis" has its history dating back to the times of oxygenation of the atmosphere, 1.2 billion years ago [4], [5].

Endosymbiosis is the phenomenon that led the foundation of the association between a two cell, where one resides inside the other. It dates back to 1.2 billion years ago, before the dawn of the Cambrian era. Up to a billion years before that, the environment was anaerobic. Photo disassociation of water resulted in formation of reactive oxygen species, letting only the metal chelated porphyrins containing cells to survive the oxidative damage (coenzymes of peroxidase and catalase). On one hand, these photoautotrophic cells, with the help of solar energy absorbing, chlorophyll like porphyrins, began producing ATP, on the other hand, heterotrophic microbes that efficiently fermented carbohydrates using their porphyrins were selected. This resulted in a gradual oxidization of the atmosphere due to the phenomenon of photosynthesis that became rampant. Ultraviolet light that helped in abiogenic production was absorbed by the ozone. Heterotrophs were forced to consume what the photoautotrophs produced. This built the foundation for the evolution of mitotic eukaryotes from prokaryotes. Surviving and replicating was the order of the time to prevent one´s own extinction. An anaerobic eukaryotic cell ingested an aerobic prokaryotic microbe, presently called the photomitochondrion, resulting in the very first evidence of obligate symbiosis [5].

Such is the extent of involvement of the symbionts in health and well-being of their hosts, scientists wondered if they are also involved in the latter's evolution. The bacterial flora might as well be called the "forgotten organ" [6]. Thus came up the "Hologenome theory of Evolution" coined by Rosenberg et al [7].

1.2 Hologenome theory of Evolution

Hologenome entails the total genetic information of the host and the bacteria. This altogether is a single dynamic entity and the unit of selection in evolution [7].

In several cases, the amount of this information in the symbionts exceeds that of the host. An average person approximately contains 5 million bacterial genes, in addition to their own 20500 genes. This community of microorganisms not only is intimately associated with the host along generations, but also, an alteration in the genomes of one or more interacting partners leads to evolution of the holobiont. Also, if an autotroph could have such a drastic effect in the fate of a eukaryotic cell, it should definitely play a substantial role in the evolution of higher organisms, now that it contributes volumes of information to the host´s already existing information, and has a profound impact on the fitness of the host [7].

In several well studied systems, it seems that evolution has led to a total dependence between the host and its associated microorganisms. A spectrum of associations lies between such obligatory interactions and the ones that have little or no effect on the evolution of interacting partners [7].

Transmission of the flora happens by two distinct mechanisms: vertical transmission, where the offspring inherits the symbionts from its mother. This is possible for the organisms that have made their way via the trans-ovarian route of the mother, eg, Aphid-*Buchnera* symbiosis. Horizontally transferred flora is the one that is acquired by the off springs in the process of growth. This occurs when the mother is tending to her off springs through physical contact and feeding them food covered with feces and sputum, such as in rumen gut microbiota. There are several intermediate routes of transmission as well. The very first endosymbiosis leading to the ingestion of mitochondria and chloroplast is a classic example of cytoplasmic inheritance. In case of plants, the seed that falls on the ground acquires the rhizosphere flora. During vegetative reproduction, for example in plants and hydra, obtaining the flora from the environment is the only way to go. Therefore, no matter what the mode of transmission is, individuals can acquire and transmit microbiota throughout life and the holobiont in the next generation is affected by the individuals that one is in close contact with [7].

As already mentioned, variations either in the host or the associated symbionts leads to evolution. Genetic differences in host occurs during sexual reproduction, mutations and

chromosomal rearrangements, whereas that of the symbionts occur by transformation, transduction, recombination, microbial amplification, horizontal gene transfer, as well as recruiting new strains in the community. The hologenome theory also points out the enormity of the host genome that takes a longer time to evolve under rapidly changing environmental conditions. So, if it was left alone to evolve, without its associated microbial genome coming to aid, it would lose its competitiveness, bringing it closer to extinction. Thus, the hologenome theory takes into account aspects of both Darwinism and Lamarckism, where the individuals evolve by selecting random variants and holobiont evolves by adapting to the environmental changes [7].

1.2.1 Controversies about the hologenome theory

The holobiont theory is highly controversial. Several researchers coined the phrase "Holes in the hologenome", stating that it is nothing but a broad generalization taking into account only certain kind of interactions [8], [9]. Although intimate associations bear a striking semblance to coevolution, but the two concepts do not always go hand in hand. Obligate relationships between host and a particular symbiont do not always guarantee a shared evolutionary history [9], [8]. The fact that the host and the symbiont are evolving no way means that they are evolving in response to one another. It is quite possible that a host is tolerant only to a particular subset of microorganisms, excluding the rest, without coevolving with them. Cases of obligate symbiosis like mitochondrion and chloroplast, or aphid- *Buchnera* where the indispensability for survival is maintained, and the symbiont genetics is inherited by the host as a part of its own, leads one to easily assume the hologenome being the unit of selection. But, one should not forget that such high-fidelity associations are not only rare, but also applies for a certain pair of host-microbe, and not the entire microbiome. Symbiont fitness and longevity does depend on that of the host, and although there is room for the hologenome to evolve together, but it is insignificant as compared to the two interacting genomes individually. A certain community of microorganisms colonizing a niche does not always mean that they are evolving to develop that niche, but they are learning to depend on the part of the environment they are in. Thus, the hologenome theory of evolution only takes care of "mutualism" type of interactions, excluding the cases where one partner poses as an antagonist to the other [9].

1.3 Colonization factors that influence the gut microbial population

Factors that determine the colonization of a certain community of microorganisms in the gut over others are very many. A stable microbial colony within a higher organism is a result of co adaptation over millennia. This has led to the evolution of molecular mechanisms to establish and maintain this host-symbiont association. In a particular study it was proven that the presence of the ccf pathway in bacteroidetes allowed its efficient colonization in the gastro intestinal tract of humans. The role of CCF proteins was suggested to be uptake and usage of glycans. Thus, they are able to utilize the dietary glycans more efficiently than others, outgrowing them and inducing a colonization resistance [10]. Microbiome studies on chickens show that both host and environmental factors shape the microbiota. Host factors like age, sex and breed; and environmental factors such as biosecurity levels, feed access, litter, and the weather have drastic effects on the gut flora [11].

The vertically transmitted microbes are acquired during passage through the birth canal or ingestion of egg shells by off springs after they hatch. The microbial composition of a pair of fraternal twins is strikingly similar as compared to infants from different parents, suggesting the influence of the reproductive tract. Also, the microbial composition of the twins remained similar temporally, suggesting the influence of the environment that an organism is exposed to [12]. Diet has a long term effect on the gut microbiota. Studies with infants showed that the ones fed with breast milk have a different microbial composition as compared to the formula fed ones. While the former has a composition of Bifidobacterium, Lactobacilli, Staphylococci, Bacteroidetes; the latter have an abundance of Proteobacteria, Clostridia and Bifidobacteria setting in late in their life cycle [13], [14]. Also, links between of Bifidobacteria with healthy weight gain, along with a predominance of clostridia and Enterococci in obese humans have been found [15].

Host organisms are equipped with efficient ways of distinguishing between commensals and pathogens. Extensive studies on microbial recognition have been performed on *Drosophila melanogaster* system. The innate immunity of Drosophila consists of pathway cascades: Toll and Imd (Immunodeficiency) pathways. The Toll pathway recognizes gram-positive bacteria and fungi, whereas the latter has recognition for the diaminopimelic acid component of the gram negative bacterial cell wall. Both the pathways produce Antimicrobial peptides via nuclear translocation of nuclear factor NF-kB. Apart from this, the reactive oxygen species (ROS) also have antimicrobial activity. It generates dual oxidase (duox) as a response to

pathogen derived uracil, which in turn activates G protein coupled receptor (GPCR) and helps releasing calcium from endoplasmic reticulum. In *D. melanogaster*, Toll pathways are more prominent in fore and hind guts, whereas the midgut, the hub of digestive activities houses the Imd pathway. There are different AMPs prevalent in different parts of the gut, along with toxins (pore-forming toxins) and dual oxidases [16], [17]. Every organism needs to achieve tolerance towards their own commensals. In order to distinguish between commensals and pathogens, *D. melanogaster* achieves immune tolerance towards its commensals by manipulating its IMD pathway. The homeobox transcription factor Caudal represses the transcription of AMP genes by binding to their promoters. In flies deficient in Caudal, shifts in microbiota has been noticed which leads to the disintegration of the epithelial layer. Another strategy by drosophila is to employ amidases that renders pro-inflammatory PGNs origination due to commensal bacteria inactive. This indirectly also modulates the IMD pathway in the favor of the commensals [18]. The DUOX system is also subject to manipulation. It is inhibited by MKP3 in presence of commensal bacteria. The interplay of the two synergistic pathways of AMPs and ROS production seem to be fine tuned to an extent where they can differentiate between commensal microbes and deleterious pathogens.

In mammals the gastrointestinal tract (GIT) consists of about 10^{14} bacteria, whose colonization starts before birth. The mucosal immune system functions differently from the systemic immune system. It undergoes several changes after bacterial colonization. It is trained by the commensal bacteria to differentiate between its own kind of bacteria and pathogens. In their complete eubiotic and stable state, they occupy the mucosal surfaces and exercise a colonization resistance towards potential pathogens. The latter basically lose the competition for adhesion space and nutrients. Indirectly, they promote a cascade of immune response as described before. The GIT microbiota plays active roles in shaping the host immune systems. Several such cases have been studied in humans. The microbiota that had previously dealt with rotaviruses by releasing secretory IgA, also counteracts the cholera toxin from *Clostridium difficle* infection [19]. Some pathogens can evade the mucosal barrier causing dysbiosis. *Porphyromonas gingivalis* escapes from the TLR2 signalling of the host. Some viruses can escape the host TLR4 cells leading to dysbiosis and inflammation [20].

1.3.1 Colonization factors in insect guts

As mentioned before, insects have evolved efficient strategies to keep a certain population as their own, while discarding the rest. *Galleria mellonella* jointly with its gut resident *Enterococcus mundtii* controls its own gut population. The host with its lysozyme, and *E. mundtii* with its mundticin together act to maintain a healthy gut population in *G. mellonella*, and if either one is knocked down; there is a predominance of pathogenic strains of serratia and staphylococci. Also, the hosts survived much longer with both the lysozyme and mundticin in action [21], [22]. The host physiology has important roles to play in determining the gut microbial population. The basic structure of the gut is mainly fore, mid and hindguts. The foregut has temporary storage for food, the midgut being the site for digestion and absorption of nutrients, whereas the hindgut sometimes houses fermentation chambers and holds feces before defecation. The peritrophic space divides the midgut into endo and ecto peritrophic spaces, and the microbes dwell in the former. The shape of the gut, whether it is a single tube like structure, as in the case of lepidoptera could have a lesser diversity of microbes as compared to the guts that have crypts, paunches, caeca as with several other arthropod families like coleoptera and hemiptera. The latter kind of structures allows microbes to persist. The hindgut of wood feeding termites has an enlarged paunch at its hindgut where bacteria and protozoa have formed a niche to contribute to oxygen, nitrogen and energy requirements for the host. Structural metamorphosis brought about by holometabolism- where the insects reorganizes their gut structure with their changing life stages from larvae, pupa and adult; moulting or shedding of the exoskeleton; and renewal of the peritrophic matrix leads to a complete elimination of the associated gut bacteria. Only at the adult stages of the insects, when they attain a final physiological form, they harbour a stable microbial population [22].

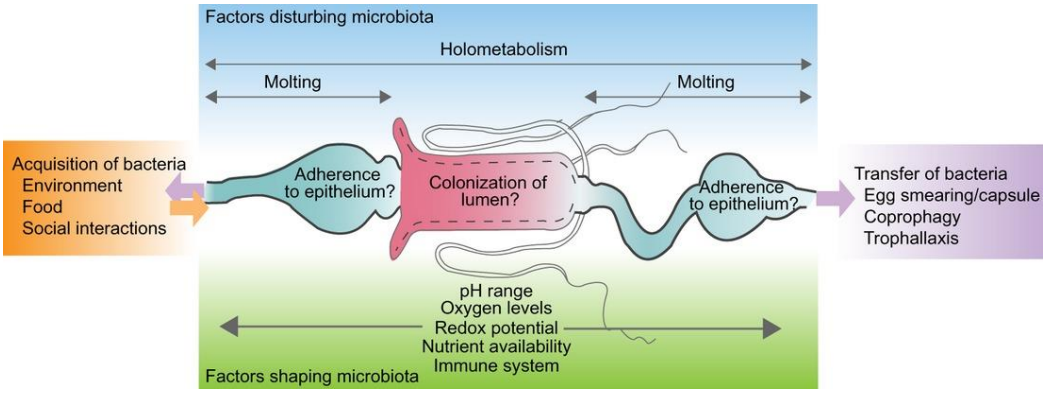

Figure 1: Factors influencing microbial colonization in the host gut: environment, diet and social interactions, physiological conditions, insect development and finally, vertical transmission to the next generation [23].

Oxygen dependence of the microbes determines whether they are colonizing inside the gut walls or in the lumen of the intestine. Anaerobic species of Clostridia occupy niches within the gut walls of *S. littoralis*, whereas facultative anaerobes like *E. mundtii* prefer to colonize the gut lumen of the larvae [24]. It starts getting anoxic towards the vicinity of the gut wall. *Clostridia* sp, an anaerobe is found within 50 µm inside the gut wall [25]. It is because of a tubular gut structure and high food throughput, there seems to be lesser roles of gut flora in digestion and nutrition provisioning of the hosts. The time taken for food throughput is also an important factor. In insects, where the major population of the gut microbiome resides in the lumen, the doubling time of the bacteria should exceed the time it takes for the food to be excreted out of the gut, to maintain a stable population [23].

pH and redox potential also selects for certain microbes over others. Although most bacteria prefer a neutral pH for growth, several also grow in high and low pH conditions in insect guts. Owing to a tannin rich diet, lepidopterans show a marked pH gradient starting from alkaline in the foregut to neutral in the hindgut. Some soil-feeding termites have extreme alkaline conditions in their guts [23]. Only microbes that are alkaline-tolerant like Firmicutes, Clostridia can grow in these compartments [26], [27]. The community generally differs widely among different compartments of such insect-guts. A more uniform spatial bacterial population can be found in non holometabolous insects without any pH gradient with the exception of termites [28]. In another example of *Pachnoda ephippiata*, the pH shifts from 8 in the foregut to 10 in the midgut and then back to 7 in the hindgut. It is the hindgut that harbors most of the bacterial population [29].

Social behaviour among families like ants, bees, wasps and termites show lots of microbial exchange by acts of tending, trophallaxis or coprophagy as compared to solitary insects where interaction happens only during mating [23]. Behavioral differences in social insects arise since different partners are endowed with different tasks. These behavioral tasks are also associated with various interacting factors like age, sex, environment and diet. The variation in gut microbiota taking all these factors into consideration has been studied in detail in honey bees. Worker bees responsible for tasks within the hive such as nursing and feeding are young and feed on pollen rich diets. Foragers who collect resources outside the colonies are older in age and feed on a protein rich diet. The worker bees have a higher relative abundance of the core bacteria, *Lactobacillus mellis*, and bifidobacteriaceae than the foragers. Both Lactobacilli and Bifidobacteria are known to process complex carbohydrates [30].

1.4 Lepidoptera and their gut microbiota

Lepidopterans make up the second most diverse insect order. They metamorphose in a holometabolous manner, meaning they have distinct physiologies in their larval, pupal and adult stages of life cycle. Their gut structure undergoes a complete reorganisation, reshedding the walls and regrowing them. Also, they have an alkaline gut. Such conditions might limit the prospects of microbes to establish a successful colonization. The gut flora has been described as transient. There also exists a dynamic nature of the gut population depending not only on the stage the lepidoptera is in, but also the environment they live in, the diet they are on since there exists both herbivorous and carnivorous caterpillars while the adult species mostly feed on nectar, and sociality, where many lepidopterans do not engage in interaction with each other except for mating [23]. *Spodoptera littoralis* and *Spodoptera frugiperda* are generalist herbivores while certain Lycaenidae larvae could be carnivorous. Also, these Lycaenidae caterpillars undergo a transition in their bacterial population when they shift from solitary to a parasitic lifestyle in their fourth larval instar. Despite the dynamicity, certain taxas do persist throughout the life cycle. Acetobacteraceae, Moraxellaceae in *Heliconius erato* and *Enterococcus mundtii* in *Spodoptera littoralis*. Their gut structure is a simple tubular structure without any intricate pouches for the bacteria to form niche, another reason behind low persistence of most bacteria in a lepidopteran gut [31].

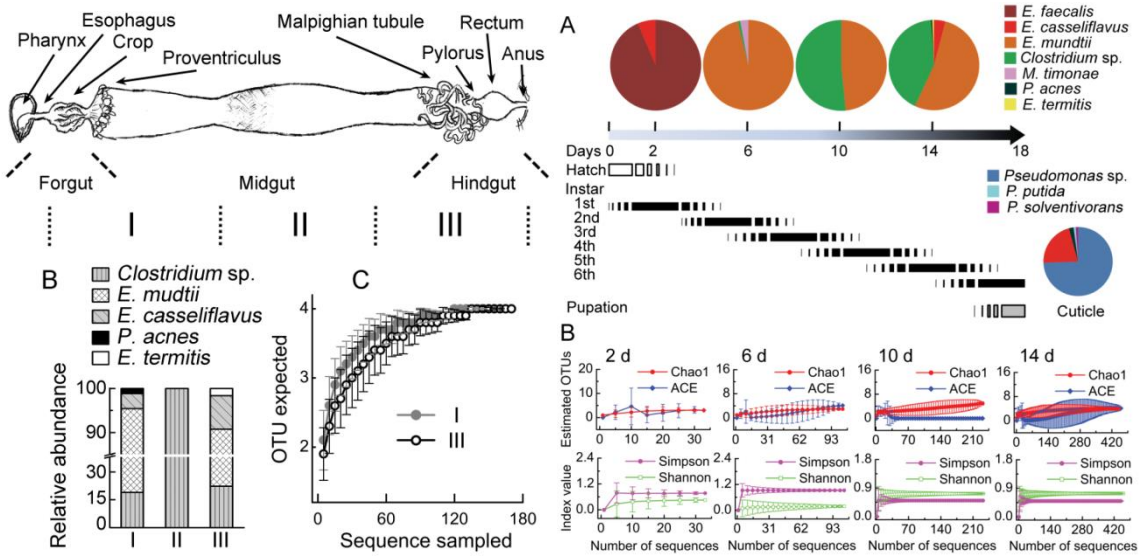

Figure 2: Spatial and temporal variation of gut bacteria in the *Spodoptera littoralis* larva, respectively. Spatially, *Clostridia* sp dominate the midgut, while temporally, *Enterococcus mundtii* and *Clostridia* sp dominate and persist in the gut along the larval stages of the insect [25].

Lepidopterans transmit their symbionts both horizontally and vertically, although scarce evidence of vertical transmissions has been reported. Vertical transmission in *Galleria mellonella* possibly happens by the symbiotic transfer via oocytes [32]. *Enterococcus mundtii*, the dominating gut bacteria of *S. littoralis* also show signs of being vertically transmitted since they were detected in the eggs of the insects. Metabolic activities of Enterococci inside the eggs have been confirmed in *Manduca Sexta* [33]. The neonates after hatching ingest the egg shells to acquire the symbionts. Enterococci, which are ubiquitous in nature, are also a dominating group in several lepidopterans. This could also hint towards them being horizontally transmitted [31].

Several of present day lepidopterans lack cellulose digesting enzymes, even while being on a herbivorous diet, indicating that the symbionts might be helping them. One common example is *P. xylostella*. Bacteria present in the larval and pupal stages of *Automeris zugana* and *Rothschildia lebeau* provide enzymes like gelatinase, caseinase, lipase, esterase, and chitinase activity to their hosts [34]. The ones present in pupae could possibly take part in the insect development and cuticle formation. Most lepidopterans chew leaves of plants, the phloem of which is devoid of nitrogen. Most of the dietary nitrogen for these insects must be coming

from their gut bacteria. Evidence of nitrogen fixing bacteria has been found in *Plutella xylostella*[35].

Besides helping in nutrient acquisition, gut bacteria of lepidopterans also play roles of defensive symbionts. *Enterococcus mundtii* in the gut of *Spodoptera littoralis* produces a class IIa bacteriocin called mundticiin. This bactericidal compound helps them exercise a colonization resistance in the gut, keeping away the early colonizers and potential pathogens: *Enterococcus fecalis* and *Enterococcus fecium* [36]. Again *E. fecalis* in the gypsy moth is known to colonize by acidifying the alkaline niches of the gut, in turn, deactivating the toxins from *B. thuringiensis* that work in an alkaline environment [37]. It has also been reported that the midgut community of *Lymantria dispar* is responsible for the virulence of *B. thuringiensis*. So, the accompanying microbial organisms play roles in pathogenic outbursts [37].

Several studies have reported the phenomenon of transgenerational immune priming, where the maternally acquired bacteria prime the immune system of the offspring so that they are able to defend themselves against pathogens. Such a phenomenon has been observed in *Trichoplusia ni*[32].

Plants elicit responses against herbivory brought about by insect pests [38], [39]. Lepidopteran pests against plants require overcoming plant defensive compounds. The gut bacteria at times have been found to manipulate such compounds. Bacteria in the gut of *P. xylostella* conjugate the plant lipophilic toxins by producing glutathione-S-transferase [40]. Secondary metabolites like terpenes are toxic against insects and bacteria because they are capable of reacting with the cell membrane and disturbing chemiosmosis [41]. Rhodococcus in the guts of gypsy moth degrade such monoterpenes at a high pH [42]. Phenolic compounds when ingested by lepidopterans produce high concentrations of Reactive Oxygen Species and that meddle with digestive activities of the insects. Enterobacteria in the lepidopteran guts are able to quench such compounds by producing enzymes like superoxide dismutase, thus preventing the ROS from damaging proteins and cells [43], [40]. Plants also employ protease inhibitors to prevent plant protein digestion by insect herbivory. For example, velvet bean caterpillar, to counterbalance proteases coming from soy-bean plants use their gut-bacterial derived proteases [44], [45]. Thus, although insects have evolved mechanisms to counteract the plant derived defences, in several cases, the contribution of their gut bacteria is commendable.

Several lepidopterans have been reported to be infected by endosymbionts like *Wolbachia* and *Spiroplasma*. These are capable of colonizing the reproductive tissue and manipulating

the host physiology to their advantage [46], [47]. The impact of wolbachia infection is several: male killing and feminization of males leading to a distorted sex ratio [48], [49], cytoplasmic incompatibility of the hosts a [50]. At times, wolbachia infection leads to longevity by provisioning riboflavin [51]. Sometimes it alters the host behaviour and immunity against pathogens, and manipulates the plant physiology in favor of the caterpillar [52]. Gut bacteria of *P. xylostella* and *H. armigera* reduces larval mortality after being exposed to the control agent Bacillus thuringiensis [53]. Such contributions of the gut microbiota lead to an immense agricultural damage every year [54].

1.5 The age of 'Omics'

Until 1994, microbiome study solely depended on culture based techniques. Scientists dealt with the downsides of it namely, slow growth of certain microbes and unculturable bacteria, without which data, any microbiome study would be incomplete. One of the earliest milestones in the field of microbiome study was sequencing the small subunit of ribosomal RNA, previously pioneered by Carl Woese in 1977. He had revealed that bacteria can evolve and show phylogenetic relationships, and based on molecular marker 16S rRNA, the same can be deciphered. Until 1980, the proposition did not convince scientists, until Norman Pace proved that cultivation of microorganisms is not necessary anymore to study their phylogeny, but simply cloning the SSU rRNA of a bacterial community was sufficient. The contributions of Woese and Pace revolutionized the culture-independent techniques to study microbial phylogeny and ecology respectively, which were previously neglected. Using this approach Wilson and Blitchington compared the levels of cultivated and uncultivated bacteria in human fecal samples in 1996. Since then, the molecular chronometer 16S rRNA has become a powerful gene to study complex communities of microbes [55].

This new technique was previously associated with Sanger sequencing. Albeit powerful, the method came with its downsides of requiring labor intensive cloning procedures and expensive sequencing steps for large scale microbiome studies. With the advent of Next Generation Sequencing (NGS) techniques, these problems were taken care of. It is not only cost effective, but also eliminates the cloning steps by relying on PCR amplification strategies with barcodes. These high-throughput methods such as shotgun sequencing that sequence the genome in its entirety have high sequencing depths that target underrepresented microbes as well. The dawn of the revolutionary sequencing methods led to the completion of

the “Human Microbiome Project” by the National Institute of Health. Genome sequencing has linked the microbial metabolic activities to human health and disease [56].

In symbioses studies, until very recently, the molecular basis was only possible of being studied on the basis of a very few genes at a time. But in the last decade, omics revolution has brought about a deep understanding of the molecular basis of symbiosis. The complete genome sequence of an organism tells us what it is capable of expressing. Mere presence of certain gene is not evidence enough that it indeed is actively expressed in a symbiotic relationship. Yet, genome sequencing gives us a lot of information about metabolic interaction between host and symbionts, for example, what kind of genomic reduction has been brought about in symbionts because of a coevolution process [57].

Transcriptomic studies give us valuable information about the several fluctuations an organism goes through while establishing a successful symbiosis. The conditions include biotic and abiotic alterations such as stressors in the gut, or that in the environment. These sudden changes are accompanied with a dramatic reorganization of gene expression to enable the organism to adapt to the perturbations. Spatial and temporal changes in interaction between the symbiont and the host can be assessed by analyzing the differential gene expression in the partners.

Proteomics or the global analysis of proteins is technically more difficult to perform than transcriptomics. Since proteins are translational products of transcripts, one would get smaller amounts of protein samples to start with although several studies have found a positive correlation between transcript and protein abundance. Also, these studies would require a protein sequence database from the same species. In cases where the analysis of post transcriptional modifications and protein-coding genes need to be addressed, proteomics is the method of choice [57].

Metabolomics or the study of the global set of metabolites cannot be deduced from the genome of the organism. One needs different methods for analyzing different classes of metabolites and most metabolites remain unidentified. In several cases, how the metabolic fingerprint differs between the host with the symbionts and without is looked into, without investment of effort in identifying the metabolites [57].

Omics approaches have indeed revolutionized and deepened the understanding of symbiotic systems. It definitely gives a clear picture of interaction among the host with its symbiotic partners, for further work to be performed on the specific aspects of the bigger picture. Of

late, as the sequence technologies are getting cheaper and bioinformatics methods becoming more user-friendly, these are the methods of choice to study symbiosis [57].

1.5.1 The practical applications of omics in insect-microbiome research.

Omics technologies have made possible volumes of research on several insect-microbial cases. Some of these are explained below:

Aphids

Aphid *Acyrthosiphon pisum -Buchnera aphidicola* relationship is a classic example of obligate mutualism where none of the partners are able to survive without the other. This bacteria, housed in special aphid cells called bacteriocytes, is able to synthesize essential amino acids- arginine, lysine, histidine, leucine, isoleucine, methionine, phenylalanine, valine, threonine and tryptophan) (Matching the supply of bacterial nutrients to the nutritional demand of the animal host). Aphids are mostly on a phloem based diets that lack in the EAAs and vitamins. Thus, they need buchnera to provide them with these nutrients. Their association began about 200 million years ago, and due to coevolutionary processes, they underwent a dramatic reduction in their genome size, by losing all the genes that are not needed in a mutualistic lifestyle. Certain strains have as small as 480 kb genome size. They only retain the genes required for biosynthesis of nutrients the host partakes as well. They lack all the regulatory genes that supervise the symbiotic process. Thus, the host here plays the dominant role in maintaining the association [58].

Cassida rubiginosa

Cassida rubiginosa houses an obligate symbiont which has a genome size of 0.27 Mb. This symbiont retained genes for functional pectinolytic metabolism targeting the two most important groups of polysaccharides: homogalacturonan and rhamnogalacturonan 1. Pectin is one of nature's most complex polysaccharides. Removal of this symbiont leads to reduced host fitness and loss in ability to degrade pectin [59].

Riptortus pedestris

Burkholderia insecticola in the gut of *Riptortus pedestris,* a notorious pest for leguminous crops occupy specialized midgut crypts and help in the host development and reproduction.

They do so by aiding in the production of hemolymph storage proteins like vitellogenin-1, hexamerin-a and hexamerin-b [60]. Aposymbiotic insects have fitness deficits- smaller sizes, prolonged developmental period and reduced egg numbers. The symbionts are involved in key cellular functions like cell division, protein biosynthesis and respiration. Transcriptomics revealed that these midgut dwelling bacteria have higher stress sensitivity owing to the cysteine-rich antimicrobial peptides produced by the host. They have pathways upregulated to assimilate the insect-nitrogen wastes and synthesize B-vitamins which are scarce in soybean plants.

Gut bacteria have also been seen responsible towards the host developing a resistance towards pesticides. *Burkholderia* in the gut of *Riptortus pedestris* is also capable of degrading organophosphate pesticide contributing towards its resistance [61].

Honey bee

The honey bee is an important model system to study host-microbe symbiosis because of the simplicity of the system and the fact that they are important pollinators threatened by population decline around the world. The hind gut of honeybee is home to a simple and specific group of bacteria. It houses five main bacterial species, all of which are habituated to survive in association with the host. They have adopted fermentation as the major metabolic pathway and utilize sugars from the carbohydrate rich diet of the host. Bees mostly feed on nectar and pollen. *Galliamella apicola* is capable of breaking down the pectin component of the inner wall of pollen as the breakdown product, galacturonate accumulate in the gut compartment predominated by *G. apicola*. The exine and intine of the pollen consists of several other components that the gut bacteria can utilize, namely, ω-hydroxy acids, flavonoids and phenolamides. Lactobacilli and *Bifidobacterium asteroides* colonize the rectum and utilize the pollen derivatives. *Frishella perrara* colonizes the entrance of the ilium and induces the scab phenomenon, an immune response leading to melanin deposition on the epithelium's cuticle lining. *Bifidobacterium asteroides* has also been shown to induce production of host derived prostaglandins and juvinile hormone derivatives. These core members exercise a colonization resistance against the parasites *Crithidia bombi* and *Lotmaria pasim*, and pathogens like *Serratia*[62].

Turtle ants

Nitrogen forms a substantial component of living cells, being the major element in nucleic acids and proteins. Herbivores face challenges in obtaining optimum levels of nitrogen in their system because of lack of accessible forms of nitrogen and because it often lacks in their preferred diets. The diet of herbivorous ants comprise of extrafloral nectar, pollen, fungi, vertebrate waste and plant wound secretion where the nitrogen content is limiting. Nitrogen provisioning in carpenter ants is done by *Blochmania* harboured intracellularly. The core microbiome of turtle ants belonging to the lineages of Burkholdriales, Opitutales and Rhizobiales recycle nitrogen from urea or uric acid sources and the host receives them in the form of essential and non-essential amino acids [63].

Termites

Termites that mostly feed on nutritionally imbalanced food source of dead plants. *Coptotermes formosanus* harbor *Klebsiella pneumoniae* and *Enterobacter agglomerans* have the ability to fix nitrogen [64]. Australian termites like *Mastotermes darwiniensis* house *Citrobacter freundii* for the same purpose [64]. Isolates of nitrogen fixing *Clostridia sp* and *Klebsiella sp* have been found in fungus cultivating species of Mactotermes [65]

2 Aims of the thesis

This thesis is an in-depth study of the gut bacteria of *Spodoptera littoralis* larvae. Since the gut environment is unfriendly, in the first part of the thesis, a reporter bacteria was chosen and its colonization strategies were followed inside the gut. In the second part, the gut conditions were altered to study the changes in the bacterial population.

The core community in the gut consists of *Enterococci, Lactobacilli* and *Clostridia*. The selection of one bacterial species over the other is quite evident throughout the lifecycle, so is the differing bacterial population and abundance among the fore, mid and hind gut of the larva. By the time the larva reaches fifth instar, *Enterococcus mundtii* persist and dominate. The gut environment dictates the persistence of its residents. There is a pH gradient from alkaline to neutral along fore to hind gut respectively, and a depleted iron condition as posed by the chelator 8-HQA (acid) produced by the insects. The aims of this thesis are two-fold:

1. **How does *Enterococcus mundtii* dominate by surviving the gut stress?**
   The method to study this particular topic is unique and has been addressed in **Chapter I**, whereas the detailed results are shown in **Chapter II**.
2. **How does 8-HQA define the microbial landscape of the *Spodoptera littoralis* gut?**
   8-HQA production was halted and the resulting bacterial population was studied in **Chapter III**.

3 Overview of the chapters

3.1 Chapter I

The Microbiome of *Spodoptera littoralis*: Development, Control and Adaptation to the Insect Host

Tilottama Mazumdar, Beng-Soon Teh and Wilhelm Boland
Max Planck Institute for Chemical Ecology, Hans-Knöll-Straße, Jena, Germany
IntechOpen: Metagenomics for gut microbes, DOI: 10.5772/intechopen.72180.
Published

This is a book chapter that describes in details a workflow that was optimized in order to study a specific bacterium, *Enterococcus mundtii* in the gut of insect host, *Spodoptera littoralis* larvae. *Enterococcus mundtii* is one of the dominant bacteria in the gut and it survives facing all odds, what the other bacteria do not. There are several stress factors operating in the gut, namely, a high alkaline pH in the foregut, iron deficiency owing to a high concentration of the iron chelator 8-hydroxyquinoline-2-carboxylic acid in the regurgitate of the insect, oxidative stresses and so on. *E. mundtii* produces a bacteriocin called mundticin that keeps out the potential pathogens namely, *Enterococcus fecalis* and *fecium* and establishes a colonization resistance. In order to investigate how this bacterium is dominating and persisting, a fluorescent reporter strain of the same bacteria was constructed by transforming it with plasmid pTRKH3-ermGFP. Next, this reporter strain was allowed to be incorporated as a gut inhabitant and change its gene expression accordingly. The changed gene expression profile was later studied after the fluorescent reporter *E. mundtii* were sorted out using a flow cytometer, followed by their transcriptomic analysis. The results from transcriptomic analysis will tell us how the gene expression profile of this symbiont changes when it dwells inside the gut of the host, as compared to when it is grown in vitro. This book chapter describes the method step wise and establishes its usage in studying any other host-symbiont pair.

Contributions: Tilottama Mazumdar and Beng Soon Teh made equal contributions to planning and writing of the manuscript and Wilhelm Boland supervised the entire process from planning the experiments to documenting them.

3.2 Chapter II

Transcriptomics reveal the survival strategies of *Enterococcus mundtii* in the gut of *Spodoptera littoralis*

Tilottama Mazumdar [1], Beng Soon Teh[2], Aishwarya Murali[4], Wolfgang Schmidt-Heck [5], Yvonne Schlenker[6], Heiko Vogel[3]and Wilhelm Boland[1]

[1, 2] Department of Bioorganic Chemistry, Max Planck Institute for Chemical Ecology, Jena, Germany

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

[4]Department of Microbiology, Friedrich Schiller University, Jena, Germany

[5]Systems Biology and Bioinformatics, Hans Knöll Institute, Jena, Germany

[6]Department of Internal Medicine, Friedrich Schiller University, Jena, Germany

Under review

This article describes the mechanisms employed by the dominating bacterium *Enterococcus mundtii* in order to survive the various stresses in the gut of its insect host, *Spodoptera littoralis* larva. To target this particular symbiont among several others in the gut, a method of selective-sorting of GFP-tagged *E. mundtii*, followed by its transcriptomic analysis was performed as already described in article I. This article reports detailed results of the change in gene expression profiles of *E. mundtii*, when it lives in the foregut and the hindgut of the larva, as compared to when it is grown in laboratory conditions.
The results shed light on several colonizing strategies of this symbiont that range from adherence to the host epithelial cells, to abating stresses like high alkaline pH in the foregut, low iron in the gut and oxidative stress. These symbionts also seem to be providing the host with lysine, which is an essential amino acid that the host may not be able to produce independently.

Contributions: TM and BT designed and performed all the experiments as proposed and supervised by WB, followed by inferring the results and writing the manuscript. AM helped

in a part of the project. YS handled the technicalities of the flow cytometry and WH analyzed the transcriptomics data. Finally, HV and WB helped in proof-reading the manuscript.

3.3 Chapter III

Role of 8-Hydroxyquinoline-2-carboxylic acid in defining the bacterial landscape in the guts of *Spodoptera littoralis* larvae

Tilottama Mazumdar [1], David Heckel[2], Shantanu Shukla[2], Sabine Hanniger[2] Aishwarya Murali[3]and Wilhelm Boland[1]

[1] Department of Bioorganic Chemistry, Max Planck Institute for Chemical Ecology, Jena, Germany

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

[3]Department of Microbiology, Friedrich Schiller University, Jena, Germany

In Preparation

Certain important factors influence the dynamic population in the guts of *S. littoralis* larvae. It is a longitudinal gut structure, a pH gradient of alkaline to neutral along the length of the gut, and presence of a putative ion chelator, 8-hydroxyquinolone-2-carboxylic acid produced by the larvae in their gut regurgitate. This compound is derived by the host from tryptophan metabolism and is present in high concentration: 0.5-5 mM. Since such compounds have been proven to have siderophoric roles, we wanted to look into the effects of this compound on the gut microbiota of *S. littoralis* larvae.

To test the hypothesis, it was required to knock down the production of the iron chelator 8-HQA. As already mentioned, 8-HQA is a product of tryptophan, generating via 3-Hydroxykynurenin. The enzyme kynurenine monooxygenase is the main enzyme bringing about its formation. The gene for this enzyme was knocked out using CRSPR/Cas9 method to reduce the production of 8-HQA to negligible amounts. The gut-microbiome of these insects were analyzed and compared with that of wild type ones to assess the roles of 8-HQA in dictating the bacterial population of the gut.

Contributions: TM performed the experiments under the supervision of DH, SS and WB. SH performed the CRSPR-Cas9 knock down. AM helped with some initial screening experiments

4 Chapter I

The microbiome of *Spodoptera littoralis*: development, control and adaptation to the insect host

-A modified approach to study colonization strategies of symbionts in insect guts, in a real-time fashion

(Published in IntechOpen: Metagenomics for gut microbes, DOI: 10.5772/intechopen.72180)

Tilottama Mazumdar, Beng-Soon Teh, Wilhelm Boland

Max Planck Institute for Chemical Ecology, Hans-Knöll-Straße 8, 07745 Jena, Germany

Abstract

The symbiotic microbial consortium in the gut of *Spodoptera littoralis* shows dramatic, but reproducible changes in line with the development of the insect from the egg via six larval instars to the pupa. Since the food is kept constant during development, factors from the insect host and certain microbial symbionts are assumed to control the composition of the microbiome. A GFP-tagged *Enterococcus mundtii*, one of the major players of the consortium, easily integrates into the microbiome and can be monitored in all gut segments at all developmental stages. The reporter organism can be recovered from the gut using a preparative low cytometry allowing subsequent RNA extraction for transcriptomic analyses. The transcriptomic profile from the fluorescent *Enterococcus* cells provides information on the adaptation of the reporter organism to the local gut conditions. The concept of using a fluorescent reporter organism that can be recovered at any time from any area of the intestinal tract will allow a holistic analysis of adaptation strategies used by the microbes to adapt to the insect gut. In combination with the analysis of transcript patterns from the gut membranes, a first insight into the molecular interaction between the insect host and the microbiome can be expected.

Keywords: *Enterococcus mundtii*, *Spodoptera litoralis*, gut microbiome, transcriptomics, flow cytometry

4.1 Introduction

The invention of a gut by multicellular organisms is one of the major inventions of evolution. The gut allows the host to exploit the metabolic and catabolic capabilities of a multitude of

microbial inhabitants to degrade and digest recalcitrant and complex organic matter. The symbionts may be also involved in the detoxification of poisonous metabolites of the food [66,67]. The membranes of the intestines carefully separate the consortium of bacterial symbionts from the host organism and prevent infection by invasive and deleterious members of the microbiome. Accordingly, the gut membranes form a complex structure of utmost importance allowing on one hand an intense exchange of nutrients along with high- and low-molecular weight (signaling) compounds, while, on the other hand, the entry of microbes and many of their macromolecular components is effectively blocked [68]. The flux of nutrients and even more complex metabolites across the membrane is controlled by transport proteins expressed in the gut membrane [22,69]. The gut microbiome provides also defense against parasites or pathogens [70-72]. The diverse functions provided by the microbial partners are vital for the insect's survival, especially in adverse ecological niches.

Although almost all organisms rely on core microbiomes [73], in many cases the gut community changes along with the insect's development. In early instars of *Spodoptera littoralis,* several *Enterococcus* spp. dominate, while in the late instars also Clostridia significantly contribute (ca. 50%) to the microbial population [74]. A core community, consisting of Enterococci, Lactobacilli, Clostridia, etc. was revealed in the insect larvae. These bacteria were constantly present in the digestion tract at relatively high frequency despite that developmental stage and diet have a great impact on shaping the bacterial communities. Clearly the insect gut selects for particular bacterial phylotypes. Enterococci are also prominent in the gut of insects such as Drosophila, ground beetle, and desert locust [75,76]. The strong dependence of the gut community on the developmental stage of the insect host may suggest that unknown low- and high molecular weight factors control the symbiotic interaction between the partners. For example, in *Drosophila melanogaster*, the immune system not only plays a central role in preventing pathogen infection, it also controls the resident bacterial population. The intestinal homeobox gene *Caudal* regulates the resident gut microbial community by repressing the nuclear factor kappa B–dependent antimicrobial peptide genes. Silencing the *Caudal* gene by RNAi resulted in the overexpression of antimicrobial peptides, which in turn reduced/altered the microbial population in the gut [77].

To monitor such development-controlled changes in the microbiome of *S. littoralis,* a fluorescent member of the gut symbionts, in particular the dominant *Enterococcus mundtii,*

appeared to be an ideal reporting organism. A GFP-labelled *E. mundtii*[78] would easily integrate into the gut community and survive adverse conditions embedded in the community of enterococci which are anyhow largely resistant to environmental stresses, such as antibiotic exposure, disinfection, desiccation, and starvation. As being fluorescent their presence can be monitored in all gut areas of the larvae and other developmental stages such as the pupae and adults. Moreover, the reporter organisms can be easily recovered from the gut and used for transcriptomic analyses. By comparison of the transcriptomes from adjacent gut tissue and of the microbes, a potential "dialogue" between the insect host and the symbiotic bacteria could be disclosed. The concept is generally applicable and can be used for a holistic analysis of host microbial interactions. The protocol of the approach based on the use of a fluorescent reporter organism, e.g GFP-tagged *E. mundtii*, is described in this chapter.

4.2 Fluorescent reporters and its applications in *in vivo* imaging

Bioluminescence and fluorescence imaging are emerging methods to monitor microorganisms in live organism. The development of fluorescent microorganisms is particularly important to allow live monitoring of its survival and persistence in the host organism. *In vivo* imaging is a popular non-invasive method to track bacterial proliferation in animals. This technique has been widely used in bacterial infection studies involving *Salmonella typhimurium*, *Listeria monocytogenes*, *Pseudomonas aeruginosa* and *Staphylococcus aureus* among others [79].

The reporter proteins are important tools to monitor gene expression from within the cells in real time and *in vivo* environment such as the gastro-intestinal tract. The criterion for genes to be selected as a reporter is the easy detection of signals from the reporter in the cells [80].The *lux* gene derived from bacteria and luciferase from firefly and click beetle (luc) are two common reporter genes used in bioluminescence imaging and fluorescence imaging are associated with the use of green and red fluorescence proteins [81,82]. Other reporter genes, for examples, the chloramphenicol acetyl-transferase gene [83], the luciferase gene of *Vibrio fischeri*[84] and the beta-glucoronidase (*gusA*) gene of *Escherichia coli*[85] have been developed.

Rat and mouse are popular model organisms to study the proliferation and colonization of lactic acid bacteria (LAB) [86,87]. LABs have been tagged with green fluorescent protein (GFP)

and mCherry to study its colonization in the intestinal tract of chicken, mouse and zebrafish [66-69,88]. In addition, GFP-labeled bacteria have been used to monitor cell activities in the activated sludge [89], survival of *E. coli* in the aquatic environments [90], during symbiotic interaction with plant [91], and during infection of macrophages [92].

