UPPSALA
UNIVERSITET

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 1121

Genome Evolution and Niche Differentiation of Bacterial Endosymbionts

KIRSTEN MAREN ELLEGAARD

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2014

ISSN 1651-6214
ISBN 978-91-554-8872-7
urn:nbn:se:uu:diva-217724

Dissertation presented at Uppsala University to be publicly examined in B42, Husargatan 3, Uppsala, Friday, 21 March 2014 at 10:15 for the degree of Doctor of Philosophy. The examination will be conducted in English. Faculty examiner: Professor Andrés Moya (University of Valencia).

Abstract
Ellegaard, K. M. 2014. Genome Evolution and Niche Differentiation of Bacterial Endosymbionts. *Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology* 1121. 57 pp. Uppsala: Acta Universitatis Upsaliensis. ISBN 978-91-554-8872-7.

Most animals contain chronic microbial infections that inflict no harm on their hosts. Recently, the gut microflora of humans and other animals have been characterized. However, little is known about the forces that shape the diversity of these bacterial communities. In this work, comparative genomics was used to investigate the evolutionary dynamics of host-adapted bacterial communities, using *Wolbachia* infecting arthropods and Lactobacteria infecting bees as the main model systems.

Wolbachia are maternally inherited bacteria that cause reproductive disorders in arthropods, such as feminization, male killing and parthenogenesis. These bacteria are difficult to study because they cannot be cultivated outside their hosts. We have developed a novel protocol employing multiple displacement amplification to isolate and sequence their genomes. Taxonomically, *Wolbachia* is classified into different supergroups. We have sequenced the genomes of *Wolbachia* strain *w*Ha and *w*No that belong to supergroup A and B, respectively, and are present as a double-infection in the fruit-fly *Drosophila simulans*. Together with previously published genomes, a supergroup comparison of strains belonging to supergroups A and B indicated rampant homologous recombination between strains that belong to the same supergroup but were isolated from different hosts. In contrast, we observed little recombination between strains of different supergroups that infect the same host.

Likewise, phylogenetically distinct members of Lactic acid bacteria co-exist in the gut of the honeybee, *Apis mellifera,* without transfer of genes between phylotypes. Nor did we find any evidence of co-diversification between symbionts and hosts, as inferred from a study of 13 genomes of *Lactobacillus kunkeei* isolated from diverse bee species and different geographic origins. Although *Lactobacillus kunkeii* is the most frequently isolated strain from the honey stomach, we hypothesize that the primary niche is the beebread where the bacteria are likely to contribute to the fermentation process.

In the human gut, the microbial community has been shown to interact with the immune system, and likewise the microbial communities associated with insects are thought to affect the health of their host. Therefore, a better understanding of the role and evolution of endosymbiotic communities is important for developing strategies to control the health of their hosts.

Keywords: niche, habitat, endosymbiont, gut microbiome, honey bee, Wolbachia, comparative, genomics

Kirsten Maren Ellegaard, Department of Cell and Molecular Biology, Molecular Evolution, Box 596, Uppsala University, SE-752 37 Uppsala, Sweden.

ISSN 1651-6214
ISBN 978-91-554-8872-7
urn:nbn:se:uu:diva-217724 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-217724)

To my Family

List of Papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.

I Ellegaard, KM.*, Klasson, L.*, Näslund, K., Bourtzis, K., Andersson, SG. (2013) Comparative Genomics of *Wolbachia* and the Bacterial Species Concept. *PLoS Genetics,* 9(4): e1003381

II Ellegaard, KM., Klasson, L., Andersson, SG. (2013): Testing the Reproducibility of Multiple Displacement Amplification on Genomes of Clonal Endosymbiont Populations. *PLoS One,* 8(11): e82319

III Ellegaard, KM.*, Tamarit, D.*, Javelind, E., Olofsson, T., Andersson, SG., Vásquez, A. Comparative Genomics of Lactic acid bacteria, isolated from the honey-stomach of *Apis mellifera. Manuscript*

IV Tamarit, D.*, Ellegaard, KM.*, Olofsson, T., Vásquez, A., Andersson, SG. Comparative Genomics of *Lactobacillus kunkeii* indicates Selection for Rapid Growth in Beebread. *Manuscript*

Reprints were made with permission from the respective publishers.

* These authors contributed equally to this work

Papers by the author not included in this thesis

1. Berglund, EC., Ellegaard, K.*, Granberg, F.*, Xie, Z.*, Maruyama, S., Kosoy, MY, Birtles, RJ., Andersson, SG. (2010) Rapid diversification by recombination in *Bartonella grahamii* from wild rodents in Asia contrasts with low levels of genomic divergence in Northern Europe and America. *Molecular Ecology,* 19(11): 2241-2255
2. Decaestecker, E.*, Labbé, P.*, Ellegaard, K., Allen, JE., Little, TJ. (2011) Candidate innate immune system gene expression in the ecological model *Daphnia*. *Developmental and comparative immunology,* 35(10): 1068-1077

* These authors contributed equally to this work

Contents

Abbreviations

ANI	Average nucleotide identity
CI	Cytoplasmic incompatibility
CRISPR	Clustered regularly interspaced short palindromic repeats
LAB	Lactic acid bacteria
MDA	Multiple displacement amplification
MLST	Multi-locus sequence typing
OTU	Operational taxonomic unit
PCR	Polymerase chain reaction
PTS	Phosphotransferase systems

Chapter 1: Introduction

> "By the means of Telescopes, there is nothing so far distant but may be represented to our view; and by the help of Microscopes, there is nothing so small as to escape our inquiry; hence there is a new visible world discovered to the understanding"
>
> - Robert Hooke, 1665 [1]

So it began, nearly 400 years ago, the discovery of the microbial world. It took another couple of centuries before the studies on bacteria began in earnest, and with the latest discoveries it is clear that we have only just scratched the surface.

Biology is not what is used to be; We have entered the "next-generation" sequencing era. Genetics has become genomics, and has in turn given rise to transcriptomics, proteomics, metabolomics, lipidomics, mobilomics.. In other words, we may say that we have entered the "omics" era. In a time when things are moving at high speed, it is a challenge to keep up with the literature and latest findings. In fact, I find that I rarely read papers, which have been published more than a decade ago.

However, some of the questions I have faced during my PhD, dabbling with various "omics" technologies, appear to have deeper roots than what was apparent at first glance. I have found myself asking very fundamental questions, such as "What is a species?" "What is a niche?" "Do species compete or cooperate?" To answer such questions, it seems appropriate to take on a longer perspective. Therefore, I have devoted the first part of this thesis to an introduction of the concepts of symbiosis, niches and bacterial species, with a focus on the historical origin of the terms, and the ramifications to current microbiology.

In chapter 5, I describe some of the known features of host-symbiont evolution from a genomic perspective, thus moving forward in time to more recent science. Finally, in chapter 6-8, I introduce the specific symbiotic bacteria, which form the basis of my thesis, as well as the papers and manuscripts we have produced so far.

Chapter 2: Symbiosis

> "We must bring all the cases where two different species live on or in one another under a comprehensive concept which does not consider the role which the two individuals play but is based on the mere co-existence and for which the term symbiosis [symbiotismus] is to be recommended"
>
> - Albert Bernhard Frank, 1877[2]

Symbiosis in retrospect

The term "symbiosis" has been attributed to the scientist Anton De Bary, who used the word for the first time in 1878 in an address with the title "The phenomena of Symbiosis" [2]. However, as shown in the quotation at the beginning of this chapter, Albert Frank had in fact already suggested a similar concept. Similarly to Frank's suggestion, De Bary defined symbiosis as "the living together of unlike named organisms". Around the same time, the belgian zoologist Pierre-Joseph van Beneden published a popular book, "Les commensaux et les Parasites" (translated to English as "Animal parasites and messmates"), in which he classified associations between "lower animals" and "higher animals" as either "parasitism", "commensalism" or "mutualism", depending on the character of the association [2].

van Beneden was thinking about relationships between different animals in his book. That "microorganisms" could also be found in close association with animals and plants had been known for some time, but the study of their functional roles had barely started. Among the first symbiotic systems described were the lichens; Notably, Simon Schwendener, who in 1868 proposed that all lichens were associations between a fungus and an alga, described this symbiosis as a "master-slave relationship". By the end of the 19th century, some examples of "cooperative living" involving microorganisms had in fact been demonstrated, such as the nitrogen-fixing "bacteroids" in the root nodules of legumes. However, the dominant view at the time was that microorganisms were parasites. Perhaps not so surprising, considering that this was also the time when eminent scientists like Louis Pasteur, Robert Koch and Joseph Lister had started gathering evidence for "the germ theory" (that diseases can be caused by "microorganisms") [3]. Important bacterial

pathogens like *Bacillus anthracis* (anthrax), *Staphylococcus, Neisseria gonorrhoeae, Salmonella typhi, Streptococcus* and *Mycobacterium tuberculosis* were all identified during the 1870s and early 1880s [2].

During the 20th century, the meaning of the term "symbiosis" has evolved in different directions, and has by some been reserved for associations beneficial for both partners [4]. Currently, the dominant opinion seems to be that the term symbiosis should apply to all kinds of associations [5], much in agreement with the original definition, and I will follow this definition in my thesis.

Today it is well known that chronic bacterial infections that inflict no evident harm on their eukaryotic hosts are everywhere. It is also known that the nature of the associations between bacteria and their hosts differs widely, making it increasingly difficult to make generalizations and classifications. Even so, several terms and classification schemes are in use, and I will describe the most common ones in this chapter, while at the same time introducing some examples of what a symbiotic association may look like.

The good, the bad and the indifferent

One intuitive way of classifying symbiotic associations is to focus on the cost-benefit character of the association. In fact, the terms used by van Beneden in his book from 1876 are still in use, and are also employed to describe the symbiotic associations involving bacteria.

A parasitic association is used to define the situation where a bacterium preys on its host, imposing a fitness cost. As an example, consider the bacteria that live on your teeth. This community consists of an estimated 1000 distinct bacterial species, some of which are likely transient, while others may form more stable associations [6]. Some of them have the ability to cause dental caries (tooth decay) by producing acid from dietary carbohydrates. Not exactly lethal, though if you postpone your visit to the dentist for too long your teeth will probably fall out, as was indeed common in past centuries.

A commensal is defined as a bacterium that does not inflict any notable harm on its host, nor does it provide any benefit. Take the bacteria on the palms of your hands. A classic teaching experiment is to provide a class of students with a set of agar-plates (solid nutrient plate, on which bacteria can grow), and ask them to put their fingers on some of the plates. More often than not, those plates that were touched will develop a diverse collection of bacterial colonies, to the horror of the unsuspecting students. Regardless of personal hygiene routines, it has been firmly established that bacteria live naturally on our skin, and may indeed be very numerous, particularly in the more humid places such as groin and armpits [7]. Whether they do anything useful there or not is still open to discussion, and various beneficial functions

have in fact been proposed [8]. However, to my best knowledge there are currently no convincing studies that have found any evidence of beneficial effects beyond correlations of various kinds. Likewise, the normal microbiota (= all microorganisms present in a habitat) found on our hands does not seem to do any harm either. So, until further notice, we may tentatively classify these bacteria as “commensals”.

Symbiotic bacteria can also be classified as “mutualists”, if they provide a benefit to their host. Sticking with ourselves as example organism, some of the bacteria in our gut likely fall into this category. The gut of a healthy adult human contains an astounding number of bacteria, roughly estimated to outnumber the human cells ten to one [7]. Large-scale investigations of these bacteria have begun quite recently, but it is already clear that the human gut microbiota is not a random collection of bacteria [9]. One known role of the gut microbiota is to contribute to the digestion of complex carbohydrates in the large intestine [10]. More recently, studies have started to emerge, which indicate that these bacteria also interact with our immune system [11]. So at least some of the bacteria in our gut can probably be classified as mutualists.

A word of caution might be appropriate here. The idea that the bacteria in our gut may be beneficial is not novel. More than 100 years ago, both Elie Metchnikoff and Henry Tissier had started to observe correlations between health and gut microbiota composition, and suggesting dietary supplements of bacteria [12]. However, when it comes to healthy human subjects, it is not straightforward to demonstrate a health benefit (!) from supplemented bacteria. Particularly, the prophylactic use of bacteria (i.e. supplying bacteria in order to prevent future disease) is problematic. In this category, we can include tablets of lactic acid bacteria taken in preparation for your vacation, and a large number of dairy products supplemented with presumably health-promoting bacteria. For these bacteria to provide a health benefit, they need to survive all the way from initial production, to super-market, to your home, through your stomach (which has a very hostile pH), and down into your gut. If they make it that far, they will have to compete with the resident population to survive, and finally provide some sort of benefit not already present. Thus, there is a long way from *in vitro* experiments in the laboratory to *in vivo* applications. In fact, the European Food Safety Authority (EFSA) has rejected thousands of health claims from the food industry, based on lack of sound experimental evidence [13]. On the other hand, there is a general agreement in the scientific community that these bacteria are most unlikely to be harmful. So if you are fond of your "microbe"-yoghurt, by all means carry on!

