Exploration of bacteria associated with Anopheles mosquitoes around the world

For the prevention of transmission of malaria

LOUISE K. J. NILSSON
Dissertation presented at Uppsala University to be publicly examined in A1:111a, BMC, Husargatan 3, Uppsala, Friday, 14 September 2018 at 09:15 for the degree of Doctor of Philosophy. The examination will be conducted in English. Faculty examiner: Professor Michael Strand (Department of Entomology, University of Georgia).

Abstract

Every year, hundreds of thousands of people die from malaria. Malaria is a disease caused by parasites, which are spread by female vector mosquitoes of the genus Anopheles. Current control measures against malaria are based on drugs against the parasites and vector control using insecticides. A problem with these measures is the development of resistance, both in the parasites against the drugs and the mosquitoes against the insecticides. Therefore, additional areas of malaria control must be explored. One such area involves the bacteria associated with the vector mosquitoes. Bacteria have been shown to affect mosquitoes at all life stages, e.g. by affecting choice of oviposition site by female mosquitoes, development of larvae and susceptibility to parasite infection in adults. Furthermore, genetic modification of symbiotic bacteria has been suggested as a means of killing the parasites in the mosquitoes. This thesis is based on four field studies and one laboratory study and aims to investigate the naturally occurring bacteria associated with different life stages of malaria mosquitoes and how they are acquired. All studies are based on amplicon sequencing of the 16S rRNA gene. We found that overall mosquitoes contain different bacterial communities. However, bacteria associated with adults reflect their life history and can predict the origin of mosquitoes. Bacteria in larvae are similar during the developmental stages but vary with breeding site. Also in larvae, the bacteria could be used to predict the origin of breeding site. Some bacteria could be related to the presence or absence of Anopheles around human habitations and the diversity of aquatic bacteria in breeding sites is large, though some taxa are common. Overall, both environmental and host-genetic factors affect the gut bacterial composition in adult females. In conclusion, this thesis contributes to increasing the knowledge of bacterial diversity associated with Anopheles mosquitoes and to provide insight into how the bacteria are acquired, which can be useful in malaria control.

Keywords: microbiota, microbiome, vector-borne disease, 16S rRNA gene, amplicon sequencing

Louise K. J. Nilsson, Department of Cell and Molecular Biology, Microbiology, Box 596, Uppsala University, SE-75124 Uppsala, Sweden.

© Louise K. J. Nilsson 2018

ISSN 1651-6214
urn:nbn:se:uu:diva-352547 (http://urn.kb.se/resolve?urn=nbn:se:uu:diva-352547)
“If you think you are too small to make a difference, you haven’t spent a night with a mosquito.”

- African proverb

Till min familj
List of Papers

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


*These authors contributed equally

Reprints were made with permission from the respective publishers.
In addition to the papers included in this thesis, the author has contributed to the following paper:


*These authors contributed equally
# Contents

Introduction ................................................................................................... 11
Malaria ...................................................................................................... 12
*Anopheles* ............................................................................................. 13
*Plasmodium* .......................................................................................... 13
Control methods ................................................................................... 14
Bacteria associated with *Anopheles* ........................................................ 15
  Adult midgut bacteria .......................................................................... 15
  Larval bacteria ..................................................................................... 17
  Breeding-water bacteria ....................................................................... 18
  Field vs laboratory ............................................................................... 18
Identification method of bacteria ......................................................... 18
Aims of the thesis ........................................................................................ 20
Materials and methods ................................................................................ 21
  Sample collection ................................................................................ 21
  DNA extraction, PCR, and MiSeq sequencing ......................................... 22
  Bioinformatics ...................................................................................... 23
Present investigations .................................................................................... 24
  Paper I | Aquatic bacteria and the presence of larvae ....................... 24
  Paper II | Bacteria associated with *Anopheles* breeding waters ........ 26
  Paper III | Bacteria associated with *Anopheles* larvae ...................... 28
  Paper IV | Bacteria associated with *Anopheles* adults ....................... 31
  Paper V | Effect of genetics and environment on mosquito gut bacteria .. 33
Meta-analysis ................................................................................................ 35
  Most abundant OTUs........................................................................... 35
  Most frequently present OTUs in *Anopheles* ...................................... 36
  Most frequently present OTUs in *Anopheles* breeding water .......... 38
General discussion.................................................................40
Conclusions and future perspectives............................................43
Svensk sammanfattning..............................................................45
Acknowledgements ..................................................................48
References ................................................................................50
# Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>16S rRNA</td>
<td>16S Ribosomal Ribonucleic Acid</td>
</tr>
<tr>
<td>ACT</td>
<td>Artemisinin-based Combination Therapy</td>
</tr>
<tr>
<td>bp</td>
<td>Base Pair</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic Acid</td>
</tr>
<tr>
<td>NMDS</td>
<td>Non-metric Multidimensional Scaling</td>
</tr>
<tr>
<td>OTU</td>
<td>Operational Taxonomic Unit</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase Chain Reaction</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organization</td>
</tr>
</tbody>
</table>
Introduction

Bacteria are a diverse group of organisms, present everywhere on Earth. They were first discovered and observed in 1676 by Antonie van Leeuwenhoek (Smit and Heniger 1975). Since then we have learnt that they are found all around us, on the ground, in the water, in the air, in extreme environments as well as in and on plants, animals and humans. It has been estimated to exist $1 \times 10^{30}$ microbial cells on earth (Gilbert et al. 2010) and in and on our bodies at least as many bacterial cells as human cells (Sender et al. 2016) with the intestine alone containing approximately 500-1000 species of bacteria (Sommer and Backhed 2013). Of the total global environmental DNA only a small fraction has been sequenced (Gilbert et al. 2010). However, with advances in high-throughput sequencing methods in the last decades, much attention has been given to investigations of bacteria associated with different environments. For example two large-scale projects, the Human Microbiome Project initiated in 2007 (NIH HMP Working Group et al. 2009) and the Earth Microbiome Project founded in 2010 (Gilbert et al. 2010; Thompson et al. 2017) with the aim of characterizing microbial communities in the human body and in different environments around the world.

Bacteria affect the environment in many ways by contributing to biogeochemical cycles, degrading and recycling essential elements like nitrogen, carbon, and phosphorus (Alongi 1994). Bacteria have also been utilized by humans for our benefit, for example in the production of food where bacteria are used in fermentation to produce cheese, yoghurt, and bread (Caplice and Fitzgerald 1999). Another area in which bacteria are used for human purposes is disease control, bacteria have for example been used in the production of pharmaceuticals for many years (Neumann and Neumann-Staubitz 2010).

Some diseases are spread by vectors, organisms that transport and spread pathogens from one host to another. As no exception these vectors are also associated with bacteria. The effect of bacteria on vectors of disease and the pathogens they carry might therefore be exploited to control disease transmission. The most recognized disease vector is probably the mosquito, which is the vector of many human diseases. This thesis includes studies of mosquitoes of the genera *Anopheles* and *Aedes*. However, the main focus is on the genus that transmits the most deadly vector borne disease, namely malaria.
Malaria

Evidence of possible cases of human malaria exists from as early as 2700 BC in a Chinese document. Several other references from other places BC also exist, including clay tablets from Mesopotamia, papyri from Egypt, and Hindu texts. The early Greeks like Hippocrates were also aware of the malaria symptoms seen in people living in marshy areas (Cox 2010). Today the disease is responsible for around 216 million cases and 445,000 deaths per year, with most cases and deaths occurring in Africa with the most vulnerable group being children under 5 years old (WHO 2017). However, malaria also occurs in many other regions of the world (Fig. 1).

Figure 1. Malaria-endemic areas in 2015. Based on a figure (Phillips et al. 2017) and a map by FreeVectorMaps.com (http://freevectormaps.com).

The symptoms of malaria include, fever, anaemia, renal failure, acidosis, and cerebral and placental malaria (Tilley et al. 2011). Malaria is usually classified as asymptomatic, uncomplicated or severe with severe malaria often being fatal. Complications are due to severe anaemia and obstruction of blood vessels by the presence of infected red blood cells in capillaries (Phillips et al. 2017). Diagnosis is based on fever and presence of parasites in the blood. The detection of parasites can be done by light microscopy of blood smears or by a rapid diagnostic test (RDT) based on detection of parasite antigens in the blood, PCR or loop-mediated isothermal amplification (LAMP), which is a type of PCR but does not require thermal cycling and electricity (Phillips et al. 2017).
Anopheles

The vectors of malaria are mosquito species of the genus Anopheles. However, not all Anopheles mosquitoes are able to spread human malaria. Of the over 450 species of Anopheles, around 70 can transmit human malaria parasites. These species are found in different areas of the world with some areas having more than one malaria-vector species present (Sinka et al. 2012). Anopheles mosquitoes able to transmit malaria are found all over the world, not just in the endemic areas shown in Fig. 1. However, transmission efficiency depends on the vector species and therefore varies worldwide. For example in sub-Saharan Africa An. gambiae is a dominant and highly effective vector (Phillips et al. 2017). Anopheles, like other genera of mosquitoes, has one aquatic life stage and one terrestrial.