4.2.1 Fluorescent proteins

GFP isolated from the jellyfish *Aequorea Victoria* is widely used as a reporter for gene expression [93]and for studying the localization and structure of living cells [94]. The GFP has a major excitation peak about 395 nm and an emission peak about 508 nm. The GFP protein contains 238 amino acids with a molecular weight of 26.9 kDa that displays green fluorescence when exposed to light in the range from blue to ultraviolet [95]. The GFP requires only oxygen but no cofactors, enzymes or substrates for chromophore formation, which is advantageous over other reporter proteins [96]. Besides, it is sensitive, non-toxic and does not affect cell growth [97,98]. In addition, the GFP protein is stable at temperature below 65˚C and pH 6-11 [99]. Since the discovery of GFP, many of its mutants have been developed with modification in spectral and folding properties, or enhanced fluorescence intensity [100-103]. The choice of a GFP variant depends on several factors like oxygen availability, pH and temperature of the environment, toxicity, multimerization and photostability[103]. The first *gfp* gene was cloned in 1992 [70] and two years later, the gene was successfully expressed in both eukaryotes and prokaryotes [97]. Apart from GFP, many variants of red fluorescent proteins, such as mCherry and tdTomato have been developed based on DsRed originally isolated from *Discosoma* sp[104]. Since then, some over 40 coral fluorescent proteins with different colors from cyan to chromo-red have been described [105].

4.2.2 Construction of a GFP fluorescent system for *E. mundtii*

LAB are widely used as probiotics due to its benefits on human and animal health by balancing the gut microbiome of the gastrointestinal tract and by eliminating pathogenic microorganisms through the production of antimicrobial peptides [106,107]. Due to the importance of LAB in many applications, it is essential to study its survival and colonization by monitoring its metabolic activities *in vivo* through the development of the fluorescent reporter microorganisms. It is important that the reporter gene in the fluorescent bacteria is stably expressed [108].

Plasmids are present in most of the members of the LAB, including *Enterococcus*, *Lactobacillus*, *Lactococcus*, *Bifidobacterium*, *Streptococcus*, *Pediococcus*, *Leuconostoc* etc.

Plasmids found in LAB vary in size (0.87 kb to more than 250 kb) and copy number (1 or more per cell) and gene content [109-112].

Enterococci harbor plasmids that are resistant to a wide range of antibiotics, including erythromycin, tetracycline, gentamicin and vancomycin [113-116]. Some of these plasmids encode bacteriocins [117-119], virulence factors [120,121], toxin [122]and sex pheromone [123]. Plasmids replicate via rolling circle replication (RCR) and theta replication [71]. Theta-replicating plasmids can carry large DNA fragments and are more stable than RCR plasmids [72]. The enterococci plasmid pAMβ1 replicates via theta-mode. In the early 90's, shuttle vectors of the pTRK family of high- and low-copy-number carrying the origin of replication of pAMβ1 for LAB and p15A for *E. coli* were developed [124]. The plasmids carrying the replicon pAMβ1 isolated from *Enterococcus faecalis*[125]have been reported to replicate in Gram-positive bacteria [126].

The choice of a good expression vector depends on several factors, such as the mode of replication, copy number and stability [127]. The expression vector used in this study is derived from a broad-host-range pTRKH3 plasmid. The replicative plasmid pTRKH3 is a shuttle vector for *E. coli*, *Enterococcus*, *Lactococcus*, *Streptococcus* and *Lactobacillus*[124]. This vector has a copy number (30-40) in *E. coli*, and a copy number (45-85) in *Lactococcus* and *Streptococcus* species [128]. The vector carries an erythromycin resistance gene, which is expressed in *E. coli* and LAB. In this chapter, we report the expression of mutated *gfp* (mgfp5) on a pTRKH3 plasmid directed by a strong constitutive promoter, erythromycin ribosomal methylase (*ermB*) [125] in *E. mundtii* (Figure 1A). The lactate dehydrogenase (*ldh*) promoter from *Lactobacillus acidophilus*[129] also has been used to direct the expression of GFP, which yielded comparable result as of *ermB*. In contrast, the use of surface-layer (*slp*) promoter from *L. acidophilus*[130] was not able to induce the expression of GFP [78]. The selection of an appropriate promoter to achieve a high level of GFP expression is of paramount importance. The nisin-inducible promoters have been used for heterologous gene expression in lactobacilli [131,132]. Nisin can be degraded within the intestinal environment which is the drawback of this inducible expression system [133,134]. Therefore, the use of constitutive or native-based promoters would be favorable as they could ensure constant production of target proteins, notably in the gut environment. Several heterologous constitutive promoters have been used for expression in lactobacilli [135-138]. Several studies using homologous promoters have been reported to achieve efficient gene expression [139,140],

as the transcriptional signal induced by native promoters is recognized by the host bacteria. Bacteria with the *gfp* gene cloned downstream of a native constitutive promoter will express GFP efficiently in broth culture (Figure 1B).

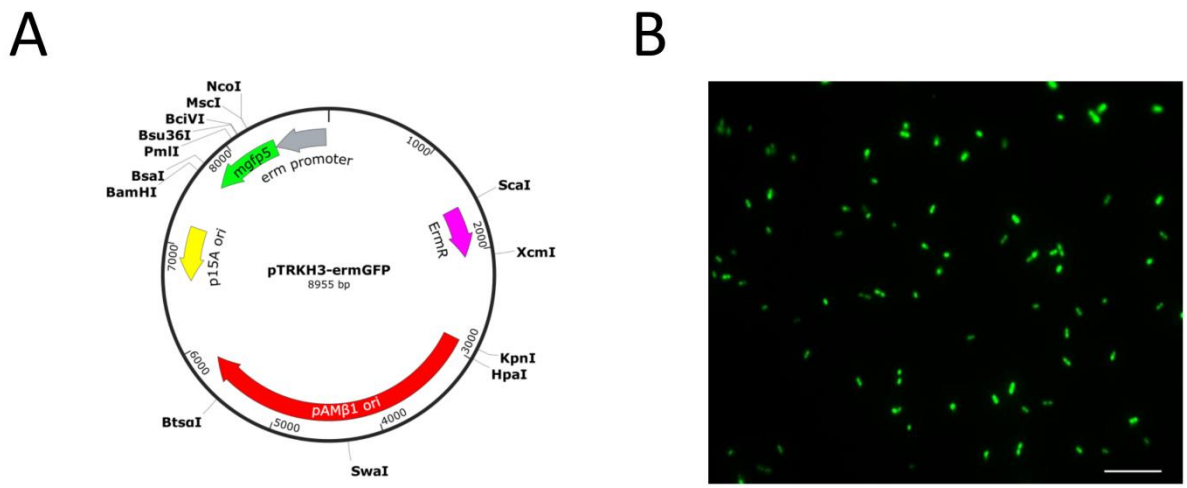

Figure 1.Construction of GFP-tagged *E. mundtii* by electroporation. (A) Plasmid map of pTRKH3 harboring *mgfp5* gene regulated by erythromycin ribosomal methylase (*ermB*) promoter. The plasmid is an *E. coli*-LAB shuttle vector with p15A and pAMβ1 origin of replications. (B) The *ermB* promoter was used to drive high expression of GFP from *E. mundtii*grown in Todd-Hewitt Bouillon (THB) broth culture. Scale bar: 10 µm.

4.2.3 Transformation of *E. mundtii* KD251 using electroporation

Several methods have been used to introduce exogenous DNA into the microbial cells which include chemical treatment, electroporation, use of biolistic gun, ultrasound, polyethylene glycol, microwave and hydrogel [73]. Of all the methods, electroporation is most efficient to transform a broad array of microorganisms [141]. The method utilizes an electric pulse that forms pores on the bacterial cell walls so that DNA can pass into the cell. In recent years, numerous lactic acid bacteria have been transformed using electroporation [142]. The success rate of electro-transformation depends on the permeability of the cell wall that allows sufficient DNA to enter the cell. In some cases, to improve electro-transformation efficiency, the cell wall is weakened by chemicals such as pretreatments with lysozyme [143,144], threonine [145,146], penicillin G [147], ethanol [148]and glycine [149,150]. The cell wall weakening chemicals are effective only for certain bacteria species and not the others [76]. It has been shown that the electro-transformation efficiency of *Lactococcus lactis* was affected by several parameters

such as the growth phase and cell density, the medium, the plasmid concentration and the electric field strength [151].

The choice of the method in the preparation of the competent cells is important for a successful transformation. The chemical preparation of competent *E. coli* cells with ice-cold calcium chloride has been reported [152], however this method yields lower transformation efficiency compared to that of the electroporation method [153]. The electrocompetent cells, the equipment and the washing buffers have to be prepared at cold temperature [154,155]. It has been reported that the electrocompetent cells could be prepared at room temperature, which leads to improved transformation efficiency for several Gram-negative bacteria [156-158].

In this chapter, we report the use of conventional method for the transformation of *E. mundtii* based on the modified protocol of *E. coli*[159]. The electrocompetent cells and electroporation protocol for *E. mundtii* are published [78]. Briefly, the bacterial cells were grown to exponential phase and are then washed with ice-cold water for two rounds to remove salts from the growth medium. The glycerol at a final concentration of 10% was added to the bacterial suspension so that the cells can be preserved and stored frozen. A concentration of plasmids between 0.15 and 0.2 µg works fine in our work. The competent cells were mixed with the plasmid DNA and were then transferred into a 0.2 cm plastic cuvette for electroporation at a pulse of 1.8 kV, 600 Ω parallel resistance and 10 µF capacitance. The pulsed cells were recovered with fresh broth medium and the cell suspension was incubated at 37˚C for 2 hours before plating on plates containing antibiotic erythromycin. After two days, bacterial transformants were screened for the plasmid-containing *gfp* gene.

4.2.4 Colonization of GFP-tagged bacterium in the gut of *S. littoralis*

It has been shown that the fluorescent reporter *E. mundtii* was integrated into the gut microbiome across all developmental stages of *S. littoralis*[78], indicating its symbiotic relationship with the insect host. The dominance and persistence of *E. mundtii* in the gut motivates us to look deeper into their gene expression system. Microorganisms have the ability to face environmental perturbances using their stress evasion system. Therefore, it is important to unravel the mechanisms used by microorganisms living within the gastro-

intestinal environment. Construction of the fluorescent reporter *E. mundtii* is one of the strategies to find out those mechanisms, since it has been possible to recover the reporter bacteria from the gut of the very same insects using the *state-of-the-art* technology of flow cytometry. Fluorescence-activated cell sorting (FACS) enabled us to selectively sort out the GFP-tagged reporter *E. mundtii* from a mixture of insect and other bacterial cells.

4.3 Fluorescent-activated cell sorting

Flow Cytometry works on the principle of separation of cells based on their intrinsic physical and chemical characteristics. It is an integration of electronics, fluidics and optics. The sample, from which the cells of interest are to be sorted, is passed through a flow cell. The sheath fluid escorts the cells down the channel, where they are encountered by a laser beam. This is where the optics system plays a role by emitting light beams of specific frequencies and wavelength. Based on the cell size and granularity, the forward scatter (FSC) and the side scatter (SSC) are measured respectively by detectors. FSC and SSC are unique for every particle. A combination of the two can differentiate between cell types in a cohort of cells. This way, the qualitative and quantitative data of a particular kind of cell can be assessed. Also, depending on whether the machine is a sorting kind or not, the cells can be isolated.

Fluorescence-activated cell sorting is an application of flow cytometry. The cells of interest are fluorescently tagged and sorted by the machine. In this case, it is the GFP-tagged fluorescent *E. mundtii* which is to be isolated from a mixture containing insect gut homogenate with other bacteria. The solution is delivered to the flow channel, carried down by the sheath fluid. The pressure from the compressor, which is adjustable, forces it down. A hydrodynamic focusing of the cells causes them to pass the laser beam: a monochromatic beam of high intensity, one at a time depending on the excitation wavelength of the fluorophore used, the laser is chosen. The scatters are then recorded. Forward scatter (FSC), the light that is refracted by the cell and continues in the same direction, tells us about the cell size. Whereas, side scattered light (SSC), the light refracted by the cells and traveling at right angles to the excitation axis, informs of the fluorescence and granularity of the cells. The more granular a cell, the more scattered light it produces. Furthermore, each cell enclosed in a droplet is assigned a charge, depending on the extent of deflection by a set of electrically charged plates [160]. After passing through the electrical field, the cells are deflected to the collection tubes. The uncharged droplets are directed to the waste. The detector system

consists of a set of photo multiplier tubes. They have specific filters to select for certain wavelengths of the beam and eliminate the rest. They are set at the excitation range to view GFP.

Once the larvae are fed with the fluorescent *E. mundtii*, the numbers that survive at various stages of life can be determined and eventually recovered for further studies. The *E. mundtii* cells are sorted for studying their transcriptome. Thus, with the advent of this technology it has become possible to focus on single cell or cells of interest, to study their function or physiological state.

4.4 RNA extraction

The GFP-tagged *E. mundtii* are sorted by the flow cytometer and collected in a RNA-protective reagent. The role of such reagents is twofold: firstly, they preserve the integrity of RNA, as it has a very short half-life, for minutes. We need the RNA to be intact and of good quality in order to process it for sequencing. Secondly, since studying the gene-expression profile is what it is aimed for, any subsequent changes cannot be allowed during the process of handling the cells. As soon as the cells are collected in a falcon tube filled with the protective reagent (RNA Protect or RNA Later), it percolates into the cells to arrest the gene expression system [161]. Additionally, the whole process is maintained at $4^{o}C$, as all metabolic activities slow down at low temperatures. The falcon tube is centrifuged at a high speed to pellet down the cells, and care is taken not to disturb it while draining out the supernatant.

RNA is very sensitive to exogenous and endogenous RNases. The entire extraction procedure is done in an RNAse free area. Moreover, RNase inhibitors are used to clean all equipment, ranging from gloves to microcentrifuge tubes to get rid of RNAse. *E. mundtii* is a gram-positive bacterium with a cell-wall containing a thick layer of peptidoglycan and lipotechoic acid, followed by a single lipid membrane. The cell wall is anchored to the membrane by diacylglycerols. To release the nucleic acid from the cell, it has to be made free of its peptidoglycan containing cell wall and membrane. Lysozyme is a glycoside hydrolase that hydrolyzes the 1, 4-beta linkages between N-acetylmuramic acid and N-acetyl-D-glucosamine residues of the peptidoglycan. Additionally, guanidium thiocyanate, beta-mercaptoethanol, and a detergent called dithiothreitol help in cell lysis and deproteinization. Proteinase K frees the RNA from the bound proteins and endogenous RNAse.

Following the lysis, the RNA is separated by density gradient centrifugation using phenol, chloroform and isoamyl alcohol, and further precipitation with ethanol. Thus, RNA is obtained from the cells of the sorted *E. mundtii*[162].

The extraction of total RNA from a very low number of bacterial cells remaining after sorting by the flow cytometer is challenging. The concentration of RNA was as low as few picograms to 50 ng. The minimum threshold quantity for a successful RNA library preparation is 100 ng. This is too low an amount to proceed with RNA sequencing. Hence, it is a prerequisite to amplify the total RNA.

Amplification of RNA

Amplification of RNA is required if the aim is to go for an effective transcriptomic profiling from a very low starting quantity of RNA. The principle is based on *in vitro* transcription. The steps are as followed:

Polyadenylation of RNA: Since it is bacterial RNA, it is devoid of a poly (A) tail. The *E. coli* poly (A) polymerase enables the addition of poly(A) tail at the ends of RNA. This stretch is required for cDNA synthesis.

Synthesis of first strand cDNA: Primers against the poly (A) stretch is used to synthesize the first strand of cDNA by reverse transcription. The primers are anchored with a bacteriophage promoter sequence: T7 oligo(dT) sequence, T3 or SP6. dNTPs are added to the reaction mix.

Synthesis of second strand cDNA: RNaseH is used to degrade the RNA from the RNA-cDNA pair. DNA polymerase is required to synthesize the second strand of cDNA. Now we have a double-stranded-cDNA with T7 promoter sequence.

Purification of cDNA: The cDNA is cleaned by removal of the fragmented RNA, enzymes, salts, which could hinder the in vitro transcription.

In Vitro Transcription: Multiple copies of antisense RNA are generated using DNA dependent RNA polymerase. Linear amplification is employed for this. Depending on the bacteriophage promoter sequence attached to the cDNA, the polymerase is selected. Promoter

specific dNTPs are added to the reaction mix. Temperature at 37^{o}C is optimum for this reaction. The reaction time depends on what extent one wants to amplify the RNA.

Purification of amplified RNA: The residual enzymes, salts, unincorporated dNTPs must be removed from the final product.

Now, the RNA has been amplified to several folds: 1-2 ug.

RNA amplification procedures have a drawback. In addition to bringing the concentration of RNA to a point where successful transcriptomic profiling is possible, it introduces certain biases. Certain amplified transcripts are at times misunderstood as duplicates and vice-versa, which could give a false positive result. In PCR-based amplification procedures, duplicates can arise from sample handling methods such as fragmentation, sequencing depth or library complexity, which unfortunately cannot be distinguished from PCR-duplicates, computationally. Removing duplicates does not improve the accuracy of quantification or the power, rather makes it worse[163].The *Taq* polymerases are more prone to introduce errors than RNA polymerases. Thus, *in vitro* transcription is favored over PCR-based amplification, although, a premature transcription termination can occur in low complexity sequences[164]. Nevertheless, linear amplification is an efficient method to follow when the starting quantity is limited [164].

4.5 Transcriptomics

At this point we have enough RNA to get a transcriptomic profiling of the bacterial cells done. Transcriptome is the entire set of genes expressed in a cell-type at a particular time point and/or condition. This is in contrast to a genome, which is the full complement of genes in a cell-type. Not all genes are constitutively induced. Information about transcripts, or genes expressed give the insights into the developmental or physiological state of the cell. It also speaks about other species of RNA, small RNAs and non-coding RNAs, novel transcripts the transcriptional start sites, splicing regions, post transcriptional modifications, 3' and 5' ends. Another aspect of transcriptomic profiling is to quantify the expressed genes. One can judge the extent of regulation of a particular gene in the given conditions. As compared to one situation, when cells behave differently in another, one can now say which genes are

differentially regulated to bring about the same. Thus, identification and quantification of genes have become possible.

In this chapter, our aim was to unravel the survival and adaptation strategies of *E. mundtii* living inside the gut of *S. littoralis* as compared to the normal laboratory conditions. This is done by cataloging the genes that are differently regulated, which make it as one of the dominant bacterial species in the gut. Therefore, genome wide RNA sequencing of the transcriptome of the gut-resident *E. mundtii* is a powerful method of choice.

4.5.1 RNASeq

Transcriptome sequencing has gone through several breakthroughs in the past years. Starting from EST-based Sanger sequencing, to the next-generation methods, it the drastically improved in sensitivity and accuracy. In contrast to the former method, whose productivity is mainly confined to the most abundant transcripts, the next-generation sequencing informs a lot more about the non-abundant ones. Thus, low expressed genes in the given situation are identified. This has been ameliorated with the advent of deep sequencing: the average number of times a nucleotide is sequenced. The deeper the sequencing, the better is the probability of detecting the less abundant transcripts. Next-generation sequencing itself has several hierarchies of its own. These days, RNAseq is more widely used than the microarrays. The former gives us a base-pair level of resolution. While microarrays can be used only when the reference genome sequence is available, RNAseq can build the transcriptome *de novo.* Also, background noise is taken better care of in case of RNAseq. These days, sequencing is not just confined to a bulk of cells. It is possible to obtain resolution up to a single cell. Naturally, the amount of RNA obtained from one single cell is in pico-gram levels, and is required to be processed as discussed above. With the increased sensitivity of the next-generation technologies, so far, Illumina allows the least starting amount of RNA.

The fragmented and adapter-ligated cDNA is let through a flow cell, which has oligonucleotides complimentary to the adapter sequences embedded in them. After hybridization, the oligonucleotides prime the polymerization process with the provided dNTPs and DNA polymerase. Each of the dNTPs is tagged with a fluorophore. As the nucleotide is incorporated, the resulting fluorescence is detected. With addition of each nucleotide, the fluorophore is released, regenerating the 3’ hydroxyl group for the next

nucleotide to join. This way, the fluorescent intensity is recorded, and converted into nucleotide identity using an algorithm.

The amplified RNA from the flow cytometry sorted fluorescent *E. mundtii* cells went through deep sequencing (Hiseq) to detect as many genes as possible to tell us the story of their survival in the gut of *S. littoralis.*

The complications arising from several different forms of RNA, alternate splicing, removal of introns, that is, the ones that are profound in eukaryotes are not required to be considered in the case bacteria. Although, there are several regulatory and non-coding RNAs in bacteria, but this particular case dictates one to follow rather straightforward approach of unraveling the upregulated and downregulated transcripts only.

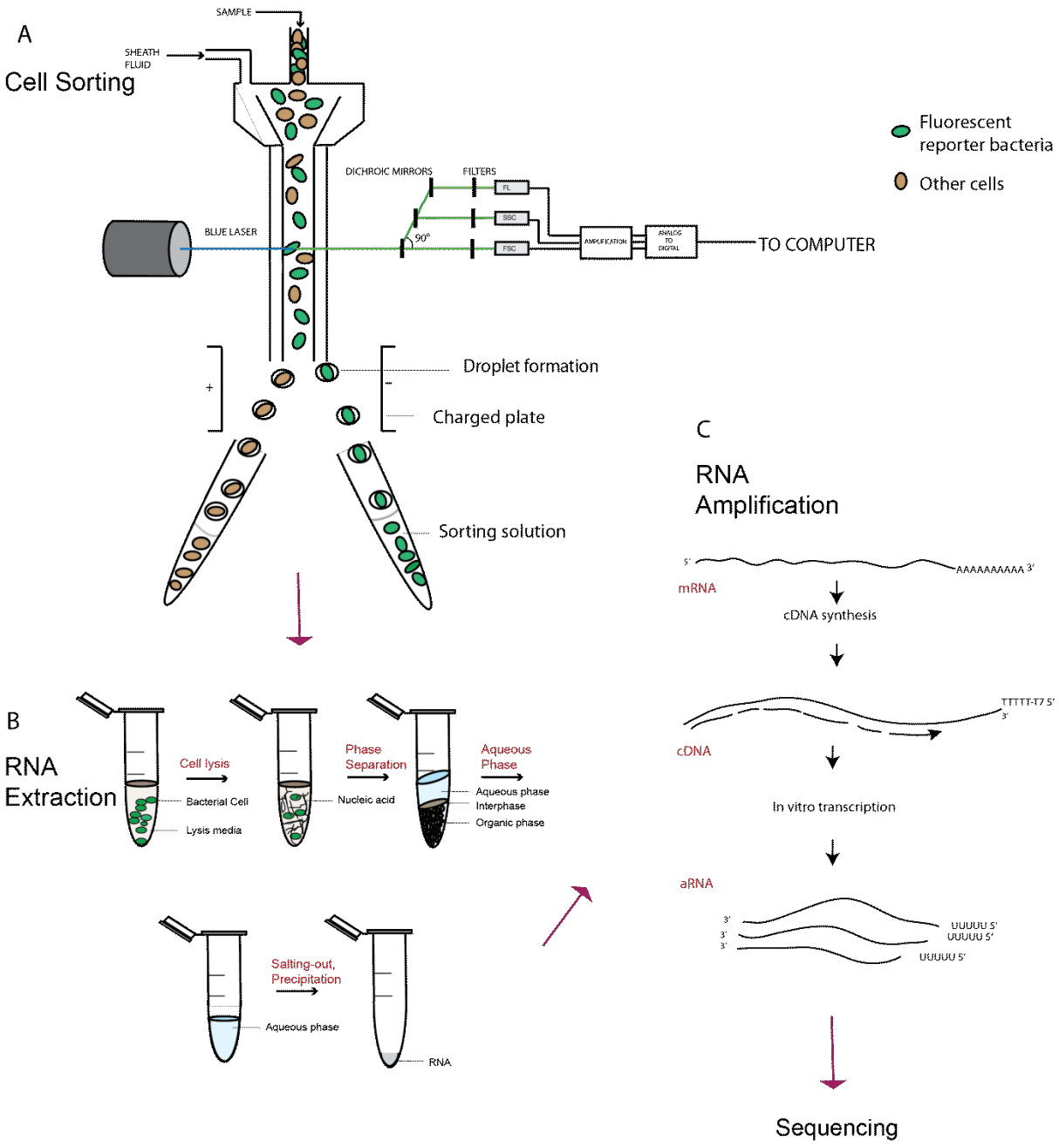

Figure 2: Overview of the workflow for bacterial RNA-seq. (A) Flow cytometry to sort fluorescent bacteria from gut homogenates. (B) Extraction of bacterial total RNA. (C) Amplification of the total RNA by *in vitro* amplification

4.5.2 Adaptation and survival strategies of *E. mundtii* in the gastro-intestinal of insect

The GFP-containing *E. mundtii* was fed to the *S. littoralis* larvae at early instars. The bacteria reporter could colonize the gut at various stages of the insect's life cycle, as seen in the fluorescent microscopic images (Figure 4).

Antimicrobial substances from insects or their resident symbionts are a survival strategy to keep pathogens at bay. It has been shown that the dominant gut bacteria, *E. mundtii* produces an antimicrobial peptide called mundticin KS, which is a stable class IIa bacteriocin. It establishes a chemical barrier, which is one of the reasons for its colonization resistance. Turns out, the early colonizers of the *S. littoralis* gut are *Enterococcus faecalis* and *Enterococcus casseliflavus*, if allowed to persist, could pose as potential pathogens for the insects. Successful antimicrobial activities against them have been shown in presence of *E. mundtii*.

The larvae were allowed to grow till fifth-instar, at which stage, the guts were homogenized to retrieve the fluorescent *E. mundtii* by flow cytometry. The RNA of these sorted bacteria was used to probe their differential behavior inside the gut. RNA sequencing and analysis of differential gene expression are performed later on.

There are numerous genes that are differentially regulated in the *E. mundtii* obtained from the gut, as compared to the *E. mundtii* grown in bacterial culture under lab conditions (Table 1). The larvae respire aerobically. Reactive oxygen species like superoxide radicals, hydrogen peroxide or hydroxyl radical are formed by partial reduction of oxygen. If not abated, they cause oxidative stress in the gut, causing damage to macromolecules. To survive the stress, the resident bacteria have to come up with means to fight it. Superoxide dismutase and catalase are effective enzymes, over produced by *E. mundtii* when inside the gut, as compared to the broth culture.

Adhesion to the host gut epithelial surface is the key to a successful colonization. Endosymbionts employ certain proteins (motifs and domains for the same). These are mostly cell-wall associated surface proteins employing certain motifs, which act as the signal peptide for attaching to the cell wall. For example, the motif called LPXTG is a sorting peptide. The endopeptidase sortase cleaves it at the site between threonine and glycine residues, and links the peptide covalently to the peptidoglycan of the cell wall [165]. There was upregulation in the genes encoding this motif and also sortase enzymes, indicating towards their attachment to

the insect gut wall and biofilm formation. The WxL domain, whose upregulation hints towards the increased colonization of the bacteria by adherence to the gut epithelium. The WxL domain proteins are also crucial for adaptation to varying environmental conditions [166].

Adaption to variable living conditions is very much attributed to "two-component systems". They form a class of signal-transduction mechanism that sense stress in the environment and get induced. The main players in the system are auto-inducing proteins, histidine protein kinase (HPK) and response regulators. Auto-inducing proteins are produced in response to stress, which interact with the HPKs. The signal is relayed to the response regulators. This cascade ultimately produces certain factors or proteins that aid in their survival in stressful environments[167]. Agr family of genes is one such system which was found upregulated in *E. mundtii* living in the insect gut.

Quorum sensing is a phenomenon where the bacterial cells interact and communicate with one another for survival. Auto-inducing proteins are also the key players for quorum sensing. In addition, several quorum sensing strategies are two-component systems as well. The AIPs accumulate in response to increase in bacterial cell density, which is followed by a signaling cascade, leading to a cooperative gene expression by the bacteria [168].

Stress proteins are a class of adaptive factors that come into play during stressful living conditions. They are general and universal stress proteins. The general stress proteins help bacteria deal with oxidative stress, heat stress, salt stress, or oxygen limitation [169]. Universal stress proteins are induced in response to temperature fluctuations, heat or oxidative stress and hypoxia. Both these protein classes displayed an upregulation in *E. mundtii* when confronted with the insect gut's living conditions [151].

Bacteria express their respective sugar transport systems depending on the types of carbon sources available. Phosphotransferase systems form a class of sugar transporters that sense the sugar source available in the environment and allow the respective transporters for fructose, glucose, mannose or cellobiose to act on it. Utilizing energy from phosphoenolpyruvate, the transport system utilizes a cascade of cytoplasmic protein components with an accompanying phosphorylation of each component [170]. These transporters are generally sugar specific and help bacteria to survive in presence of complex carbohydrate conditions, leading to their adaptation. The PTS systems in *E. mundtii* have several of these upregulation from the insect guts [171].

Table 1. Upregulation of genes and pathways in *E. mundtii*, living in the gut of *S. littoralis*

Gene/Protein	Pathway	Function
Superoxide dismutase (*SOD*)	Oxidative stress management	Quenching Reactive Oxidation Species by partial reduction of O_2^-
Catalase	Oxidative stress management	Quenching Reactive Oxidation Species, converting hydrogen peroxide to water and oxygen
LPXTG-motif cell wall anchor domain protein	Cell surface adhesion	Signal peptide cleaved by sortase for cell surface adhesion
WxL domain surface cell wall-binding protein	Cell surface adhesion	Cell surface adhesion and adaptation
Accesory gene regulator (*Agr*)	Two-component system	Virulence factor
General stress protein	Adaptation	Various stress management
Universal stress protein	Adaptation	Adaptation to diverse stress sources
Ferric (Fe^{+3}) ABC superfamily ATP binding cassette transporter (*fetC*)	Iron transport	Iron transporter permease
Phosphotransferase systems	Sugar transport	Regulates carbohydrate metabolism in diverse sources and adaption.

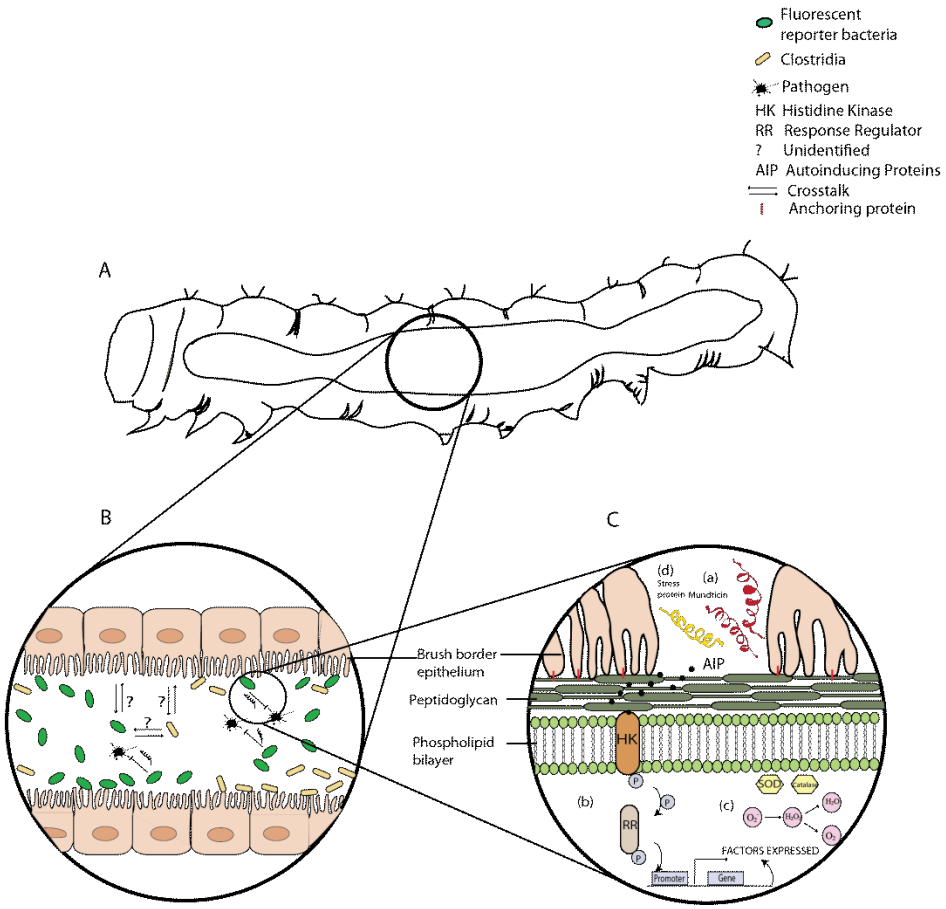

Figure 3: The gut microbiome of *S. littoralis* was predominated by *E. mundtii* and *Clostridia* sp. (A) Overview of the gut structure of fifth- instar larva of *S. littoralis*. (B) Illustration from within the gut space, which harbor major symbionts *E. mundtii*, *Clostridia* sp and other bacteria. Bacteria are adhered to the mucus layer of insect gut epithelium. Unknown interactions occur between microbe-microbe and host-microbe. (C) Some major expressed pathways by *E. mundtii* for survival in the gut. (i) Secretion of mundticin, an antimicrobial peptide to keep pathogens at bay and exerts the colonization dominance of *E. mundtii* (ii) A two-component system involving accessory gene regulator (*agr*) system, which directs a histidine kinase to phosphorylating the response regulator, leading to the activation of transcription factors required for adaption. (iii) Induction of superoxide dismutase and catalase to manage oxidative stress, leading to the conversion of superoxide radicals to water and oxygen. (iv) General or universal stress proteins to overcome different kinds of stresses, such as oxygen starvation, heat or oxidative stress.

4.6 Discussion and future prospects

Lactic acid bacteria are important in the production of fermented foods, such as dairy products. LAB are potential probiotics that provide benefits to human health [172]. The modified LAB could also be used as live vaccines or vaccine delivery systems [173]. It has been shown that the genetically modified *Lactococcus lactis* can survive and colonize the digestive tract of human [174]and gnotobiotic mice [175]. In this chapter, we report the use of GFP to tag *E. mundtii* to monitor its survival and activities in the intestinal tract of cotton leafworm, *S. littoralis*.

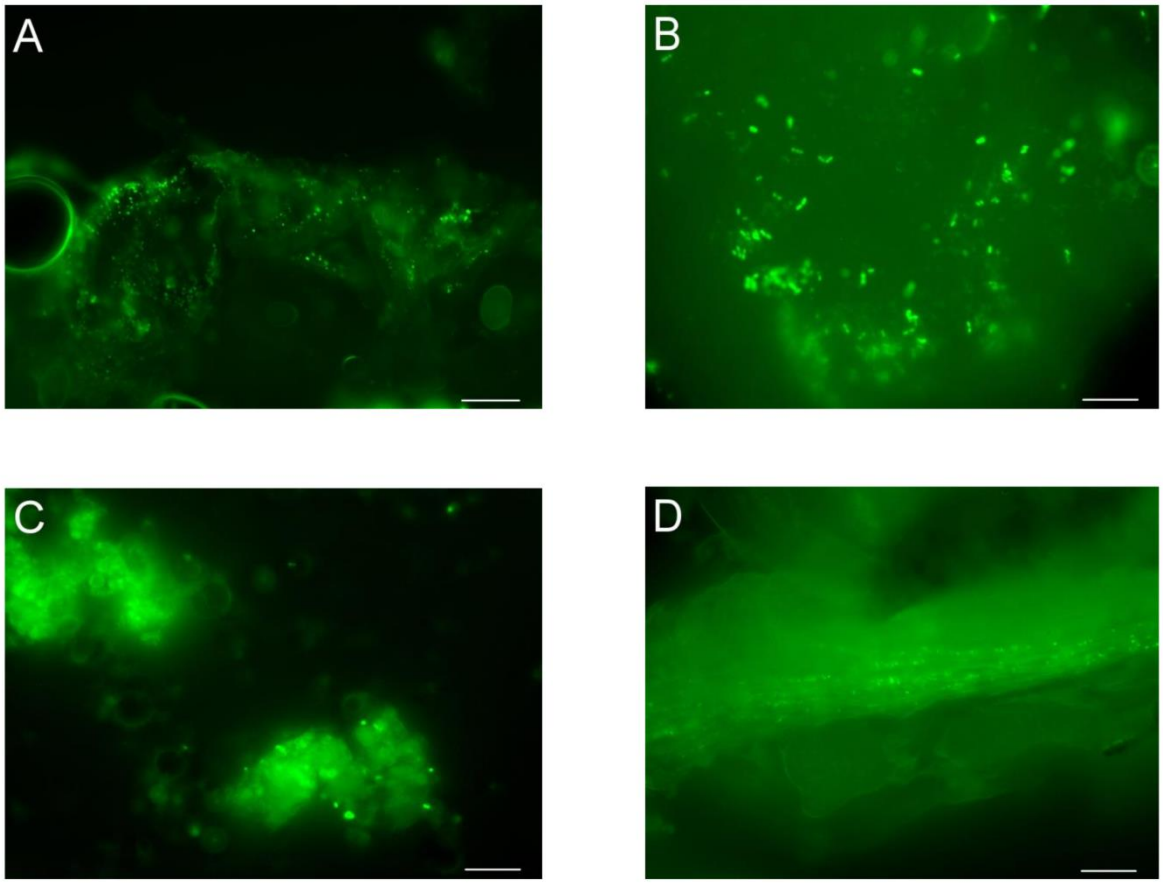

Figure 4. Localization of fluorescent *E. mundtii* in the intestinal tract of *S. littoralis* of different life stages. (A) Accumulation of bacterial cells on the peritrophic matrix separated between gut lumen and epithelium of fourth-instar larvae. (B) Bacteria clustered in the gut of fifth-instar larvae. (C) Fluorescent bacteria are still visible and colonize the tissue of pupae although no gut tissue is formed. (D) Vertical transmission of symbiont is evident as fluorescent *E. mundtii* survive first generation and colonize second generation first instar progeny. Scale bars: 10 – 20 μm.

It has been shown that spatial and temporal distribution of fluorescent *E. mundtii* was observed across all developmental stages (Figure 4), as well as in the foregut, midgut and hindgut of *S. littoralis*. The colony forming units (CFUs) data showed that the midgut houses the most abundance bacterial counts, followed by the hindgut and foregut. One interesting point to highlight is that the fluorescent *E. mundtii* were detected in the eggs of *S. littoralis*, showing the importance of symbiont transmission from one to another generation [176]. Similarly, other study also showed that fluorescent bacteria were transmitted from the gut to the eggs in *Tribolium castaneum*. The symbiotic *E. mundtii* was transmitted to the second-generation progeny, suggesting that it co-evolves together with the insect host (Figure 2). In addition, the fluorescent bacteria could be detected in fecal samples of the larvae, indicating successful passage along the intestinal tract of *S. littoralis* (data not shown). The question of how a bacterial symbiont is transmitted from one generation to the next remains to be clarified in detail. The symbiont that co-evolves with the host has a great chance to secure vertical transmission, for example the symbiosis relationship between the aphid and its endosymbiont *Buchnera aphidicola*. It has been shown that the GFP-tagged *Asaia* strain is vertically transmitted from the mother to the offspring in *Anopheles stephensi*[177]. The bacterial symbionts can be horizontally transferred via "egg smearing", a phenomenon which female stinkbugs covering the surface of her eggs with symbiotic bacteria during oviposition. The newly hatched juveniles acquire the symbionts by ingesting the egg case [178].

Several factors, including the pH, redox potential, oxygen availability, nutrient and immune system can shape the microbial composition of the gut of insects [179]. Furthermore, constant change in gut contents due to molting and metamorphosis can affect the colonization of microorganisms. Many insects have the intestinal pH in the range of 6-8, and some lepidopteran larvae have higher pH of 11-12 in their midguts [180,181]. The hindgut harbors high bacterial diversity and density in several insects such as cockroaches, crickets and termites [182-184].

Microorganisms live in the hindgut benefit by the metabolites and ions transported from the Malpighian tubules into the hindgut. The hindgut stores the nitrogenous waste and food waste possibly serve as nutrients for insect gut bacteria [179]. The hindgut also involves in water resorption [185]. The microbiota in the ileum of the hindgut of scarab beetles metabolizes plant polysaccharides into components that can be used by the insect [186].In contrast to the hindgut, the midgut is more unfavorable for microorganisms to live in. Many antimicrobial peptides[187]

and digestive enzymes (lysozymes) [188] are secreted by the midgut epithelium cells of *D. melanogaster*. The peritrophic matrix secreted by midgut epithelial cells tends to accumulate digestive enzymes and serves as barrier to separate food particles, toxins and microorganisms [189]. The high alkaline pH in the gut of lepidopteran insects could kill many microorganisms. However, the alkaline conditions favor the dominance of Firmicutes-related bacteria in the midgut of the beetle *Pachnodaephippiata*[190]. Both culture-dependent and culture-independent methods have detected the presence of *Enterococcus* in the alkaline midgut of the gypsy moth larva [191].

The mechanisms of bacterial colonization in the specific regions of the gut are not well understood. The gut of *S. littoralis* does not possess specialized structures called bacteriomes that contain endosymbionts, such as in aphids, whiteflies and other insects. The questions of how *S. littoralis* houses *E. mundtii* remain unknown as no compartment structures exist to protect the bacterium, for example the gut of pupae is totally removed. Several mosquito species that undergo metamorphosis eliminate their gut bacteria, especially in newly emerged adults [192]. The host organism selects its own microorganisms as it depends on them for growth and development. This is especially true when the bacterial symbionts of honey bees were unable to survive in the gut of bumble bees [193].

Few of the important survival strategies of *E. mundtii* are mentioned. There are obviously several other pathways that are meant for their adaptation to the differential living conditions inside the gut. A lot of other mechanisms that help the bacteria to survive in the gut have to be understood with further repetitions of the RNA sequencing.

The research remains incomplete, unless the insect's side of the story is unveiled. The symbiosis between the host and the bacteria that leads to their successful co-existence still remains a question mark, unless the insect's contribution is discovered. The future prospects would to identify the gene expression analysis of the larval gut epithelium to look for genes that regulate the gut microbiome and vice-versa.

Until then, suffice it to say, *E. mundtii* indeed is a successful and a major symbiont in the gut of *S. littoralis.* The method that we have developed here particularly looks into an indigenous bacterial species within the whole community. With further improvements and modifications, this kind of reporter system can be useful in many other species-specific interaction studies.

5 Chapter II

Transcriptomics reveal the survival strategies of *Enterococcus mundtii* in the gut of *Spodoptera littoralis*

Tilottama Mazumdar [1], Beng Soon Teh[2], Aishwarya Murali[4], Wolfgang Schmidt-Heck [5], Yvonne Schlenker[6], Heiko Vogel[3] and Wilhelm Boland[1]

[1, 2] Department of Bioorganic Chemistry, Max Planck Institute for Chemical Ecology, Jena, Germany

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

[4]Department of Microbiology, Friedrich Schiller University, Jena, Germany

[5]Systems Biology and Bioinformatics, Hans Knöll Institute, Jena, Germany

[6]Department of Internal Medicine, Friedrich Schiller University, Jena, Germany

Key words: ***Enterococcus mundtii*, *Spodoptera littoralis*, Flow cytometry, Transcriptomics, Foregut, Hindgut.**

Abstract

The complex interaction between a higher organism and its resident gut flora is a subject of immense interest in the field of symbiosis. Many insects harbor a complex community of microorganisms in their gut. Larvae of *Spodoptera littoralis,* a lepidopteran pest, house a bacterial community that varies both spatially (along the length of the gut) and temporally (during the insect's life cycle).