Returning to the cost-benefit type of definition for bacterial symbionts, some bacteria can rightfully be considered pathogens. Since these understandably tend to be subject to extensive research they are perhaps in the least need of an introduction.

The cost-benefit categorization is straightforward to understand and describe in text, but can be challenging to apply in practice. Firstly, we have a very limited knowledge of the function of many symbiotic bacteria, the gut microbiota being an excellent example. Secondly, the cost-benefit result of a bacterial encounter may differ, depending on the circumstances; A bacterium may be harmless in one host, and parasitic in another, and other environmental factors may also affect the net result. As for commensal symbiotic bacteria, one may ask whether such associations really exist, or simply reflect our patchy knowledge of bacterial ecology. Thus, some scientist use the term "commensals" to describe all symbiotic bacteria that are not known to be harmful to their host.

Intimacy and dependency in symbiosis

Bacterial symbionts vary tremendously in the intimacy of the association they form with their hosts. The bacteria I have described so far in this chapter, on our teeth, skin and gut, are all extracellular. They are attached to the various body surfaces, some of them rather tightly, but they don't actually invade our cells. Extracellular symbionts are sometimes further classified as "ectosymbionts" and "endosymbionts", to distinguish between those bacteria living on the outer surfaces of their host, or inside their host (although in some cases it is not completely straightforward to define when a specific body-part should be considered "inside" or "outside"!).

Some bacterial symbionts are intracellular, living inside the cells of their hosts. Many pathogens are known for their ability to invade host tissues and sometimes host cells too. But intracellular bacteria need not be harmful. At the extreme end, bacteria can even be crucial for the survival and reproduction of their host. A classical example of such a symbiosis can be found in aphids. Nearly all of approximately 4000 described species of aphids harbor a group of bacteria collectively known as *Buchnera aphidicola* [14]. Aphids feed exclusively on phloem sap from plants, which is a rather nutrient-poor food source. Phloem sap contains a lot of sugar, but many essential amino acids (the building blocks of proteins) are completely absent. *B. aphidicola* compensates for this deficiency by synthesizing amino acids that the insect would not otherwise ingest. In return, the bacterium receives a stable habitat and nutrients from its host. *B. aphidicola* lives inside specialized host-derived cells called "bacteriocytes", which together form an organ called the "bacteriome". In fact, it never leaves the host, but is transferred directly from mother to off-spring (vertically) via the germ-line cells.

Thus, bacterial symbionts can also be broadly classified as heritable and non-heritable, where *B. aphidicola* is a classical example of a heritable bacterium. However, vertical transmission need not be connected to the germ-line cells. In the stinkbug *Megacopta punctatissima*, the transmission of an

essential gut symbiont is ensured by the deposition of a small fecal capsule on the underside of the egg [15]. When the stinkbugs hatch, they start their life by ingesting the capsule, thus ensuring that the symbiont is not lost.

Non-heritable bacterial symbionts can be acquired from the environment, for example through physical contact or diet (horizontally). That is not to say that such bacterial acquisitions are completely random. Take the Hawaiian bobtail squid, *Euprymna scolopes* [16]; The bobtail squid is nocturnal, coming out at night to hunt for prey. To avoid being predated itself, it emits a downward light, which masks the shadow otherwise being cast by moonlight. The light, while being directed and regulated by the squid, is produced by a bacterial symbiont, *Vibrio fischeri,* which is housed in a specific light organ. *V. fisheri* is obtained in juvenile squids directly from seawater, which is teeming with all kinds of bacteria. In fact, the density of *V. fisheri* in seawater has been estimated to be around 0.1% of all the bacteria. Yet, the juvenile squids get colonized with the proper bacterium within hours of emergence.

Naturally, regardless of the intimacy of the association, it is more likely that you will acquire a bacterium from a close family member rather than from a complete stranger. Indeed, many of the bacteria in our gut are acquired at birth [17]. However, the term "heritable" bacterial symbiont normally refers to those which are inherited vertically in a very strict manner, directly in or on the embryo between generations.

As indicated by the various examples listed here, a host may be more or less dependent on its symbiotic bacteria. Therefore, another common way to classify bacterial symbionts is based on the level of this dependency. Bacteria which are required to support normal host growth and reproduction are called "obligate", while those which are not are referred to as "facultative" [18]. According to the examples given so far, *B. aphidicola* and the stinkbug symbiont are both obligate, since experimental removal of the symbiont results in abnormal host development [14, 15]. In contrast, the bobtail-squid symbiont *V. fischeri* is not truly "obligate", since healthy squids without *V. fischeri* can be raised in the laboratory [19]. However, uninfected bobtail squids have not been found in nature, and it follows that term "facultative" does not translate into "unimportant".

At this point, perhaps you have noticed that the various classification schemes of bacterial symbionts in the previous paragraphs are based on a host perspective. Confusingly, some of the same terms are also applied from the perspective of the bacterial symbiont. Thus, the bacterium *Wolbachia*, which will be described in more detail in chapter 6, is commonly described as an obligate intracellular bacterium, even though most hosts of this bacterium will cope fine in its absence. "Obligate", in this case, refers to the fact that *Wolbachia* is completely dependent on its host for survival and replication. This is actually a common feature of heritable symbionts. In other words, heritable symbionts, whether "obligate" or "facultative", are "obli-

gately symbiotic"(!) since they do not appear to have a dormant or replicative phase outside their hosts [18].

About biology and definitions

Biologists tend to have strong opinions when it comes to definitions. Anecdotally, at my licenciate defence, the very first comment I received from my opponent was a criticism of my introduction to the term symbiosis (which was considered to be too narrow). I have made an attempt to broaden my view in this chapter, but the fact remains that most of the definitions have fuzzy borders, and some are used in different contexts.

However, all of the definitions introduced in this chapter are potentially interesting to explore, as they touch upon many important aspects of symbiotic associations. In fact, the borderline cases where the terms become difficult to apply are perhaps the most interesting ones to explore. For example, the definition of symbiosis by Frank and De Bary did not specify the required closeness of the association; Arguably, very few bacteria live in complete isolation, so where should we draw the line between symbiont and free-living?

Chapter 3: Communities and niches

The observation that living organisms have distinct distributions in nature and interact with their environments is probably as old as the field of biology itself. But "the niche concept" has an evolutionary history of its own. In this chapter, I will give a brief historical overview of the niche concept, and discuss how the term connects to current microbiology.

The niche concept

As an introduction to this chapter, lets start with a simple definition:

> "An ecological niche consists of the conditions necessary to support the vital activities of a type of organism"
>
> Alley, Thomas (1982) [20]

Sufficiently vague to leave room for interpretation, while at the same time giving some idea as to what we are talking about! But what exactly is meant by "conditions"? Let's take a step back. First, it is useful to realize that all organisms have a "habitat", meaning a specific location in the environment in which they live. This place will have both abiotic and biotic features. There will be a certain temperature, humidity, elevation and pH, it may rain a lot or be very dry. Biotic features include all other organisms that occur in the habitat. Some of these organisms may be a source of food, others perhaps predators. Furthermore, other organisms may also shape the physical characteristics of the habitat; Trees may provide shelter, grazing animals will keep the plants from growing wild, and so forth. An organism living in a specific habitat will need to cope with all these factors. Therefore, the habitat says something about the lifestyle of a species, and can potentially provide important clues for understanding the function and evolution of the species.

In the early 20th century several versions of "the niche concept" started to emerge, which have had a profound influence on the thinking of biologists. Joseph Grinnell, who in 1924 was probably the first to use the word "niche" in an ecological context [21], defined the niche as the "ultimate [distribu-

tional] unit .. occupied by just one species or subspecies" [22]. Grinnell was concerned with the abiotic factors that define the distribution of a species, what we may call a "place niche". A few years later, Charles Elton made a contribution to the discussion, by defining the niche as an organism's "place in the biotic environment, its relations to food and enemies" [23]. Thus, Elton put more emphasis on the biotic features shaping the habitat of an organism, commonly referred to as a "functional niche".

Taking a starting point in the definitions by Grinell and Elton, we could take a pragmatic approach and define the niche as consisting of "all the relevant biotic and abiotic factors describing the habitat of a species". However, George Hutchinson, who is rated as one of the most important ecologists of the 20th century [24], took the niche concept to a new level in the 1950es [25]. Apart from changing the niche concept into something that could be quantified, Hutchinson made a distinction between the "fundamental niche" and the "realized niche". Briefly, the fundamental niche describes the places and conditions under which an organism can *potentially* survive, and is limited mainly by the morphological and physiological characteristics of an organism. The "realized niche" on the other hand is a narrower version of the "fundamental niche", taking into account the presence of predators and competitors, which will tend to limit the actual distribution of a species. Consequently, the "realized niche" can vary from place to place, depending on for example the presence of predators.

There is another important point to be made about Hutchinsons definition of the niche concept: Species have niches, environments do not. Consequently, the niche concept and the species concept are intimately linked.

The competitive exclusion principle

The idea that all species have a specific "place" in nature was far from novel; Among others, Charles Darwin was very influenced by this idea. During the first half of the 20th century, several scientists started to formulate this idea in a more specific way, eventually giving rise to what is now known as "the competitive exclusion principle" [26]. Briefly, the competitive exclusion principle can be stated as follows: "Complete competitors cannot coexist" [26]. Translated into plain English, two non-interbreeding populations cannot co-exist if they are functionally identical.

The first scientist who attempted to provide experimental evidence for the competitive exclusion principle was Georgy Gause. Gause performed an experiment where he placed two species of closely related protozoans in a flask containing a single bacterial culture as food source, thus forcing the two species to share "the same niche". Within a couple of days, one of the species invariably took over the culture. Gause concluded that two species

with similar ecology cannot live together in the same place, due to competition between them [27].

If we take Gause's version of competitive exclusion at face value, we would expect to find a unique combination of biotic and abiotic features defining habitats for all species. In other words, although we might go out in nature and find two organisms apparently living in the same place, we predict that a more detailed description of their niches should enable us to explain their co-existence. As such, the competitive exclusion principle does not represent a testable hypothesis, but is perhaps better understood as a conceptual model [26].

Gause's experiment, although considered a classic, was criticized by many scientists from the beginning for being an over-simplification of nature, to the point of being meaningless [26]. Natural environments tend to vary in both space and time. Spatially, in the sense that many "micro-environments" may exist side by side, with slightly different living conditions. Temporally, environments may vary with seasons, natural disasters, invasions, diseases etc. When taking such parameters into account, it is not at all apparent that competitive exclusion should occur, based on "niche overlap" [20].

Practical issues with the niche concept

One challenge connected to the description of the “niche” of a species is to decide which of all the abiotic and biotic features are most relevant. In principle, one could go on adding features to describe the niche of a species nearly *ad infinitum,* but, besides being impractical, it belies the fact that some features are obviously more important than others. In order to understand the biology of a species, we need, in addition to a reasonably detailed description of the niche of a species, to get some idea about the relative importance of different features of that niche.

From a more practical point of view, the description of a niche is also complicated by the fact that living organisms are not static, they tend to move around, whether on their own volition or via other organisms. A specific organism may well be a superior competitor in its natural niche, but that does not necessarily stop it from occasionally probing other places. Ideally, we would like to exclude such places where an organism might end up by accident. What we are looking for is in some papers referred to as “the primary niche”: The place(s) and condition(s) to which our organism of interest has *adapted*.

Competition and cooperation in the microbial world

Hutchinson defined the “realized niche” as being restricted by predators and competitors, which probably conjures images from TV-programs *a la* David Attenborough in your mind. But what does “predators” and “competitors” mean in the context of microbiology?

Bacteria, like animals and plants, may compete for food, and one could argue that all bacterial competition eventually boils down to this problem. However, it is becoming increasingly clear that bacteria have evolved many sophisticated mechanisms to secure food requirements [28].

“Scramble competition” refers to the situation where bacteria compete for the same resource based on their ability to efficiently take up and use the resource. Apart from differences in up-take mechanisms and metabolism, many bacteria are motile. Bacteria swim, twitch, glide and slide, and may even move in groups (swarming) [29]. Furthermore, they differ widely in their ability to do so. Some kinds of motility are connected to “chemotaxis”, i.e. the ability to detect and move towards (or away from) specific chemoeffectors, such as a good food resource.

However, bacteria may also take more direct action by attacking their competitors, termed “contest competition”. Many bacteria can produce antimicrobial compounds, which may specifically inhibit the growth of competitors, thus securing the preferred food source.