Plasmodium

The causative agent of malaria is a parasitic protozoan of the genus Plasmodium. There are five species of Plasmodium that spread human malaria, falciparum, vivax, ovale, malariae, and knowlesi. P. falciparum and P. vivax are the most common ones and the cause of the largest public health burden. For these two parasites, humans are the only mammalian hosts (Phillips et al. 2017). The first person to observe these parasites in the blood of malaria patients was Alphonse Laveran in 1880. Later, in 1897, Ronald Ross discovered that the malaria parasites were transmitted by infected mosquitoes in avian malaria. A year later, Grassi, Bignami, and Bastienelli showed that human malaria parasites were transmitted in the same way by mosquitoes (Cox 2010).

Life cycle

The malaria parasite life cycle requires both a vector in the form of a mosquito and a vertebrate host. When the mosquito takes a blood meal from an infected person, parasites are also ingested. Plasmodium gametocytes travel with the rest of the bloodmeal to the posterior midgut lumen of the mosquito. Gametogenesis is initiated by the mosquito environment and extracellular macro- and microgametes are produced rapidly. These gametes quickly fertilize in the bloodmeal and form diploid zygotes that undergo meiosis. The zygote transforms into a motile ookinete with the shape of a banana. The ookinetes can resist the proteases from the midgut epithelium that digest the bloodmeal and can pass through the peritrophic membrane formed around the bloodmeal by chitinase secretion. Then they cross the midgut epithelium and stop beneath the basal lamina and transform into oocysts (Baton and Ranford-Cartwright 2005). Between entering the mosquito midgut until oocyst formation the parasites are at a vulnerable stage; there is a bottleneck for the parasites to cross the midgut epithelium. From 10,000 gametocytes, only
0-5 oocysts form (Ghosh et al. 2000; Wang and Jacobs-Lorena 2013). Oocyst formation starts around 24-48h after the bloodmeal. Within each oocyst sporogony occurs where mitotic replication produces thousands of genomes. A sporoblast is formed and from this sporozoites are produced and released into the haemocoelic cavity. From there, some sporozoites end up invading the salivary glands and persist there. When the infected mosquito bites a human, some of the sporozoites are injected during the salivation part of bloodfeeding (Baton and Ranford-Cartwright 2005). The parasites then travel in the blood to the liver. Here they invade hepatocytes and within these multiply and differentiate into merozoites (Tilley et al. 2011). The thousands of merozoites are released into the bloodstream and invade red blood cells. Within the red blood cells, the parasites develop through three stages. The first stage is the ring stage, then the trophozoite stage where growth occurs, and lastly the schizont stage where division occurs. At the end of this asexual cycle that takes around 48h, the erythrocyte ruptures releasing about 20 merozoites. Some of the intraerythrocytic parasites differentiate into male and female gametocytes, which when taken up in a bloodmeal by a mosquito starts a new transmission cycle (Tilley et al. 2011).

Control methods

Methods for malaria control can be divided into two approaches, prevention including vector control, prophylaxis and vaccines, and treatment including e.g. drugs and blood transfusions (Enayati and Hemingway 2010). Prevention methods in Africa almost completely rely on insecticides in the form of longlasting insecticide-treated bednets and indoor residual spraying (Enayati and Hemingway 2010). The use of these preventive methods has proven to be very successful with the number of deaths due to malaria halving in the period 2000-2015 (Cibulskis et al. 2016). However, both longlasting insecticide-treated bednets and indoor residual spraying used mostly insecticides of the class pyrethroids. This has lead to the selection of resistance in the mosquito vectors with the result of pyrethroids no longer killing mosquitoes in some parts of Africa (Hemingway et al. 2016a). The decrease in the number of deaths can also be attributed to the use of effective drugs against \textit{P. falciparum} in the form of mainly artemisinin-based combination therapies (ACTs). However, also here resistance is developing as a problem, this time in the parasites against the drugs. In Southeast Asia, ACTs are now failing against \textit{P. falciparum} (Haldar et al. 2018). Furthermore, \textit{Anopheles} are changing their feeding behavior to bypass the bednets (Ojuka et al. 2015). This, together with the development of resistance against the existing interventions, might increase the burden of malaria again. Actually, the number of malaria cases increased by 5 million in 2016 compared to 2015 (WHO 2017). Therefore, novel methods to control malaria are needed. At the mo-
ment new tools for detection, treatment, and prevention of malaria are under development (Hemingway et al. 2016b).

One area of intervention methods being investigated is non-insecticidal vector control methods. The aim of these methods is to reduce transmission by reducing population size, shorten lifespan, or inhibit ability to transmit malaria parasites by elimination of the parasites in the vectors (Hemingway et al. 2016b). It has been shown that bacteria naturally can shorten the lifespan of *Anopheles* mosquitoes (Bahia et al. 2014; Ramirez et al. 2014) and inhibit *Plasmodium* infection in mosquitoes. For example, bacteria have been shown to have both direct inhibitory effects on the parasites by production of reactive oxygen species (Cirimotich et al. 2011) and other metabolites (Bahia et al. 2014; Ramirez et al. 2014), and indirect inhibitory effects by activating the mosquito immune system, which in turn inhibits the parasites (Meister et al. 2009). Genetic approaches have also been suggested for this purpose, both involving the vector mosquitoes themselves (transgenesis) and the bacteria associated with the mosquitoes (paratransgenesis) (Hemingway et al. 2016b). Transgenesis is the genetic transformation of the vectors to produce anti-*Plasmodium* effector molecules while paratransgenesis is the genetic transformation of mosquito-symbiotic bacteria or other microorganisms to produce anti-*Plasmodium* effector molecules (Wang and Jacobs-Lorena 2013). For the development of these methods involving bacteria, it is necessary to understand their role in mosquito physiology, effects on development and immunity, and also direct interactions with transmitted pathogens like malaria parasites. A first step is to investigate which bacteria are naturally present in *Anopheles* mosquitoes and where they come from, to further down the line be able to manipulate their functions or community composition for the purpose of malaria control.

**Bacteria associated with *Anopheles***

Bacteria associated with *Anopheles* mosquitoes have been studied at both the larval stage and the adult stage, as well as in the larval breeding water. Studies have investigated both field collected and laboratory reared *Anopheles*. Furthermore, the mosquitoes have been investigated by different methods due to technical advances in the field of DNA sequencing. This section will focus on studies investigating bacteria associated with vector mosquitoes, mainly *Anopheles*.

**Adult midgut bacteria**

Several studies have been performed with the aim of investigating the effect of bacteria on *Plasmodium* parasites in adult malaria-mosquito midguts. To investigate this, relevant bacteria present naturally in mosquito guts also
needed to be identified. In early studies, mosquitoes or dissected midguts were homogenized and cultured on agar plates and identified by biochemical analyses. For example, laboratory-reared *An. stephensi* were studied and three different isolates were identified. In the same study, Gram-negative but not Gram-positive bacteria were found to inhibit *P. falciparum* sporogonic development (Pumpuni et al. 1993). In laboratory-reared *An. stephensi*, *An. gambiae* and *An. albimanus*, 17-90% of the mosquito guts were found to contain bacteria and higher bacterial counts were associated with reduced oocyst densities in *P. falciparum* infected mosquitoes (Pumpuni et al. 1996). In field collected *An. funestus* from Kenya and *An. gambiae* s.l. from Kenya and Mali, 73 bacterial isolates from midguts corresponded to 20 genera identified by gas chromatography. Here, no correlation between gram-negative bacteria and *P. falciparum* sporozoites were observed. However, female *An. funestus* were more likely to be infected if they contained gram-positive bacteria. Only 8-29% of the mosquitoes were found to contain bacteria (Straif et al. 1998). Another study investigated field-caught *An. albimanus* from Mexico and found that *P. vivax* development could be inhibited by some bacteria, though high concentrations were needed, which also caused mosquito mortality (Gonzalez-Ceron et al. 2003).