To monitor the rapid adaptation of microbes to conditions in the gut, the GFP-tagged reporter strain of *E. mundtii*, a major player in the gut community, was constructed. After early-instar *S. littoralis* larvae were fed with the tagged microbes, these were recovered from the larval fore- and hindgut by flow cytometry. The fluorescent reporter confirmed the persistence of *E. mundtii* in the gut. RNA-sequencing of the sorted bacteria highlighted various strategies of the symbiont's survival, including upregulated pathways for tolerating alkaline stress, forming biofilms and two-component signaling systems for quorum sensing, and resisting oxidative stress. Although these symbionts depend on the host for amino acid

and fatty acids, differential regulation among various metabolic pathways points to an enriched lysine synthesis pathway of *E. mundtii* in the hindgut of the larvae.

5.1 Introduction

Insects comprise the largest phylum of arthropods on earth, according to the IUCN red list. Microorganisms are known to form symbiotic relationships with insects by supplying them with essential nutrients, protection against pathogens, and aid in digesting organic matter. They contribute significantly to insects' ability to act as potential pathogens to animals, pests or pollinators of food crops, and as cyclers of carbon and nitrogen during the decomposition of plant biomass [194].

Insects with a straight, tube-like gut usually possess a less diverse microbial population than species with invaginations and deep pouches [194]. Other factors that shape the gut population include the following: oxygen level, gut pH, the presence of digestive enzymes, antimicrobial compounds and insect diet [22,195]. Although most bacteria have an affinity for neutral pH, several acidophiles and alkalophiles have adapted to extreme pH conditions.

Vertical transmission of symbionts allows bacterial transfer (from the ovaries to the egg shells) to the next generation [196], whereas horizontal transmission occurs over the course of the life cycle, through diet and social behavior. Regardless of how bacteria are transmitted, microbial populations may be unstable during early developmental stages. For example, the gut of holometabolous insects undergoes a complete metamorphosis from pupa to adult, resulting in microbial turnover and variable microbial counts [74].

The cotton leafworm, *Spodopera littoralis,* a holometabolous lepidopteran that feeds on a broad range of plants, is a prevalent pest in the tropical and subtropical regions of the world. Larvae of this species have a longitudinal gut structure; without compartments, the gut is more able to flush out bacteria and so avoid being colonized. The simple gut structure could explain the overall low gut-bacterial density observed in Lepidoptera [194,197]. Despite the seemingly simple structure of the gut, it has a pH gradient: the anterior part and midgut of lepidopteran larvae are highly alkaline, with a pH range of 11-12 [198], but the posterior part is neutral [181]. Such a gradient might restrict the survival of many microbial species. Despite their alkaline pH, bacteria of the phylum Firmicutes, notably Enterococci and *Clostridium* sp., are found to be the core microbiome in the larval gut of *S. littoralis* [74]. In particular, *E.*

mundtii has been shown to dominate the gut of *S. littoralis* and colonize it in all developmental stages [74,195,199].

Enterococcus mundtii is a gram-positive, non-motile lactic acid bacterium, well adapted to dairy and plant environments [200]. It is found on the human naval, cow teats and the hands of milkers; in soil and in the gastrointestinal tracts of humans, animals and several species of Lepidopterans, namely *Galleria mellonela* and *Plutella xylostella*[201,202]They can exert probiotic, positive effects which have been shown in humans [203]

Antimicrobial activity has been shown for several Enterococci species, particularly *E. mundtii* isolated from a lepidopteran insect. *E. mundtii* produce an antimicrobial peptide, mundticin KS, that keeps potential pathobionts like *Enterococcus fecalis* and *Enterococcus casseliflavus* at bay. These pathobionts are apparent in first-instar larvae, but their early colonization success is brief, owing to mundticin [74,195]. Although Enterococci are thought generally to regulate insects' gut microbiome, their specific contributions remain largely unknown. Larvae of several Lepidopteran species produce high concentrations of 8-hydroxyquinoline-2-carboxylic acid, an iron chelator that is derived from tryptophan and found in the larval gut and regurgitate [204]. Since iron is one of the main elements in several metabolic pathways, such as those responsible for the quenching of reactive oxygen species, oxygen metabolism in TCA cycle, electron transport and nitrogen assimilation among others, this chelator is assumed to control the microbiome in larval guts [205].

In this paper we used GFP-tagged *E. mundtii* to visualize how the reporter microbe adapts to the gut environment of the host insect

5.2 Results

The bacterial strain *Enterococcus mundtii*, a dominant symbiont of *S. littoralis*, was employed as a reporter organism in order to follow its colonization of the insect gut. The approach provides direct information on the mode and pathways required for the bacteria to adapt to the adverse conditions encountered. GFP-tagged bacteria [176] were fed to second-instar larvae. At fifth instar, flow cytometry was used to sort the reporters to compare their gene expression with those of *E. mundtii* grown in vitro (supplementary S1).

5.2.1 Sorting of GFP-tagged *E. mundtii* cells from the gut of experimentally colonized *S. littoralis* larvae by flow cytometry

After *E. mundtii* exposed to the gut conditions of *S. littoralis* larvae were sorted and isolated using flow cytometry, their transcriptomes were compared to those of bacteria grown in Todd Hewitt Broth (THB).We chose THB-cultured *E. mundtii* grown in a shaker incubator at 37 degree Celsius and 220 RPM as a control because these are ideal, stress-free conditions. In THB, a complete medium, bacteria grow reliably, using dextrose as the source of energy. Since the *S. littoralis* foregut is alkaline and hindgut, neutral, we focused on *E. mundtii* growing at the two terminals.

From the gut homogenates containing the fluorescent reporter *E. mundtii*, 250,000 fluorescent cells were sorted by a flow cytometer. The collected cells constituted 2 to 4% of the total homogenate. In addition, for comparison, 250,000 fluorescent *E. mundtii* cells grown *in vitro* were sorted and for differential gene expression was analyzed (Fig. S1).

5.2.2 RNAseq analysis revealed many differentially expressed genes between *E. mundtii* growing in vitro vs. in vivo

To understand the mechanisms underlying the process by which *E. mundtii* adapt to (successfuly colonize) the fore and the hind guts of *S. littoralis* larvae, we analyzed gene expression between bacteria growing in vitro and in vivo. The RNA extracted from the FACS (Fluorescent Activated Cell Sorting)-sorted *E. mundtii* cells was sequenced using the Illumina Ultra-Low Input RNA kit, and the resulting 10 million short reads per treatment and replicates were processed and aligned against the fully sequenced genome of *E. mundtii* QU25 [206]. Supplementary Table S2 shows the alignment percentages of these reads against the genome.

The numbers of significantly up- and downregulated genes between *E. mundtii* cells exposed to different *S. littoralis* gut sections is shown in Table 1. Out of 2696 assembled genes, [284] and [275] genes are significantly differentially regulated (fold change = 2, $p \leq 0.05$) in *E. mundtii* in the fore- and hindgut, respectively. The density plot in Fig. S3(a) (Supplementary) shows the distribution of differentially expressed genes in foregut, hindgut and control.

There are 168 genes in common between the *E. mundtii* exposed to the fore- and hindguts that are differentially regulated when compared to the control. Most of these common genes

belong to adaptive processes required by *E. mundtii* to colonize by adhering to the gut wall, avoid stresses, and to acquire iron and complex carbohydrates. The list of these genes with functional annotation is shown in Appendix 2 (Fig.1).

To test for biological and technical variability, individual replicates were analyzed, and a PCA plot (Fig. 2) and dendrogram (Fig. S3(b)) were generated. The gene expression profiles of *E. mundtii* from the insect gut and the control form separate clusters and nodes.

	***E. mundtii* in foregut vs. control ($p < 0.05$)**	***E. mundtii* in hindgut vs.control ($p < 0.05$)**	***E. mundtii* in foregutvs. hindgut ($p < 0.05$)**
No. of genes upregulated	**91**	**71**	**28**
No. of genes downregulated	**193**	**204**	**60**
Total no. of genes ($p < 0.05$)	**284**	**275**	**88**

Table 1: Number of significantly differentially expressed genes – up- and downregulated ($p \leq 0.05$) -- in *Enterococcus mundtii* compared according to the following conditions: *E. mundtii* living in foregut vs. control, hindgut vs. control and foregut vs. hindgut.

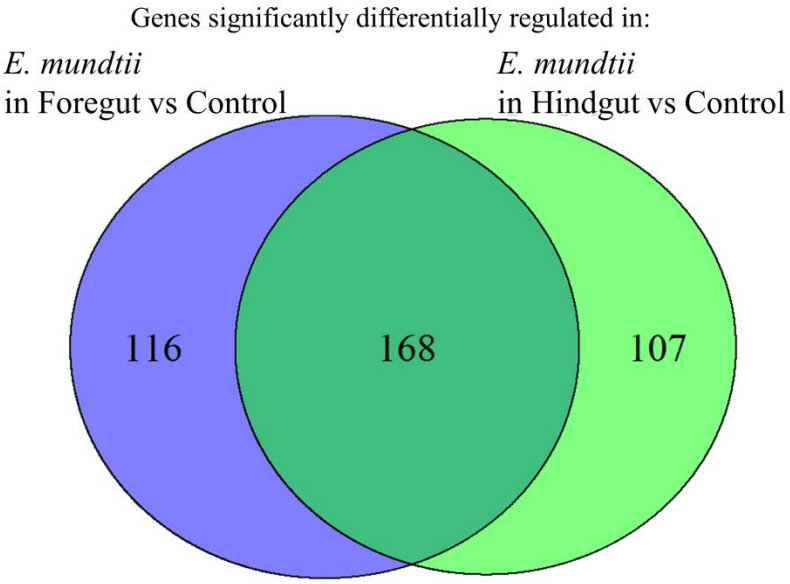

Figure 1: Venn diagram showing overlap of significantly differentially expressed genes (Appendix 2) in the following two conditions: *E. mundtii* living in foregut vs. control, and *E. mundtii* living in hindgut vs. control.

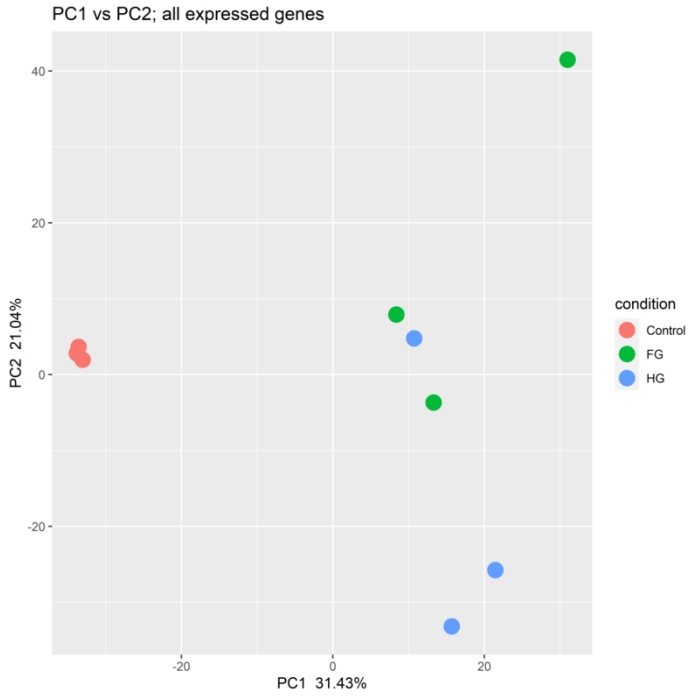

Figure 2: PCA plot showing clustering of the transcriptomic profiles among the three replicates of *E. mundtii* obtained from the foregut (FG), hindgut (HG) and control.

5.2.3 Gene enrichment analysis revealed several pathways differentially expressed between *E. mundtii* growing in vitro vs. in vivo

The differentially expressed genes were subjected to pathway analysis to determine the up- and downregulated pathways in *E. mundtii* when they are adapting to the gut conditions. Hence, the genes with functional annotation were classified according to three categories of gene ontology: molecular function, biological process and cellular component. We discuss only the category "biological processes" because it highlights the major pathways of *E. mundtii* living in the gut of the host.

To classify assembled genes with functions into different pathways, we used clusterprofiler R package. Gene annotation information of *E. mundtii* was obtained from the KEGG-FTP server and used to categorize the differentially expressed genes from our results into pathways, followed by an enrichment test by the clusterprofiler function Enricher (), (p-value cut-off= 0.05). Out of 2696 assembled genes of *E. mundtii*, 1590 were functionally annotated and classified according to pathway. Of the 284 and 275 (Table 1) differentially regulated genes (p-value cut-off= 0.05) in fore- and hind guts, respectively, 199 and 190 were functionally annotated in the category of biological processes. The pathways that are significantly enriched (p-value cut-off= 0.05) are shown in Fig. 3; percentages were calculated as such: number of genes up- or downregulated in a pathway divided by the total number of genes of that pathway that were annotated in the category. This fraction of up/down regulated genes in each significantly enriched pathway (p-value cut-off= 0.05) is shown in y-axis in Fig. 3.

The upregulated genes in both fore- and hindguts represent several pathways, including the reductive TCA cycle, nucleotoide biosynthetic process, carbohydrate metabolic process, peptidoglycan turnover, starch and sucrose metabolism and transmembrane transport (Fig. 3)

There are several notable enriched pathways in the hindgut: lysine biosynthesis via the diaminopimelate pathway might indicate the bacteria are producing the amino acid (Fig. 3, S4); cell adhesion, which could indicate that the bacteria are adhering to the host epithelium to keep from being flushed out of the host gut; and oxidative stress response.

Not only the synthesis of amino acids, such as phenylalanine, glutamate, tyrosine and tryptophan (though not lysine), but also of fatty acids (shown by the downregulation of acetyl CoA carboxylase activity, malonyl CoA biosynthetic activity) and metabolism in general

seem to be downregulated in the symbiont. Moreover, when *E. mundtii* lives in the gut, a down regulation of fatty acid biosynthesis is accompanied with enrichment of fatty acid degradation (Fig.3, Supplementary S4, S5). We hypothesize that, by obtaining these by-products from the host, symbionts avoid the energy costs associated with these processes of fatty and amino acid biosynthesis.

The genes involved in some of the important enriched pathways are discussed in detail in the next section (Fig. 4).

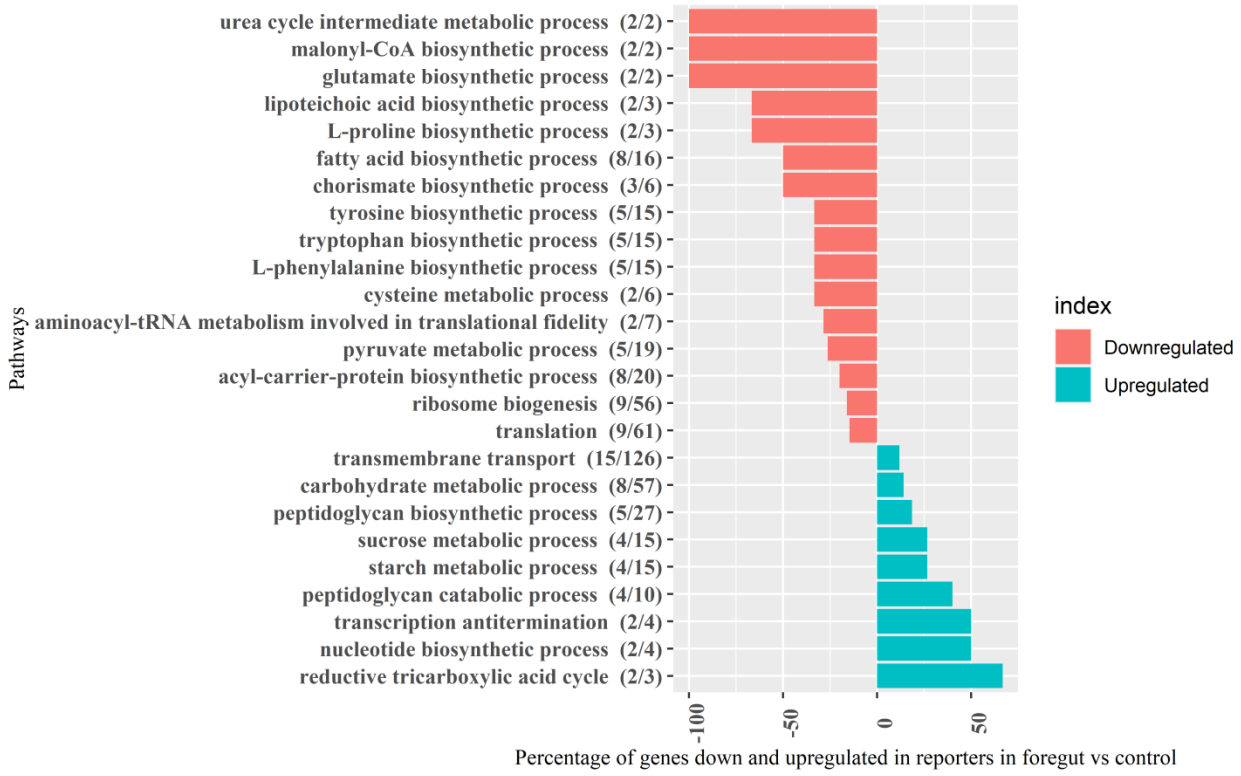

Figure 3: Summary of gene ontology classification in the category of biological processes, after an enrichment test (p-value cut-off= 0.05). The graph shows both up- and downregulation of the assembled genes of *E. mundtii,* with functional annotations, classified into enriched pathways, obtained from foregut (a) and hindgut (b), compared to genes of the

control. The percentages of each pathway refer to the percentage of genes of that particular pathway that are enriched in *E. mundtii.* The fractions of the same are denoted next to the pathways on y-axis.

5.2.4 Survival strategies of *E. mundtii* in the gut of *S. littoralis*

The differentially expressed genes that we identified are related to the adaptive strategies of *E. mundtii* in the fore- and hindguts of the larvae. We further classified these strategies in three broad categories: extracellular interactions, stress responses and metabolism, based on the results of the enriched pathways obtained in the previous section. Certain genes and their fold changes mentioned below are listed in Appendix I.

Extracellular interaction between *E. mundtii* and the gut epithelial layer of *S. littoralis*

The biological process category of gene ontology showed enrichment in the pathway of cell adhesion (Fig. 3). This motivated us to look deeper into the genes that control adherence to the host gut. Various well-characterized surface-associated proteins with conserved motifs and domains contribute to the ability of *E. mundtii* to attach itself to the gut epithelial tissue of its host. C-terminal conserved LPXTG motifs (EMQU_1297: 33-and 124-fold in the fore- and hindgut, respectively, a slight upregulation of *fms3*) and WXL domains (EMQU_0541:30- and 8-fold in the fore-and hindgut, respectively, and EMQU_0539:383-fold in the foregut).The *lysM* domain that helps in biofilm formation by is upregulated (EMQU_0157: upto 3-fold in the fore- and hindgut, respectively). The sticky matrix helps *E. mundtii* deal with stress efficiently [207,208].

Genes for chitin-binding proteins form a class of surface-associated proteins that provide adhesive properties to lactic acid bacteria so that these can adhere to the N-acetyl glucosamine component of chitin present in insects' gut epithelial cells, especially the cells lining the midgut [209]. Two of these proteins show levels as high as EMQU_0940:47- and 138-fold and EMQU_1285:25- and 69-fold, in the fore- and hindgut, respectively. Lipoproteins are placed in defined subcellular spaces formed by the plasma membrane. Their position is convenient for capturing incoming nutrients or elements such as iron. In addition, lipoproteins have been shown to help bacteria adhere to host cells [210]. EMQU_0428 is upregulated 5- and 4-fold in the fore- and hindgut, respectively. EMQU_2743 is upregulated 7-fold in the hindgut. Both are zinc transporter lipoproteins (Fig. 4(a), Appendix 1).

Stress responses of *E. mundtii* dwelling in the gut of *S. littoralis*

E. mundtii seems to be modulating their gene expression in response to the various stresses as was also seen with the upregulation of oxidative stress response in the hindgut (Fig. 3). Accordingly, they upregulate several antioxidant enzymes: superoxide dismutase (13- and 8-fold in the fore- and hindgut, respectively), catalase (EMQU_0568: 4- and 10-fold in the fore- and hindgut, respectively), NADH oxidase- peroxidase cycleEMQU_0335, 0459, 1279: up to 4-fold in the hindgut), organic hydro peroxide resistance family protein (EMQU_1453: 6-fold in the fore- and hindgut), and peptide-methionine (R)-S-oxide reductase (EMQU_0165: 3-fold in the hindgut) [211].

The *agr* two-component systems that may bring about quorum sensing in bacteria show upregulation in both the fore- and hindgut. Levels of *agrA* are upregulated about 3-fold in the hindgut and for *agrB,* about 5- and 8-fold in the fore- and hindgut, respectively.

Genes for general stress proteins (*glsB*: 32- and 97-fold; *glsB1*: 10- and 7-fold; *gls33*: 6-and 22-fold, in the fore- and hindgut, respectively) and universal stress proteins (USPs) (*uspA2*: 54-and 11-fold in the fore- and hindgut, respectively) are upregulated in *E. mundtii* in response to environmental conditions such as the presence of salt, oxygen or oxidative stresses, and toxic substances, and nutrient starvation. The expression of USPs may depend on the increased bacterial density brought about by quorum sensing [212].

Intracellular trafficking, secretion, and vesicular transport include *secE* (22- and 16- folds in the fore- and hindgut, respectively) needed for cell viability, and *virD4* (EMQU_1288: 47- and 46-fold in the fore- and hindgut, respectively) components of the type IV secretion system, all of which are upregulated [213].

Also upregulated: genes for repair proteins, such as *mutS*(EMQU_2803)and *recA* (EMQU_2752: 3-fold in the foregut) conferring DNA mismatch repair and its protection from oxidative stress; *recF* (2- and 3-fold in the fore- and hindgut, respectively) for recombination repair, whose general role is the maintenance of DNA; DNA alkylation repair protein (*alkD*) (upregulated 3-fold in the fore- and hindgut); *radA* (3-fold in the fore- and hindgut) and *radC* (3- and 6- folds in the fore- and hindgut, respectively), proteins helping in DNA repair and recombination [214]; *yafQ* (EMQU_3002) and DNA damage-induced protein J (EMQU_3001, 33- and 4- folds in the fore- and hindgut, respectively), which constitute a toxin-antitoxin system that plays a role in biofilm formation [215] (Fig. 4(b), Appendix 1).

Iron homeostasis and alkaline stress- Iron homeostasis in *E. mundtii* is important especially in environments that are iron depleted owing to the presence of compound 8-HQA. These bacteria have upregulated their *fetC* permease gene (7- fold in the foregut and 11- fold in the hindgut) to increase their ferric uptake and Fur family transcriptional regulator (EMQU_1067: 4- fold in the foregut) to maintain iron homeostasis. Adaptation that is mediated through Fur and iron uptake is common in iron-deprived environments (Fig.4(d)), Appendix 1)[216,217]. The highly alkaline pH characteristic of the larval foregut in particular is a challenge to bacteria in general but also to *E. mundtii* specifically. For example, alkaline pH has been proven to unwind the double helical structure of DNA [218]. In addition, high expression levels of the alkaline stress protein have been found in *E. mundtii* living in the alkaline foregut (5- folds), whereas its expression decreases in the neutral conditions of the hindgut (Fig.4(b))

Metabolism carried out by *E. mundtii* when they are in the gut of *S. littoralis*

Facultative anaerobes can switch between respiration and fermentation, based on oxygen availability. The expression of most glycolytic genes – for example, glucokinase (*glcK*), 1-phosphofructo kinase (*fruK*), 6-phospho-beta-glucosidase (*bglP, bglB, bglG*) phosphofructokinase A (*pfkA*) and glucose-6-phosphate isomerase in *E. mundtii* dwelling in the gut does not change much compared to the expression of genes in *E. mundtii* growing under control conditions, suggesting the glycolysis pathway is active. The same trend holds true for pyruvate dehydrogenase entering the citric acid cycle in aerobic conditions, along with lactate dehydrogenase (*ldhA* EMQU_2453). The protein that stimulates the fermentation of sugar (SfsA-EMQU_0871) under anaerobic conditions is upregulated 9- and 6-fold in the fore- and hindgut, respectively. Some alcohol dehydrogenases are upregulated to convert acetaldehyde to ethanol in the fermentation pathway (EMQU_1129:2-fold in the fore- and hindgut; EMQU_ 0525: 5- fold in the fore- and hindgut; and EMQU_0315: 3- and 4- folds in the fore- and hindgut, respectively). The acetyl CoA produced by pyruvate dehydrogenase does not significantly contribute to the production of fatty acids and amino acids, because both pathways are downregulated (Fig. 3, Supplementary S4).

Phosphotransferase systems (PTSs), which take up alternative source of sugars such as sucrose, ascorbate, mannose and, most important, cellobiose, are upregulated in *E. mundtii* in both the fore- and hindgut [219]. Cellobiose mostly comes from the plant products on which

the host is fed. The genes of at least 13 PTS cellobiose transporter-subunits are upregulated; EMQU_0876, a particular cellobiose-specific IIA component, is upregulated as high as 78- and 88-fold in the fore- and hindgut, respectively. Ascorbate is mostly taken up in the hindgut. On the other hand, fructose and lactose do not seem to be a popular source of energy (Appendix 1).

Upregulation in starch and sucrose metabolism (Fig. 3, Supplementary S4) is brought about by an increase in the sucrose-specific PTS transporter (EMQU_2136: 2- and 5- fold in the fore-and hindgut, respectively) and sucrose 6-phosphate dehydrogenase (*scrB*: 2-folds in the hindgut); and the alpha-amylase enzyme neopullanase (EMQU_1435: 52- and 30-fold in the fore- and hindgut, respectively).

Although *E. mundtii* do not seem to invest energy in synthesizing fatty or amino acids, they seem to produce lysine in the hindgut via the diaminopimelate pathway [220] (supplementary S4, Fig. 3).

Metabolism and the transport of nucleotides in *E. mundtii* living in the gut increases, as are also seen in Fig. 3.

Regarding glycerol metabolism: the *glpF* gene required for glycerol uptake is downregulated(4-fold in the foregut), whereas the genes for metabolism -- *glpO, dhaKL, glpQ*-- are also expressed, suggesting these bacteria have an alternate way of obtaining glycerol [221] (Fig.4(c), Appendix 1).

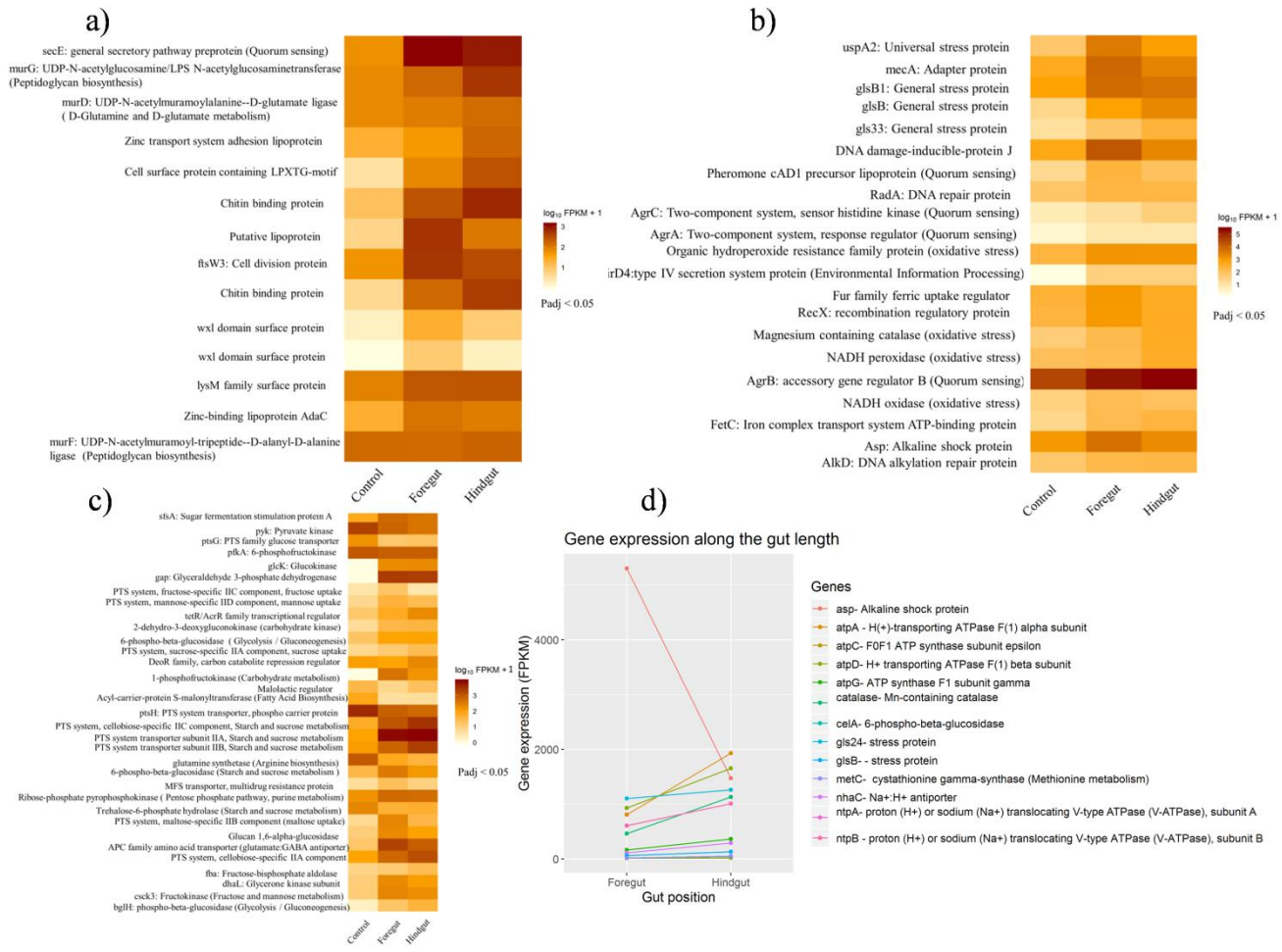

Figure 4: (a) Heatmap showing the regulation of genes involved in the attachment of *E. mundtii*, when these bacteria are in the fore- and hind-gut of *Spodoptera littoralis* larvae.(b) Heatmap showing the regulation of genes involved in the stress tolerance of *E. mundtii*, when these bacteria are in the fore- and hind-gut of *S. littoralis.*(c) Heatmap showing the regulation of genes in *E. mundtii* involved in metabolism, when these bacteria are in the fore- and hindgut of *S. littoralis.*(d) Graph showing the regulation of certain pH-related genes in *E. mundtii* living in the fore- and hindgut of the *S. littoralis* larval gut.

5.3 Discussion

This work focuses on the survival strategies of *E. mundtii* in the larval gut of *S. littoralis*, an environment threatened by stressful conditions, namely high pH, low iron content and oxidative stress. This makes it a good system to study adaptation by the symbionts in the larval gut. By sending the GFP-tagged reporter *E. mundtii* [176] to the larval gut, we were able to study how this dominant bacterium adapted to its new environment. The fluorescent

bacteria were later retrieved from the fore- and hindguts of the larvae using flow cytometry (supplementary S1, Fig. S1). To prevent any metabolic changes from occurring between the individual experimental steps of larval dissection and FACS sorting, we used RNAlater and RNAprotect reagents. Comparing the gene expression profiles of these retrieved reporters with the profiles of *E. mundtii* grown under optimal culture conditions, we were able to obtain a snapshot of the genes and the pathways that help these symbionts to survive in and adapt to the gut of *S. littoralis* larvae. The transcriptional changes found in these bacteria are an amalgamation of these factors which illustrates how *E. mundtii* is responding to stress and colonizing its host gut (Fig. 5).

For bacteria to successfuly colonize the host gut, they must prevent themselves from being flushed out of the system; adherence to gut tissue ensures they survive epithelial turnover [208] (Fig. 4). Biofilms, which are composed of adhering proteins, were first seen by FISH imaging [25]. LPXTG is a sortase-dependent site for anchoring proteins covalently attached to the peptidoglycan [222]. Lipid-anchored proteins or lipoproteins, which constitute another class of covalently associated adhesion proteins [223], are upregulated in *E. mundtii*. Wxl domains and LysM, or lysine-dependent motifs binding to the peptidoglycan, form non-covalent associations with the peptidoglycan [207,224]. Such associations occur in *Enterococcus fecalis* [225]. Chitin, a major part of the peritrophic matrix, lines the midgut epithelium of the host [226]. Chitin-binding proteins in *E. mundtii* also promote adherence to the host gut. Several bacteria, such as *L. monocytogenes*, adherent *E. coli* and *V. cholerae,* were found to initiate adhesion in the host gut by using their chitin-binding proteins [227]. Peptidoglycan turnover is a sign of active cell division (Fig. 3). Peptidoglycan biosynthetic and catabolic processes show upregulation in both the fore- and hindgut. The N-acetylmuramoyl-L-alanine amidase enzyme in the foregut helps in cell separation during division. It also aids in cell motility and establishing a symbiotic association with the host [228].

E. mundtii dwelling in the gut employ various strategies to survive adverse conditions (Fig. 5). Reactive oxygen species (ROS) result from the reduction of oxygen. Thereafter, the dismutation product of the superoxide anion (O_2^-) is hydrogen peroxide (H_2O_2). O_2^- and H_2O_2, along with the hydroxyl radical, are potent oxidants that can remove electrons from DNA, proteins, lipids, other macromolecules, which can damage both the invading and resident symbionts [229]. Lactobacilli employ enzymes such as NADH oxidase/peroxidase, superoxide dismutase and manganese-dependent catalase to counteract ROS, as was also true for *E. mundtii* [211].

Universal stress proteins are found in many bacteria; these proteins aid the adaptation of bacteria to stresses such as extreme temperature, oxidative loss, nutrient starvation and toxic agents [212]. In *E. coli*, stress proteins were first reported in fungi, archaea, plants and flies [230]. In *Burkholderia glumae*, genes that regulate universal stress protein are controlled by quorum sensing [231]. Confronted with stress, *E. mundtii* seems capable of behaving like a multicellular organism. The bacteria rely on quorum sensing as a survival strategy, aggregating on the host epithelia and forming a biofilm in the host gut. That *agrABCD* forms a two-component system and brings about quorum sensing has already been established in the Firmicutes *Staphylococcus aureus* and *Streptococcus pneumoniae* [232]. The adherence properties of *E. mundtii* may help it to form a biofilm layer on the gut wall. Thus, these two inter-related phenomena of quorum sensing and biofilm formation help bacteria to adapt to altered environments.

As discussed, 8-HQA is an iron chelator, and the larvae's ability to produce it may help them survive in an iron-depleted environment. The FetC iron complex transport permease and Fur family of transcriptional regulators may act in similar ways. FetC was found to be involved in iron homeostasis in *Apergillus fumigatus* [216]. Fur-dependent iron-acquisition system was upregulated when *Clostridium difficile* tried to infect hamsters in iron-depleted conditions [217]. It interacts with iron to determine its intercellular levels, hence bringing a halt to processes of iron dependant oxidative damage. Fur is also a major regulator of adaptation of bacteria to various hosts. They not only regulate iron homeostasis, but also mediate key adaptive responses as stress resistance, quorum sensing and biofilm formation. We presume, this could be very much the case with *E. mundtii* trying to adapt to the new living conditions of *S. littoralis* gut [233].

Alkaline shock proteins help the bacteria to adapt to extreme stress conditions [234]. Owing to the alkaline environment of the foregut [181], the *E. mundtii* living there express alkaline shock proteins as protection[181]. Such is also the case in Staphylococcus aureus[235]. Previous studies reported several genes differentially expressed in *E. faecalis* grown under alkaline conditions; similar expression patterns characterize *E. mundtii,* if alkalinity is the only factor taken into consideration. For example, we found a downregulation of methionine transport and synthesis systems, $Na^{+}H^{+}$ antiporter (NhaC family, 1-fold downregulation), upregulation of adenosine and cytidine deaminases (upto 19-fold), purine and pyrimidine metabolism. The

expression levels of Cation/H^+-related F and V-type antiporters (atp and ntp family proteins) are reduced under alkaline conditions (Fig.4(d) Appendix 1)[236] .

As facultative anaerobes, *E. mundtii* often initiate fermentation inside the host gut (Fig. 7). *E. mundtii* is found in the vicinity of the host gut surface only when some amount of oxygen is present, and that no *E. mundtii* is found in the inner layers of the anaerobic gut wall [25] highlights the low oxygen levels that characterize the gut lumen of most insects [237]. As pathway analysis clearly shows, the white-bean-based artificial diet that the host is fed on favors starch and sucrose uptake through PTS transporters and metabolic systems (Fig. 3, S4). PTS transporters help all bacteria survive environments with different levels of sugars [219]. Enriched nucleotide metabolism suggests that *E. mundtii* are striving to colonize the gut of *S. littoralis*. Previous studies with mice models showed that *E. coli* enriched their metabolism of purine and pyrimidine when colonizing the intestines of mice [238]. Although *E. mundtii* likes to reduce the energy they expend on their fatty acid and amino acid metabolism, their lysine metabolism is upregulated by bacteria living in the hindgut [239]. Whether *S. littoralis* is obtaining lysine from their symbiotic *E. mundtii* is a matter for further research. Pathway analysis shows lysine synthesis is enriched via the diaminopimelate pathway. Diaminopimelate also plays roles in peptidoglycan synthesis [240] (Fig. 3, Supplementary S4, S5).

Our data on *Enterococcus mundtii* agrees with several examples of how symbionts function in their respective host guts. Colonization of symbionts by extracellular interaction between the gut cells and the symbiont, by overcoming various stresses induced by hosts, and by changing metabolism to fit the nutrient-limiting conditions in the gut was also seen in *Snodgrasella alvi* in the gut of honey bees. Genes for biofilm formation, facing oxidative stress, fluctuating pH and repair proteins were upregulated in the symbiont in vivo [241]. The microbiota of cockroach mid-gut is also engaged in digestion of complex carbohydrates with the help of amylase enzymes, along with responding to oxidative stress by upregulating genes involving peroxidase and catalase [242]. Likewise, gut microbiota hold a record of digesting recalcitrant carbohydrates in plant or wood-feeding insects. Termites form a classic example where the cellulolytic activity of bacteria residing in the hindguts of higher termites was detected [243]. Aerotolerant intestinal symbiont *Bacteroides fragilis*, upon facing an oxidative environment, immediately react to it to prevent the immediate effects of reactive oxygen species, and also regulate their biosynthetic processes accordingly [244].

High-throughput transcriptome sequencing from tiny quantities of starting material has revealed the strategies used by *E. mundtii* to survive the gut of *S. littoralis*. Our methods can be used to study interactions between any host and its symbiont. For example, fluorescently tagged bacteria can be introduced into the insect guts in which the 8-HQA-producing gene has been knocked out. A similar method will allow us to study the behavior of the retrieved bacteria and shed light on the mechanisms of survival that underlie the exchanges between symbionts and their genes.

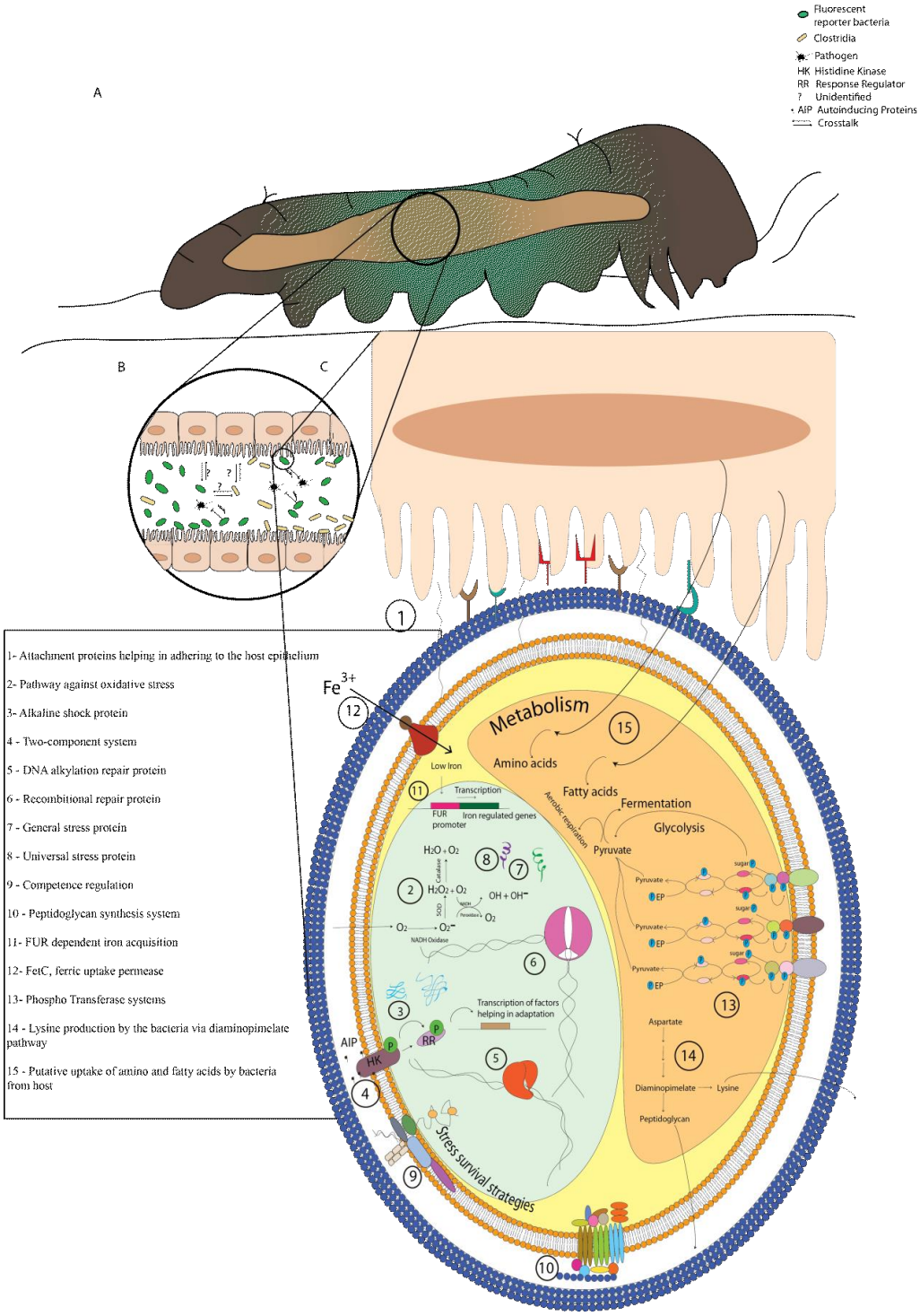

Figure 5: A snapshot of interactions between *Spodoptera littoralis* and its resident gut symbiont, *E. mundtii*. (A) An illustration of *S .littoralis* with its longitudinal gut (B) *E. mundtii* dominates in the gut along with Clostridia and keeps pathogens at bay by producing mundticiin KS. Unknown interactions occur among these two symbionts and the host gut. (C) Pathways and stress survival strategies of *E. mundtii* in the gut of *S. littoralis*.

5.4. Materials and methods

Maintenance of eggs and larvae

The eggs of *S. littoralis* were obtained from Syngenta Crop Protection Munchwielen AG (Munchwielen, Switzerland). Eggs were hatched at 14^{o} C and the larvae were maintained at 24^{o} C in an alternate 16 hours light period and 8 hours dark period. Larvae were reared on an agar-based artificial diet containing white beans, as described by Maffei et al [245].

Bacterial strain

A fluorescent strain of *E. mundtii* KD251 (isolated from the gut of *S. littoralis* in the Department of Bioorganic Chemistry) was constructed by transforming a GFP-containing expression vector pTRKH3-ermGFP, as described [199]. This strain was grown in Todd-Hewitt Bouillon(THB) (Roth, Karlsruhe, Germany) medium for both broth and 1.5% agar (Roth, Karlsruhe, Germany), and in the presence of 5 μg ml^{-1} of erythromycin (Acros Organics, NJ, USA). The strain was preserved as a glycerol stock at -80^{o} C.