Bacteria sometimes function as groups, rather than individuals. A typical example is the formation of "biofilms" [30]. A biofilm is an aggregated population of bacteria, which are immobilized and covered in a secreted extracellular matrix. When bacteria grow in this manner, they often have an increased survival as compared to when they are on their own [31]. Many pathogens grow in biofilms, and can be very difficult to eliminate with antibiotic treatment for the same reason. However, symbiotic bacteria in the gut also make biofilms, and in this case are likely contribute to the wellbeing of their host by keeping un-welcome colonizers away. Thus, another way to secure food sources can be to block the competitors out.

Natural biofilms often consist of multiple distinct species, which may also cooperate [31]. Some bacteria depend on other bacteria for colonization of surfaces. Others cooperate metabolically, where a nutrient is metabolized by one species and the waste product becomes the nutrient of another, all within the same biofilm.

Furthermore, bacteria can communicate with each other by a mechanism called “quorum sensing”, which was first discovered in *V. fischeri* some 30 years ago [32]. The bacteria export compounds akin to pheromones into the environment, and have specific proteins to detect these compounds. When the concentration reaches a certain threshold (a “quorum”) a group-response it triggered, in the case of *V.fischeri* they start to emit light. But quorum sensing can also elicit a myriad of other responses, such as for example the

production of antimicrobial compounds [33]. Thus, bacterial competition can be a cooperative effort!

Predators from a microbial perspective

A large fraction of all microbial eukaryotes have a diet consisting of bacteria, and are likely to exert a significant selection pressure, at least in some ecosystems [34]. Some bacteria are in fact predators themselves [35], and may even hunt in packs! [36]. Again, there is power in numbers, and bacteria in biofilms are typically more resistant to predation than they are on their own [37].

However, the main "predators" in the bacterial world are arguably the phages (the bacterial equivalent of viruses). Phage numbers easily exceed those of bacteria in many ecosystems, where they may cause significant mortality [38]. Even intracellular symbionts can be subject to phage attacks [39].

A typical phage life cycle involves adsorption to the host, injection of DNA into the host, replication and packaging of DNA into new phage particles and lysis of the host. Bacteria on the other hand have evolved their own mechanisms to fight off the phages. One of the most sophisticated mechanisms, termed “CRISPR” (clustered regularly interspaced short palindromic repeats), was discovered less than a decade ago, even though current estimates say that nearly half of all bacterial genomes encode them [40]. A CRISPR consists of a series of short repeats interspersed with unique “spacers” (small sequences of DNA matching phage or plasmid sequences). Upstream of the repeat region is a series of genes, which carry out the function of the system: to identify and eliminate all incoming DNA that matches any of the spacers. A single CRISPR region can contain hundreds of spacers, and effectively provides the bacterium with adaptive immunity. Usually, a complete match between spacer and target is required for the CRISPR system to function. Phages with a mutation in the target sequence therefore have a selective advantage, which may result in an “arms race”.

However, phage DNA can also be inserted into the host genome, rather than being packaged into new particles, and lyse the host at a later point, or eventually end up being trapped on the chromosome. Thus, many bacteria encode degenerate phage genes in their chromosomes. More radically, Polydnaviruses incorporated into parasitoid wasp genomes are used by the wasp to suppress the immune response of its host [41]. In contrast, the APSE virus, encoded on the genome of the facultative endosymbiont *Hamiltonella*, provides protection against parasitoids [42]. Finally, phages are vectors of horizontal gene transfer, and can therefore have a very direct effect on the evolution of bacteria as sources of genetic innovation [43].

Concluding remarks on bacteria and niches

Bacteria are capable of a wide range of interactions, ranging from complete warfare to competition and cooperation. Additionally, many bacteria live in dynamically changing environments, and are likely to be attacked by eukaryotic predators as well as phages. The relative importance (selection pressure) of these factors in natural environments has been little explored, and it is obviously a challenging question to address. However, it should be obvious that competitive exclusion is not the only force at work when it comes to interactions in microbial communities.

Chapter 4: The bacterial species concept

So far, in this thesis, I have used the terms "organism" and "species" somewhat interchangeably. Strictly speaking, an organism refers to an individual, whereas a species refers to a group of individuals. However, when it comes to bacteria, the concept of species is highly controversial. It is therefore with some trepidation that I write this chapter. My goal is to introduce the current debate, starting with a brief historical look-back, in order to set the stage for discussion.

Bacterial taxonomy then and now

As soon as bacteria were discovered, scientists began to name them, just as we name everything else. Initially, bacterial strains were named and classified based on observation, in the same manner as animals and plants. Morphological characteristics were noted, as well as their growth and survival under various experimental conditions in the laboratory. However, bacteria have a rather limited repertoire when it comes to morphology, and unlike eukaryotes do not have very informative fossil records. Therefore, while a common evolutionary origin of all bacteria was assumed, most microbiologists from the pre-sequencing times considered the bacterial tree of life as an impossible question to resolve [44].

Bacterial taxonomy entered a new era when Carl Woese began to make his catalogues of ribosomal RNA (rRNA) oligo-sequences in the early 1970s [44]. The ribosome is responsible for one of the most fundamental functions in any living system, namely the translation of RNA to protein, and is therefore both highly conserved (evolves slowly) and present in all living organisms. By comparing the short sequences of the 16S rRNA gene, and clustering them between samples, it became possible to construct some of the deeper branches of the bacterial tree of life. By the early 1980s, the Sanger sequencing method had been developed, and by the time of the new millennium the next-generation sequencing era had begun, shifting the focus from genes to genomes.

Nearly 3000 complete bacterial genomes are currently available in the public genome database on NCBI, with more than 13.000 additional genomes in various stages of completion. By analyzing and comparing genomes our understanding of what a bacterium is has changed. Bacteria replicate clonally between generations, but they are also capable of horizontal

DNA transfer and recombination between individuals. Thus, genomic diversity in bacteria is often characterized by a large variation in gene repertoire. The bacterial “pan-genome” refers to all the genes found in strains belonging to a specific “species” [45]. The gene repertoire of individual strains is commonly divided into the “core genome” (genes shared between all strains of a species) and the “accessory genome” (genes which are variably present in strains of the same species). To complicate matters further, horizontal gene transfers are not restricted to bacteria belonging to the same species, but can sometimes occur between strains otherwise considered to be very distantly related. It is primarily because of this promiscuity that the bacterial species concept is so debated.

Ideally, one would like to take ecological features into account when trying to delineate a species. However, deep sequencing of bacterial DNA extracted directly from a wide range of habitats (metagenomics) has shown us that the diversity of bacteria in nature is much larger than anyone had suspected [46], and the vast majority of these cannot be cultured in the laboratory [47]. More precisely, we don’t know how to culture them, and this situation is unlikely to change in the near future. Thus, in practice, we are currently forced to define bacterial species based on sequence data.

Ironically, despite the fact that we are now in the “genomic era” of microbiology, the 16S rRNA sequence is still widely used as a marker for species delineation in bacteria. A commonly used cut-off is a minimum of 97% sequence identity of 16S rRNA sequences between strains belonging to the same species. Sequencing a handful of "housekeeping" genes (multi-locus sequence typing) may also be employed to separate closely related groups. Some efforts have been made to take into account the complete genome information, for example a 70% cut-off in genome hybridization (as determined experimentally) or a 95% (or 99%) cut-off in average nucleotide identity (ANI) between genomes of different species.

Towards a biological species concept for bacteria

Frederic Cohan has been instrumental in advocating what we may call a bacterial version of the “biological species concept” (commonly employed for animals), which is based on reproductive isolation. Although this may sound counter-intuitive, since bacteria do not reproduce sexually like animals, Cohan argues that there are "quintessential dynamic properties", which bacterial and eukaryotic species share [48]. According to Cohan, a bacterial species (or rather, “ecotype”) should fulfill the following requirements:

1. It should fall into a well supported sequence cluster
2. It should evolve under “cohesive processes” within the species
3. It should be ecologically distinct from other species
4. It should be irreversibly separated from other species

Out of these, point one is arguably the most straightforward one to address with sequence data. Point two deserves a bit more explanation; Basically, the idea is that foreign genes or mutations, which are not beneficial for a species, will be removed by selection (individuals with such features will be “less fit”). In contrast, an advantageous mutation or novel gene could result in a “selective sweep” changing the entire population of a species. Thus, the model leans heavily on the force of periodic selection. Point three and four are directly connected to the "niche concept", as discussed in chapter 3.

While Cohan's version of the species concept is attractive in having a theoretical (and biological) basis, it does not really solve the issue of where to “draw the line”. How ecologically distinct should two populations be to qualify as species? And how do we estimate ecological distinctness based on sequence data?

Another issue is the question of whether periodic selection is sufficient to completely purge diversity from natural populations. In bacterial populations where recombination is less common than mutation, clustering is expected to occur based on genetic drift [49]. Thus, it is necessary, but not straightforward, to distinguish between transient and "irreversibly separated" sequence clusters.

Do bacterial species exist?

Comparative genomics has provided several interesting lessons about bacterial evolution, where the extreme diversity in genome evolution is perhaps one of the most remarkable. Some bacteria have huge accessory genomes, others do not. Some recombine frequently, others do not. This diversity represents another hurdle in the quest for a bacterial species concept; Is it possible to arrive at a definition that fits them all? And if not, had we better leave well enough alone?

At the extreme end, Ford Doolittle has posed an even more provocative question; Is bacterial diversity organized into discrete phenotypic and genetic clusters, or are such patterns simply a result of experimental biases and stochastic events? [48] Several studies have addressed this possibility, and most conclude that bacteria do form clusters of genetically related strains (e.g [46, 49, 50]), but the question of how to translate such clusters into species remains unresolved.

What is the future of bacterial taxonomy?

As a consequence of the controversy surrounding the species concept, many microbiologists have started to avoid using the word "species" altogether.

Particularly in metagenomic studies, words like "phylotype" and "OTU" (operational taxonomic unit) have started to replace the word "species", and it is not too hard to understand why.

At my first conference talk last summer, I felt compelled to include some sort of bacterial species concept definition to introduce my work on *Wolbachia*. After all, the species concept is a central question in the paper we have published, it is even in the title! (paper I). Knowing that I was likely entering "the lions den", I decided to use Cohan's four-point definition as a starting point for discussion, thinking that since the concept is not clearly delineated (no specific sequence cut-off), this shouldn't cause too much controversy. This naive view was shattered as soon as I finished my talk; My very first comment, from a prominent scientist in the audience, was a disagreement with the point that bacterial species should be irreversibly separated.

Having learned my lesson, I will not attempt to provide any conclusion as to what a bacterial "species" is at the end of this chapter. If you are not a biologist, you may be wondering what all the fuzz is about. Does it really matter? Do we need a "species concept" for bacteria?

The bacterial species problem is not merely a question of how we name bacteria, it is about understanding how they live and evolve. In principle, one could envisage a simple arbitrary "naming convention", without inferring that such names should translate into species. In practice, the way we name organisms tends to influence how we think about them. Even in studies where the word "phylotype" is used instead of "species", sequences belonging to the same "phylotype" are often analyzed together in a manner one would find reasonable only if they were in fact "species". Particularly in metagenomic studies, scientists are often forced to make a choice as to what represents biologically meaningful sequence clusters in the data.

Despite the next-generation sequencing "revolution", our understanding of how bacteria evolve and differentiate into different niches in natural environment is still very limited. We have taken big steps forward when it comes to answering the question "who is out there", but we have barely started on the next logical question, namely "what are they doing". With a better knowledge of how bacteria evolve and interact, it may be possible to arrive at a more appropriate and biologically relevant species concept for bacteria. It is early days for microbial ecology.

Chapter 5: Host-symbiont evolution

Evolutionary biology is essentially about understanding how different species are related to each other, and how they have adapted to their current lifestyle. These questions get a unique flavor when considering two or more species living closely together, in a symbiotic association. In this chapter, I will describe some of the remarkable findings that have been unraveled so far, concerning what the genomes of bacteria with this kind of lifestyle look like.

Co-diversification and co-evolution

Changes in the DNA occur continuously throughout the evolutionary history of a species, and the nature of such changes is at the heart of molecular evolution research. By comparing the DNA between species, we can make qualified guesses as to how they are related to each other. Two closely related species will have a recent common ancestor, and therefore have accumulated fewer changes than two more distantly related species.

What does the evolutionary history look like for species that live in symbiosis? Consider the case where a bacterium is vertically transmitted, like *B. aphidicola*, the aphid endosymbiont discussed earlier. If we construct a phylogeny of different species of aphids, we can predict that the phylogeny of their obligate symbionts will look similar. In other words, if two aphid species are closely related, so are the obligate symbionts that they carry. This pattern has in fact been observed repeatedly for obligate endosymbionts of insects, and is referred to as "co-diversification" [18]. However, the phenomenon is not restricted to insect-symbiont associations; As a remarkable example, *Helicobacter pylori*, which infects the human stomach, has an evolutionary history that mirrors human migrations in the past [51]. Overall, co-diversification between species that live together in symbiosis can be taken as evidence of a very close association.