With the development of DNA methods, *An. gambiae* s.l. and *An. funestus* from Kenya were investigated by both culture-dependent- and culture-independent methods based on sequence analysis of the 16S rRNA genes. This resulted in identification of seven genera by each method, e.g. three genera of intracellular bacteria were identified the by culture-independent method (Lindh et al. 2005). In another study utilizing 16S rRNA gene libraries from three *An. stephensi* from an insectary in Italy, one genus was found to dominate, *Asaia*. *Asaia* was also found in field collected *An. maculipennis* from Italy and *An. gambiae* from Burkina Faso and on eggs, larvae, pupae and different tissues in adults. Furthermore, it was shown to be maternally transferred. These characteristics suggested it could be a candidate for paratransgenesis (Favia et al. 2007). On the same theme of paratransgenesis, another bacteria was identified by 16S rRNA sequences as dominant in field collected *An. gambiae* s.l. in Kenya, *Thorsellia anophelis*. *T. anophelis* was also found in the breeding water and shown to tolerate an alkaline environment, which is found in mosquito guts, and be able to utilize blood for growth. These characteristics suggested it could be a suitable bacterium for paratransgenesis to control malaria (Briones et al. 2008). Other studies investigating *Anopheles* midguts for paratransgenesis candidates include three studies from Iran (Chavshin et al. 2012; Chavshin et al. 2014; Dinparast Djadid et al. 2011). All studies used culture-based methods and PCR of the 16S rRNA gene. Dinparast Djadid et al. investigated *An. stephensi* and *An. maculipennis* adult females and larvae from the laboratory and field, and breeding site water and found the majority of sequences belonged to Gammaproteobacteria. Chavshin et al. (2012) investigated 96 wild female adult
An. stephensi from two locations in Iran. Five of twelve genera were found in both regions, coastal and hilly. Chavshin et al. (2014) investigated wild An. culicifacies, 34 females and 68 larvae in three biotopes in Iran. They found diversity varying with mosquito stage and area.

Further developments in DNA sequencing methods lead to the use of pyrosequencing of the 16S rRNA gene. This resulted in a number of studies about bacteria in malaria mosquitoes (Akorli et al. 2016; Boissière et al. 2012; Coon et al. 2014; Gimonneau et al. 2014; Onchuru et al. 2016; Osei-Poku et al. 2012; Sharma et al. 2014; Tchioffo et al. 2015; Wang et al. 2011). For example, Wang et al. investigated the bacteria associated with the life history of An. gambiae in semi-natural habitats in Kenya. They found that OTU diversity was highest in the breeding water and then decreased in the order larvae, pupae, sugar-fed females and blood-fed females. Osei-Poku et al. looked at eight different mosquito species of non-bloodfed females in Kenya including Anopheles, Aedes, Culex and Mansonia. Different mosquito species had similar gut bacteria but individuals in a population had very variable bacteria. The sampling location showed no significant correlation with gut composition. However, in the study by Boissière et al., An. gambiae larvae collected in two locations in Cameroon and reared to adults in an insectary differed according to breeding site suggesting that the native aquatic breeding source determines the composition.

Larval bacteria

Anopheles larvae from different parts of the world have also been studied for bacterial composition. In Iran, 100 wild An. stephensi larvae from two locations were found to contain almost twice the number of species as adults, 40 species/strains in larvae compared to 25 in adults (Chavshin et al. 2012). In Kenya, An. gambiae larvae were also found to contain more bacteria than adults but it was also shown that they contain fewer bacterial OTUs than the breeding water (Wang et al. 2011). The finding that larvae contain fewer bacterial species than their breeding water has also been shown in other studies (Coon et al. 2014; Gimonneau et al. 2014). This suggests that the larval gut is selective. Also the observation that larvae contain more bacterial species than adults was shown in the study by Gimonneau et al. The reduction in gut bacteria from the aquatic life stages to terrestrial life stages has been suggested, for example by complete gut sterilization (Moll Coon et al. 2014 2001). However evidence for transstadial transfer of bacteria from larvae to adult mosquitoes has been presented at least in Aedes mosquitoes (Coon et al. 2014).
Breeding-water bacteria

The bacterial communities found in *Anopheles* breeding water have been reported a couple of times. Overall there is a large diversity found. Investigated sites include breeding water in Iran (Dinparast Djadid et al. 2011), field sites in Cameroon (Gimonneau et al. 2014), rearing pans in the laboratory (Coon et al. 2014) and several sites in Kenya; rice paddies (Briones et al. 2008), semi-natural habitats (Wang et al. 2011) and field sites (Onchuru et al. 2016). In Cameroon, water samples from different depth were analyzed. This showed that the bacterial composition differs between the upper surface microlayer, where *Anopheles* feed, and the subsurface layer (Gimonneau et al. 2014).

Field vs laboratory

Several studies have investigated the bacterial composition in field-collected mosquitoes and compared it to that in laboratory-reared mosquitoes. Most studies have found a difference with lower diversity in laboratory-reared mosquitoes. Gonzalez-Ceron et al. found no bacteria in insectary *An. albimanus* while some were found in field-caught *An. albimanus* from Mexico. In India, it was also found that laboratory and field-collected *An. stephensi* differed in diversity with a higher diversity in field-collected mosquitoes (Rani et al. 2009). In Kenya, laboratory and field-collected adult mosquitoes showed some similarities in gut communities at family level. However, also in this study diversity was lower in laboratory-reared mosquitoes (Wang et al. 2011). In a following study, field and laboratory mosquitoes in Cameroon differed drastically, with a loss of diversity in the laboratory mosquitoes. The composition changed from 94% Proteobacteria in field collected *An. gambiae* to 96% Flavobacteria in laboratory-reared *An. gambiae* (Boissière et al. 2012).

Identification method of bacteria

The methods used for identifying bacteria in and associated with malaria mosquitoes have changed with time. In the beginning of the development of the field, culturing of bacteria and identification based on biochemical test was used. With the development of DNA techniques, culture-independent studies became the standard as it has long been recognized that not all bacteria can be cultured (Staley and Konopka 1985). The 16S rRNA gene is now used as a phylogenetic marker and sequenced using high throughput sequencing technologies, which generates millions of reads per sample. Based on this it is easy to understand that the number of bacteria identified in mosquitoes has changed. From first believing that not all mosquitoes carry gut
bacteria, it is now known that they contain diverse bacterial communities and several bacterial species are identified in each sample.
Aims of the thesis

The overall aim of this thesis was to explore which bacteria are associated in different ways with *Anopheles* mosquitoes around the world and how these bacteria are acquired. The purpose of this was to contribute to knowledge useful for new effective interventions against malaria. The specific aims were to:

*Paper I*
- Characterize bacterial community composition in domestic water-storage containers and correlate this habitat characteristic to the absence or presence of vector mosquitoes.

*Paper II*
- Characterize and compare the bacterial community composition in *An. darlingi* breeding waters.

*Paper III*
- Investigate bacterial community composition associated with *Anopheles* larvae and the effect of breeding site and larval developmental stage.

*Paper IV*
- Investigate if bacteria associated with adult *Anopheles* mosquitoes reflect acquisition of bacteria in different environments.

*Paper V*
- Investigate the effect of environment and genetics on the bacterial gut composition in vector mosquitoes.
Materials and methods

Sample collection

Collections of samples were made in the field for studies I-IV in order to investigate populations of bacteria associated, in different ways, with malaria mosquitoes (and Aedes mosquitoes for study I) in their natural environments. For studies I and II, surface water of Anopheles and Aedes breeding sites were sampled. In study I, this was done by placing a steal mesh on the surface to collect the top layer of the water (Briones et al. 2008). In study II, this was done by using a hand dipper to scrape up the water along the surface. For study III, Anopheles larvae were collected from their breeding waters. This was performed by the use of a hand dipper technique. Examples of the collection sites are shown in Fig. 2.

Figure 2. Examples of field collection sites. Upper left: Domestic water-storage container, India (Nilsson et al. 2018). Upper right: Lake, Amazon region of Brazil. Personal photograph. Lower middle: River fringe, Ethiopia. Illustration from paper III.
For study IV, adult *Anopheles* mosquitoes were collected. This was done inside houses by mouth aspiration. Individual mosquitoes were collected and stored separately. Collections of samples for study V were performed in the laboratory where the experiment was carried out.

**DNA extraction, PCR, and MiSeq sequencing**

All samples (water, larvae, adult mosquitoes, guts of adult mosquitoes, and sugar pads) were DNA extracted, except for the water samples in study I. Different kits were used in different studies depending on availability. Extracted environmental DNA was subjected to PCR amplification targeting the 16S rRNA gene of the bacteria in the samples. The 16S rRNA gene is around 1600 nucleotides long and the molecule is one of three rRNAs in bacteria. Together they form parts of the ribosomes that are essential for protein synthesis (Olsen et al. 1986). All bacteria have this gene, which allows analysis of phylogenetic relationships. It is made up of highly conserved sequences as well as more variable ones. The conserved regions provide sequences to which PCR primers can be designed that will amplify most bacteria while the interspersed variable regions provide information for classification of the bacteria (Vetrovsky and Baldrian 2013). The PCR amplifications in all studies included in this thesis were performed by a two-step method using the primers 341F (5’-CCT ACG GGN GGC WGC AG-3’) and 805R (5’-GAC TAC HVG GGT ATC TAA TCC-3’) (Sinclair et al. 2015). These primers amplify the variable regions V3-V4 of the 16S rRNA gene, *Escherichia coli* position 341-805. In the first step of the PCR method, amplification was performed with these primers together with the extracted DNA or water from the samples as templates. In the second step of the PCR method, the PCR products from the first step were provided as the templates. These were amplified by 10 additional cycles of PCR with the same primers (341F and 805R) except for the addition of seven nucleotide-barcodes at the ends. In order to run 50 different samples as one sample on the sequencer, one of 50 unique combinations of forward and reverse primers were used per sample. These pools of 50 different samples were then purified and submitted to a sequencing facility, SNP/SEQ Technology platform in Uppsala, Sweden, where both library preparation and paired-end sequencing with 300 bp read length was performed on the Illumina MiSeq system (Illumina, San Diego, CA, USA). The raw MiSeq sequences for all studies included in this thesis are or will be deposited in the European Nucleotide Archive (ENA) database hosted by the European Bioinformatics Institute (EBI).
Bioinformatics