Introduction of the reporter bacteria into the insect microbiome

A stationary phase culture of fluorescent reporter *E. mundtii* in THB broth containing 5 μg ml^{-1} of erythromycin was grown till mid-log phase with OD_{600} ~ 0.5-0.6 at 37ºC with shaking at 220 rpm. The culture was pelleted at 5000 x g for 10 minutes at 4ºC and resuspended in distilled water. First-instar *S. littoralis* larvae (n = 120) were fed small cubes of artificial diet supplemented with two antibiotics, ampicillin (5.75 μgml^{-1}) (EMD Millipore corp., Billerica, MA, USA) and erythromycin (9.6 μgml^{-1}) for 3 days, to reduce the already existing bacterial load, before (at the second instar) being fed with100 μl from the 1:10 dilution broth (~10^{10} cells) containing fluorescent *E. mundtii* as described [199]. These larvae were allowed to grow until the fifth instar, when samples were prepared for FACS.

Sample preparation for FACS

A total of 30 fifth-instar larvae for each gut region -- foregut and hindgut -- were dissected with sterile forceps and scissors in a sterile clean bench. Following dissection, the gut tissues

were immediately submerged in 10 ml of RNAlater solution (Invitrogen, Vilnius, Lithuania). Tissues submerged in RNAlater solution were mixed with 2 ml of 6% (w/v) betaine (Sigma Aldrich, St. Louis, MO, USA) and placed on ice prior to being crushed with mortar and pestle until gut homogenates were formed. Thereafter, fluorescent *E. mundtii* were separated from the intestinal debris by filtration through 40 µm pore-size cell strainers (Falcon, NY, USA). The filtrates were then separated into aliquots of 600 µl each and kept at -80˚C for the sorting experiment.

As controls, *E. mundtii* broth cultures (10 ml, n = 3) were grown to exponential growth (OD600 ~ 0.5-0.6) and centrifuged at 5000 x g for 15 min at 4ºC to pellet the bacterial cells. Bacterial cells were washed once with sterile phosphate-buffered saline (PBS) and resuspended with RNAlater to a concentration of approximately 10^{10} CFU ml^{-1}.

Cell sorting by FACS

The gut homogenates were analyzed using BD FACSAria™ Fusion Cell Sorter (Becton Dickinson, Heidelberg, Germany). It relies on an ion laser emitting a 488 nm wavelength, and a 502 long pass filter, followed by a 530/30 band pass filter. The green fluorescent protein emits light with a peak wavelength of 530 nm. Prior to loading each sample in the FACS machine, the homogenate was thawed, and 1:5 dilution of the homogenate was made in sterile PBS, followed by vortexing for 10 seconds for proper mixing and to dislodge the bacteria from tissue. The cells were sorted at a flow rate ranging from10 µl/min to 80 µl/min. The sorting was done in a single-cell mode, and the sorted cells were collected in 5 ml sterile polypropylene round-bottom tubes (Falcon, Mexico). The cells were collected for a period of 3 hours, which corresponded to an acquisition of 6000-7000 events/sec. The flow cytometry grade of PBS buffer (Thermo Fischer, Wilmington, DE, USA) at pH of 7.4 was used as the sheath fluid. A total of ~ 250, 000 cells were sorted from each sample into 1 ml of RNA Protect solution (Qiagen, Hilden, Germany). A total of ~ 250, 000 cells were sorted from each sample of control, fore and hindgut homogenates into 1 ml of RNA Protect solution (Qiagen, Hilden, Germany)

RNA extraction and sequencing

As controls, *E. mundtii* broth cultures (10 ml, $n = 3$) were grown to exponential growth (OD_{600} ~ 0.5-0.6) and centrifuged at 5000 x g for 15 min at 4ºC to pellet the bacterial cells. Bacterial cells were washed once with sterile phosphate-buffered saline (PBS) and resuspended with the same buffer at a concentration of approximately 10^{10} CFU ml^{-1}. The FACS-sorted fluorescent bacterial cells (~ 250,000) from each foregut and hindgut were pelleted by centrifugation at 5000 x g for 10 min at 4ºC, leaving insect cell debris in the supernatant. The foregut, hindgut and *E. mundtii* cultures were each represented by three biological replicates ($n = 3$). RNA was later removed from the sorted cells prior to RNA isolation, and total RNA was isolated from the pelleted cells using the RNeasy mini kit (Qiagen, Hilden, Germany) following the manufacturer's instructions, with some modifications. Pelleted bacterial cells were lysed enzymatically for 15 min at 37ºC (enzymatic mix: 1X TE buffer, pH 8 (Applichem GmbH, Darmstadt, Germany), pH 8.0, 5 µg ml^{-1} lysozyme (Sigma Aldrich, St. Louis, MO,USA) and 50 Uml^{-1}mutanolysin (Sigma Aldrich, St. Louis, MO,USA)). All samples were DNase-treated with on-column DNase digestion per the manufacturer's protocol prior to RNA isolation. The concentration of total RNA in controls was diluted to match the bacterial concentration at the level of a single cell. RNA was further cleaned and concentrated using a concentrator kit (Zymo Research, USA) and yielding about 12 µl in final volume (~10 ng). The purified RNA was linearly amplified using MessageAmp II bacterial RNA amplification kit (Invitrogen, Vilnius, Lithuania) and 10 ng of total RNA following the manufacturer's instructions. The amplified RNA (aRNA) was concentrated by precipitation with 5M ammonium acetate. The quality and quantity of the total RNA was measured with a NanoDrop One Spectrophotometer (Thermo Scientific, Wilmington, DE, USA) and an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA, USA). RNA samples were sent to the Max Planck Genome Centre in Cologne for RNA sequencing. A total of 0.3 µg - 1 µg of amplified RNA was used for cDNA library preparation using the Ultra-Low Input RNA kit following the Illumina protocol at the Max Planck Genome Centre, Cologne. Sequencing was carried out on the HiSeq 2500 sequencer at Cologne, and a total of approximately 10 million paired-end reads (2 x 150 bp) were generated for each sample.

RNA-seq data analysis

FastQC was done for an initial quality analysis of the reads. Analysis of the reads, including trimming of adapters and differential gene expression analysis, was done on LINUX-based Command line interface, following the Tuxedo protocol [246]. The adapters were trimmed using Trimmomatic 0.36; trimmed reads were assembled using Tophat 2.1.0 and mapped to the genome of *E. mundtii* QU25 [206] using Cufflinks 2.2.0. The read counts were normalized with FPKM (fragments of kilobase of transcripts per million mapped reads) (supplementary S7), and assemblies were merged using Cuffmerge. Cuffdiff was used to compute the differentially expressed genes between *E. mundtii* from the larval gut and *E. mundtii* grown *in vitro.* Based on homology to protein families, the proteins that were predicted for *E. mundtii* were categorized under gene ontology terms (http://geneontology.org). The genes were also mapped to the KEGG database to predict the pathways (supplementary). Gene annotation information of *E. mundtii* was obtained from the KEGG-FTP server. The results of differentially expressed genes were visualized using R-package CummeRbund 2.0, on R version 3.3.3 (2017-03-06). This R-package generated all the plots: dendrograms, PCA plot and heatmaps. A fold-change of ≥ 2 was used as a threshold to analyze the differentially expressed genes. Pathway analysis was performed using the R-package, clusterProfiler[247]. The enricher () and enrichKEGG () functions performed enrichment tests with gene ontology categories and KEGG databases, respectively, and grouped enriched pathways based on the number of significantly expressed (p value cut-off= 0.05) genes in the *in vivo* conditions as compared to control [247,248].

The raw transcriptome data has been deposited to NCBI Short Read Archive (SRA). The BioProject ID is: PRJNA622409

5.5 Acknowledgments

This work was funded by the Max Planck Society and Jena School for Microbial Communications. Also, we thank Angelika Berg and Andrea Lehr for technical assistance and Emily Wheeler, Boston, for editorial assistance. This manuscript has been released as a pre-print at BioRxiv [249].

6 Chapter III

Role of 8-Hydroxyquinoline-2-carboxylic acid in defining the bacterial landscape in the guts of *Spodoptera littoralis* larvae

6.1 Introduction

The concept of gut bacteria has been attracting attention ever since people realized in 400-500 BC that several disease originate from the gut. This led to volumes of research in recent years concerning bacteria in the gut of a multitude of species to attain a deeper understanding of what roles they could be playing in a host system [250]. There seems no end, no generalized theory that explains the story of every host-microbe combination, because each pairs have a unique story of their own. All the stories taken together surely generalize a few concepts towards the direction of understanding such an interaction. All higher animals harbor an ever dynamic community of microorganisms that they have acquired along their life-cycle, the community of which they control in a way that the pathogens are eliminated. The coevolution of microbiota with their hosts has led to intimate relationships between the two, bringing about associations like mutualism, commensalism and ammensalism. Such is the proximity between these two entities that it is rightly referred to as a supraorganism, with the gut bacteria acting as a major organ, influencing the health and fitness of their host. Coevolution of certain bacteria with their host has led to the reduction of their genome, or the loss of genes, the products of which they no longer require and depend on the host for [251].

The gut bacterial community is shaped along the life cycle of the organism, jointly by its diet, behavior, environment they are in, and host-genetics. The colonization of microbes in the gut begins at birth by vertical transmission, thus commencing the genetic control [252]. The effects of host genetics on the microbiome composition although is less clear, correlations between the two were found in mammals [252] and plants [253]. Several innate and adaptive immunity genes like TLRs and INFs have been shown to be shaping the gut microbiota in mice [254].SNVs in the MEFV gene are attributed to changing the gut microbial composition in humans [255]. Bacterial composition is more similar in monozygotic twins, than in dizygotic ones [256]. Such seems the influence of genotype on the microbial composition that specific host alleles may result in presence or absence of particular microbes in a way that could determine the health of the host [254].

Physical conditions of the host gut dictate the microbial composition as well. pH of the gut compartments, their shape that influences the oxygen content, antimicrobial peptides and certain molecules produced by the hosts all lead to a presence of certain families of microbes and absence of others. AMPs and molecules secreted by the host indirectly hint towards the host genotypic control of the microbial composition. In Drosophila, the production of two different classes of AMPs are influenced by the two immune pathways the IMD (Immune deficiency) and Toll-signaling pathways, the ones that control the Gram negative and the positives respectively. Without such maintenance, the gut bacterial load would increase by ten-folds [257]. Another immune effector working for the same cause is Reactive Oxygen Species (ROS) arising from the metabolic activities of host and other bacteria in the gut [258]. Dual oxidase or DUOX-derived ROS play a major role in gut bacterial composition of mosquitoes [259].

The gut microbiota of the cotton-leaf worm *Spodoptera littoralis* has been elucidated. Certain important factors influencing the dynamic population in their guts is a longitudinal gut structure, a pH gradient of alkaline to neutral along the length of the gut, and presence of a putative ion chelator, 8-hydroxyquinolone-2-carboxylic acid produced by the larvae in their gut regurgitate. This compound is derived by the host from tryptophan metabolism and is present in high concentration: 0.5-5 mM. Quinolinic carboxylic acid derivatives are widely distributed in nature. One of their usages is quorum sensing in bacteria. 8-Hydroxy-4-methoxyquinoline-2-carboxylic acid is a siderophore at the same time in *Pseudomonas fluorescence* [260]. Quinolinic carboxylic acids have also been found in plants [261]. 8-HQA was also found to be iron chelating. ESR measurements have found that the affinity of 8-HQA towards Fe^{3+} is higher than that of Fe^{2+}. Also, this binding is a function of the gut pH [262].

Since such compounds have been proven to have siderophoric roles, we wanted to look into the effects of this compound on the gut microbiota of *S. littoralis* larvae.

At a physiological pH iron exists as insoluble $Fe(OH)_3$. At the same time, it is an important microelement. It takes part in diverse cellular processes like nucleic acid synthesis, tricarboxylic acid cycle, electron transport, secondary metabolism like production of vitamins, antibiotics, sideropores, etc, and function as cofactors in proteins and enzymes. Hence, gut bacteria should come up with efficient methods to take up iron from their surroundings.

To test the hypothesis, it was required to knock down the production of the iron chelator 8-HQA. As already mentioned, 8-HQA is a product of tryptophan, generating via 3-Hydroxykynurenin (Fig 1). The enzyme kynuerenin monooxygenase is the main enzyme bringing about its formation[263]. The gene for this enzyme was knocked out using CRSPR/Cas9 method to reduce the production of 8-HQA to negligible amounts. The gut-microbiome of these insects were analyzed and compared with that of wild type ones to assess the roles of 8-HQA in dictating the bacterial population of the gut.

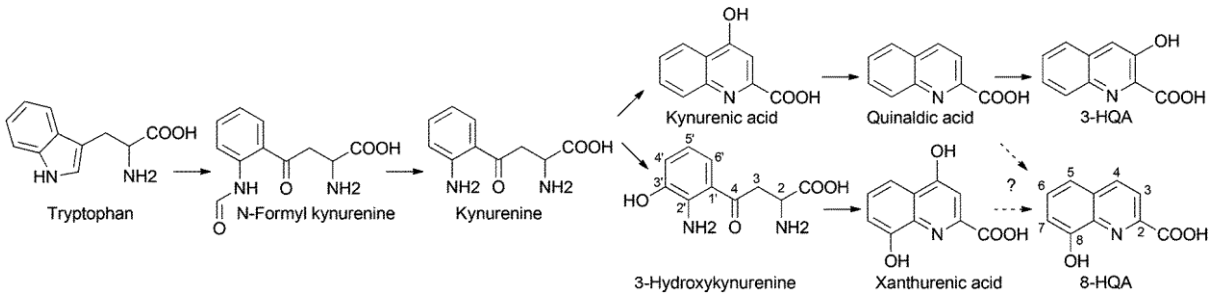

Figure 1: Biosynthesis of 8-hydroxyquinolone-2-carboxylic acid [263]

6.2 Materials and Methods

6.2.1 Insect rearing and sampling

Hatched *S. littoralis* eggs of the KMO knockout line and wild-type (WT) were obtained from the Department of Entomology, MPI-CE, Jena, Germany, where the knockout was carried out by CRISPR-Cas9 method. KMO knockout and WT larvae were grown in separate petridishes containing layer of white lima bean based artificial diet. This food was mixed with the frass from the larvae bought from Syngenta [25,176], because the previous experiments that looked at the bacterial population of *S. littoralis* larvae were conducted on this line of larvae [25], [176], whereas, the CRISPR-Cas9 knock out for Kynurenin monooxygenase (KMO) gene to inhibit the production of 8-HQA were carried out on a different line. Hence, the microbial population of the former line was allowed to colonize the latter, since it was required to compare the results of these two kinds of inhibitory methods that reduce the 8-HQA production. The larvae were separately reared and maintained based on family numbers to prevent them from inter-family mating.

5 larvae from second, third, fourth and fifth instar stages of each KMO mutants and WT lines fed with artificial diet spiked with frass from a different strain of *S. littoralis* with known bacterial population, were sampled and collected. The insects were frozen for 20 min before

the dissection, after which they were surface sterilized in 70% ethanol for three times, followed by rinsing in sterile water. All the samples were flash frozen in liquid nitrogen and were stored at -80˚C until DNA extraction.

A batch of KMO larvae was kept on an 8-HQA spiked diet (100 µl of 0.2 mg/ml 8-HQA was fed to each larvae), starting from second until the fifth instar. They are named as 5thKMO+8-HQA. 5 of such larvae were sampled at fifth instar.

6.2.2 DNA extraction

For DNA extraction, DNeasy PowerSoil kit (Qiagen) was used. Right before the extraction of DNA, each larva was put in separate PowerBead Tubes and homogenized with the help of TissueLyser LT (Qiagen) for 3 minutes at a frequency of 29 s^{-1}. The homogenate went through a 50 ug of Proteinase K (Ambion) treatment at 60 degrees overnight. The rest of the extraction was performed according to the manufacturer's instructions. The resultant DNA was purified using Zymoresearch DNA clean and concentrator-5 kit (Orange city, California). After the extraction and purification, the DNA concentrations of all the samples were measured using Nanodrop One and stored at -20˚C.

6.2.3 16S amplicon PCR

Each extracted sample DNA went through a quality control step which was a PCR amplification of the V4 region of the 16S rRNA gene. The PCR was conducted using forward F515 primer (5'-TATGGTAATTGTGTGYCAGCMGCCGCGGTAA-3') and reverse R806 primer (5'-AGTCAGCCAGCCGGACTACNVGGGTWTCTAAT-3'), using Thermofischer Scientific Taq DNA Polymerase (2 Units), 10mM dNTP (0.2 mM), 50mM $MgCl_2$ (1.5 mM) and 0.5µM of each forward and reverse primers. The PCR reaction cycle was an initial denaturation step at 94°C for 3 sec, followed 35 cycles of 94°C for 45 sec, 50°C for 30 sec, and a final elongation step at 72°C for 90 sec. Finally, an extended elongation step at 72°C for 10 min.

The DNA marker that was used to determine the band size was Gene Ruler 1 kb Plus DNA Ladder. The samples were run on 1.5% Agarose gel containing ~3-5 µl Midori green Advance DNA stain for visualization under UV illumination. The gel electrophoresis Bio-RAD Wide Mini-Sub® Cell GT chamber was used to run the samples at voltage of 150 mV

and 120 A current for 35-40 min. The gel was visualized using Gel Doc™ XR+ System and band of ~390 kb size was obtained.

6.2.4 16S amplicon sequencing and analysis

5 biological replicates of second, third fourth and fifth instar larvae from KMO and WT strains were sent for sequencing, MR DNA (Molecular Research LP), Texas. Sequencing of 2 x 300 paired-end kind was performed on Illumina MiSeq platform.

The QIIME1.0 platform was used to analyze the sequencing data of the gut microbial metagenome of wild-type and KMO knockout *S. littoralis* insects. OTUs were picked using the USEARCH method. Chimeric sequences were removed using USEARCH (version: 5.2.236). Representative sequences from each OTU were generated and mapped against the Greengenes 16S database (Greengenes Database Consortium, version13_8, 99%) to assign taxonomy. All samples were rarefied to the same number of reads. Singletons were filtered out from the OTU table. R packages: Phyloseq and ggplot2 were used to visualize the analysis of the metagenomics data. Alpha (Shannon index, simpson index, chao1 indices) and Beta diversity (NMDS using Bray-curtis distance matrices and PCoA using unifrac distances) were computed. For pairwise comparison of individual samples with the others, Pairwise Kruskal-wallis test ($P < 0.05$) was carried out for each indices of Alpha diversity.

6.2.5 Measuring levels of 8-HQA in KMO knockouts and WT larvae using GC-MS

10 fifth instar larvae each of KMO and WT conditions were starved for 5 hours and their spit were collected in individual GC vials. To collect the spit, each larva was pressed at the foregut area using a pair of forceps on a sterile petridish, and the spit was collected using a pipette. Weights of the spit were taken. The spit samples were derivatized using the method: "TBAS-Assisted Anhydrous Pentafluorobenzylation". This method involves addition of 250 ul of TBAS solution (0.1 M) in NH_4OH solution (1M), pH 10.4, drying in a SpeedVac, adding 400 µL of fresh PFBB/Amine in MeCN solution, heating at 65 degrees and mixing 100 µL decane and 3.5 mL water, and finally adding 1M sulfuric acid (1 ml) and centrifuging in a SpeedVac for phase separation.

Samples were analyzed by gas chromatography mass spectrometry (GC-MS) with an ITQ GC-ion-trap MS system (Thermo Scientific, Bremen, Germany) equipped with a fused silica capillary column ZB–5 (30 m × 0.25 mm ×0.25 µm with 10 m guard column, Zebron, Phenomenex, USA). Helium at $1mL \cdot min^{-1}$served as carrier gas with an injector temperature

of 320 °C running in splitless mode; 1 µL of sample was injected. Separation of the compounds was achieved under programmed temperature conditions (175°C for 1 min, then at 10 °C·min^{-1} to 320 °C kept for 1 min). The MS was run in EI mode (70 eV) with a scan range of 35 to 800 amu, a transfer line temperature of 320 °C, and an ion source temperature of 200 °C. Data acquisition was performed using Xcalibur 3.1(Thermo Fisher Scientific).

6.3 Results

Like already mentioned, The experimental group consisted of larvae from second, third, fourth and fifth instars, plus a sample of fifth instar larvae fed with 8-HQA. These larvae were fed with frass from a different strain of *S. littoralis,* whose bacterial composition is known. There were samples from two conditions- Kynurenin monooxygenase gene knockout (KMO) and wildtype (WT). These larvae had frass from a different strain of *S. littoralis* added to their artificial diet. This was done to colonize the larvae with a known bacterial population. A sample of fifth instar KMO larvae was fed additionally with 100 µl of 0.2 mg/ml of 8-HQA to mimic their gut environment of WT larvae (marked as 5thKMO+8HQA). The reason this concentration was chosen is because the spit of the 5thKMO+8HQA larvae gave peaks depicting similar levels of 8-HQA (area of peak of 8-HQA in the GC profile/weight of the spit of each larva) to that of wildtype larval spit when analysed in GC (Appendix 3).

6.3.1. 8-HQA measurements in KMO-knockouts as compared to WT larvae

The major aim of this work is to find a correlation of gut microbial levels with the presence and absence of 8-HQA in the guts of *S. littoralis* larvae. So, the first step was to confirm the absence of 8-HQA in the larvae knocked out of the gene of the main producer enzyme- Kynurenin monooxygenase (KMO). Figure 2 shows the relative levels of 8-HQA compared to the wild type (WT) counterparts. The KMO strains have negligible amounts of the compound.

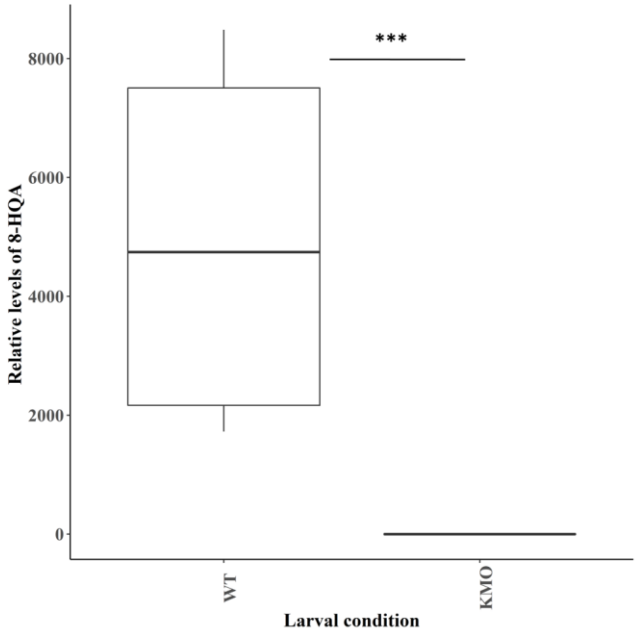

Figure 2: Relative amounts of 8-HQA mg^{-1} spit of the larvae in 5th instar wildtype (WT) and KMO gene knockout larvae.

6.3.2. Microbial diversity in WT and KMO knockouts across the life stages of *S. littoralis*

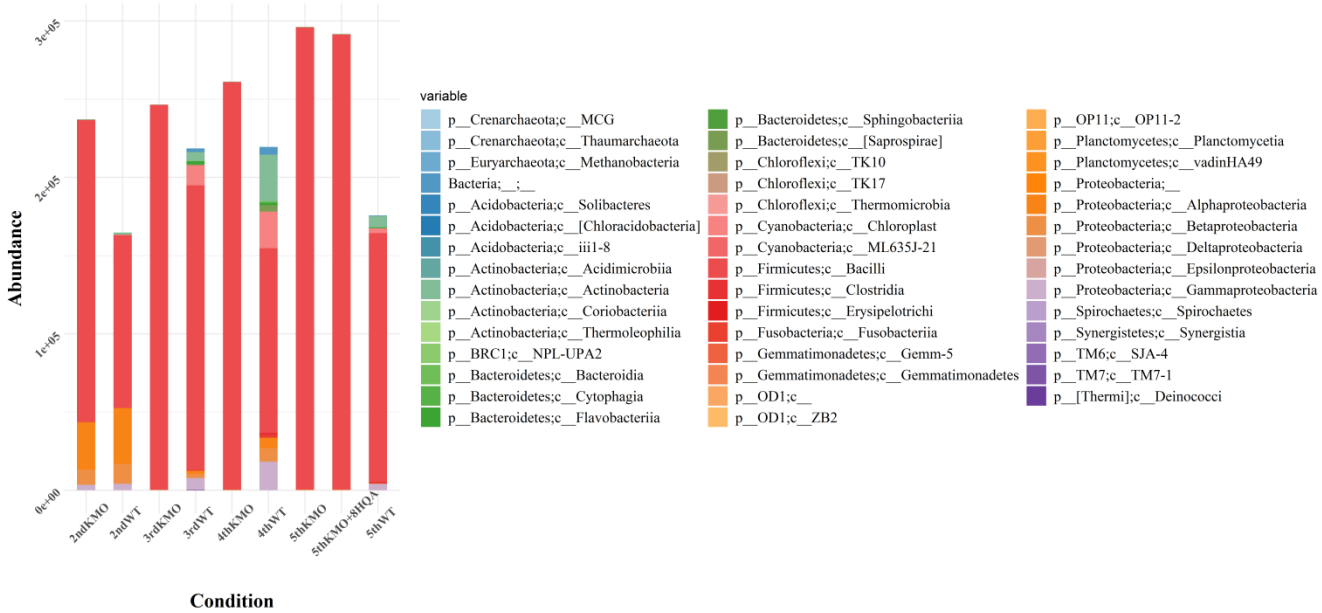

Figure 3: Comparison of relative microbial composition at the phylum level in KMO vs WT *S. littoralis* of the second, third fourth and fifth instar stages.

The KMO and WT larvae over different developmental stages showed difference in both the OTU numbers and bacterial composition among KMOs and WTs. The KMOs consistently showed a higher OTU levels than WTs. Even though the KMOs harbour more numbers of bacteria than the WTs, their bacterial composition is almost entirely dominated by bacilli of the phylum Firmicutes, especially by the family of Enterococcaceae. The frass itself had a 98% Enterococci species and the rest consisted of families of Alphaproteobacteria and Cyanobacteria (Fig. 4). Only in the second instar KMOs can one find more diversity than the KMO knockouts of other stages. This condition in particular harbors 13% Alphaproteobacteria and 4% and 1% of beta and gamma Proteobacteria respectively. On the other hand, the WTs of all the four stages show much more diversity in terms of their gut bacterial population. Although the Firmicutes are still the dominating phylum, Actinobacteria, Proteobacteria and Cyanobacteria show their presence in differing levels among the WTs of the four larval stages. Notable levels of them are- second instar shows a 19% population of alpha Proteobacteria, whereas fourth instar harbors 13% Actinobacteria and 8% Cyanobacteria. The sample- 5thKMO+8-HQA was expected to behave more like WT larvae, because the difference between KMO and WT lies in the presence of the compound in WTs. Quite contrary to our hypothesis, this particular sample showed bacteria diversity more like the KMOs, instead of WTs, suggesting that feeding 8-HQA externally does not have the same effect as producing 8-HQA (Fig. 3). This sample shows a 99% population of bacilli.

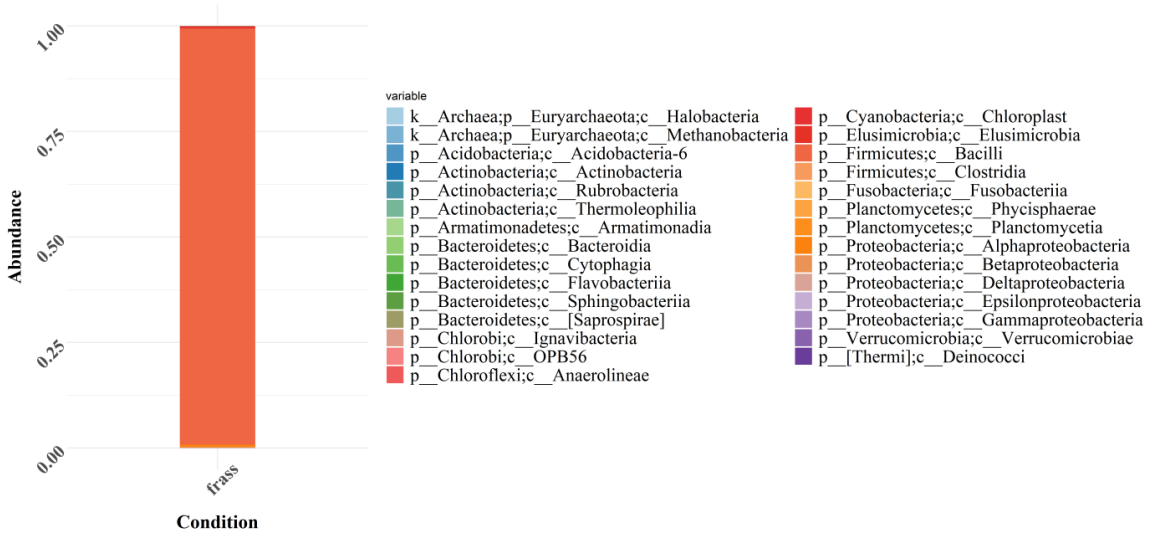

Figure 4: Relative microbial composition of the **frass** of *S. littoralis* of a different strain that was fed to the larvae from main study.

6.3.3Levels of the most abundant taxa along the developmental stages

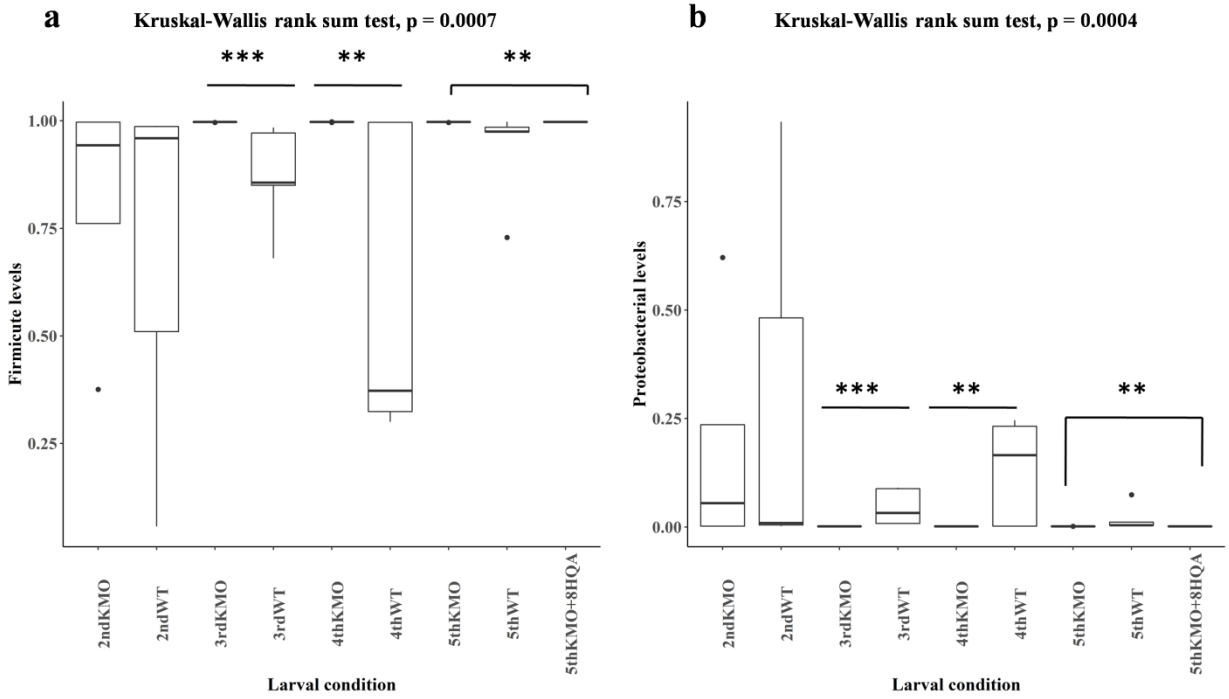

Figure 5: Levels of the most abundant taxa: a) Firmicutes and b) Proteobacteria in the KMO and WT lines of *S. littoralis*, across its developmental stages.

The most abundant phyla that showed variation between the KMOs and WTs are Firmicutes and Proteobacteria. While the levels of Firmicutes increased in absence of 8-HQA, the Proteobacterial levels decreased. Kruskal-wallis rank sum tests showed significant differences between KMOs and WTs along the developmental stages for these two phyla (p= 0.0007 and 0.0004 respectively). The differences are more apparent between the WT and KMO of third and fourth instars (Wilcoxon rank sum test- Firmicutes: 3rd and 4th, p = 0.008 and 0.03 respectively; Proteobacteria: 3rd and 4th, p = 0.007 and 0.03 respectively) (Fig. 5).

6.3.4. Alpha and Beta diversity analysis

Alpha diversity was measured on the basis of Shannon, Simpson and Observed species indices.

Kruskal wallis test shows significant difference in terms of Shannon and Simpson and observed indices in this study. Wilcoxon paired test was performed as post hoc test and it showed a significant difference between the two conditions of KMO and WT in the second and fourth instars prominently in case of Shannon and Simpson indices (species richness), and third and fifth instars when measured using observed species indices. Overall, the KMOs show a lower diversity as compared to WTs, although they lead in terms of 16S copy numbers (Fig. 6, Table 1).

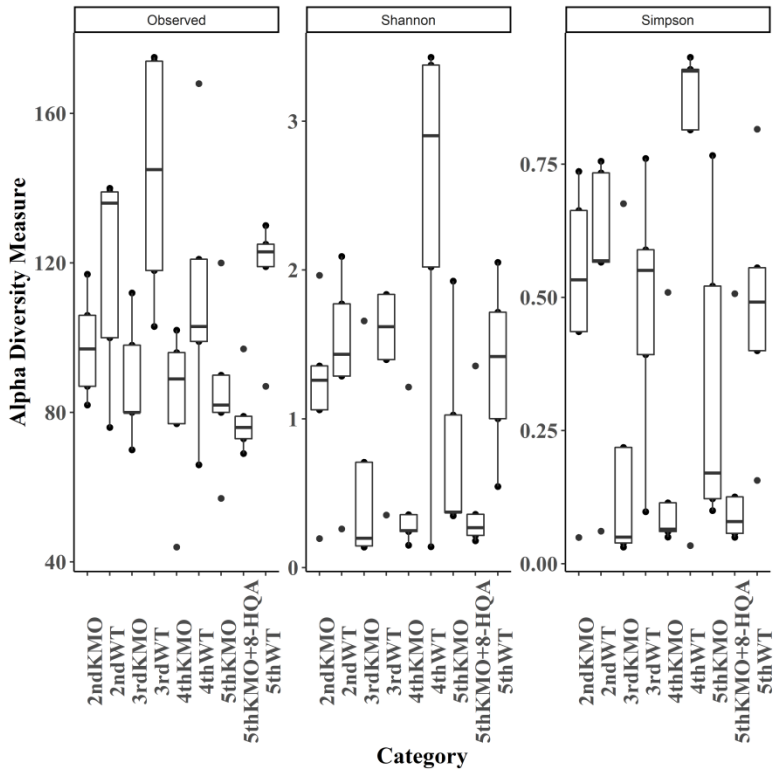

Figure 6: Box plots for comparison of alpha diversity of species in KMO and WT guts of *S. littoralis*, along their life cycle (second, third, fourth, fifth instars and fifth instar larvae fed with 8-HQA), based on the following indices: Observed richness of species, Shannon (species richness), and Simpson (species richness).

Alpha Diversity analysis for the Wt and KMO lines over the developmental stages						
Indices	**Statistical tests**					
Shannon	Kruskal-Wallis test: p = 0.03					
	Post hoc test: Wilcoxon rank sum tests for:					
	2^{nd} KMO vs WT	3^{rd} KMO vs WT	4^{th} KMO vs WT	5^{th} KMO vs WT	5^{th} KMO vs KMO+8HQA	5^{th} WT vs KMO+8HQA
	p = 0.001	p = 0.39	p = 0.007	p = 0.6	p = 0.1	p = 0.3
Simpson	Kruskal-Wallis test: p = 0.06					
	Post hoc test: Wilcoxon rank sum tests for:					
	2^{nd} KMO vs WT	3^{rd} KMO vs WT	4^{th} KMO vs WT	5^{th} KMO vs WT	5^{th} KMO vs KMO+8HQA	5^{th} WT vs KMO+8HQA
	p= 0.0015	p = 0.8	p= 0.0006	p = 0.4	p = 0.1	p = 0.05
Observed richness	Kruskal-Wallis test: p = 0.008					
	Post hoc test: Wilcoxon rank sum tests for:					
	2^{nd} KMO vs WT	3^{rd} KMO vs WT	4^{th} KMO vs WT	5^{th} KMO vs WT	5^{th} KMO vs KMO+8HQA	5^{th} WT vs KMO+8HQA
	p= 0.3095	p = 0.02	p = 0.15	p = 0.05	P = 0.4	P = 0.01

Table 1: Table showing the p-values of different alpha-diversity indices between every pair of conditions.

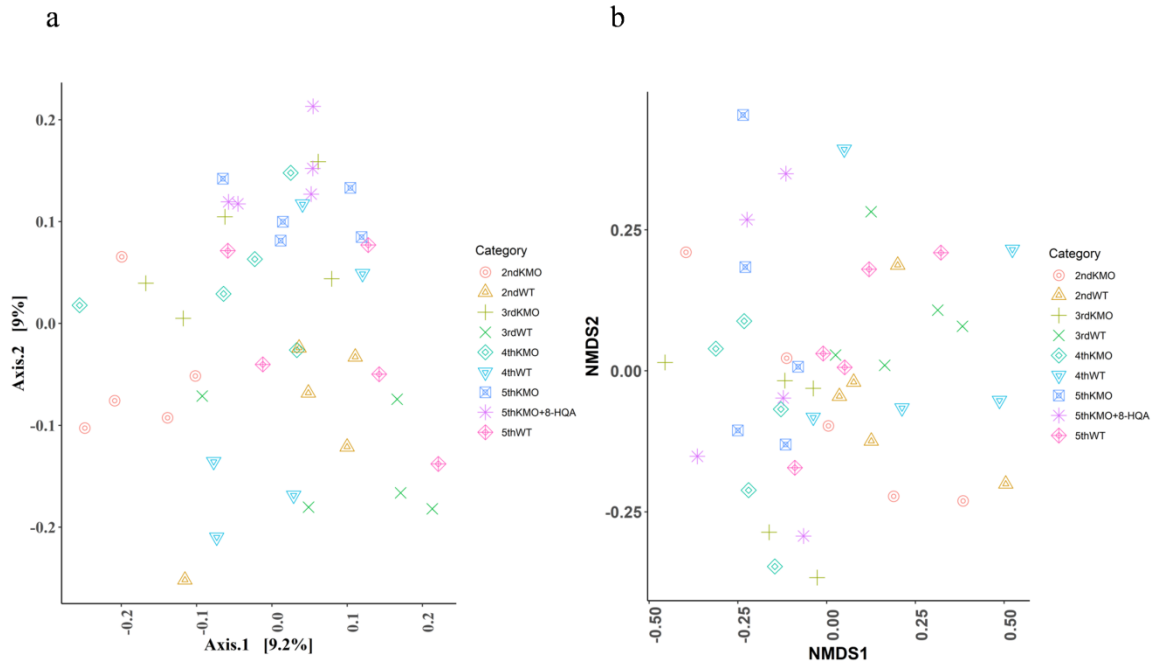

Figure 7: Beta diversity analysis between groups of species in KMO and WT *S. littoralis* guts by (a) **unweighted unifrac distances** on PCoA plot (b) **Bray-Curtis distances** on PCoA plot

Beta diversity analysis was performed using Bray Curtis distances and unifrac distances, using Principle Coordinate analysis (PCoA) and NMDS methods.

In case of the **frass-fed** samples, KMOs of all the four instars show better clustering when the beta diversity is analysed using weighted and unweighted unifrac distances on PCoA plots.

Analysis of variance using unweighted and weighted unifrac distance matrices (Adonis) gave significant results ($p < 0.05$) (Fig. 7, Table 2).

Considering all the KMOs and WTs together, they form two separate clusters on a PCoA plot according to their unweighted unifrac distances (Adonis, p=0.001) (Fig. 8)

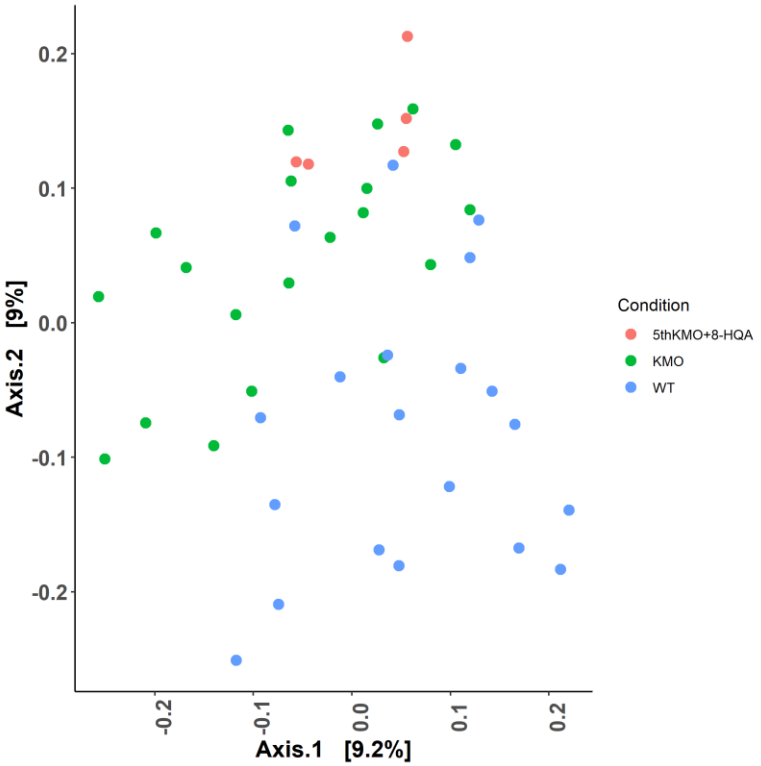

Figure 8: Beta diversity analysis between the KMO and WT lines of *S. littoralis* guts by (**unweighted unifrac distances** on PCoA plot (Adonis test p=0.001)

Beta diversity analysis for the Wt and KMO lines over the developmental stages	
Beta Diversity indices	**Adonis test**
Unifrac (unweighted)	P = 0.001
Unifrac (weighted)	P= 0.005
Bray- Curtis	P = 0.1

Table 2: Table showing p-values of Adonois tests of Beta diversity indices

6.4 Discussion

Previous studies have confirmed the presence of a high concentration of 8-Hydroxyquinoline-2-Carboxylic acid, in the guts of *Spodoptera littoralis* but none of the studies have looked into the influence of the compound on the microbiota of the host. To test how the gut microbial diversity is influenced by the presence of the compound 8-HQA, the gene for the enzyme responsible for producing the compound, Kynurenin monooxygenase (KMO) was knocked out in *S. littoralis* strains. The microbial diversity of these strains was compared to that of wild type strains (WT) that produced normal levels of the compound. The experimental larvae were fed with frass from another strain, mixed with their artificial diet. The bacterial composition of the frass was also elucidated to determine the bacterial colonization in the frass-fed larvae.

In terms of number of reads assigned in all OTUs per sample, the KMOs lead. In evenness of the distribution of the species in each OTU the WTs lead. Richness here refers to the number of different types of species represented in each condition, whereas the evenness describes how uniformly bacterial strains are distributed among each of these species in each community. In our experiment, the bacterial distribution among the detected species is non-uniform, of which the KMOs perform worse. Although the KMOs harbor higher bacterial numbers than their corresponding WTs, they are beaten by the WTs in terms of species richness and evenness, as seen from the alpha diversity analysis. Most of the reads assigned to KMOs hit the OTU of Enterococcaceae, which leaves negligible room for other phyla. The second instar KMOs although seem to house some levels of Alphaproteobacteria and clostridia, its richness and evenness still lesser than their corresponding WTs. Thus, in spite of seemingly housing a greater number of bacteria in the gut, the KMO knockouts of *Spodoptera littoralis* show a poorer diversity than the corresponding wild types.