However, the opposite is not true; There is a subtle distinction to be made between co-diversification and co-evolution. Let's return to the bobtail squid with the bioluminescent endosymbionts described earlier. In this symbiosis, the host and its symbiont do not display a clear pattern of co-diversification [52]. Yet, the bacterial symbiont has evolved specific mechanisms to facilitate colonization of the host, including modulation of host gene expression,

chemotaxis to find the right tissue to colonize, and quorum sensing to regulate bioluminescence production [16]. The host, on the other hand, has evolved mechanisms to control the growth of the symbiont, such as the expulsion of about 95% of all the bacteria every day at dawn [16]. Thus, it is possible to have co-evolution without co-diversification, particularly in cases of “open systems”, where the symbiont may spend part of its time outside the host, or change hosts frequently.

Reductive evolution and horizontal gene transfer

Bacteria which have evolved a lifestyle as intracellular endosymbionts are subject to distinct selective forces compared to free-living bacteria [53]. Firstly, the effective population size is reduced, as the bacteria are constrained to living inside their host. This in turn will make selection less efficient, when it comes to removing slightly deleterious mutations. The problem is exacerbated by the vertical transmission mode, which is likely to introduce severe population "bottlenecks" between host generations. Furthermore, those bacteria which are strictly confined to living inside host-derived cells will have limited possibilities for contact with other bacteria, and horizontal gene transfer is therefore reduced. Finally, the intracellular environment is typically very stable and rich in nutrients, removing selective constraints on genes that are no longer strictly required. Consequently, intracellular endosymbionts shrink continuously. In fact, the smallest bacterial genomes known belong to the obligate intracellular endosymbionts of insects [54].

Gene losses in intracellular endosymbionts are not random [54]. The most conspicuously absent genes in intracellular endosymbionts are perhaps those related to the formation of the cell envelope. Many intracellular endosymbionts do not have any cell wall at all, and are typically encased in host-derived membranes. Furthermore, intracellular endosymbionts tend to lose genes involved in regulation of transcription, perhaps due to the stable environment or to the host gradually taking control of the symbiotic interaction.

Intracellular bacteria can be facultative or obligate, in terms of their importance for host development (see chapter 2). Additionally, some bacteria are not strictly intracellular, in the sense that they are still able to grow as free-living bacteria too. In effect, a continuum in lifestyles exists, ranging from free-living to strictly intracellular, which correlates strongly with genome size [53].

As noted in chapter 4, bacteria are notorious for their ability to transfer DNA horizontally. Some bacterial symbionts have taken the process one step further; They can transfer DNA to the host too! *Wolbachia*, which will be introduced in the next chapter, has been implicated in many such transfers [55]. Intriguingly, the genome of the pea aphid, *Acyrthosiphon pisum*, did

not reveal any genes transmitted from its obligate endosymbiont *B. aphidicola*, but instead 12 other genes were identified which may have been transferred from *Wolbachia* [56]. Moreover, seven of these genes were found to be highly upregulated in the bacteriocytes of the aphid, where *B. aphidicola* resides! However, the functional importance of the genes is currently not known.

The cell envelope and the environment

Bacterial symbionts face a number of specific challenges compared to free-living bacteria. To begin with, they must evade the immune system of the host. The immune response of both vertebrates and invertebrates is designed to recognize bacteria based on outer surface-structures. Extracellular bacterial symbionts cannot afford to lose their cell walls, but instead display modifications of the cell envelope to avoid the host immune response. Furthermore, extracellular and cell surface components are of importance for attachments to other bacteria and host surfaces. Thus, the cell envelope of symbiotic bacteria is of great interest for the study of host-symbiont interactions.

Proteins involved in the attachment to a host ("adhesins") have been identified in many symbiotic bacteria [57, 58]. Some adhesins have very high molecular masses, and many contain numerous repeats [59, 60]. Extracellular appendages, like fimbriae and pili, can also be involved in attachment, and may modulate the host immune response [61]. Furthermore, outer surface proteins are frequently under diversifying or positive selection [62, 63].

The surface may also be more generally modified. Some bacteria produce a so-called "S-layer", which consists of identical protein subunits organized in a lattice-like layer that completely covers the surface of the bacterium [64]. Interestingly, many Lactobacilli encode several S-layer proteins, which are normally only expressed one at a time, thus potentially providing a mechanism for changing the entire surface based on environmental cues [64]. Other symbiotic bacteria modify their surface by producing an "exopolysaccharide capsule" (EPS), which can be described as a structurally varied sugar coating. Notably, in *Bifidobacterium breve,* this capsule has been shown to mediate immune response evasion, promote persistence in the gut *in vivo,* as well as protecting against colonization by pathogens [65].

Another common theme for symbiotic bacteria is the presence of proteins with domains otherwise mostly found in eukaryotes. For example, the intracellular endosymbiont *Wolbachia* is known to encode a large number of ankyrin-repeat domain proteins, which are ubiquitous in eukaryotes, but not at all common in bacteria. Similarly, we describe a family of RCC1-domain proteins in Bifidobacteria isolated from the bee gut in paper III, which is also comparatively rare in bacteria.

Most of what is known about interactions between eukaryotes and bacteria has come from the study of bacterial pathogens. However, it appears that symbiotic bacteria, whether pathogenic or mutualistic, share many of the same systems for interacting with their host [31, 58]. Which by the way really isn't a novel idea either:

> "One has a tendency to separate cases of "symbiosis" from cases of "disease" and to study them from completely different points of view. I will try to the contrary, starting with symbiosis, to understand disease...there is no absolute distinction to be made between these two orders of phenomena"
>
> - Noël Bernard, 1909 [2]

Chapter 6: *Wolbachia*

In this chapter, I will introduce the bacteria, which formed the starting point of my PhD. At the end of the chapter, I will also introduce the paper "*Wolbachia* and the bacterial species concept", which is based on a comparative genome analyses of these intriguing bacteria.

The discovery of a bacterial master manipulator of reproduction

Evidence for the essential role of microorganisms in insects started to accumulate in the 1920s, when Paul Buchner and others had begun to map their occurrence [2, 66]. It was also at this time, in 1924, that the bacterium later to be known as *Wolbachia* was first described, in "Studies on Rickettsia-like microorganisms in insects" [67]. The species was formally described and named in 1936 [68], after which it was largely ignored by the scientific community.

A few years later, a curious phenomenon was noted in mosquitoes – certain strains appeared to be incompatible with each other, in one direction; Males from "strain A" could mate with females from "strain B", but females from "strain A" did not produce viable off-spring with males from "strain B" [69, 70]. The phenomenon was further investigated in the following decades, and it became clear that the phenotype was maternally inherited.

But it was not until 1971 that a connection between *Wolbachia* and population incompatibility was proposed [71]. Compelling evidence was found a few years later, when it was shown that a simple antibiotic treatment could restore normal reproduction [72]. In the following years, the same phenomenon was described in other arthropods, and even today the list continues to grow. The prevalence of *Wolbachia* is not known, but current estimates say that more than 40% of all arthropods are likely infected [73-75]. Thus, *Wolbachia* is a prominent candidate for being the most widespread and successful endosymbiont of the insect world. Furthermore, *Wolbachia* has also been found to infect many filarial nematodes, where in some cases *Wolbachia* is an obligate endosymbiont [75].

The reproductive manipulation initially found in the mosquitoes is today known as "cytoplasmic incompatibility (CI), and has been described by a

"modification-rescue" model [76]. Although *Wolbachia* are not infecting mature sperm, their presence in male testes appears to modify the sperm in such a way that normal reproduction is unsuccessful. However, if the female is infected with an appropriate *Wolbachia* strain, she can "rescue" the modification, so that reproduction can occur normally. Infected females can also reproduce normally with uninfected males, and are therefore predicted to have a reproductive advantage in infected populations. Furthermore, when distinct *Wolbachia* strains are involved, the incompatibility may become bidirectional, in the sense that an infected female cannot reproduce successfully with a male infected with "the wrong" *Wolbachia* strain.

Cytoplasmic incompatibility is only one out of several ingenious manipulations in the repertoire of the *Wolbachia.* Other manipulations described include induction of parthenogenesis (virgin birth), male killing and feminization of male off-spring [77]. Interestingly, it has been shown in a number of studies that the same *Wolbachia* strain can cause both male killing and cytoplasmic incompatibility, depending on the host, indicating that the mechanisms of these manipulations are related [78-80]. However, despite nearly four decades of research and several complete *Wolbachia* genome sequences, the molecular factors responsible for the phenomenon remain elusive.

While *Wolbachia* is arguably the most studied reproductive manipulator of insects, the lifestyle is not unique. Several other bacteria, such as *Cardinium hertigii*, *Spiroplasma*, *Rickettsia* and *Arsenophonus nasoniae*, can cause similar phenotypes, and their scattered distribution in the tree of life indicates that this lifestyle has evolved repeatedly [18].

Current knowledge on the biology of *Wolbachia*

During the last decade, it has become apparent that *Wolbachia* are not merely reproductive manipulators.

One of the most exciting new phenotypes described for *Wolbachia* was actually found by accident [81]. A group of researchers were aiming to identify genes involved in virus resistance in *Drosophila melanogaster.* They started their study by generating mutant fly lines (P-element insertional mutagenesis), and the plan was then to screen for mutant lines with increased virus sensitivity, due to interruption of vital genes. But instead they found that most of the mutants had a higher virus resistance compared to the non-mutated control line. It turned out that an antibiotic treatment had been carried out on the control line previously, whereas the mutants were infected with *Wolbachia.* Incidentally, the same discovery was also made by another group around the same time [82].

These studies caused a major shift in the *Wolbachia* research field. Mosquitoes, which are vectors of many severe human pathogens, can have in-

creased resistance to dengue, Chikungunya, yellow fewer, West Nile viruses and malaria when they are infected with *Wolbachia* [75]. Field projects testing the use of *Wolbachia* to control dengue are currently ongoing, and other projects are being planned. Intriguingly, despite the widespread infection of *Wolbachia* in arthropods, it appears that the major vectors of human diseases, including *Anopheles* (vector of malaria), do not naturally harbor *Wolbachia.* Therefore, the hosts must first be transinfected with *Wolbachia,* which is not at all straightforward. However, the first successful transinfection of an *Anopheles* mosquito was recently published, giving new hope for the use of *Wolbachia* to control malaria [83].

Other studies have documented that *Wolbachia* may influence the fecundity of its host. A four-fold increase in fecundity was estimated in *Drosophila mauritiana* when infected with its native *Wolbachia* strain *w*Mau compared to being uninfected [84]. Similarly, during the invasion of *D.simulans* in California by the *Wolbachia* strain *w*Ri, the interaction changed from a fecundity cost to a slight benefit in the scope of 20 years [85].

That *Wolbachia* can also function as a nutritional symbiont has been shown in the bedbug *Cimex lectularius*, where it appeared to supplement the host with vitamin B [86].

In conclusion, it seems safe to assume that our knowledge of the functional roles of *Wolbachia* in nature is very far from complete, and more examples of ecological roles are likely to emerge. Incidentally, *Wolbachia* is also an excellent example of how a bacterium can be both a “parasite” and a “mutualist”, sometimes simultaneously!

Wolbachia in the fruit fly, *Drosophila simulans*

The fruit fly has been a "model organism" for research on insects for a very long time, in part due to the ease of maintaining the flies in the lab. Consequently, an enormous amount of research is available, concerning everything from evolutionary history to brain development. The fruit fly is also something of a model in *Wolbachia* research, where the interaction has been studied since 1986 [87]. *D. simulans* is known to be the host of five distinct *Wolbachia* infections, which differ phenotypically (see Table 1 for an overview).

Table 1: *Wolbachia* strains associated with *D. simulans*

Strain name	Discovery	Virus protection	Super-group	Reproductive alteration
*w*Ri	California	Intermediate	A	CI
*w*Ha	Hawaii	None	A	CI
*w*Au	Australia	High	A	None
*w*No	New Caledonia	None	B	CI
*w*Ma	Madagascar	ND	B	None

Firstly, based on crossing experiments under controlled laboratory conditions, it has been shown that the *Wolbachia* strains associated with *D.simulans* differ in their ability to cause CI, as well as the level of CI induced [88]. Three of the five strains, *w*Ha, *w*Ri and *w*No, are known to induce CI, and all three strains are bidirectionally incompatible. Furthermore, although the *w*Ma strain does not appear to cause CI, it has retained the ability to rescue CI caused by the *w*No strain [76]. While *w*No causes a moderate loss of embryos in incompatible crosses, *w*Ri can render the host nearly completely sterile.