From the fastq files produced by the sequencing facility, different bioinformatic tools were utilized in order to produce an Operational Taxonomic Unit (OTU) table and corresponding taxonomy file. The OTU table is a table showing the number of reads of each OTU there is in each sample. All following analyses are based on this information. In order to go from fastq files to an OTU table there are a number of steps that need to be taken. In summary, first the paired end reads need to be merged to create longer sequences based on both the forward and the reverse reads. Then the samples need to be demultiplexed, which means the reads are separated into the sample they came from. The primers and barcodes are removed from the reads. The reads are quality filtered to remove reads with a low quality that might be wrong. Then the reads that are left in the analysis are clustered into OTUs based on similarity of the reads. Often 97% sequence similarity is used as a threshold and singletons, OTUs with a single read, removed. Chimeras, which are artifacts produced during PCR are also removed. They are formed when an extension of a sequence is terminated too early and the resultant product anneals to another parent sequence and initiates synthesis, creating a hybrid of multiple parent sequences. This can lead to falsely high diversity and the interpretation of novel bacteria (Haas et al. 2011). From each OTU, one representative sequence is picked as the reference sequence for that OTU. Finally, taxonomy is assigned to the OTUs using a reference database with 16S rRNA gene sequences. Characterization of unknown bacteria using 16S rRNA gene sequences requires a reference database of sequences from known bacteria to compare to (Olsen et al. 1986).
Present investigations

Paper I | Aquatic bacteria and the presence of larvae

Bacteria have been shown to both attract and repel ovipositing mosquitoes (Huang et al. 2006; Ponnusamy et al. 2008; Sumba et al. 2004). Therefore, bacteria associated with breeding waters of vector mosquitoes could represent those that attract ovipositing mosquitoes, while bacteria associated with waters without mosquito larvae could represent those that repel them. Knowledge of this could potentially be utilized in vector control strategies, e.g. to direct mosquitoes to traps or repel them from breeding sites near human habitations. Our goal was therefore to investigate potential breeding sites for bacteria that differ between the presence and absence of *Anopheles* larvae. We hypothesized that the bacterial composition in sites with mosquito larvae present would represent suitable environmental conditions for the larvae while the opposite would be true for sites without larvae present.

To investigate this, we sampled water from domestic water-storage containers in a village in India both when larvae were present and when they were absent. The bacterial communities in the water samples were identified and analyzed by 16S rRNA gene sequencing. In our study sites, we found not only *Anopheles* larvae but also *Aedes* larvae. Therefore, we included them also in the analysis to see if there was any difference in bacteria associated with another genera of vector mosquitoes and if some bacteria might attract or repel both types of vector larvae. *Aedes* are vectors of human disease like dengue, yellow fever, chikungunya and Zika. The overall distribution of the identified OTUs was shown in a Venn diagram (Fig. 3).
Figure 3. Distribution of all OTUs identified in the water containers. Each OTU is represented once in the Venn diagram. In areas of overlap, the OTUs were found in water containers in both/all three categories. The area of overlap between *Aedes* present and *Anopheles* present contains OTUs found in either water with both *Aedes* and *Anopheles*, in water with *Aedes* only and *Anopheles* only, or a combination of these. n = number of water samples in each category (Nilsson et al. 2018).

We found that *Anopheles* and *Aedes* were more common when Lachnospiraceae, Synechococcaceae, Alcaligenaceae and Cryomorphaceae were present in the water. While Xanthomonadaceae, Comamonadaceae and Burkholderiaceae were more common when *Anopheles* and *Aedes* were absent from the water. When separating the mosquito genera, it was shown by indicator species analysis (Dufrene and Legendre 1997) that different OTUs were indicating the presence and absence of the different genera. However, a couple of OTUs were found to indicate both *Anopheles* and *Aedes* presence (assigned as FukuN101 in the family Microbacteriaceae and as GKS98 freshwater group in the family Alcaligenaceae) and one OTU was found to indicate both *Anopheles* and *Aedes* absence (assigned as CL500-29 marine group in the family Acidimicrobiaceae). We also found that the bacterial community composition in the water was similar in the water-storage containers based on date, more so than based on site across the sampling period.
While the breeding water of *Anopheles* has been investigated for the bacterial community composition in several studies of the Old World (Briones et al. 2008; Dinparast Djadid et al. 2011; Gimonneau et al. 2014; Nilsson et al. 2018; Onchuru et al. 2016; Wang et al. 2011), very little information exists about bacteria in *Anopheles* breeding waters in the New World. Therefore, we aimed to explore the bacterial community composition of *An. darlingi* breeding sites in the Amazon region of Brazil. *An. darlingi* is the principal malaria vector in the Amazon basin (Hiwat and Bretas 2011) where most cases of malaria on the American continent occur (Recht et al. 2017). We hypothesized that bacteria and the conditions they indicate would be similar in all locations as they all were characterized by large abundance of *An. darlingi* larvae.

To investigate this, we sampled four breeding sites and analyzed the bacterial composition within them. Analysis was based on 16S rRNA gene sequencing. We found that the sites were similar in bacterial composition to each other and to breeding sites of Old World *Anopheles* mosquitoes at high taxonomic levels. Of the total number of reads, 94% belonged to one of 36 OTUs identified in all sites. However, the bacterial composition diverged between sites at lower taxonomic levels. Furthermore, one of the sites (site 1) differed significantly in bacterial community composition from the rest (Fig. 4). This site also had a higher alpha diversity than the other sites.

To find out which OTUs were separating the sites, an indicator species analysis was performed (Dufrene and Legendre 1997). This identified some OTUs that differed between the sites. These belonged to the orders Burkholderiales for site 1, Actinomycetales for sites 2 and 3, and Clostridiales for site 4. The overall most common OTUs in all breeding sites were assigned to *Escherichia/Shigella*, *Staphylococcus*, and *Pseudomonas*. 
Figure 4. Bacterial community composition in different *An. darlingi* breeding sites at different taxonomic levels. Upper panel: Phylum level, “Other”=unknown phylum. Middle panel: Class level. Only classes making up >0.1% in any sample are named, other classes present are clustered as “Other” together with unknown classes. Lower panel: Family level. Only families making up >1% in any sample are named, other families present are clustered as “Other” together with unknown families. Illustration adapted from paper II.
During the larval stages, *Anopheles* feed on microorganisms like bacteria and detritus in the air-water interface in the surface microlayer of their breeding water (Merritt et al. 1992; Wotton et al. 1997). It has been seen that adult mosquitoes collected from different breeding sites as larvae contain different bacteria according to which breeding site they were collected from (Boissière et al. 2012; Gimonneau et al. 2014; Tchioffo et al. 2015). This implies that the environment of their breeding sites is important for the bacterial composition in the mosquitoes. Furthermore, this transfer of bacteria from larvae to adult mosquitoes has also been shown directly (Chavshin et al. 2015; Coon et al. 2014).

In this study we aimed to investigate the bacterial composition in *An. arabiensis* larvae and to investigate if the composition differs according to breeding site (both classified as individual sites and according to the aquatic type of the site) and according to larval developmental stage. We hypothesized that the individual breeding sites would have an effect on the bacteria associated with the mosquito larvae as it has been shown that adult mosquitoes have a distinct microbiota depending on what geographical area they originate from (Akorli et al. 2016; Buck et al. 2016). Furthermore, we also hypothesized that the aquatic type of the breeding site would have an effect on the bacteria associated with the mosquito larvae as it has been shown that adult mosquitoes have a distinct microbiota depending on if the breeding site is a temporary or semi-permanent site (Tchioffo et al. 2015). Our hypothesis regarding the effect of developmental stage was that early and late instar larvae would differ in their bacterial composition as this has been shown in *Culex* mosquitoes (Duguma et al. 2015). Besides analyzing these effects on bacterial community composition in larvae, we also investigated if sequencing a pool of individual larvae would be representative of all the individuals in the pool. Often samples contain more than one individual to provide enough DNA for analysis. Therefore, our aim was to see if pooled samples of individuals reflect the individuals within.