Beta diversity shifts are seen among individual conditions of KMO and WT over the developmental stages, but the prominent overlap could be explained by common bacterial taxa, like Enterococci among all the conditions. While analyzing the beta diversity of all the KMOs and WTs taken together, we notice a clear separation between them on a PCoA plot.

KMOs are supposed to have higher free iron levels since they no longer produce the iron chelator 8-HQA. Although it is necessary to study the binding affinity of ferrous and ferric ions with this compound as a future prospect of this project, but based on older studies that suggest the siderophoric roles of carboxylic quinolinic acids, we can propose this. We expect

certain metabolic changes to occur when iron is chelated vs when it is available in abundance. Iron dependent enzymes are a key to energy acquisition pathways. Change in iron levels could change the redox balance. Acetate levels were not checked in our larvae but in a similar study, acetate that resulted from a fermentative pathway, was reduced, and at the same time there was a reduction in fermentative bacteria [264].

The increase in bacterial diversity in presence of the iron-chelator, 8-HQA could mean that these taxa are exploiting other metal ion as cofactors, as was also the case in a study performed on human gut microbiome when iron availability was reduced [264].

A similar study on broiler chicken with and without iron fortified diet showed similar results. The bacterial diversity in the guts of chicken underfed on iron showed a higher bacterial diversity than the ones with excess iron in their diet [265]. So was the case in rhinos prone to iron overload disease [266]. In such studies the researchers hypothesize that a control over the iron levels could keep levels of pathogens in check [264]. Although we did not find any visible health deficits in the KMO knockouts, but earlier studies have shown the pathogenicity of the larva's own Enterococcal species- *E. fecalis* and *fecium* [25]. It has been established that *Enterococcus mundtii* is taking care of these pathogens by means of its bacteriocin, called mundticiin [36], we still do not know if 8-HQA is playing an additional role in keeping these pathogens under control. The downside of short read libraries generated by illumina is that, we do not get a species level resolution. Hence, we are unaware of the exact species of the Enterococcal population that are present in the KMO knockouts vs the WTs.

To conclude, 8-Hydroxyquinoline-2-carboxylic acid is indeed affecting the bacterial landscape of the guts of *S. littoralis* larvae. To what extent it might be beneficial to the host is still a matter of further research.

7 General Discussion

7.1 Reporter technology and advancements

Reporter technology has revolutionized the basic understanding of complex physiological phenomena like cell communication, cellular development, oncogenesis, regulation of growth and many more, those phenomena that can be related to changes in gene expression. To study the activation and inhibition of different pathways and their relationship with gene expression, a certain class of genes called reporter genes are tagged in the downstream of the promoter of the gene of choice to monitor the differential expression of the gene. This fusion is done by transforming an expression vector encoding the reporter gene to the cells. Quantifying either the mRNA or the protein product of the reporter gene, conclusions can be drawn about the expression of the gene of choice. The qualities of the best reporter genes should be: high sensitivity, high reproducibility and easily detectable. This being said, reporters have significant applications in biotechnology. Some notable reporter genes are as follows [267]:

Chloramphenicol Acetyl Transferase (CAT): It is a bacterial enzymatic gene that has been used to monitor the transgene expression in some disease models like heart disease, hepatitis B, drug resistance in bacteria [268]

Alkaline phosphatase: It is a stably expressed protein in mammals and bacteria. A spectrophotometric assay owing to the hydrolysis of AP substrates (PNPP, FADP) leads to the change in absorbance; is the method this reporter gene works [269].

Beta galactosidase: It is again a bacterial enzyme which is used as a reporter gene because of its calorimetric properties attributed by a hydrolysable substrate (ONPG) [270].

Green Fluorescent Protein (Green Fluorescent Protein (GFP): Applications, Structure, and Related Photophysical Behavior)

Bioluminescence is a phenomenon by which a living entity emits light as a result of a chemical reaction. The substrate luciferin is oxidized by the enzyme luciferase. The emitted light has functions of defense, offence or communication by the organism. Such a system is employed by species of jellyfish, shrimp, clam, insects, fish, squid, sea pansies and sea cacti. Green Fluorescent Protein was cloned for the first time from the jellyfish *Aequorea victoria*.

The bioluminescence property of Aequorea is brought about by its two proteins: Aequorin and green fluorescent protein. Aequorin is the luciferase that contains the luciferin coelenterazine. In Aequoria, the aequorin complex, on oxidizing its luciferin, undergoes a radiationless energy transfer to the GFP, which emits fluorescence. GFP is superior to the other markers described because it can analyze living cells, and fixation or sample preparation steps are not necessary. Screening of live cells is important in embryonic stem cell technology and selection of transgenic animals [271].

The last decades have seen enormous number of publications in Biology that have been successful using the GFP. It is particularly useful due to a stable structure and because its chromophore is formed in an autocatalytic cyclization that does not require a cofactor. GFP was first used to monitor protein localization in cells under fluorescent microscopy. The fact that GFP can be detected by simple visualization makes high throughput screening with GFP easier, ranging from successful cloning screens to drug screens. What's more, GFP can also help in making DNA sequences visible. This discovery has led to the study of mitotic apparatus, or the partitioning of the sister chromatids in the daughter cells. Combining with Fluorescent Activated Cell Sorting (FACS), successful clones of genes of interest can be screened from the rest [271].

Bacterial bioreporters, or bacterial cells genetically modified with one of the reporter proteins found its usage in fields of bioengineering and biotechnology. The metabolism of bacteria is used as a tool to monitor, for example, dose dependent toxicity or mutagenicity of certain potential contaminants of the environment. A particular signaling pathway is monitored, not by its inherent activity, but by an artificial output. This is achieved by visualizing the expression of reporter genes as described above [267].

7.2 Reporter gene and bio reporter technology in our research

The first part of the research has been made possible by the fusion of the reporter gene Green Fluorescent Protein to the bacteria of interest, *Enterococcus mundtii*. It was motivating to learn that this bacterium forms one of the dominating ones in the gut of *Spodoptera littoralis* larva. It exercises a colonization resistance against the early colonizers: *Enterococcus fecalis*and *Enterococcus fecium*. As it turns out, *Enterococcus mundtii* produces a class IIa bacteriocin called mundticin and it is lethal against the potential pathogens which are nothing

but the early colonizers. On the other hand, the gut environmental conditions are not the friendliest. *S. littoralis* produces a high concentration (0.5 – 5 mM) of 8-HQA (8-Hydroxyquinoline-2-Carboxylic acid). This compound is an iron chelator which reduces the availability of free iron in the gut of the host. Also, there is a marked pH gradient ranging from alkaline in the foregut to almost neutral in the hind gut. These factors make this system an ideal model for the study of how *E. mundtii* still manages to dominate and persist despite these adverse living conditions.

The next challenging aspect was to find out an efficient method which would help us selectively study one particular bacterial species out of several others in the gut. This is where we employed the usage of reporter gene GFP and converted cells of *E. mundtii* to bioreporters. A GFP containing expression vector pTRKH3-ermGFP was transformed to a population of *E. mundtii* by electroporation [176]. Successful transformants were maintained to be fed to newly hatched larvae along with their artificial diet. The larvae, after allowing growth up to fifth instar, were dissected and the guts were homogenized to retrieve the fluorescent reporter *E. mundtii* bacteria. At this point, a flow cytometer was employed. The fluorescent properties of GFP enabled a blue laser to excite it at a wavelength of 488 nm to enable the GFP-containing cells to be specifically sorted. We use these cells to study how exactly different are their living conditions inside the gut, as compared to outside (*in vitro*). RNAseq was the method of choice to study the differential gene expression of *E. mundtii* living in the gut and grown in culture conditions. In order to prevent any extraneous alteration in gene expression during the sample preparation steps, reagents like RNAlater and RNAprotect are used. Such reagents have twofold advantages: they not only percolate cells and protect the integrity of RNA, which otherwise have a very short half-life, they also act as a fixating agent and minimize changes in gene expression in these cells.

7.3 Flow cytometry making its way in host-symbiont research

Flow cytometer is definitely one of the state-of-the-art methods employed in the study of microbial ecology. It separates cell based on their intrinsic characteristics, at the same time, including electronics, fluidics and optics. The sheath fluid escorts each cell as they go through the lasers that monitor the size and the granularity of each cell. Every cell is unique based on these properties which are measured by the detector as the FSC (Forward Scatter) and SSC (Side Scatter) respectively.

In our case, the GFP containing *E. mundtii* cells are identified at two levels, their size and granularity are different from the accompanying eukaryotic cells, and their GFP differentiates them from other bacterial cells in the sample. The latter is achieved by hydrodynamic focusing and presenting each cell to a monochromatic beam interrogating them one at a time. Depending on the excitation wavelength of the fluorophore the laser wavelength is chosen. The scatters are then measured by the detector. The detector is a set of photomultiplier tubes with specific filters to select for the excitation wavelength of the fluorophore. Larger the cell, higher is the FSC, and the SSC measures how complex a cell is, along with its fluorescence. Based on these values, a charge is assigned to each cell. The charged plates deflect the cells towards the sorting tubes, to go ahead with further downstream processes.

Apart from combining reporter gene technology with flow cytometry, which can at times be time consuming, this technology can also provide first line of analysis of microbial diversity. This is achieved by phenotypic fingerprints of the microbial population. The capability of a flow cytometer of differentiating the cell to cell heterogeneity can give one an idea about the diversity of different bacteria present in a population. Such a preliminary diversity profile strongly correlated with the same results obtained from16S amplicon sequencing. Such an approach found efficient use in determining the microbial population in diverse soil samples, sediments from streams and lakes, and sludge from water filters of drinking water treatment plants [272,273].

Flow cytometry found its way in several scientific applications. In the age of growing number of antibiotic resistant strains, it is necessary to look into the susceptibility of strains against selected antibiotics. In order to study viability of cells, FCM can be used to measure criteria like changes membrane potential, change of permeability to dyes, or the presence of metabolic activity reported via a fluorescent gene [274]. FCM was used to assess the susceptibility of *E. coli* cells to mecillinam and ampicillin antibiotics. Forward and side scatters alone indicated that the DNA and protein synthesis continued but the cells did not divide. This was consistent with the action of penicillin. Likewise, susceptibility of bacterial cells to antibiotics with other modes of action such as targeting the DNA gyrase, or irreversibly binding to ribosome, FCM finds its application in accessing cell viability [275]. Not just drugs, the susceptibility of bacteria to heat, pH, oxidative stress and salt concentration was also important to investigate when *Listeria monocytogens* was found adaptive to such stresses and causing food borne diseases. A GFP tagged Sigma B gene constructs of this

bacteria studied in a FCM confirmed the role of the general stress factor Sigma B as a stress adaptive protein in such pathogens [276].

One of the revolutionizing applications of FCM is that it allows for analysing single-cells in a population. Heterogeneous populations of cells, for example, the ones making up a biofilm have different roles to play along the thickness of the biofilm [277]. Endospores of *B. cereus* at different levels of germination (*Bacillus cereus* responses to acid stress), or studies investigating virulence factors expressed by quorum sensing pathogens, require analysis at single-cell levels [278]. A lot of studies on host-symbiont interactions employ FCM. Apart from our work that studies the behaviour of a specific symbiont, *E. mundtii* inside the guts of *S. littoralis* larvae, employing the single-cell sorting capability of a flow cytometer, several other host-bacterial relationships have been investigated as well. Analysis of adherence of uropathogenic *E. coli* to vaginal cell lines have been successful using GFP constructs of the bacteria isolated from women with urinary tract infections. Counting PI stained vaginal tract cells and GFP tagged *E. coli* cells, the percentage of adherence was assessed [279]. Studying the genes that are solely expressed during a bacteria-host interaction using fluorescent constructs is an efficient way to learn adaptive and pathogenic mechanisms employed by the bacteria. In a nutshell, the technology of FCM allows for detection and counting of cells to analysing metabolic and genetic changes of the same in different conditions. It is not only results in reproducible analyses, but also is highly efficient in terms of quantifying cells present in low abundance in a population.

7.4 *Enterococcus mundtii* in the fore and hind guts of *Spodoptera littoralis* larvae

Like already explained, the precise aim of this work was to develop a method that specifically targets the symbiont of choice in a habitat. Our question was what makes *Enterococcus mundtii* dominate and survive the stress in the gut of *Spodoptera littoralis larva*. Keeping this question in mind, a new workflow was developed. The symbiont of choice was isolated from the gut and maintained, namely *Enterococcus mundtii* KD251 [176]. This strain was transformed with plasmid pTRKH3-ermGFP to make it fluorescent. Next, this reporter strain was allowed to be incorporated as a gut inhabitant and change its gene expression accordingly. The changed gene expression profile was later studied after the fluorescent reporter *E. mundtii* were sorted out using a flow cytometer, followed by their transcriptomic analysis As for the biological replicates, their clustering together in PCA plot and

dendrogram tell us that *E. mundtii* indeed behave differently in the gut of the larvae, as compared to the one grown *in vitro*.

The results we get are quite consistent with the studies already performed with other bacteria in similar stress-conditions. First and foremost, the symbiont establishes a firm adherence to the host epithelial cells. Both in the fore and the hindgut, the symbiont upregulates several attachment proteins such as LPXTG, WxL domains, chitin-binding proteins and LysM motifs that help in biofilm formation. An alkaline condition in the foregut causes similar changes in gene regulation like previous studies performed in *E. faecalis*[280]. Alkaline shock protein shows a steep upregulation when confronted with high pH in the foregut. There is a down regulation in F-type H^+ transporters (ntp), V-type Na^+H^+ transporters (atp) and the nhaC transporter. There is a decrease in proton motive force at high pH, and hence the H^+ ATPases do not operate [280]. It has already been established that 8-HQA, a potent iron chelator is present in high concentrations in the regurgitate of the insect. In order to survive in an iron-deficient environment, *E. mundtii* upregulates its FetC iron complex permease protein and FUR family transcriptional regulators, both of which aid in iron homeostasis. This bacterium has a multitude of general and universal stress proteins helping them deal with changes such as temperature, oxidative, nutrient starvation and toxic agents. Quorum sensing, dictated by two-component systems comes to rescue when bacteria is undergoing adaptive changes and division of labor in a new environment, as does tolerance towards oxidative stress brought about by the host's respiration. Upregulated phosphotransferase systems (PTS) pose helpful for the bacterium such that they can survive in diverse sugar sources. The transporters for starch and sucrose are especially upregulated since the larval host is on a plant-based diet, the symbionts are required to break down the complex carbohydrates. These symbionts, in the hindgut of the larvae, have their genes for lysine synthesis upregulated. *E. mundtii* could possibly be providing the host with this essential amino acid, although confirmatory studies are yet to be done.

The drawbacks of this method also require discussion. This work also deals with RNA Seq from tiny quantities of RNA, as only 250000 fluorescent reporter *E. mundtii* were sorted by a flow cytometer. The downstream processing of these cells starting from pelleting them down till extraction of RNA and amplification resulted in a variable quantity of RNA each time. This variation can be attributed to handling of the sorted cells for pelleting, RNA extraction and amplification. mRNA amplification on the other hand depends on a proper incorporation of poly (A) chain behind every RNA, which then is targeted for the rest of the amplification

process. This crucial step relies on relaxing the secondary structures of RNA in a denaturation step. If these steps are not uniform among replicates, introduction of variability is inevitable.

Next, the downstream processes following generation of reads: quality control and trimming produce varying number of good quality reads for each replicate. Also, the mapping percentages of these reads to the *E. mundtii* genome vary in a range of 17-55% among the 9 samples.

Having said this, now if we look at the expression of individual genes among the three biological replicates, although we find a clear trend of up and down regulation, but in many cases, the trend is not significant. Despite the clustering of the individual replicates, there exists a variation in individual gene expression among the replicates. This can be attributed to variable gene expression among single cells and low quantity of starting material, time of sampling. Also, the process of gene expression is stochastic, that is to say, there exists randomness when it comes to gene expression, leading to a variable amount of a particular transcript among cells in similar conditions. Stress response genes, in particular, tended to be noisy. As soon as the symbionts encounter stress, they start expressing stress related pathways accordingly. The time for taken for switching to an adaptive mode in a new environment may vary among cells, leading to fluctuations in levels of transcripts [281,282]

Another drawback of our approach was that we have not been able to remove the rRNAs. Even after amplifying the extracted RNA, the concentration was not high enough to go through the risk of the removal process. There existed a possibility of further RNA degradation. The percentages of rRNA reads for each replicate were about 50%. These reads did not pose any problems in our work, but can be taken care of in either of the two ways: Including another step of RNA amplification to increase the RNA concentration to a point, where it is safe to subject it to rRNA removal; or generate a higher number of reads such that the mRNAs are not neglected. Both these alternatives are expensive.

7.5 Host genetic control of microbiota

It is a well-known fact how the microbiota in an organism faces controls at so many different levels. This control commences right from their very first attempt to colonize a gut, either by vertical or horizontal transmission. The microbiome an organism acquires from their mothers on their way out of either the womb or the eggs are vertically transmitted. Such microbes are mostly obligate symbionts to the offspring. In the course of its life, the organism, depending on its diet, environment and genetic makeup, horizontally transmitted microbes add up to its already existing microbiota. Host genetics have an important role to play in determining the microbial population as well. Experiments with different combinations of in-bred mice show significant differences in microbial population [283]. Mice with mutated inflammatory signals and metabolic traits also have altered gut microbiota [284]. Deficiency of a certain *Irk4* gene makes mice more susceptible to bacterial infections [285], whereas a mouse with *Bpi* gene makes it potent towards killing gram negative bacteria [286]. Monozygotic twins have an overall similar microbiota than dizygotic ones [287]. The core microbial composition in calves correlated to the breed of the parents. Also, single nucleotide polymorphisms in mucin degrading gene in the calves correlate with breed composition, and allows for the colonization of more of mucin degrading gut bacteria [288]. In *Drosophila melanogaster*, the microbiome varies with the host nutritional phenotype and this was quantified on the basis of the host nutritional indices like protein, TAG, glycogen and glucose contents [289].

In case of our model system *Spodoptera littoralis* the gut microbiome has been elucidated in addition to investigating what the dominating bacteria *Enterococcus mundtii* is managing to live and dominate the various stress conditions in the gut of the caterpillar. Also has been established how the compound 8-hydroxyquinoline-2-carboxylic acid, a potent iron chelator is produced by the larvae in their own regurgitate. We hypothesized that this compound is controlling the gut microbial population in these insects by keeping free iron from their reach. If this is the case, then we should see a different microbial landscape in larvae devoid of this compound. Keeping this in mind, 3-kynureninmonooxygenase (KMO) was knocked out in these larvae. This gene is involved in producing 8-HQA, in the pathway starting from tryptophan.

7.6 8-HQA with and without

To test whether the compound 8-hydroxyquioline-2-carboxylic acid is a player in determining the bacterial population in the guts of *S. littoralis* larvae, it was necessary to get rid of the compound in the larvae. CRISPR-Cas9 method was employed to knock out the gene for 3-hydroxykynurenin, the major enzyme in the biochemical pathway producing the compound (Fig. 1 in Article III). Thus we had two groups of larval lines: one devoid of 8-HQA, called the KMOs and the wild types (WT). After knocking the larvae out of the 8-HQA producing gene, GC-MS analysis confirmed the absence of the compound in the regurgitate of the knockouts. Subsequently, these lines were analyzed for their bacterial composition. The workflow that was followed was this: Both lines of larvae were fed with frass from a different strain of *S. littoralis*, with known bacterial composition (99% Firmicutes of class Bacilli, and the rest were Cyanobacteria and Alphaproteobacteria) mixed to their artificial diet. 5 KMO and WT larvae each of second, third, fourth and fifth instars were sampled, sterilized and DNA was extracted from individual larvae. The bacterial DNA in the pool was identified amplifying the V4 region of the 16S rRNA gene. The PCR was conducted using forward F515 primer (5'-TATGGTAATTGTGTGYCAGCMGCCGCGGTAA-3') and reverse R806 primer (5'-AGTCAGCCAGCCGGACTACNVGGGTWTCTAAT-3'). Eventually, this region went through MiSeq sequencing to discern the bacterial population and abundances in these two lines of larvae, over the developmental stages.

The KMOs seem to house more bacteria in terms of numbers than the WTs. On the other hand, the WTs lead in terms of richness and evenness of species. Although, a majority of these bacteria belong to Enterococcaceae family in both the WTs and KMOs, the WTs still harbor a substantial numbers of Alphaproteobacteria, Actinomycetes and some Cyanobacteria, as was the composition in the frass that was fed.

Researchers have long been trying to establish a connection between iron content and microbial composition. Since 8-HQA is an iron chelator, our research somewhat points at a similar direction. In literature, there are varying accounts on this topic. Some stories found a positive correlation between iron concentration and diversity of certain bacterial taxa, while some found the same to be negative. Studies have hinted towards Lactobacilli (also Firmicutes) not requiring iron as an essential element. Mice were tested with this question in mind. One group received iron depleted diet for some days, followed by repletion, whereas the other group continually received an iron depleted diet. Lactobacilli were found in high

levels in the colonic microbiota of the mice who received the iron depleted diet [290]. Another study has shown the decrease in levels of lactobacilli with response to higher iron content. Since Lactobacilli are beneficial bacteria, their depletion could result in an onset of pathogenic species which, in a long run, could be detrimental to the host. The same study has shown abundance of Enterobacteria (Gammaproteobacteria) in iron depleted conditions, which is in line with our findings [291]. In another study with broiler chicken, iron fortified diet brought about a decrease in bacterial diversity. Rhinos showed the same when they are prone to iron overload disease [265,266]. Thus we see that organisms behave differently in iron overload conditions. In general, organisms modulate their iron concentrations to reduce the load of pathogens in their microbiota [264]. Although no fitness deficits were found in the KMO knockout larvae of *Spodoptera littoralis*, but previous studies did establish fellow Firmicutes *Enterococcus fecalis* and *fecium* pose pathogenic towards the host if not kept under control by the mundticiin produced by *Enterococcus mundtii*. There was a steep rise in Firmicute levels in the KMOs as compared to WTs. It is quite possible that this is how 8-HQA maintains the levels of iron, and consequently, keeps the bacterial population (here, Firmicutes) in check in Lepidoptera.

There are downsides to elucidating bacterial diversity using illumina short-read sequencing. Although this technology is cheaper, faster and has a higher thoughput than the last generation sequencing technologies, the upsides have been traded with high resolution species identification. We do know the status of bacterial families in our experimental conditions, but hardly are we aware of the species that show varying diversity in the same conditions. Thus, a substantial part of the story is still incomplete without getting the full picture of diversity up to the strain level [292].

Nevertheless, high-throughput illumina sequencing technologies formed the basis of this thesis and has endowed upon us a wealth of information about the gut microbiome of *Spodoptera littoralis* larvae.

8 Summary

The complex interaction amongst a higher organism and its resident gut flora is a subject of immense interest in the field of symbiosis. Insects harbor a population of gut bacteria that play roles in their growth, development and immunity. There exists a variation in the microbial population with the development of the insect.

The gut microbiota of *Spodoptera littoralis*, a Lepidopteran pest, varies spatially and temporally. The core community consists of *Enterococci, Lactobacilli* and *Clostridia*. The selection of one bacterial species over the other is quite evident throughout the lifecycle, so is the differing bacterial population and abundance among the fore, mid and hind gut of the larva. By the time the larva reaches fifth instar, *Enterococcus mundtii* persist and dominate. The gut environment dictates the persistence of its residents. There is a pH gradient from alkaline to neutral along fore to hind gut respectively, and a depleted iron condition as posed by the chelator 8-HQA (acid) produced by the insects.

We ask the following: **How does the *E. mundtii* dominate by surviving the gut stress? What kind of interaction goes on between them and their host? Finally, how does 8-HQA define the microbial landscape of the *S. littoralis* gut?**

A GFP-tagged reporter *E. mundtii* has been constructed to answer the **first two questions**. They are fed to the insects at early instars, and sorted from the gut spatially and temporally using flow-cytometry. These reporter bacteria which had integrated into the gut community of the larva must have changed their gene expression profile according to the new environment. A transcriptomic analysis of the retrieved bacteria from the host gut should answer our questions. The fluorescent reporter confirmed the persistence of *E. mundtii* in the gut. Also, RNA-sequencing of the sorted bacteria has informed us about various strategies of the symbiont's survival. The results we get are quite consistent with the studies already performed with other bacteria in similar stress-conditions. First and foremost, the symbiont establishes a firm adherence to the host epithelial cells. Both in the fore and the hindgut, the symbiont upregulates several attachment proteins such as LPXTG, WxL domains, chitin-binding proteins and LysM motifs that help in biofilm formation. An alkaline condition in the foregut causes similar changes in gene regulation like previous studies performed in *E. faecalis*. Alkaline shock protein shows a steep upregulation when confronted with high pH in the foregut. There is a down regulation in F-type H^+ transporters (ntp), V-type Na^+H^+ transporters (atp) and the nhaC transporter. There is a decrease in proton motive force at high

pH, and hence the H^+ ATPases do not operate [280]. It has already been established that 8-HQA, a potent iron chelator is present in high concentrations in the regurgitate of the insect. In order to survive in an iron-deficient environment, *E. mundtii* upregulates its FetC iron complex permease protein and FUR family transcriptional regulators, both of which aid in iron homeostasis. This bacterium has a multitude of general and universal stress proteins helping them deal with changes such as temperature, oxidative, nutrient starvation and toxic agents. Quorum sensing, dictated by two-component systems comes to rescue when bacteria is undergoing adaptive changes and division of labor in a new environment, as does tolerance towards oxidative stress brought about by the host's respiration. Upregulated phosphotransferase systems (PTS) pose helpful for the bacterium such that they can survive in diverse sugar sources. The transporters for starch and sucrose are especially upregulated since the larval host is on a plant-based diet, the symbionts are required to break down the complex carbohydrates. These symbionts, in the hindgut of the larvae, have their genes for lysine synthesis upregulated. *E. mundtii* could possibly be providing the host with this essential amino acid, although confirmatory studies are yet to be done.

To address the **third question**, a knockout line of larvae for 8-HQA (KMO) production was constructed using the CRISPR-Cas9 method, whose gut microbiota was elucidated and their differences with the wild type (WT) strains were analyzed. The results give us an idea of the importance of host-genetics in dictating the microbial composition and numbers in the guts of hosts. The KMOs seem to house more bacteria in terms of numbers than the WTs. On the other hand, the WTs lead in terms of richness and evenness of species. Although, a majority of these bacteria belong to Enterococcaceae family in both the WTs and KMOs, the WTs still harbor a substantial numbers of Alphaproteobacteria, Actinomycetes and some Cyanobacteria, as was the composition in the frass that was fed.

This thesis gives one a closer look into the gut microbiome of *Spodoptera littoralis* larvae. It focusus specifically on *Enterococcus mundtii* because it speaks for several other bacteria in microbiomes in general and throws light on the mechanisms they employ to adapt to the differing living conditions in host guts.

9 Zusammenfassung

Die komplexe Wechselwirkung zwischen einem höheren Organismus und seiner angesiedelten Darmflora ist in dem Fachgebiet der Symbiose von immensem Interesse. Insekten beherbergen eine Population von Darmbakterien, die für ihr Wachstum, ihre Entwicklung und ihre Immunität eine besondere Rolle spielen. Die Zusammensetzung dieser mikrobiellen Population variiert im Verlauf der Entwicklung des Insekts.

Die Darmmikrobiota von *Spodoptera littoralis*, einem weltweit verbreiteten Schädling, variiert räumlich und zeitlich. Die Kerngemeinschaft setzt sich aus Enterokokken, Laktobazillen und Clostridien zusammen. Die Selektion einer Bakterienart gegenüber der anderen ist während des gesamten Lebenszyklus eindeutig erkennbar, sowie das Auftreten und die Häufigkeit einzelner Populationen im Vorder-, Mittel- und Hinterdarm der Larve. Bei Erreichen des fünften Wachstumsstadiums ist *Enterococcus mundtii* die dominierende Spezies in Abhängigkeit der vorherrschenden Bedingungen innerhalb des Darmbereiches. Diese beinhalten unter anderem einen von basisch bis neutral reichenden pH Gradienten, der sich vom Vorder- bis hin zum Hinterdarm erstreckt. Darüber hinaus herrscht eine generelle Eisenarmut bedingt durch den vom Insekt produzierten Chelator 8-Hydroquinolin-2-Carboxylsäure (8-HQA).

Hieraus ergeben sich die folgenden Fragestellungen: Wie passt *E. mundtii* sich, besonders im Vergleich zu anderen Mikroben, an die gegebenen Stressbedingungen an? Wie interagieren *E. mundtii* und *S. littoralis*? Wie beeinflusst 8-HQA die mikrobielle Zusammensetzung innerhalb des Darms?

Zur Beantwortung der ersten beiden Fragen wurde ein GFP-markierter *E. mundtii* Stamm generiert, welcher in den frühen Larvenstadien über das Futter aufgenommen und weiterfolgend mittels Durchflusszytometrie räumlich und zeitlich getrennt und sortiert wurde. Durch die Integration des Reporterstamms in den Darm nahmen wir an, dass die Genexpressionsprofile der neuen Umgebung entsprechend Veränderungen beinhalten. Die Transkriptom analyse der aus dem Wirtsdarm gewonnenen Bakterien bestätigte die Dominanz des fluoreszenz-markierten *E. mundtii*. Darüber hinaus wurden mittels RNA-Sequenzierung Überlebensstrategien des aus dem Darm isolierten Symbionten erkennbar,

welche mit bereits veröffentlichten anderen Bakterien unter vergleichbaren Stressbedingungen übereinstimmen.

Eine der ersten Interaktionen zwischen *E. mundtii* und Wirt ist die Anhaftung des Symbionts an die Epithelzellen des Voder- bzw Hinterdarms des Insekts mittels Hochregulation von Bindungsproteinen (LPXTG, WxL Domänen, Chitin-bindende Proteine and LysM-Motive), welche die Bildung eines Biofilms fördern. Insbesondere die basischen Bedingungen im Vorderdarm führen zu vergleichbaren Veränderungen der Genregulation wie bereits in Studien mit *E. faecalis* beobachtet. Durch den hohen pH Wert innerhalb des Vorderdarms wird das alkalische Schockprotein stark hochreguliert einhergehend mit runterregulierten F-Typ H+ Transportern (ntp), V-Typ Na+H+ Transportern (atp) und nhaC Transportern. Darüber hinaus nimmt die Protonen bewegende Kraft („PMF“) bei steigenden pH Werten ab und stoppt die H+ ATPase-Aktivität 270.

Frühere Studien zeigen, dass Insekten einen hohen Anteil des Eisenchelators 8-HQA in ihrem Regurgitat aufweisen, welches einen drastischen Eisenmangel innerhalb des Darms verursacht. Um in diesen widrigen Bedingungen zu überleben, reguliert *E. mundtii* FetC-Eisenkomplex-Permease Proteine sowie FUR Transkriptionsregulatoren hoch, um die Eisenhomöostase zu erleichtern. *Enterococcus mundtii* verfügt zusätzlich über eine Vielzahl universeller Stressproteine, welche eine optimale Anpassung an Temperaturstress, oxidativen Stress, Nährstoffmangel sowie Giftstoffe erlauben. In Anpassung an eine neue Umgebung wird sog. Quorum Sensing aktiviert. Dieses Phänomen wird wiederum durch ein Zweikomponentensystem reguliert, welches Arbeitsteilung der Bakterien sowie Toleranz gegenüber oxidativem Stress, verursacht durch die Atmung des Wirts, ermöglicht. Eine optimale Nährstoffquelle bieten diverse vorhandene Zucker, welche durch hochregulierte Phosphotransferase Systeme (PTS) aufgenommen werden können. Aufgrund der Tatsache, dass es sich bei dem Wirt um ein herbivores Insekt handelt, nimmt dieses hauptsächlich Kohlenhydratkomplexe auf, die durch den Symbiont zersetzt werden müssen. Dies führt zu einer Aktivierung von Stärke- und Sukrosetransportern für einen effizienten Abbau. Aus dem Hinterdarmbereich entnommene Symbionten wiesen zudem eine auffällige Hochregulation von Lysin Synthese Genen auf, welche möglicherweise essentielle Aminosäuren für den Wirt bereitstellen könnten. Diese Hypothese muss jedoch durch weitere experimentelle Studien bestätigt werden.

Zur Beantwortung der dritten Frage wurden mittels der Genschere CRISPRS-Cas9 *S. littoralis* Knockout Linien mit beeinträchtigter 8-HQA Produktion (KMO) generiert. Anschließend wurde das Mikrobiom der KMO Larven analysiert und mit dem der Wildtyp (WT) *S. littoralis* Larven verglichen. Die daraus gewonnenen Erkenntnisse zeigen die Abhängigkeit der mikrobiellen Vielfalt und Häufigkeit der Symbionten von dem genetischen Material des Wirts, in diesem Fall *S. littoralis*. Verglichen mit dem WT wiesen KMO Larven eine höhere Anzahl an Bakterien auf; in Bezug auf die Artenvielfalt und Ausgeglichenheit der Spezies untereinander zeigten die WT Insekten eine bessere Balance. Obwohl ein Großteil der identifizierten Bakterien sowohl in den WTs als auch in den KMOs zur Familie der Enterococcaceae gehörte, beherbergten WT Larven zusätzlich eine beträchtliche Anzahl von Alphaproteobakterien, Actinomyceten und einigen Cyanobakterien, welche auch im Insektenkot detektiert wurden.

Diese Dissertation ermöglicht einen tieferen Einblick in das Darmmikrobiom von *Spodoptera littoralis* Larven. Innerhalb dieser Studie werden die Mechanismen und Anpassungen an erschwerte Bedingungen innerhalb des Wirtsdarms am Beispiel von *Enterococcus mundtii* analysiert. Die daraus gewonnenen Erkenntnisse dienen als repräsentatives Beispiel für weitere Symbionten und Bakterien und deren Überlebensstrategien innerhalb verschiedener Wirte.

10 References

1 Triplehorn, A. & Johnson, F. J. U. S. o. A. Borror and Delongs Introduction to the Study of Insects. Thomson brooks/cole. 263-267 (2005).

2 Abrol, D. P. in *Pollination Biology: Biodiversity Conservation and Agricultural Production* 265-310 (Springer Netherlands, 2012).

3 Pellmyr, O., Thompson, J. N., Brown, J. M. & Harrison, R. G. J. T. A. N. Evolution of pollination and mutualism in the yucca moth lineage. **148**, 827-847 (1996).

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

5 Sagan, L. J. J. o. t. b. On the origin of mitosing cells. **14**, 225-IN226 (1967).

6 O'Hara, A. M. & Shanahan, F. The gut flora as a forgotten organ. **7**, 688-693, doi:10.1038/sj.embor.7400731 (2006).

7 Zilber-Rosenberg, I. & Rosenberg, E. J. F. m. r. Role of microorganisms in the evolution of animals and plants: the hologenome theory of evolution. **32**, 723-735 (2008).

8 Douglas, A. E. & Werren, J. H. J. M. Holes in the hologenome: why host-microbe symbioses are not holobionts. **7**, e02099-02015 (2016).

9 Moran, N. A. & Sloan, D. B. J. P. b. The hologenome concept: helpful or hollow? **13**, e1002311 (2015).

10 Lee, S. M. *et al.* Bacterial colonization factors control specificity and stability of the gut microbiota. **501**, 426 (2013).

11 Kers, J. G. *et al.* Host and environmental factors affecting the intestinal microbiota in chickens. **9**, 235 (2018).

12 Palmer, C., Bik, E. M., DiGiulio, D. B., Relman, D. A. & Brown, P. O. J. P. b. Development of the human infant intestinal microbiota. **5**, e177 (2007).

13 Dreyer, J. L. & Liebl, A. L. J. H. M. J. Early colonization of the gut microbiome and its relationship with obesity. (2018).

14 Bäckhed, F. *et al.* Dynamics and stabilization of the human gut microbiome during the first year of life. **17**, 690-703 (2015).

15 Bokulich, N. A. *et al.* Antibiotics, birth mode, and diet shape microbiome maturation during early life. **8**, 343ra382-343ra382 (2016).

16 Broderick, N. A. J. P. T. o. t. R. S. B. B. S. Friend, foe or food? Recognition and the role of antimicrobial peptides in gut immunity and Drosophila–microbe interactions. **371**, 20150295 (2016).

17 Buchon, N., Silverman, N. & Cherry, S. J. N. r. i. Immunity in Drosophila melanogaster—from microbial recognition to whole-organism physiology. **14**, 796 (2014).

18 Engel, P. & Moran, N. A. J. F. m. r. The gut microbiota of insects–diversity in structure and function. **37**, 699-735 (2013).

19 Rolhion, N. & Chassaing, B. J. P. T. o. t. R. S. B. B. S. When pathogenic bacteria meet the intestinal microbiota. **371**, 20150504 (2016).

20 Lazar, V. *et al.* Aspects of gut microbiota and immune system interactions in infectious diseases, immunopathology and cancer. **9**, 1830 (2018).

21 Johnston, P. R. & Rolff, J. J. P. p. Host and symbiont jointly control gut microbiota during complete metamorphosis. **11**, e1005246 (2015).

22 Dillon, R. J. & Dillon, V. M. The gut bacteria of insects: nonpathogenic interactions. *Annual review of entomology* **49**, 71-92, doi:10.1146/annurev.ento.49.061802.123416 (2004).

23 Engel, P. & Moran, N. A. The gut microbiota of insects – diversity in structure and function. *FEMS Microbiology Reviews* **37**, 699-735, doi:10.1111/1574-6976.12025 %J FEMS Microbiology Reviews (2013).

24 Shao, Y., Arias-Cordero, E., Guo, H., Bartram, S. & Boland, W. In Vivo Pyro-SIP Assessing Active Gut Microbiota of the Cotton Leafworm, Spodoptera littoralis. *PLOS ONE* **9**, e85948, doi:10.1371/journal.pone.0085948 (2014).
25 Tang, X. *et al.* Complexity and variability of gut commensal microbiota in polyphagous lepidopteran larvae. **7** (2012).
26 Köhler, T., Stingl, U., Meuser, K. & Brune, A. J. E. m. Novel lineages of Planctomycetes densely colonize the alkaline gut of soil-feeding termites (Cubitermes spp.). **10**, 1260-1270 (2008).
27 Bignell, D. E. in *Biology of termites: a modern synthesis* 375-412 (Springer, 2010).
28 Appel, H. M. & Martin, M. M. J. J. o. C. E. Gut redox conditions in herbivorous lepidopteran larvae. **16**, 3277-3290 (1990).
29 Lemke, T., Stingl, U., Egert, M., Friedrich, M. W. & Brune, A. J. A. E. M. Physicochemical conditions and microbial activities in the highly alkaline gut of the humus-feeding larva of Pachnoda ephippiata (Coleoptera: Scarabaeidae). **69**, 6650-6658 (2003).
30 Jones, J. C. *et al.* The gut microbiome is associated with behavioural task in honey bees. *Insectes Sociaux* **65**, 419-429, doi:10.1007/s00040-018-0624-9 (2018).
31 Paniagua Voirol, L. R., Frago, E., Kaltenpoth, M., Hilker, M. & Fatouros, N. E. J. F. i. m. Bacterial symbionts in lepidoptera: Their diversity, transmission, and impact on the host. **9**, 556 (2018).
32 Freitak, D. *et al.* The maternal transfer of bacteria can mediate trans-generational immune priming in insects. **5**, 547-554 (2014).
33 Brinkmann, N., Martens, R. & Tebbe, C. C. J. A. E. M. Origin and diversity of metabolically active gut bacteria from laboratory-bred larvae of Manduca sexta (Sphingidae, Lepidoptera, Insecta). **74**, 7189-7196 (2008).
34 Pinto-Tomás, A. *et al.* Actividades enzimáticas en aislamientos bacterianos de tractos digestivos de larvas y del contenido de pupas de Automeris zugana y Rothschildia lebeau (Lepidoptera: Saturniidae). **55**, 401-415 (2007).
35 Indiragandhi, P., Anandham, R., Madhaiyan, M. & Sa, T. J. C. m. Characterization of plant growth–promoting traits of bacteria isolated from larval guts of diamondback moth Plutella xylostella (Lepidoptera: Plutellidae). **56**, 327-333 (2008).
36 Shao, Y. *et al.* Symbiont-derived antimicrobials contribute to the control of the Lepidopteran gut microbiota. **24**, 66-75 (2017).
37 Broderick, N. A., Raffa, K. F., Goodman, R. M. & Handelsman, J. J. A. E. M. Census of the bacterial community of the gypsy moth larval midgut by using culturing and culture-independent methods. **70**, 293-300 (2004).
38 Mithöfer, A. & Boland, W. J. A. r. o. p. b. Plant defense against herbivores: chemical aspects. **63**, 431-450 (2012).
39 Hilker, M. & Fatouros, N. E. J. C. o. i. p. b. Resisting the onset of herbivore attack: plants perceive and respond to insect eggs. **32**, 9-16 (2016).
40 Xia, X. *et al.* Metagenomic sequencing of diamondback moth gut microbiome unveils key holobiont adaptations for herbivory. **8**, 663 (2017).
41 Gershenzon, J. & Dudareva, N. J. N. c. b. The function of terpene natural products in the natural world. **3**, 408 (2007).
42 van der Vlugt, C. J. & van der Werf, M. J. J. A. E. M. Genetic and biochemical characterization of a novel monoterpene ε-lactone hydrolase from Rhodococcus erythropolis DCL14. **67**, 733-741 (2001).
43 Bi, J. & Felton, G. J. J. o. C. E. Foliar oxidative stress and insect herbivory: primary compounds, secondary metabolites, and reactive oxygen species as components of induced resistance. **21**, 1511-1530 (1995).

44 Visôtto, L., Oliveira, M., Ribon, A., Mares-Guia, T. & Guedes, R. J. E. e. Characterization and identification of proteolytic bacteria from the gut of the velvetbean caterpillar (Lepidoptera: Noctuidae). **38**, 1078-1085 (2009).
45 Carlini, C. R. & Grossi-de-Sá, M. F. J. T. Plant toxic proteins with insecticidal properties. A review on their potentialities as bioinsecticides. **40**, 1515-1539 (2002).
46 Ahmed, M. Z., Araujo-Jnr, E. V., Welch, J. J. & Kawahara, A. Y. J. F. i. z. Wolbachia in butterflies and moths: geographic structure in infection frequency. **12**, 16 (2015).
47 Jiggins, F., Hurst, G., Dolman, C. & Majerus, M. J. J. E. B. High-prevalence male-killing Wolbachia in the buttery Acraea encedana. **13**, 495-501 (2000).
48 Dyson, E. A. & Hurst, G. D. J. P. o. t. N. A. o. S. Persistence of an extreme sex-ratio bias in a natural population. **101**, 6520-6523 (2004).
49 Hiroki, M., Kato, Y., Kamito, T. & Miura, K. J. N. Feminization of genetic males by a symbiotic bacterium in a butterfly, Eurema hecabe (Lepidoptera: Pieridae). **89**, 167-170 (2002).
50 Sasaki, T., Kubo, T. & Ishikawa, H. J. G. Interspecific transfer of Wolbachia between two lepidopteran insects expressing cytoplasmic incompatibility: a Wolbachia variant naturally infecting Cadra cautella causes male killing in Ephestia kuehniella. **162**, 1313-1319 (2002).
51 Moriyama, M., Nikoh, N., Hosokawa, T. & Fukatsu, T. J. M. Riboflavin provisioning underlies Wolbachia's fitness contribution to its insect host. **6**, e01732-01715 (2015).
52 Kaiser, W., Huguet, E., Casas, J. m., Commin, C. & Giron, D. J. P. o. t. R. S. B. B. S. Plant green-island phenotype induced by leaf-miners is mediated by bacterial symbionts. **277**, 2311-2319 (2010).
53 Raymond, B. *et al.* A mid-gut microbiota is not required for the pathogenicity of Bacillus thuringiensis to diamondback moth larvae. **11**, 2556-2563 (2009).
54 Baxter, S. *et al.* Novel genetic basis of field-evolved resistance to Bt toxins in Plutella xylostella. **14**, 327-334 (2005).
55 Prakash, O., Jangid, K. & Shouche, Y. S. J. I. j. o. m. Carl Woese: from biophysics to evolutionary microbiology. **53**, 247-252 (2013).
56 Peterson, J. *et al.* The NIH human microbiome project. **19**, 2317-2323 (2009).
57 Chaston, J. & Douglas, A. J. T. B. B. Making the most of “omics” for symbiosis research. **223**, 21-29 (2012).
58 Gil, R., Sabater-Muñoz, B., Latorre, A., Silva, F. J. & Moya, A. J. P. o. t. N. A. o. S. Extreme genome reduction in Buchnera spp.: toward the minimal genome needed for symbiotic life. **99**, 4454-4458 (2002).
59 Salem, H. *et al.* Drastic genome reduction in an herbivore’s pectinolytic symbiont. **171**, 1520-1531. e1513 (2017).
60 Lee, J. B. *et al.* Gut symbiotic bacteria stimulate insect growth and egg production by modulating hexamerin and vitellogenin gene expression. **69**, 12-22 (2017).
61 Ohbayashi, T. *et al.* Comparative cytology, physiology and transcriptomics of Burkholderia insecticola in symbiosis with the bean bug Riptortus pedestris and in culture. **13**, 1469 (2019).
62 Bonilla-Rosso, G. & Engel, P. J. C. o. i. m. Functional roles and metabolic niches in the honey bee gut microbiota. **43**, 69-76 (2018).
63 Hu, Y. *et al.* Herbivorous turtle ants obtain essential nutrients from a conserved nitrogen-recycling gut microbiome. **9**, 964 (2018).
64 Sapountzis, P. *et al.* Potential for nitrogen fixation in the fungus-growing termite symbiosis. **7**, 1993 (2016).
65 Sarannia, T. & Kamariah, H. J. P. J. o. T. A. S. Diversity of nitrogen fixing bacteria associated with various termite species. **41**, 925-940 (2018).
66 Dowd, P. F. In situ production of hydrolytic detoxifying enzymes by symbiotic yeasts in the cigarette beetle (Coleoptera: Anobiidae). *J Econ Entomol* **82**, 396-400, doi:10.1093/jee/82.2.396 (1989).