The strains also differ in their infection density. This characteristic appears to correlate with another divergent feature, namely their ability to provide virus protection to the host [89]. Intriguingly, one study has indicated that some of the strains have distinct tissue distributions at certain time points in the development of the fly [90], but the functional importance of this phenomenon is not known.

Most of the strains occur as single-infections, but strain *w*No is nearly always found in a double-infection with strain *w*Ha. In contrast, *w*Ha also occurs as a single-infection.

Finally, the history of the association between *Wolbachia* and *D. simulans* has been investigated. The *Wolbachia* are inherited maternally through the germline cells, a feature they have in common with mitochondria. Therefore, although phylogenetic studies indicate that the *Wolbachia* occasionally get transferred horizontally, there tend to be a strong correlation between specific *Wolbachia* infections and mitochondrial haplotypes at shorter time-scales. *D. simulans* is known to have three distinct mitochondrial haplotypes, called type I-III. Each of the five *Wolbachia* strains only occur together with one specific haplotype, and the geographic distribution of these have been investigated, with the aim of reconstructing the evolutionary history of the fly. By combining these data with *Wolbachia* sequences, a number of studies have suggested that the *w*No strain is the oldest infection in *D.simulans,* while *w*Ri and *w*Au are more recent infections [91-93]. Interestingly, while *w*No has a quite limited geographic distribution, *w*Ri and *w*Au appear to be spreading around the world, indicating that the novel infections are more invasive [85, 94].

Additionally, there is also evidence to suggest that the double-infection with *w*Ha and *w*No is not novel. A very similar double-infection has been identified on the Seychelles, in the *D. simulans* sister species *Drosophila sechellia,* suggesting that the infection preceded the speciation between the hosts [95]. Combining this information with data from mitochondria and *Wolbachia,* a scenario has been proposed where the double-infected flies migrated from the Seychelles to the eastern IndoPacific, and subsequently moved to other islands where the *w*No infection was lost. If that is correct, this association is more than 200.000 years old, and begs the question of how

the host-sharing has influenced the evolution of the genomes of these symbionts.

Wolbachia genome evolution

At the time I started my PhD, only four *Wolbachia* genomes had been sequenced. However, in recent years more genomes have been added, with about 20 genomes currently available in various stages of completion [75].

Based on both MLST and comparative genomics, it has been shown that *Wolbachia* strains recombine extensively [96]. The mechanisms involved have not been established, but the high level of transposable elements encoded on the genomes could be involved [97]. Furthermore, prophages may be involved, since arthropod-associated *Wolbachia* harbor prophages [98], and phage particles have been isolated from several *Wolbachia*-infected insects [39]. Adding to that the frequent occurrence of multiple *Wolbachia* infections within hosts, and it would seem that these strains have the potential for gene innovation and intra-strain diversification, despite being intracellular endosymbionts.

The most debated group of proteins described for *Wolbachia* is arguably the ankyrin-repeat domain proteins, which are typically found in many copies in the genomes, particularly in arthropod-associated *Wolbachia* [96, 99, 100]. Given that proteins with such domains mediate protein-protein interactions in eukaryotes, they have long been prime suspects for the induction of reproductive manipulations. Indeed, an ankyrin-repeat domain protein, *ankA,* in *Anaplasma phagocytophilum* was found to be an effector protein secreted by the type IV secretion system, and thus involved in the host-pathogen interaction [101]. However, the function of these genes in *Wolbachia* is still unknown.

Taxonomy of *Wolbachia*

Wolbachia constitutes a highly diverse group of bacteria, with new strains continuously being added. As a group, *Wolbachia* are monophyletic within the alpha-proteobacteria, and are not closely related to other members of the *Anaplasmataceae*. However, within the group, several distinct phylogenetic clusters exist, termed "supergroups", which have been determined by multilocus sequence typing [77]. The possible designation of some of these groups into separate species has been debated [102], but no consensus has been reached. Originally named *Wolbachia pipientis* [68] (after the mosquito host where it was first described, *Culex pipiens*), most papers now refer to this group of bacteria simply as *Wolbachia*, in recognition of the unresolved taxonomic status [77].

In insects, most of the described *Wolbachia* strains belong to super-group A and B, which are on average approximately 97% identical at the 16S rRNA sequence. These supergroups are particularly intriguing, because they have overlapping host ranges, and often can be found in double-infections. Furthermore, members of both groups induce CI, and no group-specific functional differences have been described. However, in paper I, introduced in the next section, we suggest that there may in fact be such differences, based on a comparative genome analysis.

Paper I: *Wolbachia* and the bacterial species concept

This paper describes the results of a comparative genome analysis on *Wolbachia,* with a specific focus on supergroup comparisons. We report the sequences of two new *Wolbachia* strains, *w*Ha and *w*No, both isolated from the fruit-fly *Drosophila simulans*. The two strains belong to different supergroups (A and B), but are likely to have history of co-infecting the same host. Thus, if *Wolbachia* strains are prone to recombine and exchange genetic material across super-group boundaries, we would expect to find evidence of such activity in the genomes of these strains.

Our main finding in this paper was in fact a negative result: the genomes of the 6 strains in the study showed little evidence of recombination between supergroups. Recombination between super-groups was largely restricted to small gene fragments and phage-related genes. Notably, the two strains with a presumed history of co-infection did not display any evidence of convergence compared to the other strains studied. In contrast, we could confirm that the strains recombine extensively within both supergroups.

Which left us with a difficult question: What is the reason for the separation of the groups? Aside from co-infections in nature, recombination in phage-related genes also indicates that the strains do get in contact with each other, so something else than direct physical separation must be at play.

It is tempting to speculate that the strains have evolved different niches, and thus fulfill different roles within their host. In support for this idea, we found 33 and 24 protein families, which were conserved within each supergroup but not shared between them. Most of these genes were of unknown function, but some were good candidates for host-symbiont interaction, including outer-surface proteins in supergroup B, and *fic*-domain proteins in supergroup A. If the strains have evolved different micro-niches, this could in itself provide a barrier to recombination, since such events could be selected against (see chapter 4).

Alternatively, sequence divergence between the supergroups could provide a barrier, since bacteria overall tend to recombine less between divergent strains. Counting against this hypothesis, the recombination events that we identified between supergroups were not biased towards the less diver-

gent genes. Furthermore, if sequence divergence is the cause of the reduced gene flow between the supergroups, how did the strains diverge in the first place? Were the strains perhaps physically isolated in the past?

Regardless of the reasons, it appears that supergroup A and B have become irreversibly separated. This result calls for increased attention to the possibility of distinct biological functions of the super-groups. As a starting point, both proteomics and microscopy are promising technologies for investigating the expression of the group-specific genes, and are within reach now that we have the complete genome sequences.

Chapter 7: Working with uncultured endosymbionts

By this point in the thesis, I hope to have convinced you that bacterial endosymbionts are incredibly fascinating. However, from a research perspective there is a catch; Due to the sometimes very intimate associations between bacterium and host, they typically cannot be cultured on their own in the laboratory. This in turn means that it is very difficult to conduct experiments on them. However, things are looking up for endosymbiont research, as new technologies are continuously being developed. In this chapter, I will introduce one such methodology that I have employed myself.

The problem with DNA preparation for sequencing

While sequencing technologies continue to improve, with longer reads and more data, there seem to be less focus on the first step of the process, the preparation of DNA for sequencing. Currently, sequencing technologies by and large require DNA in the range of micrograms, which translates into very large amounts of bacteria. Without cultivation, such quantities are very difficult of obtain.

This was the first major problem I faced during my PhD. I wanted to do genomics on the endosymbiont *Wolbachia*, which cannot be cultured, and in some cases infects its host at a quite low density. While purification of symbiont DNA from the host DNA is often achievable, elaborate purification protocols also cause extensive loss of DNA. Therefore, in practice, there is a trade-off between the purity of the DNA and the quantity. To obtain sufficient DNA for sequencing, previous *Wolbachia* genome projects had used large amounts of insects for DNA extraction, and mainly been restricted to such infections with a relatively high infection density. Apart from being a very time-consuming process, the use of high numbers of insects also increases the risk of sequencing a non-clonal population. Thus, the genome of the *Wolbachia* strain *w*Pip was never closed, probably due to the presence of distinct isolates within the sequencing data [100].

Culturing bacteria is in fact not an ideal solution either. Culturing represents a highly artificial environment for most bacteria, and may therefore induce the loss of plasmids and genes. Thus, in all cases, the ideal would be

to isolate a single bacterium directly from its habitat and sequence the genome of the individual, rather than a population of bacteria.

The multiple displacement amplification method

Due to a number of coincidences, I became familiar with a method called multiple displacement amplification (MDA) very early during my PhD. MDA is a bit like a PCR reaction, it can be used to amplify DNA [103]. However, in contrast to PCR, random short primers are used, so that in principle any kind of DNA will be amplified. A special polymerase is employed, which has the capacity to do "strand-displacement", thus eliminating the need for thermal cycling, and maintaining the productivity of the polymerase.

It therefore occurred to me that it should be possible to isolate my target bacterium, *Wolbachia*, and subsequently amplify the DNA. It was in fact not a novel idea, some attempts had already been published, and I remember being particularly fascinated by the work of Hongoh *et al.* [104], who had managed to sequence the genome of an endosymbiont infecting a single protist cell infecting a termite gut! However, there were two outstanding questions; How to isolate my *Wolbachia* cells? And, can amplified DNA really be relied on for genome assembly?

MDA was known at the time to generate two kind of biases, which were of concern for genome assembly, namely amplification bias and chimera formation. Amplification bias refers to the uneven representation of the amplified DNA relative to the original DNA [105]. Chimera formation refers to the joining of two sequences of DNA during amplification, giving rise to an artificial sequence, a "chimera" [106]. However, at the time, the importance of these biases for genome assemblies had not been addressed, nor the extent of these biases under different conditions.

Despite these potential issues, I was given the "go-ahead", and in due time ended up with two *Wolbachia* genomes to analyze, the results of which can be read in paper I in this thesis.

Paper II: Testing the reproducibility of Multiple Displacement Amplification

A slightly unsatisfactory aspect of our comparative genomics publication on *Wolbachia* was that we had to omit several details concerning the protocol employed. For example, we had observed a clearly non-random trend of amplification bias, which was at odds with the majority of the literature. In the paper, this observation was reduced to a single sentence and a supple-

mentary figure. Furthermore, although we had gone to great lengths in trying to estimate the reliability of our genome assemblies, the question remained somewhat open. And how generally applicable were our results? Were we just lucky? And was the non-random bias somehow unique to *Wolbachia*?

This formed the motivation for the study published in paper II. *Bartonella australis* was chosen for the experiment by fulfilling three basic requirements: It had previously been sequenced in our lab (so we trusted the assembly!), it could be cultured on a plate, and it was in our freezer. MDA was employed on two dilution series of cells and extracted genomic DNA respectively, going all the way down to the point where we could estimate that the target genome was partially present, based on PCR on a few loci. We also included a re-sequencing of the extracted genomic DNA, in order to compare the sequencing of amplified versus unamplified DNA. And then all samples were sequenced together, on the same sequencing instrument.

Our principal finding in this paper was a confirmation of the non-random trend of multiple displacement amplification, particularly when using multiple cells as template. When template amounts got close to the single-cell level, the "amplification profile" got noisier, but over-amplified regions still coincided with amplification peaks found in samples from larger template amounts.

Furthermore, we found that the magnitude of amplification bias was negatively correlated with template amounts, but not in a completely linear manner. Particularly for low amounts of template, small increases in the number of cells had a substantial effect on the bias magnitude.

Finally, none of our assemblies contained chimeric contigs. Rather, the assemblies became increasingly fragmented with lower template amounts. Similarly, contamination was present in low-template amplifications, but basically absent when multiple cells were used. In conclusion, we found that MDA is in fact very adequate for *de novo* genome assembly of bacterial endosymbionts if a pure sample of multiple cells can be obtained.

The reasons for the non-random amplification bias are currently not known, but deserve further attention. We consider primer composition as a prominent suspect. If this is the case, improved amplifications could potentially be obtained by using diverse mixtures of primers in parallel amplifications. This should be of particular interest for those working with single-cell amplifications, where it is sometimes possible to pool samples [107].

Chapter 8: The Lactic acid bacteria

In this final chapter, I will introduce my most recent bacterial obsession, the Lactic acid bacteria. We have prepared two manuscripts on these bacteria, which I will also introduce at the end of the chapter.

The friendly bacteria which surround us

What is lactic acid? Well, are you in a place with a floor available? Not too many people around? Good! Get up on your feet, bend your legs and squad, then jump up towards the roof as high as you can. Repeat until it becomes impossible to continue (and I mean impossible), if you are in average shape that could take less than a minute. There you go! Provided that you were able to conclude the experiment without any injuries, you should now have produced a substantial amount of lactic acid in your thighs.