To do this, bacterial composition associated with individual *An. arabiensis* larvae as well as pools of individual larvae from Ethiopia were analyzed using 16S rRNA gene sequencing by MiSeq. In total, larvae from nine different breeding sites were analyzed. We found that mosquito larvae from the same breeding site contained similar bacteria, which differed from the bacteria associated with larvae from the other sites. Furthermore, we classified the breeding sites according to our interpretation of the type of aquatic habitat and found that larvae from the same type of aquatic habitat contained similar bacteria. Of the total variance, 45% could be explained by breeding site and 36% by aquatic-habitat type. We also analyzed if the bacterial composition could be used to predict the breeding site and aquatic habitat type the larvae were collected from. Random forest modeling showed that both
types of classification could be accurately predicted (Fig. 5). However, classification of mosquitoes according to breeding site was slightly better than classification according to aquatic-habitat type, with a misclassification rate of 9% compared to 11%.

Figure 5. Random forest models predicting breeding site of An. arabiensis larvae based on their associated bacterial community composition. Upper panel: Prediction of breeding site, named B, D, I, K, N, Q, T, W, Z. Lower panel: Prediction of aquatic type of breeding site. Illustrations from paper III.
The developmental stage of the larvae could on the other hand not explain the bacterial composition. It was found that the larvae were associated with similar bacterial communities as first and second instars compared to third and fourth instars (Fig. 6).

**Figure 6.** Pie charts of bacterial families in early and late developmental stages of *An. arabiensis* larvae show very similar composition. Early larvae refer to 1st and 2nd instar larvae. Late larvae refer to 3rd and 4th instar larvae. Only families represented by >1% of all reads are represented by name. Illustration adapted from paper III.

Of all the OTUs identified, only one, classified as *Thorsellia*, was common to all larvae. Regarding the representativeness of pooled samples, we found that if larvae are pooled and sequenced together, the result represents all the individual larvae in the pool.
Paper IV | Bacteria associated with *Anopheles* adults

Adult mosquitoes are associated with a large variety of bacteria. However, fundamental knowledge of where malaria mosquitoes acquire their bacteria is lacking. We hypothesized that the microbiota associated with adult mosquitoes would reflect the different environments encountered by each mosquito. Each encounter leaving its mark by contributing to the bacteria associated with the mosquito.

To investigate this, we collected *An. coluzzii* (*An. gambiae* M form) adult mosquitoes from three villages in Burkina Faso. We analyzed the whole-body bacterial composition of the mosquitoes using amplicon sequencing of the 16S rRNA gene. The mosquito samples harbored a wide range of bacterial taxa and displayed clear individual differences. However, it was also seen that the mosquitoes clustered according to village with very few exceptions (Fig. 7).

![Figure 7. NMDS plot showing distinct clustering of mosquitoes according to where *An. coluzzii* were captured. Analysis was based on all bacterial OTUs in the samples.](image)


The sub-populations in the different villages were clearly separated despite being within flight distance from each other, however, a couple of samples from the village VK3 had a bacterial flora more similar to the bacteria associated with mosquitoes of villages VK5 and VK7. Our hypothesis is that this indicates that there is some interaction between the villages and therefore suggests the existence of local populations within a meta-population net-
work. This separation of mosquitoes based on their associated bacteria was further confirmed by a random forest model. The model showed that based on the bacterial composition, the village a mosquito was captured in could in 90% of the cases be correctly predicted. The most important determinants of the random forest classification were OTUs assigned to the genera Massilia, Wolbachia, Shewanella and Acinetobacter. These four genera were likely acquired at different time points during the mosquito life history. Wolbachia is maternally transmitted via the egg and was mainly found in one village (VK5). This also confirmed the presence of Wolbachia in wild Anopheles mosquitoes, which was reported for the first time in 2014 (Baldini et al. 2014). Massilia and Shewanella were probably taken up from the breeding water. Both bacteria genera have been found to form biofilms in fresh water (Cheng et al. 2014; Liu et al. 2012), which is a common source of food for mosquito larvae (Merritt et al. 1992). Acinetobacter is common in nectar (Fridman et al. 2012) on which both male and female adults feed. The mosquitoes from one village (VK5) harbored a different set of Acinetobacter OTUs than the mosquitoes from the other villages. Since Acinetobacter species often are plant-specific (Alvarez-Perez et al. 2013; Fridman et al. 2012) it may imply different nectar sources in the different villages and that the mosquitoes have fed near where they were captured. We also found that gravid and non-gravid females contained similar bacteria, which indicates that host seeking and blood feeding does not have a major effect on the general bacterial composition. It is known that the blood meal leads to rapid proliferation of mosquito-gut bacteria and favors some species that tolerate the new conditions (Tchioffo et al. 2015). However, our result suggests the overall bacterial composition returns to pre-blood meal composition after digestion of the blood meal. This was also seen by Tchioffo et al. for some bacterial genera in An. coluzzii and An. gambiae that decreased after blood feeding but at eight days post-blood-feeding had returned to the levels found before. We also investigated if three genera of human skin bacteria (Bacillus, Brevibacterium, and Corynebacterium) that have been shown to strongly attract female An. gambiae (Verhulst et al. 2010) were differentially found in the mosquitoes based on sex and diet. However, no correlation was found. On a genus level, we found that the most abundant genera, making up more than 40% of all sequences identified to genus level, were Thorsellia, Wolbachia, Massilia, and Acinetobacter. Of these, the most abundant was Thorsellia.
Paper V | Effect of genetics and environment on mosquito gut bacteria

The question of how the gut bacteria in adult mosquitoes are determined is not fully understood. Effects of both the environment, which the mosquitoes encounter during their lifetime (Buck et al. 2016), and of vertical transfer from parent to offspring (Baldini et al. 2014; Damiani et al. 2008) have been shown. The effect of mosquito species has also been investigated and most often no difference between species has been observed (Akorli et al. 2016; Coon et al. 2016; Gimonneau et al. 2014; Tchioffo et al. 2015). However, a difference was for example shown between *Anopheles* and *Aedes* compared to *Georgecraigius* mosquitoes (Coon et al. 2014) and between two *Culex* species, however, at different collection dates (Duguma et al. 2017). In this study, we therefore aimed to investigate the effects of genetics and environment on the gut bacteria in adult vector mosquitoes.

To investigate the effect of genetics, we included two vector mosquitoes, *An. gambiae* and *Ae. albopictus*, for comparisons between species. Then we designed a laboratory study that allowed us to compare the two mosquito species reared in the same environment (either shared larval media or a shared sugar source) with the same two mosquito species reared in separate environments (separate larval media or sugar sources). Adult mosquitoes were dissected and the midguts were analyzed by 16S rRNA gene sequencing to determine the bacterial composition. In addition, the larval media and sugar sources were also analyzed by 16S rRNA gene sequencing. The study was repeated over seven generations of mosquitoes in an insectary in Lebanon.

We found that by comparing the gut bacterial compositions of all *Anopheles* mosquitoes to all *Aedes* mosquitoes across all generations there was a significant difference in gut-bacteria composition between species. Furthermore, we also found a significant effect of both larval media and sugar source on the gut bacterial compositions when comparing all mosquitoes. This indicates that both genetics and environment has an effect on the adult mosquito-midgut bacterial composition. Additional analysis showed that two OTUs were significantly different between the midguts of the two mosquito species, identified as *Elizabethkingia* and *Wolbachia*. *Elizabethkingia* was more common in *Anopheles* guts while *Wolbachia* was more common in *Aedes* guts. Furthermore, results from the parts of the study where *Anopheles* and *Aedes* shared larval medium or sugar source highlighted the importance of the genetic component in determining the gut bacterial composition. When the two species shared medium, *Elizabethkingia* was one of the most common bacteria in the medium and also found as one of the most abundant OTUs in *Anopheles*. This is in agreement with previous studies of laboratory reared *An. gambiae* where *Elizabethkingia* has been found abundantly (Boissière et al. 2012; Chouaia et al. 2010; Dong et al. 2009). However, in *Aedes*
it was almost absent in the midguts. This is in line with the above identification of differential abundance of *Elizabethkingia* in the two species. Similarly, when *Anopheles* and *Aedes* adults were fed on the same sugar source, *Burkholderia* was identified to form a major part of the reads in the sugar. However, it was much more common in *Anopheles* than in *Aedes* midguts. On the other hand, when *Anopheles* and *Aedes* shared the same larval medium, they showed a more similar bacterial gut composition to each other as adults in comparison to adults of their own species reared in a different larval medium. Similarly, when adults of the two species shared the same sugar source, they showed a more similar bacterial gut composition to each other than to individuals of their own species feeding on a different sugar source. This highlights the effect of food source (environment) on the adult gut bacteria over the effect of mosquito species (genetics). Therefore, both host genetics (different species) and environment (different food sources), act to shape the composition of the midgut bacteria in mosquitoes. To investigate if one of the two components has a stronger effect, we compared dissimilarities in bacterial gut communities between mosquitoes of the same species but feeding on independent food sources to mosquitoes of different species feeding on the same food source. The result showed no significant differences. Hence, the midgut bacteria in adult mosquitoes appear to be as strongly affected by species as by food source. Another result from this study was that even though the environment was controlled in the form of the same insectary with the same temperature and humidity, and the same food source throughout the experiment, significant changes in the bacterial composition of larval media and sugar pads as well as mosquito midgut bacteria occurred across the seven generations.
Meta-analysis

This thesis includes work on identification of bacteria associated with *Anopheles* mosquitoes around the world (Fig. 8). In this section, I will present some results from meta-analysis of the most abundant and frequently present OTUs in the papers included in the thesis.