67 Genta, F. A., Dillon, R. J., Terra, W. R. & Ferreira, C. Potential role for gut microbiota in cell wall digestion and glucoside detoxification in Tenebrio molitor larvae. *Journal of insect physiology* **52**, 593-601, doi:10.1016/j.jinsphys.2006.02.007 (2006).
68 Terra, W. R. The origin and functions of the insect peritrophic membrane and peritrophic gel. *Archives of insect biochemistry and physiology* **47**, 47-61, doi:10.1002/arch.1036 (2001).
69 Dale, C. & Moran, N. A. Molecular interactions between bacterial symbionts and their hosts. *Cell* **126**, 453-465, doi:10.1016/j.cell.2006.07.014 (2006).
70 Currie, C. R., Scott, J. A., Summerbell, R. C. & Malloch, D. Fungus-growing ants use antibiotic-producing bacteria to control garden parasites. *Nature* **398**, 701-704, doi:10.1038/19519 (1999).
71 Kroiss, J. *et al.* Symbiotic Streptomycetes provide antibiotic combination prophylaxis for wasp offspring. *Nature chemical biology* **6**, 261-263, doi:10.1038/nchembio.331 (2010).
72 Oliver, K. M., Russell, J. A., Moran, N. A. & Hunter, M. S. Facultative bacterial symbionts in aphids confer resistance to parasitic wasps. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 1803-1807, doi:10.1073/pnas.0335320100 (2003).
73 Shade, A. & Handelsman, J. Beyond the Venn diagram: the hunt for a core microbiome. *Environmental microbiology* **14**, 4-12, doi:10.1111/j.1462-2920.2011.02585.x (2012).
74 Tang, X. *et al.* Complexity and variability of gut commensal microbiota in polyphagous lepidopteran larvae. *PloS one* **7**, e36978, doi:10.1371/journal.pone.0036978 (2012).
75 Cox, C. R. & Gilmore, M. S. Native microbial colonization of Drosophila melanogaster and its use as a model of Enterococcus faecalis pathogenesis. *Infect Immun* **75**, 1565-1576, doi:10.1128/IAI.01496-06 (2007).
76 Lehman, R. M., Lundgren, J. G. & Petzke, L. M. Bacterial communities associated with the digestive tract of the predatory ground beetle, Poecilus chalcites, and their modification by laboratory rearing and antibiotic treatment. *Microbial ecology* **57**, 349-358, doi:10.1007/s00248-008-9415-6 (2009).
77 Ryu, J. H. *et al.* Innate immune homeostasis by the homeobox gene caudal and commensal-gut mutualism in Drosophila. *Science* **319**, 777-782, doi:10.1126/science.1149357 (2008).
78 Teh, B. S., Apel, J., Shao, Y. & Boland, W. Colonization of the Intestinal Tract of the Polyphagous Pest Spodoptera littoralis with the GFP-Tagged Indigenous Gut Bacterium Enterococcus mundtii. *Front Microbiol* **7**, 928, doi:10.3389/fmicb.2016.00928 (2016).
79 Doyle, T. C., Burns, S. M. & Contag, C. H. In vivo bioluminescence imaging for integrated studies of infection. *Cell Microbiol* **6**, 303-317, doi:10.1111/j.1462-5822.2004.00378.x (2004).
80 Koo, J. *et al.* A GUS/luciferase fusion reporter for plant gene trapping and for assay of promoter activity with luciferin-dependent control of the reporter protein stability. *Plant Cell Physiol* **48**, 1121-1131, doi:10.1093/pcp/pcm081 (2007).
81 Pinheiro, L. B., Gibbs, M. D., Vesey, G., Smith, J. J. & Bergquist, P. L. Fluorescent reference strains of bacteria by chromosomal integration of a modified green fluorescent protein gene. *Appl Microbiol Biotechnol* **77**, 1287-1295, doi:10.1007/s00253-007-1253-9 (2008).
82 Rhee, K. J. *et al.* Determination of spatial and temporal colonization of enteropathogenic E. coli and enterohemorrhagic E. coli in mice using bioluminescent in vivo imaging. *Gut Microbes* **2**, 34-41, doi:10.4161/gmic.2.1.14882 (2011).
83 De Vos, W. M. Gene cloning and expression in lactic streptococci. *FEMS Microbiol Rev* **46**, 281-295 (1987).
84 Eaton, T. J., Shearman, C. A. & Gasson, M. J. The use of bacterial luciferase genes as reporter genes in *Lactococcus*: regulation of the *Lactococcus lactis* subsp. *lactis* lactose genes. *J Gen Microbiol* **139**, 1495-1501, doi:10.1099/00221287-139-7-1495 (1993).
85 Platteeuw, C. G., Simons, G. & De Vos, W. M. Use of *Escherichia coli* β-glucoronidase (*gusA*) gene as a reporter gene for analysing promoters in lactic acid bacteria. *Appl Environ Microbiol* **60**, 587-593 (1994).

86 Botes, M., van Reenen, C. A. & Dicks, L. M. Evaluation of Enterococcus mundtii ST4SA and Lactobacillus plantarum 423 as probiotics by using a gastro-intestinal model with infant milk formulations as substrate. *Int J Food Microbiol* **128**, 362-370, doi:10.1016/j.ijfoodmicro.2008.09.016 (2008).
87 Duangjitcharoen, Y., Kantachote, D., Ongsakul, M., Poosaran, N. & Chaiyasut, C. Potential use of probiotic *Lactobacillus plantarum* SS2 isolated from fermented plant beverage: safety assessment and persistence in the murine gastrointestinal tract. *World J Microbiol Biotechnol* **25**, 315-321, doi:10.1007/s11274-008-9894-0 (2009).
88 Yu, Q. H., Dong, S. M., Zhu, W. Y. & Yang, Q. Use of green fluorescent protein to monitor Lactobacillus in the gastro-intestinal tract of chicken. *FEMS Microbiol Lett* **275**, 207-213, doi:10.1590/S1517-838246320140556 (2007).
89 Eberl, L. *et al.* Use of green fluorescent protein as a marker for ecological studies of activated sludge communities. *FEMS Microbiol Lett* **149**, 77-83 (1997).
90 Leff, L. G. & Leff, A. A. Use of green fluorescent protein to monitor survival of genetically engineered bacteria in aquatic environments. *Appl Environ Microbiol* **62**, 3486-3488 (1996).
91 Gage, D. J., Bobo, T. & Long, S. R. Use of green fluorescent protein to visualize the early events of symbiosis between *Rhizobium meliloti* and alfalfa (*Medicago sativa*). *J Bacteriol* **178**, 7159-7166 (1996).
92 Kremer, L., Baulard, A., Estaquier, J., Poulain-Godefroy, O. & Locht, C. Green fluorescent protein as a new expression marker in mycobacteria. *Mol Microbiol* **17**, 913-922, doi:10.1111/j.1365-2958.1995.mmi_17050913.x (1995).
93 Phillips, G. J. Green fluorescent protein--a bright idea for the study of bacterial protein localization. *FEMS Microbiol Lett* **204**, 9-18 (2001).
94 Margolin, W. Green fluorescent protein as a reporter for macromolecular localization in bacterial cells. *Methods* **20**, 62-72, doi:10.1006/meth.1999.0906 (2000).
95 Tsien, R. Y. The green fluorescent protein. *Annu Rev Biochem* **67**, 509-544, doi:10.1146/annurev.biochem.67.1.509 (1998).
96 Stepanenko, O. V., Verkhusha, V. V., Kuznetsova, I. M., Uversky, V. N. & Turoverov, K. K. Fluorescent proteins as biomarkers and biosensors: throwing color lights on molecular and cellular processes. *Curr Protein Pept Sci* **9**, 338-369 (2008).
97 Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W. & Prasher, D. C. Green fluorescent protein as a marker for gene expression. *Science* **263**, 802-805 (1994).
98 Valdivia, R. H., Hromockyj, A. E., Monack, D., Ramakrishnan, L. & Falkow, S. Applications for green fluorescent protein (GFP) in the study of host-pathogen interactions. *Gene* **173**, 47-52 (1996).
99 Bokman, S. H. & Ward, W. W. Renaturation of Aequorea gree-fluorescent protein. *Biochem Biophys Res Commun* **101**, 1372-1380, doi:10.1016/0006-291X(81)91599-0 (1981).
100 Cormack, B. P., Valdivia, R. H. & Falkow, S. FACS-optimized mutants of the green fluorescent protein (GFP). *Gene* **173**, 33-38, doi:10.1016/0378-1119(95)00685-0 (1996).
101 Pedelacq, J. D., Cabantous, S., Tran, T., Terwilliger, T. C. & Waldo, G. S. Engineering and characterization of a superfolder green fluorescent protein. *Nat Biotechnol* **24**, 79-88, doi:10.1038/nbt1172 (2006).
102 Scholz, O., Thiel, A., Hillen, W. & Niederweis, M. Quantitative analysis of gene expression with an improved green fluorescent protein. p6. *Eur J Biochem* **267**, 1565-1570 (2000).
103 Shaner, N. C., Steinbach, P. A. & Tsien, R. Y. A guide to choosing fluorescent proteins. *Nat Methods* **2**, 905-909, doi:10.1038/nmeth819 (2005).
104 Shaner, N. C. *et al.* Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. *Nat Biotechnol* **22**, 1567-1572, doi:10.1038/nbt1037 (2004).
105 Alieva, N. O. *et al.* Diversity and evolution of coral fluorescent proteins. *PloS one* **3**, e2680, doi:10.1371/journal.pone.0002680 (2008).

106 Foulquie Moreno, M. R., Sarantinopoulos, P., Tsakalidou, E. & De Vuyst, L. The role and application of enterococci in food and health. *Int J Food Microbiol* **106**, 1-24, doi:10.1016/j.ijfoodmicro.2005.06.026 (2006).
107 Bhardwaj, A. *et al.* Safety assessment and evaluation of probiotic potential of bacteriocinogenic Enterococcus faecium KH 24 strain under in vitro and in vivo conditions. *Int J Food Microbiol* **141**, 156-164, doi:10.1016/j.ijfoodmicro.2010.05.001 (2010).
108 Andreu, N., Zelmer, A. & Wiles, S. Noninvasive biophotonic imaging for studies of infectious disease. *FEMS Microbiol Rev* **35**, 360-394, doi:10.1111/j.1574-6976.2010.00252.x.
109 Zhang, W. Y. & Zhang, H. P. Genomics of lactic acid bacteria. In: Zhang, H. P, Cai, Y. M. (Eds). Lactic Acid Bacteria-Fundamentals and Practice, 1st ed. Springer Publishing Inc, New York, NY. 235-238 (2014).
110 Wang, T. T. & Lee, B. H. Plasmids in Lactobacillus. *Crit Rev Biotechnol* **17**, 227-272, doi:10.3109/07388559709146615 (1997).
111 Schroeter, J. & Klaenhammer, T. Genomics of lactic acid bacteria. *FEMS Microbiol Lett* **292**, 1-6, doi:10.1111/j.1574-6968.2008.01442.x (2009).
112 Mills, S., McAuliffe, O. E., Coffey, A., Fitzgerald, G. F. & Ross, R. P. Plasmids of lactococci - genetic accessories or genetic necessities? *FEMS Microbiol Rev* **30**, 243-273, doi:10.1111/j.1574-6976.2005.00011.x (2006).
113 Clewell, D. B., Yagi, Y., Dunny, G. M. & Schultz, S. K. Characterization of three plasmid deoxyribonucleic acid molecules in a strain of Streptococcus faecalis: identification of a plasmid determining erythromycin resistance. *J Bacteriol* **117**, 283-289 (1974).
114 Francia, M. V. & Clewell, D. B. Amplification of the tetracycline resistance determinant of pAMalpha1 in Enterococcus faecalis requires a site-specific recombination event involving relaxase. *J Bacteriol* **184**, 5187-5193, doi:10.1128/JB.184.18.5187-5193.2002 (2002).
115 Tanimoto, K. & Ike, Y. Complete nucleotide sequencing and analysis of the 65-kb highly conjugative Enterococcus faecium plasmid pMG1: identification of the transfer-related region and the minimum region required for replication. *FEMS Microbiol Lett* **288**, 186-195, doi:10.1111/j.1574-6968.2008.01342.x (2008).
116 Flannagan, S. E. *et al.* Plasmid content of a vancomycin-resistant Enterococcus faecalis isolate from a patient also colonized by Staphylococcus aureus with a VanA phenotype. *Antimicrob Agents Chemother* **47**, 3954-3959, doi:10.1128/AAC.47.12.3954-3959.2003 (2003).
117 Balla, E. & Dicks, L. M. Molecular analysis of the gene cluster involved in the production and secretion of enterocins 1071A and 1071B and of the genes responsible for the replication and transfer of plasmid pEF1071. *Int J Food Microbiol* **99**, 33-45, doi:10.1016/j.ijfoodmicro.2004.08.008 (2005).
118 Ruiz-Barba, J. L., Floriano, B., Maldonado-Barragan, A. & Jimenez-Diaz, R. Molecular analysis of the 21-kb bacteriocin-encoding plasmid pEF1 from Enterococcus faecium 6T1a. *Plasmid* **57**, 175-181, doi:10.1016/j.plasmid.2006.06.003 (2007).
119 Criado, R. *et al.* Complete sequence of the enterocin Q-encoding plasmid pCIZ2 from the multiple bacteriocin producer Enterococcus faecium L50 and genetic characterization of enterocin Q production and immunity. *Appl Environ Microbiol* **72**, 6653-6666, doi:10.1128/AEM.00859-06 (2006).
120 De Boever, E. H., Clewell, D. B. & Fraser, C. M. Enterococcus faecalis conjugative plasmid pAM373: complete nucleotide sequence and genetic analyses of sex pheromone response. *Mol Microbiol* **37**, 1327-1341, doi:mmi2072 [pii] (2000).
121 Paulsen, I. T. *et al.* Role of mobile DNA in the evolution of vancomycin-resistant Enterococcus faecalis. *Science* **299**, 2071-2074, doi:10.1126/science.1080613 (2003).
122 Brede, D. A., Snipen, L. G., Ussery, D. W., Nederbragt, A. J. & Nes, I. F. Complete genome sequence of the commensal Enterococcus faecalis 62, isolated from a healthy Norwegian infant. *J Bacteriol* **193**, 2377-2378, doi:10.1128/JB.00183-11 (2011).

123 Hirt, H. *et al.* Characterization of the pheromone response of the Enterococcus faecalis conjugative plasmid pCF10: complete sequence and comparative analysis of the transcriptional and phenotypic responses of pCF10-containing cells to pheromone induction. *J Bacteriol* **187**, 1044-1054, doi:10.1128/JB.187.3.1044-1054.2005 (2005).
124 O'Sullivan, D. J. & Klaenhammer, T. R. High- and low-copy-number Lactococcus shuttle cloning vectors with features for clone screening. *Gene* **137**, 227-231, doi:10.1016/0378-1119(93)90011-Q (1993).
125 Swinfield, T. J., Oultram, J. D., Thompson, D. E., Brehm, J. K. & Minton, N. P. Physical characterisation of the replication region of the Streptococcus faecalis plasmid pAM beta 1. *Gene* **87**, 79-90 (1990).
126 Brantl, S. & Behnke, D. Characterization of the minimal origin required for replication of the streptococcal plasmid pIP501 in Bacillus subtilis. *Mol Microbiol* **6**, 3501-3510 (1992).
127 Shareck, J., Choi, Y., Lee, B. & Miguez, C. B. Cloning vectors based on cryptic plasmids isolated from lactic acid bacteria: their characteristics and potential applications in biotechnology. *Crit Rev Biotechnol* **24**, 155-208, doi:10.1080/07388550490904288 (2004).
128 Lizier, M., Sarra, P. G., Cauda, R. & Lucchini, F. Comparison of expression vectors in Lactobacillus reuteri strains. *FEMS Microbiol Lett* **308**, 8-15, doi:10.1111/j.1574-6968.2010.01978.x (2010).
129 Kim, S. F., Baek, S. J. & Pack, M. Y. Cloning and nucleotide sequence of the Lactobacillus casei lactate dehydrogenase gene. *Appl Environ Microbiol* **57**, 2413-2417 (1991).
130 Boot, H. J. & Pouwels, P. H. Expression, secretion and antigenic variation of bacterial S-layer proteins. *Mol Microbiol* **21**, 1117-1123, doi:10.1046/j.1365-2958.1996.711442.x (1996).
131 Sheehan, V. M., Sleator, R. D., Fitzgerald, G. F. & Hill, C. Heterologous expression of BetL, a betaine uptake system, enhances the stress tolerance of Lactobacillus salivarius UCC118. *Appl Environ Microbiol* **72**, 2170-2177, doi:10.1128/AEM.72.3.2170-2177.2006 (2006).
132 Wu, C. M. & Chung, T. C. Green fluorescent protein is a reliable reporter for screening signal peptides functional in Lactobacillus reuteri. *J Microbiol Methods* **67**, 181-186, doi:10.1016/j.mimet.2006.03.009 (2006).
133 Bernbom, N. *et al.* Effects of Lactococcus lactis on composition of intestinal microbiota: role of nisin. *Appl Environ Microbiol* **72**, 239-244, doi:10.1128/AEM.72.1.239-244.2006 (2006).
134 Reunanen, J. & Saris, P. E. Survival of nisin activity in intestinal environment. *Biotechnol Lett* **31**, 1229-1232, doi:10.1007/s10529-009-9995-3 (2009).
135 Gory, L., Montel, M. C. & Zagorec, M. Use of green fluorescent protein to monitor *Lactobacillus sakei* in fermented meat products. *FEMS Microbiol Lett* **194**, 127-133, doi:10.1111/j.1574-6968.2001.tb09457.x (2001).
136 Boot, H. J., Kolen, C. P., Andreadaki, F. J., Leer, R. J. & Pouwels, P. H. The Lactobacillus acidophilus S-layer protein gene expression site comprises two consensus promoter sequences, one of which directs transcription of stable mRNA. *J Bacteriol* **178**, 5388-5394, doi:10.1128/jb.178.18.5388-5394.1996 (1996).
137 Djordjevic, G., Bojovic, B., Banina, A. & Topisirovic, L. Cloning of promoter-like sequences from Lactobacillus paracasei subsp. paracasei CG11 and their expression in Escherichia coli, Lactococcus lactis, and Lactobacillus reuteri. *Can J Microbiol* **40**, 1043-1050, doi:10.1139/m94-165 (1994).
138 Chang, T. L. *et al.* Inhibition of HIV infectivity by a natural human isolate of Lactobacillus jensenii engineered to express functional two-domain CD4. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 11672-11677, doi:10.1073/pnas.1934747100 (2003).
139 Chouayekh, H., Serror, P., Boudebbouze, S. & Maguin, E. Highly efficient production of the staphylococcal nuclease reporter in Lactobacillus bulgaricus governed by the promoter of the hlbA gene. *FEMS Microbiol Lett* **293**, 232-239, doi:10.1111/j.1574-6968.2009.01522.x (2009).

140 Fang, F. *et al.* Characterization of endogenous plasmids from Lactobacillus salivarius UCC118. *Appl Environ Microbiol* **74**, 3216-3228, doi:10.1128/AEM.02631-07 (2008).
141 Miller, J. F., Dower, W. J. & Tompkins, L. S. High-voltage electroporation of bacteria: genetic transformation of Campylobacter jejuni with plasmid DNA. *Proceedings of the National Academy of Sciences of the United States of America* **85**, 856-860 (1988).
142 Rixon, J. & Warner, P. J. Background, relevant genetic techniques and terms. In: Wood B. J. B, Warner P. J. (Eds). Genetics of lactic acid bacteria. Kluwer Academic/Plenum Publishers, New York, NY. 1-24 (2003).
143 Powell, I. B., Achen, M. G., Hillier, A. J. & Davidson, B. E. A Simple and Rapid Method for Genetic Transformation of Lactic Streptococci by Electroporation. *Appl Environ Microbiol* **54**, 655-660 (1988).
144 Rodriguez, M. C., Alegre, M. T. & Mesas, J. M. Optimization of technical conditions for the transformation of Pediococcus acidilactici P60 by electroporation. *Plasmid* **58**, 44-50, doi:10.1016/j.plasmid.2006.12.005 (2007).
145 van der Lelie, D., Bron, S., Venema, G. & Oskam, L. Similarity of minus origins of replication and flanking open reading frames of plasmids pUB110, pTB913 and pMV158. *Nucleic Acids Res* **17**, 7283-7294 (1989).
146 Dornan, S. & Collins, M. A. High efficiency electroporation of Lactococcus lactis subsp. lactis LM0230 with plasmid pGB301. *Lett Appl Microbiol* **11**, 62-64, doi:10.1111/j.1472-765X.1990.tb01275.x (1990).
147 Wei, M. Q. *et al.* An improved method for the transformation of *Lactobacillus* strains using electroporation. *J Microbiol Methods* **21**, 97-109, doi:10.1016/0167-7012(94)00038-9 (1995).
148 Assad-Garcia, J. S. *et al.* An improved protocol for electroporation of Oenococcus oeni ATCC BAA-1163 using ethanol as immediate membrane fluidizing agent. *Lett Appl Microbiol* **47**, 333-338 (2008).
149 Holo, H. & Nes, I. F. High-Frequency Transformation, by Electroporation, of Lactococcus lactis subsp. cremoris Grown with Glycine in Osmotically Stabilized Media. *Appl Environ Microbiol* **55**, 3119-3123 (1989).
150 Thompson, K. & Collins, M. A. Improvement in electroporation efficiency for Lactobacillus plantarum by the inclusion of high concentrations of glycine in the growth medium *J Microbiol Methods* **26**, 73-79, doi:10.1016/0167-7012(96)00845-7 (1996).
151 A, O. C. & McClean, S. The role of Universal Stress Proteins in Bacterial Infections. *Current medicinal chemistry*, doi:10.2174/0929867324666170124145543 (2017).
152 Hanahan, D., Jessee, J. & Bloom, F. R. Plasmid transformation of Escherichia coli and other bacteria. *Methods Enzymol* **204**, 63-113, doi:10.1016/0076-6879(91)04006-A (1991).
153 Borralho, T., Chang, Y., Jain, P., Lalani, M. & Parghi, K. Effect of electroporation versus hanahan protocols on the transformation of *Escherichia coli* HB101 with chromosomal DNA from *Escherichia coli* HB101, *Escherichia coli* B23, and *Bacillus subtilis* WB746 and the plasmid p328.5, including an analysis of competent *Escherichia coli* HB101 cellular freeze tolerance. *J Exp Microbiol Immunol* **2**, 194-200 (2002).
154 Aune, T. E. & Aachmann, F. L. Methodologies to increase the transformation efficiencies and the range of bacteria that can be transformed. *Appl Microbiol Biotechnol* **85**, 1301-1313, doi:10.1007/s00253-009-2349-1 (2010).
155 Yoshida, N. & Sato, M. Plasmid uptake by bacteria: a comparison of methods and efficiencies. *Appl Microbiol Biotechnol* **83**, 791-798, doi:10.1007/s00253-009-2042-4 (2009).
156 Tu, Q. *et al.* Room temperature electrocompetent bacterial cells improve DNA transformation and recombineering efficiency. *Scientific reports* **6**, 24648, doi:10.1038/srep24648 (2016).
157 Hu, S. *et al.* Genome engineering of Agrobacterium tumefaciens using the lambda Red recombination system. *Appl Microbiol Biotechnol* **98**, 2165-2172, doi:10.1007/s00253-013-5412-x (2014).

158 Yin, J. *et al.* A new recombineering system for Photorhabdus and Xenorhabdus. *Nucleic Acids Res* **43**, e36, doi:10.1093/nar/gku1336 (2015).
159 Dower, W. J., Miller, J. F. & Ragsdale, C. W. High efficiency transformation of E. coli by high voltage electroporation. *Nucleic Acids Res* **16**, 6127-6145, doi:10.1093/nar/16.13.6127 (1988).
160 Muller, S. & Nebe-von-Caron, G. Functional single-cell analyses: flow cytometry and cell sorting of microbial populations and communities. *FEMS Microbiol Rev* **34**, 554-587, doi:10.1111/j.1574-6976.2010.00214.x (2010).
161 Uhlenhaut, C. & Kracht, M. Viral infectivity is maintained by an RNA protection buffer. *Journal of virological methods* **128**, 189-191, doi:10.1016/j.jviromet.2005.05.002 (2005).
162 Heptinstall, J. Total RNA Isolation from Bacteria. In: Rapley R. (ed.) The Nucleic Acid Protocols Handbook. Humana Press, Totowa, NJ. 47-52, doi:10.1385/1-59259-038-1:47 (2000).
163 Parekh, S., Ziegenhain, C., Vieth, B., Enard, W. & Hellmann, I. The impact of amplification on differential expression analyses by RNA-seq. *Scientific reports* **6**, 25533, doi:10.1038/srep25533 (2016).
164 Hoeijmakers, W. A., Bartfai, R., Francoijs, K. J. & Stunnenberg, H. G. Linear amplification for deep sequencing. *Nature protocols* **6**, 1026-1036, doi:10.1038/nprot.2011.345 (2011).
165 Roche, F. M. *et al.* Characterization of novel LPXTG-containing proteins of Staphylococcus aureus identified from genome sequences. *Microbiology* **149**, 643-654, doi:10.1099/mic.0.25996-0 (2003).
166 Brinster, S., Furlan, S. & Serror, P. C-terminal WxL domain mediates cell wall binding in Enterococcus faecalis and other gram-positive bacteria. *Journal of bacteriology* **189**, 1244-1253, doi:10.1128/JB.00773-06 (2007).
167 Alm, E., Huang, K. & Arkin, A. The evolution of two-component systems in bacteria reveals different strategies for niche adaptation. *PLoS computational biology* **2**, e143, doi:10.1371/journal.pcbi.0020143 (2006).
168 Rutherford, S. T. & Bassler, B. L. Bacterial quorum sensing: its role in virulence and possibilities for its control. *Cold Spring Harbor perspectives in medicine* **2**, doi:10.1101/cshperspect.a012427 (2012).
169 Volker, U. *et al.* Analysis of the induction of general stress proteins of Bacillus subtilis. *Microbiology* **140 (Pt 4)**, 741-752, doi:10.1099/00221287-140-4-741 (1994).
170 Kotrba, P., Inui, M. & Yukawa, H. Bacterial phosphotransferase system (PTS) in carbohydrate uptake and control of carbon metabolism. *Journal of bioscience and bioengineering* **92**, 502-517, doi:10.1016/S1389-1723(01)80308-X (2001).
171 Romano, A. H., Eberhard, S. J., Dingle, S. L. & McDowell, T. D. Distribution of the phosphoenolpyruvate: glucose phosphotransferase system in bacteria. *J Bacteriol* **104**, 808-813 (1970).
172 Gasson, M. J. Progress and potential in the biotechnology of lactic acid bacteria. *FEMS Microbiol Rev* **12**, 3-19, doi:10.1111/j.1574-6976.1993.tb00010.x (1993).
173 Wells, J. M., Wilson, P. W., Norton, P. M., Gasson, M. J. & Le Page, R. W. Lactococcus lactis: high-level expression of tetanus toxin fragment C and protection against lethal challenge. *Mol Microbiol* **8**, 1155-1162, doi:10.1111/j.1365-2958.1993.tb01660.x (1993).
174 Klijn, N., Weerkamp, A. H. & de Vos, W. M. Genetic marking of Lactococcus lactis shows its survival in the human gastrointestinal tract. *Appl Environ Microbiol* **61**, 2771-2774 (1995).
175 Gruzza, M., Duval-Iflah, Y. & Ducluzeau, R. Colonization of the digestive tract of germ-free mice by genetically engineered strains of Lactococcus lactis: study of recombinant DNA stability. *Microb Releases* **1**, 165-171 (1992).
176 Teh, B.-S., Apel, J., Shao, Y. & Boland, W. J. F. i. m. Colonization of the intestinal tract of the polyphagous pest Spodoptera littoralis with the GFP-tagged indigenous gut bacterium Enterococcus mundtii. **7**, 928 (2016).

177 Favia, G. *et al.* Bacteria of the genus Asaia stably associate with Anopheles stephensi, an Asian malarial mosquito vector. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 9047-9051, doi:10.1073/pnas.0610451104 (2007).
178 Funkhouser, L. J. & Bordenstein, S. R. Mom knows best: the universality of maternal microbial transmission. *PLoS Biol* **11**, e1001631, doi:10.1371/journal.pbio.1001631 (2013).
179 Engel, P. & Moran, N. A. The gut microbiota of insects - diversity in structure and function. *FEMS Microbiol Rev* **37**, 699-735, doi:10.1111/1574-6976.12025 (2013).
180 Wieczorek, H., Beyenbach, K. W., Huss, M. & Vitavska, O. Vacuolar-type proton pumps in insect epithelia. *The Journal of experimental biology* **212**, 1611-1619, doi:10.1242/jeb.030007 (2009).
181 Funke, M. *et al.* Rapid hydrolysis of quorum-sensing molecules in the gut of lepidopteran larvae. *Chembiochem : a European journal of chemical biology* **9**, 1953-1959, doi:10.1002/cbic.200700781 (2008).
182 Cruden, D. L. & Markovetz, A. J. Microbial ecology of the cockroach gut. *Annu Rev Microbiol* **41**, 617-643, doi:10.1146/annurev.mi.41.100187.003153 (1987).
183 Santo Domingo, J. W., Kaufman, M. G., Klug, M. J. & Tiedje, J. M. Characterization of the Cricket Hindgut Microbiota with Fluorescently Labeled rRNA-Targeted Oligonucleotide Probes. *Appl Environ Microbiol* **64**, 752-755 (1998).
184 Bauer, S., Tholen, A., Overmann, J. & Brune, A. Characterization of abundance and diversity of lactic acid bacteria in the hindgut of wood- and soil-feeding termites by molecular and culture-dependent techniques. *Arch Microbiol* **173**, 126-137, doi:10.1007/s002039900120 (2000).
185 Chapman, R. F., Simpson, S. J. & Douglas, A. E. *The Insects: Structure and Function*. 5th edn, (Cambridge University Press, 2013).
186 Huang, S. W., Zhang, H. Y., Marshall, S. & Jackson, T. A. The scarab gut: A potential bioreactor for bio-fuel production. *Insect Sci* **17**, 175-183, doi:10.1111/j.1744-7917.2010.01320.x (2010).
187 Lemaitre, B. & Hoffmann, J. The host defense of Drosophila melanogaster. *Annu Rev Immunol* **25**, 697-743, doi:10.1146/annurev.immunol.25.022106.141615 (2007).
188 Shanbhag, S. & Tripathi, S. Epithelial ultrastructure and cellular mechanisms of acid and base transport in the Drosophila midgut. *The Journal of experimental biology* **212**, 1731-1744, doi:10.1242/jeb.029306 (2009).
189 Shao, L., Devenport, M. & Jacobs-Lorena, M. The peritrophic matrix of hematophagous insects. *Archives of insect biochemistry and physiology* **47**, 119-125, doi:10.1002/arch.1042 (2001).
190 Egert, M., Wagner, B., Lemke, T., Brune, A. & Friedrich, M. W. Microbial community structure in midgut and hindgut of the humus-feeding larva of Pachnoda ephippiata (Coleoptera: Scarabaeidae). *Appl Environ Microbiol* **69**, 6659-6668, doi:10.1128/AEM.69.11.6659-6668.2003 (2003).
191 Broderick, N. A., Raffa, K. F., Goodman, R. M. & Handelsman, J. Census of the bacterial community of the gypsy moth larval midgut by using culturing and culture-independent methods. *Appl Environ Microbiol* **70**, 293-300, doi:10.1128/AEM.70.1.293-300.2004 (2004).
192 Moll, R. M., Romoser, W. S., Modrzakowski, M. C., Moncayo, A. C. & Lerdthusnee, K. Meconial peritrophic membranes and the fate of midgut bacteria during mosquito (Diptera: Culicidae) metamorphosis. *J Med Entomol* **38**, 29-32 (2001).
193 Kwong, W. K., Engel, P., Koch, H. & Moran, N. A. Genomics and host specialization of honey bee and bumble bee gut symbionts. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 11509-11514, doi:10.1073/pnas.1405838111 (2014).
194 Engel, P. & Moran, N. A. The gut microbiota of insects–diversity in structure and function. *FEMS microbiology reviews* **37**, 699-735 (2013).

195 Shao, Y. *et al.* Symbiont-Derived Antimicrobials Contribute to the Control of the Lepidopteran Gut Microbiota. *Cell chemical biology* **24**, 66-75, doi:10.1016/j.chembiol.2016.11.015 (2017).
196 Lee, J. B. *et al.* Bacterial cell motility of Burkholderia gut symbiont is required to colonize the insect gut. *Febs Lett* **589**, 2784-2790, doi:10.1016/j.febslet.2015.08.022 (2015).
197 Hammer, T. J., Janzen, D. H., Hallwachs, W., Jaffe, S. P. & Fierer, N. Caterpillars lack a resident gut microbiome. *Proceedings of the National Academy of Sciences of the United States of America* **114**, 9641-9646, doi:10.1073/pnas.1707186114 (2017).
198 Wieczorek, H., Beyenbach, K. W., Huss, M. & Vitavska, O. Vacuolar-type proton pumps in insect epithelia. *Journal of Experimental Biology* **212**, 1611-1619 (2009).
199 Teh, B.-S., Apel, J., Shao, Y. & Boland, W. Colonization of the intestinal tract of the polyphagous pest Spodoptera littoralis with the GFP-tagged indigenous gut bacterium Enterococcus mundtii. *Frontiers in microbiology* **7**, 928 (2016).
200 Magni, C. *et al.* Draft genome sequence of Enterococcus mundtii CRL1656. *Journal of bacteriology* **194**, 550, doi:10.1128/JB.06415-11 (2012).
201 Johnston, P. R. & Rolff, J. J. P. p. Host and symbiont jointly control gut microbiota during complete metamorphosis. **11** (2015).
202 Ishag, H. Z., Xiong, Q., Liu, M., Feng, Z. & Shao, G. E. coli recA gene improves gene targeted homologous recombination in Mycoplasma hyorhinis. *Journal of microbiological methods* **136**, 49-56 (2017).
203 Hanchi, H., Mottawea, W., Sebei, K. & Hammami, R. The Genus Enterococcus: Between Probiotic Potential and Safety Concerns—An Update. *Frontiers in microbiology* **9** (2018).
204 Pesek, J., Svoboda, J., Sattler, M., Bartram, S. & Boland, W. Biosynthesis of 8-hydroxyquinoline-2-carboxylic acid, an iron chelator from the gut of the lepidopteran Spodoptera littoralis. *Organic & biomolecular chemistry* **13**, 178-184, doi:10.1039/c4ob01857e (2015).
205 Silva, B. & Faustino, P. An overview of molecular basis of iron metabolism regulation and the associated pathologies. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease* **1852**, 1347-1359 (2015).
206 Shiwa, Y. *et al.* Complete genome sequence of Enterococcus mundtii QU 25, an efficient L-(+)-lactic acid-producing bacterium. *DNA research* **21**, 369-377 (2014).
207 Voronina, O. L. *et al.* Burkholderia contaminans Biofilm Regulating Operon and Its Distribution in Bacterial Genomes. *BioMed research international* **2016** (2016).
208 Otto, M. Physical stress and bacterial colonization. *FEMS microbiology reviews* **38**, 1250-1270, doi:10.1111/1574-6976.12088 (2014).
209 Tellam, R. L., Wijffels, G. & Willadsen, P. Peritrophic matrix proteins. *Insect biochemistry and molecular biology* **29**, 87-101 (1999).
210 Hancock, L. E., Murray, B. E. & Sillanpaa, J. in *Enterococci: From Commensals to Leading Causes of Drug Resistant Infection* (eds M. S. Gilmore, D. B. Clewell, Y. Ike, & N. Shankar) (2014).
211 Tomusiak-Plebanek, A. *et al.* Lactobacilli with superoxide dismutase-like or catalase activity are more effective in alleviating inflammation in an inflammatory bowel disease mouse model. *Drug design, development and therapy* **12**, 3221 (2018).
212 Kim, H., Goo, E., Kang, Y., Kim, J. & Hwang, I. Regulation of universal stress protein genes by quorum sensing and RpoS in Burkholderia glumae. *Journal of bacteriology* **194**, 982-992, doi:10.1128/JB.06396-11 (2012).
213 Rao, C. V. S., De Waelheyns, E., Economou, A. & Anne, J. Antibiotic targeting of the bacterial secretory pathway. *Biochimica et biophysica acta* **1843**, 1762-1783, doi:10.1016/j.bbamcr.2014.02.004 (2014).
214 Ivanov, E. L. & Haber, J. E. DNA repair: RAD alert. *Current Biology* **7**, R492-R495 (1997).

215 Kurasz, J. E. *et al.* Genotoxic, Metabolic, and Oxidative Stresses Regulate the RNA Repair Operon of Salmonella enterica Serovar Typhimurium. *Journal of bacteriology* **200**, e00476-00418, doi:10.1128/jb.00476-18 (2018).
216 Haas, H. Iron–a key nexus in the virulence of Aspergillus fumigatus. *Frontiers in microbiology* **3**, 28 (2012).
217 Ho, T. D. & Ellermeier, C. D. Ferric Uptake Regulator Fur control of Putative Iron Acquisition Systems in Clostridium difficile. *Journal of bacteriology*, JB. 00098-00015 (2015).
218 Ageno, M., Dore, E. & Frontali, C. The alkaline denaturation of DNA. *Biophysical Journal* **9**, 1281 (1969).
219 Kotrba, P., Inui, M. & Yukawa, H. Bacterial phosphotransferase system (PTS) in carbohydrate uptake and control of carbon metabolism. *Journal of Bioscience and Bioengineering* **92**, 502-517 (2001).
220 Pavelka, M. S. & Jacobs, W. R. Biosynthesis of diaminopimelate, the precursor of lysine and a component of peptidoglycan, is an essential function of Mycobacterium smegmatis. *Journal of bacteriology* **178**, 6496-6507, doi:DOI 10.1128/jb.178.22.6496-6507.1996 (1996).
221 Ran, S., Liu, B., Jiang, W., Sun, Z. & Liang, J. Transcriptome analysis of Enterococcus faecalis in response to alkaline stress. *Frontiers in microbiology* **6**, 795, doi:10.3389/fmicb.2015.00795 (2015).
222 Navarre, W. W. & Schneewind, O. Proteolytic cleavage and cell wall anchoring at the LPXTG motif of surface proteins in Gram-positive bacteria. *Molecular microbiology* **14**, 115-121 (1994).
223 Sengupta, R. *et al.* The role of cell surface architecture of lactobacilli in host-microbe interactions in the gastrointestinal tract. *Mediators of inflammation* **2013** (2013).
224 Boekhorst, J., Wels, M., Kleerebezem, M. & Siezen, R. J. The predicted secretome of Lactobacillus plantarum WCFS1 sheds light on interactions with its environment. *Microbiology* **152**, 3175-3183 (2006).
225 Brinster, S., Furlan, S. & Serror, P. C-terminal WxL domain mediates cell wall binding in Enterococcus faecalis and other gram-positive bacteria. *Journal of bacteriology* **189**, 1244-1253 (2007).
226 Campbell, P. M., Cao, A. T., Hines, E. R., East, P. D. & Gordon, K. H. Proteomic analysis of the peritrophic matrix from the gut of the caterpillar, Helicoverpa armigera. *Insect biochemistry and molecular biology* **38**, 950-958, doi:10.1016/j.ibmb.2008.07.009 (2008).
227 Tran, H. T., Barnich, N. & Mizoguchi, E. Potential role of chitinases and chitin-binding proteins in host-microbial interactions during the development of intestinal inflammation. *Histology and histopathology* **26**, 1453 (2011).
228 Vermassen, A. *et al.* Cell wall hydrolases in bacteria: insight on the diversity of cell wall amidases, glycosidases and peptidases towards peptidoglycan. **10**, 331 (2019).
229 Paes, M. C., Oliveira, M. B. & Oliveira, P. L. Hydrogen peroxide detoxification in the midgut of the blood-sucking insect, Rhodnius prolixus. *Archives of insect biochemistry and physiology* **48**, 63-71, doi:10.1002/arch.1058 (2001).
230 Nyström, T. & Neidhardt, F. C. Cloning, mapping and nucleotide sequencing of a gene encoding a universal stress protein in Eschericha coli. *Molecular microbiology* **6**, 3187-3198 (1992).
231 Kim, H., Goo, E., Kang, Y., Kim, J. & Hwang, I. Regulation of universal stress protein genes by quorum sensing and RpoS in Burkholderia glumae. *Journal of bacteriology* **194**, 982-992 (2012).
232 Cvitkovitch, D. G., Li, Y.-H. & Ellen, R. P. Quorum sensing and biofilm formation in Streptococcal infections. *The Journal of clinical investigation* **112**, 1626-1632 (2003).