Lactic acid is produced in a metabolic process called fermentation. In chemical terms, fermentation is the production of acids, gases or alcohol from sugars, without the use of oxygen. Thus, in your body, this will not occur until your muscles start running low on oxygen, hence the necessity of the strenuous little exercise I proposed. However, several species of bacteria and yeast employ fermentation as their favorite pathway for converting sugar to energy, and some are in fact incapable of using oxygen at all.

At this point, you may not be favorably impressed by the idea of fermentation; Put your feet up, breathe, relax, and have a beer. That's right, where would we be without fermentation? Wine and beer, sour foods ranging from pickled cucumbers to cheese, yoghurt and sourdough.. Fermentation has been used for thousands of years, to help us preserve and improve our food.

The name "Lactic acid bacteria" (LAB) refers to a highly diverse group of bacteria, which have in common precisely the ability to produce lactic acid from sugar by fermentation. While these bacteria have a long history in our food industry, they were first described by the famous Louis Pasteur back in 1857 [3, 108]. Pasteur observed that properly aged wine contained little spherical yeast cells. In contrast, when the wine turned sour, there was a proliferation of bacterial cells. Naturally, Pasteur proposed that the wine should be heated gently, which would help get rid of the bacteria.

However, the fact that bacteria could produce lactic acid was first demonstrated by Joseph Lister, who published his findings in a paper in the "Trans-

actions of the pathological society of London (1877-1878) [109]. Lister had a bacterial culture, which he had observed closely in the microscope. The culture contained several morphologically distinct bacteria, but one kind was dominant, and he named that one "*Bacterium lactis"*. He prepared 16 tubes of milk, which he pre-heated to eliminate all resident bacteria. Then he diluted his culture to a point "calculated .. to contain on the average a single *Bacterium lactis*", and added one such volume to the first ten tubes, and a bit more to the last ones. One by one, the milk in the tubes curdled, starting with the tubes to which he had added the largest volume of bacteria. Within five days, five of the tubes with the lowest inoculate had curdled too, but the other five remained liquid, even four months later. Incidentally, apart from demonstrating that the souring of the milk was caused by the added bacterium, his experiment was the first method developed for isolating and growing a pure culture of a bacterium, and is therefore justifiably considered a landmark in the history of microbiology [109]. The bacterium employed by Joseph Lister is today known by the name *Lactococcus lactis,* and is one of the key organisms in the early stages of cheese production. In fact, it was recently proposed to be the "official state microbe" by the number-one cheese-producing American state Wisconsin, but that's another story...[110].

Other important Lactic acid bacteria were described and cultured in the following decades. Henry Tissier isolated bifidobacteria from the feces of breast-fed infants in 1899 [111], and Ernst Moro isolated *Lactobacillus acidophilus* from feces of children in 1900 [112], thus demonstrating another important habitat of Lactic acid bacteria, namely the human gut.

Today, Lactic acid bacteria are known to occur in a large number of habitats of economical and health importance: as natural inhabitants of food products, including fruit, vegetables, meat and dairy products, and as colonizers of mucosal surfaces in animals, such as gut and vagina [111, 113]. More recently, it has become clear that some social insects, like the honeybee, also harbor Lactic acid bacteria, we will return to that point later in this chapter.

Taxonomy and genomics of Lactic acid bacteria

"Lactic acid bacteria" is a biological definition of a group of bacteria, rather than taxonomical, originating from the pre-Woese era (see chapter 4). In fact, the group consists of bacteria, which are very distantly related to each other. The two genera *Lactobacillus* and *Bifidobacterium*, which have been the focus of some of the work in my thesis, belong to two different phyla (major bacterial divisions). Furthermore, the genus *Lactobacillus* is paraphyletic, meaning that the most recent common ancestor of the genus has descendants, which are named differently. In other words, from a taxonomic point of view, the Lactic acid bacteria constitute a bit of a mess.

Considering the relatively early discovery of the LAB, and the ease of culturing several of them, comparative genomic studies have been initiated quite recently. In a broad comparative analysis of genomes from the order *Lactobacillales,* which includes the *Lactobacillus* genus, Makarova *et al.* found that the evolution of these bacteria was characterized by an extensive gene loss, likely due to the adaptation to nutrient-rich environments [113]. However, some acquisitions of genes involved in sugar and amino acid metabolism were also noted.

The bifidobacteria are notable in having a large proportion of their genomes dedicated to carbohydrate metabolism and transport, some of which are likely to have been horizontally acquired [114]. However, in general the genomes of the bifidobacteria display a large degree of synteny (conservation of gene order) between strains.

The insect-associated Lactic acid bacteria

Reserach on Lactic acid bacteria in insects has started very recently. Bifidobacterial species were in fact isolated from honeybees more than 4 decades ago, but have not been much studied. Similarly, the first *Lactobacillus* species from *Apis mellifera* was described only last year [115].

However, with the launching of the human microbiome project, interest in other gut microbiomes has also increased. Furthermore, bees are of considerable economic interest. You may think of them as pollinating flowers in your garden, but in fact nearly a third of all crops require bee pollination. Without pollination, the apple tree does not produce any apples. Without bees, we would be left with crops that are wind-pollinated, such as wheat, rice and corn, and the vegetable shelf in the supermarket would be very dull indeed. So we have good reason to take an interest in the health status of the bees.

Bees made the headlines in 2006, when huge losses started to occur all around the world [116]. Bees were found to leave the hive in the morning to go out and forage, and then disappear. The condition was named "colony collapse disorder", and inspired a number of documentaries with captivating titles such as "The silence of the bees" and "Queen of the sun: what are the bees trying to tell us?" Actually, bees have been known to disappear in the past too. What made the headlines was the sheer magnitude of the losses. Some bee-keepers reported losses of up 80% [116]. Several studies were initiated to identify the source of the phenomenon, and several factors have been implicated, including viruses, mites, pesticides and malnutrition but none seem to provide the complete answer.

The gut microbiota of the bees has been investigated in several metagenomic studies, and consistently seem to contain 8 phylotypes (see chapter 4), which completely dominate the community, although relative abundances

appear to vary [117, 118]. Two of the phylotypes are affiliated with the genus *Lactobacillus*, while one is classified as *Bifidobacterium*. The role of this community is currently unknown, but in a manner analogous to the human gut microbiome, it has been proposed that the bee microbiome may protect against pathogens. Additionally, it has been hypothesized that the bacteria may help to degrade complex compounds, otherwise potentially toxic to the bee. Considering both the recent decline of honeybees, and the fact that some beekeepers use antibiotics in bee management, this question does deserve further attention.

From a more general point of view, given all the parallels to the human gut microbiome, the bee gut microbiome represents an excellent model system for investigating the dynamics of such communities. How stable are the strains over time? How do they manage to avoid competitive exclusion? Have they evolved distinct niches, and if so, are these niches spatial or functional? There is a long way to go, but as with all other things in biology, we must start with a careful description of the components of our ecosystem. In the two manuscripts, III and IV, we take a first step in this direction, by analyzing the genome sequences of 24 strains, which were isolated from the stomach of the honeybee.

Paper III: Comparative genomics of Lactic acid bacteria

In this paper, we compare 11 genomes of Lactic acid bacteria isolated from the stomach of the honeybee, all from the same apiary in Sweden. These 11 strains represent three phylotypes, which are part of the core gut microbiota of the honeybee, and as such have been found in several metagenomic studies from around the world.

As a first point of interest, we wanted to see if we could find any unique gene repertoire associated with being a Lactic acid bacterium living in the gut of an insect. We did not find any protein families that were conserved between all strains, and absent in previously sequenced LAB genomes. However, we did find that each group contained conserved protein families, which did not have homologues in other LAB. Based on domain predictions, some of these could be inferred to be outer surface proteins. Furthermore, gene flow between the groups appeared to be restricted. Therefore, it seems likely that the phylotypes represent three distinct niches in the insect gut. However, in order to elucidate their exact roles in the gut, we need more experimental data. Where precisely do the bacteria attach in the gut? Do they co-locate? And what exactly are they doing?

We also found that each phylotype group had a large pan-genome, where proteins putatively involved in carbohydrate metabolism were dominant. This finding is interesting, because it shows that the superficial "simplicity" of the core gut microbiota, as estimated from the low number of phylotypes,

is perhaps exaggerated. It also begs the question of how the community is maintained. Notably, a metagenome study on the gut of worker bees from Arizona identified all four of the phylotype strains from one of the groups in the same sample, indicating that these strains form a stable community.

Paper IV: Comparative genomics of *Lactobacillus kunkeii*

In the honey stomach, the most frequently isolated LAB strain is *Lactobacillus kunkeei*, a strain which was originally isolated from a bad wine fermentation. This strain can be found in many *Apis* species around the world, and even in more distantly related bee species, indicating a strong association with the bees. In this paper, we compare 13 genomes of *L. kunkeei*, isolated from the stomach of honeybees from very diverse geographic locations.

A principal question was whether we could find evidence of co-diversification between these bacteria and the honeybee. This appeared not to be the case. The question of co-evolution therefore remains somewhat open, particularly when considering that *L. kunkeei* also has been isolated from flowers, and beebread and honey in the bee-hive. Incidentally, an excellent example of how difficult it can be to find the niche of a bacterium!

In the genomes, the most conspicuous protein family was a block of huge genes, with representatives in all the strains. No sequences in the databases are similar enough to give clues about the function of these genes, not even at the domain level. However, by comparing the sequences between genomes, it is clear that these genes are prone to recombination, and diversifying evolution. Given the size of the proteins, they are likely to fulfill an important role for these bacteria, as they can be predicted to be costly to produce. The "giant genes" are therefore good candidates for elucidating the niche of the strains.

L. kunkeei is not included in the core bee gut microbiome, where it has only been found occasionally, and at very low levels. In contrast, it has been found in beebread and honey. Beebread is an important protein source, particularly for bee larvae, as it is made from fermented pollen. Pollen, while rich in protein, contains very hard cell walls, which must be broken down to extract the contents. We therefore speculate that *L. kunkeei* is involved in this process in the beehive. In support for an adaptation to this particular habitat, we found that genes involved in amino acid catabolism tended to be located close to the origin, possibly allowing for a high expression level.

If *L. kunkeei* turns out to be a key player in the production of beebread, it fulfills a very important function in the beehive, and represents an interesting example of "microbial farming".

Svensk sammanfattning

De flesta djur bär på bakterier. I människokroppen finns till exempel tio gånger fler bakterieceller än de eukaryota celler vi är uppbyggda av, och bara i tarmarna väger bakteriecellerna flera kilon De flesta av dessa bakterier orsakar dock ingen skada hos sina värddjur, utan många tros till och med vara bra för oss. I den här avhandlingen har jag studerat arvsmassan hos bakterier som anpassat sig till insekter. Jag har använt mig av två olika modellsystem, där det ena utgörs av *Wolbachia* som är den mest vanligt förekommande bakterien i insekter och det andra utgörs av mjölksyrabakterier som anpassat sig till bin. I insekter finns bakterien *Wolbachia* bland annat i de reproduktiva vävnaderna, där den lever inuti värdcellerna. *Wolbachia* är mest känd för att den kan påverka reproduktionen hos sin värd för att främja sin egen överlevnad och spridning. Ibland kan en värd vara infekterad med flera olika *Wolbachia*-stammar, som har etablerat ett stabilt samhälle. I tarmen hos honungsbin finns flera olika arter av mjölksyrabakterier, men det är ännu oklart vilken roll dessa bakterier spelar. I den första delen av tarmen, honungs-magen, hittas ofta en specifik art, *Lactobacillus kunkeii*, men den funktionella betydelsen av bakterien i denna miljö har inte kunnat fastställas.

I denna avhandling har jag studerat evolutionen av dessa bakterier baserat på jämförelser av deras arvsmassor. Vi har bestämt arvsmassan för två nya stammar av Wolbachia, båda isolerade från bananflugan Drosophila simulans. Ett nytt protokoll har även utvecklats för att isolera dessa Wolbachia-stammars DNA från ett fåtal embryon av bananflugan. En jämförande analys av arvsmassan från de två nya stammarna, tillsammans med tidigare publicerade arvsmassor från andra Wolbachia-stammar genomfördes. De analyserade stammarna kan delas in i två olika grupper och i vår analys fann vi att Wolbachia stammar som tillhör samma grupp utbyter DNA med varandra väldigt ofta även om bakterierna isolerats från olika arter av insekter. Däremot var utbytet starkt reducerad mellan olika genetiska grupper, även om stammarna som jämfördes hade isolerats från samma insekt.

På liknande sätt jämfördes arvsmassorna från mjölksyrabakterier som isolerats från honungs-magen hos honungsbin. Dessa visade sig representera tre grupper av bakterier som tidigare hittats i tarmen hos honungsbin. Stammarna är alla isolerade från samma värd och provtagningsställe, men trots det fann vi få tecken på gen-utbyte mellan de tre grupperna. Varje grupp bar på unika egenskaper, som sannolikt är av betydelse för interaktionen med deras värd.