Figure 8. Map showing the locations of sample-collection sites for the studies included in this thesis. Created using Office templates.

Most abundant OTUs

In all papers (I-V), we have investigated the bacterial communities associated in some way with mosquitoes, mainly of the genus *Anopheles*. In papers I-IV, we investigated bacterial communities associated with samples collected in the field to investigate the naturally occurring bacteria. Paper V on the other hand, was a laboratory study where we tried to control the conditions to analyze the effect of environment and genetics on the mosquito-gut bacteria. The studies span different life stages of mosquitoes, from the breeding water to larvae to adults from different continents in the world. Nevertheless, investigating if there are any similarities seen in bacterial composition might reveal some general patterns. Summarizing the most abundant OTUs in each
paper and parts of each paper (Table 1), it can be seen that the same genus of bacteria (*Thorsellia*) is found most abundant in both *Anopheles* larvae and adults from the field. However, in the laboratory *Elizabethkingia* is the most abundant OTU. The Most abundant OTUs with origin in water differ in assigned taxonomy between the study sites. The abundance of the most abundant OTUs in the studies from India is much lower than in the other studies suggesting a higher diversity of bacteria in those sites. However this might be explained by the methods used. In the study from India the OTUs were classified into OTUs at a level of 99% similarity while in the other studies the classification was at a level of 97% similarity.

*Table 1. Most abundant OTUs in papers I-V*

<table>
<thead>
<tr>
<th>Paper</th>
<th>Study site</th>
<th>Origin</th>
<th>Most abundant OTUa</th>
<th>Abundance (%)b</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>India</td>
<td>Water-all containers</td>
<td><em>Acinetobacter</em></td>
<td>3.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Water-without larvae</td>
<td>Gammaprooeobacteria</td>
<td>4.6</td>
</tr>
<tr>
<td>II</td>
<td>Brazil</td>
<td>Water (field sites)</td>
<td><em>Escherichia/Shigella</em></td>
<td>22.0</td>
</tr>
<tr>
<td>III</td>
<td>Ethiopia</td>
<td>Larvae</td>
<td><em>Thorsellia</em></td>
<td>10.9</td>
</tr>
<tr>
<td>IV</td>
<td>Burkina Faso</td>
<td>Adults</td>
<td><em>Thorsellia anophelis</em></td>
<td>10.2</td>
</tr>
<tr>
<td>V</td>
<td>Lebanon (laboratory)</td>
<td>Adult guts (<em>Anopheles</em>)</td>
<td><em>Elizabethkingia</em></td>
<td>35.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Adult guts (<em>Aedes</em>)</td>
<td><em>Wolbachia</em></td>
<td>43.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Water (rearing pans)</td>
<td><em>Microbacterium</em></td>
<td>31.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sugar (cotton pads)</td>
<td><em>Pantoea</em></td>
<td>32.1</td>
</tr>
</tbody>
</table>

a The taxonomic classification of the most abundant OTUs are given.
b Percentage of the total number of reads in the study that is assigned to each OTU.

Besides comparing the most abundant OTUs in the datasets, it is interesting to compare the most frequent OTUs in the datasets as these might differ. The most frequent OTUs might represent bacteria that are strongly associated with a particular type of samples, forming the core bacteria.

**Most frequently present OTUs in *Anopheles***

In papers III-V, we investigated bacteria associated with *Anopheles* larvae or adult mosquitoes. An interesting question is if there are any bacterial species that are always present in *Anopheles* mosquitoes. These might for example be targeted for paratransgenesis. Bacteria associated with both larvae, field collected adults and laboratory reared adults are compared here (Table 2).
### Table 2. Top five most frequently present OTUs in *Anopheles* mosquito samples

<table>
<thead>
<tr>
<th>Paper Study site</th>
<th>Origin</th>
<th>Most frequent OTUs&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Frequency&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Abundance (%)&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>III Ethiopia</td>
<td>Larvae</td>
<td>Proteobacteria; Gammaproteobacteria; Enterobacteriales; <em>Thorsellia</em></td>
<td>54/54</td>
<td>10.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Proteobacteria; Betaproteobacteria; Burkholderiales; Burkholderiales; <em>Ralstonia</em></td>
<td>53/54</td>
<td>10.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Actinobacteria; Actinobacteria (class); Micrococcales; Micrococcaceae; <em>Arthrobacter</em></td>
<td>53/54</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Proteobacteria; Alphaproteobacteria; Rhodobacteriales; <em>Rhodobacteraceae</em></td>
<td>51/54</td>
<td>2.1</td>
</tr>
<tr>
<td>IV Burkina Faso</td>
<td>Adults</td>
<td>Proteobacteria; Gammaproteobacteria</td>
<td>29/29</td>
<td>8.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Firmicutes; Bacilli; Bacillales; Staphylococcaceae; <em>Staphylococcus</em></td>
<td>29/29</td>
<td>1.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Actinobacteria; Actinobacteria (class); Propionibacteriales; <em>Propionibacterium</em></td>
<td>29/29</td>
<td>1.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Proteobacteria; Gammaproteobacteria; Alteromonadales; Alteromonadaceae; <em>Marinobacter</em></td>
<td>29/29</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Firmicutes; Bacilli; Bacillales; Staphylococcaceae; <em>Staphylococcus</em></td>
<td>29/29</td>
<td>0.5</td>
</tr>
<tr>
<td>V Lebanon</td>
<td>Adults (laboratory)</td>
<td>Bacteroidetes; Flavobacteriaceae; <em>Elleithkingia</em></td>
<td>93/96</td>
<td>35.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Firmicutes; Bacilli; Bacillales; Staphylococcaceae; <em>Staphylococcus</em></td>
<td>92/96</td>
<td>1.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriales; <em>Enterobacter</em></td>
<td>90/96</td>
<td>7.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Actinobacteria; Actinobacteria; Propionibacteriales; <em>Propionibacterium</em></td>
<td>90/96</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Proteobacteria; Betaproteobacteria; Burkholderiales; Burkholderiales; <em>Burkholderia-Paraburkholderia</em></td>
<td>89/96</td>
<td>3.8</td>
</tr>
</tbody>
</table>

<sup>a</sup> The taxonomic classification of the most frequent OTUs are given. The lowest taxonomic classifications are highlighted in bold.

<sup>b</sup> The number of samples in which the OTU is found/the total number of samples in the study.

<sup>c</sup> Percentage of the total number of reads in the study that is assigned to each OTU.
Two of the most frequently present OTUs (Table 2) were found to be the same as the most abundant OTUs (Table 1) in the same studies, identified as *Thorsellia* and *Elizabethkingia*. In the study from Burkina Faso, *Thorsellia* that was found as most abundant was not present in all samples. Instead, several other OTUs were identified in those samples. A couple of OTUs from adult field-collected mosquitoes and laboratory-reared mosquitoes were identified as belonging to the same genera, *Staphylococcus* and *Propionibacterium*. The finding of these could indicate a close relationship between them and *Anopheles* mosquitoes.

Most frequently present OTUs in *Anopheles* breeding water

In papers I, II, and V, we investigated bacteria associated with *Anopheles* breeding water. These studies represent very different environments, domestic water-storage containers, lakes and laboratory rearing pans. Common bacteria might indicate an association with *Anopheles* mosquitoes, and in the case of the field studies could indicate bacteria beneficial for *Anopheles* larvae. The similarities found include the presence of *Acinetobacter* and Sphingomonadales in both field sites and Microbacteriaceae in domestic water-storage containers and the laboratory (Table 3).