233 Porcheron, G. & Dozois, C. M. Interplay between iron homeostasis and virulence: Fur and RyhB as major regulators of bacterial pathogenicity. *Veterinary microbiology* **179**, 2-14 (2015).
234 Balaji, S. & Krishnan, M. V. In Silico analysis of Alkaline Shock proteins in Enterobacteria. *Journal of Proteomics & Bioinformatics* **2**, 21-37 (2008).
235 Kuroda, M., Ohta, T. & Hayashi, H. Isolation and the gene cloning of an alkaline shock protein in methicillin-resistant Staphylococcus aureus. *Biochemical and biophysical research communications* **207**, 978-984 (1995).
236 Wei, J. & Jingping, L. Transcriptome analysis of Enterococcus faecalis in response to alkaline stress. *Frontiers in microbiology* **6**, 795 (2015).
237 Johnson, K. S. & Barbehenn, R. V. Oxygen levels in the gut lumens of herbivorous insects. *Journal of Insect Physiology* **46**, 897-903 (2000).
238 Vogel-Scheel, J., Alpert, C., Engst, W., Loh, G. & Blaut, M. Requirement of purine and pyrimidine synthesis for colonization of the mouse intestine by Escherichia coli. *Appl Environ Microbiol* **76**, 5181-5187, doi:10.1128/AEM.00242-10 (2010).
239 Russell, C. W. *et al.* Matching the supply of bacterial nutrients to the nutritional demand of the animal host. *Proceedings of the Royal Society B: Biological Sciences* **281**, 20141163 (2014).
240 Pavelka, M. & Jacobs, W. R. Biosynthesis of diaminopimelate, the precursor of lysine and a component of peptidoglycan, is an essential function of Mycobacterium smegmatis. *Journal of bacteriology* **178**, 6496-6507 (1996).
241 Powell, J. E., Leonard, S. P., Kwong, W. K., Engel, P. & Moran, N. A. Genome-wide screen identifies host colonization determinants in a bacterial gut symbiont. *Proceedings of the National Academy of Sciences* **113**, 13887-13892 (2016).
242 Zhang, J., Zhang, Y., Li, J., Liu, M. & Liu, Z. Midgut transcriptome of the cockroach Periplaneta americana and its microbiota: digestion, detoxification and oxidative stress response. *PloS one* **11** (2016).
243 Tokuda, G. & Watanabe, H. Hidden cellulases in termites: revision of an old hypothesis. *Biology letters* **3**, 336-339 (2007).
244 Sund, C. J. *et al.* The Bacteroides fragilis transcriptome response to oxygen and H2O2: the role of OxyR and its effect on survival and virulence. *Molecular microbiology* **67**, 129-142 (2008).
245 Maffei, M., Bossi, S., Spiteller, D., Mithofer, A. & Boland, W. Effects of feeding Spodoptera littoralis on lima bean leaves. I. Membrane potentials, intracellular calcium variations, oral secretions, and regurgitate components. *Plant physiology* **134**, 1752-1762, doi:10.1104/pp.103.034165 (2004).
246 Trapnell, C. *et al.* Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature protocols* **7**, 562-578, doi:10.1038/nprot.2012.016 (2012).
247 Yu, G., Wang, L.-G., Han, Y. & He, Q.-Y. J. O. a. j. o. i. b. clusterProfiler: an R package for comparing biological themes among gene clusters. **16**, 284-287 (2012).
248 Boyle, E. I. *et al.* GO::TermFinder—open source software for accessing Gene Ontology information and finding significantly enriched Gene Ontology terms associated with a list of genes. *Bioinformatics* **20**, 3710-3715, doi:10.1093/bioinformatics/bth456 %J Bioinformatics (2004).
249 Mazumdar, T. *et al.* Survival strategies of Enterococcus mundtii in the gut of Spodoptera littoralis: a live report. *bioRxiv*, 2020.2002.2003.932053, doi:10.1101/2020.02.03.932053 (2020).
250 Holmes, E., Li, J. V., Marchesi, J. R. & Nicholson, J. K. J. C. m. Gut microbiota composition and activity in relation to host metabolic phenotype and disease risk. **16**, 559-564 (2012).

251 McCutcheon, J. P., Moran, N. A. J. G. b. & evolution. Functional convergence in reduced genomes of bacterial symbionts spanning 200 My of evolution. **2**, 708-718 (2010).
252 Blekhman, R. *et al.* Host genetic variation impacts microbiome composition across human body sites. **16**, 191 (2015).
253 Wagner, M. R. *et al.* Host genotype and age shape the leaf and root microbiomes of a wild perennial plant. **7**, 1-15 (2016).
254 Spor, A., Koren, O. & Ley, R. J. N. R. M. Unravelling the effects of the environment and host genotype on the gut microbiome. **9**, 279-290 (2011).
255 Khachatryan, Z. A. *et al.* Predominant role of host genetics in controlling the composition of gut microbiota. **3** (2008).
256 Dąbrowska, K. & Witkiewicz, W. J. F. i. m. Correlations of host genetics and gut microbiome composition. **7**, 1357 (2016).
257 Douglas, A. E. The molecular basis of bacterial-insect symbiosis. *J Mol Biol* **426**, 3830-3837, doi:10.1016/j.jmb.2014.04.005 (2014).
258 Cirimotich, C. M., Ramirez, J. L., Dimopoulos, G. J. C. h. & microbe. Native microbiota shape insect vector competence for human pathogens. **10**, 307-310 (2011).
259 Oliveira, J. H. M. *et al.* Blood meal-derived heme decreases ROS levels in the midgut of Aedes aegypti and allows proliferation of intestinal microbiota. **7** (2011).
260 Mossialos, D. *et al.* Quinolobactin, a New Siderophore ofPseudomonas fluorescens ATCC 17400, the Production of Which Is Repressed by the Cognate Pyoverdine. **66**, 487-492 (2000).
261 Schennen, A. & Hölzl, J. J. P. m. 6-Hydroxykynurensäure, die erste N-haltige Verbindung aus den Blättern von Ginkgo biloba. **52**, 235-236 (1986).
262 Gama, S. *et al.* Thermodynamic study on 8-hydroxyquinoline-2-carboxylic acid as a chelating agent for iron found in the gut of Noctuid larvae. **42**, 8062-8073 (2018).
263 Pesek, J. *et al.* Biosynthesis of 8-hydroxyquinoline-2-carboxylic acid, an iron chelator from the gut of the lepidopteran Spodoptera littoralis. **13**, 178-184 (2015).
264 Parmanand, B. A. *et al.* A decrease in iron availability to human gut microbiome reduces the growth of potentially pathogenic gut bacteria; an in vitro colonic fermentation study. **67**, 20-27 (2019).
265 Reed, S., Neuman, H., Glahn, R. P., Koren, O. & Tako, E. J. P. O. Characterizing the gut (Gallus gallus) microbiota following the consumption of an iron biofortified Rwandan cream seeded carioca (Phaseolus Vulgaris L.) bean-based diet. **12** (2017).
266 Roth, T. L. *et al.* Reduced Gut Microbiome Diversity and Metabolome Differences in Rhinoceros Species at Risk for Iron Overload Disorder. **10**, 2291 (2019).
267 Jiang, T., Xing, B., Rao, J. J. B. & Reviews, G. E. Recent developments of biological reporter technology for detecting gene expression. **25**, 41-76 (2008).
268 Rajamanickam, C. & Jeejabai, R. Evaluation of the cardiac isoform of alpha2-macroglobulin as a factor inducing cardiac hypertrophy. *Methods in molecular medicine* **112**, 261-275, doi:10.1385/1-59259-879-x:261 (2005).
269 Schenborn, E. & Groskreutz, D. Reporter gene vectors and assays. *Molecular Biotechnology* **13**, 29-44, doi:10.1385/MB:13:1:29 (1999).
270 Leahy, M., Vaughan, P., Fanning, L., Fanning, S. & Sheehan, D. Purification and some characteristics of a recombinant dimeric rhizobium meliloti beta-galactosidase expressed in escherichia coli. *Enzyme and microbial technology* **28**, 682-688, doi:10.1016/s0141-0229(01)00314-3 (2001).
271 Misteli, T. & Spector, D. L. Applications of the green fluorescent protein in cell biology and biotechnology. *Nature biotechnology* **15**, 961-964, doi:10.1038/nbt1097-961 (1997).
272 Props, R. *et al.* Measuring the biodiversity of microbial communities by flow cytometry. **7**, 1376-1385 (2016).

273 Frossard, A., Hammes, F. & Gessner, M. O. J. F. i. m. Flow cytometric assessment of bacterial abundance in soils, sediments and sludge. **7**, 903 (2016).
274 Ambriz-Aviña, V., Contreras-Garduño, J. A. & Pedraza-Reyes, M. J. B. r. i. Applications of flow cytometry to characterize bacterial physiological responses. **2014** (2014).
275 Steen, H. B., Jernaes, M. W., Skarstad, K. & Boye, E. in *Methods in cell biology* Vol. 42 477-487 (Elsevier, 1994).
276 Gandhi, M. & Chikindas, M. L. J. I. j. o. f. m. Listeria: a foodborne pathogen that knows how to survive. **113**, 1-15 (2007).
277 Pamp, S. J., Sternberg, C. & Tolker-Nielsen, T. J. C. P. A. T. J. o. t. I. S. f. A. C. Insight into the microbial multicellular lifestyle via flow-cell technology and confocal microscopy. **75**, 90-103 (2009).
278 Henke, J. M. & Bassler, B. L. J. J. o. b. Quorum sensing regulates type III secretion in Vibrio harveyi and Vibrio parahaemolyticus. **186**, 3794-3805 (2004).
279 Stapleton, A. E. *et al.* Precise and rapid assessment of Escherichia coli adherence to vaginal epithelial cells by flow cytometry. **50**, 31-37 (2002).
280 Wei, J. & Jingping, L. J. F. i. m. Transcriptome analysis of Enterococcus faecalis in response to alkaline stress. **6**, 795 (2015).
281 Elowitz, M. B., Levine, A. J., Siggia, E. D. & Swain, P. S. J. S. Stochastic gene expression in a single cell. **297**, 1183-1186 (2002).
282 Davis, K. M. & Isberg, R. R. J. B. Defining heterogeneity within bacterial populations via single cell approaches. **38**, 782-790 (2016).
283 Org, E. *et al.* Genetic and environmental control of host-gut microbiota interactions. **25**, 1558-1569 (2015).
284 Henao-Mejia, J. *et al.* Inflammasome-mediated dysbiosis regulates progression of NAFLD and obesity. **482**, 179-185 (2012).
285 Suzuki, N. *et al.* Severe impairment of interleukin-1 and Toll-like receptor signalling in mice lacking IRAK-4. **416**, 750-754 (2002).
286 Wittmann, I. *et al.* Murine bactericidal/permeability-increasing protein inhibits the endotoxic activity of lipopolysaccharide and gram-negative bacteria. **180**, 7546-7552 (2008).
287 Goodrich, J. K. *et al.* Human genetics shape the gut microbiome. **159**, 789-799 (2014).
288 Fan, P. *et al.* Host genetic effects upon the early gut microbiota in a bovine model with graduated spectrum of genetic variation. **14**, 302-317 (2020).
289 Chaston, J. M., Dobson, A. J., Newell, P. D. & Douglas, A. E. J. A. E. M. Host genetic control of the microbiota mediates the Drosophila nutritional phenotype. **82**, 671-679 (2016).
290 Dostal, A. *et al.* Iron depletion and repletion with ferrous sulfate or electrolytic iron modifies the composition and metabolic activity of the gut microbiota in rats. **142**, 271-277 (2012).
291 Zimmermann, M. B. *et al.* The effects of iron fortification on the gut microbiota in African children: a randomized controlled trial in Cote d'Ivoire. **92**, 1406-1415 (2010).
292 Johnson, J. S. *et al.* Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. **10**, 1-11 (2019).

11 Supplementary Material

S1: **Flow cytometry**

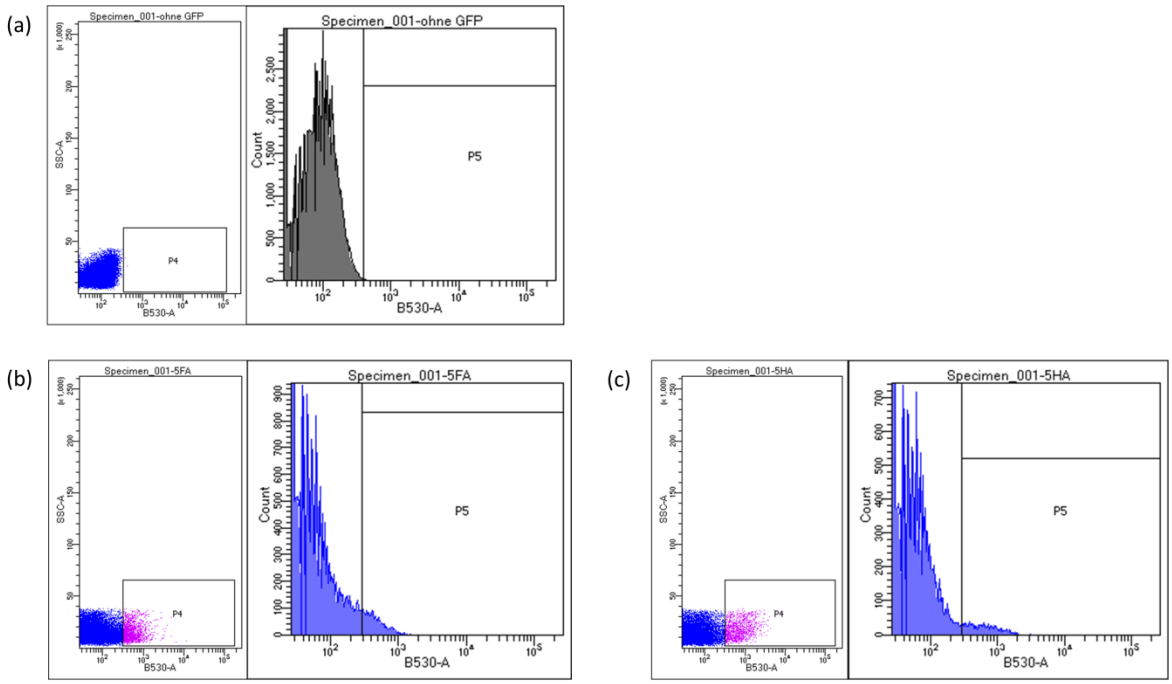

Figure S1: Sorting profiles of *Enterococcus mundtii*-pTRKH3 from (a) control (without GFP), (b) foregut homogenate and (c) hindgut homogenate. P4 and P5 correspond to the density of fluorescent cells detected by the scatters, and the cells that are actually sorted, respectively. The purple dots correspond to the GFP-containing *Enterococcus mundtii* that was sorted using the flow cytometer.

S2: Table showing the mapping percentages of the reads obtained from *Enterococcus mundtii* living in the foregut, hind gut of *Spodoptera littoralis* larvae, and the same bacteria grown *in vitro.* The mapping was done against the whole genome of *Enterococcus mundtii* QU25

Sample	Replicates	# of reads from Trimmomatic; after adapter trimming	% alignment by Tophat
Foregut	F1	13949493	17.55
	F2	10562548	58.80
	F3	10623050	73.40
Hindgut	H1	9491522	27.25
	H2	9161970	47.10
	H3	10187317	55.50
Control	C1	9386723	41.55
	C2	9854743	38.70
	C3	9733545	48.25

S3: Distribution of the transcriptomic profiles of *E. mundtii* obtained from the foregut (FG), hindgut (HG) of *S. littoralis* and the ones growing *in vitro* control conditions (FL)

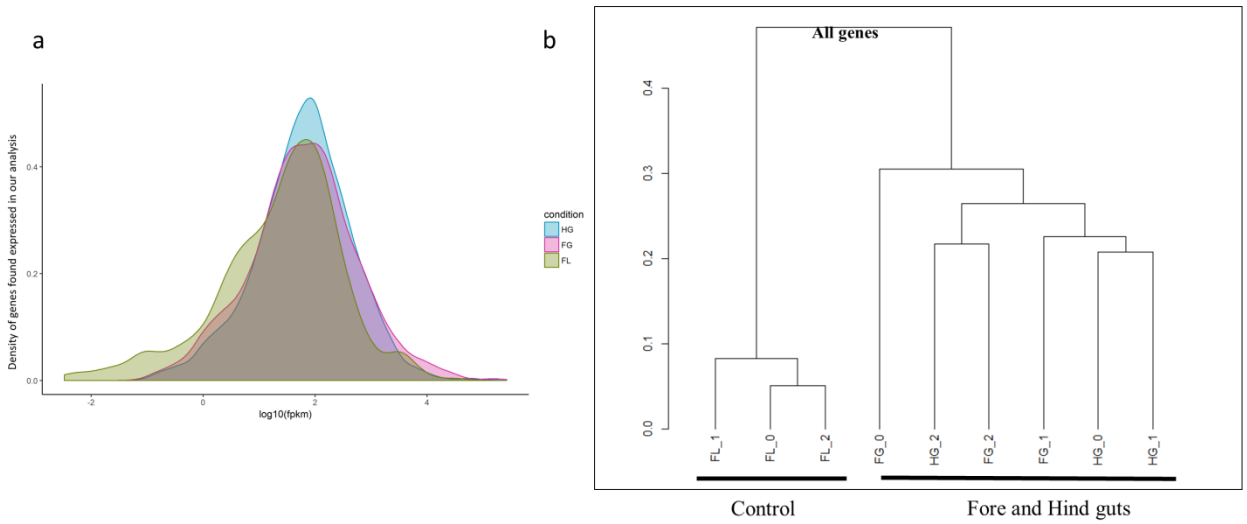

Figure S3: (a) Density plot showing the distribution of the differentially expressed genes in foregut (FG), hindgut (HG) or control (FL)

(b) **:** Dendrogram showing hierarchical clustering among the three replicates of transcriptomes analyzed from *E. mundtii* obtained from the foregut, hindgut and control. The gene expression profiles from the foreguts (FGs) and the hindguts (HGs) partially overlap and cluster away from the control (FL) profile

S4: **Annotations based on KEGG database**

Annotations using the Kegg database show a similar trend as the gene ontology category of Biological Processes. EnrichKEGG () function of the clusterprofiler R package performed over-representation test and annotated 938 genes of *E. mundtii* in the 68 and 72 genes (7.2% and 7.6%, respectively) are upregulated in *E. mundtii* when they are living in the fore and hindguts respectively. The same for downregulated genes are 63 and 57 (6.7% and 6%) respectively. The ones that are significantly enriched are shown in figure (S4).

Degradation of aromatic compounds by the *E. mundtii* could be occurring due to the lignin content in the plant diet. The enrichment of starch and sugar metabolism remains similar.

Enrichment in napthalenedegradation and chloroalkane and chloroalkene degradation mirror the high activity of alcohol dehydrogenases, which carries out steps in bacterial glycolysis and fermentation.

There are marked downregulations in several amino acid and fatty acid biosynthetic pathways, biosynthesis of antibiotics, propionate metabolism and biosynthesis of secondary metabolites; although the hindgut dwelling *E. mundtii* seem to biosynthesize lysine. It is quite

possible that the lysine produced in the hindgut gets reabsorbed by the larval system through their hindgut intestinal cells.

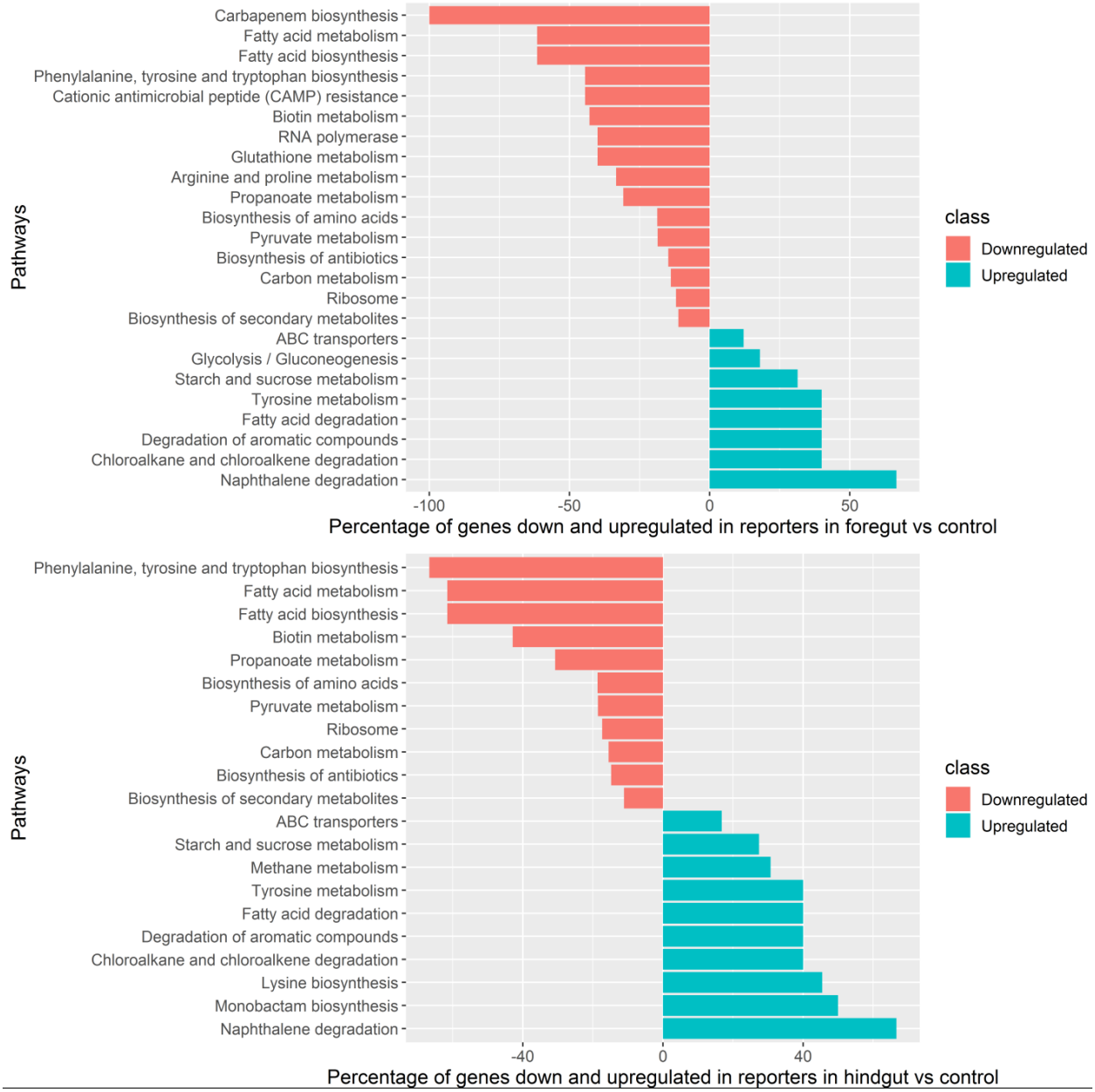

Figure S5: Kegg Orthology classification of assembled unigenes, annotated from transcriptomic information given by *E. mundtii* obtained from fore (a) and hindguts (b) respectively. The graphs show both upand down- regulationof the assembled genes, as compared to control. The percentages of each pathway refers to the percentage of genes of that particular pathway that are enriched in *E. mundtii*

S5: Enriched pathways of *E. mundtii* living in the hindgut as compared to foregut

When the gene expression of *Enterococcus mundtii* living in two regions of the gut (fore and hind) was compared, the only important enriched pathway belongs to lysine biosynthesis, which is seen to be upregulated in the hindgut as compared to the foregut, according to the Biological Process category of gene ontology (Fig. S5)

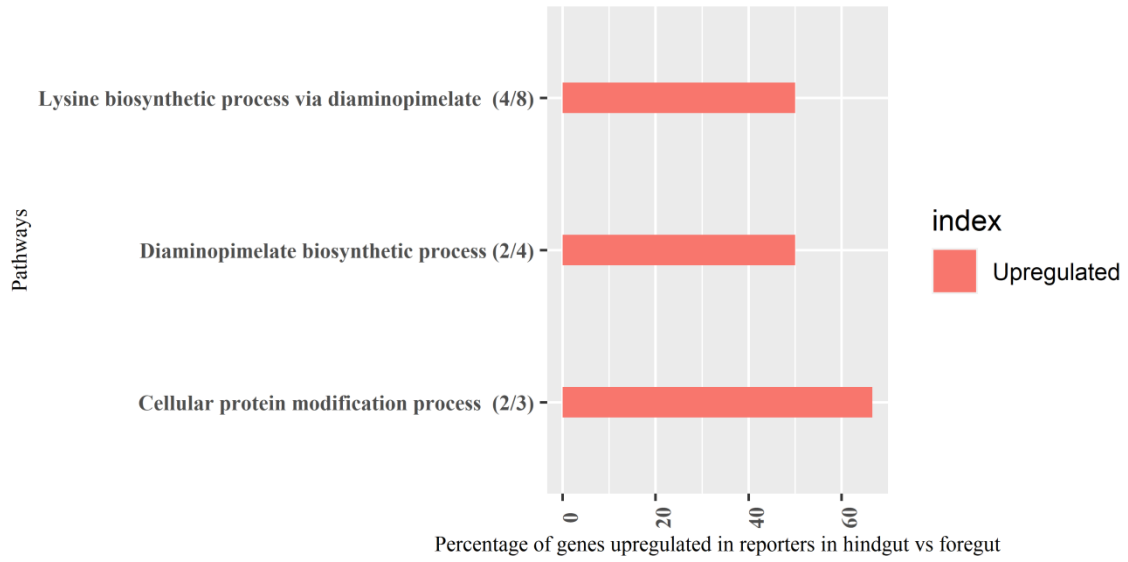

Figure S5: Summary of gene ontology classification in the category of biological processes. The graph shows both up- and downregulation of the assembled genes of *E. mundtii* obtained from hindgut as compared to the ones from foregut. The percentages of each pathway refer to the percentage of genes of that particular pathway that are enriched in *E. mundtii.*

12 Appendix

12.1 Appendix 1:

Some selected genes that are differentially regulated when *Enterococcus mundtii* is dwelling in the fore-and hindguts of *Spodoptera littoralis* vs. control

	E. mundtii **living in: Foregut vs Control**									
	current_genes@annotation$gene_short_name	**gene_id**	**sample_1**	**sample_2**	**value_1**	**value_2**		**log2_fold_**	**p_value**	**q_value**	**significant**
Extracellular interaction	wxl domain containing proteins										
	EMQU_0043	XLOC_000	FG	Control	8.94313	13.8503	0.645699	-0.63107	0.2984	0.554407	no
	EMQU_0485	XLOC_000	FG	Control	7.51487	1.21054	6.207866	2.6341	0.03575	0.153679	no
	EMQU_0487	XLOC_000	FG	Control	22.8062	23.3514	0.976652	-0.03408	0.9516	0.974827	no
	EMQU_0539	XLOC_000	FG	Control	7.03728	0.018348	383.5385	8.58323	0.06	0.215365	no
	EMQU_0541	XLOC_000	FG	Control	23.112	0.768304	30.08184	4.91082	0.00005	0.001175	yes
	lpxtg motif containing proteins										
	fms3	XLOC_000	FG	Control	22.9629	15.079	1.52284	0.60677	0.3373	0.589299	no
	EMQU_1297	XLOC_001	FG	Control	73.7189	2.22187	33.17876	5.05219	0.4396	0.68306	no
	lysM, biofilm forming protein										
	EMQU_0517	XLOC_000	FG	Control	257.543	85.74	3.003767	1.58677	0.0045	0.037119	yes
	EMQU_1080	XLOC_000	FG	Control	1468.54	518.8	2.830648	1.50114	0.0164	0.089816	no
	Accessory gene regulator: two-component system										
	agrB	XLOC_001	FG	Control	180974	34990.9	5.17203	2.37073	0.0307	0.138327	no
	EMQU_1481 agrA	XLOC_001	FG	Control	7.91384	2.07278	3.817984	1.93281	0.0869	0.272407	no
	Cell surface anchor proteins										
	EMQU_0445,EMQU_0446,ebpCfm	XLOC_000	FG	Control	4.86095	7.60765	0.638956	-0.64621	0.55865	0.769515	no
	EMQU_0540	XLOC_000	FG	Control	2.45222	0.018742	130.8409	7.03167	1	1	no
	EMQU_1048,ftsW3	XLOC_000	FG	Control	540.559	53.618	10.08167	3.33366	0.00005	0.001175	yes
	Chitin binding proteins										
	EMQU_0940	XLOC_000	FG	Control	163.522	3.46965	47.12925	5.55855	0.00235	0.023188	yes
	EMQU_1285	XLOC_000	FG	Control	272.066	10.6683	25.50228	4.67256	0.00005	0.001175	yes
Stress survival strategies	**oxidative stress**										
	EMQU_0929 sod	XLOC_001	FG	Control	946.264	73.8632	12.81103	3.67931	0.05655	0.206952	no
	EMQU_0568 catalase	XLOC_001	FG	Control	163.414	45.7908	3.568708	1.8354	0.0025	0.024208	yes
	EMQU_0459 NADH peroxidase	XLOC_000	FG	Control	148.947	111.411	1.336915	0.418914	0.4206	0.668188	no
	EMQU_1279 NADH Oxidase	XLOC_000	FG	Control	818.533	565.966	1.446258	0.532324	0.30245	0.558232	no
	EMQU_1453 ohr	XLOC_000	FG	Control	1364.84	234.808	5.812579	2.53918	0.0016	0.017484	yes
	EMQU_0165	XLOC_001	FG	Control	75.9495	38.8089	1.957012	0.968651	0.1688	0.407999	no
	stress proteins										
	glsB	XLOC_001	FG	Control	773.485	23.9101	32.34972	5.01568	0.003	0.027298	yes
	glsB1	XLOC_001	FG	Control	7635.66	740.723	10.30839	3.36575	0.00005	0.001175	yes
	gls33	XLOC_001	FG	Control	77.1609	12.6029	6.122472	2.61412	0.0026	0.024467	yes
	uspA2	XLOC_000	FG	Control	3942.33	72.4201	54.43696	5.76651	0.00005	0.001175	yes
	rpmG2,secE	XLOC_002	FG	Control	1293.71	59.3845	21.78531	4.44528	0.04355	0.174536	no
	EMQU_1288	XLOC_000	FG	Control	38.6307	0.81305	47.51331	5.57026	0.00465	0.037955	yes
	Repair proteins										
	EMQU_2803	XLOC_002	FG	Control	36.9518	38.3836	0.962698	-0.05485	0.93955	0.968756	no
	recA	XLOC_002	FG	Control	2029.41	722.416	2.809199	1.49016	0.015	0.085347	no
	recF,yaaA	XLOC_000	FG	Control	974.208	389.591	2.500592	1.32227	0.16515	0.40266	no
	alkD	XLOC_000	FG	Control	181.389	64.5425	2.810381	1.49077	0.01925	0.099925	no
	EMQU_2770 radA	XLOC_002	FG	Control	222.489	75.9783	2.928323	1.55007	0.0024	0.023483	yes
	EMQU_0695,EMQU_0697,EMQU_0698,EMQU_0699,EMQU_0700,valS (radC)	XLOC_000	FG	Control	14127.1	4588.23	3.078987	1.62246	0.52985	0.75026	no
	EMQU_3001 dnaJ,EMQU_3002 YafQ	XLOC_002	FG	Control	18351.1	559.087	32.82334	5.03665	0.00005	0.001175	yes
	Alkaline stress										
	asp,nusB	XLOC_000	FG	Control	6009.24	1141.25	5.26549	2.39657	0.0016	0.017484	yes
	EMQU_2152 nhaC	XLOC_001	FG	Control	19.2674	41.5033	0.464238	-1.10706	0.109	0.314687	no
Metabolism	EMQU_0667 adenosine deaminase	XLOC_000	FG	Control	285.117	14.5191	19.63737	4.29554	0.00005	0.001175	yes
	EMQU_2472 glucosamine-6-phosphate deaminase	XLOC_002	FG	Control	1806.1	84.4369	21.38994	4.41886	0.00005	0.001175	yes
	metC	XLOC_001	FG	Control	2.60686	58.5393	0.044532	-4.48902	0.0006	0.008529	yes
	6-phospho-beta-glucosidases										
	bglB	XLOC_000	FG	Control	2.94125	4.03119	0.729623	-0.45478	0.63355	0.822402	no
	bglH	XLOC_000	FG	Control	16.5532	0.993126	16.66777	4.05899	0.0056	0.043007	yes
	bglG2	XLOC_000	FG	Control	17.3353	15.5342	1.115944	0.158264	0.81695	0.913648	no
	bglF	XLOC_000	FG	Control	9.17342	23.6807	0.38738	-1.36818	0.02545	0.120842	no
	pfkA phosphofurctokinase A	XLOC_000	FG	Control	830.107	1043.88	0.795213	-0.33059	0.5029	0.733644	no
	ldhA lactose dehydrogenase	XLOC_001	FG	Control	41.7702	10.7938	3.869833	1.95227	0.0147	0.084774	no
	sfsA sugar fermentation stimulation protein	XLOC_000	FG	Control	627.444	71.2775	8.802834	3.13797	0.00005	0.001175	yes
	alcohol dehydrogenases										
	fdh	XLOC_002	FG	Control	118.885	33.7184	3.52582	1.81796	0.0016	0.017484	yes
	EMQU_0525,EMQU_0526	XLOC_000	FG	Control	455.45	100.681	4.523694	2.1775	0.0045	0.037119	yes
	EMQU_1129	XLOC_000	FG	Control	32.8955	19.2005	1.713263	0.776746	0.2463	0.505018	no
	EMQU_0450,EMQU_0452,dapB,dapF (lysine biosynthesis)	XLOC_000	FG	Control	63.2449	46.3105	1.365671	0.449611	0.5162	0.741362	no
	EMQU_0175,EMQU_0176,EMQU_0177,EMQU_0178 (lysine transport)	XLOC_001	FG	Control	1177.89	565.646	2.08238	1.05824	0.4756	0.711119	no

Phospho Transferase Systems

EMQU_0169 trehalose-specific IIA component	XLOC_000:FG	Control	568.571	439.437	1.293862	0.371683	0.48035	0.71493	no
EMQU_0285 PTS family porter component IIA	XLOC_000:FG	Control	3.80208	2.70527	1.405435	0.491014	0.5166	0.741362	no
EMQU_0286 ascorbate-specific IIB component	XLOC_000:FG	Control	9.35016	4.67965	1.998047	0.998591	0.60405	0.804196	no
EMQU_0287 PTS system ascorbate-specific transporter subunit IIC	XLOC_000:FG	Control	5.83015	2.94888	1.977073	0.983367	0.3209	0.577118	no
EMQU_0307 cellobiose-specific IIB component	XLOC_000:FG	Control	356.938	183.373	1.946513	0.960894	0.1086	0.313726	no
EMQU_0308 cellobiose-specific IIA component	XLOC_000:FG	Control	644.87	89.5451	7.201622	2.84832	0.00015	0.003064	yes
EMQU_0309 cellobiose-specific IIC component	XLOC_000:FG	Control	282.52	138.844	2.034802	1.02489	0.05615	0.206457	no
EMQU_0390 sugar-specific IIA component	XLOC_000:FG	Control	251.418	5.39258	46.62295	5.54297	0.00005	0.001175	yes
EMQU_0463,EMQU_0464 fructose-specific IIB component	XLOC_000:FG	Control	1.96445	0.195792	10.03335	3.32673	1	1	no
EMQU_0465, fructose-specific IIA component	XLOC_000:FG	Control	6.71826	0.889271	7.554795	2.91739	0.2735	0.525538	no
EMQU_0470 mannose-specific IIB component	XLOC_000:FG	Control	26.9057	3.65529	7.360757	2.87985	0.03475	0.151046	no
EMQU_0471,EMQU_0472 mannose-specific IIC, II D component	XLOC_000:FG	Control	20.4329	4.99818	4.088068	2.03142	0.04515	0.179259	no
EMQU_0657 cellobiose-specific IIB component	XLOC_000:FG	Control	3.313	0.474112	6.9878	2.80484	1	1	no
EMQU_0720 cellobiose-specific IIC component	XLOC_001:FG	Control	8.57765	6.03448	1.42144	0.507354	0.50875	0.737806	no
EMQU_0875 cellobiose-specific IIB component	XLOC_000!FG	Control	686.39	115.917	5.921392	2.56594	0.00065	0.008968	yes
EMQU_0876 cellobiose-specific IIA component	XLOC_000!FG	Control	7454.4	95.3937	78.14353	6.28805	0.00005	0.001175	yes
EMQU_0877cellobiose-specific IIC component	XLOC_000!FG	Control	1142.83	60.5217	18.88298	4.23901	0.00005	0.001175	yes
EMQU_1390 cellobiose-specific IIC component	XLOC_001:FG	Control	1.90791	1.88248	1.013509	0.019361	1	1	no
EMQU_2136 sucrose-specific IIA component	XLOC_001!FG	Control	12.4744	5.65732	2.205002	1.14077	0.1206	0.331402	no
EMQU_2183 cellobiose-specific IIC component	XLOC_001!FG	Control	26.5973	5.97553	4.451036	2.15414	0.0068	0.049385	yes
EMQU_2184,EMQU_2185,EMQU_2186 beta-glucoside-specific IIA component	XLOC_001!FG	Control	69.8086	120.015	0.581666	-0.78174	0.30935	0.564575	no
mtlF mannitol-specific IIA component	XLOC_002!FG	Control	10.6931	6.63801	1.610889	0.687854	0.68255	0.852248	no
EMQU_2415 cellobiose-specific IIC component	XLOC_002!FG	Control	9.68388	2.91282	3.324572	1.73317	0.3361	0.588168	no
EMQU_2463,EMQU_2464 mannose-specific IIC, IID component	XLOC_002!FG	Control	43.0103	5.85687	7.343564	2.87648	0.0063	0.046548	yes
EMQU_2465 mannose-specific IIA component	XLOC_002!FG	Control	98.6587	2.15572	45.76601	5.51621	0.1644	0.40209	no
EMQU_2571 ascorbate-specific IIC component	XLOC_002:FG	Control	18.4545	7.5075	2.458142	1.29757	0.08615	0.270782	no
EMQU_2572,EMQU_2573 ascorbate-specific IIAB component	XLOC_002:FG	Control	9.10346	8.51933	1.068565	0.095675	0.9117	0.958302	no
EMQU_2609 sucrose-specific IIC component	XLOC_002:FG	Control	93.3782	21.2684	4.390467	2.13437	0.0002	0.003772	yes
EMQU_2677 cellobiose-specific IIC component	XLOC_002:FG	Control	19.8547	0.105524	188.1534	7.55577	0.21425	0.47155	no
celB PTS system cellobiose transporter subunit IIB	XLOC_002:FG	Control	219.915	76.572	2.872003	1.52206	0.02125	0.106523	no
EMQU_2816 fructose-specific IIB component	XLOC_002:FG	Control	3.58546	2.32851	1.539809	0.622746	0.69335	0.858647	no
EMQU_2817 fructose-specific IIA component	XLOC_002:FG	Control	42.1034	2.67551	15.73659	3.97605	0.0114	0.070807	no
EMQU_2819,EMQU_2820nitrogen regulatory IIA component	XLOC_002:FG	Control	10.0406	3.4275	2.929424	1.55061	0.47735	0.712142	no