Vi jämförde dessutom arvsmassorna från ett tiotal olika stammar av bakterien *Lactobacillus kunkeei*, som isolerats från honungs-magen hos olika arter av bin från olika delar av världen. Utifrån denna analys fann vi inga bevis för att dessa stammar var strikt associerade med den bi-art som de isolerats från. Baserat på våra studier av arvsmassan från de olika stammarna föreslår vi att *Lactobacillus kunkeei* har specialiserat sig på att leva i bi-bröd, som består av nektar och pollen och som används som mat för larverna i bisamhället.

I den mänskliga tarmen har den mikrobiella miljön visats interagera med immunsystemet, och det är sannolikt att mikrobiella samhällen som är förknippade med insekter också påverkar hälsan hos sin värd. En bättre förståelse av den roll som dessa bakteriella samhällen spelar är mycket viktig för att utveckla långsiktiga strategier for hälsa och välmående hos oss människor och våra tamdjur, och inte att förglömma, våra bisamhällen med sina stora ekonomiska värden.

Acknowledgements

7 years has gone, since my father and I drove to Sweden with all my worldly belongings. So many amazing people have come and gone, and I am immensely grateful for having met you all!

Siv, I wanted you as my supervisor from the first time I heard you give a talk. Your enthusiasm is contagious to say the least, as is your never-failing optimism. Thank you for being you, and for letting me be me. **Lisa**, cookie Klasson, how would I have managed without you? Thank you for your endless patience with my endless bioinformatic problems, and for always being there to help and share your knowledge. Thanks for all the fabulous and hilarious times we have had, I forgive you for the margarita bowl! **Daniel**, it has been a real pleasure to work with you. Even though I worry about your sleeping hours and your eating habits (popcorn for dinner?), I must admit that I admire your ambition and perfectionism for all things science. Keep making multi-coloured figures Dani, that´s who you are. **FeiFei**, thanks for inviting me to your home, again and again, I probably dont deserve it, but I certainly enjoy it. **Kasia**, my room-mate, colleague and friend, thank you for being all that, and so damn nice too. **Ann-Sofie** and **Kristina**, you have taught me nearly everything I know in the lab. Thank you for being so patient, competent, helpful and incredibly nice. Dr. **Johan V**., I really don´t know how we will manage without you. **Minna**, you are coming back soon, right? Stressed is indeed desserts backwards, and we need you here to make the point. **Mayank** and **Wei**, I am very glad we got you safely back from Portugal :) Thanks for letting me have a home on your computers too. And **The Wabi-team**, **Johan R** and **Pall**, I will beat you both in the next run, you´ll see. **Björn N.**, you were an inspiration for me, even as a master-student. **Lionel,** thank you for being so incredibly helpful and competent, always! **Jan**, the book-club was great, thanks for taking the time, and for caring about teaching. **Jessin**, all the best to you, I believe you can do anything you want. **Ajith,** thanks for making social activities obligatory, I always wanted to do that, but you made it! **Yu**, I hope we will share a new office too, it´s going to be super. **Patrik**, you are next! Good luck! And all **the "cell-singles"!** Thanks to **Thijs**, **Jimmy**, **Anja**, **Joran**, **Anders** and **Roel**, for crowding up our journal club and coffee room, you are making the department of molecular evolution grand!

I also wish to thank all my past colleagues, so many great people, too many to include here. Even so; **Eva, Fredrik** and **Olga,** you were my first supervisors, and I wouldn´t be here today if you hadnt done such a great job. **Jennifer,** Dr. Ast! Thanks for all the laughs and all the great times! **Emelie** and **Johan W**., thank you for all the hard work you did for us this summer!

Alejandra and **Tobias**, thank you for sharing your data with us, and introducing the fabulous bee endosymbionts.

I would also like to thank everyone from **Uppsala Kyokushin karate-club**; you were another excellent reason to stay in Sweden.

Camille (Dr. Körner) og **Ibber** (Dr. Möller-Hansen), husmor-klubben! Tänk at det skulle komme så vidt! Jeg er så utroligt glad for at vi har holdt kontakten, og for at I er mine venner :)

Mor og Far, Birgitte og Niels, min famile. Tak fordi I altid er der for mig. Denne bog er tilegnet Jer.

References

1. Gest, H., *The discovery of microorganisms by Robert Hooke and Antoni Van Leeuwenhoek, fellows of the Royal Society.* Notes Rec R Soc Lond, 2004. **58**(2): p. 187-201.
2. Sapp, J., *Evolution by association: A history of symbiosis*. 1994: Oxford University Press.
3. Gaynes, R.P., *Germ Theory: Medical Pioneers in Infectious Diseases*. 2011, Washington, DC, USA: ASM Press.
4. smith, D.C., *Symbiosis research at the end of the millenium.* Hydrobiologia, 2001. **461**: p. 49-54.
5. Relman, D.A., *'Til death do us part': coming to terms with symbiotic relationships. Forward.* Nat Rev Microbiol, 2008. **6**(10): p. 721-4.
6. Wade, W.G., *The oral microbiome in health and disease.* Pharmacol Res, 2013. **69**(1): p. 137-43.
7. Consortium, T.H.M.P., *Structure, function and diversity of the healthy human microbiome.* Nature, 2012. **486**(7402): p. 207-214.
8. Trivedi, B., *Microbiome: The surface brigade.* Nature, 2012. **492**(7429): p. S60-1.
9. Ley, R.E., et al., *Worlds within worlds: evolution of the vertebrate gut microbiota.* Nat Rev Microbiol, 2008. **6**(10): p. 776-88.
10. Bergman, E.N., *Energy contributions of volatile fatty acids from the gastrointestinal tract in various species.* Physiol Rev, 1990. **70**(2): p. 567-90.
11. Mazmanian, S.K., et al., *Staphylococcus aureus sortase, an enzyme that anchors surface proteins to the cell wall.* Science, 1999. **285**(5428): p. 760-3.
12. Schrezenmeir, J. and M. de Vrese, *Probiotics, prebiotics, and synbiotics--approaching a definition.* Am J Clin Nutr, 2001. **73**(2 Suppl): p. 361S-364S.
13. Katan, M.B., *Why the European Food Safety Authority was right to reject health claims for probiotics.* Benef Microbes, 2012. **3**(2): p. 85-9.
14. Shigenobu, S. and A.C. Wilson, *Genomic revelations of a mutualism: the pea aphid and its obligate bacterial symbiont.* Cell Mol Life Sci, 2011. **68**(8): p. 1297-309.
15. Fukatsu, T. and T. Hosokawa, *Capsule-transmitted gut symbiotic bacterium of the Japanese common plataspid stinkbug, Megacopta punctatissima.* Appl Environ Microbiol, 2002. **68**(1): p. 389-96.

16. Norsworthy, A.N. and K.L. Visick, *Gimme shelter: how Vibrio fischeri successfully navigates an animal's multiple environments.* Front Microbiol, 2013. **4**: p. 356.
17. Koenig, J.E., et al., *Succession of microbial consortia in the developing infant gut microbiome.* Proc Natl Acad Sci U S A, 2011. **108 Suppl 1**: p. 4578-85.
18. Moran, N.A., J.P. McCutcheon, and A. Nakabachi, *Genomics and evolution of heritable bacterial symbionts.* Annu Rev Genet, 2008. **42**: p. 165-90.
19. Claes, M.F. and P.V. Dunlap, *Aposymbiotic culture of the sepiolid squid Euprymna scolopes: role of the symbiotic bacterium Vibrio fischeri in host animal growth, development, and light organ morphogenesis.* J Exp Zool, 2000. **286**(3): p. 280-96.
20. Alley, T.R., *Competition theory, evolution, and the concept of an ecological niche.* Acta Biotheor, 1982. **31**(3): p. 165-79.
21. Vandermeer, J.H., *Niche theory.* Annual Review of Ecology and Systematics, 1972. **3**: p. 107-132.
22. Grinnell, J., *Geography and evolution.* Ecology, 1924. **5**: p. 225-229.
23. Elton, C.S., *Animal Ecology*. 1927: University of Chicago Press.
24. Kingsland, S., *Modelling Nature: Episodes in the History of Population Ecology*. Second ed. 1995, USA: The University of Chicago Press.
25. Hutchinson, G.E., *Concluding remarks.* Cold Spring Harbor Symposia on Quantitative Biology, 1957. **22**(2): p. 415-427.
26. Hardin, G., *The competitive exclusion principle.* Science, 1960. **131**(3409): p. 1292-7.
27. Gause, G.F., *The Struggle for Existence*. 1934, Baltimore: Williams and Wilkins.
28. Hibbing, M.E., et al., *Bacterial competition: surviving and thriving in the microbial jungle.* Nat Rev Microbiol, 2010. **8**(1): p. 15-25.
29. Kearns, D.B., *A field guide to bacterial swarming motility.* Nat Rev Microbiol, 2010. **8**(9): p. 634-44.
30. Abee, T., et al., *Biofilm formation and dispersal in Gram-positive bacteria.* Curr Opin Biotechnol, 2011. **22**(2): p. 172-9.
31. Elias, S. and E. Banin, *Multi-species biofilms: living with friendly neighbors.* FEMS Microbiol Rev, 2012.
32. Eberhard, A., et al., *Structural identification of autoinducer of Photobacterium fischeri luciferase.* Biochemistry, 1981. **20**(9): p. 2444-9.
33. Miller, M.B. and B.L. Bassler, *Quorum sensing in bacteria.* Annu Rev Microbiol, 2001. **55**: p. 165-99.
34. Jousset, A., *Ecological and evolutive implications of bacterial defences against predators.* Environ Microbiol, 2012. **14**(8): p. 1830-43.
35. Sockett, R.E., *Predatory lifestyle of Bdellovibrio bacteriovorus.* Annu Rev Microbiol, 2009. **63**: p. 523-39.

36. Berleman, J.E. and J.R. Kirby, *Deciphering the hunting strategy of a bacterial wolfpack.* FEMS Microbiol Rev, 2009. **33**(5): p. 942-57.
37. Friman, V.P., S.P. Diggle, and A. Buckling, *Protist predation can favour cooperation within bacterial species.* Biol Lett, 2013. **9**(5): p. 20130548.
38. Weinbauer, M.G., *Ecology of prokaryotic viruses.* FEMS Microbiol Rev, 2004. **28**(2): p. 127-81.
39. Kent, B.N. and S.R. Bordenstein, *Phage WO of Wolbachia: lambda of the endosymbiont world.* Trends Microbiol, 2010. **18**(4): p. 173-81.
40. Sorek, R., C.M. Lawrence, and B. Wiedenheft, *CRISPR-mediated adaptive immune systems in bacteria and archaea.* Annu Rev Biochem, 2013. **82**: p. 237-66.
41. Bezier, A., et al., *Polydnaviruses of braconid wasps derive from an ancestral nudivirus.* Science, 2009. **323**(5916): p. 926-30.
42. Moran, N.A., et al., *The players in a mutualistic symbiosis: insects, bacteria, viruses, and virulence genes.* Proc Natl Acad Sci U S A, 2005. **102**(47): p. 16919-26.
43. Canchaya, C., et al., *Phage as agents of lateral gene transfer.* Curr Opin Microbiol, 2003. **6**(4): p. 417-24.
44. Sapp, J. and G.E. Fox, *The singular quest for a universal tree of life.* Microbiol Mol Biol Rev, 2013. **77**(4): p. 541-50.
45. Mira, A., et al., *The bacterial pan-genome:a new paradigm in microbiology.* Int Microbiol, 2010. **13**(2): p. 45-57.
46. Venter, J.C., et al., *Environmental genome shotgun sequencing of the Sargasso Sea.* Science, 2004. **304**(5667): p. 66-74.
47. Stewart, E.J., *Growing unculturable bacteria.* J Bacteriol, 2012. **194**(16): p. 4151-60.
48. Cohan, F.M., *What are bacterial species?* Annu Rev Microbiol, 2002. **56**: p. 457-87.
49. Fraser, C., W.P. Hanage, and B.G. Spratt, *Recombination and the nature of bacterial speciation.* Science, 2007. **315**(5811): p. 476-80.
50. Acinas, S.G., et al., *Fine-scale phylogenetic architecture of a complex bacterial community.* Nature, 2004. **430**(6999): p. 551-4.
51. Wirth, T., A. Meyer, and M. Achtman, *Deciphering host migrations and origins by means of their microbes.* Mol Ecol, 2005. **14**(11): p. 3289-306.
52. Dunlap, P.V., et al., *Phylogenetic analysis of host-symbiont specificity and codivergence in bioluminescent symbioses.* Cladistics, 2007. **23**: p. 507-523.
53. Toft, C. and S.G. Andersson, *Evolutionary microbial genomics: insights into bacterial host adaptation.* Nat Rev Genet, 2010. **11**(7): p. 465-75.
54. McCutcheon, J.P. and N.A. Moran, *Extreme genome reduction in symbiotic bacteria.* Nat Rev Microbiol, 2012. **10**(1): p. 13-26.
55. Dunning Hotopp, J.C., *Horizontal gene transfer between bacteria and animals.* Trends Genet, 2011. **27**(4): p. 157-63.