<table>
<thead>
<tr>
<th>Paper</th>
<th>Study site</th>
<th>Most frequent OTUs</th>
<th>Frequency</th>
<th>Abundance (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>India</td>
<td>Proteobacteria; <em>Gammaproteobacteria</em></td>
<td>22/24</td>
<td>1.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Proteobacteria; <em>Gammaproteobacteria</em></td>
<td>17/24</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Proteobacteria; Gammaproteobacteria; Pseudomonadales; Moraxellaceae; <em>Acinetobacter</em></td>
<td>16/24</td>
<td>3.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Proteobacteria; Alphaproteobacteria; Sphingomonadales; Sphingomonadaceae; <em>Novosphingobium</em></td>
<td>15/24</td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Actinobacteria; Actinobacteria (class); Micrococcaceae; <em>Microbacteriaceae</em></td>
<td>15/24</td>
<td>0.3</td>
</tr>
<tr>
<td>II</td>
<td>Brazil</td>
<td>Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; <em>Escherichia/Shigella</em></td>
<td>16/16</td>
<td>22.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Firmicutes; Bacilli; Bacillales; Staphylococcaceae; <em>Staphylococcus</em></td>
<td>16/16</td>
<td>19.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; <em>Pseudomonas</em></td>
<td>16/16</td>
<td>12.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Proteobacteria; Gammaproteobacteria; Pseudomonadales; Moraxellaceae; <em>Acinetobacter</em></td>
<td>16/16</td>
<td>2.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Proteobacteria; Alphaproteobacteria;</td>
<td>16/16</td>
<td>2.2</td>
</tr>
<tr>
<td>Country</td>
<td>Taxonomic Classification</td>
<td>Number of Samples</td>
<td>Percentage</td>
<td></td>
</tr>
<tr>
<td>---------------</td>
<td>------------------------------------------------------------------------------------------</td>
<td>-------------------</td>
<td>------------</td>
<td></td>
</tr>
<tr>
<td>Lebanon (laboratory)</td>
<td>Sphingomonadales; <strong>Sphingomonadaceae</strong>&lt;br&gt;Bacteroidetes; Flavobacteria; Flavobacteriales; Flavobacteriaceae; <em>Elleithkingia</em>&lt;br&gt;Proteobacteria; Betaproteobacteria; Burkholderiales; Comamonadaceae; <strong>Comamonas</strong>&lt;br&gt;Actinobacteria; Actinobacteria; Micrococcales; Microbacteriaceae; <strong>Microbacterium</strong>&lt;br&gt;Bacteroidetes; Sphingobacteria; Sphingobacteriales; Sphingobacteriaceae; <strong>Sphingobacterium</strong>&lt;br&gt;Proteobacteria; Betaproteobacteria; Burkholderiales; <strong>Oxalobacteraceae</strong></td>
<td>21/21 18.3</td>
<td>21/21 12.3</td>
<td>21/21 11.9</td>
</tr>
</tbody>
</table>

a The taxonomic classification of the most frequent OTUs are given. The lowest taxonomic classifications are highlighted in bold.

b The number of samples in which the OTU is found/the total number of samples in the study.

c Percentage of the total number of reads in the study that is assigned to each OTU.
General discussion

We investigated breeding waters of Anopheles (mainly An. stephensi) and Aedes (mainly Ae. aegypti) in domestic water-storage containers in India and of natural breeding sites of An. darlingi in the Amazon region of Brazil. Bacteria associated with breeding waters of vector mosquitoes could represent those that attract ovipositing mosquitoes, opposite to bacteria that might repel them. Knowledge of this could be applied to direct mosquitoes to traps or repel them from breeding sites near human habitations.

The families of bacteria found to be associated with vector larvae in domestic water-storage containers in India were however not found (except four reads) in the An. darlingi breeding sites in Brazil. Of the bacterial OTUs identified as indicators of breeding water with An. stephensi and An. subpictus in India, mainly one of the families those indicator OTUs were assigned to (Pseudomonadaceae) was found in Brazil. The others were either not found or only found by a few reads. Furthermore, two families identified in India as more common in the absence of larvae (Comamonadaceae and Burkholderiaceae) were found in the breeding waters with Anopheles in Brazil, though mainly in site 1. These differences in bacterial composition in water with Anopheles larvae in the two studies are probably due to the different types of water in the studies. In India, chlorinated tap water in domestic water-storage containers compared to in Brazil, lakes/dams. The differences might also be linked to different Anopheles species being investigated. Also, the overall composition of the bacterial communities might be more important than individual bacterial species, which might provide the same functions for the larvae. Other factors besides bacteria are also important for mosquito larvae and may differ between species, e.g. higher temperature for An. gambiae and higher ammonium and phosphate for Ae. aegypti (Onchuru et al. 2016). Furthermore, we found that bacterial composition changed with date in India, suggesting a dynamic bacterial composition. Related to this, the one site in Brazil that differed most from the rest was sampled at a different time (3 months earlier), though no difference in temperature and rainfall was seen between the dates.

In laboratory mosquito-rearing pans, the water contained bacteria, which was shown to affect the composition in the mosquito guts. However, it was also found that the mosquito species affected the bacterial composition in the rearing pans, suggesting there is a bidirectional effect. This correlation of bacteria found in the breeding water and in the mosquitoes is seen clearly as
the most frequent OTU in *Anopheles* rearing pans was also the most frequent and abundant OTU in *Anopheles* guts (Tables 1-3).

Similar to bacterial composition in different field sites in Brazil, the bacterial composition in larvae from different sites differed between breeding site of origin. Three sites in Brazil contained similar bacteria while the third differed significantly. From the finding in Ethiopia that larvae from similar types of aquatic habitats also contain similar bacteria perhaps the sites in Brazil could be classified according to some aquatic feature that would separate site 1 from the rest of the sites. It would be interesting to also study the larvae from these sites to see if they would separate in the same way as the water samples did, based on bacterial composition.

Bacteria associated with adult *Anopheles* mosquitoes were investigated both in the field in Burkina Faso and in the laboratory in Lebanon. From the meta-analysis, it was found that in Burkina Faso, *Thorsellia* was the most abundant OTU while in the laboratory *Elizabethkingia* was the most abundant and frequent (Tables 1 and 2). Interestingly, *Thorsellia* was also found to be the most abundant and most frequent genus identified in Ethiopian field collected larvae making up >10% of the total number of reads in both larvae and adults. This suggests a strong association between *Anopheles* and *Thorsellia*. In addition, *Thorsellia* was found in the breeding sites in Brazil, though only with a few reads. *Thorsellia* is a genus of bacteria that previously has been identified only in vector mosquitoes (mainly in *Anopheles*) and their breeding waters (Akorli et al. 2016; Briones et al. 2008; Buck et al. 2016; Chavshin et al. 2014; Duguma et al. 2013; Kampfer et al. 2006; Kampfer et al. 2015; Lindh et al. 2005; Ngo et al. 2015; Rani et al. 2009; Segata et al. 2016; Wang et al. 2011). Besides being present in the guts of mosquitoes, *Thorsellia* has also been found in the reproductive tissues of *Anopheles* suggesting the possibility of vertical transmission (Segata et al. 2016). Based on this, *Thorsellia* could be a candidate for paratransgenesis for control of malaria.

On the other hand, *Elizabethkingia* has often been found to be associated with laboratory reared *Anopheles* (Boissière et al. 2012; Chouaia et al. 2010; Dong et al. 2009). The bacteria associated with laboratory-reared mosquitoes have been found to be less diverse than in the field (Boissière et al. 2012). Therefore, one should be careful when investigating bacterial communities associated with laboratory-reared mosquitoes as they might differ from wild mosquito populations.

Another mosquito life event that might affect the gut-bacterial composition is the adult feeding. Both males and females feed on sugar sources while females also feed on blood. Gut-bacterial composition might reflect feeding on different nectar sources. In the laboratory study we found that the bacterial composition in the sugar pads had a greater effect on the adult gut microbiota than the water in the rearing pans. This might have been observed in the study of adult mosquitoes from Burkina Faso. There we found *Acineto-
bacter, which has been found in nectar, more abundant in males than in females overall which could reflect their different food sources. Furthermore, different OTUs of *Acinetobacter* were identified in mosquitoes from one of the villages, which could reflect acquisition from a different plant. The effect of blood feeding has been shown to be a decrease in the number of OTUs (Coon et al. 2014; Tchioffo et al. 2015) while certain bacteria proliferate greatly (Wang et al. 2011), increasing the overall number of bacteria (Pumpuni et al. 1996). However, we found no difference in bacterial community composition in bloodfed and non-bloodfed female mosquitoes, suggesting no major impact of host seeking and blood feeding on the overall bacterial composition.
Conclusions and future perspectives

In order to eradicate malaria, several strategies will most likely be needed. In addition to insecticides, drugs and vaccines, bacteria might be useful. Bacteria can have many effects on vector mosquitoes, from directing ovipositing mosquitoes to development of larvae to affecting human pathogens in the adult mosquitoes. The study of bacteria associated with vector mosquitoes at all these stages can therefore be useful for development of new intervention methods to prevent transmission of malaria, as well as other diseases, spread by mosquitoes. From this thesis, we conclude that:

I Correlations between certain bacteria and presence, as well as absence, of *Anopheles* and *Aedes* larvae in domestic water-storage containers can be found. However, future studies are needed to investigate these bacteria in more detail to confirm if they have any functions in attracting or repelling vector mosquitoes.

II *An. darlingi* larvae can develop in surface water with different bacteria, based on our findings of diverging bacterial communities in their breeding sites. However, the bacteria common to all sites might contribute to the suitability of the habitats. Further studies on bacterial communities in *An. darlingi* breeding waters are needed to conclude if there are any connections between breeding site bacteria and the presence of *An. darlingi*.