	E. mundtii living in: Hindgut vs Control									
	current_genes@annotation$gene_short_name	gene_id	sample_1	sample_2	value_1	value_2	fold chang	log2_fold_	p_value	q_value	significant
Extracellular interaction	wxl domain containing proteins										
	EMQU_0043	XLOC_000	HG	Control	15.3819	13.8503	1.110582	0.151316	0.7891	0.901228	no
	EMQU_0485	XLOC_000	HG	Control	2.12834	1.21054	1.758174	0.814084	1	1	no
	EMQU_0487	XLOC_000	HG	Control	6.0516	23.3514	0.259154	-1.95	0.00985	0.063281	no
	EMQU_0539	XLOC_000	HG	Control	0.847213	0.018348	46.17392	5.52901	1	1	no
	EMQU_0541	XLOC_000	HG	Control	6.43246	0.768304	8.372285	3.06562	0.00025	0.004446	yes
	lpxtg motif containing proteins										
	fms3	XLOC_000	HG	Control	17.234	15.079	1.142914	0.192725	0.78315	0.899065	no
	EMQU_1297	XLOC_001	HG	Control	276.842	2.22187	124.5986	6.96115	0.3246	0.579274	no
	lysM, biofilm forming protein										
	EMQU_0517	XLOC_000	HG	Control	264.527	85.74	3.085223	1.62537	0.0017	0.018405	yes
	EMQU_1080	XLOC_000	HG	Control	962.047	518.8	1.85437	0.89093	0.09995	0.29737	no
	Accessory gene regulator: two-component system										
	agrB	XLOC_001	HG	Control	265058	34990.9	7.575055	2.92125	0.0029	0.026595	yes
	EMQU_1481 agrA	XLOC_001	HG	Control	8.24085	2.07278	3.975748	1.99123	0.11025	0.316487	no
	Cell surface anchor proteins										
	EMQU_0445,EMQU_0446,ebpCfm	XLOC_000	HG	Control	23.2154	7.60765	3.051586	1.60956	0.033	0.145005	no
	EMQU_0540	XLOC_000	HG	Control	1.89396	0.018742	101.0543	6.65898	1	1	no
	EMQU_1048,ftsW3	XLOC_000	HG	Control	323.47	53.618	6.032862	2.59284	0.00005	0.001175	yes
	Chitin binding proteins										
	EMQU_0940	XLOC_000	HG	Control	479.383	3.46965	138.1647	7.11025	0.0023	0.022985	yes
	EMQU_1285	XLOC_000	HG	Control	736.496	10.6683	69.03593	6.10928	0.00005	0.001175	yes
Stress survival strategies	**oxidative stress**										
	EMQU_0929 sod	XLOC_001	HG	Control	555.04	73.8632	7.514432	2.90966	0.2024	0.456865	no
	EMQU_0568 catalase	XLOC_001	HG	Control	439.995	45.7908	9.608808	3.26436	0.00005	0.001175	yes
	EMQU_0459 NADH peroxidase	XLOC_000	HG	Control	414.128	111.411	3.717119	1.89419	0.0005	0.007544	yes
	EMQU_1279 NADH Oxidase	XLOC_000	HG	Control	1237.25	565.966	2.186085	1.12835	0.03185	0.141465	no
	EMQU_1453 ohr	XLOC_000	HG	Control	1282.44	234.808	5.461654	2.44934	0.00005	0.001175	yes
	EMQU_0165	XLOC_001	HG	Control	118.394	38.8089	3.050692	1.60913	0.0131	0.07795	no
	stress proteins										
	glsB	XLOC_001	HG	Control	2321.49	23.9101	97.09244	6.60129	0.0023	0.022985	yes
	glsB1	XLOC_001	HG	Control	5358.13	740.723	7.233649	2.85472	0.00005	0.001175	yes
	gls33	XLOC_001	HG	Control	271.018	12.6029	21.50442	4.42656	0.00005	0.001175	yes
	uspA2	XLOC_000	HG	Control	834.712	72.4201	11.52597	3.52682	0.00005	0.001175	yes
	rpmG2,secE	XLOC_002	HG	Control	961.897	59.3845	16.19778	4.01772	0.0471	0.184186	no
	EMQU_1288	XLOC_000	HG	Control	37.685	0.81305	46.35016	5.5345	0.00465	0.037955	yes
	Repair proteins										
	EMQU_2803	XLOC_002	HG	Control	50.1773	38.3836	1.307259	0.386545	0.55445	0.767504	no
	recA	XLOC_002	HG	Control	975.438	722.416	1.350244	0.43322	0.38235	0.631494	no
	recF,yaaA	XLOC_000	HG	Control	1206.08	389.591	3.095759	1.6303	0.1044	0.304749	no
	alkD	XLOC_000	HG	Control	196.911	64.5425	3.050873	1.60922	0.00535	0.041496	yes
	EMQU_2770 radA	XLOC_002	HG	Control	207.496	75.9783	2.73099	1.44942	0.0103	0.065187	no
	EMQU_0695,EMQU_0697,EMQU_0698,EMQU_0699,EMC	XLOC_000	HG	Control	27719.6	4588.23	6.041458	2.5949	0.2593	0.51344	no
	EMQU_3001 dnaJ,EMQU_3002 YafQ	XLOC_002	HG	Control	2342.23	559.087	4.189384	2.06674	0.0023	0.022985	yes
	Alkaline stress										
	asp,nusB	XLOC_000	HG	Control	1959.03	1141.25	1.716565	0.779526	0.1671	0.405356	no
	EMQU_2152 nhaC	XLOC_001	HG	Control	57.8409	41.5033	1.393646	0.478863	0.42925	0.674824	no
Metabolism	EMQU_0667 adenosine deaminase	XLOC_000	HG	Control	184.412	14.5191	12.70134	3.66692	0.00005	0.001175	yes
	EMQU_2472 glucosamine-6-phosphate deaminase	XLOC_002	HG	Control	764.978	84.4369	9.059759	3.17947	0.00005	0.001175	yes
	metC	XLOC_001	HG	Control	13.8188	58.5393	0.23606	-2.08277	0.00755	0.052625	no
	6-phospho-beta-glucosidases										
	bglB	XLOC_000	HG	Control	5.53878	4.03119	1.373981	0.458364	0.55325	0.76682	no
	bglH	XLOC_000	HG	Control	42.8138	0.993126	43.11014	5.42995	0.00225	0.022728	yes
	bglG2	XLOC_000	HG	Control	20.4446	15.5342	1.316103	0.396274	0.533	0.751903	no
	bglF	XLOC_000	HG	Control	14.0362	23.6807	0.592727	-0.75456	0.20075	0.453798	no
	pfkA phosphofurctokinase A	XLOC_000	HG	Control	818.513	1043.88	0.784106	-0.35088	0.468	0.704485	no
	ldhA lactose dehydrogenase	XLOC_001	HG	Control	10.5677	10.7938	0.979053	-0.03054	0.97025	0.982221	no
	sfsA sugar fermentation stimulation protein	XLOC_000	HG	Control	418.669	71.2775	5.873789	2.55429	0.00005	0.001175	yes
	alcohol dehydrogenases										
	fdh	XLOC_002	HG	Control	141.433	33.7184	4.194535	2.06851	0.0004	0.006363	yes
	EMQU_0525,EMQU_0526	XLOC_000	HG	Control	508.209	100.681	5.047715	2.33563	0.0003	0.005084	yes
	EMQU_1129	XLOC_000	HG	Control	44.0849	19.2005	2.296029	1.19914	0.07545	0.246597	no
	EMQU_0450,EMQU_0452,dapB,dapF (lysine biosynthesis	XLOC_000	HG	Control	218.8	46.3105	4.72463	2.2402	0.002	0.020787	yes
	EMQU_0175,EMQU_0176,EMQU_0177,EMQU_0178 (lysi	XLOC_001	HG	Control	437.44	565.646	0.773346	0.370814	0.7637	0.891068	no
	Phospho Transferase Systems										
	EMQU_0169 trehalose-specific IIA component	XLOC_000	HG	Control	548.463	439.437	1.248104	0.319737	0.55645	0.768138	no
	EMQU_0285 PTS family porter component IIA	XLOC_000	HG	Control	26.195	2.70527	9.682952	3.27545	0.0003	0.005084	yes
	EMQU_0286 ascorbate-specific IIB component	XLOC_000	HG	Control	39.2145	4.67965	8.379793	0.20854	0.0299	0.136166	no
	EMQU_0287 PTS system ascorbate-specific transporter su	XLOC_000	HG	Control	8.82057	2.94888	2.991159	1.58071	0.21705	0.473808	no
	EMQU_0307 cellobiose-specific IIB component	XLOC_000	HG	Control	1018.06	183.373	5.551853	2.47297	0.00005	0.001175	yes
	EMQU_0308 cellobiose-specific IIA component	XLOC_000	HG	Control	1379.44	89.5451	15.40497	3.94532	0.00005	0.001175	yes
	EMQU_0309 cellobiose-specific IIC component	XLOC_000	HG	Control	799.11	138.844	5.755452	2.52493	0.00005	0.001175	yes
	EMQU_0390 sugar-specific IIA component	XLOC_000	HG	Control	33.6888	5.39258	6.247251	2.64322	0.0001	0.002165	yes
	EMQU_0463,EMQU_0464 fructose-specific IIB componen	XLOC_000	HG	Control	1.09225	0.195792	5.578624	2.47991	1	1	no
	EMQU_0465, fructose-specific IIA component	XLOC_000	HG	Control	3.0058	0.889271	3.380072	1.75705	0.21655	0.473052	no
	EMQU_0470 mannose-specific IIB component	XLOC_000	HG	Control	13.1486	3.65529	3.597143	1.84685	0.1632	0.401044	no
	EMQU_0471,EMQU_0472 mannose-specific IIC, II D comp	XLOC_000	HG	Control	18.1993	4.99818	3.641185	1.86441	0.06205	0.220858	no
	EMQU_0657 cellobiose-specific IIB component	XLOC_000	HG	Control	3.3736	0.474112	7.115618	2.83099	1	1	no
	EMQU_0720 cellobiose-specific IIC component	XLOC_001	HG	Control	8.35631	6.03448	1.384761	0.469637	0.5708	0.777638	no
	EMQU_0875 cellobiose-specific IIB component	XLOC_000	HG	Control	2197.46	115.917	18.95718	4.24467	0.00005	0.001175	yes
	EMQU_0876 cellobiose-specific IIA component	XLOC_000	HG	Control	8451.05	95.3937	88.59128	6.46909	0.00005	0.001175	yes
	EMQU_0877cellobiose-specific IIC component	XLOC_000	HG	Control	3101.49	60.5217	51.24592	5.67937	0.00005	0.001175	yes
	EMQU_1390 cellobiose-specific IIC component	XLOC_001	HG	Control	4.4595	1.88248	2.368949	1.24425	0.31645	0.570872	no
	EMQU_2136 sucrose-specific IIA component	XLOC_001	HG	Control	25.4856	5.65732	4.504889	2.17149	0.0006	0.008529	yes
	EMQU_2183 cellobiose-specific IIC component	XLOC_001	HG	Control	92.9659	5.97553	15.55777	3.95956	0.00005	0.001175	yes
	EMQU_2184,EMQU_2185,EMQU_2186 beta-glucoside-sp	XLOC_001	HG	Control	363.602	120.015	3.029638	1.59915	0.03995	0.164767	no
	mtlF mannitol-specific IIA component	XLOC_002	HG	Control	29.2225	6.63801	4.402298	2.13826	0.1725	0.412796	no
	EMQU_2415 cellobiose-specific IIC component	XLOC_002	HG	Control	35.5197	2.91282	12.19427	3.60813	0.0568	0.207542	no
	EMQU_2463,EMQU_2464 mannose-specific IIC, IID comp	XLOC_002	HG	Control	26.1215	5.85687	4.459976	2.15704	0.0311	0.139458	no
	EMQU_2465 mannose-specific IIA component	XLOC_002	HG	Control	15.2398	2.15572	7.069471	2.8216	0.1997	0.451974	no
	EMQU_2571 ascorbate-specific IIC component	XLOC_002	HG	Control	28.6278	7.5075	3.813227	1.93101	0.0254	0.120727	no
	EMQU_2572,EMQU_2573 ascorbate-specific IIAB compor	XLOC_002	HG	Control	26.2621	8.51933	3.082649	1.62417	0.0494	0.190075	no
	EMQU_2609 sucrose-specific IIC component	XLOC_002	HG	Control	67.8642	21.2684	3.190847	1.67394	0.0094	0.061402	no
	EMQU_2677 cellobiose-specific IIC component	XLOC_002	HG	Control	6.44759	0.105524	61.1007	5.93312	0.21425	0.47155	no
	celB PTS system cellobiose transporter subunit IIB	XLOC_002	HG	Control	237.039	76.572	3.095635	1.63023	0.0281	0.130251	no
	EMQU_2816 fructose-specific IIB component	XLOC_002	HG	Control	12.8735	2.32851	5.528643	2.46692	0.09655	0.289836	no
	EMQU_2817 fructose-specific IIA component	XLOC_002	HG	Control	52.9449	2.67551	19.78871	4.3066	0.0072	0.051022	no
	EMQU_2819,EMQU_2820nitrogen regulatory IIA compon	XLOC_002	HG	Control	9.10528	3.4275	2.656537	1.40955	0.2707	0.522519	no

	***E. mundtii* living in: Foregut vs Hindgut**									
	current_genes@annotation$gene_short_name	**gene_id**	**sample_1**	**sample_2**	**value_1**	**value_2**	**fold chang**	**log2_fold_**	**p_value**	**q_value**	**significant**
Extracellular interaction	wxl domain containing proteins										
	EMQU_0043	XLOC_000	HG	FG	15.3819	8.94313	1.719968	0.782382	0.1909	0.440908	no
	EMQU_0485	XLOC_000	HG	FG	2.12834	7.51487	0.283217	-1.82002	0.0648	0.226351	no
	EMQU_0487	XLOC_000	HG	FG	6.0516	22.8062	0.265349	-1.91404	0.0143	0.083185	no
	EMQU_0539	XLOC_000	HG	FG	0.847213	7.03728	0.120389	-3.05422	0.2545	0.509327	no
	EMQU_0541	XLOC_000	HG	FG	6.43246	23.112	0.278317	-1.8452	0.0077	0.053038	no
	lpxtg motif containing proteins										
	fms3	XLOC_000	HG	FG	17.234	22.9629	0.750515	-0.41405	0.56015	0.770309	no
	EMQU_1297	XLOC_001	HG	FG	276.842	73.7189	3.755373	1.90896	0.55205	0.76682	no
	lysM, biofilm forming protein										
	EMQU_0517	XLOC_000	HG	FG	264.527	257.543	1.027118	0.038598	0.94215	0.970153	no
	EMQU_1080	XLOC_000	HG	FG	962.047	1468.54	0.655104	-0.61021	0.3236	0.579211	no
	Accessory gene regulator: two-component system										
	agrB	XLOC_001	HG	FG	265058	180974	1.464619	-0.55052	0.61685	0.812907	no
	EMQU_1481 agrA	XLOC_001	HG	FG	8.24085	7.91384	1.041321	-0.05842	0.9467	0.97227	no
	Cell surface anchor proteins										
	EMQU_0445,EMQU_0446,ebpCfm	XLOC_000	HG	FG	23.2154	4.86095	4.775898	2.25577	0.1136	0.322591	no
	EMQU_0540	XLOC_000	HG	FG	1.89396	2.45222	0.772345	-0.37269	1	1	no
	EMQU_1048,ftsW3	XLOC_000	HG	FG	323.47	540.559	0.598399	-0.74082	0.2172	0.473808	no
	Chitin binding proteins										
	EMQU_0940	XLOC_000	HG	FG	479.383	163.522	2.931612	1.5517	0.0197	0.101809	no
	EMQU_1285	XLOC_000	HG	FG	736.496	272.066	2.707049	1.43672	0.0218	0.108351	no
Stress survival strategies	**oxidative stress**										
	EMQU_0929 sod	XLOC_001	HG	FG	555.04	946.264	0.586559	-0.76965	0.70635	0.865	no
	EMQU_0568 catalase	XLOC_001	HG	FG	439.995	163.414	2.692517	1.42895	0.02685	0.125955	no
	EMQU_0459 NADH peroxidase	XLOC_000	HG	FG	414.128	148.947	2.780372	1.47528	0.0088	0.05838	no
	EMQU_1279 NADH Oxidase	XLOC_000	HG	FG	1237.25	818.533	1.511546	-0.59603	0.26305	0.516492	no
	EMQU_1453 ohr	XLOC_000	HG	FG	1282.44	1364.84	0.939627	-0.08983	0.8926	0.950358	no
	EMQU_0165	XLOC_001	HG	FG	118.394	75.9495	1.558852	-0.64048	0.35505	0.60649	no
	stress proteins										
	glsB	XLOC_001	HG	FG	2321.49	773.485	3.001338	1.58561	0.01375	0.081198	no
	glsB1	XLOC_001	HG	FG	5358.13	7635.66	0.701725	-0.51102	0.3689	0.619069	no
	gls33	XLOC_001	HG	FG	271.018	77.1609	3.512375	1.81244	0.00395	0.033589	yes
	uspA2	XLOC_000	HG	FG	834.712	3942.33	0.211731	-2.2397	0.00255	0.024439	yes
	rpmG2,secE	XLOC_002	HG	FG	961.897	1293.71	0.743518	-0.42756	0.6624	0.83972	no
	EMQU_1288	XLOC_000	HG	FG	37.685	38.6307	0.975519	-0.03576	0.95335	0.975499	no
	Repair proteins										
	EMQU_2803	XLOC_002	HG	FG	50.1773	36.9518	1.357912	0.44139	0.54795	0.765974	no
	recA	XLOC_002	HG	FG	975.438	2029.41	0.480651	-1.05694	0.067	0.23092	no
	recF,yaaA	XLOC_000	HG	FG	1206.08	974.208	1.238011	0.308029	0.7138	0.868934	no
	alkD	XLOC_000	HG	FG	196.911	181.389	1.085573	0.118456	0.8465	0.928926	no
	EMQU_2770 radA	XLOC_002	HG	FG	207.496	222.489	0.932612	-0.10065	0.85085	0.931897	no
	EMQU_0695,EMQU_0697,EMQU_0698,EMQU_0699,EMQU_0700,valS (radC)	XLOC_000	HG	FG	27719.6	14127.1	1.962158	0.972442	0.6702	0.843798	no
	EMQU_3001 dnaJ,EMQU_3002 YafQ	XLOC_002	HG	FG	2342.23	18351.1	0.127634	-2.96991	0.0039	0.033468	yes
	Alkaline stress										
	asp,nusB	XLOC_000	HG	FG	1959.03	6009.24	0.326003	-1.61704	0.0241	0.116442	no
	EMQU_2152 nhaC	XLOC_001	HG	FG	57.8409	19.2674	3.002009	1.58593	0.03375	0.14766	no
Metabolism	EMQU_0667 adenosine deaminase	XLOC_000	HG	FG	184.412	285.117	0.646794	-0.62862	0.43325	0.677243	no
	EMQU_2472 glucosamine-6-phosphate deaminase	XLOC_002	HG	FG	764.978	1806.1	0.423552	-1.23938	0.07335	0.242273	no
	metC	XLOC_001	HG	FG	13.8188	2.60686	5.300937	2.40625	0.01765	0.094127	no
	6-phospho-beta-glucosidases										
	bglB	XLOC_000	HG	FG	5.53878	2.94125	1.883138	0.91314	0.3336	0.585898	no
	bglH	XLOC_000	HG	FG	42.8138	16.5532	2.586436	1.37096	0.0434	0.174485	no
	bglG2	XLOC_000	HG	FG	20.4446	17.3353	1.179362	0.238009	0.7399	0.881508	no
	bglF	XLOC_000	HG	FG	14.0362	9.17342	1.530095	0.613616	0.32805	0.581788	no
	pfkA phosphofurctokinase A	XLOC_000	HG	FG	818.513	830.107	0.986033	-0.02029	0.96785	0.981531	no
	ldhA lactose dehydrogenase	XLOC_001	HG	FG	10.5677	41.7702	0.252996	-1.98281	0.00615	0.045802	yes
	sfsA sugar fermentation stimulation protein	XLOC_000	HG	FG	418.669	627.444	0.667261	-0.58368	0.2961	0.551797	no
	alcohol dehydrogenases										
	fdh	XLOC_002	HG	FG	141.433	118.885	1.189662	0.250554	0.6601	0.837799	no
	EMQU_0525,EMQU_0526	XLOC_000	HG	FG	508.209	455.45	1.115839	0.15813	0.8279	0.92083	no
	EMQU_1129	XLOC_000	HG	FG	44.0849	32.8955	1.34015	0.422392	0.52195	0.74335	no
	EMQU_0450,EMQU_0452,dapB,dapF (lysine biosynthesis)	XLOC_000	HG	FG	218.8	63.2449	3.459567	1.79059	0.0047	0.038031	yes
	EMQU_0175,EMQU_0176,EMQU_0177,EMQU_0178 (lysine transport)	XLOC_001	HG	FG	437.44	1177.89	0.371376	-1.42905	0.3527	0.604241	no
	Phospho Transferase Systems										
	EMQU_0169 trehalose-specific IIA component	XLOC_000	HG	FG	26.195	568.571	0.046072	-0.05195	0.92805	0.964553	no
	EMQU_0285 PTS family porter component IIA	XLOC_000	HG	FG	39.2145	3.80208	10.31396	2.78443	0.0001	0.002165	yes
	EMQU_0286 ascorbate-specific IIB component	XLOC_000	HG	FG	8.82057	9.35016	0.94336	2.06832	0.30965	0.564575	no
	EMQU_0287 PTS system ascorbate-specific transporter subunit IIC	XLOC_000	HG	FG	1018.06	5.83015	174.6199	0.597339	0.5207	0.7427	no
	EMQU_0307 cellobiose-specific IIB component	XLOC_000	HG	FG	1379.44	356.938	3.864649	1.51208	0.0157	0.087624	no
	EMQU_0308 cellobiose-specific IIA component	XLOC_000	HG	FG	799.11	644.87	1.23918	1.097	0.0987	0.295342	no
	EMQU_0309 cellobiose-specific IIC component	XLOC_000	HG	FG	33.6888	282.52	0.119244	1.50004	0.0103	0.065187	no
	EMQU_0390 sugar-specific IIA component	XLOC_000	HG	FG	1.09225	251.418	0.004344	-2.89975	0.00005	0.001175	yes
	EMQU_0463,EMQU_0464 fructose-specific IIB component	XLOC_000	HG	FG	3.0058	1.96445	1.530097	-0.84682	1	1	no
	EMQU_0465, fructose-specific IIA component	XLOC_000	HG	FG	13.1486	6.71826	1.957144	-1.16034	0.57035	0.777251	no
	EMQU_0470 mannose-specific IIB component	XLOC_000	HG	FG	18.1993	26.9057	0.676411	-1.033	0.4106	0.657969	no
	EMQU_0471,EMQU_0472 mannose-specific IIC, II D component	XLOC_000	HG	FG	3.3736	20.4329	0.165106	-0.16701	0.8431	0.926716	no
	EMQU_0657 cellobiose-specific IIB component	XLOC_000	HG	FG	8.35631	3.313	2.522279	0.026148	1	1	no
	EMQU_0720 cellobiose-specific IIC component	XLOC_001	HG	FG	2197.46	8.57765	256.1844	-0.03772	0.9647	0.981531	no
	EMQU_0875 cellobiose-specific IIB component	XLOC_000	HG	FG	8451.05	686.39	12.31232	1.67874	0.0089	0.05871	no
	EMQU_0876 cellobiose-specific IIA component	XLOC_000	HG	FG	3101.49	7454.4	0.416062	0.181037	0.74865	0.884201	no
	EMQU_0877cellobiose-specific IIC component	XLOC_000	HG	FG	4.4595	1142.83	0.003902	1.44036	0.0243	0.117166	no
	EMQU_1390 cellobiose-specific IIC component	XLOC_001	HG	FG	25.4856	1.90791	13.35786	1.22489	0.34075	0.593284	no
	EMQU_2136 sucrose-specific IIA component	XLOC_001	HG	FG	92.9659	12.4744	7.452535	1.03072	0.14605	0.373597	no
	EMQU_2183 cellobiose-specific IIC component	XLOC_001	HG	FG	363.602	26.5973	13.67064	1.80542	0.00495	0.03944	yes
	EMQU_2184,EMQU_2185,EMQU_2186 beta-glucoside-specific IIA component	XLOC_001	HG	FG	29.2225	69.8086	0.418609	2.38088	0.0004	0.006363	yes
	mtlF mannitol-specific IIA component	XLOC_002	HG	FG	35.5197	10.6931	3.32174	1.4504	0.45775	0.697133	no
	EMQU_2415 cellobiose-specific IIC component	XLOC_002	HG	FG	26.1215	9.68388	2.697421	1.87496	0.3924	0.641473	no
	EMQU_2463,EMQU_2464 mannose-specific IIC, IID component	XLOC_002	HG	FG	15.2398	43.0103	0.354329	-0.71944	0.33095	0.583875	no
	EMQU_2465 mannose-specific IIA component	XLOC_002	HG	FG	28.6278	98.6587	0.29017	-2.69461	0.07305	0.241965	no
	EMQU_2571 ascorbate-specific IIC component	XLOC_002	HG	FG	26.2621	18.4545	1.423073	0.633447	0.39525	0.64321	no
	EMQU_2572,EMQU_2573 ascorbate-specific IIAB component	XLOC_002	HG	FG	67.8642	9.10346	7.45477	1.5285	0.03865	0.161607	no
	EMQU_2609 sucrose-specific IIC component	XLOC_002	HG	FG	6.44759	93.3782	0.069048	-0.46044	0.462	0.700413	no
	EMQU_2677 cellobiose-specific IIC component	XLOC_002	HG	FG	237.039	19.8547	11.93868	-1.62265	0.04795	0.186419	no
	celB PTS system cellobiose transporter subunit IIB	XLOC_002	HG	FG	12.8735	219.915	0.058539	0.108175	0.8801	0.946259	no
	EMQU_2816 fructose-specific IIB component	XLOC_002	HG	FG	52.9449	3.58546	14.76656	1.84418	0.3427	0.594687	no
	EMQU_2817 fructose-specific IIA component	XLOC_002	HG	FG	9.10528	42.1034	0.21626	0.330555	0.71985	0.872438	no
	EMQU_2819,EMQU_2820nitrogen regulatory IIA component	XLOC_002	HG	FG	9.10528	10.0406	0.906846	-0.14107	0.91355	0.958643	no

12.2 Appendix 2:

Genes that are significantly differentially regulated in *Enterococcus mundtii* when they are dwelling in both fore and hindguts of *Spodoptera littoralis*.

Cufflinks ID	**Genes that are up regulated in *Enterococcus mundtii* when they are dwelling in both fore and hindguts of *Spodoptera littoralis***
XLOC_000016	EMQU_0019 EMQU_0019K07238 zinc transporter, ZIP family | (GenBank) gufA-like protein zinc transporter
XLOC_000029	EMQU_0040 Hypothetical protein
XLOC_000100	EMQU_0140,EMQU_0141,fetC
XLOC_000242	dhaL
XLOC_000246	EMQU_0390PTS system, sugar-specific IIA component [EC:2.7.1.-] K20107 PTS system, maltose-specific IIB component
XLOC_000255	EMQU_0406 no KO assigned | (GenBank) Hypothetical proteinthetical protein
XLOC_000257	EMQU_0408 no KO assigned | (GenBank) Hypothetical proteinthetical protein
XLOC_000269	EMQU_0428 K09815 zinc transport system substrate-binding protein | (GenBank) zinc-binding lipoprotein AdcA
XLOC_000270	EMQU_0433 no KO assigned | (GenBank) Hypothetical proteinthetical protein
XLOC_000273	EMQU_0436K16509 regulatory protein spx | (GenBank) arsenate reductase
XLOC_000297	EMQU_0474K01223 6-phospho-beta-glucosidase [EC:3.2.1.86] | (GenBank) bglH; beta-glucosidase
XLOC_000311	EMQU_0491 K10212 glycosyl-4,4'-diaponeurosporenoate acyltransferase [EC:2.3.1.-] | (GenBank) conserved Hypothetical proteinthetical protein
XLOC_000327	EMQU_0517 no KO assigned | (GenBank) lysM family surface protein
XLOC_000333	EMQU_0525 K00121 S-(hydroxymethyl)glutathione dehydrogenase / alcohol dehydrogenase [EC:1.1.1.284 1.1.1.1] | (GenBank) S-(hydroxymethyl)glutathione dehydrogenase
XLOC_000336	EMQU_0531 Hypothetical proteinthetical protein
XLOC_000345	EMQU_0541 no KO assigned | (GenBank) wxL domain surface protein
XLOC_000356	EMQU_0557 K19973 manganese/zinc transport system ATP-binding protein
XLOC_000366	EMQU_0571 K05340 glucose uptake protein
XLOC_000375	EMQU_0583 K00948 ribose-phosphate pyrophosphokinase [EC:2.7.6.1] | (GenBank) ribose-phosphate pyrophosphokinase
XLOC_000396	EMQU_0608 K01223 6-phospho-beta-glucosidase (glucolysis)
XLOC_000430	EMQU_0658 no KO assigned | (GenBank) metal-dependent hydrolase
XLOC_000437	EMQU_0667K11991 tRNA(adenine34) deaminase
XLOC_000467	EMQU_0724 K02343 DNA polymerase III subunit gamma/tau
XLOC_000519	EMQU_0833no KO assigned | (GenBank) integral membrane protein
XLOC_000533	EMQU_0851 K03925 MraZ protein | (GenBank) cell division protein MraZ
XLOC_000534	EMQU_0852 K03438 16S rRNA (cytosine1402-N4)-methyltransferase
XLOC_000548	EMQU_0871 K06206 sugar fermentation stimulation protein A | (GenBank) sfsA; sugar fermentation stimulation protein
XLOC_000551	EMQU_0875 K02760 PTS system, cellobiose-specific IIB component [EC:2.7.1.196 2.7.1.205] | (GenBank) PTS system transporter subunit IIB
XLOC_000552	EMQU_0876 K02759 PTS system, cellobiose-specific IIA component [EC:2.7.1.196 2.7.1.205] | (GenBank) PTS system transporter subunit IIA
XLOC_000553	EMQU_0877 K02761 PTS system, cellobiose-specific IIC component | (GenBank) PTS family lactose/cellobiose porter component IIC
XLOC_000591	EMQU_0940 K21713 lytic chitin monoxygenase [EC:1.14.99.53] | (GenBank) chitin binding protein
XLOC_000628	EMQU_1000 pmrA; MFS family major facilitator multidrug/cation transporter
XLOC_000669	uspA2
XLOC_000675	EMQU_1048 K05837 rod shape determining protein RodA | (GenBank) cell cycle protein FtsW, FtsW3
XLOC_000682	EMQU_1061 K08974 putative membrane protein | (GenBank) transmembrane protein
XLOC_000713	EMQU_1097 no KO assigned | (GenBank) degV family protein
XLOC_000726	EMQU_1126 K06889 uncharacterized protein | (GenBank) alpha/beta hydrolase
XLOC_000815	EMQU_1283 putative lipoprotein
XLOC_000817	EMQU_1285 chitin binding protein
XLOC_000820	EMQU_1288 K03205 type IV secretion system protein VirD4 | (GenBank) conjugal transfer protein
XLOC_000823	EMQU_1292 no KO assigned | (GenBank) trsE protein
XLOC_000824	EMQU_1294putative secreted cell wall protein
XLOC_000900	EMQU_1453 no KO assigned | (GenBank) organic hydroperoxide resistance family protein
XLOC_001110	EMQU_2844 K20379 sex pheromone cAD1 | (GenBank) pheromone cAD1 precursor lipoprotein
XLOC_001118	EMQU_2889 no KO assigned | (GenBank) XRE family transcriptional regulator
XLOC_001139	EMQU_0073 K03784 purine-nucleoside phosphorylase [EC:2.4.2.1] | (GenBank) purine nucleoside phosphorylase
XLOC_001185	EMQU_0317 oleate hydratase
XLOC_001241	EMQU_0568 K07217 Mn-containing catalase
XLOC_001399	EMQU_1411 no KO assigned | (GenBank) degV family protein
XLOC_001413	EMQU_1433 K15771 arabinogalactan oligomer / maltooligosaccharide transport system permease protein | (GenBank) malC; maltose/maltodextrin ABC superfamily ATP binding cassette transporter, membrane protein
XLOC_001414	EMQU_1434 K15770 arabinogalactan oligomer / maltooligosaccharide transport system substrate-binding protein | (GenBank) malX; maltose/maltodextrin ABC superfamily ATP binding cassette transporter, binding protein
XLOC_001415	EMQU_1435 K01208 cyclomaltodextrinase / maltogenic alpha-amylase / neopullulanase [EC:3.2.1.54 3.2.1.133 3.2.1.135] | (GenBank) neopullulanase (Starch and sucrose metabolism)
XLOC_001420	EMQU_1444 no KO assigned | (GenBank) N-acetylmuramoyl-L-alanine amidase
XLOC_001434	glsB1
XLOC_001445	glsB
XLOC_001446	gls33
XLOC_001449	EMQU_1495fad; short chain dehydrogenase/reductase family oxidoreductase
XLOC_001452	EMQU_1498 general stress protein A
XLOC_001654	EMQU_1797no KO assigned | (GenBank) marR family transcriptional regulator
XLOC_001915	EMQU_2157 no KO assigned | (GenBank) HAD-superfamily hydrolase
XLOC_001926	EMQU_2172 K01223 6-phospho-beta-glucosidase [EC:3.2.1.86] | (GenBank) 6-phospho-beta-glucosidase (glycolysis)
XLOC_001930	EMQU_2178 K00874 2-dehydro-3-deoxygluconokinase [EC:2.7.1.45] | (GenBank) carbohydrate kinase (pentose phosphate pw)
XLOC_001932	EMQU_2180 K01042 L-seryl-tRNA(Ser) seleniumtransferase [EC:2.9.1.1] | (GenBank) pyridoxal phosphate-dependent enzyme, EMQU_2181K01465 dihydroorotase
XLOC_001934	EMQU_2183K02761 PTS system, cellobiose-specific IIC component
XLOC_001936	EMQU_2187 K03488 beta-glucoside operon transcriptional antiterminator | (GenBank) bglG family transcriptional antiterminator
XLOC_002009	EMQU_2292 no KO assigned | (GenBank) gp46
XLOC_002094	EMQU_2454 no KO assigned | (GenBank) NAD dependent epimerase/dehydratase family protein
XLOC_002100	EMQU_2470no KO assigned | (GenBank) permease
XLOC_002101	EMQU_2471 K07507 putative Mg2+ transporter-C (MgtC) family protein | (GenBank) mgtC; mgtC family magnesium (Mg2+) transporter-C
XLOC_002102	EMQU_2472 K02564 glucosamine-6-phosphate deaminase [EC:3.5.99.6] | (GenBank) glucosamine-6-phosphate isomerase
XLOC_002142	EMQU_2540 K02055 putative spermidine/putrescine transport system substrate-binding protein | (GenBank) spermidine/putrescine ABC superfamily ATP binding cassette transporter
XLOC_002167	EMQU_0315K00121 S-(hydroxymethyl)glutathione dehydrogenase / alcohol dehydrogenase [EC:1.1.1.284 1.1.1.1] | (GenBank) fdh; S-(hydroxymethyl)glutathione dehydrogenase (glycolysis)
XLOC_002217	EMQU_2663 K00847 fructokinase [EC:2.7.1.4] | (GenBank) cscK3; fructokinase Fructose and mannose metabolism
XLOC_002218	EMQU_2664 K01191 alpha-mannosidase [EC:3.2.1.24] | (GenBank) alpha-mannosidase
XLOC_002221	EMQU_2671 K17318 putative aldouronate transport system substrate-binding protein | (GenBank) multiple sugar ABC transporter substrate-binding protein
XLOC_002222	EMQU_2672 K17320 putative aldouronate transport system permease protein | (GenBank) ypcH; carbohydrate ABC superfamily ATP binding cassette transporter, membrane protein
XLOC_002223	EMQU_2673 K17319 putative aldouronate transport system permease protein | (GenBank) sugar ABC transporter permease
XLOC_002292	EMQU_2731K00942 guanylate kinase [EC:2.7.4.8] | (GenBank) gmk; guanylate kinase (Purine metabolism)
XLOC_002494	EMQU_3001 K07473 DNA-damage-inducible protein J | (GenBank) DNA-damage-inducible protein J

Cufflinks ID	Genes that are down regulated in *Enterococcus mundtii* when they are dwelling in both fore and hindguts of *Spodoptera littoralis*
XLOC_000058	rplB large subunit ribosomal protein L2
XLOC_000063	rpsQ small subunit ribosomal protein S17
XLOC_000064	rplN
XLOC_000065	rplX
XLOC_000066	rplE
XLOC_000068	rpsH
XLOC_000070	rplR
XLOC_000071	rpsE
XLOC_000268	EMQU_0426K01226 trehalose-6-phosphate hydrolase [EC:3.2.1.93] | (GenBank) alpha amylase family protein
XLOC_000278	EMQU_0443K02356 elongation factor P | (GenBank) efp; elongation factor P
XLOC_000319	EMQU_0504K01736 chorismate synthase [EC:4.2.3.5] | (GenBank) chorismate synthase
XLOC_000321	EMQU_0506 K00800 3-phosphoshikimate 1-carboxyvinyltransferase [EC:2.5.1.19] | (GenBank) 3-phosphoshiki EMQU_0508 K04518 prephenate dehydratase [EC:4.2.1.51] | (GenBank) prephenate dehydratase
XLOC_000460	EMQU_0711K03713 MerR family transcriptional regulator, glutamine synthetase represso
XLOC_000461	EMQU_0712K01915 glutamine synthetase (Arginine biosynthesis)_
XLOC_000474	EMQU_0736K02346 DNA polymerase IV
XLOC_000491	EMQU_0763 Hypothetical proteinthetical prot
XLOC_000547	EMQU_0866 K01870 isoleucyl-tRNA synthetase [EC:6.1.1.5] | (GenBank) ileS
XLOC_000562	EMQU_0897 K00067 dTDP-4-dehydrorhamnose reductase [EC:1.1.1.133] | (GenBank) dTDP-4-dehydrorhamnose reductase
XLOC_000598	EMQU_0955K08483 phosphotransferase system, enzyme I, PtsI [EC:2.7.3.9] | (GenBank) phosphoenolpyruvate-protein phosphotransferase
XLOC_000600	EMQU_0964 K00645 [acyl-carrier-protein] S-malonyltransferase [EC:2.3.1.39] | (GenBank) acyl-carrier-protein S-malonyltransferase
XLOC_000601	EMQU_0965K00059 3-oxoacyl-[acyl-carrier protein] reductase [EC:1.1.1.100] | (GenBank) fabG; 3-oxoacyl-ACP reductase
XLOC_000602	EMQU_0966 K09458 3-oxoacyl-[acyl-carrier-protein] synthase II [EC:2.3.1.179] | (GenBank) 3-oxoacyl-(acyl carrier protein) synthase II,
XLOC_000603	EMQU_0967 K02160 acetyl-CoA carboxylase biotin carboxyl carrier protein | (GenBank) acetyl-CoA carboxylase, biotin carboxyl carrierprotein
XLOC_000604	EMQU_0969 K01961 acetyl-CoA carboxylase, biotin carboxylase subunit [EC:6.4.1.2 6.3.4.14] | (GenBank) accC; biotin carboxylase
XLOC_000605	EMQU_0971 K01962 acetyl-CoA carboxylase carboxyl transferase subunit alpha [EC:6.4.1.2 2.1.3.15] | (GenBank) acetyl-CoA carboxylase subunit alpha, accD
XLOC_002361	EMQU_2832 K03046 DNA-directed RNA polymerase subunit beta' [EC:2.7.7.6] | (GenBank) rpoC; DNA-directed RNA polymerase subunit beta prime
XLOC_000622	EMQU_0993 K01868 threonyl-tRNA synthetase [EC:6.1.1.3] | (GenBank) threonyl-tRNA synthetase
XLOC_000631	EMQU_1004pyk; pyruvate kinase
XLOC_001066	EMQU_2596 K06607 myo-inositol catabolism protein IolS [EC:1.1.1.-] | (GenBank) oxidoreductase
XLOC_001122	EMQU_0015 no KO assigned | (GenBank) ABC transporter ATP-binding protein
XLOC_001141	EMQU_0077 K01834 2,3-bisphosphoglycerate-dependent phosphoglycerate mutase [EC:5.4.2.11] | (GenBank) gpmA; phosphoglycerate mutase (glycolysis)
XLOC_001255	EMQU_0640 K06147 ATP-binding cassette, subfamily B, bacterial | (GenBank) ABC transporter ATP-binding protein/permease
XLOC_001291	EMQU_0868 K00036 glucose-6-phosphate 1-dehydrogenase [EC:1.1.1.49 1.1.1.363] | (GenBank) zwf; glucose-6-phosphate 1-dehydrogenase (pentose phosphate pathway)
XLOC_001297	EMQU_0916 beta-lactamase
XLOC_001309	EMQU_0990 K03657 DNA helicase II / ATP-dependent DNA helicase PcrA [EC:3.6.4.12] | (GenBank) ATP-dependent DNA helicase
XLOC_002170	EMQU_2578 K16511 adapter protein MecA 1/2 | (GenBank) mecA; competence negative regulator MecA
XLOC_002225	EMQU_2675no KO assigned | (GenBank) MOP superfamily multidrug/oligosaccharidyl-lipid/polysaccharide flippase transporter
XLOC_002232	EMQU_2683 K01869 leucyl-tRNA synthetase [EC:6.1.1.4] | (GenBank) leuS; leucyl-tRNA synthetase
XLOC_002309	EMQU_2753 K03742 nicotinamide-nucleotide amidase [EC:3.5.1.42] | (GenBank) competence damage-inducible protein A
XLOC_002361	EMQU_2832 K03046 DNA-directed RNA polymerase subunit beta' [EC:2.7.7.6] | (GenBank) rpoC; DNA-directed RNA polymerase subunit beta prime

12.3 Appendix 3:

GC measurement of spit for levels of 8-HQA from WTs; and KMOs after fed with serially diluted concentrations of commercial 8-HQA. The highlighted cells have similar levels of 8-HQA in the spit: WT and KMOs fed with 0.2 mg/ml of 8-HQA.

GC measurement of spit for levels of 8-HQA from WTs; and KMOs after fed with serially diluted concentrations of commercial 8-HQA

mass= 368, Retention Time= 9.34-9.41

condition	tube no.	weight of vial (mg)	vial+spit (mg)	weight of spit (mg)	area of peak	area/spitWeight	average
	1	2361,99	2362,16	0,17	0	0	
	2	2407,28	2410,11	2,83	nd		
	3	2381,4	2390,31	8,91	15383	1726,487093	
	4	2366,31	2376,06	9,75	22554	2313,230769	
	5	2411,48	2413,85	2,37	17013	7178,481013	
Control: Wild types	6	2396,98	2404,92	7,94	67365	8484,256927	4925,61395
	7	2374,01	2389,5	15,49	0	0	
	8	2355,89	2373,43	17,54	0	0	
	9	2394,86	2410,1	15,24	0	0	
	10	2396,35	2413,66	17,31	0	0	
	11	2376,32	2389,8	13,48	0	0	
	12	2357,22	2361,01	3,79	0	0	
	13	2384,21	2396,35	12,14	0	0	
	14	2379,5	2408,04	28,54	0	0	
	15	2367,66	2379,92	12,26	0	0	
	16	2386,08	2398,78	12,7	0	0	
	17	2411,93	2426,93	15	0	0	
KMOs fed with plain water	18	2359	2362,76	3,76	0	0	0
	19	2358,15	2366,91	8,76	593486	67749,54338	
	20	2363,08	2371,51	8,43	924319	109646,382	
	21	2357,12	2384,33	27,21	1081471	39745,35097	
	22	2413,42	2445,39	31,97	1308421	40926,52487	
	23	2382,95	2436,71	53,76	1380206	25673,4747	
	24	2360,12	2370,84	10,72	1358438	126719,9627	
	25	2388,02	2399,29	11,27	1969110	174721,3842	
	26	2372,64	2380,97	8,33	2480059	297726,1705	
	27	2416,51	2498,24	81,73	6687747	81827,32167	
	28	2356,12	2424,27	68,15	640618	9400,117388	
	29	2386,34	2418,53	32,19	921029	28612,27089	
KMO knockouts fed with 3mg/ml 8-HQA	30	2384,49	2394,8	10,31	8934953	866629,7769	155781,5233
	31	2378,15	2430,6	52,45	3336	63,60343184	
	32	2403,22	2411,94	8,72	467912	53659,63303	
	33	2399,3	2405,62	6,32	872684	138082,9114	
	34	2365,99	2387,56	21,57	1377757	63873,75985	
	35	2364,77	2397,7	32,93	1349813	40990,37352	
	36	2384,19	2396,84	12,65	509753	40296,67984	
	37	2367,54	2376,61	9,07	285449	31471,77508	
	38	2358,6	2375,7	17,1	2047780	119753,2164	
	39	2356,52	2377,91	21,39	348984	16315,28752	
	40	2381,64	2401,73	20,09	1847776	91974,91289	
	41	2351,03	2355,88	4,85	21893	4514,020619	
KMO knockouts fed with (1:5 of 3mg/ml) 8-HQA	42	2395,64	2398,45	2,81	29808	10607,82918	50967,00023
	43	2388,77	2405,58	16,81	118333	7039,440809	
	44	2364,62	2396,31	31,69	626584	19772,2941	
	45	2382,39	2403,74	21,35	120301	5634,70726	
	46	2363,9	2378,99	15,09	796034	52752,41882	
	47	2406,54	2434,11	27,57	99912	3623,939064	
	48	2373,02	2382,83	9,81	14158	1443,221203	
	49	2386,39	2422,84	36,45	155147	4256,433471	
	50	2395,88	2414,05	18,17	309609	17039,57072	
	51	2397,13	2405,26	8,13	128328	15784,50185	
	52	2407,32	2422,98	15,66	31725	2025,862069	
	53	2368,74	2379,5	10,76	131440	12215,61338	
KMO knockouts fed with (1:10 of 3mg/ml) 8-HQA	54	2379,57	2410,87	31,3	159163	5085,079872	12222,75688
	55	2365,68	2398,71	33,03	159970	4843,172873	
	56	2358,73	2386,88	28,15	87578	3111,119005	
	57	2364,49	2381,58	17,09	70254	4110,825044	
	58	2396,92	2439,21	42,29	22298	527,2641286	
	59	2414,14	2429,11	14,97	227339	15186,30595	
	60	2385,54	2349,19	-36,35	13591		
	61	2346,52	2393,32	46,8	0		
	62	2380,76	2393,31	12,55	35485	2827,49004	
	63	2347,87	2363,52	15,65	29438	1881,022364	
	64	2387,76	2393,09	5,33	3526	661,5384615	
	65	2381,56	2397,02	15,46	102958	6659,637775	
KMO knockouts fed with (1:15 of 3mg/ml) 8-HQA (0.2 mg/ml)	66	2388,92	2417,34	28,42	43970	1547,149894	4135,552553

13 Acknowledgement

I would like to extend my heartfelt gratitude to Professor Dr. Wilhelm Boland for giving me the opportunity to work in this esteemed institute. His supervision has been a guiding light all throughout the process of my PhD. All my other supervisors, Professor Dr. Erika Kothe, Professor Dr. David Heckel and Dr. Shantanu Shukla have played major roles in bringing the projects to completion.

My gratitude goes out to all my collaborators: Dr. Wolfgang Schmidt, Yvonne Schlenker, Dr. Martin Froester, Professor Dr. Peter Zipfel, Dr. Sabine Haenniger and colleagues from the former department of Bioorganic Chemistry- Beng Soon Teh, Vincensius Oetama, Pol Alonso, Abith Vattekatte, Dr. Maritta Kunert, Dr. Stefan Bartram, Anja David, Anja Meents, Andrea Lehr and Angelika Berg for contributing to the projects, without whose help, the projects would not be successful.

Supervising my Master Students- Aishwarya Murali and Mehedi Hasan Fuad made me see entire different perspectives to my projects; hence it has been an enriching experience working together with them. In addition, their contributions have been endless.

Apart from my professional life, my friends in Jena have been an amazing support system throughout the process. I would personally like to thank Ravindra, Prerna, Sagar, Amol, Kohulan, Anja, Vincensius, Nanxia, Sin, Burak, Alberto, Prasad, Shraddha, Samir, Ghada, Ramya and Manasa for making my stay in Jena so enjoyable. Lastly, it goes without saying that contributions of my family in India are immense. Their love and trust formed my pillar of strength through thick and thin.

14 Ehrenwörtliche Erklärung

Entsprechend der geltenden, mir bekannten Promotionsordnung der Fakultät für Biowissenschaften der Friedrich-Schiller-Universität Jena erkläre ich, dass ich die vorliegende Dissertation eigenständig angefertigt und alle von mir benutzten Hilfsmittel und Quellen angegeben habe. Personen, die mich bei der Auswahl und Auswertung des Materials sowie bei Fertigstellung der Manuskripte unterstützt haben, sind am Beginn eines jeden Kapitel genannt. Es wurde weder die Hilfe einer kommerziellen Promotionsvermittlung in Anspruch genommen, noch haben Dritte für Arbeiten, welche im Zusammenhang mit dem Inhalt der vorliegenden Dissertation stehen, geldwerte Leistungen erhalten. Die vorgelegte Dissertation wurde außerdem weder als Prüfungsarbeit für eine staatliche oder andere wissenschaftliche Prüfung noch als Dissertation an einer anderen Hochschule eingereicht. Weiterhin wurde keine ähnliche oder andere Abhandlung als Dissertation anderswo eingereicht.

Tilottama Mazumdar, Jena, 26th June 2020