56. Nikoh, N., et al., *Bacterial genes in the aphid genome: absence of functional gene transfer from Buchnera to its host.* PLoS Genet, 2010. **6**(2): p. e1000827.
57. Schneewind, O. and D.M. Missiakas, *Protein secretion and surface display in Gram-positive bacteria.* Philos Trans R Soc Lond B Biol Sci, 2012. **367**(1592): p. 1123-39.
58. Juge, N., *Microbial adhesins to gastrointestinal mucus.* Trends Microbiol, 2012. **20**(1): p. 30-9.
59. Edelman, S.M., et al., *Identification of a high-molecular-mass Lactobacillus epithelium adhesin (LEA) of Lactobacillus crispatus ST1 that binds to stratified squamous epithelium.* Microbiology, 2012. **158**(Pt 7): p. 1713-22.
60. Siezen, R., et al., *Lactobacillus plantarum gene clusters encoding putative cell-surface protein complexes for carbohydrate utilization are conserved in specific gram-positive bacteria.* BMC Genomics, 2006. **7**: p. 126.
61. Turroni, F., et al., *Role of sortase-dependent pili of Bifidobacterium bifidum PRL2010 in modulating bacterium-host interactions.* Proc Natl Acad Sci U S A, 2013. **110**(27): p. 11151-6.
62. Jiggins, F.M., G.D. Hurst, and Z. Yang, *Host-symbiont conflicts: positive selection on an outer membrane protein of parasitic but not mutualistic Rickettsiaceae.* Mol Biol Evol, 2002. **19**(8): p. 1341-9.
63. Rikihisa, Y., *Anaplasma phagocytophilum and Ehrlichia chaffeensis: subversive manipulators of host cells.* Nat Rev Microbiol, 2010. **8**(5): p. 328-39.
64. Hynonen, U. and A. Palva, *Lactobacillus surface layer proteins: structure, function and applications.* Appl Microbiol Biotechnol, 2013. **97**(12): p. 5225-43.
65. Fanning, S., et al., *Bifidobacterial surface-exopolysaccharide facilitates commensal-host interaction through immune modulation and pathogen protection.* Proc Natl Acad Sci U S A, 2012. **109**(6): p. 2108-13.
66. Sapp, J., *Paul Buchner (1886-1978) and hereditary symbiosis in insects.* Int Microbiol, 2002. **5**(3): p. 145-50.
67. Hertig, M. and S.B. Wolbach, *Studies on Rickettsia-Like Micro-Organisms in Insects.* J Med Res, 1924. **44**(3): p. 329-374 7.
68. Hertig, M., *The rickettsia Wolbachia pipientis (gen. et. sp. n) and associated inclusions of the mosquito, Culex pipiens.* Parasitology, 1936. **28**: p. 453-486.
69. Engelstadter, J. and A. Telschow, *Cytoplasmic incompatibility and host population structure.* Heredity (Edinb), 2009. **103**(3): p. 196-207.
70. Marshall, J.F., *The british mosquitoes*. 1938, London: Trustees of the British Museum.
71. Yen, J.H. and A.R. Barr, *New hypothesis of the cause of cytoplasmic incompatibility in Culex pipiens L.* Nature, 1971. **232**(5313): p. 657-8.

72. Yen, J.H. and A.R. Barr, *The etiological agent of cytoplasmic incompatibility in Culex pipiens.* J Invertebr Pathol, 1973. **22**(2): p. 242-50.
73. Jeyaprakash, A. and M.A. Hoy, *Multiple displacement amplification in combination with high-fidelity PCR improves detection of bacteria from single females or eggs of Metaseiulus occidentalis (Nesbitt) (Acari: Phytoseiidae).* J Invertebr Pathol, 2004. **86**(3): p. 111-6.
74. Hilgenboecker, K., et al., *How many species are infected with Wolbachia?--A statistical analysis of current data.* FEMS Microbiol Lett, 2008. **281**(2): p. 215-20.
75. LePage, D. and S.R. Bordenstein, *Wolbachia: Can we save lives with a great pandemic?* Trends Parasitol, 2013. **29**(8): p. 385-93.
76. Zabalou, S., et al., *Multiple rescue factors within a Wolbachia strain.* Genetics, 2008. **178**(4): p. 2145-60.
77. Werren, J.H., L. Baldo, and M.E. Clark, *Wolbachia: master manipulators of invertebrate biology.* Nat Rev Microbiol, 2008. **6**(10): p. 741-51.
78. Hornett, E.A., et al., *You can't keep a good parasite down: evolution of a male-killer suppressor uncovers cytoplasmic incompatibility.* Evolution, 2008. **62**(5): p. 1258-63.
79. Jaenike, J., *Spontaneous emergence of a new Wolbachia phenotype.* Evolution, 2007. **61**(9): p. 2244-52.
80. Sasaki, T., N. Massaki, and T. Kubo, *Wolbachia variant that induces two distinct reproductive phenotypes in different hosts.* Heredity (Edinb), 2005. **95**(5): p. 389-93.
81. Teixeira, L., A. Ferreira, and M. Ashburner, *The bacterial symbiont Wolbachia induces resistance to RNA viral infections in Drosophila melanogaster.* PLoS Biol, 2008. **6**(12): p. e2.
82. Hedges, L.M., et al., *Wolbachia and virus protection in insects.* Science, 2008. **322**(5902): p. 702.
83. Bian, G., et al., *Wolbachia invades Anopheles stephensi populations and induces refractoriness to Plasmodium infection.* Science, 2013. **340**(6133): p. 748-51.
84. Fast, E.M., et al., *Wolbachia enhance Drosophila stem cell proliferation and target the germline stem cell niche.* Science, 2011. **334**(6058): p. 990-2.
85. Weeks, A.R., et al., *From parasite to mutualist: rapid evolution of Wolbachia in natural populations of Drosophila.* PLoS Biol, 2007. **5**(5): p. e114.
86. Hosokawa, T., et al., *Wolbachia as a bacteriocyte-associated nutritional mutualist.* Proc Natl Acad Sci U S A, 2010. **107**(2): p. 769-74.
87. Hoffmann, A.A., M. Turelli, and G.M. Simmons, *Unidirectional incompatibility between populations of Drosophila simulans.* Evolution, 1986. **40**(4): p. 692-701.

88. Mercot, H. and S. Charlat, *Wolbachia infections in Drosophila melanogaster and D. simulans: polymorphism and levels of cytoplasmic incompatibility.* Genetica, 2004. **120**(1-3): p. 51-9.
89. Osborne, S.E., et al., *Variation in antiviral protection mediated by different Wolbachia strains in Drosophila simulans.* PLoS Pathog, 2009. **5**(11): p. e1000656.
90. Veneti, Z., et al., *Heads or tails: host-parasite interactions in the Drosophila-Wolbachia system.* Appl Environ Microbiol, 2004. **70**(9): p. 5366-72.
91. Solignac, M., *Mitochondrial DNA in the Drosophila melanogaster complex.* Genetica, 2004. **120**(1-3): p. 41-50.
92. Ballard, J.W., *Sequential evolution of a symbiont inferred from the host: Wolbachia and Drosophila simulans.* Mol Biol Evol, 2004. **21**(3): p. 428-42.
93. Dean, M.D. and J.W. Ballard, *High divergence among Drosophila simulans mitochondrial haplogroups arose in midst of long term purifying selection.* Mol Phylogenet Evol, 2005. **36**(2): p. 328-37.
94. Kriesner, P., et al., *Rapid sequential spread of two Wolbachia variants in Drosophila simulans.* PLoS Pathog, 2013. **9**(9): p. e1003607.
95. Charlat, S., et al., *Evolution of Wolbachia-induced cytoplasmic incompatibility in Drosophila simulans and D. sechellia.* Evolution, 2002. **56**(9): p. 1735-42.
96. Klasson, L., et al., *The mosaic genome structure of the Wolbachia wRi strain infecting Drosophila simulans.* Proc Natl Acad Sci U S A, 2009. **106**(14): p. 5725-30.
97. Cerveau, N., et al., *Short- and long-term evolutionary dynamics of bacterial insertion sequences: insights from Wolbachia endosymbionts.* Genome Biol Evol, 2011. **3**: p. 1175-86.
98. Bordenstein, S.R. and J.J. Wernegreen, *Bacteriophage flux in endosymbionts (Wolbachia): infection frequency, lateral transfer, and recombination rates.* Mol Biol Evol, 2004. **21**(10): p. 1981-91.
99. Wu, M., et al., *Phylogenomics of the reproductive parasite Wolbachia pipientis wMel: a streamlined genome overrun by mobile genetic elements.* PLoS Biol, 2004. **2**(3): p. E69.
100. Klasson, L., et al., *Genome evolution of Wolbachia strain wPip from the Culex pipiens group.* Mol Biol Evol, 2008. **25**(9): p. 1877-87.
101. Lin, M., et al., *Anaplasma phagocytophilum AnkA secreted by type IV secretion system is tyrosine phosphorylated by Abl-1 to facilitate infection.* Cell Microbiol, 2007. **9**(11): p. 2644-57.
102. Lo, N., et al., *Taxonomic status of the intracellular bacterium Wolbachia pipientis.* Int J Syst Evol Microbiol, 2007. **57**(Pt 3): p. 654-7.
103. Binga, E.K., R.S. Lasken, and J.D. Neufeld, *Something from (almost) nothing: the impact of multiple displacement amplification on microbial ecology.* ISME J, 2008. **2**(3): p. 233-41.

104. Hongoh, Y., et al., *Complete genome of the uncultured Termite Group 1 bacteria in a single host protist cell.* Proc Natl Acad Sci U S A, 2008. **105**(14): p. 5555-60.
105. Pinard, R., et al., *Assessment of whole genome amplification-induced bias through high-throughput, massively parallel whole genome sequencing.* BMC Genomics, 2006. **7**: p. 216.
106. Lasken, R.S. and T.B. Stockwell, *Mechanism of chimera formation during the Multiple Displacement Amplification reaction.* BMC Biotechnol, 2007. **7**: p. 19.
107. Blainey, P.C., *The future is now: single-cell genomics of bacteria and archaea.* FEMS Microbiol Rev, 2013. **37**(3): p. 407-27.
108. Pasteur, L., *Report on the lactic-acid fermentation.* Molecular medicine, 1857. **1**(6): p. 599-601.
109. Santer, M., *Joseph Lister: first use of a bacterium as a 'model organism' to illustrate the cause of infectious disease of humans.* Notes Rec R Soc Lond, 2010. **64**(1): p. 59-65.
110. Davey, M., *And now, a State Microbe*, in *The New York Times*2010.
111. Lee, J.H. and D.J. O'Sullivan, *Genomic insights into bifidobacteria.* Microbiol Mol Biol Rev, 2010. **74**(3): p. 378-416.
112. Weirich, A. and G.F. Hoffmann, *Ernst Moro (1874-1951)--a great pediatric career started at the rise of university-based pediatric research but was curtailed in the shadows of Nazi laws.* Eur J Pediatr, 2005. **164**(10): p. 599-606.
113. Makarova, K., et al., *Comparative genomics of the lactic acid bacteria.* Proc Natl Acad Sci U S A, 2006. **103**(42): p. 15611-6.
114. Ventura, M., et al., *Genomics of Actinobacteria: tracing the evolutionary history of an ancient phylum.* Microbiol Mol Biol Rev, 2007. **71**(3): p. 495-548.
115. Killer, J., et al., *Lactobacillus apis sp. nov., from the stomach of honeybees (Apis mellifera) having in vitro inhibitory effect on causative agents of American and European Foulbrood.* Int J Syst Evol Microbiol, 2013.
116. Evans, J.D. and R.S. Schwarz, *Bees brought to their knees: microbes affecting honey bee health.* Trends Microbiol, 2011. **19**(12): p. 614-20.
117. Martinson, V.G., et al., *A simple and distinctive microbiota associated with honey bees and bumble bees.* Mol Ecol, 2011. **20**(3): p. 619-28.
118. Moran, N.A., et al., *Distinctive gut microbiota of honey bees assessed using deep sampling from individual worker bees.* PLoS One, 2012. **7**(4): p. e36393.

Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 1121

Editor: The Dean of the Faculty of Science and Technology

A doctoral dissertation from the Faculty of Science and Technology, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology.

Distribution: publications.uu.se
urn:nbn:se:uu:diva-217724

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2014