III Bacteria associated with *An. arabiensis* larvae can be explained to a large extent by the mosquito-breeding site. Both in terms of individual sites but also in terms of classification according to the type of site. For future studies, it would be interesting to investigate mosquito larvae from more breeding sites of similar aquatic types and even further apart to see if this pattern can be seen on a larger scale. If so, this might make it possible to determine from what type of breeding site a mosquito originates from based on its gut bacteria. Then vector control measures could be directed towards those sites.

IV Bacteria can reflect the village mosquitoes are collected from and can be used to assign the mosquitoes to different populations. This finding that bacteria can be used to gain information on population structures could be useful in aiding vector interventions, as information about the spatial structure of vector populations is important for choosing effective intervention methods. The data suggest that the mosquitoes form local meta-
populations in neighboring villages with restricted overlap. If this is true for other regions as well, intensified village-by-village malaria intervention could be more applicable than was considered previously. Future studies are however needed to support the generality of this finding. As well as studies on environmental microbiota for more environment-specific information. This type of method determining where an insect is coming from could be extended into other areas of research as well.

V By controlling the environmental conditions in a laboratory setting, we could attribute differences in bacterial community composition in adult female-mosquito guts to both the environment in the form of food source and host genetics in the form of mosquito genus. When comparing the relative influence of each, no significant difference was found suggesting equal contribution of genetics and environment on the midgut microbiota.
Svensk sammanfattning


I denna avhandling undersöker vi bakteriesamhällen i vektormyggor, i huvudsak malariamyggor, och i vattnet där äggen kläcks och larverna och pupporna utvecklas innan de är färdigutvecklade vuxna myggor. Vårt mål var att lära oss mer om vilka bakterier som finns naturligt inuti myggorna.
och i deras omgivning och vad som påverkar vilka bakterier som utgör myggnas tarmflora. För att undersöka hela bakteriesamhällen i myggnorna och deras vatten använde vi oss av en molekylärbiologisk teknik där vi amplifis-

I artikel I undersökte vi bakteriesamhällen i vattentankar i hushåll i Indien som ibland innehöll malariamyggor och myggor av ett annat vektormygg-
släkte kallat Aedes som sprider sjukdomar som dengue och Zika. Här fann vi att förekomsten av vissa bakterietyper korrelerade med förekomsten av vektormygger medan andra korrelerade med frånvaron av myggnorna. Möj-
ligtvis skulle det kunna tyda på någon repellerande eller attraherande effekt av bakterierna på myggnorna som lägger ägg, vilket skulle kunna utnyttjas för att minska kontakten med människor.

I artikel II undersökte vi bakteriesamhällen i vatten i Brasilien, Amazo-
nas, kända för förekomst av malariamygglarver. Här fann vi att tre av fyra sjöar innehöll liknande bakterier medan den fjärde skilde sig åt. Detta kan tyda på att Anopheles darlingi, malariamyggarten vi undersökte, kan utveck-
las i vatten med olika typer av bakterier.

I artikel III undersökte vi bakteriesamhällen i malariamygglarver i Etio-
pien från olika platser. Här fann vi att larver från samma plats innehöll lik-
nande bakterier som skiljde sig från larver från andra platser. Dessutom fann vi att om vi klassificerade platserna i kategorier enligt typ av plats så inne-
höll larverna också liknande bakterier som skiljde sig från bakterierna i lar-
verna från de annorlunda klassificerade platserna. Vi kunde även visa att det var möjligt att förutsäga vilken plats en mygga kom ifrån baserat på dess bakterier.

I artikel IV undersökte vi vuxna myggor fångade i Burkina Faso i tre olika byar. Här fann vi också att myggor från samma plats, by, innehöll lik-
nande bakterier som skiljde sig åt från bakterierna i myggor från de andra byarna. Också här gick det att förutsäga från vilken by en mygga kom base-
rat på dess bakterier. Dessutom såg vi att det verkade som myggorna bildade lokala populationer i en metapopulation med begränsad rörelse mellan popu-
lationerna, byarna. Om detta stämmer skulle det kanske vara möjligt att fo-
kusera på malariakontroll i ett område i taget.

I artikel V undersökte vi på laboratorium effekten av arv och miljö på bakterierna i malariamyggor och Aedes myggor. Vi fann att både arv (i form av myggart) och miljö (i form av diet) påverkade vilka bakterier som fanns i magen hos vuxna honor. Det gick inte att säga vilken effekt som var starkast. Bakterier i sockerlösning hade också en större påverkan på bakterierna i myggnorna än bakterier i larvernas vatten.

En intressant upptäckt var att den bakterie som återkom som vanligast i både larver och vuxna myggor var av samma släkt, Thorsellia. Detta är en
släkt av bakterier som enbart hittats i vektormyggor och deras vatten vilket skulle göra den till en kandidat för paratransgenes.

Sammanfattningsvis bidrar denna avhandling till en utökad kunskap om mångfalden av bakterier associerade med malariamyggor och en bättre förståelse för hur bakteriessamhällen i myggor bestäms, och på detta sätt bidra till utvecklingen av nya metoder för bekämpning av malaria.
Acknowledgements

First and foremost, I would like to thank my main supervisor Olle. Without you I would probably never have become a PhD student. So thank you for believing in me in the first place and along the way. It has definitely been a journey, taking me around the world, such a privilege. I much appreciate your constant positivity, support, and for showing me a good work-life balance. I am grateful for all I have learnt.

I would also like to thank my co-supervisors. Sebastian, besides Olle you are the only person starting and finishing this project with me. Thank you for sticking around, your good ideas and nice company in Manaus. Staffan. I would very much like to thank you for taking me (and Olle) into your group. Thank you for your kindness, problem-solving attitude and enthusiasm. It has been a pleasure to join your group even though I did not know anything about Giardia. I would also like to thank a previous co-supervisor, Eva. I appreciated your support and encouragement coming from the same background.

During the years I have had a lot of collaborators and co-authors of papers and manuscripts. I would like to thank you all for your contributions. Especially I would like to mention a few. Calle, thank you for your clever ideas and patience with trying to explain them to me. Stefan, thank you for your good advice on matters dealing with sequencing and related areas. Moritz, thank you for your help with bioinformatics. The team in Brazil, thank you for welcoming me on my many visits. Osvaldo, thank you for your knowledge and help from distance. Sally and Mike, it has been nice to meet you and work with you. Hopefully we can publish our paper soon. Also, I would like to acknowledge all the people who have collected and processed the samples in the field and the lab for these studies, several of whom I have never met.

I started as a PhD student at SLU but I will finish as a PhD at UU. At SLU I would like to thank Emilia and Helena in the DNA lab for help and good company. I would also like to thank my fellow PhD students for the discussions and memories. At UU I would like to thank all the people at Micro where I spent the last year(s) of my studies, it is an inspiring place. I would especially like to thank the Giardia group for making me feel welcome and making the work environment a lot of fun. Laura, Dimitra, Jingyi, Feifei, Ásgeir, Showgy, Jana, Sascha, Inês, Eva and Anders, you are all really nice people and scientists.
To be able to write this thesis, there are many other aspects that are important besides the actual work. Therefore, I would like to thank some people for their part. First I would like to thank not a person but a club, UARS. When first deciding to become a PhD student there were many PhD students in the club and the advice from some of you helped me decide to go for it. Rowing is a sport that demands a lot of focus, training, and dedication with ups and downs, just like working on this thesis. Having other things than work to focus on has helped me over the years. Malin, you are the best rowing partner. However, I do not row with you because you are the best (that is just a bonus) but because you are my good friend. This is probably why we have shared so many other things besides rowing during my years as a PhD student. Thank you for all the laughs and good times. Jenna and Maria, thank you for all the nice lunches together. I have much enjoyed our conversations about PhD-student life as well as life in general.

My family, without you I would not be here. Mamma och pappa, thank you for always believing in me, supporting me, pushing me, caring and giving me the freedom to do what I want. Martin, thank you for being so strong, smart, caring and reliable even when it is not easy, I am very proud to be your big sister. Gunilla and Helmuth, thank you for always being there to help. Last but not least Billy, thank you for being part of my life. You make everything better, including me.
References


Briones, A. M., et al. (2008), Thorsellia anophelis is the dominant bacterium in a Kenyan population of adult Anopheles gambiae mosquitoes, ISME J, 2 (1), 74-82.


Chavshin, A. R., et al. (2015), Malpighian tubules are important determinants of Pseudomonas transstadial transmission and longtime persistence in Anopheles stephensi, Parasite Vector, 8, 36.


Kampfer, Peter, et al. (2015), Proposal of *Thorsellia kenyensis* sp. nov. and *Thorsellia kandunguensis* sp. nov., isolated from larvae of *Anopheles arabiensis*, as members of the family Thorselliaceae fam. nov', *Int J Syst Evol Microbiol*, 65 (2), 444-51.


Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 1691

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. (Prior to January, 2005, the series was published under the title “Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology”.)

Distribution: publications.uu.se
urn:nbn:se:uu:diva-352547